Affine Processes in Finance
-Numerical Approximation, Simulation and Model Properties-

Kyoung-Kuk Kim

Submitted in the partial fulfillment of the requirements
for the Degree of Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2008

Abstract

Affine Processes in Finance; Numerical Approximation, Simulation and Model Properties

Kyoung-Kuk Kim

This thesis deals with theoretical and numerical questions related to affine jump-diffusion models used in finance. In more detail, we look at three different classes within the affine jump-diffusion class.

The first is the Heston stochastic volatility model which has been used extensively since its first introduction by Heston (1993). To price financial derivatives with complex payoff structures, we have to resort to the Monte Carlo simulation. We propose new simulation schemes for the Heston model based on the squared Bessel bridge decomposition. These new methods perform well in different parameter settings and they are compared with two other existing methods, first, the exact scheme of Broadie and Kaya (2006) and, second, the QE method of Andersen (2005).

The second question is about the tail behavior of the canonical affine diffusion processes which were introduced by Dai and Singleton (2000) in the context of financial econometrics to study the term structure of interest rates. We show that the canonical models have light tails or exponential bounded tails, and the explicit conditions that guarantee light tails are given. Moreover, we prove that there exists a unique limiting stationary distribution for each canonical model and the regions of finite exponential moments of such stationary distributions are determined by the stability region of the dynamical system associated with a given model.

We further go into the detailed analysis of the dynamical system of a canonical affine diffusion process. We prove that the stability region of such a dynamical system can be represented by the union of stable sub-manifolds under some mild conditions, and also derive some partial differential equation of which solution is blow-up times of the dynamical system. Through an asymptotic analysis of those blow-up times, we calculate the implied volatility asymptotics for options with short maturities and extreme strikes based on Lee (2004).

The third and final question involves the general affine jump-diffusion models. It is computationally too expensive to apply numerical integration schemes to compute vanilla option prices in an affine jump-diffusion model which does not have an explicit Fourier transform formula. To extend the category of models that can be tested in financial econometrics, we apply the well known saddlepoint technique to affine jump-diffusion models. After we develop the basic idea and review some known saddlepoint techniques, we test them for the Heston model, the model of stochastic volatility with jumps (SVJ) and the Scott model. Implementation details and some modifications of existing methods are also given.

Contents

List of Figures

List of Tables

Acknowledgements

All work in this dissertation would not have been possible without the guidance of my advisor Professor Paul Glasserman. My deepest gratitude goes to him. Besides the insights that he showed, the inspiration that he gave, and the knowledge that he taught, I have learnt how to motivate myself, how to question, and how to research.

As a student who came from pure mathematics background, I was able to feel the joy of applied sciences for the past four years and am proud of being a part of this community. I sincerely thank my doctoral committee members: Prof. Mark Broadie, Prof. Assaf Zeevi, Prof. Steve Kou, and Prof. Rama Cont for all their teaching and advice. Also, Prof. Sergei Savin, Prof. Ward Whitt, Prof. Costis Maglaras, and Prof. Ioannis Karatzas motivated me to delve into the fascinating fields of probability and management science.

My life at Columbia has been fun and enjoyable, largely due to my friends. Discussions with them gave me new and fruitful insights. I want to thank Matulya Bansal, Itay Gurvich, KunSoo Park, Wei Ke, Wanmo Kang, Dongyoup Lee, Jeong-Seog Song, Sung-Uk Kim, and Kwang-Hyun Kim for their friendship.

I would like to extend my deepest gratitude to my family. My parents and younger brother have always believed in me and supported all the decisions that I made.

To Mom, Dad, and my brother

Chapter 1

Introduction

The theory of asset pricing, particularly that of financial derivatives, has been developed for the last three decades since Black and Scholes (1973), and Merton (1973). As there are many good sources that account for detailed theories (see, e.g., Duffie 2001 or Musiela and Rutkowski 2005), we just briefly recall the fundamental pricing equation; for a contingent claim V which matures at T, its price at time $0 \leq t \leq T$ is given by

$$V_t = B_t \mathbb{E}\left[\frac{V_T}{B_T}\middle|\mathcal{F}_t\right] \tag{1.1}$$

where B is the numeraire asset and $\mathbb{E}$ is the expectation under a measure in which the process V/B is a martingale. We call this measure a martingale measure associated with B. Here $\mathcal{F}$ stands for the filtration to which V and B are adapted.

This "relative pricing" approach to derivative pricing is very popular in financial engineering. For example, we can choose the nemeraire to be a bank account which returns risk free interest with continuous compounding or to be the price of a bond. In these cases, the associated martingale measures are called the risk-neutral measure and the forward measure, respectively. However, if markets are incomplete, then there could be many other equivalent martingale measures that make V/B a martingale. In this thesis, however, we always assume that we start with the risk-neutral measure and thus avoid any complication related to market incompleteness.

To apply (1.1) to a derivative of interest, we need a model that describes the price

movement of the underlying asset (or assets) S of the contingent claim V. There is a universe of stochastic models for this purpose and it is conventional to set S as some specific semimartingale. For example, S could be a Lévy process or it could be a solution to a stochastic differential equation (SDE)

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t \tag{1.2}$$

with W being a multi-dimensional Brownian motion. Proceeding one more step, we could add jumps $J(S_t, t)$ to (1.2). In such a case, by imposing specific parametric forms on $\mu(S_t, t)$, $\sigma(S_t, t)$ and $J(S_t, t)$ we obtain some nice properties that are very useful in derivative pricing as in, for example, Black and Scholes (1973), Cox (1975), Heston (1993) or Kou (2002). In this regard, there is one important class of stochastic processes called affine jump-diffusions. The variety of models that fall into this class is explained in Section 5.1. In these models, the log of asset prices $X_t = \log S_t$ is given by a solution to the following SDE:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + d\left(\sum_{i=1}^{N(t)} V_i\right) \tag{1.3}$$

where

$$\mu(x) = K_0 + K_1 x, \quad K_0 \in \mathbb{R}^n,\ K_1 \in \mathbb{R}^{n\times n},\ x \in \mathbb{R}^n$$
$$(\sigma(x)\sigma(x)^\top)_{ij} = H_{0ij} + H_{1ij} \cdot x, \quad H_{0ij} \in \mathbb{R},\ H_{1ij} \in \mathbb{R}^n$$

and $N(t)$ is a Poisson random variable with intensity process $\lambda(X_t) = l_0 + l_1 \cdot X_t$ for $l = (l_0, l_1) \in \mathbb{R} \times \mathbb{R}^n$, V_i's are independent and identically distributed (iid) random variable that stand for jump sizes. Moreover, the numeraire, a bank account B, is also modeled as

$$B_t = \exp\left(\int_0^t (\rho_0 + \rho_1 \cdot X_s)ds\right).$$

The usefulness of affine jump-diffusion models lies in the fact that there exists an explicit Fourier transform formula by which we can compute the cumulative distribution function

of X_t:

$$\phi(\theta) = \mathbb{E}\exp(i\theta \cdot X_t) = \exp(\alpha(T-t) + \beta(T-t) \cdot X_t) \tag{1.4}$$

where α, β are solutions of some ordinary differential equations. We refer to Duffie et al. (2000) and Duffie et al. (2003) for the detailed analysis of affine jump-diffusion processes.

We investigate theoretical and numerical questions related to affine jump-diffusion processes in the subsequent chapters. More specifically, we study affine jump-diffusions at three levels. At the simplest level, there is the Heston stochastic volatility model (Heston 1993) which is a two-dimensional model consisting of the stock price process S_t and the variance process V_t:

$$\begin{aligned} dS_t &= \mu S_t dt + S_t \sqrt{V_t} dW_t^1, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma \sqrt{V_t} dW_t^2 \end{aligned}$$

with (W^1, W^2) being a two-dimensional Brownian motion with correlation ρ. It is still a very popular model in financial engineering and has been widely applied to various kinds of markets such as bonds, equities and indices. Considering the complexity of derivatives that exist today, e.g., barrier options, bermudan options etc., Monte Carlo simulation is a widely used pricing method with great popularity. In Chapter 2, we study an efficient simulation method of the Heston model, which builds on the exact simulation scheme of Broadie and Kaya (2006). This method is based on a series expansion of the integral of the variance process conditional on the endpoints, $\left(\int_0^t V_s ds | V_0, V_t\right)$.

At somewhat intermediate level, we study canonical affine diffusion processes introduced by Dai and Singleton (2000). These were used for the study of term structure of interest rates and have the following coefficients: in (1.3),

$$\mu(X_t) = -A^\top(\Theta - X_t), \quad \sigma(X_t)\sigma(X_t)^\top = \left(\begin{array}{c|c} D_1 & \\ \hline & D_2 \end{array}\right)$$

where A is a matrix with some special conditions and D_i's are some diagonal matrices of which entries are affine functions of X_t. See Section 3.2 for details on the parametric

restrictions on the models. The main question we look at regarding canonical affine diffusion processes is when the exponential moments of X are finite. The finiteness of exponential moments of a random variable U is closely related to the tail behavior of the distribution of U. For example, if a non-negative random variable U has $\mathbb{E}e^{\theta U} < \infty$ for any $\theta \in \mathbb{R}$, then U is light tailed and if $\mathbb{E}e^{\theta U} = \infty$ for any $\theta > 0$, then U is heavy tailed. In between, U has an exponentially bounded tail. In Chapter 3, we prove that X admits a unique limiting stationary distribution, say X_∞, and the set of vectors θ such that $\theta \cdot X_\infty$ has finite exponential moments coincide with the stability region of the dynamical system that α, β satisfy. Moreover, we show the necessary and sufficient conditions on θ for $\theta \cdot X_t$ to be Gaussian. More background on the concepts like stability region or dynamical system can be found in Section 3.4.1.

In the following chapter, more detailed study of the dynamical system of α and β is carried out. Especially, the stability region of the system and associated partial differential equations are discussed. These questions are important because it has an implication in the context of option pricing. For example, Lee (2004) showed how the critical exponents p^* or q^* such that $\mathbb{E}e^{(p^*+1)\theta_t \cdot X_t}$ or $\mathbb{E}e^{-q^*\theta_t \cdot X_t}$ become infinite are related to the slopes of implied volatility curves at extreme strikes, while the asset price process is given by $\theta_t \cdot X_t$ for a deterministic function of time θ_t. In Section 4.5, we compute these slopes explicitly for options with extreme strikes and small maturities through an asymptotic analysis of the stability region of the dynamical system of α and β.

At the most general level, we deal with affine jump-diffusions (1.3). Even though the transform formula (1.4) is available, only simple models that have closed form α, β have been studied particularly in financial econometrics. This is mainly because it becomes too time consuming otherwise. In more detail, a probability $\mathbb{P}(\theta \cdot X_t > y)$ is calculated via the Fourier inversion formula

$$\mathbb{P}(\theta \cdot X_t > y) = \frac{1}{2\pi i} \int_{\tau - i\infty}^{\tau + i\infty} \phi(z\theta) e^{-zy} \frac{dz}{z}, \quad \tau > 0 \tag{1.5}$$

and we use a numerical integration scheme to calculate this integral. However, if $\phi(\cdot)$ is not available in closed form, i.e., α and β are not solvable analytically, then we have

to solve differential equations associated with α and β numerically at each evaluation point of the above integral. This certainly causes a big problem if one wants to use a more general affine jump-diffusion model than simple ones. In this regard, an efficient numerical approximation of option prices is attractive and we apply the well known statistical method called the saddlepoint technique to affine jump-diffusions. Introduced in statistics by Daniels (1954) and applied to derivative pricing by Rogers and Zane (1999), this technique is essentially asymptotic expansions of contour integrals such as (1.5) in the complex plane and has been in use, for example, in option pricing, risk management and credit derivatives. We will see in Chapter 5 how much computational efficiency is obtained by applying saddlepoint approximations in the affine jump-diffusion setting and this opens a possibility of testing more complex models in financial econometrics.

Chapter 2

Gamma Expansion of the Heston Stochastic Volatility Model

Approximate simulation methods for the Heston stochastic volatility model are proposed. Based on the squared Bessel bridge decomposition in Pitman and Yor (1982), the integral of the variance process $\int_0^t V_s ds$ conditional on the endpoints can be simulated by generating three independent random variables. Computational gain is due to, first, that we avoid the inverse transform method and, second, that we reduce the computation of Bessel functions as much as possible.

2.1 Introduction

Among many stochastic models used in quantitative finance, Heston's stochastic volatility model is still one of the most popular models among practitioners. By introducing stochasticity to the volatility process of the asset price of interest, the model made it possble to explain the implied volatility skews to some extent. Now it is applied to many different kinds of financial intstruments including bonds, equities and indices. On the other hand, we observed the very fast growth of financial markets in terms of the size and the complexity of derivatives people trade. So risk management as well as derivative pricing is ever more important. This means that efficient calculation of prices and greeks is a vital factor in quantitative finance.

Regarding the Heston model, Monte Carlo simulation as a method of pricing and hedging is still very important despite the availability of the closed form solutions of vanilla option prices because we do not and cannot expect to have such closed form solutions for many exotic path dependent derivatives. Until recently, discretization methods have been the default approaches in Monte Carlo simulation of the Heston model. This class of methods includes the Euler scheme, the Milstein scheme and other schemes with higher order of convergence. Kloeden and Platen (1999) explain various methods in their textbook and concrete numerical investigation was conducted by Kahl and Jäckel (2006). However, as noted in literature (see, e.g., Andersen 2005 or Broadie and Kaya 2006), these methods lose their appeal when it comes to the Heston model with not-benign model parameters, causing problems with negative variance, which might generate significant biases. The recent discretization method proposed by Andersen (2005) attracted attention as his method avoids such problems and works reasonably well in different market situations, i.e. in a wide range of model parameter values while maintaining computational efficiency.

A totally different approach was pioneered by Broadie and Kaya (2006). Without any discretization, their method is an exact scheme and produces no bias. Based on the inverse transform method, the key step is the computation of the characteristic function of $\left(\int_s^t V_u du | V_s, V_t\right)$, the integral of the variance process V conditional on the endpoints V_s and V_t. Even though this exact scheme recovers the usual convergence rate of Monte Carlo

simulation (so outperforms the Euler scheme), people find it hard to implement in practice because of the computational time that it requires. This computational burden is due to, first, that we need to implement a root finding algorithm to apply the inverse transform method, which in turn requires many computations of the characteristic function and, second, that the characteristic function involves two evaluations of the modified Bessel function of the first kind, which is a solution of second order ordinary differential equation and represented by an infinite series.

In this chapter, we propose yet another approximate simulation method of the Heston model along the line of Broadie and Kaya (2006). The bottom line is to use mathematical properties of the squared Bessel bridge investigated by Pitman and Yor (1982, 2000). Based on their decomposition of the squared Bessel bridge, we prove that the conditional path integral $\left(\int_0^t V_s ds | V_0, V_t\right)$ is decomposed into the sum of three independent random variables. Moreover, each of these random variables admit series expansions using Poisson, exponential, gamma and Bessel random variables. We also test a simulation approach that uses a single Beta random variables.

The chapter is constructed as follows. In Section 2.2, we present our main result. In the following two sections, we review the exact scheme of Broadie and Kaya (2006) and detail our approximate simulation scheme. Numerical results are given in Section 2.5. Section 2.6 deals with the case when we are given a non-equidistant time grid. Section 2.7 concludes.

2.2 Main Result

The Heston model is a two-dimensional stochastic process (S_t, V_t) which satisfies the following SDE:

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{V_t}\left(\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2\right) \tag{2.1}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^1 \tag{2.2}$$

where (W^1, W^2) is two-dimensional standard Brownian motion. The variance V process is called the Cox-Ingersoll-Ross (CIR) process and has been studied extensively and used for

term structure modeling since its first introduction by Cox et al. (1985). It is known that the distribution of V_t for a given initial value V_0 and time t follows the noncentral chi-square distribution,

$$V_t \overset{d}{=} \frac{\sigma^2(1-e^{-\kappa t})}{4\kappa}\chi'^2_\delta\left(\frac{4\kappa e^{-\kappa t}}{\sigma^2(1-e^{-\kappa t})}V_0\right), \quad t>0, \quad \delta=\frac{4\kappa\theta}{\sigma^2}. \tag{2.3}$$

However, from the viewpoint of the Monte Carlo simulation of the Heston model, it is important to simulate the path integral of V, $\int_0^t V_s ds$, not just V_t. In more detail, a simple application of Itô's formula to $\log S_t$ shows

$$S_t = S_0 \exp\left(\mu t - \frac{1}{2}\int_0^t V_s ds + \rho\int_0^t \sqrt{V_s}dW_s^1 + \sqrt{1-\rho^2}\int_0^t \sqrt{V_s}dW_s^2\right)$$

and thus, given $\int_0^t V_s ds$ and $\int_0^t \sqrt{V_s}dW_s^1$,

$$\log\frac{S_t}{S_0} \overset{d}{=} \mathcal{N}\left(\mu t - \frac{1}{2}\int_0^t V_s ds + \rho\int_0^t \sqrt{V_s}dW_s^1, (1-\rho^2)\int_0^t V_s ds\right) \tag{2.4}$$

as W^2 is independent of the V process. From (2.2) we also have

$$\int_0^t \sqrt{V_s}dW_s^1 = \frac{1}{\sigma}\left\{V_t - V_0 - \kappa\theta t + \kappa\int_0^t V_s ds\right\}.$$

Hence, if the joint distribution of $\left(V_t, \int_0^t V_s ds\right)$ is known and can be simulated efficiently, then the simulation of S_t given (S_0, V_0) is an easy task. This is the approach taken by Broadie and Kaya (2006) and their novel method will be explained more in Section 2.3. As the distribution of V_t is explicitly known, our focus is on $\left(\int_0^t V_s ds | V_0, V_t\right)$. Therefore, in this section, we investigate some properties of this conditional path integral, and they are based on the squared Bessel bridge decomposition studied by Pitman and Yor (1982) and on a close look at the characteristic function of the integral.

Remark If we define a process $B(t)$ by

$$V_t = e^{-\kappa t}B\left(\frac{e^{\kappa t}-1}{4\kappa/\sigma^2}\right), \tag{2.5}$$

then B becomes a δ-dimensional squared Bessel process which satisfies

$$dB(t) = \delta dt + 2\sqrt{B(t)}dW(t), \quad W\left(\frac{e^{\kappa t}-1}{4\kappa/\sigma^2}\right) = \int_0^t \frac{\sigma}{2} e^{\kappa s/2} dW_s^1.$$

It is well known that the above SDE has the unique strong solution for each $\delta \geq 0$ and $B_0 = V_0 \geq 0$. Moreover, if $\delta = 0$, then 0 is an absorbing state. If $\delta \in (0,2)$, then 0 is reached almost surely, but instantaneously reflecting. See p.439 of Revuz and Yor (1999) for details.

Remark From (2.5), we observe that the conditional law $(V_s, 0 \leq s \leq t | V_0, V_t)$ can be defined using the conditional law $(B_s, 0 \leq s \leq t | B_0, B_t)$. A reader can consult p.446 of Pitman and Yor (1982) for more information about the definition of the conditional law of the squared Bessel bridge.

Before proceed, recall the definition of a Bessel random variable which we denote by $BES(\nu, z)$ with $\nu > -1$ and $z > 0$ (see Yuan and Kalbfleisch 2000). It is a random variable X taking non-negative integer values with probabilities

$$p_n(\nu, z) := \mathbb{P}(X = n) = \frac{(z/2)^{2n+\nu}}{I_\nu(z) n! \Gamma(n+\nu+1)}, \quad n \geq 0$$

where $I_\nu(z)$ is the modified Bessel function of the first kind. We will drop ν and z when there is no source of confusion. Now we state our main result. The conditional path integral of the CIR process can be decomposed into three independent random variables, all of which admit series representations.

Theorem 2.2.1 *The distribution of $\int_0^t V_s ds$ conditional on endpoints V_0, V_t admits a decomposition:*

$$\left(\int_0^t V_s ds \mid V_0 = v_0,\ V_t = v_t\right) \stackrel{d}{=} X_1 + X_2 + X_3 \equiv X_1 + X_2 + \sum_{j=1}^{\eta} Z_j$$

where X_i's are independent random variables, Z_j's are i.i.d. copies of Z and η is an independent Bessel random variable with parameters $\nu = \delta/2 - 1$ and

$$z = \frac{2\kappa/\sigma^2}{\sinh(\kappa t/2)} \sqrt{v_0 v_t}.$$

Moreover, X_1, X_2 and Z have the following representations:

$$X_1 = \sum_{n=1}^{\infty} \frac{1}{\gamma_n} \sum_{j=1}^{N_n} Exp_j(1), \quad X_2 = \sum_{n=1}^{\infty} \frac{1}{\gamma_n} \Gamma_n(\delta/2, 1), \quad Z = \sum_{n=1}^{\infty} \frac{1}{\gamma_n} \Gamma_n(2, 1)$$

where

$$\lambda_n = \frac{16\pi^2 n^2}{\sigma^2 t(\kappa^2 t^2 + 4\pi^2 n^2)}, \quad \gamma_n = \frac{\kappa^2 t^2 + 4\pi^2 n^2}{2\sigma^2 t^2}$$

and N_n's are independent Poisson random variables with mean $(v_0 + v_t)\lambda_n$, $Exp_j(1)$'s i.i.d. Exponential random variables with rate 1 and $\Gamma_n(k, \theta)$'s independent gamma random variables with shape parameter k and scale parameter θ.

Proof We prove the result in two steps. First, we show that $(\int_0^t V_s ds | V_0, V_t)$ can be decomposed into the sum of three random variables and second, each of those random variables has the series representation above based on its Laplace transform.

The first step is a simple application of a result in Pitman and Yor (1982), p.456. Fix $t > 0$ and define a process $\{A_s\}_{0 \le s \le 1}$ by

$$A_s = \frac{4}{\sigma^2 t} V_{st}. \tag{2.6}$$

Then, it is easy to see that A solves a stochastic differential equation

$$dA_s = (\delta + 2aA_s)ds + 2\sqrt{A_s} dW_s$$

with $a = -\kappa t/2$ and a standard Brownian motion W. This is called a δ-dimensional squared Ornstein-Ulhenbeck (OU) process with parameter a. Let us denote the conditional law

$$(A_s, 0 \le s \le 1 | A_0 = x, A_1 = y) \tag{2.7}$$

by $\{A^{\delta,1}_{x,y}(s)\}_{0 \le s \le 1}$ or simply $A^{\delta,1}_{x,y}$. Pitman and Yor (1982) showed that this squared OU bridge has the following decomposition:

$$A^{\delta,1}_{x,y} \stackrel{d}{=} A^{0,1}_{x,0} + A^{0,1}_{0,y} + A^{\delta,1}_{0,0} + A^{4\eta,1}_{0,0}$$

where the four squared OU bridges on the right hand side are independent processes and η is an independent Bessel random variable with $\nu = \delta/2 - 1$ and $z = \sqrt{xy}a/\sinh(a)$. Here, $A^{0,1}_{0,y}$ should be understood as the law of the time-reversed process of $A^{0,1}_{y,0}$ because 0 is an absorbing state for a 0-dimensional squared OU process.

From the above decomposition of the squared OU bridge, we get

$$\int_0^1 A^{\delta,1}_{x,y}(s)ds \overset{d}{=} \int_0^1 A^{0,1}_{x,0}(s)ds + \int_0^1 A^{0,1}_{0,y}(s)ds + \int_0^1 A^{\delta,1}_{0,0}(s)ds + \int_0^1 A^{4\eta,1}_{0,0}(s)ds.$$

But, the second term on the right hand side is same in distribution as $\int_0^1 A^{0,1}_{y,0}(s)ds$ by the definition of $A^{0,1}_{0,y}$. On the other hand, a family of the conditional laws $\{A^{f,1}_{x,0}\}_{x\geq 0, f\geq 0}$ parameterized by x, f has an additivity property

$$A^{f,1}_{x,0} + A^{f',1}_{x',0} \overset{d}{=} A^{f+f',1}_{x+x',0}$$

which is a direct consequence of a similar additivity property of the squared Bessel bridges and the transformation (6.b) in Pitman and Yor (1982). Therefore, we have

$$\int_0^1 A^{\delta,1}_{x,y}(s)ds \overset{d}{=} \int_0^1 A^{0,1}_{x+y,0}(s)ds + \int_0^1 A^{\delta,1}_{0,0}(s)ds + \int_0^1 A^{4\eta,1}_{0,0}(s)ds.$$

Finally, we also observe that the last term on the right side can be expressed as

$$\int_0^1 A^{4\eta,1}_{0,0}(s)ds \overset{d}{=} \sum_{j=1}^{\eta} \int_0^1 \left(A^{4,1}_{0,0}\right)^{(j)}(s)ds$$

thanks to the above additivity property. Here, $(A^{4,1}_{0,0})^{(j)}$'s are i.i.d. copies of $A^{4,1}_{0,0}$.

Now, the remaining step is to convert the decomposition of $\int_0^1 A^{\delta,1}_{x,y}(s)ds$ into that of $(\int_0^t V_s ds | V_0, V_t)$. It is obvious from (2.6) and (2.7) that

$$\left(\int_0^t V_s ds | V_0 = v_0, V_t = v_t\right) = \frac{\sigma^2 t^2}{4} \int_0^1 A^{\delta,1}_{x,y}(s)ds$$

where $x = 4v_0/(\sigma^2 t)$ and $y = 4v_t/(\sigma^2 t)$. Then, the first part of the proof is complete by

defining

$$\begin{aligned} X_1 &= \frac{\sigma^2 t^2}{4} \int_0^1 A^{0,1}_{x+y,0}(s)ds, \\ X_2 &= \frac{\sigma^2 t^2}{4} \int_0^1 A^{\delta,1}_{0,0}(s)ds, \\ Z &= \frac{\sigma^2 t^2}{4} \int_0^1 A^{4,1}_{0,0}(s)ds. \end{aligned}$$

Since $a = -\kappa t/2$, $x = 4v_0/(\sigma^2 t)$ and $y = 4v_t/(\sigma^2 t)$, we have

$$z = \frac{a}{\sinh(a)}\sqrt{xy} = \frac{2\kappa/\sigma^2}{\sinh(\kappa t/2)}\sqrt{v_0 v_t}.$$

Let us turn to the second statement of the theorem which turns out to be useful in Monte Carlo simulation of the Heston model in later sections. As essential tools, we record the Laplace transforms of X_1, X_2 and Z in the next lemma.

Lemma 2.2.1 *The Laplace transforms Φ^1, Φ^2, Φ^Z of X_1, X_2 and Z are given as follows: for $b \geq 0$,*

$$\Phi^1(b) = \exp\left(\frac{(v_0+v_t)}{\sigma^2}\left(\kappa \coth\frac{\kappa t}{2} - L\coth\frac{Lt}{2}\right)\right), \tag{2.8}$$

$$\Phi^2(b) = \left(\frac{L}{\kappa}\cdot\frac{\sinh \kappa t/2}{\sinh Lt/2}\right)^{\delta/2}, \tag{2.9}$$

$$\Phi^Z(b) = \left(\frac{L}{\kappa}\cdot\frac{\sinh \kappa t/2}{\sinh Lt/2}\right)^{2} \tag{2.10}$$

where $L = \sqrt{2\sigma^2 b + \kappa^2}$.

Proof The proof is a straightforward calculation based on the Laplace transforms of squared Bessel bridges in Revuz and Yor (1999) and the change of measure formula (6.d) in Pitman and Yor (1982).

Recall that a 0-dimensional squared Bessel process B defined by $dB_t = 2\sqrt{B_t}dW_t$ with W being standard Brownian motion under $\tilde{\mathbb{Q}}$ has the Laplace transform

$$\tilde{\mathbb{E}}\left[\exp\left(-\frac{b^2}{2}\int_0^1 B_s ds\right)\middle| B_0 = x,\ B_1 = 0\right] = \exp\left(\frac{x}{2}(1 - b\coth b)\right)$$

for $b \in \mathbb{R}$ and $x \geq 0$ where $\tilde{\mathbb{E}}$ means expectation under $\tilde{\mathbb{Q}}$. Using the change of measure

formula (6.d) in Pitman and Yor (1982), for $b \geq 0$ we have (with $a = -\kappa t/2$)

$$\begin{aligned}
\Phi^1(b) &= \mathbb{E}\left[\exp\left(-b\frac{\sigma^2 t^2}{4}\int_0^1 A^{0,1}_{x+y,0}(s)ds\right)\right] \\
&= \tilde{\mathbb{E}}\left[\exp\left(-\left(\frac{b\sigma^2 t^2}{4}+\frac{a^2}{2}\right)\int_0^1 A^{0,1}_{x+y,0}(s)\right)\right] \div \tilde{\mathbb{E}}\left[\exp\left(-\frac{a^2}{2}\int_0^1 A^{0,1}_{x+y,0}(s)ds\right)\right] \\
&= \exp\left(\frac{(v_0+v_t)}{\sigma^2}\left(\kappa\coth\frac{\kappa t}{2} - L\coth\frac{Lt}{2}\right)\right)
\end{aligned}$$

where $L = \sqrt{2\sigma^2 b + \kappa^2}$. This is exactly how the Laplace transform of $\left(\int_0^t V_s ds \middle| V_0, V_t\right)$ is produced in Broadie and Kaya (2006).

As for X_2 and Z, recall the Laplace transform

$$\tilde{\mathbb{E}}\left[\exp\left(-\frac{b^2}{2}\int_0^1 B_s ds\right)\middle| B_0 = B_1 = 0\right] = \left(\frac{b}{\sinh b}\right)^{f/2}$$

of the f-dimensional squared Bessel bridge such that $dB_t = f dt + 2\sqrt{B_t}dW_t$, W being a standard Brownian motion under $\tilde{\mathbb{Q}}$. By proceeding similarly as above, we obtain the desired results. ∎

Another very useful tool is the following infinite product in p.22 of Pitman and Yor (2000):

$$\prod_{n=1}^{\infty}\left(1+\frac{x^2}{\pi^2 n^2}\right)^{-1} = \frac{x}{\sinh x} \tag{2.11}$$

by which they presented the squared Bessel bridge with zero endpoints as an infinite sum of gamma random variables. This observation will be revisited in Section 2.6.

It is well known that $\prod(1+a_n)$ with $a_n \neq -1$ converges simultaneously with $\sum \log(1+a_n)$ (using the principal branch in $\mathbb{C}$ if necessary) and that this product absolutely converges if and only if $\sum |a_n|$ does. See Ahlfors (1979), p.192, Theorems 5, 6. These facts and (2.11) imply that

$$\sum_{n=1}^{\infty}\log\left(1+\frac{x^2}{\pi^2 n^2}\right) = -\log\frac{x}{\sinh x}.$$

Since the left hand side is uniformly convergent on compact intervals, we can take term-

wise derivatives to deduce

$$\sum_{n=1}^{\infty} \frac{2x^2}{x^2 + \pi^2 n^2} = x \coth x - 1.$$

Now, for two real values $x \geq y \geq 0$, we get

$$x \coth x - y \coth y = \sum_{n=1}^{\infty} \frac{2x^2}{x^2 + \pi^2 n^2} - \sum_{n=1}^{\infty} \frac{2y^2}{y^2 + \pi^2 n^2} = \sum_{n=1}^{\infty} \frac{2\pi^2 n^2 (x^2 - y^2)}{(x^2 + \pi^2 n^2)(y^2 + \pi^2 n^2)}.$$

Plugging $x = Lt/2$, $y = \kappa t/2$ in this formula (with $L = \sqrt{2\sigma^2 b + \kappa^2}$) and rearranging terms,

$$\frac{1}{\sigma^2}\left(\kappa \coth \frac{\kappa t}{2} - L \coth \frac{Lt}{2}\right) = -\sum_{n=1}^{\infty} \frac{16\pi^2 n^2}{\sigma^2 t(\kappa^2 t^2 + 4\pi^2 n^2)} \cdot \frac{b}{b + \frac{\kappa^2 t^2 + 4\pi^2 n^2}{2\sigma^2 t^2}} = -\sum_{n=1}^{\infty} \frac{\lambda_n b}{b + \gamma_n}.$$

We turn to the infinite sum in the statement. Define

$$X_1' = \sum_{n=1}^{\infty} \frac{1}{\gamma_n} \sum_{j=1}^{N_n} Exp_j(1).$$

This random variable is well defined because the sum of variances $\sum_{n=1}^{\infty} 2(v_0 + v_t)\lambda_n/\gamma_n^2$ is obviously finite and, thus, the infinite sum converges almost surely. Therefore, for $b \geq 0$, by the Dominated Convergence Theorem, we have

$$\begin{aligned}
\log \mathbb{E} e^{-bX_1'} &= \log \mathbb{E} \exp\left(-b \sum_{n=1}^{\infty} \frac{1}{\gamma_n} \sum_{j=1}^{N_n} Exp_j(1)\right) \\
&= \sum_{n=1}^{\infty} \log \mathbb{E} \exp\left(-\frac{b}{\gamma_n} \sum_{j=1}^{N_n} Exp_j(1)\right) \\
&= -\sum_{n=1}^{\infty} \frac{(v_0 + v_t)\lambda_n b}{b + \gamma_n}.
\end{aligned}$$

Hence, by the uniqueness of the Laplace transform, $X_1 \stackrel{d}{=} X_1'$ and we can set X_1 to be the series representation in the statement without any loss of generality.

As for X_2, from (2.9) and (2.11) we have

$$\begin{aligned}
\log \mathbb{E}e^{-bX_2} &= \frac{\delta}{2}\left(\log\frac{Lt/2}{\sinh Lt/2} - \log\frac{\kappa t/2}{\sinh \kappa t/2}\right) \\
&= \frac{\delta}{2}\left\{-\sum_{n=1}^{\infty}\log\left(1+\frac{L^2t^2}{4\pi^2n^2}\right)+\sum_{n=1}^{\infty}\log\left(1+\frac{\kappa^2t^2}{4\pi^2n^2}\right)\right\} \\
&= -\frac{\delta}{2}\sum_{n=1}^{\infty}\log\left(1+\frac{b}{\gamma_n}\right) \\
&= \sum_{n=1}^{\infty}\log\mathbb{E}\exp\left(-\frac{b}{\gamma_n}\Gamma_n(\delta/2,1)\right) \\
&= \log\mathbb{E}\exp\left(-b\sum_{n=1}^{\infty}\frac{1}{\gamma_n}\Gamma_n(\delta/2,1)\right)
\end{aligned}$$

where the random variable in the last expression is well defined by the same reason as above. The expansion of Z is a special case with $\delta = 4$. The proof is complete. ∎

Remark We note that X_1, X_2 and Z belong to certain classes of infinitely divisible distributions. As noted in Bondesson (1982), the class $\mathcal{T}_2$ is described as the set of distributions of weak limits of finite convolutions of Poisson mixtures of exponential distributions and the class of generalized gamma convolutions (g.g.c.) is the set of distributions of weak limits of finite convolutions of exponential distributions. The class of g.g.c. is again a subset of $\mathcal{T}_2$. Therefore, X_1, X_2, Z are in $\mathcal{T}_2$ and, in particular, X_2, Z in g.g.c. So, X_3 is the mixture of g.g.c. distributions with the Bessel law as the mixing distribution. See also Steutel and van Harn (2004) for an extensive study of infinitely divisible distributions.

In contrast, the distribution V_t is that of the Poisson mixture of gamma random variables as easily deduced from its Laplace transform.

2.3 Monte Carlo Simulation of the Heston Model

The exact simulation scheme of the Heston model developed by Broadie and Kaya (2006) exploits the explicit characteristic function formula of the squared Bessel bridge and the facts observed in Section 2.2. Briefly reviewing its procedure,

- Simulate V_t given V_0 using a noncentral chi-square distribution as in (2.3).

- Simulate $\int_0^t V_s ds$ given V_0 and V_t using the inverse transform method, i.e., for $U \sim unif[0,1]$ we find x such that $\mathbb{P}\left(\int_0^t V_s ds \leq x | V_0, V_t\right) = U$ using a root-finding algorithm. This cumulative distribution function (CDF) is computed by the Fourier inversion integrals since there is a closed form expression for the Fourier transform of $\left(\int_0^t V_s ds | V_0, V_t\right)$.
- Simulate S_t given S_0 using (2.4).

This exact simulation method recovers $O(s^{-1/2})$ convergence of an unbiased Monte Carlo estimator unlike other discretization methods. Here s means a user's computational budget. However, as noted in Andersen (2005), this method is computationally expensive and so loses some practical appeal.

We focus on the second step of the Broadie-Kaya scheme and aims to improve the computational efficiency of the simulation of $\left(\int_0^t V_s ds | V_0, V_t\right)$ by applying our main result, Theorem 2.2.1. In applying the series expansions of X_i's, we have to truncate them at some level $n = K$. Proposition 2.3.1 is useful in this regard. For notational simplicity, we introduce three random variables

$$X_1^K = \sum_{n=K+1}^{\infty} \frac{1}{\gamma_n} \sum_{j=1}^{N_n} Exp_j(1), \quad X_2^K = \sum_{n=K+1}^{\infty} \frac{1}{\gamma_n} \Gamma_n(\delta/2, 1), \quad Z^K = \sum_{n=K+1}^{\infty} \frac{1}{\gamma_n} \Gamma_n(2, 1)$$

and we also denote gamma random variables that match the mean and the variance of each of the above three random variables by Γ^i for $i = 1, 2, 3$. The next result shows asymptotic decay rates of these means and variances and they are useful in the proof of Proposition 2.3.1.

Lemma 2.3.1 *As K increases,*

$$\begin{aligned}
\mathbb{E}X_1^K &\sim \frac{2(v_0 + v_t)t}{\pi^2 K}, \quad Var(X_1^K) \sim \frac{2(v_0 + v_t)\sigma^2 t^3}{3\pi^4 K^3}, \\
\mathbb{E}X_2^K &\sim \frac{\delta\sigma^2 t^2}{4\pi^2 K}, \quad Var(X_2^K) \sim \frac{\delta\sigma^4 t^4}{24\pi^4 K^3}, \\
\mathbb{E}Z^K &\sim \frac{\sigma^2 t^2}{\pi^2 K}, \quad Var(Z^K) \sim \frac{\sigma^4 t^4}{6\pi^4 K^3}.
\end{aligned}$$

Proof Observe that

$$\begin{aligned} Var(X_1^K) &= (v_0+v_t)\sum_{n=K+1}^{\infty}\frac{2\lambda_n}{\gamma_n^2} = (128(v_0+v_t)\pi^2\sigma^2t^3)\sum_{n=K+1}^{\infty}\frac{n^2}{(\kappa^2t^2+4\pi^2n^2)^3} \\ &\sim (128(v_0+v_t)\pi^2\sigma^2t^3)\int_K^{\infty}\frac{y^2}{(4\pi^2y^2)^3}dy = \frac{2(v_0+v_t)\sigma^2t^3}{3\pi^4K^3}. \end{aligned}$$

All other asymptotics are similarly obtained. ∎

Proposition 2.3.1 *For a random variable $V = X_1^K, X_2^K, Z^K$ and the corresponding Γ^i, the following asymptotic normality holds:*

$$\frac{V-\mathbb{E}(V)}{\sqrt{Var(V)}} \Longrightarrow \mathcal{N}(0,1), \quad \frac{\Gamma^i-\mathbb{E}(V)}{\sqrt{Var(V)}} \Longrightarrow \mathcal{N}(0,1) \quad \text{as } K\uparrow\infty.$$

Moreover, the distance between V and Γ^i decreases faster than the convergence of V to the normal distribution $\mathcal{N}^i := \mathcal{N}(\mathbb{E}(V), Var(V))$ in the following sense: for all b in a neighborhood of the origin,

$$0 \le \log\mathbb{E}e^{|b|V} - \log\mathbb{E}e^{|b|\Gamma^i} \le \log\mathbb{E}e^{|b|V} - \log\mathbb{E}e^{|b|\mathcal{N}^i}$$

for all sufficiently large K values.

Proof We will prove the statements for $V = X_1^K$ because other cases can be proven in a similar fashion.

From Theorem 2.2.1, it is easy to derive that for all b in a small neighborhood of the origin,

$$\log\mathbb{E}e^{bX_1^K} = \sum_{n=K+1}^{\infty}\frac{(v_0+v_t)\lambda_n b}{\gamma_n - b} = \sum_{n=K+1}^{\infty}\sum_{m=1}^{\infty}(v_0+v_t)\lambda_n\left(\frac{b}{\gamma_n}\right)^m$$

and one can readily show that the double sequence is absolutely convergent for each fixed b and K. This yields

$$\log\mathbb{E}\exp\left(b\frac{X_1^K-\mathbb{E}X_1^K}{\sqrt{Var(X_1^K)}}\right) = \frac{b^2}{2} + \sum_{m=3}^{\infty}\sum_{n=K+1}^{\infty}(v_0+v_t)\lambda_n\left(\frac{b}{\gamma_n\sqrt{Var(X_1^K)}}\right)^m$$

and the asymptotic normality for X_1^K follows if the double sequence in the above expression

converges to zero as a consequence of Lemma 15.15 of Kallenberg (2002). To see this convergence, first we observe that for fixed $s > 0$,

$$\begin{aligned}
\sum_{m=3}^{\infty}\sum_{n=K+1}^{\infty}\lambda_n\frac{s^m}{\gamma_n^m} &\leq \frac{4}{\sigma^2 t}\sum_{m=3}^{\infty}\sum_{n=K+1}^{\infty}\frac{s^m}{\gamma_n^m} = \frac{4}{\sigma^2 t}\sum_{m=3}^{\infty}\sum_{n=K+1}^{\infty}\frac{(2\sigma^2t^2s)^m}{(\kappa^2t^2+4\pi^2n^2)^m} \\
&\leq \frac{4}{\sigma^2 t}\sum_{m=3}^{\infty}\int_K^{\infty}\frac{(2\sigma^2t^2s)^m}{(\kappa^2t^2+4\pi^2y^2)^m}dy \leq \frac{4}{\sigma^2 t}\sum_{m=3}^{\infty}\int_K^{\infty}\frac{(2\sigma^2t^2s)^m}{(4\pi^2y^2)^m}dy \\
&\leq \frac{4}{\sigma^2 t}\sum_{m=3}^{\infty}\left(\frac{\sigma^2t^2s}{2\pi^2}\right)^m\frac{1}{(2m-1)K^{2m-1}}.
\end{aligned}$$

Then, we get

$$\begin{aligned}
\sum_{m=3}^{\infty}\sum_{n=K+1}^{\infty}\lambda_n\left(\frac{|b|}{\gamma_n\sqrt{Var(X_1^K)}}\right)^m &\leq \frac{4}{\sigma^2 t}\sum_{m=3}^{\infty}\left(\frac{\sigma^2t^2|b|}{2\pi^2\sqrt{Var(X_1^K)}}\right)^m\frac{1}{(2m-1)K^{2m-1}} \\
&\leq \frac{4K}{\sigma^2 t}\sum_{m=3}^{\infty}\left(\frac{\sigma^2t^2|b|}{2\pi^2K^2\sqrt{Var(X_1^K)}}\right)^m \\
&= \frac{4K}{\sigma^2 t}\left(\frac{\sigma^2t^2|b|}{2\pi^2K^2\sqrt{Var(X_1^K)}}\right)^3 \\
&\quad \div\left(1-\frac{\sigma^2t^2|b|}{2\pi^2K^2\sqrt{Var(X_1^K)}}\right)
\end{aligned}$$

where the last equality holds for all sufficiently large K's and it is easy to see that the last formula goes to zero as K increases, utilizing Lemma 2.3.1.

By matching the mean and the variance, one can get the shape parameter k and the scale parameter ζ of Γ^1,

$$k = \frac{\left(\mathbb{E}X_1^K\right)^2}{Var(X_1^K)} \sim \frac{6(v_0+v_t)K}{\sigma^2 t}, \quad \zeta = \frac{Var(X_1^K)}{\mathbb{E}X_1^K} \sim \frac{\sigma^2t^2}{3\pi^2K^2}.$$

From

$$\log \mathbb{E}e^{b\Gamma^1} = -k\log(1-\zeta b) = \sum_{m=1}^{\infty}\frac{k}{m}(\zeta b)^m,$$

we get

$$\log \mathbb{E} \exp\left(b \frac{\Gamma^1 - \mathbb{E}X_1^K}{\sqrt{Var(X_1^K)}} \right) = \frac{b^2}{2} + \sum_{m=3}^{\infty} \frac{k}{m} \left(\frac{\zeta b}{\sqrt{Var(X_1^K)}} \right)^m$$

and the asymptotic normality for Γ^1 follows because

$$\sum_{m=3}^{\infty} \frac{k}{m} \left(\frac{\zeta |b|}{\sqrt{Var(X_1^K)}} \right)^m \leq k \left(\frac{\zeta |b|}{\sqrt{Var(X_1^K)}} \right)^3 \div \left(1 - \frac{\zeta |b|}{\sqrt{Var(X_1^K)}} \right) \to 0.$$

Moreover, we have

$$\begin{aligned} \log \mathbb{E} e^{|b| X_1^K} - \log \mathbb{E} e^{|b| \Gamma^1} &= (v_0 + v_t) \sum_{m=3}^{\infty} \left(\sum_{n=K+1}^{\infty} \frac{\lambda_n}{\gamma_n^m} \right) |b|^m - \sum_{m=3}^{\infty} \left(\frac{k \zeta^m}{m} \right) |b|^m \\ &= (v_0 + v_t) \sum_{m=3}^{\infty} \sum_{n=K+1}^{\infty} \frac{\lambda_n}{\gamma_n^m} (1 - R_{n,m}) |b|^m \end{aligned}$$

with

$$R_{n,m} = \frac{k \zeta^m / (m(v_0 + v_t))}{\sum_{n=K+1}^{\infty} \lambda_n / \gamma_n^m}.$$

But, we observe that

$$\begin{aligned} \sum_{n=K+1}^{\infty} \frac{\lambda_n}{\gamma_n^m} &= \frac{16\pi^2}{\sigma^2 t} \sum_{n=K+1}^{\infty} \frac{n^2 (2\sigma^2 t^2)^m}{(\kappa^2 t^2 + 4\pi^2 n^2)^{m+1}} \sim \frac{4}{\sigma^2 t} \sum_{n=K+1}^{\infty} \frac{(2\sigma^2 t^2)^m}{(4\pi^2 n^2)^m} \\ &\sim \frac{4}{\sigma^2 t} \left(\frac{\sigma^2 t^2}{2\pi^2} \right)^m \int_K^{\infty} \frac{1}{y^{2m}} dy = \frac{4}{\sigma^2 t} \left(\frac{\sigma^2 t^2}{2\pi^2} \right)^m \frac{1}{(2m-1) K^{2m-1}} \end{aligned}$$

and $k\zeta^m / (m(v_0 + v_t)) \sim 6/(m\sigma^2 t K^{2m-1}) \left(\sigma^2 t^2 / (3\pi^2) \right)^m$. These imply

$$R_{n,m} \sim \frac{6m - 3}{2m} \left(\frac{2}{3} \right)^m$$

as K increases. Thus,

$$\begin{aligned}\log \mathbb{E}e^{|b|X_1^K} - \log \mathbb{E}e^{|b|\Gamma^1} &= (v_0+v_t)\sum_{m=3}^{\infty}\sum_{n=K+1}^{\infty}\frac{\lambda_n}{\gamma_n^m}(1-R_{n,m})|b|^m \\ &\leq (v_0+v_t)\sum_{m=3}^{\infty}\sum_{n=K+1}^{\infty}\frac{\lambda_n}{\gamma_n^m}|b|^m \\ &= \log \mathbb{E}e^{|b|X_1^K} - \log \mathbb{E}e^{|b|\mathcal{N}^1}\end{aligned}$$

for all large K's. ∎

From the simulation point of view, Theorem 2.2.1 and Proposition 2.3.1 suggest three different approaches. We can simply truncate series expansions of X_i's at some fixed level K, adjust the remaining terms by a normal random variable or by a gamma random variable. But, we take on the last idea as our main approach, namely the gamma approximation. After truncating at level K, we approximate the remaining summation by a single gamma random variable:

$$X_1 \approx \sum_{n=1}^{K}\frac{1}{\gamma_n}\sum_{j=1}^{N_n} Exp_j(1) + \Gamma^1$$

and similarly for X_2 and Z. The gamma approximation has an advantage over a normal approximation because it never generates a negative value in addition to the faster convergence in the sense of Proposition 2.3.1. In the next section, we describe more detailed simulation procedure of each random variable.

Remark It is easy to see from (2.8), (2.9) that X_1 and X_2 are non-negative Lévy processes with time parameters $v_0 + v_t$ and δ, i.e. subordinators. From the series expansions of X_1 and X_2, we get their Lévy densities

$$\rho_1(x) = \sum_{n=1}^{\infty}\lambda_n\gamma_n e^{-\gamma_n x}, \quad \rho_2(x) = \sum_{n=1}^{\infty}\frac{1}{2x}e^{-\gamma_n x}.$$

There are a few simulation methods for Lévy processes or infinitely divisible distributions. Especially for subordinators, a method of Rosiński (see Cont and Tankov 2003) is to construct a series represetation of a subordinator using a function $U(x) := \int_x^{\infty}\rho_i(y)dy$. However, the computation of $U^{-1}(y)$, which is essential in his method, is cumbersome

Table 2.1: Model parameters.

	case I	case II	case III	case IV
κ	0.5	0.3	1	6.2
θ	0.04	0.04	0.09	0.02
σ	1	0.9	1	0.6
ρ	-0.9	-0.5	-0.3	-0.7

Table 2.2: Time for 100 samples when we use $\Phi_1(b)$ and when we insert a dummy $I_\nu(z)$.

$v_0 + v_t$	no $I_\nu(z)$	one $I_\nu(z)$
2	0.57	2.08
0.2	6.67	35.08
0.02	76.59	381.77

because a closed form of $U(x)$ using elementary functions is not available.

On the other hand, Bondesson (1982) proposed a general simulation approach to infinitely divisible distributions using shot noise distributions. This method also derives an infinite series representation of the distribution of interest. But, again, since $\rho_i(x)$'s are infinite sums, we need to truncate the Lévy densities first, and compute the corresponding shot noise representations and truncate them again.

Letting those approaches be open possibilities, we focus on our series expansions of X_1, X_2 and Z, and take a simple approach just by truncating those series at some fixed K.

Throughout the rest of the chapter, four different parameter settings are used given in Table 2.1. The first three cases are taken from Andersen (2005) and case IV is set close to estimated parameters in Duffie et al. (2000). According to those papers, case I is related to long-dated FX options, case II to long-dated interest rate options, case III to equity options and case IV to S&P 500 index options. Andersen (2005) explains the reason for the particular choices as because they are challenging and practically relevant.

2.4 Simulation Procedure

2.4.1 Simulation of X_1

Exact Simulation. As the Laplace transform of X_1 is available in closed form, one can try the exact simulation scheme using the inverse transform method as in Broadie and Kaya (2006). The algorithm consists of two steps; first, generate $U \sim unif[0,1]$ and, second, find $x \geq 0$ such that $\mathbb{P}(X_1 \leq x) = U$. In the second step, there are two iterations involved. One is the root-finding procedure to find x and the other is the calculation of the CDF of X_1. One can use, for example, the algorithm described in Abate and Whitt (1992) to calculate the CDF. In our implementation of this exact method, we use the Abate-Whitt algorithm and `fzero` function in MATLAB for a root-finding procedure. The CDF is calculated by

$$\mathbb{P}(X_1 \leq x) \approx \frac{hx}{\pi} + \frac{2}{\pi}\sum_{k=1}^{N} \frac{\sin hkx}{k} \mathrm{Re}(\Phi^1(-ihk)), \quad h = \frac{2\pi}{x + u_\epsilon} \tag{2.12}$$

where $u_\epsilon = \mu_{X_1} + m\sigma_{X_1}$ with μ_{X_1}, σ_{X_1} the mean and the standard deviation of X_1, and this controls the discretization error and m is set to be not less than 5. Also, the truncation error is handled by stopping the iteration at $k = N$ such that $|\Phi^1(-ihN)|/N < \pi\epsilon'/2$ with $\epsilon' = 10^{-5}$ in this chapter. This discussion followed the implementation of the Abate-Whitt algorithm in Broadie and Kaya (2006).

On the other hand, the main computational load of the Broadie-Kaya scheme is the inclusion of the modified Bessel function of the first kind $I_\nu(z)$ in the Fourier transform of $\left(\int_0^t V_s ds | V_0, V_t\right)$. To see the effect of the inclusion of this Bessel function, we include one dummy calculation of $I_\nu(z)$, which does not affect $\Phi^1(b)$, and compare the results. We record the mean and the variance of X_1,

$$\mu_{X_1} = (v_0 + v_t)\mu^*_{X_1}, \quad \sigma^2_{X_1} = (v_0 + v_t){\sigma^*_{X_1}}^2$$

where $\mu^*_{X_1}$ and ${\sigma^*_{X_1}}^2$ are the mean and the variance of X_1 with $(v_0 = 1, v_t = 0)$:

$$\begin{aligned}
\mu^*_{X_1} &= \frac{1}{\kappa}\coth\left(\frac{\kappa t}{2}\right) - \frac{t}{2}\mathrm{csch}^2\left(\frac{\kappa t}{2}\right), \\
{\sigma^*_{X_1}}^2 &= \frac{\sigma^2}{\kappa^3}\coth\left(\frac{\kappa t}{2}\right) + \frac{\sigma^2 t}{2\kappa^2}\mathrm{csch}^2\left(\frac{\kappa t}{2}\right) - \frac{\sigma^2 t^2}{2\kappa}\coth\left(\frac{\kappa t}{2}\right)\mathrm{csch}^2\left(\frac{\kappa t}{2}\right).
\end{aligned}$$

Note that these can be computed and stored in the initialization of the Monte Carlo simulation if we work with an equidistant time grid and fixed parameters κ, σ.

The initial guess x_0 for `fzero` is set as

$$x_0 = F_N^{-1}(U), \quad F_N(x) = \mathbb{P}(N \le x), \quad N \overset{d}{=} \mathcal{N}\left(\mu_{X_1}, \sigma^2_{X_1}\right)$$

and $x_0 = 0.01 \times \mu_{X_1}$ if $F_N^{-1}(U)$ is negative. The tolerance level is set equal to 10^{-5}. This is same as in Broadie and Kaya (2006).

The results are shown in Table 2.2. Parameters κ, σ from case I in Table 2.1 and time step $t = 1$ are used. The value of $v_0 + v_t$ vary along each simulated path because on each path and at each time grid point a new V_t is generated given V_0. Thus, we choose to take three different levels 0.02, 0.2, 2 for $v_0 + v_t$. The results show that the simulation time is quite sensitive to $v_0 + v_t$ and also to the inclusion of $I_\nu(z)$. This at least gives us a hint about a drawback of the Broadie-Kaya scheme.

Gamma Approximation. We assume that the relevant parameters κ, θ and σ are fixed and that we work with an equidistant time grid. Here we set $t = 1$. Then, $\{\lambda_n\}$ and $\{\gamma_n\}$ can be tabulated in the initialization of the Monte Carlo simulation as well as $\mu^*_{X_1}$, $\sigma^*_{X_1}$, $\mathbb{E}X_1^K$ and $Var(X_1^K)$ for each K. The last two values then determine the shape and the scale parameters of Γ^1. We choose to calculate those numbers for all $K < 100$ because the simulation time of the gamma approximation of X_1 would be too big if $K > 100$.

It turns out that the simulation times are very sensitive to the level of K. Table 2.3 shows the results of gamma approximations with $K = 1$, $K = 20$ with $v_0 + v_t = 0.02$ using MATLAB. The case $K = 20$ takes more than ten times as much time as $K = 1$ case. However, gamma approximations seem to work reasonable well even for small K. Figure 2.1 demonstrates the CDFs of gamma approximations of X_1 for $K = 1$ and $K = 20$, and they are fairly close

Table 2.3: Time for 10,000 samples with $v_0 + v_t = 0.02$ in Gamma approximation

case	K=1	K=20
I	1.66	22.23
II	1.88	23.14
III	1.93	23.18
IV	1.86	24.27

Table 2.4: Time for 10,000 samples with $K = 1$ in Gamma approximation

$v_0 + v_t$	case I
2	3.73
0.2	2.24
0.02	1.86

to each other. On the other hand, Table 2.4 report the simulation times for $K = 1$ with different $v_0 + v_t$ values. The gamma approximation is not as sensitive to $v_0 + v_t$ as to K. Even though we do not report, smaller time steps give better results. Also, we do not report the computation time for $\{\lambda_n\}$, $\{\gamma_n\}$ and others as they can be done very fast.

Beta Approximation. We also test a simple idea that has been applied to the simulation of a random variable with state space $[0, 1]$.

Let us define a random variable X_1^{base} by the Laplace transform: for $b > 0$,

$$\mathbb{E}\exp(-bX_1^{base}) = \exp\left(\frac{v_{base}}{\sigma^2}\left(\kappa\coth\frac{\kappa t}{2} - L\coth\frac{Lt}{2}\right)\right)$$

where $L = \sqrt{2\sigma^2 b + \kappa^2}$ and v_{base} is some fixed positive number. Then, it follows that

$$\mathbb{E}\exp(-bX_1) = \left[\mathbb{E}\exp\left(-bX_1^{base}\right)\right]^{\frac{v_0+v_t}{v_{base}}} = \left[\mathbb{E}\exp\left(-bX_1^{base}\right)\right]^h \prod_{j=1}^{K} \mathbb{E}\exp\left(-bX_1^{base}\right)$$

where $K := \lfloor (v_0 + v_t)/v_{base} \rfloor$ and $h := (v_0 + v_t)/v_{base} - K$. The first component is the Laplace transform of X_1 with $v_{base} \cdot h$ instead of $v_0 + v_t$, say $\tilde{X}_1$, and thus we can write

$$X_1 \overset{d}{=} \tilde{X}_1 + \sum_{j=1}^{K} X_{1,j}^{base}$$

Table 2.5: Time for 10,000 samples in Beta approximation

$v_0 + v_t$	case I	case II	case III	case IV
0.02	0.7	0.69	0.7	0.88
0.05	0.7	0.7	0.72	0.7
0.23	0.71	0.71	0.96	0.74
2.02	0.91	0.92	0.86	1.13

Table 2.6: Time for tabulation with $v_{base} = \theta$ in Beta approximation

case	tabulation
I	6.08
II	4.85
III	2.63
IV	2.51

and here $X_{1,j}^{base}$'s are i.i.d. copies of X_1^{base}.

The idea of the beta approximation is that we simulate X_1^{base} from a pre-calculated table and use a beta random variable to approximate $\tilde{X}_1$. More explicitly, we set

$$\tilde{X}_1 \approx B(hk, (1-h)k) \cdot X_1^{base}, \quad k = \frac{\left(\mathbb{E}X_1^{base}\right)^2}{Var(X_1^{base})}$$

where $B(hk, (1-h)k)$ is an independent beta random variable. This choice of k makes the first two moments of both sides of the approximation coincide. This approach is based on the classical result that the above approximation becomes exact if X_1 (and X_1^{base}, $\tilde{X}_1$) is a gamma random variable, and on the proximity between X_1 and a gamma random variable as observed above. It has been noted in literature that beta distributions can approximate distributions with values in $[0, 1]$ with sometimes great accuracy, e.g. see Springer (1979).

It is clear that the smaller v_{base} is, the better the results as $\tilde{X}_1$ becomes negligible. However, smaller v_{base} increases computation time because K increases. In this chapter, v_{base} is set equal to θ as the simulated v_t values would move around the long run mean θ. Tables 2.5, 2.6 show the simulation times of beta approximations for each case in Table 2.1 and the time for tabulation of X_1^{base} distribution using MATLAB. Beta approximations are faster than gamma approximations, but the tabulation takes much time. However, we note

Table 2.7: Time for tabulation of X_2

case	tabulation
I	1.41
II	1.87
III	0.36
IV	0.13

that this tabulation is done once in the initialization of the Monte Carlo simulation and so this computational burden becomes negligible as the number of simulated paths increases or the time grid becomes more dense. The proximity of the CDFs of beta approximations and the true CDFs is shown in Figure 2.2 for case I. Other three cases reveal a similar level of performance.

2.4.2 Simulation of X_2 and X_3

Simulation of X_2. We can employ all the approaches for the simulation of X_2 as done for X_1. However, there is a simpler method using the tabulation idea. If we fix parameters κ, θ, σ (so fixed δ) and time step t, then the simulation of X_2 does not depend on any intermediate simulated V_t values. Therefore, once we make a distribution table in the initialization of the Monte Carlo simulation, we simply generate a uniform random variable U and get X_2 from the table by the inverse transform method and linear interpolation. For example, we first compute $F_{X_2}(i) = \mathbb{P}(X_2 \leq x_i)$ with

$$x_i = w\mu_{X_2} + \frac{i-1}{M}(u_\epsilon - w\mu_{X_2}), \quad i = 1, \ldots, M+1$$

with $u_\epsilon = \mu_{X_2} + m\sigma_{X_2}$ and w is some small positive number (we set $M = 200$ and $w = 0.01$). Then, second, compute a vector ζ such that

$$\zeta_j = \inf\left\{i : \frac{j-1}{J} \leq F_{X_2}(i)\right\}, \quad j = 1, \ldots, J.$$

This vector helps to identify the index i such that $F_{X_2}(i-1) < U \leq F_{X_2}(i)$ with U drawn from a uniform distribution. After finding this i, the linear interpolation part is straightforward.

Table 2.8: Time for the sequential search method

z	$\delta = 0.1$	$\delta = 2$	$\delta = 6$
1	1.555	1.429	1.303
5	1.461	1.415	1.522
10	1.644	1.553	1.432
50	2.035	1.967	1.878
100	2.425	2.324	2.355
200	3.323	3.194	3.232
300	4.312	4.299	4.238

We set $J = 100$.

However, there is one complication in the computation of the CDF of X_2. In computing (2.9) with $b \in i\mathbb{R}$, MATLAB (and other numerical packages) uses the complex logarithm with the principal branch $(-\pi, \pi]$. This eventually leads to a discontinuity of $\Phi^2(-ib)$ as b moves along the real line and thus to a discontinuous CDF of X_2. This kind of discontinuity is also observed in Broadie and Kaya (2006). Therefore, we need to keep track of how many times $\frac{L}{\kappa} \cdot \frac{\sinh \kappa t/2}{\sinh Lt/2}$ rotates around the origin as b varies. In our implementation, we do this by adding 2π whenever $\Phi^2(-ihk)$ (with h as in the Abate-Whitt algorithm and $k = 1, \ldots, N$ for a truncation level N) crosses the negative real axis, moving from the second quadrant to the third quadrant. Table 2.7 shows the times for tabulation for case I to case IV using MATLAB.

We compute the mean and the variance for a reader's convenience:

$$\mu_{X_2} = \delta \mu^*_{X_2}, \quad \sigma^2_{X_2} = \delta {\sigma^*_{X_2}}^2$$

where $\mu^*_{X_2}$ and ${\sigma^*_{X_2}}^2$ are given by

$$\begin{aligned} \mu^*_{X_2} &= \frac{\sigma^2}{4\kappa^2}\left(-2 + \kappa t \coth\left(\frac{\kappa t}{2}\right)\right) \\ {\sigma^*_{X_2}}^2 &= \frac{\sigma^4}{8\kappa^4}\left(-8 + 2\kappa t \coth\left(\frac{\kappa t}{2}\right) + \kappa^2 t^2 \operatorname{csch}^2\left(\frac{\kappa t}{2}\right)\right). \end{aligned}$$

To compute a precise distribution table, we set $u_\epsilon = \mu_{X_2} + 12\sigma_{X_2}$. Time step is set to be 1.

Simulation of X_3. By the same reason above, we simulate Z from a pre-computed

distribution table. We note that there is no complication of the complex logarithm for Z because (2.10) has the exponent 2. Clearly, we have

$$\mu_Z = 4\mu^*_{X_2}, \quad \sigma^2_Z = 4{\sigma^*_{X_2}}^2.$$

To simulate X_3, we need to generate the Bessel random variable $\eta = BES(\nu, z)$ with

$$\nu = \frac{\delta}{2} - 1, \quad z = \frac{2\kappa/\sigma^2}{\sinh(\kappa t/2)}\sqrt{v_0 v_t}.$$

Since v_0, v_t vary on each simulated path, we generate η at each time grid point on each path. Several authors studied the simulation of Bessel random variables. Devroye (2002) suggested three algorithms using the acceptance–rejection approach by facilitating an upper bound of the probability distribution of a Bessel random variable. Iliopoulos and Karlis (2003) also suggested some acceptance-rejection algorithms, which use properties of Bessel law studied in Yuan and Kalbfleisch (2000). But, we employ a simple sequential search method (Iliopoulos and Karlis 2003 dealt this approach as well), which is based on the following recursive relation:

$$p_{n+1} = \frac{z^2}{4(n+1)(n+1+\nu)}p_n, \quad p_0 = \frac{(z/2)^\nu}{I_\nu(z)\Gamma(\nu+1)}$$

and we return a value $n(U)$ such that

$$\sum_{n<n(U)} p_n < U \leq \sum_{n\leq n(U)} p_n.$$

It turns out that the computing time is not sensitive to δ, but to z-value. As z increases, the computing time increases as well. See Table 2.8. However, the typical z-values that arise in cases I-IV stay small. Indeed, if we set $v_0 = v_t = \theta$, then z-values are 0.16, 0.2, 0.35 and 0.06, respectively.

2.5 Numerical Tests

We apply the beta and the gamma approximations to European call options and compare the results with those of other two methods. The first one is the exact scheme of Broadie and Kaya (2006) and the second one is the QE method of Andersen (2005). Parameters are from Table 2.1. Other parameters are as follows:

S_0	100
strike	100
maturity	1(yr)
v_0	θ

We set $v_{base} = \theta$ for beta approximations and the truncation level $K = 1$ and 10 for gamma approximations.[1]

So far we demonstrated numerical results using MATLAB. However, it becomes too time consuming when it comes to simulation with a large number of trials. From now on, all numerical results are obtained using programs coded in the C programming language and compiled by Microsoft Visual C++ 6.0 in the release mode. Execution files are run on a personal desktop computer with Intel Pentium 4 CPU 3.20 GHz and 1.0GB of RAM. The numbers of sample payoffs are 10K, 40K, 160K, 640K, 2560K and 10240K.

The first comparison In Broadie and Kaya (2006), they compared the exact method with the Euler scheme and found that the exact method exhibits better performance. The Euler scheme, in some cases, is very slow in decreasing the simulation bias. See p.222 of their paper.

The simulation biases of the beta and the gamma approximations are shown in Table 2.10. They are obtained using 1 billion number of simulation trials. The numbers in the parentheses are the standard errors. The starred biases mean they are not statistically significant at the level of two standard deviations.

Figure 2.3 shows the performance of each method. Apparently, as the number of

[1] As for the Broadie-Kaya scheme, Özgur Kaya provided us with the code. We are grateful for his help.

simulation trials increases the bias dominates the RMSE (root mean square error) in the case of beta approximation and gamma approximation with $K = 1$. However, gamma approximation with $K = 10$ achieves the same level of convergence rate of the exact method while shortening the computation time by the factor of 10^2 to 10^3.

The tabulation times for gamma and beta approximations are reported in Table 2.9. As the number of simulation trials increases, the computational burden for tabulation becomes relatively negligible.

The second comparison Andersen (2005) compared his QE method with various discretization methods and showed that the QE method outperforms others. In our numerical tests, we set $\gamma_1 = \gamma_2 = 0.5$, which are parameters used in the QE method, not the first two of $\{\gamma_n\}$, and $\psi_C = 1.5$ (same as in Andersen 2005). We do not implement the martingale correction scheme as we are dealing with the at-the-money options. See Andersen (2005) for the details of the QE method and other variants.

Even though any theoretical convergence rate of the QE method is not given in Andersen (2005), Figure 2.4 shows approximate decay rates of biases. Corresponding numbers are given in Table 2.11 and 1 billion simulation trials are used. Convergence rates are different in different cases; the next table presents the average difference of log biases in each case.

I	−1.28
II	−1.88
III	−1.63
IV	−1.27

Following optimal allocation rule discussed in Duffie and Glynn (1995), this means the convergence rate of the RMSE (when optimally allocated) would be approximately $O(s^{-r})$ with $r = 0.36, 0.39, 0.38$ and 0.36, respectively. Figure 2.5 demonstrates these observations. The dotted lines are the simulation results with time step size 1/8 and 1/32. As one can see, in cases II & III the QE method does better up to 160,000 simulation trials. However, in most of other cases the gamma approximation with $K = 10$ shows a better performance. The gamma

Table 2.9: Computation times for tabulation in Beta and Gamma approximations.

case	I	II	III	IV
Beta	1.06	1.16	0.7	0.64
Gamma	0.69	0.81	0.44	0.36

approximation also has a faster convergence rate. Table 2.12 summarizes simulation results and the numbers in the column for the QE method are the best performing cases.

Additional numerical tests are reported at the end of this chapter. We look at in-the-money (ITM) and out-of-the-money (OTM) European calls with strike 80 and 120, respectively. Figures 2.6 – 2.8 show how much effective the Gamma and Beta approximations are compared to the exact method and the QE method, including the biases of the QE method with different time steps. Corresponding numbers are given in Tables 2.16 – 2.18. Similar figures and tables for OTM calls are also provided.

However, we note that in some cases the Gamma approximation with $K = 10$ performs not as well as at-the-money calls. Those cases are ITM case I, OTM case I, and OTM case II. Especially, the simulation biases reported in Tables 2.13, 2.16 are much bigger than those of the QE method. More numerical tests imply that increasing K does not help to reduce the simulation biases. It turns out that these large biases come from tabulation. In our implementation of the Abate-Whitt algorithm (2.12), we used $u_\epsilon = \mu_V + 12\sigma_V$, $V = X_2, Z$ and set the truncation error $\epsilon' = 10^{-5}$. Instead, we increase u_ϵ to $\mu_V + 14\sigma_V$ and set $\epsilon' = 10^{-7}$. This yields much less simulation biases of the Gamma approximation with $K = 10$ as reported in Table 2.19, which are similar to the biases of the QE method with the time step size 1/32. Figure 2.12 shows the performance of the approximation for those three cases.

Figure 2.1: Convergence of Gamma approximations: $v_0 + v_t = 0.02$.

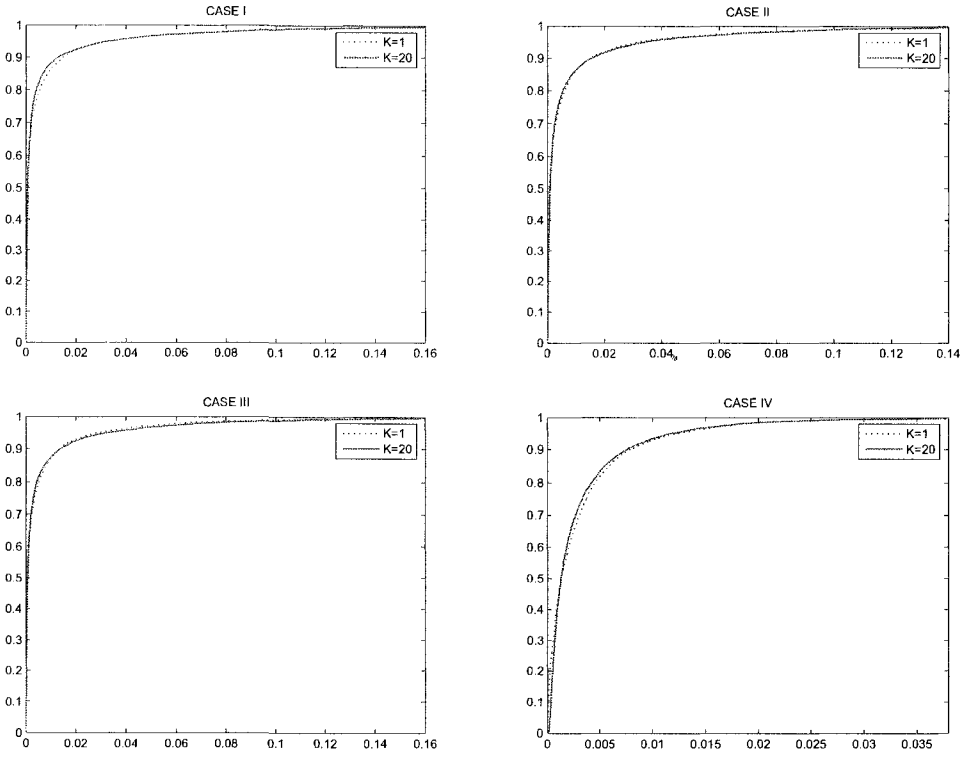


Figure 2.2: Beta approximations (dotted) and true CDFs (solid) : case I

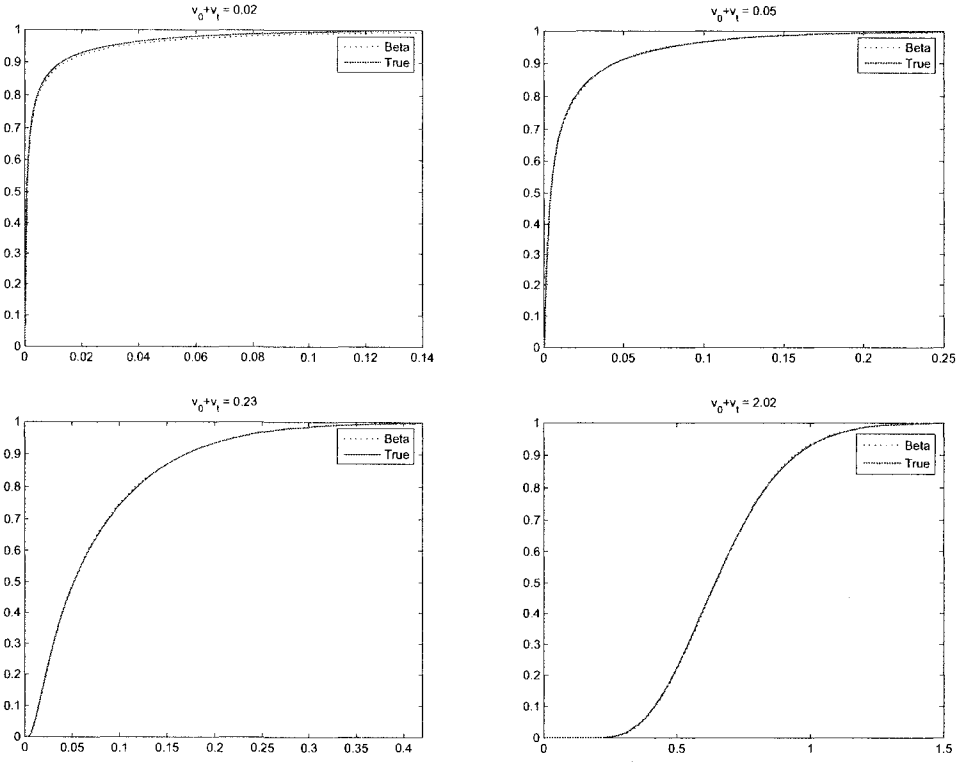


Figure 2.3: Convergence of Simulation Methods for European Call.

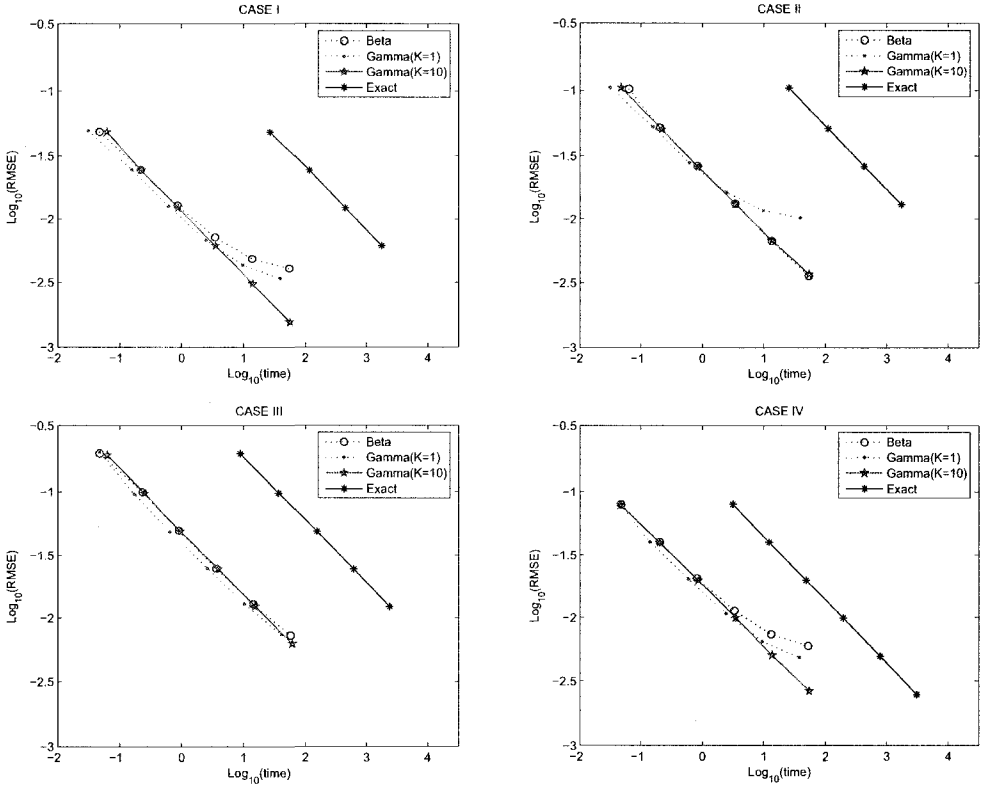


Figure 2.4: Convergence of biases of the QE method.

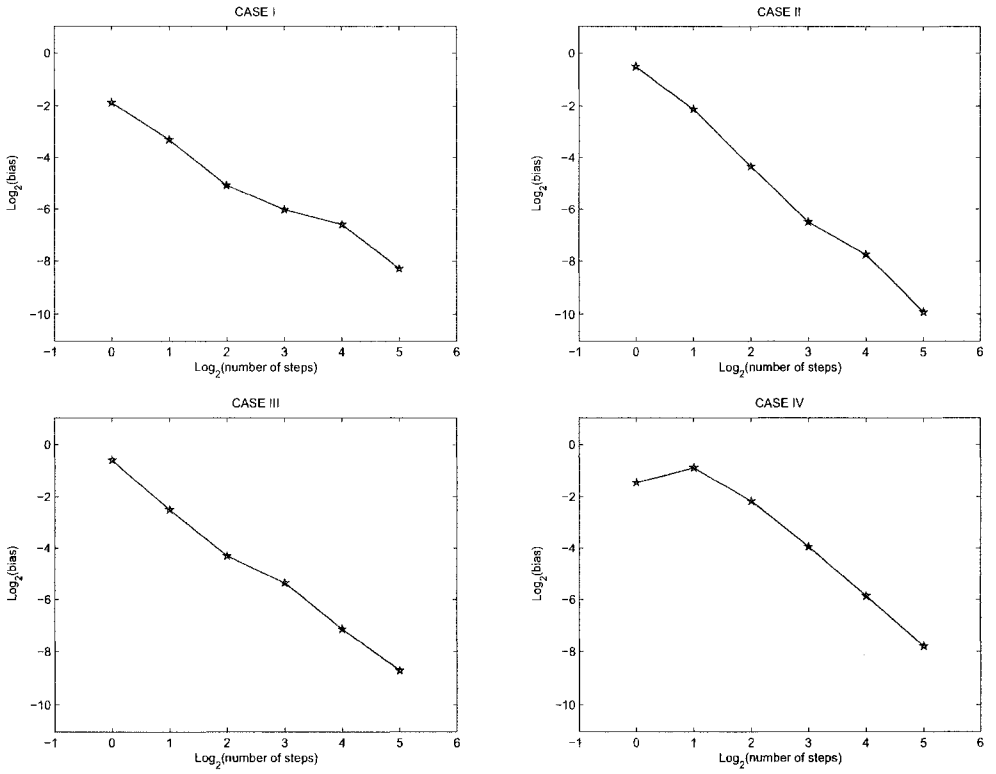


Figure 2.5: Comparison of Gamma approximation with $K = 10$ and the QE method.

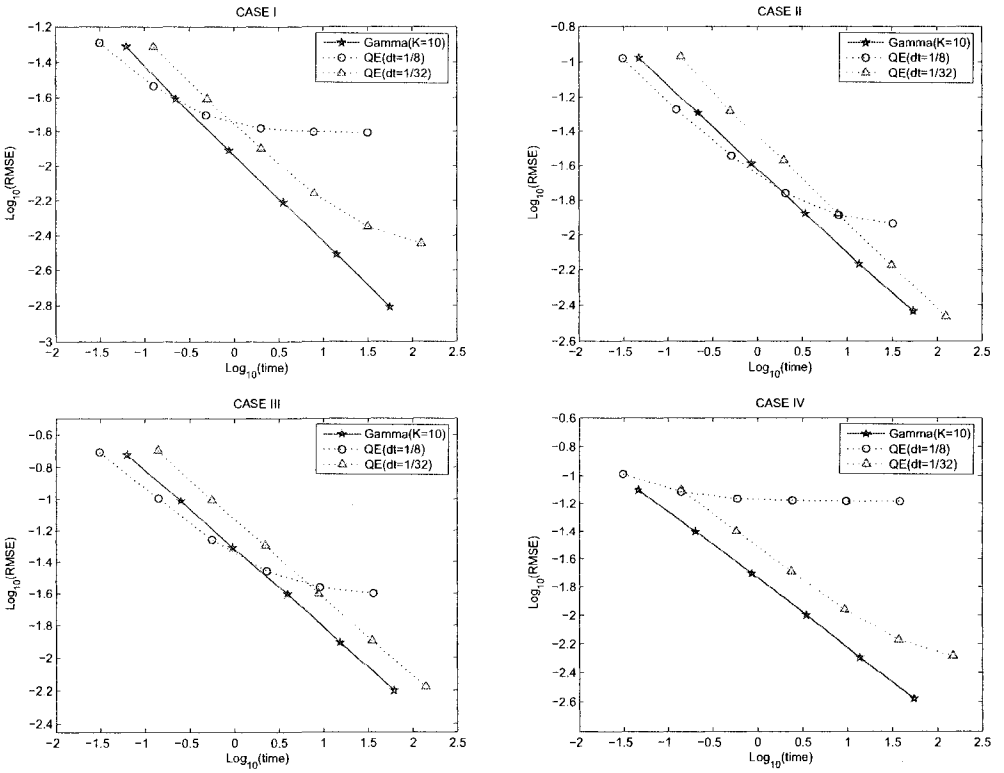


Table 2.10: Biases of the Beta and Gamma approximations (numbers with * are not statistically significant at two standard deviations level)

	case I	case II	case III	case IV
Beta	-0.00375 (0.00016)	0.00140 (0.00033)	-0.00394 (0.00062)	0.00544 (0.00025)
Gamma(K=1)	0.00303 (0.00015)	0.00969 (0.00033)	-0.00409 (0.00062)	0.00417 (0.00025)
Gamma(K=10)	-0.00027* (0.00015)	-0.00164 (0.00033)	-0.00121* (0.00062)	0.00093 (0.00025)

Table 2.11: Biases of the QE method

steps	case I	case II	case III	case IV
1	-0.27287 (0.00018)	0.71166 (0.00033)	0.66345 (0.00061)	0.36716 (0.00026)
2	-0.10063 (0.00016)	0.22785 (0.00033)	0.17704 (0.00062)	0.54419 (0.00026)
4	-0.02954 (0.00016)	0.04896 (0.00033)	0.05043 (0.00062)	0.22331 (0.00025)
8	-0.01540 (0.00016)	0.01108 (0.00033)	0.02422 (0.00062)	0.06485 (0.00025)
16	-0.01034 (0.00016)	0.00463 (0.00033)	0.00709 (0.00062)	0.01704 (0.00025)
32	-0.00322 (0.00015)	0.00104 (0.00033)	0.00237 (0.00062)	0.00451 (0.00025)

Table 2.12: Simulation Results for cases I-IV

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.047	0.0489	0.031	0.0498	0.062	0.0489	8	0.031	0.0513	26.859	0.0486
40,000	0.218	0.0246	0.156	0.0247	0.219	0.0246	8	0.125	0.0291	117.421	0.0244
160,000	0.859	0.0128	0.609	0.0127	0.875	0.0123	8	0.484	0.0197	450.781	0.0122
640,000	3.484	0.0072	2.484	0.0068	3.562	0.0061	32	7.906	0.0069	1775.984	0.0061
2,560,000	13.766	0.0048	9.781	0.0043	14.109	0.0031	32	31.734	0.0044		
10,240,000	54.735	0.0040	38.875	0.0034	56.265	0.0016	32	126.093	0.0036		

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.063	0.1028	0.031	0.1052	0.047	0.1050	8	0.031	0.1039	25.609	0.1047
40,000	0.203	0.0527	0.156	0.0533	0.218	0.0507	8	0.125	0.0532	110.734	0.0515
160,000	0.828	0.0264	0.609	0.0280	0.859	0.0257	8	0.515	0.0285	426.734	0.0262
640,000	3.406	0.0132	2.484	0.0163	3.438	0.0132	8	2.046	0.0172	1735.578	0.0130
2,560,000	13.422	0.0067	9.765	0.0117	13.594	0.0068	16	15.796	0.0080		
10,240,000	53.547	0.0036	38.875	0.0102	54.156	0.0037	32	124.328	0.0034		

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.047	0.1944	0.047	0.1977	0.063	0.1883	8	0.031	0.1950	8.906	0.1937
40,000	0.234	0.0994	0.172	0.0960	0.250	0.0973	8	0.141	0.1006	36.984	0.0980
160,000	0.906	0.0497	0.640	0.0487	0.953	0.0489	8	0.562	0.0549	156.218	0.0489
640,000	3.672	0.0249	2.641	0.0249	3.906	0.0248	16	4.421	0.0255	621.109	0.0245
2,560,000	14.469	0.0129	10.375	0.0130	15.406	0.0124	16	17.546	0.0142	2397.000	0.0123
10,240,000	57.625	0.0073	41.328	0.0074	61.453	0.0063	32	139.484	0.0066		

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.047	0.0794	0.046	0.0788	0.046	0.0784	8	0.031	0.1021	3.140	0.0795
40,000	0.203	0.0400	0.140	0.0401	0.203	0.0394	16	0.296	0.0431	12.265	0.0397
160,000	0.813	0.0206	0.593	0.0202	0.859	0.0198	16	1.172	0.0261	48.765	0.0198
640,000	3.329	0.0113	2.422	0.0107	3.484	0.0099	32	9.312	0.0109	195.296	0.0099
2,560,000	13.172	0.0074	9.500	0.0065	13.718	0.0050	32	37.031	0.0067	779.859	0.0049
10,240,000	52.484	0.0060	37.828	0.0048	54.718	0.0026	32	148.172	0.0051	3106.250	0.0025

2.6 Extension to Non-Equidistant Time Grids

So far, we have described approximate methods for a fixed time step size. This becomes problematic when one needs to simulate stock prices on non-equidistant time grids. To resolve this problem, we introduce a method that is similar to the previous one, but independent of the time step size.

Suppose we need to compute the expectation under a risk neutral measure $\mathbb{Q}$

$$\mathbb{E}^{\mathbb{Q}}\left[p\left(S_0,\ldots,S_m\right)\right]$$

where p is a discounted payoff function depending on stock prices S_i's at times $0 = t_0 < \cdots < t_m = T$ for a given maturity T. From now on, we assume that (2.1)–(2.2) are given under $\mathbb{Q}$ with $\mu = r - d$ risk free interest rate minus dividend rate. We introduce a new process $A_t := V_{4t/\sigma^2}$, which satisfies

$$dA_t = \left(\delta - \frac{4\kappa}{\sigma^2}A_t\right)dt + 2\sqrt{A_t}d\hat{W}_t^1$$

where $\hat{W}_t$ is also a standard Brownian motion under $\mathbb{Q}$. Similarly, define $S'_t = S_{4t/\sigma^2}$. Next, we define a new probability measure $\mathbb{P}$ by

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \exp\left(\int_0^t q\sqrt{A_s}d\hat{W}_t^1 - \frac{q^2}{2}\int_0^t A_s ds\right)$$

with $q = 2\kappa/\sigma^2$. Then, by the Girsanov Theorem, $dW^P := d\hat{W}_t^1 - q\sqrt{A_t}dt$ becomes a Brownian motion under $\mathbb{P}$ and thus

$$dA_t = \delta dt + 2\sqrt{A_t}dW_t^P$$

a squared Bessel process with dimension δ. In summary,

$$\mathbb{E}^{\mathbb{Q}}\left[p\left(S_0,\ldots,S_m\right)\right] = \mathbb{E}^{P}\left[p\left(S'_0,\ldots,S'_m\right)\frac{d\mathbb{Q}}{d\mathbb{P}}\right]$$

where

$$\begin{aligned}
\frac{dS'_t}{S'_t} &= \left(\frac{4\mu}{\sigma^2} + \frac{2q}{\sigma}A_t\right)dt + \frac{2}{\sigma}\sqrt{A_t}\left(\rho dW_t^P + \sqrt{1-\rho^2}d\hat{W}_t^2\right), \\
dA_t &= \delta dt + 2\sqrt{A_t}dW_t^P, \\
\frac{d\mathbb{Q}}{d\mathbb{P}} &= \exp\left(-\int_0^t q\sqrt{A_s}dW_t^P - \frac{q^2}{2}\int_0^t A_s ds\right)
\end{aligned}$$

and $S'_i = S_{4t_i/\sigma^2}$. If we apply the procedure of the Broadie-Kaya scheme, it becomes clear that we need to simulate $\left(\int_s^t A_u du | A_s, A_t\right)$ and all other parts remain same as before except some trivial changes in coefficients.

If we define $A'_u = A_{s+\alpha u}/\alpha$ for some $\alpha > 0$, then A' is still a δ-dimensional squared Bessel process and

$$\left(\int_s^t A_u du \middle| A_s = a_1, A_t = a_2\right) = \left(\alpha^2 \int_0^{(t-s)/\alpha} A'_u du \middle| A'_0 = \frac{a_1}{\alpha}, A'_{(t-s)/\alpha} = \frac{a_2}{\alpha}\right).$$

Therefore, by setting $\alpha = t - s$ we just need to simulate

$$\left(\int_0^1 A'_u du | A'_0 = a_1/(t-s), A'_1 = a_2/(t-s)\right).$$

This can be done by decomposing this conditional distribution into the sum of three independent random variables as done in Theorem 2.2.1 as a straightforward application of the squared Bessel bridge decomposition of Pitman and Yor (1982). The proof of the following theorem is omitted.

Theorem 2.6.1 *The distribution of $\int_0^1 A'_u du$ conditional on endpoints A'_0, A'_1 admits a decomposition:*

$$\left(\int_0^1 A'_u du \mid A'_0 = a'_0, A'_1 = a'_1\right) \stackrel{d}{=} Y_1 + Y_2 + Y_3 \equiv Y_1 + Y_2 + \sum_{j=1}^{\eta'} Z'_j$$

where Y_i's are independent random variables, Z'_j's are i.i.d. copies of Z' and η' is an independent Bessel random variable with parameters $\nu = \delta/2 - 1$ and $z = \sqrt{a'_0 a'_1}$. Moreover, Y_1, Y_2 and Z' have

the following representations:

$$Y_1 = \sum_{n=1}^{\infty} \frac{2}{\pi^2 n^2} \sum_{j=1}^{N_n} Exp_j(1), \quad Y_2 = \sum_{n=1}^{\infty} \frac{2}{\pi^2 n^2} \Gamma_n(\delta/2, 1), \quad Z' = \sum_{n=1}^{\infty} \frac{2}{\pi^2 n^2} \Gamma_n(2, 1)$$

where N_n's are independent Poisson random variables with mean $a'_0 + a'_1$, $Exp_j(1)$'s i.i.d. Exponential random variables with rate 1 and $\Gamma_n(k, \theta)$'s independent gamma random variables with shape parameter k and scale parameter θ.

The series expansion of Y_2 is taken from p.21 of Pitman and Yor (2000), which is a direct consequence of (2.11). We can also derive Lévy densities of Y_1 and Y_2 as done for X_1 and X_2. Now we can apply gamma and beta approximations to $\int_0^1 A'_u du$ and tabulation is free of the time step size. In other words, we make distribution tables once all parameters κ, θ, σ are fixed and use them for any time grid.

Remark Note that $\delta = 4\kappa\theta/\sigma^2$ is the only value used in tabulation which is associated with model parameters. Suppose we use the above series expansion to simulate Y_2 instead of tabulation. In other words, we apply the beta or gamma approximation to Y_2 as well. This means that we are free of any model parameter and the beta or gamma approximation requires once-in-a-lifetime tabulation. This extends the possibility of the proposed approximate schemes to efficient calibration of model parameters to market prices. Detailed investigation in this direction remains as a future research.

Remark One can extend gamma and beta approximations to variants of the Heston model such as the SVJ or SVCJ models. The extensions are straightforward and explained well in Broadie and Kaya (2006), so we do not deal with this issue in this chapter.

2.7 Conclusion

We showed a series expansion of the conditional path integral of the variance process $\left(\int_0^t V_s ds | V_0, V_t\right)$ in the Heston stochastic volatility model. This path integral is decomposed into the sum of three independent random variables and each of them expands as an infinite sum of the Poisson mixture of exponential random variables or an infinite sum of gamma

random variables. Based on this result, we proposed a new Monte Carlo simulation scheme of the Heston model. The basic procedure is same as that of the exact scheme of Broadie and Kaya. However, we simulate $\left(\int_0^t V_s ds | V_0, V_t\right)$ by simulating Poisson, exponential, gamma and Bessel random variables which appear in the series expansions. We also tested the beta approximation which uses a single beta random variable and tabulation of the CDF of some base random variable.

In our implementation, we used the gamma and the beta approximations for X_1, but the tabulation idea is used for the other two random variables X_2 and X_3 as long as the model parameters and the time step size are fixed. This pre-caching needs to be done in the initialization of Monte Carlo simulation. In this pre-caching procedure, one can avoid using a given time step size by facilitating the Girsanov theorem. After some change of measure, it turns out that we only need parameters κ, θ and σ to be fixed. One can apply this approach to the derivatives of which payoffs depend on the time grid with different step sizes.

The numerical results show that the beta and the gamma approximations work better than the exact method, while they exhibit similar performance as Andersen's QE scheme in some cases. However, in all cases considered here, the gamma approximation shows a larger convergence rate.

Figure 2.6: Convergence of Simulation Methods for In-the-money European Call.

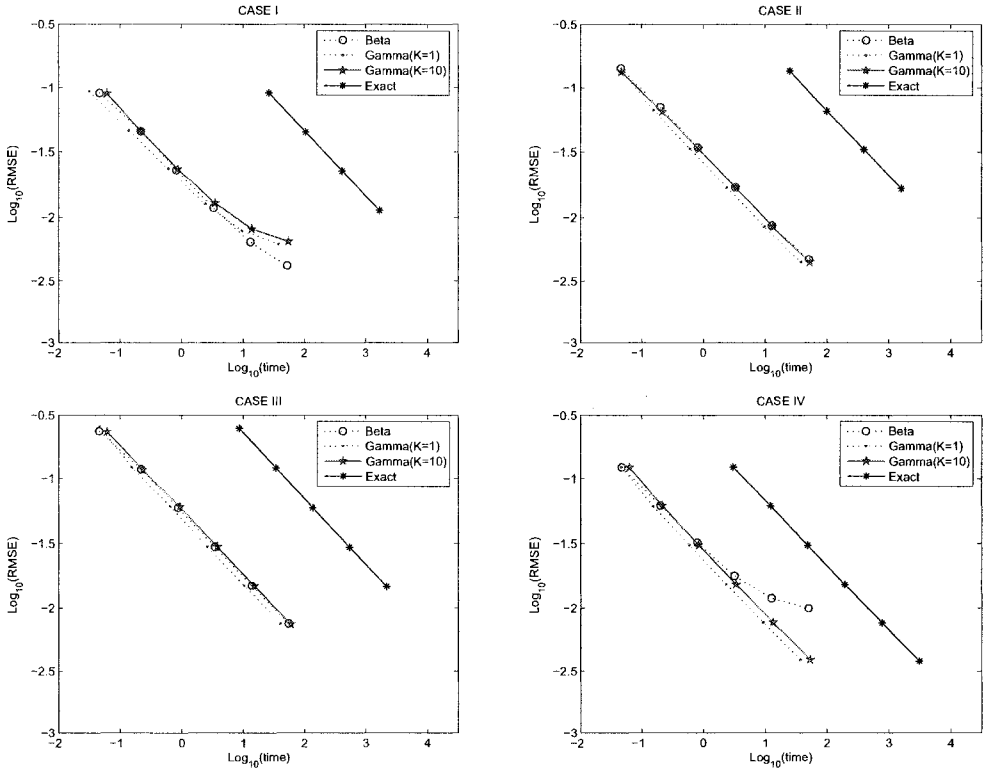


Figure 2.7: Convergence of biases of the QE method for In-the-money Call.

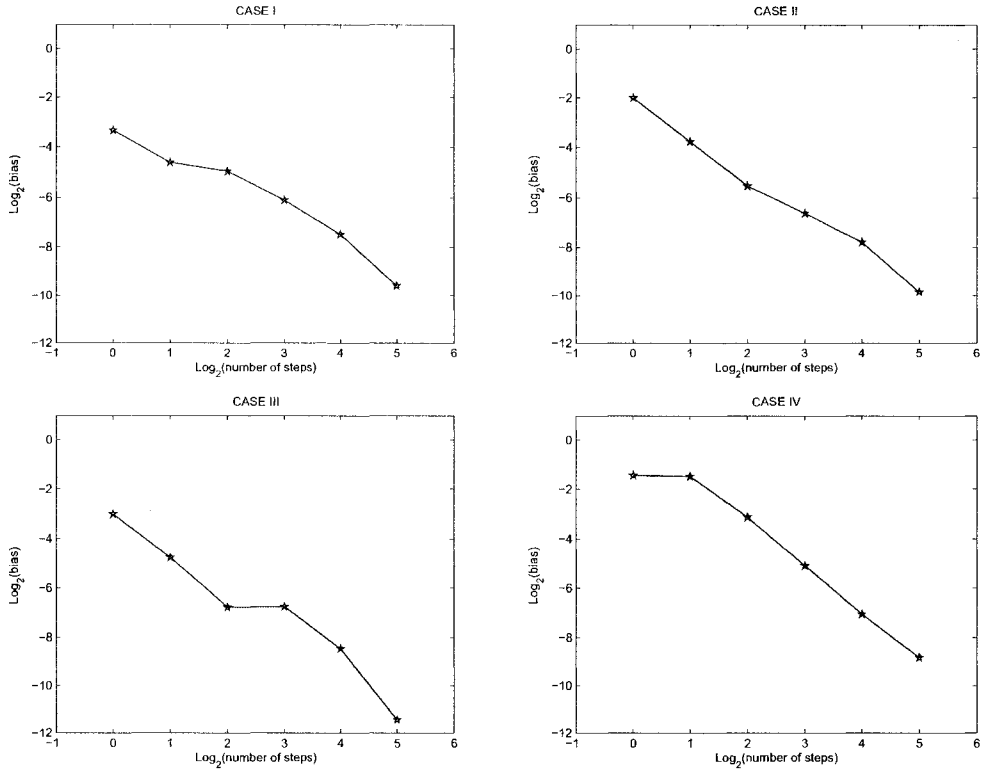


Figure 2.8: Comparison of Gamma approximation with $K = 10$ and the QE method for In-the-money Call.

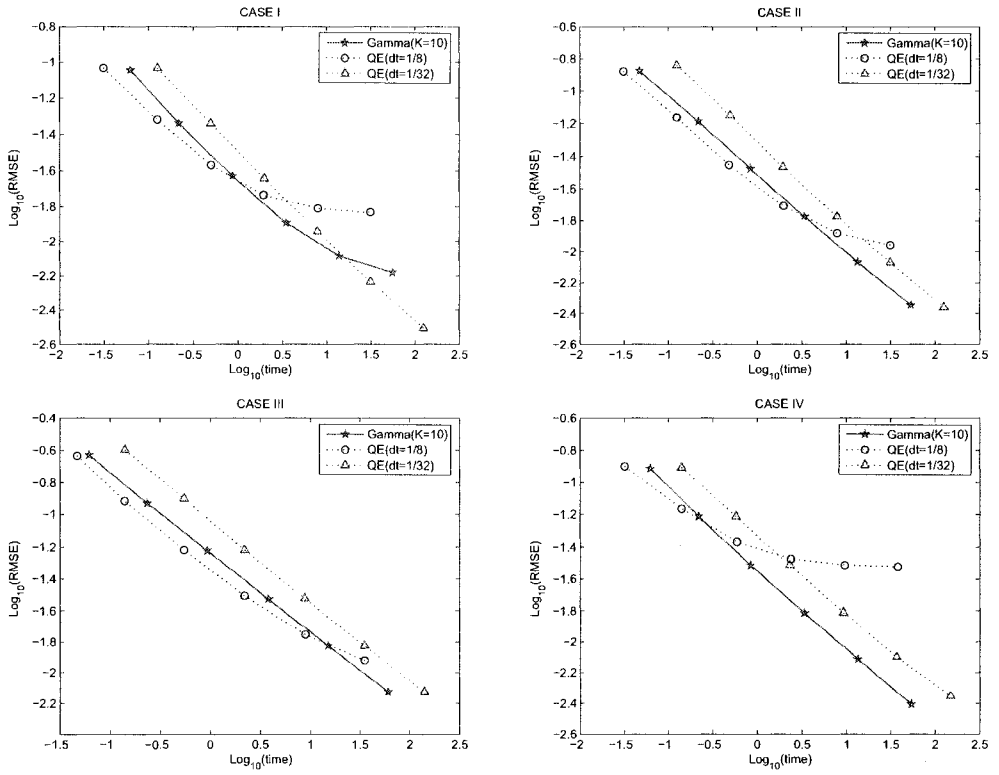


Table 2.13: Biases of the Beta and Gamma approximations for In-the-money Call (numbers with * are not statistically significant at two standard deviations level)

	case I	case II	case III	case IV
Beta	0.00311 (0.00029)	-0.00222 (0.00042)	0.00109* (0.00076)	0.00916 (0.00039)
Gamma(K=1)	0.00558 (0.00029)	-0.00157 (0.00042)	0.00092* (0.00076)	-0.00080 (0.00039)
Gamma(K=10)	0.00592 (0.00029)	-0.00158 (0.00042)	0.00035* (0.00076)	0.00088 (0.00039)

Table 2.14: Biases of the QE method for In-the-money Call.

steps	case I	case II	case III	case IV
1	-0.09749 (0.00029)	-0.24833 (0.00045)	-0.12422 (0.00077)	0.36833 (0.00040)
2	-0.04036 (0.00029)	-0.07305 (0.00043)	-0.03718 (0.00076)	0.35778 (0.00040)
4	-0.03164 (0.00029)	-0.02167 (0.00043)	-0.00918 (0.00076)	0.11472 (0.00040)
8	-0.01440 (0.00029)	-0.01017 (0.00043)	-0.00933 (0.00076)	0.02966 (0.00039)
16	-0.00554 (0.00029)	-0.00449 (0.00043)	-0.00285 (0.00076)	0.00757 (0.00039)
32	-0.00130 (0.00029)	-0.00109 (0.00043)	-0.00036* (0.00076)	0.00221 (0.00039)

Table 2.15: Simulation Results for cases I-IV with Option Strike 80.

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.047	0.0906	0.032	0.0932	0.062	0.0898	2	0.015	0.1002	26.734	0.0904
40,000	0.218	0.0456	0.140	0.0461	0.218	0.0456	2	0.032	0.0612	106.109	0.0451
160,000	0.828	0.0229	0.609	0.0234	0.875	0.0234	8	0.500	0.0269	424.032	0.0226
640,000	3.297	0.0118	2.406	0.0127	3.485	0.0128	16	3.953	0.0126	1695.296	0.0113
2,560,000	13.203	0.0065	9.610	0.0080	13.890	0.0082	32	31.313	0.0058		
10,240,000	52.969	0.0042	38.453	0.0063	55.562	0.0066	32	125.266	0.0031		

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.046	0.1412	0.047	0.1358	0.047	0.1323	4	0.016	0.1420	25.031	0.1349
40,000	0.203	0.0707	0.157	0.0669	0.219	0.0650	4	0.063	0.0715	100.250	0.0658
160,000	0.813	0.0341	0.609	0.0336	0.843	0.0333	4	0.235	0.0403	401.328	0.0331
640,000	3.297	0.0170	2.407	0.0169	3.359	0.0169	8	1.969	0.0196	1605.359	0.0167
2,560,000	12.953	0.0087	9.625	0.0086	13.391	0.0085	16	15.875	0.0095		
10,240,000	51.531	0.0048	38.578	0.0045	53.563	0.0045	32	125.484	0.0043		

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.046	0.2347	0.047	0.2557	0.062	0.2332	2	0.008	0.2573	8.609	0.2461
40,000	0.219	0.1180	0.156	0.1217	0.234	0.1168	2	0.032	0.1278	34.547	0.1204
160,000	0.875	0.0589	0.656	0.0601	0.937	0.0598	4	0.266	0.0608	138.250	0.0592
640,000	3.453	0.0298	2.593	0.0299	3.797	0.0298	4	1.078	0.0313	552.969	0.0296
2,560,000	13.860	0.0150	10.422	0.0150	15.141	0.0149	4	4.250	0.0175	2211.969	0.0149
10,240,000	55.437	0.0075	40.875	0.0075	60.750	0.0075	16	70.750	0.0080		

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.047	0.1218	0.046	0.1221	0.063	0.1216	8	0.032	0.1257	3.031	0.1225
40,000	0.204	0.0617	0.156	0.0609	0.219	0.0612	8	0.141	0.0681	12.141	0.0612
160,000	0.797	0.0319	0.594	0.0306	0.844	0.0306	16	1.172	0.0315	48.547	0.0306
640,000	3.172	0.0178	2.328	0.0153	3.375	0.0153	16	4.704	0.0171	194.109	0.0153
2,560,000	12.656	0.0119	9.297	0.0077	13.469	0.0077	32	37.031	0.0080	776.422	0.0076
10,240,000	50.594	0.0099	37.219	0.0039	53.828	0.0039	32	148.141	0.0044	3106.515	0.0038

Figure 2.9: Convergence of Simulation Methods for Out-of-the-money European Call.

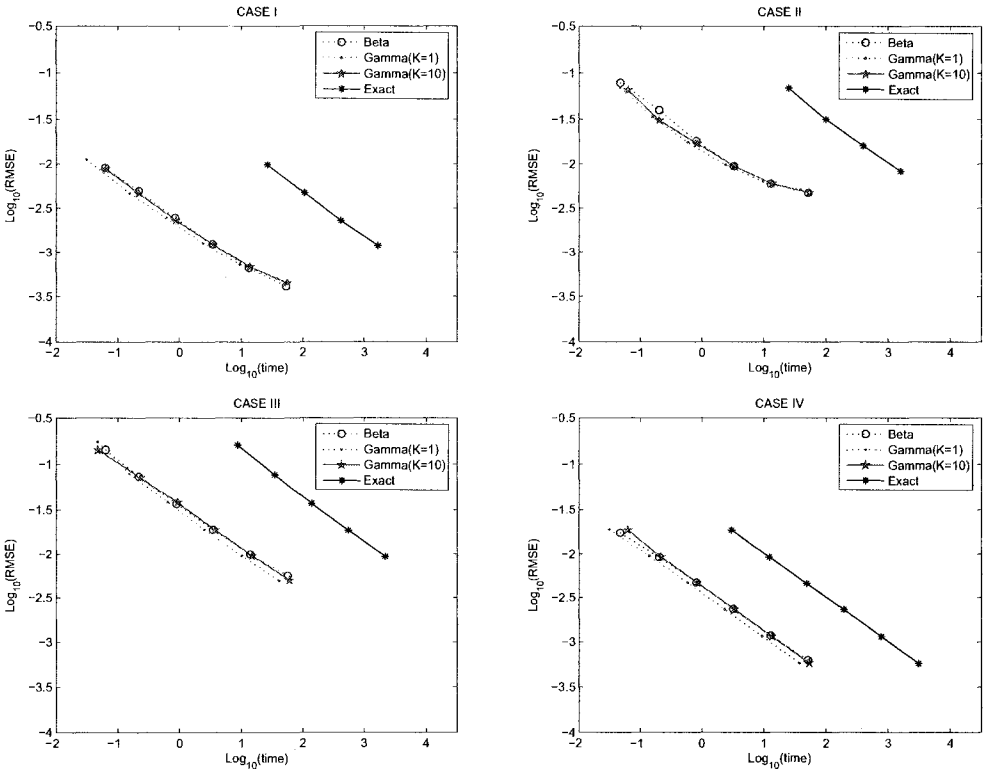


Figure 2.10: Convergence of biases of the QE method for Out-of-the-money Call.

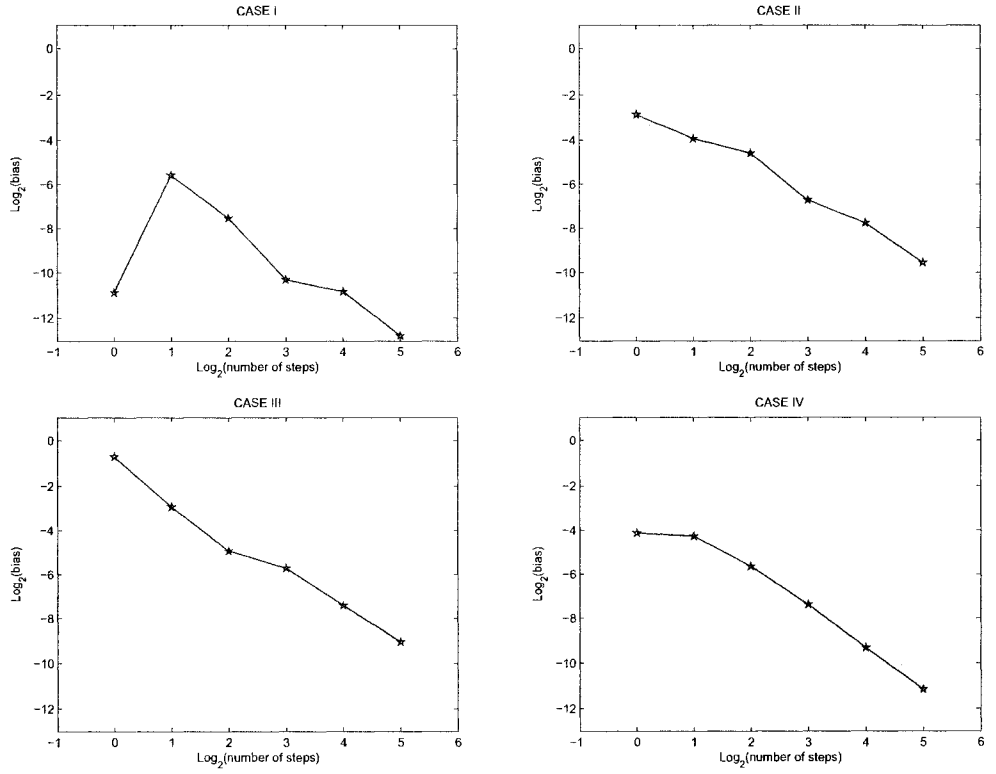


Figure 2.11: Comparison of Gamma approximation with $K = 10$ and the QE method for Out-of-the-money Call.

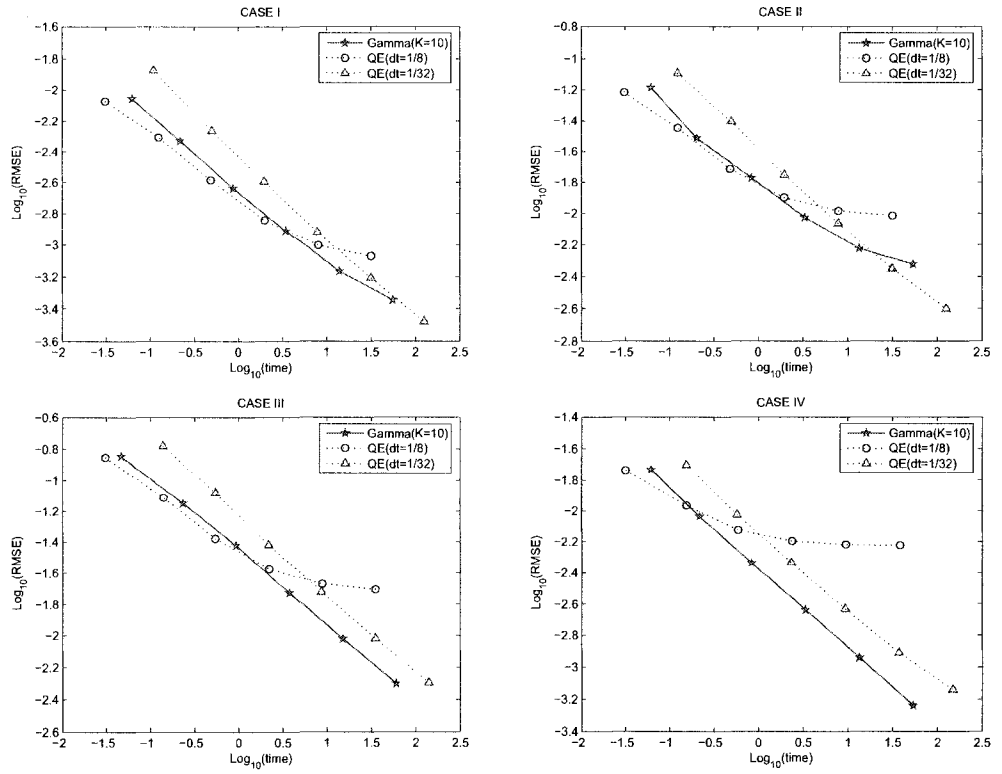


Table 2.16: Biases of the Beta and Gamma approximations for Out-of-the-money Call (numbers with * are not statistically significant at two standard deviations level)

	case I	case II	case III	case IV
Beta	-0.00028 (0.00003)	-0.00419 (0.00021)	-0.00316 (0.00047)	0.00023 (0.00006)
Gamma(K=1)	0.00040 (0.00003)	0.00468 (0.00021)	-0.00149 (0.00047)	-0.00005* (0.00006)
Gamma(K=10)	-0.00034 (0.00003)	-0.00424 (0.00021)	-0.00178 (0.00047)	0.00006* (0.00006)

Table 2.17: Biases of the QE method for Out-of-the-money Call.

steps	case I	case II	case III	case IV
1	0.00053 (0.00002)	0.13745 (0.00016)	0.61036 (0.00044)	0.05741 (0.00006)
2	-0.02091 (0.00003)	-0.06580 (0.00020)	0.13036 (0.00046)	0.05177 (0.00006)
4	-0.00540 (0.00003)	-0.04162 (0.00021)	0.03286 (0.00047)	0.02005 (0.00006)
8	-0.00079 (0.00003)	-0.00944 (0.00021)	0.01910 (0.00047)	0.00592 (0.00006)
16	-0.00055 (0.00003)	-0.00460 (0.00021)	0.00586 (0.00047)	0.00157 (0.00006)
32	-0.00014 (0.00003)	-0.00133 (0.00021)	0.00189 (0.00047)	0.00044 (0.00006)

Table 2.18: Simulation Results for cases I-IV with Option Strike 120.

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.062	0.0090	0.031	0.0112	0.063	0.0088	8	0.031	0.0085	26.640	0.0097
40,000	0.219	0.0049	0.156	0.0047	0.219	0.0047	8	0.125	0.0049	106.063	0.0047
160,000	0.844	0.0025	0.610	0.0024	0.875	0.0023	8	0.485	0.0026	423.984	0.0023
640,000	3.438	0.0012	2.407	0.0013	3.469	0.0012	8	2.000	0.0014	1694.922	0.0012
2,560,000	13.329	0.0007	9.609	0.0007	13.891	0.0007	32	31.531	0.0006		
10,240,000	53.203	0.0004	38.469	0.0005	55.593	0.0005	32	126.047	0.0003		

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.047	0.0779	0.047	0.0692	0.063	0.0651	8	0.031	0.0610	25.031	0.0680
40,000	0.203	0.0397	0.157	0.0336	0.204	0.0306	8	0.125	0.0356	100.266	0.0315
160,000	0.812	0.0181	0.593	0.0174	0.844	0.0169	8	0.484	0.0193	401.594	0.0161
640,000	3.266	0.0094	2.407	0.0097	3.344	0.0094	16	3.984	0.0097	1605.062	0.0083
2,560,000	12.953	0.0060	9.625	0.0063	13.438	0.0060	32	31.516	0.0044		
10,240,000	51.891	0.0047	38.515	0.0051	53.671	0.0047	32	126.094	0.0025		

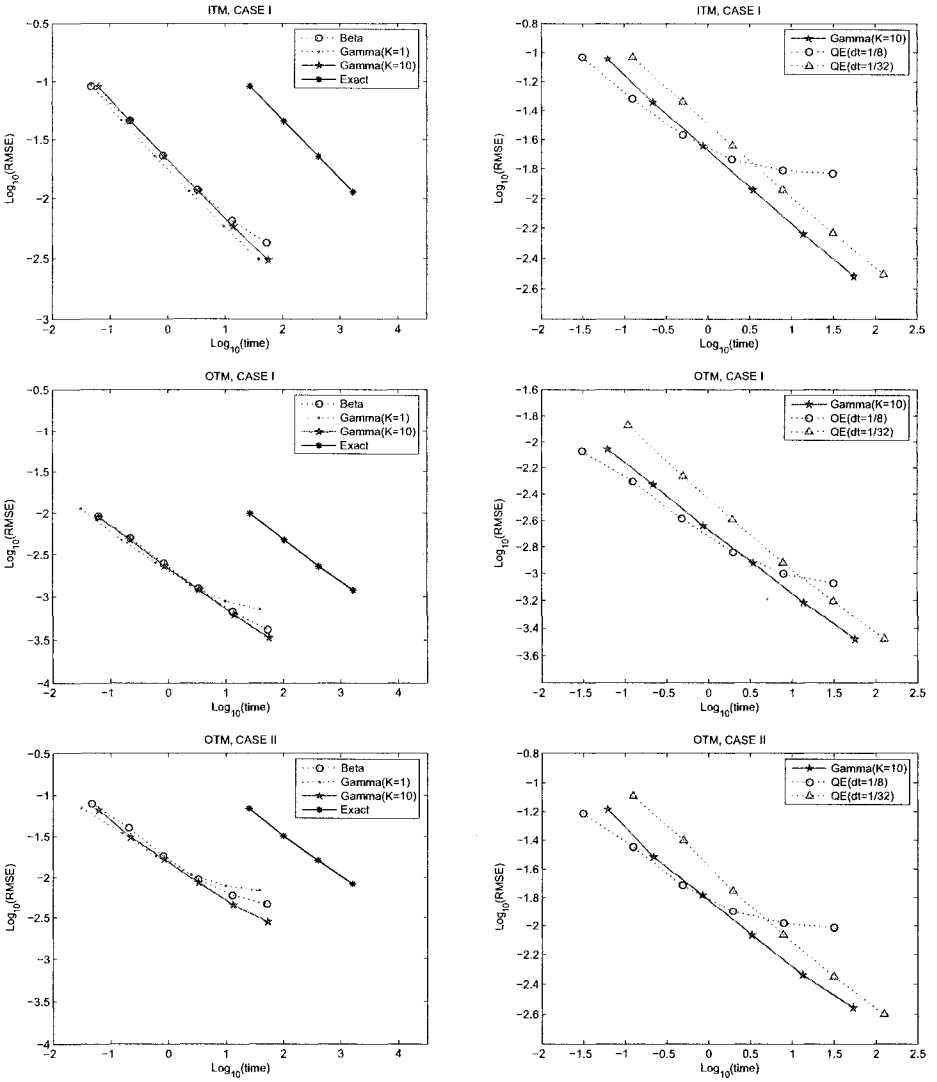
No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.063	0.1424	0.047	0.1733	0.047	0.1419	4	0.015	0.1617	8.625	0.1583
40,000	0.219	0.0728	0.156	0.0782	0.234	0.0711	4	0.063	0.0814	34.532	0.0760
160,000	0.890	0.0365	0.640	0.0382	0.937	0.0376	8	0.546	0.0417	138.234	0.0366
640,000	3.485	0.0188	2.546	0.0188	3.797	0.0187	16	4.469	0.0197	552.859	0.0184
2,560,000	14.062	0.0099	10.234	0.0095	15.156	0.0095	16	17.843	0.0110	2211.125	0.0093
10,240,000	56.156	0.0057	40.890	0.0049	60.687	0.0050	32	141.407	0.0050		

No. of trials	Beta		Gamma(K=1)		Gamma(K=10)			QE		Exact	
	time	RMSE	time	RMSE	time	RMSE	steps	time	RMSE	time	RMSE
10,000	0.047	0.0172	0.031	0.0188	0.062	0.0185	8	0.032	0.0183	3.032	0.0185
40,000	0.203	0.0091	0.141	0.0094	0.219	0.0092	8	0.156	0.0108	12.140	0.0092
160,000	0.813	0.0047	0.593	0.0046	0.844	0.0046	16	1.187	0.0049	48.532	0.0046
640,000	3.203	0.0023	2.328	0.0023	3.375	0.0023	16	4.719	0.0028	194.062	0.0023
2,560,000	12.796	0.0012	9.312	0.0012	13.547	0.0011	32	37.422	0.0012	776.141	0.0011
10,240,000	51.266	0.0006	37.235	0.0006	53.891	0.0006	32	149.828	0.0007	3105.703	0.0006

Table 2.19: Biases of the Gamma approximations with $K = 10$ and more accurate tabulation.

	Tabulation Time	Bias
ITM, CASE I	1.500	0.00105 (0.00029)
OTM, CASE I	1.484	-0.00014 (0.00003)
OTM, CASE II	2.031	-0.00178 (0.00021)

Figure 2.12: Comparison of Gamma approximation with $K = 10$ and the QE method with more accurate tabulation.



Chapter 3

Moment Explosions and Stationary Distributions in Affine Diffusion Models

Many of the most widely used models in finance fall within the affine family of diffusion processes. The affine family combines modeling flexibility with substantial tractability, particularly through transform analysis; these models are used both for econometric modeling and for pricing and hedging of derivative securities. We analyze the tail behavior, the range of finite exponential moments, and the convergence to stationarity in affine models, focusing on the class of canonical models defined by Dai and Singleton (2000). We show that these models have limiting stationary distributions and characterize these limits. We show that the tails of both the transient and stationary distributions of these models are necessarily exponential or Gaussian; in the non-Gaussian case, we characterize the tail decay rate for any linear combination of factors. We also give necessary and sufficient conditions for a linear combination of factors to be Gaussian. Our results follow from an investigation into the stability properties of the systems of ordinary differential equations associated with affine diffusions.

3.1 Introduction

The affine family of diffusion models includes many of the most widely used models in finance. The affine framework offers substantial modeling flexibility and a high degree of tractability, particularly through Laplace or Fourier transforms. Examples of affine diffusions include the Ornstein-Uhlenbeck (OU) process, the square-root diffusion associated with the Cox-Ingersoll-Ross (CIR) interest rate model (Cox et al. 1985), the Heston (1993) stochastic volatility model, the interest rate models of Brown and Schaefer (1994) and Longstaff and Schwartz (1992), and the Duffie and Kan (1996) family of term structure models. Affine models are used both for econometric modeling of time series data and for pricing and hedging of derivative securities.

Duffie et al. (2000) develop a transform analysis for affine jump-diffusions in a very general setting. They derive generalized characteristic functions associated with these models and show that these are exponentials of affine functions of the state variables; the coefficients of these affine functions are characterized as solutions to ordinary differential equations (ODEs). Duffie et al. (2003) characterize regular affine processes and their associated differential equations. Dai and Singleton (2000) define equivalence classes of affine models that are invariant under certain affine transformations, and they define a canonical model within each class. See Singleton (2006) for an extensive discussion of the estimation of these models.

In this paper, we study the tail behavior of affine diffusions and their stationary distributions. We focus on canonical models and show that these models do indeed have limiting stationary distributions. We characterize the tail behavior of the transient and stationary distributions of these models, and we show that the tails are always exponential or Gaussian. This in turn allows us to characterize the range of finite moments for asset price processes constructed from affine diffusions.

We obtain our results through an analysis of the stability of the ODEs that determine the transforms associated with an affine model. To illustrate the connection between tail behavior and transforms, let X be a positive-valued random variable and let $\phi(\theta) = \mathbb{E}[\exp(\theta X)]$ denote its moment generating function (the mapping $\theta \to \phi(-\theta)$ is its Laplace

transform). We can distinguish various types of tail behavior for X based on properties of $\phi(\theta)$ for $\theta \geq 0$: If $\phi(\theta) = \infty$ for all $\theta > 0$, then X is heavy-tailed; if $\phi(\theta)$ is finite all for $\theta \in [0, \theta_0)$, for some $\theta_0 > 0$, then the tail of X is exponentially bounded; if, in addition, $\phi(\theta) = \infty$ for all $\theta > \theta_0$, then the tail is exponentially bounded both above and below, so X has an exponential tail; and if $\phi(\theta) < \infty$ for all $\theta \geq 0$, then X is light-tailed. Similar statements apply to a two-sided random variable through consideration of both positive and negative values of θ. When we refer to the tails of a random vector $X \in \mathbb{R}^n$, we mean the tails of random variables of the form $u \cdot X$, $u \in \mathbb{R}^n$, with $u \cdot X$ denoting the scalar product of u and X.

Consider, now, an OU process

$$dY_t = a(b - Y_t)\,dt + \sigma\,dW_t, \tag{3.1}$$

with $a, \sigma > 0$ and $b \geq 0$, or a CIR process

$$dY_t = a(b - Y_t)\,dt + \sigma\sqrt{Y_t}\,dW_t \tag{3.2}$$

with, in addition, $2ab > \sigma^2$ and $Y_0 > 0$. In either case, take Y_0 fixed, for simplicity. Then, in the case of (3.1), Y_t has a Gaussian distribution for all $t > 0$ and a stationary Gaussian limit distribution as $t \to \infty$; in particular, Y_t has light tails for all t. In the case of (3.2), Y_t has a scaled noncentral chi-square distribution for all $t > 0$ and a stationary limit with a gamma distribution; thus, Y_t has an exponential tail for all t.

Our results extend this simple illustration to the full range of canonical affine models. We establish the existence of limiting stationary distributions, and we show that any linear combination of the state variables has either an exponential tail or a Gaussian distribution. The dynamics of a canonical affine model cannot produce heavy-tailed distributions, nor can they produce non-Gaussian light-tailed distributions; the same holds for any affine model obtained from a canonical model through an affine transformation. As a point of contrast, we note that GARCH models typically generate heavy-tailed marginal distributions, even when driven by light-tailed innovations; see Basrak et al. (2002).

The tail behavior of an affine process determines the maximal moments in an asset-price model constructed from the affine process. More explicitly, suppose the process Y takes values in $\mathbb{R}^n$, and construct a price process $P_t = \exp(a_t + u_t \cdot Y_t)$, where a_t is a scalar function of time, and u_t is an $\mathbb{R}^n$-valued function of time. The points

$$\underline{\theta}_t = \inf\{\theta \in \mathbb{R} : \mathbb{E}[P_t^\theta] < \infty\} \quad \text{and} \quad \overline{\theta}_t = \sup\{\theta \in \mathbb{R} : \mathbb{E}[P_t^\theta] < \infty\}$$

coincide with the endpoints of the interval of convergence of the moment generating function of $u_t \cdot Y_t$. We use the structure of the transform of Y_t to characterize these points. It follows from our investigation that the interval $(\underline{\theta}_t, \overline{\theta}_t)$ shrinks (or, more precisely, does not expand) as t increases. Inverting the dependence on t leads to the smallest t at which $\mathbb{E}[P_t^\theta]$ becomes infinite, for fixed θ. This is the problem of finding the *moment explosion time* studied by Andersen and Piterbarg (2007) in the Heston model. Through results of Lee (2004), the extremal values $\underline{\theta}_t$, $\overline{\theta}_t$ determine the asymptotic slope of the implied volatility curve for options on P_t.

We derive our results through an analysis of the ODEs that arise in the transform analysis of affine models. We show that the moment generating function of $u \cdot Y_t$, $u \in \mathbb{R}^n$, is infinite at θ precisely if the solution to the ODE for Y explodes by time t from initial condition θu. It follows that Y_t has exponential tails if the solution remains finite on $[0, t]$ from all initial conditions in a neighborhood of the origin, and Y_t has light tails if this holds for all initial condition in $\mathbb{R}^n$. The limiting behavior of the distribution of Y_t is determined by the behavior of the ODEs as $t \to \infty$. By characterizing the stability of the ODEs, we show that $\{Y_t, t \geq 0\}$ has a limiting distribution that does not depend on Y_0, and that this limiting distribution is, in fact, stationary for Y. The tails of this stationary distribution are determined by the stability region of the ODE for Y; properties of the stability region are themselves of some interest, as we illustrate through examples. Our final result shows that a linear combination of the components of Y_t is light-tailed only if it is Gaussian, and we characterize which linear combinations have this property through the model parameters defining Y.

The rest of this chapter is organized as follows. Section 3.2 reviews the dynamics and

parametric restrictions for canonical affine models and states our main results. Section 3 illustrates these results with examples. Sections 3.4 to 3.6, develop the analysis and proofs underlying our results. Section 3.4 includes relevant background on the theory of dynamical systems. We conclude in Section 3.7.

3.2 Main Results

The canonical affine models introduced byDai and Singleton (2000) follow equations of the form

$$dY_t = -A^\top(\Theta - Y_t)dt + \sqrt{diag(F_t)}dW_t, \tag{3.3}$$

evolving on $\mathbb{R}^n$ and driven by an n-dimensional standard Brownian motion W. Here, F_t is an affine function of Y_t, also taking values in $\mathbb{R}^n$, and $diag(F_t)$ denotes the $n \times n$ diagonal matrix whose diagonal entries are the components of F_t. The interpretation of the process Y depends on the application. For example, in some models, one defines a short rate process r_t by setting $r_t = u_0 + u_1 \cdot Y_t$, for some $u_0 \in \mathbb{R}$ and some $u_1 \in \mathbb{R}^n$; other models define an asset price process P_t by setting $\log(P_t) = a_t + b_t \cdot Y_t$, for some deterministic functions a and b.

The canonical specification of Dai and Singleton (2000) imposes additional restrictions on (3.3). To state these, we introduce some notational conventions to be used throughout the paper. For vectors or matrices a and b, we write $a \geq b$ if every entry of a is at least as large as the corresponding entry of b; we write $a > b$ if $a \geq b$ and $a \neq b$; and we write $a \gg b$ if every entry of a is strictly larger than the corresponding entry of b. We set $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x \geq 0\}$ and $\mathbb{R}^m_{++} = \{x \in \mathbb{R}^m : x \gg 0\}$, with the dimension of the zero vector determined by context. We write $|x|$ for the Euclidean norm of the vector x.

In the Dai and Singleton (2000) classification, the canonical model $\mathbb{A}_m(n)$ partitions the state vector Y as (Y^v, Y^d), with Y^v evolving on $\mathbb{R}^m_+$ and Y^d on $\mathbb{R}^{n-m}$, as a consequence of restrictions imposed on (3.3). The components of Y^v are called volatility factors, and the components of Y^d are called dependent factors. We use the superscripts v and d more generally to indicate partitions of vectors and matrices associated with the partioning of

Y. Thus, we often write a vector $u \in \mathbb{R}^n$ as (u^v, u^d), with u^v having m components and u^d having $n - m$ components. The parameters of a canonical model $\mathbb{A}_m(n)$ are required to satisfy conditions **(C1)**–**(C4)**, below. Dai and Singleton (2000) and Singleton (2006) explain the econometric identification issues that motivate these conditions.

(C1) The matrix A has the block form

$$A = \begin{pmatrix} A^v & A^c \\ 0 & A^d \end{pmatrix},$$

and it has real and strictly negative eigenvalues.

(C2) The off-diagonal entries of A^v are nonnegative.

(C3) The vector $\Theta = (\Theta^v, \Theta^d)$ has $\Theta^d = 0$, $\Theta^v \geq 0$, and $(-A^\top \Theta)^v \gg 0$.

(C4) The vector $F_t = (F_t^v, F_t^d)$ satisfies

$$F_t^v = Y_t^v, \quad F_t^d = e + (B^c)^\top Y_t^v,$$

where e is a vector of 1s and B^c is a matrix in $\mathbb{R}_+^{m \times (n-m)}$.

The eigenvalue condition in **(C1)** ensures mean reversion in Y. It implies (through, e.g., p.62 of Horn and Johnson 1990) that A^v and A^d also have strictly negative eigenvalues, in view of the block triangular form of A. Together, **(C1)** and **(C2)** imply that $-A^v$ is an M-matrix (as defined, e.g., in Berman and Plemmons 1994). The vector Θ represents the long-run mean of Y. We could rewrite (3.3) in terms of

$$\Lambda = -A^\top \Theta. \tag{3.4}$$

Indeed, if we specify Λ^v rather than Θ, with $\Lambda^v \gg 0$, then the fact that $-A^v$ is an M-matrix guarantees (see p.137 of Berman and Plemmons 1994: inverse-positivity of M-matrix) that we can find a $\Theta^v \geq 0$ for which $-A^{v\top} \Theta^v = \Lambda^v$; in fact, we can take $\Theta^v = -(A^{v\top})^{-1} \Lambda^v$. If we then set $\Theta^d = 0$ and $\Lambda^d = (-A^\top \Theta)^d = (A^{c\top})(A^{v\top})^{-1} \Lambda^v$, we complete the specification of Λ in

a manner consistent with **(C3)** and (3.4). Thus, we can choose either Θ or Λ in specifying the model.

Condition **(C4)** requires that only the volatility factors Y^v appear inside the square root in (3.3), which is natural, given that the components of Y^d will be allowed to become negative. The form of F_t^v implies that the volatility factors are correlated only through the matrix A in the drift of Y. Cheridito et al. (2006) show that the diffusion matrix of any affine diffusion on $\mathbb{R}_+^m \times \mathbb{R}^{n-m}$ can be diagonalized through an affine transformation if $m \leq 1$ or $m \geq n-1$ (in particular, if $n \leq 3$); but they also provide examples for which no such transformation exists.

To illustrate this modeling framework, we formulate a stochastic volatility model in the class $\mathbb{A}_1(2)$ — that is, a two-factor model with a single volatility factor. We write the state vector as $Y = (Y^v, Y^d)$, with dynamics

$$dY_t^v = (m_1 + pY_t^v)dt + \sqrt{Y_t^v}dW_t^1 \tag{3.5}$$

$$dY_t^d = (m_2 + qY_t^v + rY_t^d)dt + \sqrt{1 + sY_t^v}dW_t^2, \tag{3.6}$$

for some constants m_1, m_2, p, q, r and s. The restrictions of the general model $\mathbb{A}_m(n)$ require $m_1 > 0$, $p < 0$, $q \geq 0$, $r < 0$, $s \geq 0$, and $qm_1 = pm_2$. We can then construct an asset-price process P_t by setting

$$\log(P_t) = a_t + 2b_tY_t^v + 2c_tY_t^d, \tag{3.7}$$

for some deterministic functions a_t, b_t and c_t. We will apply our general results to the moments of P_t in the next section and illustrate the qualitatively different behavior produced by different ranges of parameter values in the model.

The model (3.3) has associated with it a system of ODEs on $\mathbb{R}^n$ specified by

$$\begin{pmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix} = \begin{pmatrix} A^v & A^c \\ 0 & A^d \end{pmatrix} \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} + \begin{pmatrix} I & B^c \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1^2(t) \\ \vdots \\ x_n^2(t) \end{pmatrix}. \tag{3.8}$$

We will write this system more compactly as

$$\dot{x} = f_0(x) = Ax + B(x_1^2, \ldots, x_n^2), \quad x(0) = u, \tag{3.9}$$

with B the corresponding block matrix in (3.8), and the initial condition $u \in \mathbb{R}^n$ included here for future reference. We will see that, for any initial condition u, the system (3.9) admits a unique solution on a time interval $[0, t)$, for some $t > 0$. But, the solution may blow up in finite time and fail to exist beyond some finite time τ. We discuss this point in greater detail in Section 3.4.1.

The analysis in Duffie et al. (2000) leads to the representation

$$\mathbb{E}\left[\exp(2u \cdot Y_t)\right] = \exp\left(2\int_0^t \Lambda \cdot x(s)ds + 2\int_0^t |x^d(s)|^2 ds + 2x(t) \cdot Y_0\right), \tag{3.10}$$

with x solving (3.9) and Λ as in (3.4), at least under some regularity conditions. Our first result asserts the validity of this formula (even in the infinite case) without further conditions and adds a stronger conclusion:

Theorem 3.2.1 *The transform formula (3.10) holds in the sense that if either side is well-defined and finite, then the other is also finite and equality holds. Moreover, the right side of (3.10) is well-defined and finite if and only if the solution of (3.9) exists at time t. Consequently, for any $t \geq 0$, the right side of (3.10) is finite for any vector u in a neighborhood of the origin.*

This result connects the stability of the ODE (3.9) with the tail behavior of Y_t:

Corollary 3.2.1 *Consider the system in (3.9) with initial condition $x(0) = \theta u/2$, $\theta > 0$. If the solution x exists at t, then*

$$\limsup_{y\to\infty} \frac{1}{y} \log \mathbb{P}(u \cdot Y_t > y) \leq -\theta.$$

If the solution explodes before t, then

$$\limsup_{y\to\infty} \frac{1}{y} \log \mathbb{P}(u \cdot Y_t > y) \geq -\theta.$$

For any $t \geq 0$, the solution x exists at t for all sufficiently small $|\theta| > 0$.

Corollary 3.2.1 describes the tail behavior of Y_t: the last statement of the corollary and the first limsup together imply that for any u and any $\epsilon > 0$, we have

$$\mathbb{P}(u \cdot Y_t > y) \leq e^{-(\theta-\epsilon)y},$$

for some $\theta > 0$ and all sufficiently large y. Thus, $u \cdot Y_t$ has an exponentially bounded right tail and, with an obvious modification to the argument, an exponentially bounded left tail as well.

A further consequence of Theorem 3.2.1 is a comparison of the tails of the volatility factors of two models. For processes Y^1 and Y^2 on $\mathbb{R}^m_+$, if $\mathbb{E}\exp(u \cdot Y^1_t) \geq \mathbb{E}\exp(u \cdot Y^2_t)$ for all $u \in \mathbb{R}^m_+$, then Y^1_t has heavier tails than Y^2_t. We give conditions for such a comparison for processes in $\mathbb{A}_m(m)$.

Corollary 3.2.2 *Let Y^i be a process in $\mathbb{A}_m(m)$ with parameters A^i and Λ^i, $i = 1, 2$.*

1. *Suppose $A^1 = A^2$ and $Y^1_0 = Y^2_0$; then $\mathbb{E}\exp(2u \cdot Y^1_t) \geq \mathbb{E}\exp(2u \cdot Y^2_t)$ for all $u \in \mathbb{R}^m_+$ and $t \geq 0$ if and only if $\Lambda^1 \geq \Lambda^2$.*

2. *Suppose $\Lambda^1 = \Lambda^2$ and $Y^1_0 = Y^2_0 = Y_0$; then $\mathbb{E}\exp(2u \cdot Y^1_t) \geq \mathbb{E}\exp(2u \cdot Y^2_t)$ for all $(u, Y_0) \in \mathbb{R}^m_+ \times \mathbb{R}^m_+$ and $t \geq 0$ if and only if $A^1 \geq A^2$.*

Our next result considers the limit as $t \to \infty$. Define the stability region S of the ODE (3.9) to be the set of initial conditions u for which the solution $x(t)$ exists for all $t \geq 0$ and $\lim_{t\to\infty} x(t) = 0$ if $x(0) = u$.

Theorem 3.2.2 *The process Y has a unique stationary distribution, which is also the limiting distribution of Y_t, as $t \to \infty$, for any Y_0. Moreover, if Y_∞ has the stationary distribution of Y and we define*

$$\mathcal{S} = \{u \in \mathbb{R}^n : \mathbb{E}\exp(2u \cdot Y_\infty) < \infty\},$$

then $\mathcal{S}$ coincides with S, the stability region of the system (3.9). This set contains a neighborhood of the origin.

By arguing as in Corollary 3.2.1, we conclude that $u \cdot Y_\infty$ has exponentially bounded tails for all $u \in \mathbb{R}^n$. As a consequence of our analysis, we will identify the distribution of Y_∞ through its moment generating function.

Theorems 3.2.1 and 3.2.2 preclude the possibility of heavy tails for Y_t and Y_∞ — any linear combination of the components of Y_t or Y_∞ has tails that are bounded by some exponential decay. We turn next to the possibility of light tails — tails that decay faster than any exponential. The Gaussian subfamily of canonical affine models (which corresponds to taking $m = 0$ and thus removing all volatility factors) demonstrates that such light-tailed models are indeed possible within the canonical affine framework. Our next result shows that the Gaussian case is the *only* light-tailed case among canonical models. More precisely, we show that if the moment generating function of $u \cdot Y_t$ is finite for all $\theta \in \mathbb{R}$, then the distribution of $u \cdot Y_t$ is Gaussian.

Before stating the theorem, we review some facts from linear algebra. By choosing an appropriate basis, we can transform A^d into a Jordan canonical form; in other words, there exists an invertible matrix P such that $P^{-1}A^dP = J$, and J is a block diagonal Jordan matrix. (The columns of P are eigenvectors or generalized eigenvectors of A^d.) Let $\lambda_1, \ldots, \lambda_k$ denote the distinct eigenvalues of A^d, and let a_{λ_i} denote the algebraic multiplicity of λ_i, which is the multiplicity of $(x - \lambda_i)$ in the characteristic polynomial of A^d. The matrix J can then be chosen to have k diagonal blocks of the form $\lambda_i I_i + N_i$, $i = 1, \ldots, k$, with I_i the identity matrix and N_i a nilpotent matrix, both of dimension $a_{\lambda_i} \times a_{\lambda_i}$. The entries of N_i immediately above its main diagonal take the values 0 or 1, and all other entries of N_i are equal to 0.

We introduce a special matrix W to state our last theorem. For this, we select the q-th row of P if there exists some p with $B^c_{pq} \neq 0$, $q = 1, \ldots, n - m$. Denoting the row vectors thus

extracted from P by $w_1, \ldots, w_l$, we define

$$W := \left(\begin{array}{c} w_1 \\ \vdots \\ w_l \\ \hline \\ A^c P \end{array} \right) = \left[\begin{array}{c|c|c} W_1 & \cdots & W_k \end{array} \right].$$

In the block decomposition on the right, W_1 consists of the first a_{λ_1} columns of W, W_2 consists of the next a_{λ_2} columns, and so on. Similarly, we define

$$\tilde{u} := P^{-1} u^d = \left(\begin{array}{c} \tilde{u}^1 \\ \vdots \\ \tilde{u}^k \end{array} \right), \quad \tilde{u}^i \in \mathbb{R}^{a_{\lambda_i}}.$$

Theorem 3.2.3 *Assume that a Jordan canonical form J of A^d is given as above. Then for any given $t > 0$ and $u \in \mathbb{R}^n$, the following holds: $\mathbb{E} \exp(2\theta u \cdot Y_t) < \infty$ for all $\theta \in \mathbb{R}$ if and only if $u^v = 0$ and*

$$W_i N_i^l \tilde{u}^i = 0, \quad l = 0, \ldots, a_{\lambda_i} - 1, \quad i = 1, \ldots, k. \tag{3.11}$$

Moreover, $u \cdot Y_t$ has a Gaussian distribution if and only if these conditions hold.

Since the multiplicities of the roots of the characteristic polynomial of A^d are sensitive to the coefficients of the polynomial, small changes in the entries of A^d can make it diagonalizable. For diagonalizable A^d, (3.11) reduces to

$$W_i \tilde{u}^i = 0, \qquad i = 1, \ldots, k. \tag{3.12}$$

Conditions (3.11) and (3.12) may seem surprisingly complicated, but we will illustrate their significance and application through examples in the next section. A more intuitive approach to checking whether a linear combination of factors has a Gaussian distribution would be to check if each of the factors is Gaussian; individual factors might then be

checked recursively, as follows: no volatility factor is Gaussian, no dependent factor that has a volatility factor in its drift or diffusion coefficient is Gaussian, no dependent factor that has a non-Gaussian dependent factor in its drift is Gaussian, and so on. Our examples will show that this approach cannot cover all cases because of special cancellations that can occur; nevertheless, Theorem 3.2.3 does support sufficient conditions of this type, as we will show in the next corollary. These conditions become necessary when each eigenvalue of A^d has a geometric multiplicty of 1, a restriction that effectively rules out certain cancellations. The geometric multiplicity g_{λ_i} of an eigenvalue λ_i is the dimension of the eigenspace associated with λ_i.

We make precise the recursive procedure sketched above through a directed graph G on the coordinates of the dependent factors. Introduce an edge $(i, i+i)$ in G if $J_{i,i+1} = 1$. Call a node j of the graph *restricted* with respect to a matrix M if $M_{ij} \neq 0$ for some i. Extend this property to other nodes by saying that j is restricted if it is reachable from a restricted node through a directed path in G. For any matrix D, let $\mathbf{1}_D$ denote the matrix with $(\mathbf{1}_D)_{ij} = 1$ if $D_{ij} \neq 0$ and 0 otherwise.

Corollary 3.2.3 *A sufficient condition for (3.11) is that $\tilde{u}_j = 0$ for all j restricted with respect to $\mathbf{1}_{A^c P} + \mathbf{1}_{B^c}\mathbf{1}_P$. This condition becomes necessary if $g_{\lambda_i} = 1$ for all $i = 1, \ldots, k$.*

3.3 Examples and Applications

3.3.1 Stochastic Volatility: A Simple Case

To illustrate our results, we begin by considering the stochastic volatility model (3.5)–(3.7), based on the $\mathbb{A}_1(2)$ dynamics in (3.5)–(3.6). Through (3.10), moments of P_T are given by

$$\mathbb{E}[P_T^\theta] = \exp\left(a_T\theta + 2\int_0^T (m_1 x_1(t) + m_2 x_2(t))\,dt + 2\int_0^T x_2(t)^2 dt + 2(x_1(T)Y_0^v + x_2(T)Y_0^d)\right), \tag{3.13}$$

where (x_1, x_2) solves the ODE

$$\dot{x}_1 = px_1 + qx_2 + x_1^2 + sx_2^2, \quad \dot{x}_2 = rx_2, \tag{3.14}$$

with initial condition $(x_1(0), x_2(0)) = (\theta b_T, \theta c_T)$.

We begin with the simple case $q = s = 0$, in which the ODE for x_1 reduces to a scalar quadratic differential equation. We digress briefly to record properties of this scalar system because it will be an important tool at several points in our analysis.

Consider, then, the scalar quadratic ODE $\dot{x} = \alpha x^2 + \beta x + \gamma$, with $\alpha > 0$. Let $D = \beta^2 - 4\alpha\gamma$, and denote by η_1 and η_2 the two solutions of $\alpha x^2 + \beta x + \gamma = 0$. The following properties of the solution x, which are easily derived from its closed form, are also used in Andersen and Piterbarg (2007). If $D > 0$ with $\eta_1 < \eta_2$, then

$$x(t) \to \eta_1 \text{ as } t \to \infty, \text{ if } x(0) < \eta_2;$$
$$x(t) \equiv \eta_1 \text{ or } \eta_2, \text{ if } x(0) = \eta_1 \text{ or } \eta_2, \text{ respectively;}$$
$$x(t) \to \infty \text{ as } t \to \tau, \text{ if } x(0) > \eta_2,$$

with

$$\tau = \frac{1}{\alpha(\eta_2 - \eta_1)} \log \frac{x(0) - \eta_1}{x(0) - \eta_2}. \tag{3.15}$$

If $D = 0$, then

$$x(t) \to -\frac{\beta}{2\alpha} \text{ as } t \to \infty, \text{ if } x(0) < -\beta/2\alpha;$$
$$x(t) \equiv -\frac{\beta}{2\alpha}, \text{ if } x(0) = -\beta/2\alpha;$$
$$x(t) \to \infty \text{ as } t \to \tau, \text{ if } x(0) > -\beta/2\alpha,$$

with

$$\tau = \frac{1}{x(0) - \beta/2\alpha}$$

If $D < 0$, then

$$x(t) \to \infty \text{ as } t \to \tau = \frac{1}{\sqrt{-D}} \left(\pi - 2 \tan^{-1} \frac{2\alpha x(0) + \beta}{\sqrt{-D}} \right).$$

These cases are illustrated in Figure 3.1. Consider, in particular, the first case, $D > 0$. The two roots are equilibrium points — points at which $\dot{x} = 0$. The root η_1 is a stable equilibrium for the ODE; $x(t)$ moves toward η_1 from any initial condition less than η_1 or

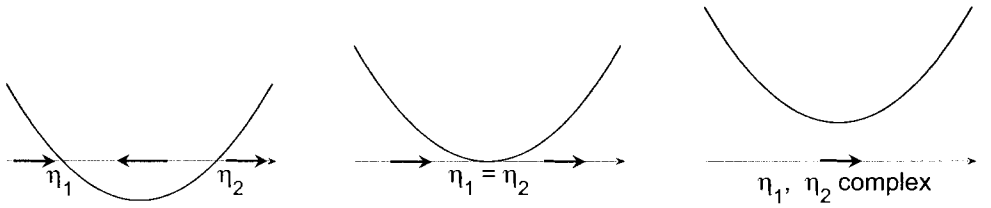


Figure 3.1: Qualitative behavior of $\dot{x} = \alpha x^2 + \beta x + \gamma$ with equilibria η_1, η_2

between the two roots, so the stability region for the system is

$$S = \{x : x < \eta_2\}.$$

In contrast, η_2 is an unstable equilibrium, and x blows up in finite time τ if $x(0) > \eta_2$. The set S_T consists of all initial conditions from which x continues to exist throughout $[0, T)$. From the expression for the explosion time τ in (3.15), we find that

$$S_T = \{x : x \leq (\eta_2 e^{\alpha T(\eta_2 - \eta_1)} - \eta_1)/(e^{\alpha T(\eta_2 - \eta_1)} - 1)\}.$$

We can now apply this to (3.13). In the case $q = s = 0$, the solution x_1 in (3.14) becomes infinite at $\tau = (\log(\theta b_T + p) - \log(\theta b_T))/p$, if $\theta b_T > -p$; otherwise, $x_1(t)$ is finite for all t and converges exponentially to zero. In other words, if $\theta b_T < -p/(1 - e^{pT})$, then the right side of (3.13) is finite; the second coordinate x_2 is always finite and integrable. We therefore conclude that

$$\sup\{\theta : \mathbb{E}[P_T^\theta] < \infty\} = \begin{cases} \frac{-p}{b_T(1-e^{pT})}, & \text{if } b_T > 0; \\ \infty, & \text{if } b_T \leq 0; \end{cases}$$

$$\inf\{\theta : \mathbb{E}[P_T^\theta] < \infty\} = \begin{cases} -\infty, & \text{if } b_T \geq 0; \\ \frac{-p}{b_T(1-e^{pT})}, & \text{if } b_T < 0. \end{cases}$$

We can illustrate these properties through the following sets:

$$\begin{aligned} S &= \{(x,y) : \lim_{t\to\infty} \mathbb{E}\exp(2xY_t^v + 2yY_t^d) < \infty\} \\ S_T &= \{(x,y) : \mathbb{E}\exp(2xY_t^v + 2yY_t^d) < \infty,\ \forall t \in [0,T)\}. \end{aligned}$$

Theorems 3.2.1, 3.2.2 imply that these sets coincide, respectively, with the set S of initial conditions for which the solution to (3.14) exists for all time and converges to zero, and the set S_T for which the solution exists throughout $[0,T)$. Rewriting S and S_T above in terms of p and T, we get

$$\begin{aligned} S &= (-\infty, -p) \times \mathbb{R} \\ S_T &= (-\infty, -p/(1-e^{pT})] \times \mathbb{R}. \end{aligned}$$

If $(\theta b_T, \theta c_T) \in S_T^o$ (the interior of S_T), then (3.13) is finite; if $(\theta b_T, \theta c_T) \in S$, then (3.13) is finite for all T. The left panel of Figure 3.2 illustrates the boundaries of these sets. The parabola shows the values of $\dot{x}_1 = px_1 + x_1^2$ in (3.14) as a function of x_1. The larger of the two solutions to the equation $\dot{x}_1 = 0$ determines the upper limit of the stability region for x_1 (as in Figure 3.1), so ∂S passes through this point. As T decreases, ∂S_T shifts to left.

We can also see from the figure that $(\theta b_T, \theta c_T)$ lies outside S_T for some (and then all) sufficiently large $\theta > 0$ or $\theta < 0$, unless (b_T, c_T) lies on the vertical axis. Thus, P_T^θ has infinite expectation for some θ unless $b_T = 0$. When $b_T = 0$, $\log(P_T) = a_T + 2c_T Y_T^d$ has a Gaussian distribution, and thus does indeed have finite moments of all orders. This is a simple graphical description of the conditions in Theorem 3.2.3 for this example.

3.3.2 Stochastic Volatility: Further Cases

We continue to work with the basic model (3.5)–(3.7), but now take $s > 0$, $q = 0$, and $p = r < 0$. In this case, the function $\xi(t) := e^{-pt}x_1(t)/\sqrt{sx_2(0)^2}$ solves $\dot{\xi}/(\xi^2+1) = \sqrt{sx_2(0)^2}e^{pt}$. Then, we have

$$\tan^{-1}(\xi(t)) - \tan^{-1}(\xi(0)) = \sqrt{sx_2^2(0)}\left(e^{pt}-1\right)/p.$$

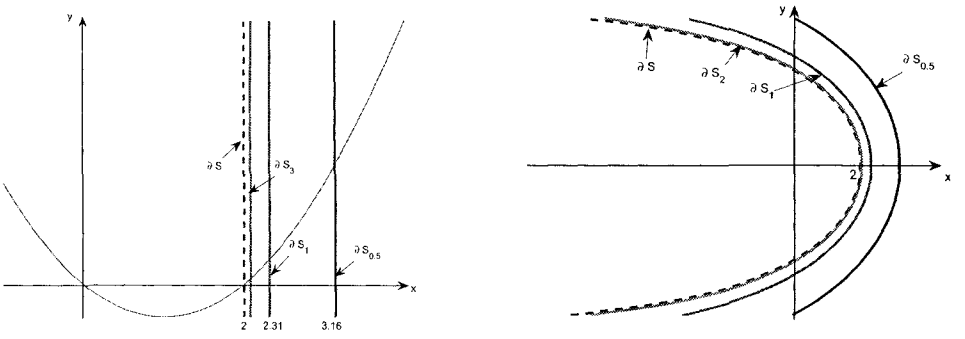


Figure 3.2: Boundaries of S and S_T for $\mathbb{A}_1(2)$ models. The left panel has parameters $p = -2$, $q = s = 0$; the right panel has $p = r = -2$, $q = 0$, $s = 1$.

Therefore,

$$x_1(t) = \sqrt{sx_2^2(0)}e^{pt}\tan\left(\sqrt{sx_2^2(0)}\left(e^{pt}-1\right)/p + \tan^{-1}\left(x_1(0)/\sqrt{sx_2^2(0)}\right)\right), \quad x_2(t) = x_2(0)e^{pt}.$$

Then,

$$S = \{(x,y) : x < \sqrt{sy^2}\tan(\pi/2 + \sqrt{sy^2}/p)\}$$

and

$$S_T = \{(x,y) : x \leq \sqrt{sy^2}\tan(\pi/2 + \sqrt{sy^2}(1-e^{pT})/p)\}.$$

These sets are illustrated in the right panel of Figure 3.2. For any nonzero point (b_T, c_T), the line defined by the points $(\theta b_T, \theta c_T)$ as θ ranges over $\mathbb{R}$ crosses the boundary of S_T twice, once with θ positive and once with θ negative. If (b_T, c_T) is in the interior of S_T, then these values of θ are the extremal moments $\overline{\theta}_T$ and $\underline{\theta}_T$ as a consequence of Theorem 3.2.1. In particular, $\mathbb{E}[P_T^\theta]$ becomes infinite for all sufficiently large positive or negative θ. The log price $\log(P_T)$ is never Gaussian.

We next consider the effect of varying $r < 0$, which is the coefficient on Y_t^d in the expression for dY_t^d in (3.6), while fixing $s > 0$, $q = 0$ and $p < 0$. We can represent $x_1(t)$ in terms a function $\psi(l)$ by setting

$$\psi'\left(-\frac{\sqrt{k}e^{rt}}{r}\right) = \frac{1}{\sqrt{k}e^{rt}}\left(x_1(t) + \frac{p}{2}\right)\psi\left(-\frac{\sqrt{k}e^{rt}}{r}\right), \tag{3.16}$$

with $k = sx_2(0)^2$. The function $\psi(l)$ solves a second order ODE,

$$l^2\psi''(l) + l\psi'(l) + \left(l^2 - \left(\frac{p}{2r}\right)^2\right)\psi(l) = 0. \tag{3.17}$$

It follows that $\psi(l)$ is a linear combination of Bessel functions of the first and second kinds, respectively; see, e.g., p.748 of Polyanin and Zaitsev (2003) for properties of the solution. Since any multiple of $\psi(l)$ satisfies (3.16), we can set $\psi(l)$ as the solution to (3.17) for $l \in (0, -\sqrt{k}/r]$ with $\psi(-\sqrt{k}/r) = \sqrt{k}$, which then satisfies $\psi'(-\sqrt{k}/r) = x_1(0) + p/2$. Since $S = \{x(0) : \lim_{t\to\infty} x(t) = 0\}$, from (3.16) we get

$$S = \left\{x(0) : \lim_{l\downarrow 0} \frac{l\psi'(l)}{\psi(l)} = -\frac{p}{2r}\right\} = \left\{x(0) : -\sqrt{k}/r < \text{the first zero of } \psi(l)\right\}.$$

A similar analysis can be carried out for $s = 0$, $q > 0$ and $p < 0$ case. Figure 3.3 shows the boundary of S for different values of r. The left panel has $q = 0$ and $s = 1$; the right panel has $q = 1$ and $s = 0$. In both cases, the stability region becomes smaller as r approaches zero, indicating that $Y_\infty = (Y^v_\infty, Y^d_\infty)$ has heavier (though still exponentially bounded) tails at smaller values of $|r|$. This is to be expected from the role of r in the dynamics (3.5)–(3.6) of the model.

The two panels of Figure 3.3 show an interesting contrast. In the right panel, we see that a line of the form $\{\theta u : \theta \in \mathbb{R}\}$, $u \in S \cap \mathbb{R}^2_{++}$, crosses the boundary of S just once, at some $\theta > 0$; in the left panel, such a line would cross the boundary of S at both a positive and negative value of θ, as noted in our discussion of Figure 3.2. This reflects an interesting distinction between two ways the volatility factor Y^v can influence the dependent factor Y^d. When Y^v appears in the diffusion coefficient of Y^d (the left panel, with $q = 0$, $s \neq 0$), it makes both the right and left tails of $u \cdot Y_\infty$ exponential, $u \in \mathbb{R}^2_{++}$; when Y^v appears only in the drift of Y^d (the right panel, with $q \neq 0$, $s = 0$), one tail of $u \cdot Y_\infty$ is exponential, but the other is light. The figure has $q > 0$, so the right tail is the exponential one; taking $q < 0$ would reflect the figure about the horizontal axis, corresponding to an exponential left tail.

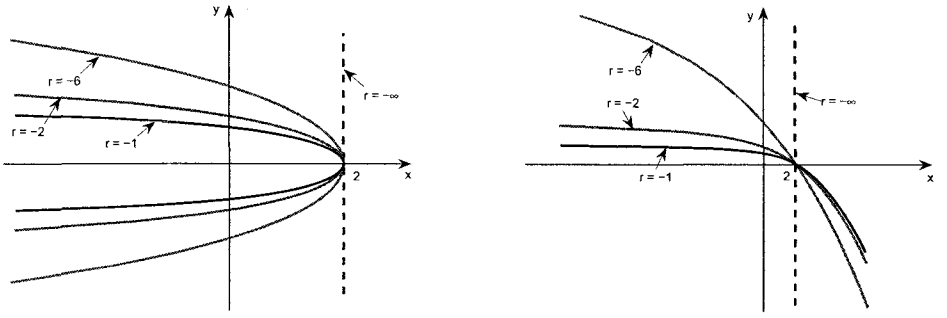


Figure 3.3: Stability boundaries for $\mathbb{A}_1(2)$ models. The left panel has parameters $p = -2$, $q = 0$, $s = 1$; the right panel has $p = -2$, $q = 1$, $s = 0$.

3.3.3 Two Volatility Factors

Our next example is a model in $\mathbb{A}_2(2)$:

$$\begin{aligned} dY_t^1 &= (m_1 + pY_t^1 + rY_t^2)\,dt + \sqrt{Y_t^1}\,dW_t^1 \\ dY_t^2 &= (m_2 + qY_t^1 + sY_t^2)\,dt + \sqrt{Y_t^2}\,dW_t^2. \end{aligned}$$

This can be viewed as a two-factor CIR model; it also belongs to the family of continuous-state branching processes, as explained in Duffie et al. (2003). The associated system of ODEs is

$$\dot{x}_1 = px_1 + qx_2 + x_1^2 \tag{3.18}$$

$$\dot{x}_2 = rx_1 + sx_2 + x_2^2. \tag{3.19}$$

To satisfy the restrictions on the A matrix in (3.3), we require $p, s < 0$, $q, r \geq 0$, and $ps - qr > 0$.

The ODEs (3.18)–(3.19) do not admit a closed-form solution, but we can investigate the qualitative behavior of the system and illustrate this behavior graphically. (We review relevant background on dynamical systems in Section 3.4.1.) Figure 3.4 shows the vector field defined by (3.18)–(3.19) with $p = -3$, $q = 1$, $r = 1/2$, and $s = -1$. The two parabolic curves are the points in the plane satisfying $\dot{x}_1 = 0$ in (3.18) and $\dot{x}_2 = 0$ in (3.19). At the intersections of the two parabolic curves we have $(\dot{x}_1, \dot{x}_2) = 0$, making these equilibrium

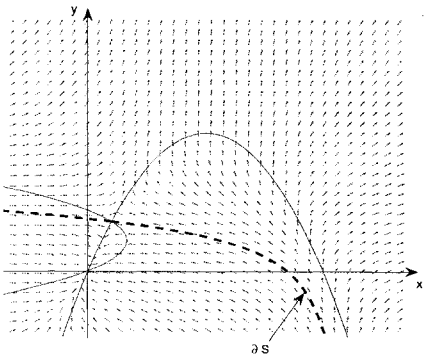


Figure 3.4: Vector field of an $\mathbb{A}_2(2)$ model and ∂S with , $p = -3$, $q = 1$, $r = 0.5$ and $s = -1$.

points; there are two equilibrium points in the example of Figure 3.4, one of which is the origin. The origin is a stable equilibrium: the system approaches the origin from all initial conditions in a neighborhood of the origin. Indeed, the system approaches the origin from all initial conditions in the stability region S, whose boundary ∂S is indicated by a dashed line in the figure. If $x(0)$ lies outside of S, the system explodes, in the sense that $|x(t)| \to \infty$.

The other point of intersection of the two parabolas is an unstable equilibrium: there are initial conditions arbitrarily close to this point from which the system will approach either the origin or infinity. (In the language of dynamical systems, this is a hyperbolic equilibrium of type 1, and therefore unstable; see, Section 3.4.1 and, e.g., Chiang et al. 1988 for background.) Associated with the unstable equilibrium is a stable manifold — a curve in the plane of initial conditions from which the system moves toward the unstable equilibrium. This curve is contained within ∂S.

From Theorem 3.2.2, we know that the points u in S are precisely the points for which $\mathbb{E}[\exp(2u \cdot Y_\infty)]$ is finite. Because S contains a neighborhood of the origin, any linear combination of the components of Y_∞ has exponentially bounded tails. For $u \in S \cap \mathbb{R}^2_{++}$, the line $\{\theta u : \theta \in \mathbb{R}\}$ crosses ∂S just once, at some $\theta > 0$, so $\mathbb{E}[\exp(\theta u \cdot Y_\infty)]$ becomes infinite at for all sufficiently large $\theta > 0$ but remains finite for all $\theta < 0$. In other words, $u \cdot Y_\infty$ has an exponential right tail and a light left tail (in fact, $u \cdot Y_\infty$ is nonnegative).

Figure 3.5 illustrates the behavior of this system for other parameter values. The left panel of the figure shows an example with three equilibrium points, and the right

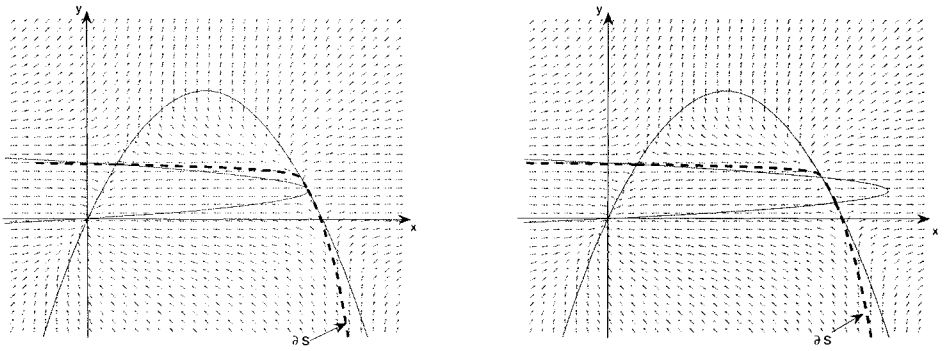


Figure 3.5: The stability boundary for $\mathbb{A}_2(2)$ models. The left panel has parameters $p = -3$, $q = 1$, $r = 0.089$, $s = -1$; the right panel has the same parameters, except with $r = 0.07$.

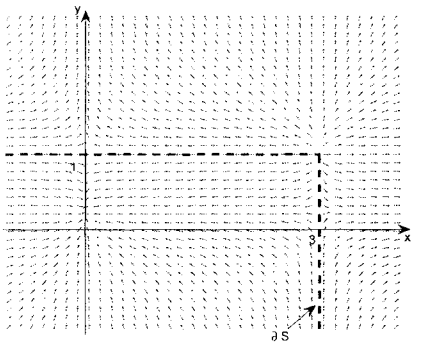


Figure 3.6: The stability boundary for $\mathbb{A}_2(2)$ with $p = -3$, $q = r = 0$, $s = -1$.

panel shows one with four equilibrium points. In both cases, the origin is the only stable equilibrium. Figure 3.6 shows a degenerate case with $q = r = 0$. Here, equations (3.18) and (3.19) decouple, and the stability of each reduces to the analysis of the scalar quadratic differential equation in Section 3.3.1.

3.3.4 Gaussian Conditions

In Theorem 3.2.3, we gave conditions under which $u \cdot Y_t$ and $u \cdot Y_\infty$ have finite moments of all orders, and we noted that these conditions also determine when Y_t and Y_∞ are Gaussian. From the perspective of the associated ODEs, $u \cdot Y_\infty$ has finite moments of all orders precisely if the ODE solution exists for all $t \geq 0$, from all initial conditions θu, $\theta \in \mathbb{R}$;

in other words, the stability region S includes all multiples u. We now illustrate these properties with examples.

Consider the following family of models in $\mathbb{A}_1(3)$:

$$dY_t^1 = (\Lambda_1 - Y_t^1)\,dt + \sqrt{Y_t^1}\,dW_t^1 \tag{3.20}$$

$$dY_t^2 = (\Lambda_2 + aY_t^1 - Y_t^2)\,dt + dW_t^2 \tag{3.21}$$

$$dY_t^3 = (\Lambda_3 + bY_t^1 + cY_t^2 - Y_t^3)\,dt + dW_t^3. \tag{3.22}$$

The model has Y^1 as volatility factor and Y^2 and Y^3 as dependent factors. The matrix A has the form

$$A = \left(\begin{array}{c|cc} A^v & A^c & \\ \hline & & \\ & A^d & \end{array}\right) = \left(\begin{array}{c|cc} -1 & a & b \\ \hline 0 & -1 & c \\ 0 & 0 & -1 \end{array}\right),$$

and $B^c = 0$ because the volatility factor Y^1 does not appear in the diffusion coefficient of either Y^2 or Y^3.

Since A^d is already block diagonal, it is easy to check that

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1/c \end{pmatrix}$$

if $c \neq 0$, and $P = I_2$ if $c = 0$. Condition (3.11) becomes

$$a\,u_2 + b\,u_3 = 0, \quad a\,c\,u_3 = 0. \tag{3.23}$$

The case $c = 0$ reduces to (3.12). Theorem 3.2.3 requires $u^v = 0$, so we must have $u_1 = 0$.

We consider several cases for the parameters a, b, and c.

$a = 0$: We can satisfy (3.23) with any u that is a multiple of $(0, 1, 0)$; i.e., with $u \cdot Y_t = u_2 Y_t^2$. This is also evident from the fact that Y^2 is an Ornstein-Uhlenbeck (OU) process when $a = 0$. If we also have $b = 0$, then u_2 and u_3 are both free in (3.23) and, indeed, (Y^2, Y^3) is a Gaussian process.

$c = 0$, $a \neq 0$, $b \neq 0$: Condition (3.12) is satisfied by taking $u = (0, 1/a, -1/b)$, or any multiple thereof. From (3.20)–(3.22), we see that neither Y^2 nor Y^3 is Gaussian — each has the volatility factor Y^1 in its drift. Nevertheless, the linear combination $u^d \cdot Y^d$ is Gaussian. We can also see this by noting that

$$\begin{aligned} d(u^d \cdot Y_t^d) &= (m - \frac{1}{a}Y_t^2 + \frac{1}{b}Y_t^3)\,dt + \frac{1}{a}dW_t^2 - \frac{1}{b}dW_t^3 \\ &= -u^d \cdot Y_t^d\,dt + \frac{1}{a}dW_t^2 - \frac{1}{b}dW_t^3, \end{aligned}$$

with $m = (\Lambda_2/a) - (\Lambda_3/b) = 0$, in light of (3.4); thus $u^d \cdot Y^d$ is an OU process constructed from non-Gaussian processes. This example illustrates why Corollary 3.2.3 cannot cover all cases.

$c \neq 0$, $a \neq 0$: (3.23) requires $u_2 = u_3 = 0$; thus, no $u \cdot Y$ is Gaussian, except the degenerate case $u \equiv 0$. If $b = 0$, then the equation for Y^3 in (3.22) has no direct dependence on a volatility factor, but it fails to be Gaussian because it depends on Y^2 which depends on Y^1. This is also a consequence of Corollary 3.2.3; the first coordinate of $\tilde{u} = P^{-1}u^d = (u_2 \quad c\,u_3)$ is restricted with respect to $\mathbf{1}_{A^cP}$ and the second coordinate has a directed path from the first coordinate.

In this example, the conclusion of the first case ($a = 0$) and that of the third case ($c \neq 0$, $a \neq 0$) coincide with what one would expect based on the intuitive approach to checking for Gaussian distributions outlined after (3.12) and formalized in Corollary 3.2.3. However, the second case ($c = 0$, $a \neq 0$, $b \neq 0$) shows that the intuitive approach cannot cover all cases. The necessary and sufficient conditions in Theorem 3.2.3 capture the possibility of a Gaussian distribution resulting from a cancellation of factors, as in this example.

3.4 Analysis of Quadratic Dynamical Systems

3.4.1 Definitions and Terminology

In this section, we establish some properties of the ODE system (3.9), in particular viewing it as defining a mapping from the initial condition u to the solution $x(t)$ at time t. We begin

by reviewing some definitions and basic properties from the theory of dynamical systems; additional background can be found in Hirsch and Smale (1974) and Chiang et al. (1988).

Consider, then, an equation

$$\dot{x} = f(x) \tag{3.24}$$

defined by a C^r function $f : W \to E$, with $W \subset E$ open and E a normed vector space. For each $u \in W$, there is a unique solution to (3.24), with $x(0) = u$, defined on a maximal open time interval $I(u) \subset \mathbb{R}$. For $t \in I(u)$, we denote this solution either by $x(t)$ or $\Phi_t(u)$; the notation $\Phi_t(u)$ makes explicit the dependence on the initial condition u. Also, the uniqueness of the solution allows us to write, for example,

$$\Phi_{s+t}(u) = \Phi_s(\Phi_t(u)),$$

for t and $s + t$ in $I(u)$. In particular, Φ_{-t} is the inverse of Φ_t.

Define

$$\Omega = \{(t, u) \in \mathbb{R} \times W : t \in I(u)\};$$

then Φ is a mapping from Ω to W. Standard properties of dynamical systems imply that Ω is open in $\mathbb{R} \times W$ and Φ is C^r if f is C^r, for $0 \leq r \leq \infty$. In fact, Φ is analytic in t and u as long as $\Phi_t(u)$ stays in the domain of analyticity of f.

Let τ denote the (possibly infinite) right endpoint of the interval $I(u)$. If $\tau < \infty$, then for any compact set $K \subset W$, there is a $t \in I(u)$ with $\Phi_t(u) \neq K$; in other words, the solution escapes the domain of definition in finite time, and τ is the "blow-up time" from u.

An *equilibrium point* of (3.24) is a point $\eta \in W$ at which $f(\eta) = 0$. An equilibrium point η is called *hyperbolic* if every eigenvalue of the Jacobian of f at η has a nonzero real part. The *type* of an equilibrium point is the number of eigenvalues (counted according to their multiplicity) with positive real parts. The stable manifold of a hyperbolic equilibrium is the set of points $u \in W$ for which $\Phi_t(u) \to \eta$ as $t \to \infty$; the unstable manifold is the set of $u \in W$ for which $\Phi_{-t}(u) \to \eta$ as $t \to \infty$. A hyperbolic equilibrium η_0 of type zero is a stable equilibrium; this means that its stable manifold contains a neighborhood of η_0 or, equivalently, that its unstable manifold consists solely of η_0. It is also a standard fact that

this stable manifold of η_0 is an open set.

For the system (3.9) associated with a canonical affine model, the origin is a hyperbolic equilibrium of type zero and thus a stable equilibrium. The origin is, in fact, a unique stable equilibrium (see Lemma 4.3.1). We denote its stable manifold by S and call this the *stability region* of the dynamical system. Part of the content of Theorem 3.2.2 is that the stable manifold of the origin determines the range of finite moments of the limiting stationary distribution of the model.

As an aside, we note that the unstable equilibrium in Figure 3.4 is of type 1; the equilibrium at the point of tangency of the two parabolic curves in the left panel of Figure 3.5 fails to be hyperbolic; and, in the right panel of Figure 3.5, the four equilibrium points defined by the four points of intersection of the two curves have types 0, 1, 2, and 1 when taken in clockwise order, starting from the origin. The type-2 equilibrium is a source: its stable manifold consists solely of the point itself.

3.4.2 Solution Properties

Our analysis of the dynamical system (3.9) makes extensive use of comparison theorems, and these in turn prove to be very useful in establishing some distributional properties of Y. The comparison results rely on a concept of quasi-monotonicity. Under the componentwise ordering of vectors introduced in Section 3.2, we call a function $f : \mathbb{R}^n \to \mathbb{R}^n$ *quasi-monotone increasing* if, whenever $x \leq y$ and $x_k = y_k$ for some k, then $f_k(x) \leq f_k(y)$. A mapping $x \mapsto Ax$ defined by a matrix A is thus quasi-monotone increasing if and only if $A_{ij} \geq 0$ whenever $i \neq j$. Suppose that f defined on $\mathbb{R}^n$ is quasi-monotone increasing and locally Lipschitz continuous. Let $x(t), y(t) : [a, b] \to \mathbb{R}^n$ be differentiable functions such that

$$\dot{x}(t) - f(x(t)) \leq \dot{y}(t) - f(y(t)), \quad \forall\, t \in [a, b];$$

then it follows from Volkmann (1972) that

$$x(a) \leq y(a) \Rightarrow x(t) \leq y(t) \quad \forall t \in [a, b]. \tag{3.25}$$

When $n = 1$, this reduces to a standard comparison result for scalar differential equations.

The relevance of this result to our setting comes from property **(C2)**, which makes A^v quasi-monotone, and the fact that the mapping $(x_1, \ldots, x_n) \mapsto (x_1^2, \ldots, x_n^2)$ is also quasi-monotone. Through (3.25), we arrive at the following comparison property for the solution Φ to (3.9):

Lemma 3.4.1 *For any $u \in \mathbb{R}^n$ and $\theta > 1$, we have*

$$\theta\Phi_t(u) \leq \Phi_t(\theta u),$$

for all $t \geq 0$ at which both sides are well-defined.

The proofs of this result and the next two lemmas are deferred to the appendix.

For later reference, we also record the following results on the decay of solutions. See, e.g., Chapter 7 of Verhulst (1996). For the system (3.9), there exist positive constants C, δ, and μ such that

$$|\Phi_t(u)| \leq C|u|e^{-\mu t} \tag{3.26}$$

for all $|u| \leq \delta$, and

$$|\Phi_t^d(u)| \leq C|u^d|e^{-\mu t}, \tag{3.27}$$

for all $u \in \mathbb{R}^n$. The constant $-\mu$ can be chosen to be the eigenvalue of A of smallest magnitude.

Lemma 3.4.2 *For each $u \in \mathbb{R}^n$, the trajectory $\{\Phi_t(u) : t \in [0, \tau)\}$ of (3.9) is bounded below.*

Lemma 3.4.3 *Suppose $|\Phi_t(u)| \to \infty$ as $t \to \tau$, for some $\tau \leq \infty$. Then $\int_0^t \Lambda \cdot \Phi_s(u)ds \to \infty$ as $t \to \tau$.*

3.4.3 Proof of Theorem 3.2.1 and its Corollaries

In light of the expression that appears in the exponent of (3.10), it is natural to introduce the notation

$$\Psi_t(u) = \int_0^t \Lambda \cdot \Phi_s(u)ds + \int_0^t |\Phi_s^d(u)|^2 ds + \Phi_t(u) \cdot Y_0.$$

For $(t,u) \in \Omega$, $\Phi_s(u)$ is bounded for $s \in [0,t]$, so $\Psi_t(u)$ is well-defined and finite. As part of the proof of Theorem 3.2.1, we will show that $\Psi_t(u)$ blows up at τ precisely if $\Phi_t(u)$ does.

Proof of Theorem 3.2.1 We first show that the finiteness of $\Phi_t(u)$ is equivalent to that of $\Psi_t(u)$. One direction is trivial: if $(t,u) \in \Omega$, then $\Phi_s(u)$ is bounded for $s \in [0,t]$ and thus $\Psi_t(u)$ is finite. To show the converse, observe that $|\Phi_t^d(u)|$ is bounded on $t \in \mathbb{R}_+$ (by (3.27)) and $\Phi_t(u)$ is bounded below for its entire life span $t \in [0,\tau)$ (by Lemma 3.4.2). It follows that $\Phi_t(u) \cdot Y_0 = \Phi_t^v(u) \cdot Y_0^v + \Phi_t^d(u) \cdot Y_0^d$ is also bounded below because $Y_0^v \geq 0$. It thus follows from Lemma 3.4.3 and the continuity of $\Phi_t(u)$ (as a function of t) that

$$\Phi_t(u) \text{ blows up at time } \tau \Leftrightarrow \Psi_t(u) \text{ blows up at } \tau. \tag{3.28}$$

Next, we show that if $\Psi_t(u)$ is finite, then $\mathbb{E}\exp(2u \cdot Y_t)$ is also finite and equality holds in (3.10). Duffie et al. (2003) define regular affine Markov processes and show that there are necessary and sufficient conditions for parameters of an affine model to ensure regularity, namely, admissibility. They also show that the transform formula holds true for all $(t,u) \in \mathbb{R}_+ \times \mathbb{C}_-^m \times i\mathbb{R}^{n-m}$ for affine models with admissible parameters. It is not hard to check that canonical affine models satisfy the admissibility condition. And the processes generated by them are conservative, as defined in Duffie et al. (2003). This follows easily from Proposition 9.1 in Duffie et al. (2003); we note that the generalized Riccati equation (2.14) with (2.15) in Duffie et al. (2003) is (3.9) in the canonical case.

Now suppose $\Psi_t(u)$ is finite. Since the process Y is conservative regular affine, by Lemma A.1.2 we can invoke Theorem 2.16 in Duffie et al. (2003) and conclude that $\mathbb{E}\exp(2u \cdot Y_t)$ is finite and the transform formula holds.

We now prove the converse of the main statement of the theorem. Suppose, then, that $\mathbb{E}\exp(2u \cdot Y_t) < \infty$ for some $t > 0$ and $u \in \mathbb{R}^n$. Because the origin is a stable equilibrium and its stability region S is open (see Section 3.4.1), there is a $\theta_0 \in (0,1)$ such that $\theta_0 u \in S$. But if $\theta_0 u \in S$, then $\lim_{s\to\infty} \Phi_s(\theta_0 u) = 0$, and it follows that $\sup_s |\Phi_s(\theta_0 u)| < \infty$. We may then define a positive θ^* by setting

$$\theta^* = \sup\{\theta > 0 : \frac{1}{\theta}\int_0^t \Lambda \cdot \Phi_s(\theta u) ds < \infty\}, \tag{3.29}$$

the supremum taken over those $\theta > 0$ for which $\Phi_t(\theta u)$ is well-defined — i.e., those for which $t \in I(\theta u)$. (If $\Phi_s(\theta u)$ blows up before t, then the integral in (3.29) is infinite.)

If $\theta^* > 1$, then $\Phi_t(u)$ is finite, and we have already shown that this implies that $\Psi_t(u)$ is finite, and we have also shown that (3.10) holds in this case. To complete the proof, we will show that $\theta^* \leq 1$ leads to a contradiction.

Suppose, then, that $\theta^* \leq 1$. Because $\Lambda^v \gg 0$ and $\Phi_s^d(u)$ is linear in the initial condition u, Lemma 3.4.1 implies that the function

$$\theta \mapsto \frac{1}{\theta} \int_0^t \Lambda \cdot \Phi_s(\theta u) ds, \quad \theta \in [\theta_0, \theta^*)$$

is increasing. This implies that

$$\lim_{\theta \uparrow \theta^*} \int_0^t \Lambda \cdot \Phi_s(\theta u) ds = \infty.$$

Also by Lemma 3.4.1, we have

$$\frac{1}{\theta_0} \Phi_s(\theta_0 u) \leq \frac{1}{\theta} \Phi_s(\theta u)$$

for all $(\theta, s) \in R \equiv [\theta_0, \theta^*) \times [0, t]$. Since $\Phi_s(\theta_0 u)$ is bounded below (by Lemma 3.4.2), $\Phi_s(\theta u)$ is bounded below uniformly on R. Moreover, the solution $\Phi_s^d(\theta u)$ to the linear part of (3.9) is uniformly bounded above as well on R, as is easily deduced from (3.27). Thus,

$$\Psi_t(\theta u) \geq \int_0^t \Lambda \cdot \Phi_s(\theta u) ds + K,$$

for some constant K and all $\theta \in [\theta_0, \theta^*)$. It follows that $\lim_{\theta \uparrow \theta^*} \Psi_t(\theta u) = \infty$.

However, for any $\theta \in (0, \theta^*)$, we have $\Psi_t(\theta u) < \infty$, which we already know implies that (3.10) holds at θu, so

$$\exp(2\Psi_t(\theta u)) = \mathbb{E} \exp(2\theta u \cdot Y_t) \leq \left(\mathbb{E} \exp(2u \cdot Y_t)\right)^\theta < \infty,$$

by Jensen's inequality. This implies that $\limsup_{\theta \uparrow \theta^*} \Psi_t(\theta u) < \infty$. But this is a contradiction,

so we must in fact have $\theta^* > 1$.

The last assertion of the theorem now follows directly from the fact that the stability region S of the origin is open. ∎

Proof of Corollary 3.2.1 The indicated tail properties are standard consequences of finite moment generating functions, but we include a brief proof for completeness. From the inequality $1\{z > y\} \leq \exp(\theta(z-y))$, $\theta \geq 0$, we get $\mathbb{P}(u \cdot Y_t > y) \leq \exp(-\theta y)\mathbb{E}\exp(\theta u \cdot Y_t)$, from which the first limsup follows. Suppose now that

$$\limsup_{y\to\infty} \frac{1}{y} \log \mathbb{P}(u \cdot Y_t > y) \leq -\theta - \epsilon,$$

for some $\epsilon > 0$. Then $\mathbb{P}(u \cdot Y_t > y) \leq \exp(-(\theta+\epsilon)y)$ for all sufficiently large y, and so

$$\theta \int_{-\infty}^{\infty} e^{\theta y} \mathbb{P}(u \cdot Y_t > y)\, dy < \infty.$$

With the change of variables $x = \exp(\theta y)$, this becomes

$$\int_0^{\infty} \mathbb{P}(\exp(\theta u \cdot Y_t) > x)\, dx = \mathbb{E}\exp(\theta u \cdot Y_t).$$

The last statement in the corollary is an easy consequence of the fact that the stability region of (3.9) contains a neighborhood of the origin. ∎

Proof of Corollary 3.2.2 In the case of $\mathbb{A}_m(m)$, the vector field $f_o(x)$ of (3.9) is quasi-monotone increasing. We may therefore apply the comparison result in (3.25) with the trivial solution $x \equiv 0$ to conclude that $\Phi_t(u) \geq 0$, for all $t \geq 0$, for any $u \geq 0$.

Fix a $u \geq 0$. If $\Phi_s(u)$ blows up at or before t, then there is nothing to prove because both expectations are infinite. If $\Phi_t(u)$ is finite, then the transform formula (3.10) holds due to Theorem 3.2.1. It follows from (3.10) and the nonnegativity of $\Phi_t(u)$ that the ordering of Λ^1 and Λ^2 implies the ordering of the $\mathbb{E}\exp(2u \cdot Y_t^i)$, $i = 1, 2$. Conversely, if $\mathbb{E}\exp(2u \cdot Y_t^1) \geq \mathbb{E}\exp(2u \cdot Y_t^2)$ for all $u \in \mathbb{R}_+^m$ and $t \geq 0$, then we get

$$\Lambda^1 \cdot u = \lim_{t\downarrow 0} \frac{1}{t}\Lambda^1 \cdot \int_0^t \Phi_s(u)ds \geq \Lambda^2 \cdot u = \lim_{t\downarrow 0} \frac{1}{t}\Lambda^2 \cdot \int_0^t \Phi_s(u)ds.$$

Since this holds for any $u \geq 0$, $\Lambda^1 \geq \Lambda^2$.

For the second statement of the corollary, we write $x(t)$ for a solution to (3.9) with A^1 for A, and $y(t)$ for a solution with A^2. Suppose $u \geq 0$ and $Y_0 \geq 0$ are given. Then, if $A^1 \geq A^2$, we have

$$\dot{x} - (A^2 x + (x_1^2, \ldots, x_m^2)) = (A^1 - A^2)x \geq 0 = \dot{y} - (A^2 y + (y_1^2, \ldots, y_m^2)),$$

the inequality following from the fact that $x(t) \geq 0$ for $x(0) = u \geq 0$. Thus, $x(t) \geq y(t)$ and so the inequality for exponential moments follows from (3.10), because x, y, Λ and Y_0 are nonnegative. Conversely, if the inequality holds for all nonnegative u and Y_0, then

$$\lim_{t \downarrow 0} \frac{1}{t}\left(\frac{1}{2}\log \mathbb{E}\exp(2u \cdot Y_t^1) - u \cdot Y_0\right) \geq \lim_{t \downarrow 0} \frac{1}{t}\left(\frac{1}{2}\log \mathbb{E}\exp(2u \cdot Y_t^2) - u \cdot Y_0\right)$$

yields $((A^1 - A^2)u) \cdot Y_0 \geq 0$. Since Y_0 is an arbitrary vector in $\mathbb{R}_+^m$, $(A^1 - A^2)u \geq 0$, and this in turn implies $A^1 \geq A^2$. ∎

3.5 Convergence to Stationarity

In this section, we use the transform formula (3.10) and our analysis of the ODE (3.9) to prove that a canonical affine model has a unique limiting distribution, that this limiting distribution is stationary, and that the domain of the moment generating function of this limiting stationary distribution coincides with the stability region of the associated dynamical system.

As a first step in our analysis, we show that the moment generating function of Y_t converges, as $t \to \infty$, precisely on the stability region.

Lemma 3.5.1 *Let S be the stability region of the system (3.9). Then,*

$$S = \{u \in \mathbb{R}^n : \lim_{t \to \infty} \mathbb{E}\exp(2u \cdot Y_t) < \infty\}.$$

Proof Suppose $u \in S$. Then, as in (3.26), $\Phi_t(u)$ converges to the origin exponentially as

$t \to \infty$; we may therefore define

$$t_\delta = \inf\{t : |\Phi_t(u)| \le \delta\} < \infty.$$

Let μ and C be as in (3.26). Then, for $t \ge t_\delta$,

$$\begin{aligned}\int_0^t |\Lambda \cdot \Phi_s(u)|ds &\le \int_0^t |\Lambda| \cdot |\Phi_s(u)|ds \\ &\le \int_0^{t_\delta} |\Lambda| \cdot |\Phi_s(u)|ds + C\delta|\Lambda| \int_{t_\delta}^t e^{-\mu(s-t_\delta)}ds.\end{aligned}$$

The last integral converges to a finite value as $t \to \infty$. The integrability of $|\Phi_t^d(u)|^2$ as a function of t follows similarly from (3.27). Therefore, $\lim_{t\to\infty} |\Psi_t(u)| < \infty$, and thus Theorem 3.2.1 implies

$$\lim_{t\to\infty} \mathbb{E} \exp(2u \cdot Y_t) = \lim_{t\to\infty} \exp(2\Psi_t(u)) = \exp(2\Psi_\infty(u)) < \infty. \tag{3.30}$$

For the converse, suppose $u \notin S$. If $\Phi_t(u)$ blows up in finite time τ, then $\lim_{t\to\tau} \exp(2\Psi_t(u)) = \infty$, as shown in (3.28), so no further argument is required in this case. Assume that $\Phi_t(u)$ exists for all $t \ge 0$. Since S open and it contains the origin, we can choose $k > 1$ sufficiently large that $u/k \in S$. Then Lemma 3.4.1 implies $k\Phi_t(u/k) \le \Phi_t(u)$ for all t. This implies that

$$\liminf_{t\to\infty} \int_0^t \Phi_{s,i}(u)ds \ge c_i := \int_0^\infty k\Phi_{s,i}(u/k)ds,$$

for some real number c_i, for each $i \in \{1, \ldots, m\}$. We also have

$$\liminf_{t\to\infty} \Phi_t(u) \ge \liminf_{t\to\infty} k\Phi_t(u/k) = 0.$$

But this liminf cannot be the zero vector; if it were, $\Phi_t(u)$ would reach S in finite time and then converge to 0, which would contradict the fact that $u \notin S$. Thus some component i of $\Phi_t(u)$ has a positive liminf, and i must be in $\{1, \ldots, m\}$ because $\Phi_t^d(u)$ converges to zero. As

a consequence,

$$\liminf_{t\to\infty} \int_0^t \Lambda^v \cdot \Phi_s^v(u)ds \geq \sum_{j\neq i} \Lambda_j c_j + \liminf_{t\to\infty} \int_0^t \Lambda_i \Phi_{s,i}(u)ds = \infty.$$

It follows that $\liminf_{t\to\infty} \Psi_t(u) = \infty$ and thus $\liminf_{t\to\infty} \mathbb{E}\exp(2u \cdot Y_t) = \infty$. ∎

Proof of Theorem 3.2.2 We start by showing that the sequence $\{Y_t\}$ is tight (as defined, for example, in Chung 2001, p.90). For this, we need to show $\lim_{r\to\infty} \sup_t \mathbb{P}(|Y_t| > r) = 0$. But we have

$$\begin{aligned}
\mathbb{P}(|Y_t| > r) &\leq \mathbb{P}\Big(\bigcup_i \{|Y_{t,i}| > r/\sqrt{n}\}\Big) \leq \sum_i \mathbb{P}(|Y_{t,i}| > r/\sqrt{n}) \\
&= \sum_i \Big\{\mathbb{P}(Y_{t,i} > r/\sqrt{n}) + \mathbb{P}(-Y_{t,i} > r/\sqrt{n})\Big\} \\
&= \sum_i \Big\{\mathbb{P}(e^{2\delta Y_{t,i}} > e^{2\delta r/\sqrt{n}}) + \mathbb{P}(e^{-2\delta Y_{t,i}} > e^{2\delta r/\sqrt{n}})\Big\} \\
&\leq \sum_i \Big\{\frac{\mathbb{E}e^{2\delta Y_{t,i}}}{e^{2\delta r/\sqrt{n}}} + \frac{\mathbb{E}e^{-2\delta Y_{t,i}}}{e^{2\delta r/\sqrt{n}}}\Big\},
\end{aligned}$$

where δ is a positive constant such that $B_\delta(0) \subset S$. From Lemma 3.5.1, we get $\sup_t \mathbb{E}\exp(\pm 2\delta Y_{t,i}) \leq M_i < \infty$, for some M_i, for each i. Therefore,

$$\sup_t \mathbb{P}(|Y_t| > r) \leq 2\sum_i M_i \exp(-2\delta r/\sqrt{n})$$

which converges to zero as $r \to \infty$.

Because the sequence $\{Y_t\}$ is tight, it is relatively compact (Chung 2001, p.90), so each subsequence $\{Y_{t'}\}$ contains a further subsequence $\{Y_{t''}\}$ converging weakly to some limiting random vector Y^a. Since we have $\sup_{t''} \mathbb{E}\exp(2u \cdot Y_{t''}) < \infty$, for any $u \in B_\delta(0)$ (by Lemma 3.5.1) and since $Y_{t''} \Rightarrow Y^a$, Theorem 4.5.2 in Chung (2001) implies that

$$\lim_{t''\to\infty} \mathbb{E}\exp(2\theta u \cdot Y_{t''}) = \mathbb{E}\exp(2\theta u \cdot Y^a), \quad \forall \theta \in (0,1). \tag{3.31}$$

Equality continues to hold if we replace θu by u because $B_\delta(0)$ is open: we can find $u' \in B_\delta(0)$ such that $u = \theta u'$ for some $\theta \in (0,1)$ and then apply (3.31) at u'. From (3.30) we know that

the original sequence $\{Y_t\}$ satisfies $\lim_{t\to\infty} \mathbb{E}\exp(2u\cdot Y_t) = \exp(2\Psi_\infty(u))$ for $u \in B_\delta(0)$, so the same limit applies to $\{Y_{t''}\}$. Applying the same argument to any other weakly convergent subsequence of $\{Y_t\}$, say with limit Y^b, we find that

$$\mathbb{E}\exp(2u\cdot Y^a) = \exp(2\Psi_\infty(u)) = \mathbb{E}\exp(2u\cdot Y^b), \quad \forall u \in B_\delta(0).$$

But the distribution of a random vector is uniquely determined by its moment generating function in a neighborhood of the origin, so $Y^a \sim Y^b$. Since every convergent subsequence has the same limiting distribution, the original sequence $\{Y_t\}$ also converges to Y^a in distribution, so we now denote Y^a by Y_∞. We have shown that $\mathbb{E}\exp(2u\cdot Y_t) \to \mathbb{E}\exp(2u\cdot Y_\infty)$ for all $u \in B_\delta(0)$. Our next step will be to show that this holds for all $u \in S$, and to show that $\mathbb{E}\exp(2u\cdot Y_\infty) = \infty$ if $u \notin S$.

For any $u \in S$, we can find $u' \in S$ and $\theta \in (0,1)$ with $u = \theta u'$, because S is an open set containing the origin. We know that $Y_t \Rightarrow Y_\infty$ and, by Lemma 3.5.1, that $\sup_t \mathbb{E}\exp(2u'\cdot Y_t)$ is finite. It follows from Theorem 4.5.2 of Chung (2001) that $\mathbb{E}\exp(2u\cdot Y_t) \to \mathbb{E}\exp(2u\cdot Y_\infty)$, so we conclude that $S \subseteq \{u : \mathbb{E}\exp(2u\cdot Y_\infty) < \infty\}$.

We prove the opposite inclusion by contradiction. For this, suppose that $u \notin S$ and that $\mathbb{E}\exp(2u\cdot Y_\infty) < \infty$. Define

$$\theta^* = \sup\{\theta \in [0,1] : \theta u \in S\};$$

then $\theta^* > 0$ and $\theta^* u$ is on ∂S, the topological boundary of S, because S is open and $u \notin S$. Fix a $\theta_0 \in (0,\theta^*)$, so that $\theta_0 u \in S$, and set $g(t) = \Phi_t(\theta_0 u)/\theta_0$. Lemma 3.4.1 implies that $\Phi_t(\theta u) \geq \theta g(t)$, for all $t \geq 0$ and all $\theta \in [\theta_0, \theta^*)$. Consider the trajectory of $\Phi_t(\theta^* u)$. We claim that $\tau = \infty$. To see this, choose a $\theta \in (\theta_0, \theta^*)$. Then, for each $i \in \{1,\ldots,m\}$,

$$\begin{aligned} x_i^2 + \sum_j A_{ij}x_j + \sum_j B_{ij}x_j^2 &\geq x_i^2 + A_{ii}x_i + \theta\sum_{j\neq i} A_{ij}g_j(t) \\ &\geq x_i^2 + A_{ii}x_i + \theta M \end{aligned}$$

where $x(t) = \Phi_t(\theta u)$ and M is a lower bound of the summation. Next, we define a new

function y starting at t_0 by

$$\dot{y} = y^2 + A_{ii}y + \theta M, \quad y(t_0) = x_i(t_0).$$

If $y(t_0)$ is sufficiently large, then $y(t)$ blows up in finite time (see Section 3.3.1) and so does $x_i(t)$. Suppose $\tau < \infty$. Then, it is possible to choose θ close to θ^* and $t_0 < \tau$ such that some $x_i(t_0)$ becomes large enough to make $y(t)$ blow up in finite time. This is a contradiction to $\theta u \in S$.

Therefore, we have $\lim_{t\to\infty} \Psi_t(\theta^* u) = \infty$ as shown in the proof of Lemma 3.5.1. On the other hand, we have

$$\begin{aligned} \int_0^\infty \Lambda^v \cdot (\Phi_t^v(\theta^* u) - \theta^* g^v(t))dt &= \int_0^\infty \lim_{\theta\uparrow\theta^*} \Lambda^v \cdot (\Phi_t^v(\theta u) - \theta g^v(t))dt \\ &\leq \liminf_{\theta\uparrow\theta^*} \int_0^\infty \Lambda^v \cdot \Phi_t^v(\theta u)dt - \theta^* \int_0^\infty \Lambda^v \cdot g^v(t)dt \end{aligned} \tag{3.32}$$

where the equality comes from the continuity of the flow Φ and the inequality is from Fatou's lemma. Since $\Lambda^v \cdot g^v(t)$ and $\Phi_t^d(\theta^* u)$ are integrable, $\lim_{t\to\infty} \Psi_t(\theta^* u) = \infty$ implies that the left side of (3.32) is infinite. Therefore, $\liminf_{\theta\uparrow\theta^*} \int_0^\tau \Lambda^v \cdot \Phi_t^v(\theta u)dt = \infty$. But for $\theta \in (0, \theta^*)$, $\theta u \in S$ and utilizing Jensen's inequality,

$$\exp(2\Psi_\infty(\theta u)) = \mathbb{E}\exp(2\theta u \cdot Y_\infty) \leq \left(\mathbb{E}\exp(2u \cdot Y_\infty)\right)^\theta < \infty.$$

Therefore, $\limsup_{\theta\uparrow\theta^*} \Psi_\infty(\theta u) < \infty$ and this is a contradiction.

To conclude the proof, we need to show that the limiting distribution is a stationary distribution. Suppose, therefore, that $Y_0 \sim Y_\infty$. Then for any $u \in S$, by taking a conditional

expectation,

$$\begin{aligned}
\mathbb{E}\exp(2u\cdot Y_t) &= \mathbb{E}\exp\Big(2\int_0^t \Lambda\cdot\Phi_s(u)ds + 2\int_0^t |\Phi_s^d(u)|^2 ds + 2\Phi_t(u)\cdot Y_0\Big)\\
&= \exp\Big(2\int_0^t \Lambda\cdot\Phi_s(u)ds + 2\int_0^t |\Phi_s^d(u)|^2 ds\Big)\mathbb{E}\exp(2\Phi_t(u)\cdot Y_0)\\
&= \exp\Big(2\int_0^t \Lambda\cdot\Phi_s(u)ds + 2\int_0^t |\Phi_s^d(u)|^2 ds\Big)\\
&\quad\times\exp\Big(2\int_0^\infty \Lambda\cdot\Phi_s(\Phi_t(u))ds + 2\int_0^\infty |\Phi_s^d(\Phi_t(u))|^2 ds\Big)\\
&= \exp\Big(2\int_0^\infty \Lambda\cdot\Phi_t(u)dt + 2\int_0^\infty |\Phi_t^d(u)|^2 dt\Big) \qquad (3.33)\\
&= \mathbb{E}\exp(2u\cdot Y_\infty).
\end{aligned}$$

Because the distribution of a random vector is determined by the values of its moment generating function in a neighborhood of the origin, we conclude that Y_t has the distribution of Y_∞ whenever Y_0 does. ∎

Observe that (3.33) gives the moment generating function of Y_∞ and thus characterizes the stationary distribution of Y_t.

From the preceding proof, we see that the distribution of Y_∞ is determined by the behavior of the dynamical system (3.9) on the stable manifold S of the stable equilibrium at the origin: the fact that $\Phi_t(u) \to 0$ for $u \in S$ is crucial to the convergence of $\Psi_t(u)$ and thus the moment generating function of $u\cdot Y_t$. This raises the question of whether other, unstable equilibria play any role in the stochastic behavior of the basic model (3.3). Our next result illustrates a setting in which they do.

Proposition 3.5.1 *Suppose that η is a hyperbolic equilibrium of system (3.9) of type less than n. Then for any u in the stable manifold of η, we have*

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}\exp(2u\cdot Y_t) = 2\Lambda\cdot\eta. \qquad (3.34)$$

Proof If u lies on the stable manifold of η, then $\lim_{t\to\infty}\Phi_t(u) = \eta$, so $\Psi_t(u)$ is well defined for all $t \geq 0$. The limit on the left side of (3.34) is given by the limit of $2\Psi_t(u)$ as $t \to \infty$; i.e.,

by

$$\lim_{t\to\infty}\frac{2}{t}\int_0^t \Lambda\cdot\Phi_s(u)ds + \lim_{t\to\infty}\frac{2}{t}\int_0^t |\Phi_s^d(u)|^2 ds + \lim_{t\to\infty}\frac{2}{t}\Phi_t(u)\cdot Y_0.$$

The last term is clearly zero, and the second term also vanishes because

$$\frac{1}{t}\int_0^t |\Phi_s^d(u)|^2 ds \le \frac{1}{t}\int_0^t C^2|u^d|^2\exp(-2\mu s)ds = \frac{1}{t}C^2|u^d|^2\frac{1-\exp(-2\mu t)}{2\mu}\to 0,$$

in light of (3.27). The first limit is $2\Lambda\cdot\eta$. ∎

The condition in the proposition on the equilibrium's type ensures the existence of a stable manifold. An equilibrium of type n is a source, an example of which appears in the right panel of Figure 3.5, at the upper right intersection of the two curves. The limit in (3.34) arises in the definition of the rate function used in the Gärtner-Ellis Theorem (see, e.g., Dembo and Zeitouni 1998). The behavior in (3.34) is somewhat pathological because the limit, viewed as a function of u, fails to be a closed convex function. As a consequence, the Gärtner-Ellis Theorem does not apply to the sequence $\{Y_t/t\}$.

Theorem 3.2.2 characterizes the set of u for which $\mathbb{E}\exp(2u\cdot Y_\infty)$ is finite and identifies this set with the stability region S of (3.9). The problem of describing the boundary of S has attracted considerable attention. Genesio et al. (1985) survey methods using a Lyapunov approach; Chiang et al. (1988) characterize ∂S in terms of stable submanifolds of unstable equilibria. Chapter 4 establishes a similar result for the quadratic system (3.9).

Theorem 3.2.2 raises the question of characterizing the region in which Y_t has finite exponential moments, for finite t; that is, characterizing

$$S_t = \{u\in\mathbb{R}^n : \mathbb{E}\exp(2u\cdot Y_s) < \infty,\ \forall\, s\in[0,t)\}.$$

This set coincides with the set of initial conditions u for which the solution $\Phi_s(u)$ exists throughout $[0,t)$. Directly from the definition of S_t, we see that S_t shrinks as t increases; that S_t is convex follows from Hölder's inequality. Beyond these basic properties, it is generally more difficult to characterize S_t than S, at least from the perspective of the dynamical system (3.9). Theorem 3.2.3 and the analysis in the next section give some results in this direction.

3.6 Gaussian Conditions

Lemma 3.6.1 *For any $t > 0$ and $u \in \mathbb{R}^n$, $\mathbb{E}\exp(2\theta u \cdot Y_t) < \infty$ for all $\theta \in \mathbb{R}$ if and only if*

$$u^v = 0, \quad A^c x^d(s) = 0, \quad B^c(x^2_{m+1}(s), \ldots, x^2_n(s)) = 0$$

for all $s \geq 0$, where x^d is the solution to $\dot{x} = A^d x$ with $x(0) = u^d$. Moreover, in this case, $u \cdot Y_t$ has a Gaussian distribution.

Proof See the appendix. ∎

Proof of Theorem 3.2.3 In writing $P^{-1}A^dP = J$, we may assume P is chosen to give J the specific form described before the statement of the theorem. We further assume that the k distinct eigenvalues of A^d are numbered in decreasing order, $\lambda_k < \cdots < \lambda_1 < 0$.

Define $y(t) = P^{-1}x(t)$, where x is the solution to $\dot{x} = A^d x$ with $x(0) = u^d$. Then y satisfies $\dot{y} = Jy$ with $y(0) = \tilde{u}$ and $\tilde{u} = P^{-1}u^d$. Let y^i denote the block of y corresponding to the i-th block $J_i = \lambda_i I_i + N_i$ of J. We use this notation similarly for other vectors. In other words, if the $a_{\lambda_i} \times a_{\lambda_i}$ matrix J_i runs through coordinates $(p+1, p+1), \ldots, (p+a_{\lambda_i}, p+a_{\lambda_i})$ of J, then v^i of $v \in \mathbb{R}^n$ is $(v_{p+1}, \ldots, v_{p+a_{\lambda_i}})$. Since we have $\dot{y}^i = J_i y^i$, $y^i(0) = \tilde{u}^i$, the solution is expressed as follows:

$$y^i(t) = \exp(\lambda_i t)\left[I_i + \sum_{l=1}^{a_{\lambda_i}-1} \frac{t^l}{l!} N_i^l\right]\tilde{u}^i.$$

Suppose that $w^\top y \equiv 0$ for some $w \in \mathbb{R}^n$. Then $\sum_{i=1}^k {w^i}^\top y^i \equiv 0$. If we divide this by $\exp(\lambda_1 t)$, which has the smallest magnitude among eigenvalues, and send $t \to \infty$, then $\exp(-\lambda_1 t){w^1}^\top y^1 \equiv 0$; otherwise, we equate one exponentially decreasing function with a polynomial, which is absurd. By applying the same procedure with other λ_i's, we conclude that ${w^i}^\top y^i \equiv 0$ for each i. Consequently, $w^\top y \equiv 0$ is equivalent to

$${w^i}^\top N_i^l \tilde{u}^i = 0, \quad i = 1, \ldots, k, \quad l = 0, \ldots, a_{\lambda_i} - 1. \tag{3.35}$$

This observation implies that the first two conditions in Lemma 3.6.1 are equivalent to requiring that $u^v = 0$ and that (3.35) holds for all ${w^i}^\top$ that are rows of A^cP. As for the third

condition in Lemma 3.6.1, we note that $x_q = \sum_l P_{ql} y_l \equiv 0$ if there exists some p such that $B^c_{pq} \neq 0$. Therefore, (3.11) follows. ∎

Proof of Corollary 3.2.3 Choose any block J_i of J and $\tilde{u}^i$. By construction, J_i is itself a block diagonal matrix consisting of Jordan blocks associated with λ_i; each Jordan block has a 1 in every entry immediately above the main diagonal. Let Q be any Jordan block of J_i and $\tilde{u}^Q$ the corresponding block of $\tilde{u}^i$ with dimension d, say. Then, the following condition becomes a sufficient condition that induces (3.35):

$$w^{Q^\top} N^l \tilde{u}^Q = 0, \quad l = 0, \ldots, d-1, \quad \forall Q$$

where N is Q less the diagonal part. But, then, this is just

$$w^{Q^\top} \begin{pmatrix} \tilde{u}^Q_1 \\ \vdots \\ \tilde{u}^Q_{d-1} \\ \tilde{u}^Q_d \end{pmatrix} = 0, \quad w^{Q^\top} \begin{pmatrix} \tilde{u}^Q_2 \\ \vdots \\ \tilde{u}^Q_d \\ 0 \end{pmatrix} = 0, \quad \ldots, w^{Q^\top} \begin{pmatrix} \tilde{u}^Q_d \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0.$$

Therefore, an equivalent statement is that if j is a coordinate with $w^Q_j \neq 0$, then $\tilde{u}^Q_j = \tilde{u}^Q_{j+1} = \cdots = \tilde{u}^Q_d = 0$.

The directed graph G in this case consists of paths such as $n \to n+1 \to \cdots \to n+d-1$ if Q starts at the coordinate (n, n). If j is restricted with respect to $\mathbf{1}_{A^cP} + \mathbf{1}_{B^c}\mathbf{1}_P$, then $(A^cP)_{ij} \neq 0$ or $B^c_{iq} P_{qj} \neq 0$ for some i, q. This in turn means that $w_j \neq 0$ where w is the i-th row of A^cP or the q-th row of P, and thus $\tilde{u}_j = 0$. In this case, the observation in the previous paragraph requires that any other components of $\tilde{u}$ that have a directed path from $\tilde{u}_j$ in G are also zero.

If $g_{\lambda_i} = 1$ for all i, then there is only one Jordan block for each λ_i and thus Q coincides with J_i. Therefore, the condition above becomes necessary, too. ∎

Corollary 3.2.3 essentially means that we achieve a non-Gaussian distribution for $u \cdot Y_t$ as long as it has some dependence on one or some of volatility driving factors by including them in the dynamics or by including a factor that depends on volatility factors. Of course,

u has to be outside the closed set specified by (3.11). The vectors in this set cancel out the effects of the volatility factors in $u \cdot Y_t$. The next examples illustrate this feature in more detail.

Example $\mathbb{A}_m(n)$ with diagonal A^d. In this case, we have

$$dY_j^d(t) = \left(\Lambda_j^d + \sum_k A_{kj}^c Y_k^v + A_{jj}^d Y_j^d(t)\right)dt + \sqrt{1 + \sum_k B_{kj}^c Y_k^v} dW_j^d(t)$$

and

$$\begin{aligned} d\left(u^d \cdot Y^d(t)\right) &= \left(u^d \cdot \Lambda^d + \sum_k \left(\sum_j u_j^d A_{kj}^c\right) Y_k^v + \sum_j u_j^d A_{jj}^d Y_j^d(t)\right)dt \\ &\quad + \sum_j u_j^d \sqrt{1 + \sum_k B_{kj}^c Y_k^v} dW_j^d(t) \end{aligned} \tag{3.36}$$

For $u \cdot Y$ not to have any dependence on Y^v, we must have $u^v = 0$,

$$\sum_j u_j^d A_{kj}^c = 0, \quad k = 1, \ldots, m$$

and $u_j^d = 0$ whenever there exists k such that $B_{kj}^c \neq 0$. However, these conditions are not enough to remove all the dependence on Y^v. For example, suppose A^d is given by

$$A^d = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

Then, (3.36) becomes

$$d\left(u^d \cdot Y^d(t)\right) = \left(u^d \cdot \Lambda^d + \lambda_1\left(u^d \cdot Y^d(t)\right) + (\lambda_2 - \lambda_1)u_3^d Y_3^d(t)\right)dt + \sum_{j \notin \mathcal{J}} u_j^d dW_j^d(t)$$

where $\mathcal{J}$ is a set of coordinates that are restricted with respect to $\mathbf{1}_{B^c}$. Therefore, if Y_3^d has a volatility factor in its drift or diffusion, then $u^d \cdot Y^d$ is not free of Y^v effects. This kind of additional dependency is captured by (3.11).

Example $\mathbb{A}_m(m+2)$. This class of models has two dependent factors. We consider the case in which A^d has only one eigenvalue λ with $g_\lambda = 1$. The other possible case is diagonal and is similar to the example above but with a lower dimension. Let $P = (v_1 \quad v_2)$ be the non-singular matrix of an eigenvector and a generalized eigenvector in a Jordan canonical form of A^d and let $P^{-1}u = (a, b)$. We write

$$A^d P = P \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad L = \begin{pmatrix} I^v & 0 \\ 0 & P^\top \end{pmatrix}.$$

Next we apply an invariant affine transformation as defined in Dai and Singleton (2000), $Y \mapsto LY$. Then the dynamics of Y^v are the same as the original and that of $\tilde{Y} = P^\top Y^d$ becomes

$$d\tilde{Y}_t = \left(P^\top \Lambda^d + (A^c P)^\top Y_t^v + J^\top \tilde{Y}_t\right) dt + P^\top \sqrt{diag(F_t^d)} dW_t^d.$$

Denoting $\tilde{Y}$ by $(\tilde{Y}_1, \tilde{Y}_2)$,

$$\begin{aligned} d\begin{pmatrix} \tilde{Y}_1(t) \\ \tilde{Y}_2(t) \end{pmatrix} &= P^\top \Lambda^d dt + (A^c P)^\top Y_t^v dt + \begin{pmatrix} \lambda \tilde{Y}_1(t) \\ \tilde{Y}_1(t) + \lambda \tilde{Y}_2(t) \end{pmatrix} dt \\ &\quad + \begin{pmatrix} (v_1)_1 \sqrt{1 + \sum B_{k1}^c Y_k^v} dW_2(t) + (v_1)_2 \sqrt{1 + \sum B_{k2}^c Y^v} dW_3(t) \\ (v_2)_1 \sqrt{1 + \sum B_{k1}^c Y_k^v} dW_2(t) + (v_2)_2 \sqrt{1 + \sum B_{k2}^c Y_k^v} dW_3(t) \end{pmatrix} \end{aligned}$$

Note that $u \cdot Y_t = \tilde{u} \cdot \tilde{Y}_t = a\tilde{Y}_1(t) + b\tilde{Y}_2(t)$ (we assume $u^v = 0$). Now suppose $a \neq 0$. Then, $u \cdot Y_t$ has a dependence on Y^v unless $(A^c P)_{k1} = 0$ and $B_{ki}^c = 0$ for all k whenever $(v_1)_i \neq 0$. This is the same as asking whether coordinate 1 is restricted with respect to $\mathbf{1}_{A^c P} + \mathbf{1}_{B^c} \mathbf{1}_P$. A similar argument applies to the case $b \neq 0$ regarding the second coordinate.

If $a = 0$ but $b \neq 0$, then we still have to consider the dependence of $\tilde{Y}_1$ on Y^v because $\tilde{Y}_2$ is correlated with $\tilde{Y}_1$ through the drift term. This means that $u \cdot Y_t$ has dependence on Y^v if coordinate 1 is restricted. It is clear from the dynamics of $\tilde{Y}$ that the final dynamics induce a Gaussian distribution after we remove the dependence on Y^v.

3.7 Conclusion

We have established three general results for affine models. Our first result confirms the validity of the transform representation without further conditions and shows that the range of exponents for which the transform is finite at time t coincides with the set of initial conditions from which the ODE solution exists on $[0, t]$. Based on this result, we are able to investigate the properties of affine models by analyzing the associated differential equations. As an example, we gave two comparison criteria for processes in $\mathbb{A}_m(m)$.

Our second result establishes the existence of a limiting stationary distribution and characterizes this limit through its transform; the tail behavior of the limiting distribution is determined by the stability region of the associated dynamical system.

Our last result gives necessary and sufficient conditions for a linear combination of factors to have a Gaussian distribution and shows that any non-Gaussian linear combination has exponential tails. Essentially, a Gaussian distribution is obtained by removing from a linear combination all the dependence on the volatility factors, but the precise conditions that achieve this turn out to be subtle.

Chapter 4

Stability Analysis of Riccati Differential Equations related to Affine Diffusion Models

We study a class of generalized Riccati differential equations associated with canonical affine diffusion processes. As seen in Chapter 3, the generalized Riccati equations determine the Fourier transform of the diffusion's transition law. We investigate stable regions of the dynamical systems and analyze their blow-up times. We discuss the implication of applying these results to affine diffusions and, in particular, to option pricing theory.

4.1 Introduction

In this chapter, we study the stability properties of the quadratic differential equations (3.9) associated with canonical affine diffusion processes. For convenience, we recall that

$$\dot{x}(t) = f_0(x) = Ax + B\begin{pmatrix} x_1^2 \\ \vdots \\ x_n^2 \end{pmatrix}, \quad x(0) = u \in \mathbb{R}^n \tag{4.1}$$

where A and B are given as

$$A = \begin{pmatrix} A^v & A^c \\ 0 & A^d \end{pmatrix}, \quad B = \begin{pmatrix} I & B^c \\ 0 & 0 \end{pmatrix}$$

with A^v, $I \in \mathbb{R}^{m\times m}$ $(m \leq n)$ and other matrices belonging to Euclidean spaces with appropriate dimensions and parametric conditions on them as in Section 3.2. Our objectives are first, to study the stable regions of (4.1) and second, to investigate the blow-up phenomena of solutions.

The determination of stable regions of stable equilibria in non-linear dynamical systems holds significance in various contexts and there have been many theoretical and numerical solution approaches to this question (see, e.g., Chiang and Fekih-Ahmed 1996, Chiang et al. 1988, Genesio et al. 1985, Vannelli and Vidyasagar 1985 and references therein). The techniques used in this area vary according to a specific problem of interest: for example, Levin (1994), Tibken (2000) for polynomial systems and Cheng et al. (2004), Saha et al. (1997) for power systems to name a few. On the other hand, the escape of a solution to infinity, or the blow-up of a solution in finite time has also been widely studied and Baris et al. (2006), Crouch and Pavon (1987), Getz and Jacobson (1977), Martin (1981), Sasagawa (1982) address this issue for certain classes of quadratic differential equations.

The link between the diffusions (3.3) and the ordinary differential equations (ODEs) (4.1) is the Fourier transform formula as formulated and generalized in Duffie et al. (2000) and Duffie et al. (2003) for a larger class of stochastic processes. We note that (4.1) is a special case of generalized Riccati equations as defined in Duffie et al. (2003). Relevant

backgrounds are provided in the next section.

This chapter begins by reviewing some notation and concepts from the theory of dynamical systems in Section 4.2. The following three sections characterize the boundaries of stability regions and the regions in which solutions exist at time t. Then, the results are applied in the option pricing context. Section 4.6 concludes.

4.2 Model Description and Background

Throughout this chapter, we will use the notational conventions introduced in Section 3.2 including the orderings on $\mathbb{R}^n$ and $\mathbb{R}^{n\times n}$; n denotes the dimension of the system (4.1), $m \leq n$ such that $A^v \in \mathbb{R}^{m\times m}$ and we have for any vectors or matrices a and b,

$$a \geq b \Longleftrightarrow a_{ij} \geq b_{ij}$$
$$a > b \Longleftrightarrow a \geq b,\ a \neq b$$
$$a \gg b \Longleftrightarrow a_{ij} > b_{ij}.$$

And for $a \in \mathbb{R}^n$, we define $a^v = (a_1, \ldots, a_m)$ and $a^d = (a^d_1, \ldots, a^d_{n-m}) = (a_{m+1}, \ldots, a_n)$. Similarly, if a is an n by n matrix, then a^v is the upper-left m by m block and a^d is the lower-right $n-m$ by $n-m$ block so that the notation for A^v and A^d in (4.1) matches. Also, we write $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x \geq 0\}$, $\mathbb{R}^m_{++} = \{x \in \mathbb{R}^m : x \gg 0\}$ (similarly for matrices), and $|x|$ is the usual Euclidean norm of a vector x and 0 is the zero vector (or the zero matrix) with an appropriate dimension which should be clear from the context. Parametric restrictions on (4.1) are also given in Section 3.2.

As in Section 3.4.1, we define $I(u)$ as the maximal open interval of existence of a solution to (4.1) with $x(0) = u$ and define

$$\tau : \mathbb{R}^n \to (0, \infty], \quad \tau(u) = \sup I(u).$$

To specify the initial condition, we write $\Phi_t(u)$ for $x(t)$. If $I(u) = (a, b)$, then $|x(t)|$ becomes infinite as $t \to b$. Recall that when η is an equilibrium and its Jacobian $J(\eta)$ has eigenvalues

(counting multiplicity) with k positive real part, η is said to be of *type k*. If η is hyperbolic and type 0, then η is a stable equilibrium. If η is hyperbolic but its type is positive, then it is unstable and denoted by UEP for short .

As shown in Theorem 3.2.1, (3.10) holds true unconditionally as long as either side is well-defined and finite. Therefore, studying the blow-up phenomena of the dynamical system (4.1) is equivalent to studying the finiteness of the exponential moments of the process $\{u \cdot Y_t\}$. In Section 4.4, we characterize S in terms of the stable sub-manifolds of hyperbolic equilibria on the stability boundary ∂S (see Chiang et al. 1988 for a general approach in this direction). Similarly as in Chapter 3, we define

$$S_t = \{u \in \mathbb{R}^n : \mathbb{E}\exp(2u \cdot Y_s) < \infty, \quad \forall s \in [0,t)\},$$

and this coincides with a set in which a solution to (4.1) with $x(0) = u$ exists in $[0,t)$. In other words, $S_t = \{u \in \mathbb{R}^n : \tau(u) \geq t\}$.

4.3 Properties of the System and Blow-up Times

A great deal of work has been performed on the analytical or numerical computation of stability regions, e.g., see Genesio et al. (1985) for a compact survey. In particular, Chiang et al. (1988) showed that the boundary of a stability region can be represented as the union of the stable sub-manifolds of hyperbolic equilibria on the stability boundary ∂S under certain conditions. Inspired by this, we demonstrate that a similar result can be pursued under a slightly altered assumption in our case.

Before we proceed, observe that the assumptions on A and A^v make $-A^v$ a nonsingular M-matrix (see Berman and Plemmons 1994). This induces two nice properties that are used in the proofs of our results: first, $-(A^v)^{-1} \geq 0$ and second, $-A^v x \geq 0$ implies $x \geq 0$.

Another immediate consequence of the assumptions on A is some qualitative behavior of the system (4.1). Observe that the origin is an equilibrium, because $f_o(0) = 0$. And since the Jacobian of f_o at 0 is A and A has negative eigenvalues, 0 is a stable equilibrium. On the other hand, the system (4.1) has a linear part, $\dot{x}^d = A^d x^d$. Since A is block triangular, A^d

also has negative eigenvalues. Therefore, there are positive constants C and μ such that

$$|x^d(t)| \leq C|x^d(0)|e^{-\mu t}. \tag{4.2}$$

This implies that the stability region S of the origin, which is invariant and open, contains $\{0\} \times \mathbb{R}^{n-m}$. Lemma 4.3.1 provides information about equilibria of (4.1).

Lemma 4.3.1 *For the system (4.1), the following statements hold:*

1. *The number of equilibria is finite.*
2. *If η is an equilibrium, then $\eta^d = 0$ and $\eta^v \in \prod_{i=1}^{m}[0, -A_{ii}]$.*
3. *The origin is the only stable equilibrium.*

Proof *Part 1:* We show that the set

$$X = \left\{x \in \mathbb{C}^n : Ax + B(x_1^2, \ldots, x_n^2) = 0\right\}$$

is compact in the usual Euclidean topology. Since elements of an affine algebraic set in $\mathbb{C}^n$ are finitely many, the number of equilibria which are elements of X is finite. See Lemma 12.4.3 of Sommese and Wampler (2005).

By the definitions of A and B,

$$X = \left\{(x^v, x^d) \in \mathbb{C}^n : A^v x^v + A^c x^d + (x_1^2, \ldots, x_m^2) + B^c(x_{m+1}^2, \ldots, x_n^2) = 0,\ A^d x^d = 0\right\}.$$

Since A^d is invertible, $x^d = 0$. Thus, $X = \left\{(x^v, 0) : A^v x^v + (x_1^2, \ldots, x_m^2) = 0\right\}$. Therefore, X is compact in $\mathbb{C}^n$ if and only if $X' = \left\{x \in \mathbb{C}^m : A^v x + (x_1^2, \ldots, x_m^2) = 0\right\}$ is compact in $\mathbb{C}^m$. Rewriting the equation via $x = \alpha + i\beta$, here $i = \sqrt{-1}$, we get

$$A^v\alpha + \begin{pmatrix} \alpha_1^2 - \beta_1^2 \\ \vdots \\ \alpha_m^2 - \beta_m^2 \end{pmatrix} = 0, \quad A^v\beta + 2\begin{pmatrix} \alpha_1\beta_1 \\ \vdots \\ \alpha_m\beta_m \end{pmatrix} = 0.$$

Suppose $\{|\alpha| : \alpha + i\beta \in X'\}$ is unbounded. Then we can choose α with $\max_k |\alpha_k|$ being arbitrarily large. Let us assume that $i = i(\alpha)$ is the index that $|\alpha_i|$ achieve the maximum magnitude among $|\alpha_1|, \ldots, |\alpha_m|$. We then observe

$$\beta_i^2 = \alpha_i^2 + \sum_k A_{ik}^v \alpha_k \geq \alpha_i^2 - \sum_k |A_{ik}^v|\,|\alpha_k| \geq \alpha_i^2 - \sum_k |A_{ik}^v|\,|\alpha_i| \geq \alpha_i^2 - mM\,|\alpha_i|$$

where $M = \max_{j,k} |A_{jk}^v|$. Also, for any index j,

$$\beta_j^2 = \alpha_j^2 + \sum_k A_{jk}^v \alpha_k \leq \alpha_j^2 + \sum_k |A_{jk}^v|\,|\alpha_k| \leq \alpha_i^2 + \sum_k |A_{jk}^v|\,|\alpha_i| \leq \alpha_i^2 + \sum_k |A_{jk}^v|\,\alpha_i^2 = C_j \alpha_i^2$$

where $C_j = \sum_k |A_{jk}^v| + 1$. Since we have $2\alpha_i\beta_i = -\sum_k A_{ik}^v \beta_k$, using above two inequalities,

$$2|\alpha_i|\,|\beta_i| \leq \sum_k |A_{ik}^v|\,|\beta_k| \leq M \sum_k |\beta_k| \leq M \sum_k \sqrt{C_k}|\alpha_i|$$

and from the first inequality, $4\alpha_i^2\beta_i^2 \geq 4\alpha_i^2(\alpha_i^2 - mM|\alpha_i|)$. Therefore,

$$4\alpha_i^2(\alpha_i^2 - mM|\alpha_i|) \leq (M \sum_k \sqrt{C_k})^2 \alpha_i^2.$$

And this implies $|\alpha_i|$ cannot be arbitrarily large, which is a contradiction.

On the other hand, we showed above that $\beta_j^2 \leq C_j \alpha_i^2$ for any index j if $|\alpha_i| = \max_k |\alpha_k|$. Thus $\{|\beta| : \alpha + i\beta \in X'\}$ is also bounded. Consequently, X' is compact because X' is clearly closed as a zero set of finitely many polynomials, in addition to being bounded.

Part 2: An equilibrium η is a solution of $A\eta + B(\eta_1^2, \ldots, \eta_m^2) = 0$. This implies $\eta^d = 0$ and $\eta^v = -(A^v)^{-1}(\eta_1^2, \ldots, \eta_m^2)$. Recall that $-A^v$ (or equivalently, the transpose of it) is an M-matrix. Then as mentioned at the beginning of this section, $-(A^v)^{-1} \geq 0$ and thus $\eta^v \geq 0$. For each $i = 1, \ldots, m$, $\eta_i^2 + A_{ii}\eta_i = -\sum_{k \neq i} A_{ik}\eta_k$. Since the off-diagonal entries of A^v are non-negative and $\eta \geq 0$, the right side is not positive. Therefore, $0 \leq \eta_i \leq -A_{ii}$.

Part 3: Assume that a non-zero equilibrium η is a stable equilibrium (thus hyperbolic)

and consider its Jacobian $J(\eta) = A + 2Bdiag(\eta)$. Then

$$J(\eta)\eta = (\eta_1^2, \ldots, \eta_m^2, 0, \ldots, 0) \geq 0.$$

And $J(\eta)$ looks like

$$J(\eta) = \begin{pmatrix} A^v + 2diag(\eta^v) & A^c \\ 0 & A^d \end{pmatrix}.$$

The eigenvalues of $J(\eta)$ are those of $J(\eta)^v$ and A^d. Since η is a stable equilibrium, $J(\eta)$ is nonsingular and every eigenvalue of $J(\eta)^v$ has negative real part. Note also that the off-diagonal entries of $J(\eta)^v$ are non-negative. This implies that $-J(\eta)^v$ is a nonsingular M-matrix (see p.135 of Berman and Plemmons 1994). However, this cannot happen as the following argument shows.

Suppose $-J(\eta)^v$ is an M-matrix. We showed above that $J(\eta)^v\eta^v \geq 0$. Thus $-J(\eta)^v(-\eta^v) \geq 0$ and this, in turn, implies $-\eta^v \geq 0$. However, this together with part 2 leads to $\eta = 0$, which is a contradiction to our assumption that η is non-zero. ∎

Recall that we introduced the blow-up regions S_t in Section 4.2. We prove some topological properties of S_t's that are related to the characterization of S and ∂S.

Lemma 4.3.2 *Suppose that two real numbers M, c are given satisfying $|u^d| \leq M$, $c \leq \min_{j=1,\ldots,m} u_j$. If u_i for fixed $i \in \{1, \ldots, m\}$, fixing everything else, is sufficiently large, then (4.1) blows up in finite time. Moreover, $\tau(u)$ can be bounded above by a function of M, c and u_i, and this bound can be made arbitrarily small by increasing u_i.*

Proof It is shown in Lemma A.2.1 that $\min_{j=1,\ldots,m} x_j(t)$ is well defined and it is bounded below by some function $v(t)$. And this dynamics of $v(t)$ depends on the bound of $|u^d|$, here M, and $v(0) = c$ which can be set as any value less than $\min_{j=1,\ldots,m} u_j$. Then the trajectory of $v(t)$ is bounded below, say by $L = L(M, c)$. By (4.2), $|x^d(t)|$ is bounded by $C|u^d|$. Then for any

$i \in \{1, \ldots, m\}$, we have

$$\begin{aligned} \dot{x}_i &= x_i^2 + A_{ii}x_i + \sum_{k \neq i} A_{ik}x_k + \sum_{k=m+1}^{n} B_{ik}x_k^2 \\ &\geq x_i^2 + A_{ii}x_i + \sum_{k=1,\neq i}^{m} A_{ik}L + \sum_{k=m+1}^{n} A_{ik}x_k \\ &\geq x_i^2 + A_{ii}x_i + C_iL - K \end{aligned}$$

where $C_i = \sum_{k=1,\neq i}^{m} A_{ik}$ and $K = C|u^d| \max_i \sum_{k=m+1}^{n} |A_{ik}|$.

Then we define a new function y by an ODE

$$\dot{y} = y^2 + A_{ii}y + C_iL - K, \quad y(0) = u_i. \tag{4.3}$$

With $D_i := A_{ii}^2 - 4(C_iL - K)$, the function y blows up in finite time $\overline{\tau}$ if $y(0)$ is large. If $D_i = 0$, then y blows up at $\overline{\tau} = (y(0) - A_{ii}/2)^{-1}$ if $y(0) > -A_{ii}/2$. And if $D_i < 0$, then it does so at time

$$\overline{\tau} = \frac{1}{\sqrt{-D_i}} \left(\pi - 2 \tan^{-1} \frac{2y(0) + A_{ii}}{\sqrt{-D_i}} \right).$$

Finally, if $D_i > 0$, then

$$\overline{\tau} = \frac{1}{\sqrt{D_i}} \log \frac{y(0) - \eta_1}{y(0) - \eta_2}$$

where η_i are two equilibria of (4.3). We see that in any case the blow-up time goes to zero as $y(0) = u_i$ increases. Since we have $x_i(t) \geq y(t)$, $\overline{\tau}$ is an upper bound of the blow-up time of x_i, and consequently an upper bound of $\tau(u)$. ∎

Lemma 4.3.3 *$\tau(u) : \mathbb{R}^n \to (0, \infty]$ is continuous.*

Proof Suppose that $\{u_k\}$ is a sequence of vectors converging to u but $\lim_k \tau(u_k) = \tau^* > \tau(u)$. Since $\lim_{t \uparrow \tau(u)} \max_j \Phi_{t,j}(u) = \infty$ and since Φ is continuous in Ω, we can find $t' = t'(N) < \tau(u)$ and $k = k(N)$ for any given large N such that $N < \max_j \Phi_{t',j}(u_k) < \infty$.

Note that some component $x_i(t)$ of a solution $x(t)$ to (4.1) never decreases if the initial condition $x_i(0)$ is sufficiently large where $i \in \{1, \ldots, m\}$. To see this, find a function $v(t)$ as in

Lemma A.2.1 that is bounded below and $x_i(t) \geq v(t)$. Observe that

$$\dot{x}_i = x_i^2 + \sum_k A_{ik}x_k + \sum_{k=m+1}^{n} B_{ik}x_k^2 \geq x_i^2 + A_{ii}x_i + \sum_{k\neq i,=1}^{m} A_{ik}L - K$$

where $L = L(M,c)$ is a lower bound of v and $K = \sum_{k=m+1}^{n} C|A_{ik}| \cdot |x^d(0)|$ and C is a constant in (4.2). Therefore, it is enough to have a large $x_i(0)$ that makes the right side positive.

This observation implies that for any given large N, we can choose some $t' < \tau(u)$, k and i such that $\Phi_{t',i}(u_k) > N$ and this does not decrease from time t'. Thus, $\Phi_{\tau(u),i}(u_k) > N$. Since this is true for any large N, we conclude $\limsup_k \Phi_{\tau(u),i}(u_k) = \infty$. Now consider $\Phi_t(u_k)$ starting from $\tau(u)$. Then, by Lemma 4.3.2, the blow up time of $\{\Phi_t(u_k)\}_{t\geq\tau(u)}$ can be arbitrarily close to $\tau(u)$ by selecting a large k. This is a contradiction to $\tau^* > \tau(u)$.

To prove the converse, suppose that $\lim_k \tau(u_k) = \tau^* < \tau(u)$. We take $t \in (\tau^*, \tau(u))$. Then $\Phi_t(u)$ is finite. Since Ω is open and $(t,u) \in \Omega$, (t,u_k) belongs to Ω for all large k's. Thus $\Phi_t(u_k)$ is finite and this is a contradiction to $\tau(u_k) < \tau(u)$ for large k's. ∎

Lemma 4.3.4 *Suppose $\tau(u) < \infty$. Then $\tau(\theta u) < \tau(u)$ for $\theta > 1$.*

Proof Let $y(t) = \Phi_t(\theta u)/\theta$ and $x(t) = \Phi_t(u)$. Then (3.4.1) implies that we always have $y(t) \geq x(t)$. We note that $y(t)$ satisfies $\dot{y} = Ay + \theta B(y_1^2, \ldots, y_n^2)$, $y(0) = u$. One implication of this is that $y^d(t) = x^d(t)$. Let $\tilde{x}(t) = \max_k x_k(t)$, which is a well-defined piecewise differentiable function in a similar way as in the proof of Lemma A.2.1. Then by assumption, $\tilde{x}$ blows up in finite time, say τ. Therefore, in the following argument we can assume that the initial value of $\tilde{x}(t)$ starting from time t_0, $\tilde{x}(t_0)$, is a sufficiently large positive real number and $|x^d(t)|/\tilde{x}(t)$ is sufficiently small, say less than ϵ whenever $t \geq t_0$. And we note that if we start from this sufficiently large initial value, then $\tilde{x}$ never decreases during its entire life span $[t_0, \tau)$. Finally, note that we can find a number M such that $|x^d(t)| \leq M$ for any t.

On this footing, we claim $\int_{t_0}^{\tau(u)} \tilde{x}(t)dt = \infty$. Let $\tilde{x} = x_i$ in some interval $I \subset [t_0, \tau(u))$. Then,

$$\frac{\dot{\tilde{x}}}{\tilde{x}} = \sum_{k=1}^{m} A_{ik}\frac{x_k}{\tilde{x}} + \sum_{k=m+1}^{n} A_{ik}\frac{x_k}{\tilde{x}} + \tilde{x} + \sum_{k=m+1}^{n} B_{ik}\frac{x_k^2}{\tilde{x}} \leq \sum_{k=1}^{m} A_{ik} + \tilde{x} + C \leq \tilde{x} + C_1$$

for some constant C, C_1 independent of i. This implies that

$$\log \tilde{x}(t) \leq \log \tilde{x}(t_0) + \int_{t_0}^{t} \tilde{x}(s)ds + C_1(t - t_0),$$

which proves the claim.

Now we define $\gamma(t) = \tilde{y}(t)/\tilde{x}(t) \geq 1$, which is also piecewise C^1, and $\tilde{y}(t)$ means $y_i(t)$ whenever $\tilde{x}(t) = x_i(t)$. We observe in the interval which $\tilde{x} = x_i(t)$ with $i \in \{1, \ldots, m\}$, γ satisfies

$$\begin{aligned}
\dot{\gamma} &= \frac{\dot{y}_i x_i - y_i \dot{x}_i}{x_i^2} \\
&= \frac{x_i(\theta y_i^2 + \sum_k A_{ik} y_k + \theta \sum_{k=m+1}^{n} B_{ik} y_k^2) - y_i(x_i^2 + \sum_k A_{ik} x_k + \sum_{k=m+1}^{n} B_{ik} x_k^2)}{x_i^2} \\
&= \theta x_i \gamma^2 - y_i + \sum_{k \neq i, =1}^{m} A_{ik}\Big(\frac{y_k}{x_i} - \frac{x_k}{x_i}\gamma\Big) + \sum_{k=m+1}^{n} A_{ik}\frac{x_k}{x_i}(1-\gamma) + \sum_{k=m+1}^{n} B_{ik}\frac{x_k^2}{x_i}(\theta - \gamma) \\
&\geq \theta x_i \gamma^2 - y_i + \sum_{k \neq i, =1}^{m} A_{ik}\Big(\frac{x_k}{x_i} - \frac{x_k}{x_i}\gamma\Big) + \sum_{k=m+1}^{n} |A_{ik}|\epsilon(1-\gamma) + \sum_{k=m+1}^{n} B_{ik} M\epsilon(1-\gamma) \\
&= \theta x_i \gamma^2 - x_i \gamma + \sum_{k \neq i, =1}^{m} A_{ik}\Big(\frac{x_k}{x_i}\Big)\Big(1-\gamma\Big) + \sum_{k=m+1}^{n} |A_{ik}|\epsilon(1-\gamma) + \sum_{k=m+1}^{n} B_{ik} M\epsilon(1-\gamma) \\
&\geq \theta x_i \gamma^2 - x_i \gamma + \sum_{k \neq i, =1}^{m} A_{ik}\Big(1-\gamma\Big) + \sum_{k=m+1}^{n} |A_{ik}|\epsilon(1-\gamma) + \sum_{k=m+1}^{n} B_{ik} M\epsilon(1-\gamma)
\end{aligned}$$

and this can be written as

$$\dot{\gamma} \geq \theta x_i \gamma^2 - x_i \gamma + C(1-\gamma) \tag{4.4}$$

where C is an appropriate non-negative constant. In the first inequality, we used that $A_{ik} \geq 0$ for $k \in \{1, \ldots, m\} \backslash \{i\}$, $y \geq x$, $\gamma \geq 1$, $y^d = x^d$, $|x^d|/\tilde{x} \leq \epsilon$, $|x^d| \leq M$, $B_{ik} \geq 0$ and x_i is very large. In the second inequality, we utilized the fact that in the interval we are considering, $x_k \leq x_i$ for all k and $\gamma \geq 1$. Here we note that if $\gamma(t_0) = 1$ with $\tilde{x}(t_0)$ being a large positive number, then $\dot{\gamma}(t_0) \geq (\theta - 1)x_i(t_0) > 0$, so $\gamma(t_0 + \epsilon) > 1$ for small ϵ. Therefore, we can assume

$\gamma(t_0) > 1$ from the beginning. Then we see

$$\dot{\gamma} \geq x_i\gamma^2 - x_i\gamma + C\big(1-\gamma\big) = (\gamma-1)(x_i\gamma - C) \geq (\gamma-1)(x_i - C) = (\gamma-1)(\tilde{x} - C).$$

Recall that $\tilde{x}(t_0)$ is assumed to be sufficiently large and thus $\tilde{x}$ never decreases in $[t_0, \tau)$. Now the above inequality with $\gamma(t_0) > 1$ implies

$$\frac{d}{dt}\ln(\gamma - 1) \geq \tilde{x} - C.$$

Thus $\gamma(t) - 1 \geq (\gamma(t_0) - 1)\exp(\int_{t_0}^{t}(\tilde{x}(s) - C)ds)$. Therefore, $\gamma \to \infty$ as t approaches τ. Hence, we can assume $\gamma(t_0) > 2$ by shifting the starting point, i.e., by taking even larger t_0 as a starting point.

Inequality (4.4) also yields

$$\dot{\gamma} \geq \tilde{x}\gamma^2 - \tilde{x}\gamma + C(1-\gamma) \geq \tilde{x}\gamma^2 - \tilde{x}\gamma - C\gamma \geq \tilde{x}\gamma^2 - 2\tilde{x}\gamma.$$

This and $\gamma(t_0) > 2$ imply

$$\frac{d}{dt}\ln\frac{\gamma - 2}{\gamma} \geq 2\tilde{x}.$$

And this leads to

$$\gamma(t)\Big(1 - \frac{\gamma(t_0) - 2}{\gamma(t_0)}\exp(2\int_{t_0}^{t}\tilde{x}(s)ds)\Big) \geq 2.$$

Therefore, γ blows up strictly before t reaches τ. In other words, $y(t)$ blows up strictly before $x(t)$ does. ∎

Proposition 4.3.1 *The blow-up region S_t is a closed convex proper subset of $\mathbb{R}^n$. The topological boundary of S_t is given by $\partial S_t = \{u \in S_t : \tau(u) = t\}$ and $S_{t'} \subset S_t^o$ for $t' > t$.*

Proof Recall that S_t is a set in which a solution to (4.1) with $x(0) = u$ exists in $[0, t)$. In other words,

$$S_t = \{u : \mathbb{E}\exp(2u \cdot Y_s) < \infty, \quad \forall s \in [0, t)\} = \{u : \tau(u) \geq t\}.$$

For $u, v \in \mathbb{R}^n$ and $\lambda \in (0,1)$, by Hölder's inequality,

$$\mathbb{E}\exp\left(2(\lambda u + (1-\lambda)v)\cdot Y_s\right) \le \left[\mathbb{E}\exp(2u\cdot Y_s)\right]^{\lambda}\left[\mathbb{E}\exp(2v\cdot Y_s)\right]^{1-\lambda}.$$

Therefore, $\lambda u + (1-\lambda)v \in S_t$ whenever $u, v \in S_t$. Since $\tau(u)$ is continuous in u by Lemma 4.3.3, S_t is closed and it is also proper by Lemma 4.3.2; if $S_t = \mathbb{R}^n$, then we simply choose $u = (0,\ldots,0,u_i,0,\ldots,0)$ and let u_i go to infinity.

To prove the second statement, let $T = \{u \in S_t : \tau(u) = t\}$. Since $\tau(u)$ is continuous in u, T is closed. For each $u_0 \in T$ and in any small open ball U centered at u_0, we can choose two positive real numbers $\theta_1 > 1$ and $\theta_2 < 1$ such that $\theta_i u_0 \in U$. By Lemma 4.3.4,

$$\tau(\theta_1 u_0) < \tau(u_0) = t < \tau(\theta_2 u_0).$$

Thus, $\theta_2 u_0 \in S_t \backslash T = \{u \in S_t : \tau(u) > t\}$, which is open, and $\theta_1 u_0 \in S_t^c$, the complement of S_t which is also open. Since $S_t \backslash T$ is an open subset of S_t, it is included in S_t^o. Conversely, any $u \in S_t^o$ is in $S_t \backslash T$ because we can find some $\theta > 1$ such that $\theta u \in S_t$ and consequently $\tau(u) > \tau(\theta u) \ge t$. Therefore, we conclude that $S_t \backslash T = S_t^o$ and $T = \partial S_t$. Hence, for $t' > t$,

$$S_{t'} = \{u : \tau(u) \ge t'\} \subset \{u : \tau(u) > t\} = S_t^o.$$

∎

4.4 Characterization of the Stability Boundary

From Proposition 4.3.1, we conclude that $S_\infty := \bigcap_t S_t$ is closed and convex. Since $S \subset S_t$ for all t, we have $\overline{S} \subset S_\infty$. The next theorem is our first main result.

Theorem 4.4.1 *Suppose that every bounded trajectory of (4.1) converges to an equilibrium. Then for a hyperbolic equilibrium η, we have $\eta \in \partial S$ if and only if $W_\eta^s \subset \partial S$. Moreover, $\overline{S} = S_\infty$.*

Proof One direction is trivial. For the other direction, suppose $\eta \in \partial S$. Choose a point $u \in W_\eta^s$. If it is on ∂S, then there is nothing to prove. If it is in S, then it converges to 0. So this case cannot happen. Thus we assume $u \notin \overline{S}$.

Let u' be a point on the intersection of ∂S and a line segment connecting u and the origin. Then there is $\theta \in (0,1)$ such that $u' = \theta u$. Then $\Phi_t(u') \leq \theta \Phi_t(u)$ by (3.4.1) (they exist at any time t because $\lim_{t\to\infty} \Phi_t(u) = \eta$ and $\Phi_t(u')$ cannot escape $\mathbb{R}^n$ in finite time: it is bounded above by $\Phi_t(u)$ and bounded below by Lemma A.2.1). This implies $\lim_t \Phi_t(u') \leq \theta\eta$. By assumption, $\lim_t \Phi_t(u')$ is a non-zero equilibrium on ∂S (∂S is an invariant set). Let us call this η'. By this and Lemma 4.3.1, we have $0 \leq \eta' \leq \theta\eta$, $\eta^d = \eta'^d = 0$ and η^v, η'^v are solutions of $A^v x + (x_1^2, \ldots, x_m^2) = 0$. Then,

$$-\theta^2 A^v \eta^v = \theta^2(\eta_1^2, \ldots, \eta_m^2) \geq ({\eta'_1}^2, \ldots, {\eta'_m}^2) = -A^v \eta'^v.$$

This means $-A^v(\theta^2\eta^v - \eta'^v) \geq 0$. Therefore, $\theta^2\eta^v \geq \eta'^v$ thanks to the fact that $-A^v$ is a nonsingular M-matrix. Repeated application of this procedure yields $\theta^{2k}\eta \geq \eta'$ for any integer k. Since $\theta < 1$ and $\eta' \geq 0$, $\eta' = 0$. But, this is a contradiction to the assumption that η' is on ∂S because ∂S does not contain the origin.

Let us prove the second statement. Suppose $u \in S_\infty \backslash \overline{S}$. We claim that $\{\Phi_t(u) : t \geq 0\}$ is bounded in $\mathbb{R}^n$. We know that each component of $\Phi_t(u)$ is bounded below by some number, say c, and $|\Phi_t^d(u)|$ is bounded by some number M. Suppose it is not bounded above. Since $\Phi_t^d(u)$ converges to zero, there is some $i \in \{1, \ldots, m\}$ such that $\{\Phi_{t,i}(u)\}_{t\geq 0}$ is not bounded above. Then, Lemma 4.3.2 implies that $\Phi_{t,i}(u)$ blows up in finite time. This is a contradiction because $u \in S_\infty$ and thus $\Phi_t(u)$ exists at all times t. Since the trajectory is bounded, it converges to an equilibrium point by assumption and the equilibrium must be non-zero. However, this cannot happen by the same argument as above. Therefore, $\overline{S} = S_\infty$. ∎

Corollary 4.4.1 *Suppose that every bounded trajectory of (4.1) converges to an equilibrium and that the system has hyperbolic equilibria only. Then,*

$$\partial S = \partial S_\infty = \bigcup_{\eta \neq 0} W_\eta^s.$$

Proof Since a non-zero equilibrium η is in $S_\infty \backslash S$, it is on ∂S by the previous proposition. Then, the result is immediate. ∎

The corollary above implies that the stable manifolds of the UEPs of type 1 determine $\partial S = \partial S_\infty$ except for a set of measure zero (recall $\dim W^s_\eta = n - k$ for an UEP η of type k). There are many numerical methods addressing how to compute the stable manifolds of equilibria. But, we will not discuss this problem in this paper. An interested reader can consult Cheng et al. (2004) or Saha et al. (1997), for example.

It is a "generic" property (in the sense that a property holds true in the countable intersection of open dense subsets) to admit hyperbolic equilibria only for C^r ($r \geq 1$) vector fields (see Chiang et al. 1988 or Smale 1967 and references therein). The next examples show that the models which satisfy the first assumption of Corollary 4.4.1 are not empty, but rather ample.

Example $\mathbb{A}_2(n)$. Let $A^v = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. After simple calculations, we find that the following conditions are necessary and sufficient for $-A^v$ to be a non-singular M-matrix.

$$a < 0,\ b \geq 0,\ c \geq 0,\ d < 0,\ ad - bc > 0.$$

From these conditions, it is straightforward to determine conditions for the system to have two, three or four equilibria. And one can check that three equilibria case happens only when the two parabolas $x^2 + ax + by = 0$ and $y^2 + cx + dy = 0$ are tangent in the (x, y)-plane.

Berlinskiĭ's Theorem (see Chicone and Shafer 1983) has an implication about the hyperbolicity of an equilibrium of our system. It states that if X is a quadratic vector field in the plane with two relatively prime quadratic polynomials, which is, by the way, satisfied by our system, and if X has four equilibria, then the Jacobian determinant at each one is nonzero and every saddle point is hyperbolic. Moreover, if the quadrilateral with vertices at the critical points is convex, then two opposite vertices are saddles and the other two are anti-saddles (nodes, foci or centers).

Lemma 4.4.1 *In* $\mathbb{A}_2(n)$, *every bounded trajectory is an equilibrium point or converges to an equilibrium.*

Proof First we observe that $\Phi_t(u_1) \leq \Phi_t(u_2)$ whenever $u_1^v \leq u_2^v$ and $u_1^d = u_2^d$. To see this, we

define

$$x(t) = \Phi_t^v(u_1), \quad y(t) = \Phi_t^v(u_1), \quad z(t) = \Phi_t^d(u_1) = \Phi_t^d(u_2).$$

Then, they satisfy

$$\dot{x} - f(x) = A^c z + B^c \begin{pmatrix} z_1^2 \\ \dots \\ z_{n-m}^2 \end{pmatrix} = \dot{y} - f(y)$$

where $f(x) = A^v x + (x_1^2, \dots, x_m^2)$, which is quasi-monotone increasing and locally Lipschitz. Since $x(0) = u_1^v \leq u_2^v = y(0)$, the result follows from (3.25).

For the discussion that follows next, we refer to Chapter 4 of Verhulst (1996) for the results that are related. Suppose we have a bounded trajectory which is not an equilibrium, neither does it converge to an equilibrium. Then, the set of the limit points of the trajectory, the ω-limit set Γ is invariant, compact, connected and not empty. Moreover, there is a minimal (i.e., closed, invariant and nonempty with no smaller subsets with these properties) subset $K \subset \Gamma$. Since $\Phi_t^d(u) \to 0$ as $t \to \infty$, K is decomposed as $K_0 \times \{0\}$ with $K_0 \subset \mathbb{R}^2$. This means that K_0 itself is a minimal set of the system

$$\dot{x} = \begin{pmatrix} x_1^2 \\ x_2^2 \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix} x.$$

Then by the Poincaré-Bendixon theorem, K_0 is a periodic orbit. Note that this orbit is not self-intersecting because this curve cannot have two different derivatives at an intersection. Now the Jordan Curve Theorem implies that

$$\mathbb{R}^2 \backslash K_0 = K_0^o \cup K_0^c$$

where K_0^o is the inside of the orbit and K_0^c is the outside of the orbit and they are open. Choose one point, say $p \in K_0^o$ and define

$$K_1 = \{p + x : x \in \mathbb{R}_+^2\} \cap K_0, \quad K_2 = \{p + x : x \in \mathbb{R}_-^2\} \cap K_0.$$

Clearly K_i's are compact and non-empty. Let us now choose the maximizer of $\max_{u\in K_2}|u-p|$, say u_2. Then there is no $u \in K_0$ such that $u < u_2$. Also we choose any $u_1 \in K_1$. Then we have $u_2 < u_1$. Since K_0 is a closed orbit, there is some t_0 such that $\Phi^v_{t_0}(u_1) = u_2$. By the observation made at the beginning of the proof,

$$\Phi^v_{t_0}(u_2) \le \Phi^v_{t_0}(u_1) = u_2.$$

By the choice of u_2, the left hand side (which is in K_0 by invariant property) is not strictly less than u_2. Therefore, $\Phi^v_{t_0}(u_2) = u_2$. However, this implies that t_0 is a multiple of the period of K_0 but $\Phi^v_{t_0}(u_1) = u_2 \neq u_1$. This is a contradiction. ∎

Proposition 4.4.1 *In $\mathbb{A}_2(n)$, every equilibrium is hyperbolic unless there are three equilibria. Moreover, if we have two equilibria, then one is the origin as a focus and the other is on the stability boundary of the origin as an UEP of type 1. And if we have four equilibria, then we have one focus at the origin, an UEP of type 1, an UEP of type 2 and an UEP of type 1 in the clockwise order in $\mathbb{R}^2 \times \{0\}$.*

Proof We consider $bc \neq 0$ case only. Other cases can be analyzed similarly. An equilibrium point is determined by

$$x^2 + ax + by = 0 \tag{4.5}$$

$$y^2 + cx + dy = 0 \tag{4.6}$$

Suppose $\eta = (\eta_1, \eta_2, 0, \ldots, 0)$ is an equilibrium point. Then the Jacobian of the system at this point is

$$J(\eta) = \begin{pmatrix} A^v & A^c \\ 0 & A^d \end{pmatrix} + 2 \left(\begin{array}{cc|c} \eta_1 & 0 & \\ 0 & \eta_2 & \\ \hline & & 0 \end{array}\right).$$

Since the eigenvalues of $J(\eta)$ are those of $J(\eta)^v$ and A^d, and since A^d has negative eigenvalues, it is enough for us to study the eigenvalues of $J(\eta)^v$. The characteristic polynomial of $J(\eta)^v$ is $P(\lambda) = \det(J(\eta)^v - \lambda I) = \lambda^2 - (p+q)\lambda + pq - bc$ where $p = a + 2\eta_1$ and $q = d + 2\eta_2$. Then the

determinant of this quadratic polynomial is $D = (p-q)^2 + bc > 0$. Therefore, η is hyperbolic if and only if 0 is not an eigenvalue of $J(\eta)^v$, i.e., $pq = bc$.

Consider the two parabolas (4.5), (4.6). From each, we get

$$\frac{dy}{dx}(\eta) = -\frac{p}{b}, \quad \frac{dy}{dx}(\eta) = -\frac{c}{q}.$$

Therefore, two parabolas become tangent at η if and only if $pq = bc$. Then we have exactly three equilibria. Hence, all critical points are hyperbolic except three equilibria case.

Now suppose we have two equilibria. Recall that ∂S is a closed invariant subset of $\mathbb{R}^n$ of dimension $n-1$ (see Chiang et al. 1988). The trajectory $\{\Phi_t(u) : t \geq 0\}$ is bounded for $u \in \partial S$ and it is contained in ∂S. This is because first, $\partial S \subset S_\infty$ so $\Phi_t(u)$ does not blow up in finite time and second, any $|\Phi_{t,i}(u)|$ cannot be arbitrarily large due to Lemma 4.3.2. Therefore, by Lemma 4.4.1, it converges to an equilibrium. Since ∂S does not contain the origin, it is the other equilibrium point, say η. In other words,

$$\lim_{t\to\infty} \Phi_t(u) = \eta$$

whenever $u \in \partial S$. But ∂S is of dimension $n-1$ and this means $\dim W^s(\eta) \geq n-1$. However, η cannot be a focus because the origin is the only stable equilibrium as implied by Lemma 4.3.1. Thus $\dim W^s(\eta) = n-1$ and so η is an UEP of type 1.

To prove the last statement, suppose that there are four equilibria. Then, two parabolas (4.5) and (4.6) have four solutions and since they form a convex quadrilateral, Berlinskiĭ's Theorem applies and we conclude that two opposite vertices are saddles and the other two are antisaddles. Since the origin is the only stable equilibrium by Lemma 4.3.1, we have one focus at the origin, a saddle, a source and a saddle clockwise. This implies that the number of negative eigenvalues (we know that they have real eigenvalues) of the Jacobian $J(\eta)^v$ are 2, 1, 0 and 1, respectively. Therefore, the Jacobian $J(\eta)$ of the original system has n, $n-1$, $n-2$ and $n-1$ number of negative eigenvalues, respectively. ∎

Example $\mathbb{A}_m(n)$ with symmetric A^v. Suppose that a given trajectory is bounded. Then we know that the limit set lies in $\mathbb{R}^m_+ \times \{0\}$. However, when A^v is symmetric, the system

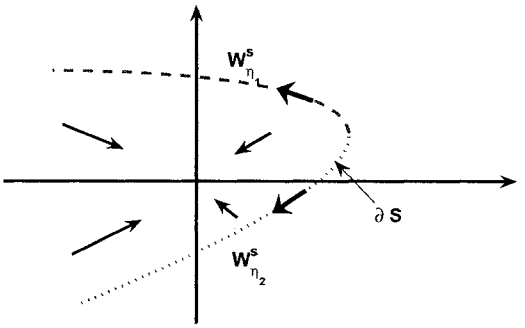


Figure 4.1: A general picture of ∂S with two hypothetical UEPs of type 1.

$\dot{x} = A^v x + (x_1^2, \ldots, x_m^2)$ becomes a gradient system. In other words, $\dot{x} = \nabla V(x)$ with

$$V(x) = \frac{1}{2} x^\top A^v x + \frac{1}{3} \sum_i x_i^3.$$

Let Γ be the ω-limit set of the trajectory. Then, Γ is invariant, compact, connected, and not empty. Since it is in $\mathbb{R}^m_+ \times \{0\}$, we can decompose it as $\Gamma_0 \times \{0\}$, where Γ_0 is a compact invariant and connected set of the above gradient system. However, it is well known that a point in a limit set of a gradient system is an equilibrium. See p.203 of Hirsch and Smale (1974). Therefore, Γ_0 consists of equilibria. Since (4.1) has only a finite number of equilibria, Γ_0 and so Γ is a single equilibrium and we conclude that every bounded trajectory converges to an equilibrium.

Based on these results, we can draw a simple picture of the stability region of (4.1). Figure 4.1 shows ∂S consisting of two stable sub-manifolds of UEPs of type 1. The intersection of two sub-manifolds can be thought of as a source or a sub-manifold with dimension less than $n - 1$.

4.5 Asymptotic Behavior of Blow-up Times and Application

In this section, we describe the blow-up times of (4.1) as a solution of a partial differential equation (PDE). Recall that $\partial S_t = \{u : \tau(u) = t\}$. In other words, ∂S_t's are the level sets of the function $\tau(u) : \mathbb{R}^n \to (0, \infty]$. Since $S_t = \{u : \tau(u) \geq t\}$ and S_t is convex, we conclude that

$\tau(u)$ is quasi-concave. (This is a standard result in convex analysis.) However, $\tau(u)$ is not necessarily differentiable as we see in the next example:

$$\dot{x}_1 = ax_1 + x_1^2, \quad \dot{x}_2 = bx_2 + x_2^2 \tag{4.7}$$

with $x(0) = u$ and a, b are negative. In this case, $S_t = (-\infty, -a/(1-e^{at})] \times (-\infty, -b/(1-e^{bt})]$ and ∂S_t is not smooth at the vertex.

Proposition 4.5.1 *The continuous function $\tau|_{S_\infty^c} : S_\infty^c \to (0, \infty)$ is quasi-concave, and is differentiable almost everywhere satisfying*

$$\nabla\tau(u) \cdot f_o(u) = -1 \tag{4.8}$$

for $u \in S_\infty^c$ with a boundary condition $\tau(u) = \infty$ on ∂S_∞. Conversely, a function τ satisfying (4.8) and a condition that $\lim_{t\uparrow} \tau(\Phi_t(u)) = 0$ for all $u \in S_\infty^c$ is unique.

Proof Differentiability is a direct application of the result by Crouziex (1982). For a fixed $u \in S_\infty^c$, we can choose a small positive h such that $\Phi_h(u)$ is finite because $I(u)$ is an open set containing 0. Then we have

$$\tau(u) - h = \tau(\Phi_h(u)).$$

By differentiating this with respect to h at $h = 0$, we get (4.8). The boundary condition is obvious.

To prove the last statement, suppose $\tau_1(u)$ and $\tau_2(u)$ are two solutions of (4.8). Then, we can construct $\tau(u) = \tau_1(u) - \tau_2(u)$ on S_∞^c and this satisfies $\nabla\tau(u) \cdot f_o(u) = 0$. For any $u \in S_\infty^c$ and a positive $h \in I(u)$,

$$\tau(\Phi_h(u)) - \tau(u) = \int_0^h \nabla\tau(\Phi_s(u)) \cdot f_o(\Phi_s(u))ds = 0.$$

Therefore, $\tau(u)$ is constant on each trajectory $\{\Phi_t(u) : t \geq 0\}$ for $u \in S_\infty^c$. Since $\tau(\Phi_t(u))$ converges to zero as t approaches the blow-up time of $\Phi_t(u)$ by the assumption, $\tau(u)$ must be identically zero. ∎

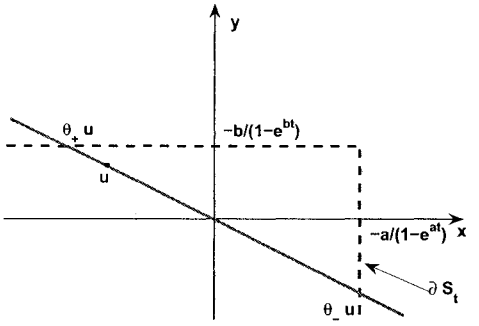


Figure 4.2: Inverse of $\tau(u)$ for (4.7).

We are also interested in determining the *critical multiplier* θ such that $\sup I(\theta u) = t$ for given u and t. This is the inverse function of $\tau(\theta u) = t$, but this function $\theta(u,t)$ may not be single-valued. For example, consider (4.7). If u belongs to the second quadrant, then there are positive and negative θ such that $\theta u \in \partial S_t$. If u is in the first quadrant, then there is only one such θ. And there is no θ for any t if $u = 0$. See Figure 4.2. Nevertheless, the next theorem says $\theta(u,t)$ is well-defined at least locally under some regularity conditions and it becomes a solution of some PDEs.

Proposition 4.5.2 *Assume that $\tau(u)$ is continuously differentiable on some neighborhood of $u_0 \in \partial S_{t_0}$ and $\nabla\tau(u_0) \cdot u_0 \neq 0$. Then, there is some open neighborhood $U \times I$ of (u_0, t_0) such that $\theta : U \times I \to (0,\infty)$ is well-defined and it satisfies*

$$\nabla\theta \cdot u = -\theta, \quad \partial_t\theta = \frac{1}{\theta}\nabla\theta \cdot f_o(\theta u), \quad \theta(u_0,t_0) = 1.$$

Proof Suppose that $\tau(u)$ is C^1 on some neighborhood of $u_0 \in \partial S_{t_0}$. Consider a function $\phi(u,t,\theta) := \tau(\theta u) - t$ defined on some neighborhood of $(u_0,t_0,1)$. Since $\left(\partial\phi/\partial\theta\right)(u_0,t_0,1) = \nabla\tau(u_0) \cdot u_0$ and this is nonzero by assumption, we conclude that there exists a C^1 function $\theta(u,t)$ defined on some neighborhood $U \times I$ of (u_0,t_0) such that

$$\tau(\theta(u,t)u) = t \tag{4.9}$$

by the Implicit Function Theorem. Clearly, $\theta(u_0,t_0) = 1$.

If we take a derivative with respect to t in (4.9), then we get $\partial_t \theta \nabla\tau(\theta u)\cdot u = 1$. This implies $\nabla\tau(\theta u)\cdot u = 1/\partial_t u \neq 0$ on $U\times I$. And if we do the differentiation with respect to u, then

$$(\nabla\tau(\theta u)\cdot u)\nabla\theta + \theta\nabla\tau(\theta u) = 0. \tag{4.10}$$

Multiplying (4.10) by u, we have $(\nabla\tau(\theta u)\cdot u)(\nabla\theta\cdot u + \theta) = 0$. Since $\nabla\tau(\theta u)\cdot u \neq 0$, we conclude that $\nabla\theta\cdot u = -\theta$ on $U\times I$. On the other hand, if we multiply (4.10) by $f_o(\theta u)$, then from (4.8), $(\nabla\tau(\theta u)\cdot u)\nabla\theta\cdot f_o(\theta u) = \theta$ and the result follows because $\partial_t\theta = (\nabla\tau(\theta u)\cdot u)^{-1}$. ∎

Equation (4.8) and the first equation in Theorem 4.5.2 look similar to the Zubov equation for the stability region of an equilibrium (see, e.g., Genesio et al. 1985). Concisely, if C^1 functions $\xi(u)\in[0,1]$ and $\phi(u)\geq 0$ satisfy $\nabla\xi(u)\cdot f_o(u) = -\phi(u)(1-\xi(u))$, then $S = \{u : \xi(u) = 1\}$. Or equivalently,

$$\nabla\tilde{\xi}(u)\cdot f_o(u) = -\phi(u), \quad \tilde{\xi} := -\log(1-\xi)$$

and thus $S = \{u : \tilde{\xi}(u) = \infty\}$. There have been many results concerning approximation methods for the Zubov equation. However, we do not pursue this direction in this article. Instead, we prove a limiting behavior of $\tau(u/t)$ and $\theta(u,t)$ near $t = 0$ under some mild conditions.

Theorem 4.5.1 *Let $\{u_t\}$ be a sequence of vectors in $\mathbb{R}^n$ that converges to u as $t\downarrow 0$. Suppose that $u^v > 0$ or $a := B^c(u^2_{m+1},\ldots,u^2_n)\neq 0$. Then,*

$$\lim_{t\downarrow 0}\frac{1}{t}\tau\left(\frac{u_t}{t}\right) = \lim_{t\downarrow 0} t\,\theta(u_t,t) = \xi$$

where $\xi = \min_i \tau_i$. And each τ_i for $i = 1,\ldots,m$ is given by

$$\tau_i = \frac{1}{\sqrt{a_i}}\left(\frac{\pi}{2} - \tan^{-1}\frac{u_i}{\sqrt{a_i}}\right) \quad \text{if} \quad a_i > 0, \quad \tau_i = \frac{1}{u_i} \quad \text{if} \quad u_i > 0,\ a_i = 0$$

or $\tau_i = \infty$ otherwise.

Proof Consider the following system of ODEs:

$$\dot{x}(s) = B\begin{pmatrix} x_1(s)^2 \\ \vdots \\ x_n(s)^2 \end{pmatrix}, \quad x(0) = u.$$

It is straightforward to see that the above system explodes at time ξ given in the statement. And the assumptions imposed on u make ξ finite. Next, we consider a perturbed system with a parameter t:

$$\dot{y}(s) = tAy(s) + B\begin{pmatrix} y_1(s)^2 \\ \vdots \\ y_n(s)^2 \end{pmatrix}.$$

A solution of this system is continuous in t, s and $y(0)$ as noted in p.44 of Lefschetz (1957). Let us denote the solution with $y(0) = u_t$ by $y(s;\, t)$. Since u_t converges to u, $\lim_{t\downarrow 0} y(s;\, t) = x(s)$ if $x(s)$ exists.

Let ξ_t be the blow-up time of $y(\cdot\,;\, t)$. We claim $\xi_t \to \xi$ as $t \downarrow 0$. Suppose $\xi^* = \lim_{k\to\infty} \xi_{t_k} > \xi$ for some convergent sequence $\{\xi_{t_k}\}$ with $\lim_{k\to\infty} t_k = 0$. Observe that we have

$$\dot{y}^v - tA^v y^v = \begin{pmatrix} y_1^2 \\ \vdots \\ y_m^2 \end{pmatrix} + B^c \begin{pmatrix} y_{m+1}^2 \\ \vdots \\ y_n^2 \end{pmatrix} + tA^c y^d \geq tA^c y^d = tA^c z^d = \dot{z}^v - tA^v z^v$$

where $\dot{z} = tAz$ with $z(0) = u_t$. Invoking (3.25), we conclude that $y(s;\, t) \geq z(s;\, t) = \exp(tAs)u_t$ for any s and t.

Since A has negative eigenvalues, (4.2) implies $|z(s;\, t)| \leq C|u_t| \leq C(|u| + 1)$ for some positive constant C and all sufficiently small t's. Since $x(s)$ blows up at ξ, $\lim_{s\uparrow\xi} x_i(s) = \infty$ for some $i \in \{1, \ldots, m\}$. By the assumption $\xi^* > \xi$, $y(\xi;\, t_k)$ is finite and well-defined for all sufficiently large k and we get

$$\lim_{k\to\infty} y_i(\xi;\, t_k) = \infty$$

as $y(s;\, t)$ converges to $x(s)$. But we have, as in the proof of Lemma 4.3.2,

$$\begin{aligned}\dot{y}_i &= y_i^2 + tA_{ii}y_i + t\sum_{k\neq i} A_{ik}y_k + \sum_{k=m+1}^{n} B_{ik}y_k^2 \\ &\geq y_i^2 + tA_{ii}y_i + t\sum_{k\neq i} A_{ik}z_k \\ &\geq y_i^2 + tA_{ii}y_i - tK\end{aligned}$$

where $K = C(|u| + 1)\sum_{k\neq i} A_{ik}$. Assume that a new function $w(s)$ starting at ξ satisfies

$$\dot{w}(s) = w^2 + t_k A_{ii} w(s) - t_k K - 1, \quad w(\xi) = y_i(\xi;\, t_k).$$

Then, $y(s;\, t_k) \geq w(s)$ on $[\xi, \xi_{t_k})$ and we deduce $\xi_{t_k} \leq \tau_w$, the blow-up time of the function w.

Let us denote the two equilibria for $w(s)$ by $\alpha > \beta$ and choose t_k so small that $y_i(\xi;\, t_k) > \alpha$. Then $w(s)$ starting at ξ blows up in finite time

$$\tau_w = \xi + \frac{1}{\sqrt{t_k^2 A_{ii}^2 + 4t_k K + 4}} \log \frac{w(\xi) - \beta}{w(\xi) - \alpha}.$$

Since $\lim_{k\to\infty} y(\xi;\, t_k) = \infty$, $\xi_{t_k} \leq \tau_w \to \xi$ as t_k decreases. This is a contradiction and as a consequence, a limit of every convergent subsequence of $\{\xi_t\}$ is less than or equal to ξ.

Now we suppose $\xi^* < \xi$. Since $x(s) = \lim_{t\downarrow 0} y(s;\, t)$ and $x(s)$ is well-defined for any $s \in [0, \xi)$, $y(s;\, t_k)$ is also finite for sufficiently small t_k's for each s in that interval. However, this is clearly a contradiction to the assumption that $\xi_{t_k} \sim \xi^* < \xi$ for small t_k's. Therefore, $\xi^* = \xi$. Since this is true for every convergent subsequence of $\{\xi_t\}$, we conclude $\lim_{t\downarrow 0} \xi_t = \xi$.

To prove the main statement, we define $\zeta(s)$ by

$$\zeta(s) := \frac{1}{t} y\left(\frac{s}{t};\, t\right).$$

Then, ζ satisfies (4.1) with $\zeta(0) = u_t/t$ and, by definition, $\tau(u_t/t) = t\,\xi_t$. By the previous argument,

$$\frac{1}{t}\tau\left(\frac{u_t}{t}\right) \to \xi.$$

If we denote $\tau(u_t/t)$ by $\tilde{\tau}_t$, then $\theta(u_t, \tilde{\tau}_t) = 1/t$. Since $\tilde{\tau}_t \to 0$, we can re-parameterize u_t as a function of $v := \tilde{\tau}_t$, say $\tilde{u}_v := u_t$, thus

$$v\theta(\tilde{u}_v, v) = \tilde{\tau}_t \theta\left(u_t, \tilde{\tau}_t\right) = \frac{\tilde{\tau}_t}{t} \to \xi.$$

Since this holds for any sequence $\tilde{u}_v$ converging to u, we can replace it by the original sequence u_v. Therefore,

$$t\theta(u_t, t) \to \xi.$$

∎

In modern financial economics and econometrics, one important subject is the study of financial instruments called derivatives. To explain complex phenomena observed in the markets, there are a large number of stochastic models developed by researchers and affine diffusion processes have been applied successfully in this regard. Especially in the option pricing theory, it is typical to model the log value of the underlying asset price P_t of a contingent claim as

$$\log P_s = a_s + 2b_s \cdot Y_s$$

where a_s and b_s are deterministic functions of time and Y is an affine diffusion process (see, e.g., Duffie et al. 2003). One of the most popular derivatives is a call option and its price is the value of the right to buy a stock (or any underlying asset) at pre-determined time T, maturity, and at fixed price K, strike. The call price is given by

$$C(K, T) = B_0 \mathbb{E}(P_T - K)^+$$

where B_0 is the price of a bond maturing at T (see Lee 2004). It is standard to analyze this price in terms of Black-Scholes implied volatility $\sigma(x, T)$ which is defined implicitly by

$$C(e^x \mathbb{E}P_T, T) = (B_0 \mathbb{E}P_T)\left\{\Phi(d_+) - e^x\,\Phi(d_-)\right\}, \quad d_\pm = \frac{-x}{\sigma(x,T)\sqrt{T}} \pm \frac{\sigma(x,T)\sqrt{T}}{2}$$

where $\Phi(y) := \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$. The complexity of $\sigma(x,T)$ makes its explicit analysis hard. Rather, there is a stream of literature dealing with asymptotic behavior of $\sigma(x,T)$ as $x \to \pm\infty$ or $T \to 0$ or ∞ under some specific models of P_s. For example, see Benaim and Friz (2006) or Lee (2004). In particular, Lee (2004) proved a nice asymptotic relationship between $\sigma(x,T)$ and the critical exponents p^*, q^* in any modeling setting:

$$\limsup_{x\to\infty} \frac{\sigma^2(x,T)}{|x|/T} = \psi(p^*), \quad \limsup_{x\to-\infty} \frac{\sigma^2(x,T)}{|x|/T} = \psi(q^*) \tag{4.11}$$

where

$$p^* = \sup\{p : \mathbb{E}P_T^{1+p} < \infty\}, \quad q^* = \sup\{q : \mathbb{E}P_T^{-q} < \infty\}, \quad \psi(x) = 2 - 4(\sqrt{x^2+x} - x).$$

In our setting, $p^* + 1$ and q^* are merely the critical multipliers $\theta(b_T, T)$ and $\theta(-b_T, T)$, respectively. Therefore, a solution to the PDEs in Proposition 4.5.2 has a direct implication on $\sigma(x,T)$ for large K's.

Now suppose that $\lim_{t\downarrow 0} b_t = b$ and b satisfies the assumptions in Theorem 4.5.1:

$$b^v > 0 \quad \text{or} \quad B^c(b_{m+1}^2, \ldots, b_n^2) \neq 0.$$

Since Ω is open, we can choose an open ball U centered at $(0,b)$ so that $(t, b_t) \in U$ for all small t. This implies $\theta(b_t, t) > 1$ and so p^* is well-defined and positive. Let $\xi(b)$ be the value corresponding to ξ in Theorem 4.5.1. Then, we get

$$\limsup_{x\to\infty} \frac{\sigma^2(x,T)}{|x|} = \frac{\psi(\theta(b_T,T)-1)}{T} = \frac{\theta(b_T,T)\psi(\theta(b_T,T)-1)}{T\theta(b_T,T)} \sim \frac{1}{2\xi(b)} \text{ as } T \downarrow 0$$

where the first equality comes from (4.11) and the approximation is from Theorem 4.5.1 and the fact that $\lim_{x\to\infty} \psi(x)x = 1/2$. A similar conclusion can be drawn for q^* as long as $-b$ satisfies the assumptions in Theorem 4.5.1. Empirically, the tail slopes of implied volatility $\sigma(x,T)$ are bigger for shorter maturity options (e.g., see Duffie et al. 2000). The above observation means that this tail slope, however, cannot be arbitrarily large even for extremely small maturities under the canonical affine diffusion models.

4.6 Conclusion

We have shown that the stability boundaries of Riccati differential equations which arise in financial econometrics can be expressed as unions of stable sub-manifolds of equilibria on the stability boundaries under the assumption that every bounded trajectory converges to an equilibrium. Since we have only one stable equilibrium while all other equilibria, which are finitely many, are contained in some compact set in $\mathbb{R}^m_+ \times \{0\}$, a general picture of stability regions is obtained.

The blow-up regions of our system are defined via the blow-up times $\tau(u)$ and the boundaries of blow-up regions are level sets of $\tau(u)$. The function $\tau(u)$ turns out to be continuous and quasi-concave, and it solves a PDE similar to the Zubov equation. The critical multipliers $\theta(u,t)$ such that $\tau(\theta u) = t$ satisfy another PDE, and both functions possess an asymptotic property that has an implication on implied volatilities for options with extreme strikes and small maturities in the option pricing theory.

Chapter 5

Saddlepoint Approximations for Affine Jump-Diffusion Models

Affine jump-diffusion (AJD) processes constitute a large and widely used class of continuous-time asset pricing models that balance tractability and flexibility in matching market data. The prices of e.g., bonds, options, and other assets in AJD models are given by extended pricing transforms that have an exponential-affine form; these transforms have been characterized in great generality by Duffie et al. (2000). Calculating model prices requires inversion of these transforms, and this has limited the application of AJD models to the comparatively small subclass for which the transforms are available in closed form. This article seeks to widen the scope of AJD models amenable to practical application through approximate transform inversion techniques. More specifically, we develop the use of saddlepoint approximations for AJD models. These approximations facilitate the calculation of prices in AJD models whose transforms are not available explicitly. We derive and test several alternative saddlepoint approximations and find that they produce accurate prices over a wide range of parameters.

5.1 Introduction

Affine jump-diffusion (AJD) processes constitute a large class of continuous-time asset pricing models that balance tractability and flexibility in matching market data. In an AJD model, the drift vector, the diffusion matrix and the jump intensity all have affine dependence on the state vector. As shown by Duffie et al. (2000), this restriction leads to considerable tractability in term structure modeling and option pricing, while at the same time allowing model features like state-dependent conditional variances and flexible correlations between state variables that are absent from simpler models. The objective of this article is to further expand the scope of tractable AJD models through the use of approximate transform inversion techniques.

The AJD family of models includes many widely used special cases, such as the Gaussian model of Vasicek (1977), the square-root diffusion of Cox et al. (1985), the Heston (1993) stochastic volatility model, and extensions of these models to include jumps. AJD processes have been used extensively in empirical work, including, for example, Bakshi et al. (1997), Bates (1996, 2000), Broadie et al. (2007), Chernov (2003), Duffie et al. (1997), Duffie and Singleton (1997), Eraker (2004), Eraker et al. (2003) and Pan (2002). The yield factor models of Dai and Singleton (2000) and Duffie and Kan (1996) fall within the AJD family. Duffie et al. (2003) develop the theoretical foundations of AJD processes. A detailed account of the econometric aspects of AJD models is given in Singleton (2006).

As demonstrated in Duffie et al. (2000) (henceforth DPS), the tractability of AJD models lies in the special form taken by a wide class of transforms, including various Fourier and Laplace transforms as special cases. These transforms have an exponential-affine form, meaning that they are exponentials of affine functions of the state vector; the coefficients of these affine functions are in some cases available explicitly and, more generally, can be characterized through solutions of ordinary differential equations. Through their transform analysis, DPS derive what could be viewed as a far-reaching generalization of the Black-Scholes formula for option prices. This makes the AJD family of models particularly attractive for empirical studies that combine option prices with time series data on underlying prices or rates. Studies of this type include Andersen et al. (2002), Bakshi et al.

(1997), Bates (1996, 2000, 2003), Broadie et al. (2007), Chen and Scott (2002), Chernov (2003), Chernov and Ghysels (2000), Eraker (2004), Eraker et al. (2003) and Pan (2002).

Despite the many examples of studies using AJD models, the models used in empirical work have remained limited to a relatively small subclass for which the pricing transforms are available in closed form. This restriction appears to be driven more by convenience of implementation than by considerations of empirical validity. In the general framework of DPS, the pricing transforms are characterized in terms of solutions of ordinary differential equations (ODEs). The AJD models used in practice (such as those of Cox et al. 1985 and Heston 1993) are those for which these ODEs can be solved explicitly, thus providing explicit expressions for the pricing transforms. In this setting, each model-price calculation requires the numerical inversion of a closed-form transform, which can be accomplished with relatively modest computational effort.

For more general AJD models — those for which the pricing transforms are not available in closed form — each price calculation requires, in principle, embedding the numerical solution of a system of ODEs within a numerical inversion routine. Numerical transform inversion is a numerical integration problem that typically uses hundreds or thousands of evaluations of the transform, and each such function evaluation requires the solution of a system of ODEs. It is the impracticality of this combination that has limited the application of AJD models to the most tractable cases.

In this article, we develop the use of saddlepoint approximations as alternatives to numerical transform inversion in order to widen the scope of practical AJD models. The saddlepoint method is rooted in asymptotic expansions for evaluating contour integrals in the complex plane. It was introduced in statistics by Daniels (1954) to approximate the probability density function of the sum of independent random variables. Lugannani and Rice (1980) derive a saddlepoint approximation for the distribution function. See Daniels (1987) and Jensen (1995) for overviews of applications in statistics. Rogers and Zane (1999) apply saddlepoint approximations to option pricing; applications in credit risk include Dembo et al. (2004), Gordy (2002), Martin et al. (2001), and Yang et al. (2006). Aït-Sahalia and Yu (2006) derive saddlepoint approximations for transition densities of continuous-time Markov processes with applications to statistical inference. In the affine framework,

Collin-Dufresne and Goldstein (2002) use Edgeworth expansions for swaption pricing. Saddlepoint approximations also have potential applicability to risk management in the setting of Duffie and Pan (2001).

Saddlepoint approximations rely on the solution to an equation defined by the derivative of the transform to be inverted; this solution is the saddlepoint. We investigate various ways of computing or approximating the saddlepoint in the setting of AJD models. We also compare alternative versions of saddlepoint approximations for price calculations. We find that saddlepoint approximations do indeed provide an effective way to calculate prices in AJD models whose ODEs do not admit explicit solutions.

This chapter consists of six sections. After this introductory section, in Section 2 we present the extended transforms of AJD models that are necessary in calculating the derivatives used in the approximations. In Section 3, we review the saddlepoint method and associated approximations, and we explain how the saddlepoint method applies to AJD models. In Section 4, we propose an alternative saddlepoint method that relies on associated partial differential equations (PDEs) derived using convex duality. We test the approximations numerically in Section 5, and find that saddlepoint techniques yield surprisingly small relative errors over a wide range of parameters. We conclude the paper in Section 6.

5.2 Affine Jump-diffusion Model and Extended Transforms

We start by reviewing basic facts about AJD processes. Following the notation in DPS, an AJD process $X \in \mathbb{R}^n$ is defined as a solution of the stochastic differential equation (SDE)

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + dZ_t$$

where W is an $(\mathcal{F}_t)$-adapted Brownian motion in $\mathbb{R}^n$, $\mathcal{F}_t$ stands for the σ-field of information sets available up to time t, and Z is a pure jump process whose jumps have a fixed probability distribution ν on $\mathbb{R}^n$ and arrive with intensity $\lambda(X_t)$. The asset price of interest, S_t, at time t is assumed to be $(\bar{a}_t + \bar{b}_t \cdot X_t)\exp(a_t + b_t \cdot X_t)$ for deterministic $\bar{a}_t$, $\bar{b}_t$, a_t and b_t; for simplicity

we assume $S_t = e^{d \cdot X_t}$. The more general case can be reduced to this case at the expense of introducing time-dependency in the characteristics of X defined below. The dynamics of other assets, stochastic interest rates or stochastic volatility can be included as coordinates of the vector-valued process X. The functional forms of $\mu(X_t)$, $\sigma(X_t)$, $\lambda(X_t)$ and the interest rate $r(X_t)$ are specified as follows:

$$\mu(x) = K_0 + K_1 x, \quad K_0 \in \mathbb{R}^n,\ K_1 \in \mathbb{R}^{n \times n},\ x \in \mathbb{R}^n$$

$$(\sigma(x)\sigma(x)^\top)_{ij} = H_{0ij} + H_{1ij} \cdot x, \quad H_{0ij} \in \mathbb{R},\ H_{1ij} \in \mathbb{R}^n$$

$$\lambda(x) = l_0 + l_1 \cdot x, \quad l = (l_0, l_1) \in \mathbb{R} \times \mathbb{R}^n$$

$$r(x) = \rho_0 + \rho_1 \cdot x, \quad \rho = (\rho_0, \rho_1) \in \mathbb{R} \times \mathbb{R}^n$$

$$\theta(c) = \int_{\mathbb{R}^n} \exp(c \cdot z) dv(z) \quad \text{for } c \in \mathbb{C}^n, \text{ "jump transform".}$$

The process X is said to have the characteristic (K, H, l, θ, ρ).

The state variable X_t at time t takes values in a domain $D \subset \mathbb{R}^n$ on which the process is defined. For instance, $(\sigma(X_t)\sigma(X_t)^\top)_{ii}$ should be non-negative for each i. A discussion of the state space D and constraints on the charateristic of X can be found in Chapter 5 of Singleton (2006), and Duffie et al. (2003) deal with this issue in a more general framework. The definition above implies that the process X is Markovian and that when a jump occurs, its jump size is independent of the jump arrival rate or the past history of X.

In DPS, the authors prove that certain Fourier-type transforms of an AJD process can be found by solving the following set of ODEs:

$$\dot{\beta}(t) = -\rho_1 + K_1^\top \beta(t) + \frac{1}{2}\beta(t)^\top H_1 \beta(t) + l_1(\theta(\beta(t)) - 1) \tag{5.1}$$

$$\dot{\alpha}(t) = -\rho_0 + K_0 \cdot \beta(t) + \frac{1}{2}\beta(t)^\top H_0 \beta(t) + l_0(\theta(\beta(t)) - 1) \tag{5.2}$$

$$\dot{B}(t) = K_1^\top B(t) + \beta(t)^\top H_1 B(t) + l_1 \nabla\theta(\beta(t)) B(t) \tag{5.3}$$

$$\dot{A}(t) = K_0 \cdot B(t) + \beta(t)^\top H_0 B(t) + l_0 \nabla\theta(\beta(t)) B(t) \tag{5.4}$$

with $\beta(0) = u$, $\alpha(0) = 0$, $B(0) = v$, $A(0) = 0$ for some $u \in \mathbb{C}^n$, $v \in \mathbb{R}^n$, with $\nabla\theta(c)$ a row vector. These transforms facilitate the pricing of many financial derivatives such as

European calls or puts, quanto options, Asian options and others using Fourier inversion. To apply saddepoint techniques, we will need ODEs that characterize cumulant generating functions (CGFs) and their derivatives. See DPS for the proof of the next theorem.

Theorem 5.2.1 (DPS) *Suppose the system of ODEs (5.1)–(5.4) has a unique solution and the other technical conditions in Duffie et al. (2000), p.1351, hold. Then*

$$\begin{aligned}
\psi_0(u, X_t, t, T) &= \mathbb{E}\left[\exp\left(-\int_t^T r(X_s)ds\right)e^{u\cdot X_T}\middle|\mathcal{F}_t\right] \\
&= e^{\alpha(T-t)+\beta(T-t)\cdot X_t} \\
\psi_1(v, u, X_t, t, T) &= \mathbb{E}\left[\exp\left(-\int_t^T r(X_s)ds\right)(v\cdot X_T)e^{u\cdot X_T}\middle|\mathcal{F}_t\right] \\
&= \psi_0(u, X_t, t, T)\Big(A(T-t) + B(T-t)\cdot X_t\Big)
\end{aligned}$$

where $u \in \mathbb{C}^n$, $v \in \mathbb{R}^n$, $t \le T$ and the process X has the characteristic (K, H, l, θ, ρ).

The integral that we shall consider in later sections is $\mathbb{E}[\exp(-\int_t^T r(X_s)ds)(b\cdot X_T)^k e^{(a+zb)\cdot X_T} \mid \mathcal{F}_t]$ for some $a, b \in \mathbb{R}^n$ and $z \in \mathbb{R}$. When $k = 0$ and $t = 0$, it becomes $\psi_0(a + zb, X_0, 0, T) = \exp(\alpha(T, z) + \beta(T, z)\cdot X_0)$. Note that here we include z to express the dependence of α, β on z through the initial conditions $\alpha(0, z) = 0$, $\beta(0, z) = a + zb$. If $k = 1$, $t = 0$, then by Theorem 5.2.1 we get $\psi_1(b, a + zb, X_0, 0, T) = (A(T, z) + B(T, z)\cdot X_0)\exp(\alpha(T, z) + \beta(T, z)\cdot X_0)$ with initial conditions $A(0, z) = 0$, $B(0, z) = b$. Provided we can interchange differentiation and expectation in

$$\frac{\partial\psi_0(a + zb, X_t, t, T)}{\partial z} = \frac{\partial}{\partial z}\mathbb{E}\left[\exp\left(-\int_t^T r(X_s)ds\right)e^{(a+zb)\cdot X_T}\middle|\mathcal{F}_t\right],$$

viewing ψ_0 as a function of two variables z and t, we get

$$\frac{\partial\alpha(T-t, z)}{\partial z} + \frac{\partial\beta(T-t, z)}{\partial z}\cdot X_t = A(T-t, z) + B(T-t, z)\cdot X_t$$

for all t and X_t, so we conclude $\partial\alpha(t, z)/\partial z = A(t, z)$, $\partial\beta(t, z)/\partial z = B(t, z)$. One condition that justifies the interchange of differentiation and integration is the finiteness of ψ_0 for some interval $z \in (-l, l)$ containing 0 as an interior point. This can be proved by the Dominated Convergence Theorem and the Mean Value Theorem; see, e.g., page 43 of Shreve (2004). By

repeating the same argument, one can calculate the k-th partial derivative of ψ_0, $\partial^k\psi_0/\partial z^k$, by interchanging the order of differentiation and integration without changing the interval in which $\partial^k\psi_0/\partial z^k$ becomes finite.

Through this line of reasoning, we arrive at Theorem 5.2.2, below, and the following new set of ODEs:

$$\begin{aligned}\dot{D}(t) &= K_1^\top D(t) + \beta(t)^\top H_1 D(t) + l_1\nabla\theta(\beta(t))D(t) \qquad (5.5)\\ &\quad + B(t)^\top H_1 B(t) + l_1 B(t)^\top \nabla^2\theta(\beta(t))B(t)\\ \dot{C}(t) &= K_0 \cdot D(t) + \beta(t)^\top H_0 D(t) + l_0\nabla\theta(\beta(t))D(t) \qquad (5.6)\\ &\quad + B(t)^\top H_0 B(t) + l_0 B(t)^\top \nabla^2\theta(\beta(t))B(t)\end{aligned}$$

with $\alpha(t), \beta(t), A(t), B(t), \nabla(\theta(c))$ as before, $C(0) = 0, D(0) = 0$, and $(\nabla^2\theta(c)_{i,j}) = (\int e^{c\cdot z} z_i z_j d\nu(z))$ the Hessian of $\theta(c)$. We also need the following technical conditions, which extend conditions in DPS. The proof of Theorem 5.2.2 is based on showing that a certain process is a martingale; these conditions are useful in verifying the martingale property.

Definition 5.2.1 *(K, H, l, θ, ρ) is well-behaved at (v, u, T) if ODEs (5.1)–(5.6) are solved uniquely* [1]*, if θ is twice differentiable at $\beta(t)$ for all $t \le T$, and if the following conditions are satisfied:*

$$\begin{aligned}&(i) \quad \mathbb{E}\Big[\int_0^T |\gamma(t)\lambda(X_t)|dt\Big] < \infty, \text{ where } \gamma(t) = \Big(\Phi_t'(\theta(\beta_t) - 1) + 2\Phi_t\nabla\theta(\beta_t)B_t\\ &\quad + \Psi_t B_t^\top \nabla^2\theta(\beta_t)B_t + \Psi_t'(\theta(\beta_t) - 1) + \Psi_t\nabla\theta(\beta_t)D_t\Big)\\ &(ii) \quad \mathbb{E}\Big[\Big(\int_0^T \eta(t)\cdot\eta(t)dt\Big)^{1/2}\Big] < \infty, \text{ where } \eta(t) = \Big(\Phi_t'\beta_t^\top + 2\Phi_t B_t^\top + \Psi_t'\beta_t^\top + \Psi_t D_t^\top\Big)\sigma(X_t)\\ &(iii) \quad \mathbb{E}\Big[\big|\Phi_T' + \Psi_T'\big|\Big] < \infty\end{aligned}$$

Here Φ_t, Φ_t', Ψ_t, Ψ_t' are processes defined in the appendix and $\beta_t = \beta(T - t)$, $B_t = B(T - t)$, $D_t = D(T - t)$ for notational convenience. The next theorem is a natural extension of Theorem 5.2.1 and will play a key role in later sections.

[1] Conditions that ensure this are presented in Duffie et al. (2003) in a more general framework.

Theorem 5.2.2 *Suppose (K, H, l, θ, ρ) is well-behaved at (v, u, T). Then*

$$\begin{aligned}\psi_2(v,u,X_t,t,T) &= \mathbb{E}\left[\exp\left(-\int_t^T r(X_s)ds\right)(v\cdot X_T)^2 e^{u\cdot X_T} \mid \mathcal{F}_t\right] \\ &= \psi_0(u,X_t,t,T) \\ &\quad \times\Big((A(T-t)+B(T-t)\cdot X_t)^2 + (C(T-t)+D(T-t)\cdot X_t)\Big)\end{aligned}$$

where $v \in \mathbb{R}^n$, $u \in \mathbb{C}^n$, $t \le T$, the process X has the characteristic (K, H, l, θ, ρ).

Proof See the appendix. ∎

Again assuming that we can interchange the order of differentiation and expectation (for example, supposing $|\psi_0| < \infty$ for all $z \in (-l, l)$ for some l and treating ψ_0 as a function of z and t), we have

$$\begin{aligned}\frac{\partial^2\psi_0(a+zb,X_t,t,T)}{\partial z^2} &= \mathbb{E}\left[\exp\left(-\int_t^T r(X_s)ds\right)(b\cdot X_T)^2\, e^{(a+zb)\cdot X_T}\middle|\mathcal{F}_t\right] \\ &= \psi_2(b,a+zb,X_t,t,T);\end{aligned}$$

and from this we conclude

$$\frac{\partial^2\alpha(t,z)}{\partial z^2} = C(t,z), \quad \frac{\partial^2\beta(t,z)}{\partial z^2} = D(t,z).$$

These transforms can be continued as long as we are working with a sufficiently well behaved AJD process. Indeed, it is easy to find a pattern in the related ODEs. From the relationships above between α, β, A, B, C and D and the corresponding ODEs (5.1)–(5.6), we observe that if we have a set of ODEs for the k-th derivative of ψ_0, then we get a new set of ODEs for the $(k+1)$-th derivative just by differentiating the previous ODEs with respect to the variable z.[2] For a rigorous proof we would need to define suitable processes as in

[2]This leads us to conjecture the functional form of $\mathbb{E}[\exp(-\int_t^T r(X_s)ds)(b\cdot X_T)^N e^{(a+zb)\cdot X_T} \mid \mathcal{F}_t]$ should be

$$\sum_{(m_1,\ldots,m_N):\sum km_k=N} \frac{N!}{m_1!m_2!\cdots m_N!}\psi_0(a+zb,X_t,t,T)\prod_{j:m_j\neq 0}\left(\frac{\partial^j\alpha}{j!\partial z^j}(T-t,z)+\frac{\partial^j\beta}{j!\partial z^j}(T-t,z)\cdot X_t\right)^{m_j}$$

from the Faà di Bruno's formula and the ODEs satisfied by $\partial^j\alpha/\partial z^j$, $\partial^j\beta/\partial z^j$ can be derived by applying the same formula to the ODEs (5.1), (5.2).

Theorem 5.2.2 and give some extended conditions to make the Brownian part and the jump part martingales. We write the next set of ODEs for later use.

Theorem 5.2.3 *Under the conditions in the appendix we have*

$$\begin{aligned}\psi_3(v,u,X_t,t,T) &= \mathbb{E}\left[\exp\left(-\int_t^T r(X_s)ds\right)(v\cdot X_T)^3 e^{u\cdot X_T}\middle|\mathcal{F}_t\right]\\ &= \psi_0(u,X_t,t,T)\\ &\quad\times\Big((A(T-t)+B(T-t)\cdot X_t)^3+3(A(T-t)+B(T-t)\cdot X_t)\\ &\quad\times(C(T-t)+D(T-t)\cdot X_t)+(E(T-t)+F(T-t)\cdot X_t)\Big)\end{aligned}$$

where $v\in\mathbb{R}^n$, $u\in\mathbb{C}^n$, $t\le T$, the process X has the characteristic (K,H,l,θ,ρ) and

$$\begin{aligned}\dot F(t) &= K_1^\top F(t)+\beta(t)^\top H_1F(t)+l_1\nabla\theta(\beta(t))F(t)\\ &\quad+3B(t)^\top H_1D(t)+3l_1B(t)^\top\nabla^2\theta(\beta(t))D(t)+l_1\int_{\mathbb{R}^n}e^{z\cdot\beta(t)}(z\cdot B(t))^3d\nu(z) \qquad (5.7)\\ \dot E(t) &= K_0\cdot F(t)+\beta(t)^\top H_0F(t)+l_0\nabla\theta(\beta(t))F(t)\\ &\quad+3B(t)^\top H_0D(t)+3l_0B(t)^\top\nabla^2\theta(\beta(t))D(t)+l_0\int_{\mathbb{R}^n}e^{z\cdot\beta(t)}(z\cdot B(t))^3d\nu(z) \qquad (5.8)\end{aligned}$$

with $\alpha(t)$, $\beta(t)$, $A(t)$, $B(t)$, $C(t)$, $D(t)$, $\nabla(\theta(c))$, $\nabla^2(\theta(c))$ as before, and $E(0)=0$, $F(0)=0$.

Proof See the appendix. ∎

5.3 Saddlepoint Approximation and Option Pricing

5.3.1 Option Pricing

When we price options with the log of underlying asset following an AJD process, $S_t=e^{d\cdot X_t}$, the basic building block is

$$G_{a,b}(y;X_0,T)=\mathbb{E}\left[\exp\left(-\int_0^T r(X_s)ds\right)e^{a\cdot X_T}\mathbf{1}_{\{b\cdot X_T\le y\}}\right]$$

so that, as shown in DPS, a European call option price, for example, can be calculated as follows:

$$\begin{aligned} C(T,c) &= \mathbb{E}\left[\exp\left(-\int_0^T r(X_s)ds\right)(e^{d\cdot X_T}-c)^+\right] \\ &= \mathbb{E}\left[\exp\left(-\int_0^T r(X_s)ds\right)(e^{d\cdot X_T}-c)\mathbf{1}_{\{d\cdot X_T\geq \ln c\}}\right] \\ &= G_{d,-d}(-\ln c\,; X_0, T) - c\, G_{0,-d}(-\ln c\,; X_0, T). \end{aligned}$$

To facilitate the application of saddlepoint approximations, we will express this as a difference of two probabilities, after some possible scaling and change of measure. This will reduce the calculation of the option price to the task of calculating those probabilities. To this end, first suppose the characteristic (K, H, l, θ, ρ) of the AJD process X is well-behaved at (b, a, T). Then there exist $\tilde{\alpha}(t)$, $\tilde{\beta}(t)$ solving the ODEs (5.1), (5.2) in Theorem 5.2.1 with the boundary conditions $\tilde{\alpha}(0) = 0$, $\tilde{\beta}(0) = a$. On the other hand, it is easy to show, as noted in DPS, that

$$\xi_t = \exp\left(-\int_0^t r(X_s)ds\right)e^{\tilde{\alpha}(T-t)+\tilde{\beta}(T-t)\cdot X_t}$$

is a positive martingale, using Itô's formula and (5.1), (5.2). So an equivalent probability measure $\mathbb{Q}$ given by $d\mathbb{Q}/d\mathbb{P} = \xi_T/\xi_0$ is well defined. Also note that from the definition of ψ_0 in Section 2, $\psi_0(a, X_0, 0, T) = \mathbb{E}[\exp(-\int_0^T r(X_s)ds)e^{a\cdot X_T}] = \xi_0$. Thus the random variable $Y := b \cdot X_t$ has a moment generating function under $\mathbb{Q}$ given by

$$\begin{aligned} e^{\mathcal{K}(z)} &= \mathbb{E}^{\mathbb{Q}}\left[e^{zY}\right] = \frac{1}{\xi_0}\mathbb{E}\left[\exp\left(-\int_0^T r(X_s)ds\right)e^{(a+zb)\cdot X_T}\right] \\ &= \frac{\psi_0(a+zb, X_0, 0, T)}{\psi_0(a, X_0, 0, T)} = \exp\left(\alpha(T,z) - \alpha(T,0) + (\beta(T,z) - \beta(T,0))\cdot X_0\right) \end{aligned}$$

where $\alpha(t,z)$, $\beta(t,z)$ denote the solutions of (5.1), (5.2) with $\alpha(0,z) = 0$, $\beta(0,z) = a + zb$ so that $\tilde{\alpha}(t) = \alpha(t,0)$, $\tilde{\beta}(t) = \beta(t,0)$.

The CGF of Y is $\mathcal{K}(z)$ under $\mathbb{Q}$. Unless Y is a constant almost surely, Y has a positive variance and so $\mathcal{K}(z)$ is strictly convex in z. Proposition 5 in DPS implies that X is again an

AJD process under $\mathbb{Q}$ with the characteristic $(K^{\mathbb{Q}}, H, l^{\mathbb{Q}}, \theta^{\mathbb{Q}})$ where

$$K_0^{\mathbb{Q}}(t) = K_0 + H_0\tilde{\beta}(T-t), \quad K_1^{\mathbb{Q}}(t) = K_1 + H_1\tilde{\beta}(T-t),$$
$$l_0^{\mathbb{Q}}(t) = l_0\theta(\tilde{\beta}(T-t)), \quad l_1^{\mathbb{Q}}(t) = l_1\theta(\tilde{\beta}(T-t)),$$
$$\theta^{\mathbb{Q}}(c,t) = \theta(c+\tilde{\beta}(T-t))/\theta(\tilde{\beta}(T-t)).$$

Finally we note that $G_{a,b}(y; X_0, T) = \mathbb{E}[\exp(-\int_0^T r(X_s)ds)e^{a\cdot X_T}1_{\{b\cdot X_T \le y\}}] = \xi_0\mathbb{Q}(Y \le y)$. So the option-pricing problem is reduced to the calculation of the cumulative distribution function (CDF) $\mathbb{Q}(Y \le y)$ or its complement $\mathbb{Q}(Y > y)$.

In the AJD setting, this tail probability can be represented through the Fourier inversion formula,

$$\mathbb{Q}(Y > y) = \frac{1}{2\pi i}\int_{\tau-i\infty}^{\tau+i\infty} e^{(\mathcal{K}(z)-zy)}\frac{dz}{z}, \quad \tau > 0.^3$$

Numerical calculation of this integral requires evaluation of the integrand at hundreds or thousands of points. Unless $\mathcal{K}(z)$ is available in closed form, we would need to solve the ODEs (5.1), (5.2) numerically at each evaluation point. This computational burden limits the scope of AJD models amenable to practical application and motivates our investigation of approximations. In the next subsection, we review the saddlepoint method and explain how we apply this method to option pricing in AJD models.

Remark For European call options, a simpler calculation is possible. To simplify the measure transform, suppose the short rate is a constant r. Then the option price is given by

$$\begin{aligned} C(T,c) &= \mathbb{E}\left[e^{-rT}(S_T - c)^+\right] \\ &= e^{-rT}\left\{\mathbb{E}e^{X_T} - \mathbb{E}\left[e^{X_T} \wedge c\right]\right\} \\ &= e^{-rT}\left\{e^{\mathcal{K}(1)} - c\,\mathbb{P}(X_T + Y > \ln c)\right\} \end{aligned}$$

where $S_T = e^{X_T}$, Y is exponentially distributed with unit mean, independent of X_T, and

[3]This can be shown using the Plancherel Theorem and the Dominated Convergence Theorem (see the appendix of Rogers and Zane (1999)).

$e^{\mathcal{K}(z)} = \mathbb{E}[e^{zX_T}]$. So

$$\mathbb{E}\left[e^{z(X_T+Y)}\right] = e^{\mathcal{K}(z)}\frac{1}{1-z} = e^{\mathcal{K}(z)-\ln(1-z)}, \quad z < 1.$$

This means we need to calculate only one tail probability. If we want to use the Fourier inversion formula, this reduces the workload by almost a half. A similar but different use of exponential density functions was made in Butler and Wood (2004) to approximate the moment generating functions of truncated random variables.

5.3.2 Saddlepoint Approximation

Daniels (1954) introduced the saddlepoint method to statistics in order to approximate the probability density function (PDF) of the mean of i.i.d. random variables X_i's. Assuming we know the CGF $\mathcal{K}(z)$ where $e^{\mathcal{K}(z)} = \mathbb{E}[e^{zX_1}]$, the PDF $f_n(\overline{x})$ of $\overline{X} = \sum_1^n X_i/n$ is given by

$$f_n(\overline{x}) = \frac{n}{2\pi i}\int_{\tau-i\infty}^{\tau+i\infty} e^{n(\mathcal{K}(z)-z\overline{x})}dz, \quad \text{for any } \tau \in \{x \in \mathbb{R} : |\mathcal{K}(x)| < \infty\}.$$

Daniels (1954) used the method of steepest descent to expand this contour integral. The saddlepoint $\hat{z}$ is defined by the saddlepoint equation $\mathcal{K}'(\hat{z}) = \overline{x}$; the modulus of the integrand is minimized along the real axis at $\hat{z}$ and maximized at $\hat{z}$ along the contour parallel to the imaginary axis passing through $\hat{z}$. So, the region outside a neighborhood of the saddlepoint contributes little to the integration, and we get Daniels' formula through a Taylor expansion of the exponent $\mathcal{K}(z) - z\overline{x}$ around $\hat{z}$. (The method of steepest descent is explained in Chapter 7 of Bleistein and Handelsman (1975).)

Lugannani and Rice (1980) approximated tail probabilities rather than densities. The following form of the Lugannani-Rice (LR) formula can be found in Daniels (1987):

$$\mathbb{P}(\overline{X} > \overline{x}) = 1 - \Phi(\sqrt{n}\hat{w}) + \phi(\sqrt{n}\hat{w})\left\{\frac{b_0}{n^{1/2}} + \frac{b_1}{n^{3/2}} + o(n^{-3/2})\right\} \tag{5.9}$$

where $b_0 = 1/\hat{u} - 1/\hat{w}$, $b_1 = (\lambda_4/8 - 5\lambda_3^2/24)/\hat{u} - \lambda_3/(2\hat{u}^2) - 1/\hat{u}^3 + 1/\hat{w}^3$ and $\hat{w} = \text{sgn}(\hat{z})\sqrt{2(\hat{z}y - \mathcal{K}(\hat{z}))}$, $\hat{u} = \hat{z}\sqrt{\mathcal{K}''(\hat{z})}$, $\lambda_3 = \mathcal{K}^{(3)}(\hat{z})/\mathcal{K}''(\hat{z})^{3/2}$, $\lambda_4 = \mathcal{K}^{(4)}(\hat{z})/\mathcal{K}''(\hat{z})^{4/2}$. When $\overline{x} = \mathbb{E}[X_1] = \mathcal{K}'(0)$, the

formula reduces to

$$\mathbb{P}(\overline{X} > \mathcal{K}'(0)) = \frac{1}{2} - \frac{\lambda_3(0)}{6\sqrt{2\pi n}} + O(n^{-3/2}). \tag{5.10}$$

Here Φ, ϕ are the CDF and the PDF of the standard normal distribution, respectively. We will use this formula with $n = 1$ and b_0 in test cases. The accuracy of the approximation (5.9) for small n depends on the proximity of the underlying distribution to the normal distribution. Wood et al. (1993) study the saddlepoint approximation with a non-normal distribution replacing Φ and ϕ for a better approximation. We will test such a variant with a stochastic volatility jump-diffusion model using a gamma distribution as the base distribution in the approximation.

To apply the LR formula (5.9), we need to find the solution $\hat{z}$ of the saddlepoint equation $\mathcal{K}'(z) = y$ for some given real number y and compute $\mathcal{K}(\hat{z})$ and its derivatives. In an AJD setting, from Section 2 we have

$$\begin{aligned}
\mathcal{K}(z) &= \alpha(T,z) - \alpha(T,0) + (\beta(T,z) - \beta(T,0)) \cdot X_0 \\
\mathcal{K}'(z) &= A(T,z) + B(T,z) \cdot X_0 \\
\mathcal{K}''(z) &= C(T,z) + D(T,z) \cdot X_0, \quad \text{etc.},
\end{aligned}$$

and these functions can be evaluated by solving a set of ODEs, the size of which depends on the order of derivatives one wants to compute. Once $\hat{z}$ is found, each system of ODEs need only be solved once. The total number of ODE solutions required depends on the approximation chosen through the number of derivatives of $\mathcal{K}(\hat{z})$ used. In contrast, numerical inversion of the characteristic function requires the solution of ODEs (5.1), (5.2) for each evaluation point in the numerical integration. Finding $\hat{z}$ is therefore critical to the method.

Under rather mild conditions, the saddlepoint equation $\mathcal{K}'(z) = y$ has a unique root. We will, in particular, impose the following two conditions on the AJD process X, option

maturity T and real vectors a, b.

Assumption 1 *There exists an $l > 0$ such that $|\psi_0(a + zb, X_0, 0, T)| < \infty$ for all $z \in (-l, l)$.*

Assumption 2 *The CGF $\mathcal{K}(z)$ of $b \cdot X_T$ is strictly convex and steep at the boundary of*

$$\mathcal{D} = \{z \in \mathbb{R} : |\mathcal{K}(z)| < \infty\} .$$

Unless $b \cdot X_T$ is constant almost surely, $\mathcal{K}(z)$ is strictly convex and the convexity of $\mathcal{K}(z)$ implies that $\mathcal{D}$ is an interval. Steepness means $\lim_{z \to v} \mathcal{K}'(z) = -\infty$ and $\lim_{z \to u} \mathcal{K}'(z) = \infty$ where $v = \inf \mathcal{D}$ and $u = \sup \mathcal{D}$ (see Barndorff-Nielsen (1978) for more details). These assumptions are conditions on the tails of the random variable $b \cdot X_T$. Assumption 1 allows us to interchange the order of differentiation and integration as discussed in Section 2. Assumption 2 ensures the existence of a unique solution of the saddlepoint equation for any given $y \in \mathbb{R}$ and is not restrictive in practice.

Remark Although we focus on AJD models, the same approximations can be applied to quadratic term structure models (see, e.g., Leippold and Wu (2002) or Cheng and Scaillet (2002)) where extended transforms are again given by systems of ODEs. We also note that such systems of equations can be derived by re-writing quadratic term structure models as AJD models as observed in Cheng and Scaillet (2002), Proposition 3.

5.3.3 Approximating the Saddlepoint

As already noted, solving the saddlepoint equation is a key step in applying the saddlepoint method. Numerical solution of the equation might require many iterations, each iteration requiring evaluation of the derivative of the CGF. This could be problematic in high-dimensional models without a closed-form CGF. The approximations to the saddlepoint $\hat{z}$ discussed in this section address this difficulty.

Several authors have addressed the problem of analytically intractable CGFs. Easton and Ronchetti (1986) approximate $\mathcal{K}(z)$ by

$$\tilde{\mathcal{K}}(z) = \mu z + \frac{1}{2}\sigma^2 z^2 + \frac{1}{6}\kappa_3 z^3 + \frac{1}{24}\kappa_4 z^4$$

using the first four cumulants, and use $\tilde{z}$ for which $\tilde{\mathcal{K}}'(\tilde{z}) = y$ instead of the true saddlepoint $\hat{z}$. This approximate saddlepoint equation for $\tilde{\mathcal{K}}$ might have multiple roots, so Wang (1992) modifies this method and uses

$$\tilde{\mathcal{K}}(z;b) = \mu z + \frac{1}{2}\sigma^2 z^2 + \left(\frac{1}{6}\kappa_3 z^3 + \frac{1}{24}\kappa_4 z^4\right) g_b(z)$$

where $g_b(z) = \exp(-\kappa_2 b^2 z^2/2)$ with a properly chosen constant $b > 0$.

Starting from a Taylor expansion of $\mathcal{K}'(z)$ around $z = 0$, Lieberman (1994) presents a series reversion of the saddlepoint equation $\mathcal{K}'(\hat{z}) = y$ as a power series in $(y-\mu)/\sigma^2$. When expanded to third order, this yields

$$\hat{z}_3 = \frac{y-\mu}{\sigma^2} - \frac{\kappa_3}{2\sigma^2}\left(\frac{y-\mu}{\sigma^2}\right)^2 + \left(\frac{\kappa_3^2}{2\sigma^4} - \frac{\kappa_4}{6\sigma^2}\right)\left(\frac{y-\mu}{\sigma^2}\right)^3 \tag{5.11}$$

as an approximation to the exact saddlepoint $\hat{z}$. Here, $(y-\mu)/\sigma^2$ is the first iteration of a Newton-Raphson algorithm starting from $z_0 = 0$. Lieberman (1994) then derives a saddlepoint approximation based on $\hat{z}_3$. With $\hat{v}_3 = \hat{z}_3\sqrt{n\mathcal{K}''(\hat{z}_3)}$, $\hat{\lambda}_3 = \mathcal{K}^{(3)}(\hat{z}_3)/\mathcal{K}''(\hat{z}_3)^{3/2}$, $\hat{\lambda}_4 = \mathcal{K}^{(4)}(\hat{z}_3)/\mathcal{K}''(\hat{z}_3)^{4/2}$ and $H(x) = \mathbf{1}_{\{x>0\}} + \frac{1}{2}\mathbf{1}_{\{x=0\}}$, Lieberman's approximation is

$$\begin{aligned}
\mathbb{P}(\overline{X} > y) \;=\; & H(-\hat{v}_3) + \exp\left(n(\mathcal{K}(\hat{z}_3) - y\hat{z}_3) + \frac{\hat{v}_3^2}{2}\right) \\
& \times\Bigg[\left(H(\hat{v}_3) - \Phi(\hat{v}_3)\right)\left(1 - \frac{\hat{\lambda}_3\hat{v}_3^3}{6\sqrt{n}} + \frac{1}{n}\left(\frac{\hat{\lambda}_4\hat{v}_3^4}{24} + \frac{\hat{\lambda}_3^2\hat{v}_3^6}{72}\right)\right) \\
& + \phi(\hat{v}_3)\left(\frac{\hat{\lambda}_3(\hat{v}_3^2-1)}{6\sqrt{n}} - \frac{1}{n}\left(\frac{\hat{\lambda}_4(\hat{v}_3^3-\hat{v}_3)}{24} + \hat{\lambda}_3^2\frac{\hat{v}_3^5 - \hat{v}_3^3 + 3\hat{v}_3}{72}\right)\right)\Bigg]\left(1 + O(n^{-3/2})\right).
\end{aligned} \tag{5.12}$$

We will test this idea of an approximate saddlepoint. We will see that Lieberman's method is not uniformly accurate over a large range of strikes because the error in Lieberman's approximate saddlepoint, $\hat{z}_3$ which is an expansion in terms of $(y-\mu)/\sigma^2$, becomes large as y increases.

We propose an improvement that proceeds one more step. We expand $\mathcal{K}'(z)$ around

$z = \hat{z}_3$ (rather than $z = 0$) to third order to get

$$\tilde{z}_3 = \tilde{z} + \frac{y - \mathcal{K}'(\hat{z}_3)}{\mathcal{K}''(\hat{z}_3)} - \frac{\mathcal{K}'''(\hat{z}_3)}{2\mathcal{K}''(\hat{z}_3)}\left(\frac{y - \mathcal{K}'(\hat{z}_3)}{\mathcal{K}''(\hat{z}_3)}\right)^2 + \left(\frac{\mathcal{K}^{(3)}(\hat{z}_3)^2}{2\mathcal{K}''(\hat{z}_3)^2} - \frac{\mathcal{K}^{(4)}(\hat{z}_3)}{6\mathcal{K}''(\hat{z}_3)}\right)\left(\frac{y - \mathcal{K}'(\hat{z}_3)}{\mathcal{K}''(\hat{z}_3)}\right)^3. \quad (5.13)$$

Note that (5.13) reduces to (5.11) if $\hat{z}_3$ is replaced by zero. Evaluation of $\tilde{z}_3$ uses the same set of ODEs which are used to get $\hat{z}_3$; we do not need higher order derivatives of $\mathcal{K}(z)$ or any extra set of ODEs for (5.13). To evaluate (5.13), we solve one set of ODEs associated with $\mathcal{K}(z)$ through $\mathcal{K}^{(4)}(z)$ twice to get $\hat{z}_3$, and then solve the same set of ODEs to get $\tilde{z}_3$.

In our numerical tests, we will test the effectiveness of using the approximate saddlepoints $\hat{z}_3$ and $\tilde{z}_3$ in the LR formula (5.9) in place of the exact value $\hat{z}$. The approximations $\hat{z}_3$ and $\tilde{z}_3$ can also be used to initialize the root-finding procedure to solve for $\hat{z}$, and we will test this idea with $\hat{z}_3$.

5.4 A Dual PDE and Approximate Saddlepoint Method

In this section, we show that the problem of solving the saddlepoint equation can be transformed from a root-finding problem into a matter of a function evaluation through a duality relation.

Recall $\mathcal{K}(z) = \alpha(T,z) - \alpha(T,0) + (\beta(T,z) - \beta(T,0)) \cdot X_0$ with $\alpha(0,z) = 0, \beta(0,z) = a + zb$. Let us express the ODEs (5.1)–(5.4) as

$$\frac{\partial}{\partial t}\beta(t,z) = L_\beta(\beta), \quad \frac{\partial}{\partial t}\alpha(t,z) = L_\alpha(\beta), \quad \frac{\partial}{\partial t}B(t,z) = L_B(B,\beta), \quad \frac{\partial}{\partial t}A(t,z) = L_A(B,\beta)$$

where $L_\cdot$ is the operator corresponding to each function; for example, $L_\beta(x) = -\rho_1 + K_1^\top x + x^\top H_1 x/2 + l_1(\theta(x) - 1)$. Now define

$$\mathcal{H}(t,x,z) = \alpha(t,z) + \beta(t,z) \cdot x$$

so that $\mathbb{E}[e^{-\int_t^T r_s ds} e^{(a+zb)\cdot X_T} \mid \mathcal{F}_t, X_t = x] = e^{\alpha(T-t,z)+\beta(T-t,z)\cdot x} = e^{\mathcal{H}(T-t,x,z)}$ implies

$$\mathcal{H}(T, X_0, z) = \mathcal{K}(z) + \alpha(T,0) + \beta(T,0) \cdot X_0.$$

The function $\mathcal{H}(t,x,z)$ is convex in z, and strictly convex as long as $b \cdot X_t$ is not constant almost surely. This allows us to apply a technique developed by Jonsson and Sircar (2002) in their analysis of a partial hedging strategy. We define the convex dual

$$\mathcal{H}^*(t,x,y) := \sup_z \{yz - \mathcal{H}(t,x,z)\}.$$

The supremum should be understood to be taken over the set $\mathcal{D}(t) = \{z \in \mathbb{R} : |\mathcal{H}(t,x,z)| < \infty\}$. (Indeed, $\alpha(t,z)$ and $\beta(t,z)$ can take infinite values, as illustrated by Andersen and Piterbarg 2007.) Note that $\mathcal{D}(t)$ is defined analogously to $\mathcal{D}$ in Assumption 2. We similarly define $v(t) = \inf \mathcal{D}(t)$ and $u(t) = \sup \mathcal{D}(t)$. The next proposition tells us that under this assumption the solution of the saddlepoint equation $\hat{z}$ is actually a partial derivative of $\mathcal{H}^*(t,x,y)$ and that $\mathcal{H}^*$, $\hat{z}$ jointly satisfy some PDEs.

Proposition 5.4.1 *Suppose Assumption 1 holds, and suppose that Assumption 2 holds for all $t \in (0,T]$. Then $\mathcal{H}^*(t,x,y)$ for $(t,x,y) \in (0,T] \times \mathbb{R}^n \times \mathbb{R}$ can be expressed as $\mathcal{H}^*(t,x,y) = y\hat{z}(t,x,y) - \mathcal{H}(t,x,\hat{z}(t,x,y))$ where $\hat{z}(t,x,y)$ is the unique solution of $(\partial\mathcal{H}/\partial z)(t,x,z) = y$ for each (t,x,y). In addition, $\hat{z}(t,x,y)$ is a continuously differentiable function with $(\partial\mathcal{H}^*/\partial y)(t,x,y) = \hat{z}(t,x,y)$ and*

$$\frac{\partial \mathcal{H}^*}{\partial t} = -L_\alpha(-\nabla_x \mathcal{H}^*) - L_\beta(-\nabla_x \mathcal{H}^*) \cdot x, \tag{5.14}$$

$$\frac{\partial \hat{z}}{\partial t} = -L_A(-\nabla_x \hat{z}, -\nabla_x \mathcal{H}^*) - L_B(-\nabla_x \hat{z}, -\nabla_x \mathcal{H}^*) \cdot x. \tag{5.15}$$

Proof Consider a function $\Gamma(t,x,y,z) := (\partial\mathcal{H}/\partial z)(t,x,z) - y$. By the steepness of $\mathcal{K}(z)$ and the relation $\mathcal{H}(T,X_0,z) = \mathcal{K}(z) + \alpha(T,0) + \beta(T,0) \cdot X_0$, $\Gamma(t_0,x_0,y_0,z) = 0$ has a unique solution $\hat{z}$ for each $(t_0,x_0,y_0) \in (0,T] \times \mathbb{R}^n \times \mathbb{R}$. If $(\partial\Gamma/\partial z)(t_0,x_0,y_0,\hat{z}) \neq 0$, then the Implicit Function Theorem implies that we can find a small neighborhood B of (t_0,x_0,y_0) and a unique continuously differentiable function $\hat{z}(t,x,y)$ such that $\hat{z}(t_0,x_0,y_0) = \hat{z}$ and $\Gamma(t,x,y,\hat{z}(t,x,y)) = 0$ for all $(t,x,y) \in B$. By patching these neighborhoods together throughout the domain of (t,x,y), we confirm the smoothness of $\hat{z}(t,x,y)$.

To see why $\partial\Gamma/\partial z = \partial^2\mathcal{H}/\partial z^2$ does not vanish, we observe that $e^{\mathcal{H}(t,x,z)} = \mathbb{E}_x[\Lambda]$ where $\Lambda = e^{-\int_0^t r_s ds} e^{a \cdot X_t} e^{z(b \cdot X_t)}$ and the subscript x on the expectation indicates the initial condition $X_0 = x$. Similarly, $(\partial\mathcal{H}/\partial z)e^{\mathcal{H}} = \mathbb{E}_x[(b \cdot X_t)\Lambda]$ and $\left((\partial\mathcal{H}/\partial z)^2 + \partial^2\mathcal{H}/\partial z^2\right)e^{\mathcal{H}} = \mathbb{E}_x[(b \cdot X_t)^2\Lambda]$.

Suppose $(\partial^2\mathcal{H}/\partial z^2)(t,x,z) = 0$ for some (t,x,z). Then by the Cauchy-Schwarz inequality,

$$\begin{aligned}\left(\frac{\partial\mathcal{H}}{\partial z}\right)^2 e^{2\mathcal{H}} &= \mathbb{E}_x\left[(b\cdot X_t)\Lambda\right]^2 = \mathbb{E}_x\left[(b\cdot X_t)\Lambda^{1/2}\Lambda^{1/2}\right]^2 \\ &\leq \mathbb{E}_x\left[(b\cdot X_t)^2\Lambda\right]\mathbb{E}_x\left[\Lambda\right] = \left(\frac{\partial\mathcal{H}}{\partial z}\right)^2 e^{2\mathcal{H}}.\end{aligned}$$

Therefore, $(b\cdot X_t)\Lambda^{1/2}$ and $\Lambda^{1/2}$ should be linearly dependent. However, this implies $b\cdot X_t$ is constant almost surely, which contradicts the assumptions.

From the existence of a unique solution $\hat{z}$ of the saddlepoint equation, we have

$$\mathcal{H}^*(t,x,y) = y\hat{z}(t,x,y) - \mathcal{H}(t,x,\hat{z}(t,x,y)).$$

The differentiability of $\hat{z}(t,x,y)$ enables us to take its partial derivatives to derive

$$\begin{aligned}\frac{\partial\mathcal{H}^*}{\partial t}(t,x,y) &= y\frac{\partial\hat{z}}{\partial t}(t,x,y) - \frac{\partial\mathcal{H}}{\partial t}(t,x,\hat{z}) - \frac{\partial\mathcal{H}}{\partial z}(t,x,\hat{z})\frac{\partial\hat{z}}{\partial t}(t,x,y) \\ &= -\frac{\partial\mathcal{H}}{\partial t}(t,x,\hat{z}(t,x,y)).\end{aligned} \tag{5.16}$$

Taking partial derivatives with respect to x and y yields

$$\begin{aligned}\nabla_x\mathcal{H}^*(t,x,y) &= -\nabla_x\mathcal{H}(t,x,\hat{z}(t,x,y)) = -\beta(t,\hat{z}), \\ \frac{\partial\mathcal{H}^*}{\partial y} &= \hat{z}(t,x,y).\end{aligned} \tag{5.17}$$

By definition, we have $\mathcal{H}(t,x,z) = \alpha(t,z) + \beta(t,z)\cdot x$ and thus $\partial\mathcal{H}/\partial t = L_\alpha(\beta) + L_\beta(\beta)\cdot x$. Plugging this into (5.16) and using (5.17), we get (5.14).

Now to derive (5.15) we first recall that $\partial\alpha/\partial z = A$, $\partial\beta/\partial z = B$, $\partial A/\partial z = C$, $\partial B/\partial z = D$, and so $(\partial\mathcal{H}/\partial z)(t,x,\hat{z}) = A(t,\hat{z}) + B(t,\hat{z})\cdot x = y$. Taking partial derivatives with respect to y, x and t for both sides of this equation gives

$$C(t,\hat{z})\frac{\partial\hat{z}}{\partial y} + (D(t,\hat{z})\cdot x)\frac{\partial\hat{z}}{\partial y} = 1, \tag{5.18}$$

$$C(t,\hat{z})\nabla_x\hat{z} + (D(t,\hat{z})\cdot x)\nabla_x\hat{z} + B(t,\hat{z}) = 0, \tag{5.19}$$

$$L_A(B,\beta) + L_B(B,\beta)\cdot x + (C(t,\hat{z}) + (D(t,\hat{z})\cdot x))\frac{\partial\hat{z}}{\partial t} = 0. \tag{5.20}$$

The first equation (5.18) implies $\partial\hat{z}/\partial y \neq 0$ and thus $C(t,\hat{z}) + D(t,\hat{z}) \cdot x = 1/(\partial\hat{z}/\partial y)$. We combine this with (5.19) and (5.20) to get

$$B(t,\hat{z}) = -\frac{\nabla_x \hat{z}}{\partial\hat{z}/\partial y},$$
$$\frac{\partial\hat{z}}{\partial t} = \frac{\partial\hat{z}}{\partial y}(-L_A(B,\beta) - L_B(B,\beta)\cdot x).$$

Since L_A and L_B are linear in their first arguments, the last equation is equivalent to (5.15). ∎

The PDEs (5.14), (5.15) help identify $\hat{z}$ (or $\mathcal{K}(z)$) when they are easy to calculate numerically. However, the boundary behavior of $\mathcal{H}^*$ as $t \to 0$ can be tricky because there is a possible discontinuity at $t = 0$. For example, suppose that X is a one-dimensional Lévy process (within the AJD class) with continuously differentiable characteristic exponent $\psi(z)$ such that $\psi(z) = \infty$ outside of (a,b) for some real values $a < 0 < b$. This particular example was suggested by one referee to Glasserman and Kim (2008). We further assume that the risk free interest rate is zero. Then,

$$\mathbb{E}[e^{zX_t}|X_0 = x] = \exp\left(t\psi(z) + xz\right), \quad \mathcal{H}(t,x,z) = t\psi(z) + xz$$

and thus $\mathcal{D}(t) = (a,b)$, but $\mathcal{D}(0) = (-\infty,\infty)$. However, this leads to

$$\mathcal{H}^*(t,x,y) = (y-x)\hat{z} - t\psi(\hat{z}), \quad \hat{z} = \psi'^{-1}\left(\frac{y-x}{t}\right)$$

and $\mathcal{H}^*(0,x,y) = \infty \cdot \mathbf{1}_{y\neq x}$. By sending t to zero, we can see that if $y > x$, then $\hat{z} \to b$. Since $\psi(z^*)$ becomes positive for $\hat{z}$ sufficiently close to b, we get

$$\limsup_{t\downarrow 0} \mathcal{H}^*(t,x,y) \leq (y-x)b.$$

Similarly, if $y < x$, then $\hat{z} \to a$ and

$$\limsup_{t\downarrow 0} \mathcal{H}^*(t,x,y) \leq (y-x)a.$$

This discontinuity complicates numerical solution of the PDEs (5.14), (5.15). We can get around this by advancing time a little bit, say until $\epsilon > 0$. In other words, since $\mathcal{H}^*(\epsilon, x, y)$ is a finite function, we initiate a numerical scheme from this function to get $\hat{z}(T, x, y)$. In the case of a diffusion process, we approximate $\mathcal{H}^*(\epsilon, x, y)$ by using an Euler approximation for the diffusion from time 0 to time ϵ. This leads to a quadratic approximation for $\mathcal{H}$ and then a quadratic approximation for $\mathcal{H}^*$. We illustrate this in the next example, which deals with the Heston model.

Example Suppose the dynamics of the underlying process $(X, v) \in \mathbb{R}^2$ are given as follows:

$$dX_t = (r + uv_t)dt + \sqrt{v_t}dW_t^1, \tag{5.21}$$

$$dv_t = (a - bv_t)dt + \sigma\sqrt{v_t}dW_t^2, \tag{5.22}$$

where the correlation of the two Brownian motions W^1, W^2 is ρ. The characteristics of this affine diffusion model are

$$K_0 = \begin{pmatrix} r \\ a \end{pmatrix}, \; K_1 = \begin{pmatrix} 0 & u \\ 0 & -b \end{pmatrix}, \; H_0 = 0,$$

$$H_{1,11} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; H_{1,12} = H_{1,21} = \begin{pmatrix} 0 \\ \sigma\rho \end{pmatrix}, \; H_{1,22} = \begin{pmatrix} 0 \\ \sigma^2 \end{pmatrix}.$$

As shown in the next section, the SDEs (5.21) and (5.22) include the Heston model as a special case. With an asset price $S_t = e^{X_t}$ and $e^{\mathcal{H}(t,x,v,z)} = \mathbb{E}[e^{z(1,0)\cdot(X_t,v_t)}] = \exp(\alpha(t,z) + \beta_1(t,z)x + \beta_2(t,z)v)$, Proposition 5.4.1 implies

$$\begin{aligned} \frac{\partial \mathcal{H}^*}{\partial t}(t, X, v, y) &= K_0 \cdot \nabla_{(X,v)}\mathcal{H}^* + (K_1^\top \nabla_{(X,v)}\mathcal{H}^*) \cdot (X, v) \\ &\quad -\frac{1}{2}\left(\left(\nabla_{(X,v)}\mathcal{H}^*\right)^\top H_1 \nabla_{(X,v)}\mathcal{H}^*\right) \cdot (X, v) \\ &= r\frac{\partial \mathcal{H}^*}{\partial X} + a\frac{\partial \mathcal{H}^*}{\partial v} + \left(u\frac{\partial \mathcal{H}^*}{\partial X} - b\frac{\partial \mathcal{H}^*}{\partial v}\right)v \\ &\quad -\frac{1}{2}\left(\left(\frac{\partial \mathcal{H}^*}{\partial X}\right)^2 + 2\sigma\rho\frac{\partial \mathcal{H}^*}{\partial X}\frac{\partial \mathcal{H}^*}{\partial v} + \sigma^2\left(\frac{\partial \mathcal{H}^*}{\partial v}\right)^2\right)v. \end{aligned} \tag{5.23}$$

As discussed above, we use an ϵ approximation to implement a numerical method for

this PDE. First, we freeze the drift and volatility coefficient during a small time interval $[0, \epsilon]$ (this is the Euler approximation), to get $X_\epsilon \approx X_0 + (r + uv_0)\epsilon + \sqrt{v_0}W_\epsilon^1$. This then leads us to

$$\begin{aligned}\exp\left(\mathcal{H}(\epsilon, x, v, z)\right) &= \mathbb{E}\left[e^{zX_\epsilon} \mid (X_0, v_0) = (x, v)\right] \\ &\approx \mathbb{E}\left[\exp\left(zx + z(r + uv)\epsilon + z\sqrt{v}W_\epsilon^1\right)\right] \\ &= \exp\left(zx + z(r + uv)\epsilon + \frac{1}{2}z^2 v\epsilon\right).\end{aligned}$$

Finally, we get $\mathcal{H}^*(\epsilon, x, v, y) \approx \hat{z}y - \hat{z}\,(x + (r + uv)\epsilon) - \frac{1}{2}\hat{z}^2 v\epsilon$ where $\hat{z} = (y - x - (r + uv)\epsilon)/(\epsilon v)$. Since we have $\beta_1(t, z) = z$ for this model, we either compute $(\partial\mathcal{H}^*/\partial y)(T, X_0, v_0, y)$ or $(\partial\mathcal{H}^*/\partial X)(T, X_0, v_0, y)$ where T is the option maturity. We then use this as an approximate saddlepoint.

Figure 5.1 shows the graphs of this quadratic approximation for $\mathcal{H}$ and $\mathcal{H}^*$ for the following parameter values: $r = 3\%$, $u = -1/2$, $a = 0.08$, $b = 2$, $\sigma = 0.2$, $X_0 = \log(100)$, $v_0 = 0.04$ and $\epsilon = 0.01$. The strike K varies from 60 to 140 and $S_0 = 100$. Note that $\mathcal{H}(0, X_0, v_0, z) = X_0 \cdot z$ and $\mathcal{H}^*(0, X_0, v_0, y) = \infty \cdot 1_{\{y \neq X_0\}}$. The log-moneyness is defined by $\log(K/S_0) = y - \log(S_0)$. The approximation $\mathcal{H}_0$ of $\mathcal{H}$ captures the true curve well in the middle, but fails to do so where $\mathcal{H}$ explodes. Likewise, the approximation $\mathcal{H}_0^*$ of $\mathcal{H}^*$ performs well when the strike K is close to S_0, but produces larger errors as it moves away from S_0.

5.5 Test Cases

In this section we test the performance of saddlepoint approximation technique, for the Heston model, a stochastic volatility jump-diffusion (SVJ) model and the Scott model. Particularly, we look at the following methods:

LR method	equation (5.9) with numerical calculation of the saddlepoint $\hat{z}$
Lieberman method	equation (5.12)
L-LR method	equation (5.9) with $\hat{z}$ approximated using $\hat{z}_3$ in (5.11)
App-LR method	equation (5.9) with $\hat{z}$ approximated using $\tilde{z}_3$ in (5.13)
PDE method	equation (5.9) with $\hat{z}$ approximated using PDEs (5.14), (5.15)

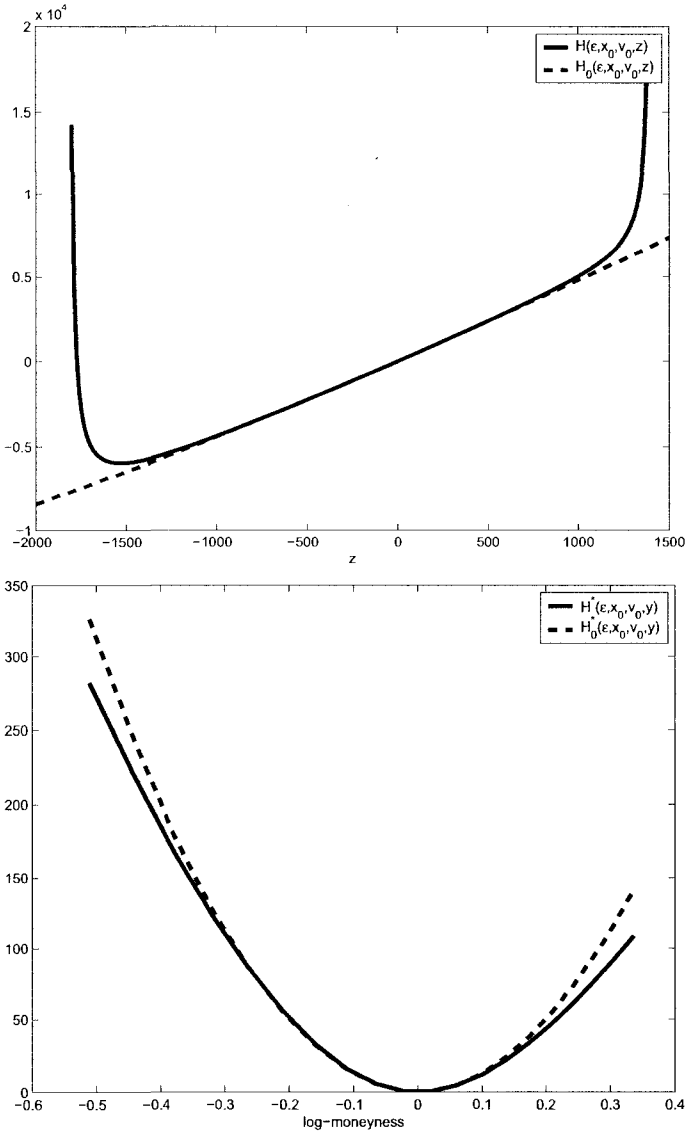


Figure 5.1: Graphs of $\mathcal{H}(\epsilon, X_0, v_0, z)$, $\mathcal{H}^*(\epsilon, X_0, v_0, y)$ and their quadratic approximations.

In applying equation (5.9), we exclude b_1 and higher order terms as their inclusion does not consistently improve the results. The motivation for testing the last four methods lies in avoiding potentially time-consuming calculation of $\hat{z}$. All tables can be found at the end of this chapter.

5.5.1 Heston Model

In the Heston model (Heston 1993), the pricing transforms are available in closed form, so no approximations are necessary. We use this as a test case for the approximations precisely because the tractability of the model allows us to compare the approximations with values computed through transform inversion.

The stock price and the volatility in the Heston model under a risk-neutral measure are assumed to follow

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{v_t} S_t dW_t^1, \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dW_t^2 \end{aligned}$$

where r is the constant interest rate and (dW_t^1, dW_t^2) is a 2-dimensional Brownian motion with $< dW_t^1, dW_t^2 >= \rho dt$. We define $X_t = \log S_t$ and apply Itô's formula to X_t to get an AJD process (X, v) with

$$dX_t = \left(r - \frac{1}{2}v_t\right)dt + \sqrt{v_t} dW_t^1$$

and v is as above. See the appendix for the characteristic of this process. The price of a European call option is then given by

$$C(T,c) = \mathbb{E}\left[e^{-rT}(S_T - c)^+\right] = S_0\mathbb{Q}(X_T > \ln c) - c\, e^{-rT}\mathbb{P}(X_T > \ln c),$$

where $\mathbb{Q}$ is defined by the measure transform $d\mathbb{Q}/d\mathbb{P} = e^{-rT}e^{X_T - X_0}$, which corresponds to taking S_T as numeraire asset. The dynamics of (X, v) can be written as

$$\begin{aligned} dX_t &= (r + v_t/2)dt + \sqrt{v_t} dW_t^{1,Q}, \\ dv_t &= (\kappa\theta - (\kappa - \rho\sigma)v_t)dt + \sigma\sqrt{v_t} dW_t^{2,Q}, \end{aligned}$$

where $W^{1,\mathbb{Q}}$ and $W^{2,\mathbb{Q}}$ are standard Brownian motions under $\mathbb{Q}$ with correlation parameter ρ. The CGF of X_T under $\mathbb{P}$ is defined by $e^{\mathcal{K}(z)} = \mathbb{E}[e^{zX_T}] = \exp(\alpha(T) + \beta(T) \cdot (X_0, v_0))$ where $\beta(0) = (z, 0)$, $\alpha(0) = 0$. Through Heston (1993), we have an explicit solution for the CGF of X_T given by

$$\begin{aligned}
&\mathcal{K}(z) = C + Dv_0 + zX_0 \\
&C = rzT + \frac{\kappa\theta}{\sigma^2}\left\{(\kappa - \rho\sigma z + d)T - 2\ln\left[\frac{1 - ge^{dT}}{1 - g}\right]\right\} \\
&D = \frac{\kappa - \rho\sigma z + d}{\sigma^2}\left[\frac{1 - e^{dT}}{1 - ge^{dT}}\right] \\
&g = \frac{\kappa - \rho\sigma z + d}{\kappa - \rho\sigma z - d} \\
&d = \sqrt{(\rho\sigma z - \kappa)^2 - \sigma^2(-z + z^2)}.
\end{aligned}$$

Again the idea is that when this kind of analytic solution is not available, we use the associated ODEs to find the saddlepoint and apply the saddlepoint method. How many calculations does this require? Let us suppose, for simplicity, that the computation times in solving ODEs for (α, β), (A, B) or (C, D) are approximately the same, say τ. Although the dimensions of the ODEs will grow exponentially as we differentiate repeatedly, we are interested in ODEs associated with $\mathcal{K}^{(j)}$ for $j = 4$ at most. Also some special structure of the models helps to simplify the equations. For example, $B_1(t) = 1$, $D_1(t) = F_1(t) = H_1(t) = 0$ in the Heston model. With the assumption of constant τ, we can compare the computational loads of different saddlepoint approximations. The computing time to approximate $G_{a,b}(y; X_0, T) = \xi_0\mathbb{Q}(Y \le y)$ using the LR method is about $\tau + 2k\tau + 3\tau$, where k is the number of iterations to solve the saddlepoint equation numerically. Here the first term is for $e^{\mathcal{K}(0)} = \xi_0$ and the last term is for $\mathcal{K}(\hat{z})$, $\mathcal{K}''(\hat{z})$. On the other hand, the time needed to apply the Lieberman method is then about $5\tau + 5\tau$ because we have to find $\mathcal{K}(0), \ldots, \mathcal{K}^{(4)}(0)$ and evaluate $\mathcal{K}(\tilde{z}), \ldots, \mathcal{K}^{(4)}(\hat{z}_3)$, while the L-LR method would require approximately $5\tau + 3\tau$ because we evaluate only up to $\mathcal{K}''(\hat{z}_3)$. The time for the App-LR method is $10\tau + 3\tau$. In each case, the most time-consuming step is getting an accurate or approximate saddlepoint, and the computational load of this step determines the efficiency of the approximation. It will become clear in our examples that the cost of this step depends

60	70	80	90	100	110	120	130	140
48	45	42	36	20	35	39	43	45
13	9	7	7	5	6	8	12	15

Table 5.1: Average number of function evaluations in the numerical solution of the saddlepoint equation in the Heston model, by strike price. The first row corresponds to initializing the root-finding procedure at zero; the second row corresponds to starting at Lieberman's approximate saddlepoint.

on option moneyness and maturity.

Numerical Results

The LR method. The initial asset price S_0 is set equal to 100, the strike c varies from 60 to 140 and the option maturity T is in the range of 0.1 to 2 years. Other parameters are given by $S_0 = 100$, $v_0 = 4\%$, $\kappa = 2$, $\theta = 4\%$, $\sigma = 0.2$, $\rho = 20\%$, $r = 3\%$. We solve the saddlepoint equation numerically by using the `fzero` function in MATLAB (which uses a bisection and interpolation algorithm) and solving the ODEs (5.1)–(5.4) at each iteration. Table 5.1 shows the average number of iterations in this step for each strike. Initializing `fzero` at the approximate saddlepoint $\hat{z}_3$ in (5.11) reduces the number of iterations by 66%–84%. Table 5.5 shows the relative errors of the LR method with respect to the accurate prices shown in the upper half.[4] The relative errors are less than 0.1% over the whole range considered.

The Lieberman Method and the L-LR method. Tables 5.6 and 5.7 show the relative errors of the Lieberman method and the L-LR method, respectively. As mentioned earlier, the approximate saddlepoint $\hat{z}_3$ incurs large errors as y (log of strike) moves away from the mean μ. So the Lieberman method works best for at-the-money (ATM) options while the L-LR method yields the smallest errors for deep in-the-money (ITM) calls. Also, we find that relative errors are enormous in the upper right part of the tables, but the out-of-the-money (OTM) call prices in that section are very small, so even small absolute errors become very large relative errors.

The App-LR method. In Table 5.8, we use the App-LR method. This method solves the

[4]The analytic prices in Table 5.4 for the Heston model and the SVJ model are produced using the program SecPrcV2.7 by Mark Broadie, Ozgur Kaya and Guy Shahar. They employed a modified trapezoidal-type routine for transform inversion. We thank Mark Broadie for providing us with a copy of this program.

ODEs for $\partial\alpha^j/\partial z^j$, $\partial\beta^j/\partial z^j$, $j = 0,\ldots,4$, one more time, but it reduces the relative errors a lot compared to the Lieberman method and the L-LR method. An important advantage of this method is that, while keeping the errors small, we solve the ODEs a fixed number of times. Using a root-finding iteration like `fzero` requires solving ODEs an unpredictable number of times.

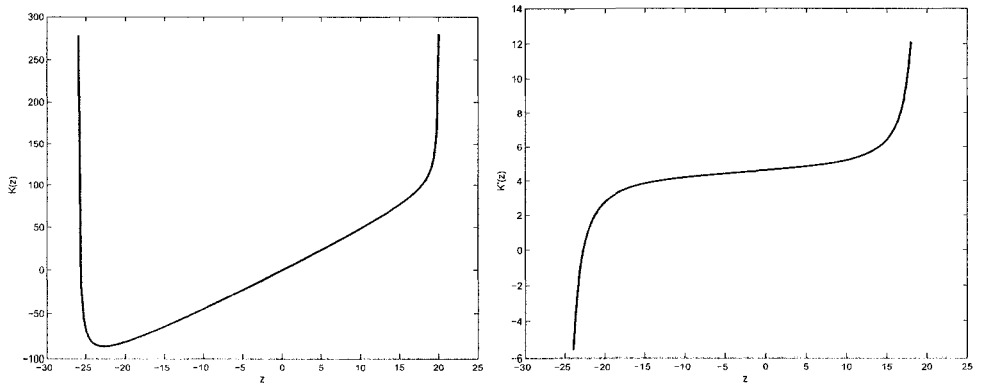
In light of the greater accuracy of the App-LR method compared with the Lieberman method and the L-LR method, in the subsequent examples we restrict attention to the LR method and the App-LR method.

Dependence of Approximation on Saddlepoint. The results above have the implication that the accuracy of saddlepoint approximations largely depends on how well we approximate the saddlepoint itself. To illustrate this more clearly, we display the shapes of the curves $\mathcal{K}(z)$ and $\mathcal{K}'(z)$ in Figure 5.2.[5] The shape of $\mathcal{K}'(z)$ looks approximately cubic. This suggests the following approach: solve ODEs (5.1)–(5.4) for some fixed values of z and for a fixed maturity, and apply a cubic spline interpolation to get an approximation for $\mathcal{K}(z)$.[6] The results are reported in Table 5.9. In most cases, the relative errors are close to the values from the LR method in Table 5.5 except in the upper right section of the table where we have small option prices. However, this approximation has an exceptionally large relative error at $T = 1.9$, $c = 110$. This again shows the importance of accurate evaluation of the saddlepoint. Any user who wants to adopt this approach should be very careful regarding this matter. One advantage of this spline approach is, first, the time for computation is relatively small (in the example, it resolves ODEs (5.1)–(5.4) 30 times for each maturity) and, second, a single approximation can be used for options with the same maturity but different strikes.

The PDE method. We also test the idea of an approximate saddlepoint in Section 5.4. Because of the non-linearity of the PDE (5.23), we use an explicit finite difference method to solve it numerically. Table 5.10 reports results for this PDE method. Relative errors are small for ITM calls ($c = 60, 70, 80$) with values less than 2%. But they become close to 8% for $c = 90$ and 100% for $c = 100$. This poor performance arises in part because we set the

[5]The graph of $\mathcal{K}(z)$ shows the moment generating function explodes around $20 + \epsilon$ and $-25 - \epsilon$.
[6]`interp1` in MATLAB

Figure 5.2: Graphs of $\mathcal{K}(z)$ and $\mathcal{K}'(z)$ with $T = 1$ in the Heston model



number of time steps n equal to 20 and ϵ to be $T - 0.02$ only. However, if one tries to make further refinements, increasing n or decreasing ϵ, then function values tend to explode easily. It was this instability that led us to restrict the range of strikes and maturities tested as shown in the table. This method is fast in finding an approximate saddlepoint, but better numerical methods are necessary to reduce huge relative errors for non-ITM options.

5.5.2 SVJ Model

As in Bates (1996), the asset price and volatility processes in the SVJ model under a risk-neutral measure $\mathbb{P}$ are as follows:

$$\frac{dS_t}{S_{t-}} = (r - \lambda k)dt + \sqrt{v_t}dW_t^1 + (\xi_{N_{t-}} - 1)dN_t,$$
$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2$$

where N is a Poisson process with rate λ and the ξ_i's are i.i.d. lognormal random variables with mean μ_J and variance σ_J^2. Since $\{e^{-rt}S_t\}$ is a martingale under the risk-neutral measure, this condition gives the relation $k = e^{\mu_J + \sigma_J^2/2} - 1$. Also, W^1 and W^2 are standard Brownian motions with correlation parameter ρ as in the Heston model. We define $X_t = \log S_t$ as usual and then Itô's formula yields

$$dX_t = (r - \lambda k - v_t/2)dt + \sqrt{v_t}dW_t^1 + \eta_{N_{t-}}dN_t,$$

60	70	80	90	100	110	120	130	140
47	43	41	35	18	35	40	43	44
27	16	9	7	6	6	6	7	8

Table 5.2: Average number of function evaluations used in the numerical solution of the saddlepoint equation for each strike in the SVJ model. The first row initiates the root-finding at zero and the second row initiates it at Lieberman's approximate saddlepoint.

where $\eta_i \sim N(\mu_J, \sigma_J^2)$. The characteristic of this AJD process (X, v) is given in the appendix. Its CGF $\mathcal{K}(z)$ under $\mathbb{P}$ is defined by $e^{\mathcal{K}(z)} = \mathbb{E}[e^{zX_T}]$.

A European call option price is, with a new probability measure $\mathbb{Q}$ defined by $d\mathbb{Q}/d\mathbb{P} = e^{X_T - \mathcal{K}(1)}$,

$$C(T,c) = e^{-rT}\left\{e^{\mathcal{K}(1)}\mathbb{Q}(X_T > \ln c) - c\,\mathbb{P}(X_T > \ln c)\right\}.$$

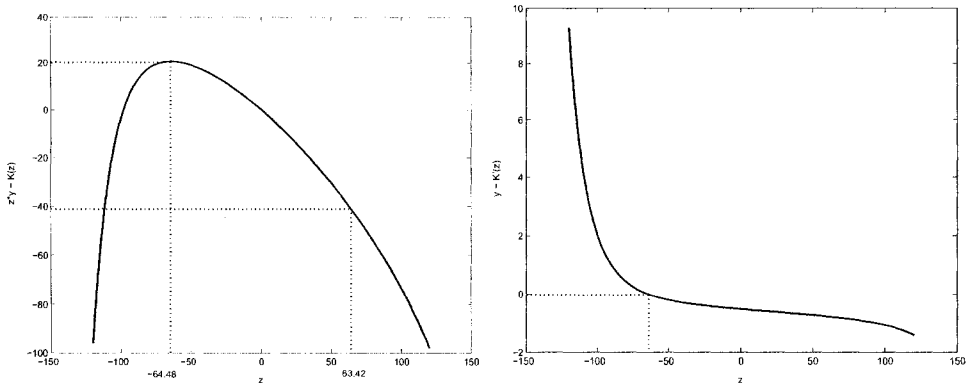
And $e^{\mathcal{K}_Q(z)} = \mathbb{E}^Q[e^{zX_T}] = e^{\mathcal{K}(1+z)-\mathcal{K}(1)}$, $\mathcal{K}_Q(z)$ denoting the CGF of X_T under $\mathbb{Q}$. From this relation between $\mathcal{K}_Q(z)$ and $\mathcal{K}(z)$, the solution $\tilde{z}$ of $\mathcal{K}_Q'(z) = y$ is given by $\hat{z} - 1$ with $\mathcal{K}'(\hat{z}) = y$.

Numerical Results

The LR method. As in Section 5.1.1, we test the LR method and compare the results with analytical option prices. Table 5.2 shows the effectiveness of using the approximate saddlepoint $\hat{z}_3$ in (5.11) as a starting point for the root-finding routine for the saddlepoint equation. The average number of function evaluations for each strike is reduced considerably, as we noted in the Heston model. We use the parameters $r = 3\%$, $\kappa = 2$, $\theta = 4\%$ (long run mean volatility = 20%), $v_0 = 4\%$ (initial volatility = 20%), $\sigma = 20\%$, $\rho = -20\%$, $\mu_J = -3\%$, $\sigma_J = 2\%$, $\lambda = 100\%$, $S_0 = 100$. Table 5.11 present the true option prices. Table 5.12 shows that the relative errors of the LR method are less than 0.4% in the whole region.

The App-LR method. With the same parameters, the App-LR method produces small relative errors close to those of Table 5.12, as reported in Table 5.13, except the one fairly extreme case of $T = 0.1$ and $c = 60$. The reason that the method fails for this case is that the approximate saddlepoint, $\hat{z}_3 = 23.9788$, from (5.11) is too far from the true saddlepoint, $\hat{z} = -64.4843$, resulting in the huge error of the modified approximate saddlepoint, $\tilde{z}_3 =$

Figure 5.3: Graphs of $yz - \mathcal{K}(z)$, $y - \mathcal{K}'(z)$ where $y = \ln c$, $T = 0.1$, $c = 60$ in the SVJ model



63.4224, from (5.13). In fact, this error makes $\hat{w}$ in the LR formula (5.9) imaginary. More precisely, $\hat{z}y - \mathcal{K}(\hat{z})$ becomes negative, as illustrated in Figure 5.3. (One could address this problem by checking if $\hat{z}y - \mathcal{K}(\hat{z})$ is positive and reverting to a root-finding iteration if it is not.) This indicates the potential limitation of the application of the App-LR method when a call option is deep ITM with a short maturity. We will see a similar pattern in the Scott model.

Sensitivity of Approximation. With the option strike 100, $T = 0.1$ and $c = 100$, the effects of λ, μ_J, and σ_J are shown in Figures 5.4, 5.5 and 5.6. As the jump arrival rate λ increases from 0 to 200%, relative errors increase linearly up to 0.085%. As the mean of the jump size μ_J decreases from 0 to −20%, relative errors make a smooth curve with a peak of 1.4% at $\mu_J = -16\%$. The volatility of the jump size has the biggest effect, making the relative error more than 10% as σ_J grows.[7] However, empirical values found in the literature stay small enough for the LR method to produce small relative errors. More specifically, as Broadie et al. (2007) summarize in their paper, Eraker et al. (2003), Andersen et al. (2002), Chernov et al. (2003) and Eraker (2004) report 4.07%, 1.95%, 0.7% and 6.63% for σ_J, respectively. Broadie et al. (2007) report σ_J between 9% and 10% when a risk premium for σ_J is assumed to exist.

[7]Figure 5.6 shows the relative errors grow as σ_J becomes larger. Numerical values are obtained from (5.9) with $n = 1$ and b_0 only or (5.10) if $\hat{z}$ is close to zero (in our case, (5.10) is used if $\hat{z} < 10^{-4}$). Indeed, when $\sigma_J = 14\%$, we have $\hat{z} = 4.38 \times 10^{-5}$ and (5.9) yields a 184.86% relative error while (5.10) gives a relative error of 6.97%.

Figure 5.4: Effect of the jump arrival rate in the SVJ Model

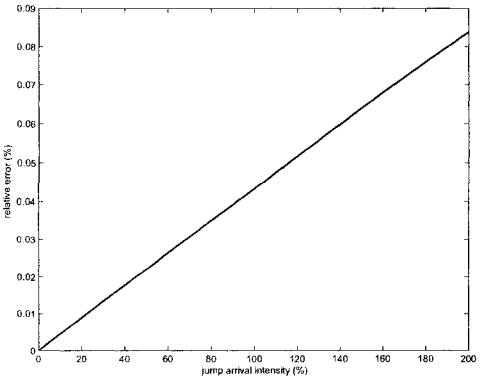


Figure 5.5: Effect of the mean of the jump size in the SVJ Model

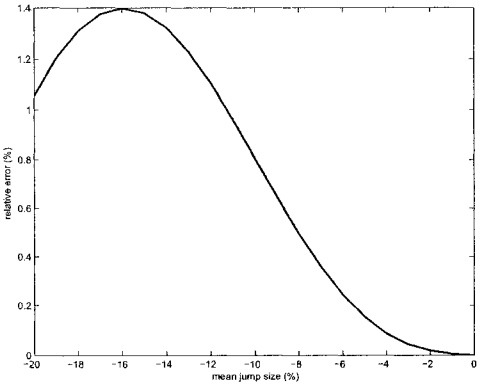
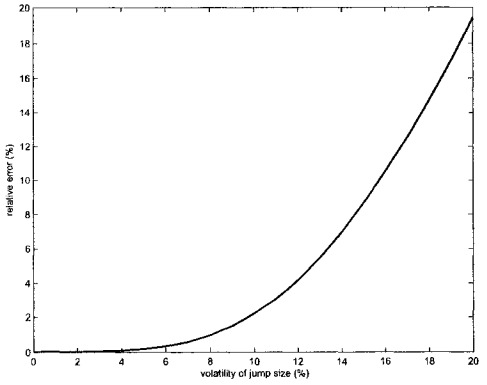


Figure 5.6: Effect of the volatility of the jump size in the SVJ Model



Nonnormal-based Approximation. The added skewness due to the jump component in the SVJ model makes the saddlepoint approximation using a gamma distribution for the base distribution attractive. We test this method for two strikes in Table 5.14.[8] The gamma-based approximation is better for $c = 90$, but not for $c = 100$. This result reasserts the conclusion of Wood et al. (1993), "... any gains are likely to be small when the normal-based approximation does well."

5.5.3 Scott Model

As the last test case, we apply the methods to the jump-diffusion model with stochastic volatility and stochastic interest rates in Scott (1997). Under a risk-neutral measure $\mathbb{P}$, the dynamics of the state variables are given by

$$\begin{aligned} dX_t &= (r_t - \lambda k - \sigma^2 y_t^1/2)dt + \sigma\sqrt{y_t^1}dW_t + \eta_{N_t-}dN_t, \\ dy_t^1 &= \kappa_1(\theta_1 - y_t^1)dt + \sigma_1\sqrt{y_t^1}dW_t^1, \\ dy_t^2 &= \kappa_2(\theta_2 - y_t^2)dt + \sigma_2\sqrt{y_t^2}dW_t^2 \end{aligned}$$

where W_t, W_t^1, W_t^2 are Brownian motions with $< dW_t, dW_t^1 >= \rho dt$, $< dW_t, dW_t^2 >= 0$, $r_t = y_t^1 + y_t^2$, $\eta_i \overset{iid}{\sim} N(\mu_J, \sigma_J^2)$ and $k = e^{\mu_J + \sigma_J^2/2} - 1$. The stock price S_t is $\exp(X_t)$.

The characteristics for this model are given in the appendix. A function $\mathcal{K}(z)$ is defined by $e^{\mathcal{K}(z)} = \mathbb{E}[e^{-\int_0^T r_s ds} e^{zX_T}]$. Note that $\mathcal{K}(z)$ is not the CGF of X_T under $\mathbb{P}$. The European call option price is

$$C(T,c) = e^{\mathcal{K}(1)}\mathbb{Q}_1(X_T > \ln c) - c\, e^{\mathcal{K}(0)}\mathbb{Q}_2(X_T > \ln c),$$

where the probability measures $\mathbb{Q}_i$, $i = 1, 2$, are defined by $d\mathbb{Q}_1/d\mathbb{P} = e^{-\int_0^T r_s ds} e^{X_T - \mathcal{K}(1)}$ and $d\mathbb{Q}_2/d\mathbb{P} = e^{-\int_0^T r_s ds - \mathcal{K}(0)}$, so that

$$e^{\mathcal{K}_{Q_1}(z)} = \mathbb{E}^{Q_1}\left[e^{zX_T}\right] = e^{\mathcal{K}(1+z)-\mathcal{K}(1)}, \quad e^{\mathcal{K}_{Q_2}(z)} = \mathbb{E}^{Q_2}\left[e^{zX_T}\right] = e^{\mathcal{K}(z)-\mathcal{K}(0)}$$

[8] The PDF of a gamma distribution Gamma(k, θ) is expressed as $f(x, k, \theta) = x^{k-1}e^{-(x/\theta)}/(\Gamma(k)\theta^k)$ for $x > 0$, the shape parameter k and the scale parameter θ. We use a chi-square distribution $\chi^2(\nu)$ of which PDF is that of Gamma$(\nu/2, 2)$ where ν is the degree of freedom. In Table 5.14, ν is set equal to 4. Other values for ν have similar results.

60	70	80	90	100	110	120	130	140
47	45	43	38	29	34	42	43	46
21	17	10	7	7	5	7	8	10

Table 5.3: Average number of function evaluations used in the numerical solution of the saddlepoint equation for each strike in the Scott model. The first row initiates the root-finding at zero and the second row initiates it at Lieberman's approximate saddlepoint.

and $\mathcal{K}_{Q_i}$ is the CGF of X_T under $\mathbb{Q}_i$. The saddlepoint equation is given by $\mathcal{K}'_{Q_1}(\tilde{z}) = \mathcal{K}'(1+\tilde{z}) = y$ for $\mathbb{Q}_1$ and $\mathcal{K}'_{Q_2}(\hat{z}) = \mathcal{K}'(\hat{z}) = y$ for $\mathbb{Q}_2$. So implementing the LR method requires solving $\mathcal{K}'(z) = y$ only.

There are two ways to use the App-LR method. One is to use this method for each of $\mathbb{Q}_i(X_T > \ln c)$, $i = 1,2$, trying to approximate the corresponding saddlepoints separately. The other is to set the approximation of $\tilde{z}$ equal to the approximation of $\hat{z}$ minus one, based on the relation $\tilde{z} = \hat{z} - 1$. Using this consistent approximation requires solving half as many ODEs. In more detail, the first method solves the ODEs for $\partial\alpha^j/\partial z^j$, $\partial\beta^j/\partial z^j$, $j = 0,\ldots,4$, four times to get two approximate saddlepoints $\hat{z}$ and $\tilde{z}$, while the latter one solves the same ODEs just twice. In our tests, the second method produces smaller errors, particularly at short maturities.

Numerical Results

The LR method. We use the same range of parameters for maturity and strike. Additional parameters are set as follows: $S_0 = 100$, $y_0^1 = \theta_1 = 3\%$, $y_0^2 = \theta_2 = 2\%$, $\kappa_1 = 5$, $\kappa_2 = 0.4$, $\sigma = 1$, $\sigma_1 = 0.23$, $\sigma_2 = 0.1$, $\rho = -26\%$, $\mu_J = -4\%$, $\sigma_J = 1\%$, $\lambda = 100\%$. The analytical values in Table 5.15 were computed using Fourier inversion, using the `quad` function in MATLAB with a large interval for the numerical integration. Different integration intervals give different values, but we find the errors to be very small. Again in Table 5.3, we find that initiating `fzero` at the approximate saddlepoint $\hat{z}_3$ in (5.11) helps to reduce the computation time for solving the saddlepoint equation, and in Table 5.16 we observe small relative errors (less than 0.1% in most cases) for the LR method with respect to the analytical valuation.

The App-LR method. Table 5.17 shows results of the App-LR method. As noted in

Section 5.2.1, we see that the method is not applicable to some deep ITM calls with short maturities. There are also two big errors in the upper right part of the table that do not have counterparts in the SVJ model. These errors, however, disappear when we use the second implementation, setting the approximation of $\tilde{z}$ equal to the approximation of $\hat{z}$ minus one. We find that this method dominates the first method throughout the whole region considered. See Table 5.18. Even though this second method still cannot be applied to some deep ITM calls with short maturities, it produces relative errors very close to those of the LR method.

5.6 Conclusion

When a closed-form solution for the characteristic function in an affine jump-diffusion model is not available, transform inversion combining numerical integration with hundreds or thousands of ODE solutions can be very time consuming. We have seen that saddlepoint approximations can be an effective alternative computational tool for calculating prices in affine jump-diffusion models.

In saddlepoint approximations, we find that accurate calculation of the saddlepoint is the most critical and often the most challenging task. We can address this issue either by solving the saddlepoint equation numerically or by obtaining an approximate saddlepoint. Results in this paper can be summarized as follows:

- The LR method (the Lugannani-Rice formula with a numerical solution of the saddlepoint equation) yields the smallest relative errors, ranging from 0.0% – 0.3% in most cases for the models considered here.

- Initiating a root-finding iteration at the approximate saddlepoint $\hat{z}_3$ of Lieberman substantially reduces the number of iterations.

- The App-LR method (the LR formula with an improved series approximation to the saddlepoint) gives small relative errors close to those of the LR method. However, it gives poor results for some deep ITM options with short maturities.

- For ATM options, the LR method dominates. For OTM or ITM options, the App-LR method is better, considering speed and accuracy together.

- If speed is of greater concern than accuracy, then it is best to use the Lieberman method for ATM options and to use the L-LR method for ITM options.

In our numerical tests, we have considered a wide range of strikes and maturities. Empirical work with AJD models generally focuses on a much more limited range, and this further supports the use of saddlepoint approximations.

Table 5.4: Heston Model: Analytic option prices

	60	**70**	**80**	**90**	**100**	**110**	**120**	**130**	**140**
0.1	40.180	30.210	20.240	10.366	2.662	0.236	0.008	1.7E-04	2.4E-06
0.2	40.359	30.419	20.490	10.940	3.840	0.826	0.121	0.014	0.001
0.3	40.538	30.628	20.769	11.546	4.773	1.455	0.353	0.075	0.015
0.4	40.716	30.840	21.077	12.141	5.581	2.068	0.656	0.190	0.052
0.5	40.894	31.057	21.403	12.714	6.308	2.658	0.997	0.349	0.119
0.6	41.072	31.279	21.743	13.266	6.979	3.224	1.360	0.544	0.212
0.7	41.250	31.507	22.089	13.799	7.608	3.770	1.734	0.764	0.330
0.8	41.429	31.739	22.440	14.313	8.202	4.297	2.115	1.004	0.469
0.9	41.608	31.976	22.793	14.812	8.769	4.807	2.499	1.260	0.626
1.0	41.789	32.217	23.145	15.296	9.313	5.304	2.884	1.527	0.799
1.1	41.970	32.460	23.497	15.767	9.837	5.787	3.269	1.803	0.985
1.2	42.153	32.706	23.848	16.225	10.343	6.258	3.653	2.087	1.182
1.3	42.336	32.954	24.196	16.673	10.834	6.719	4.035	2.377	1.388
1.4	42.521	33.204	24.542	17.111	11.312	7.170	4.414	2.671	1.604
1.5	42.706	33.455	24.885	17.539	11.777	7.612	4.792	2.969	1.827
1.6	42.893	33.707	25.225	17.959	12.231	8.046	5.167	3.270	2.056
1.7	43.079	33.959	25.562	18.370	12.675	8.472	5.539	3.574	2.291
1.8	43.267	34.211	25.896	18.774	13.109	8.890	5.908	3.879	2.532
1.9	43.455	34.463	26.227	19.171	13.535	9.302	6.275	4.185	2.776
2.0	43.644	34.716	26.554	19.561	13.953	9.707	6.638	4.492	3.025

Table 5.5: Heston Model: Relative errors of the LR method

	60	70	80	90	100	110	120	130	140
0.1	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.001%	0.002%
0.2	0.000%	0.000%	0.000%	0.000%	-0.001%	-0.001%	0.000%	0.002%	0.005%
0.3	0.000%	0.000%	0.000%	-0.001%	-0.004%	-0.005%	-0.003%	0.002%	0.008%
0.4	0.000%	0.000%	0.000%	-0.003%	-0.008%	-0.011%	-0.008%	-0.001%	0.008%
0.5	0.000%	0.000%	-0.001%	-0.005%	-0.013%	-0.018%	-0.015%	-0.006%	0.004%
0.6	0.000%	0.000%	-0.001%	-0.007%	-0.017%	-0.025%	-0.023%	-0.014%	-0.002%
0.7	0.000%	0.000%	-0.002%	-0.009%	-0.022%	-0.031%	-0.032%	-0.024%	-0.012%
0.8	0.000%	0.000%	-0.003%	-0.011%	-0.025%	-0.037%	-0.040%	-0.035%	-0.024%
0.9	0.000%	-0.001%	-0.004%	-0.013%	-0.028%	-0.042%	-0.048%	-0.045%	-0.036%
1.0	0.000%	-0.001%	-0.004%	-0.014%	-0.030%	-0.045%	-0.054%	-0.055%	-0.048%
1.1	0.000%	-0.001%	-0.005%	-0.016%	-0.031%	-0.048%	-0.059%	-0.063%	-0.060%
1.2	0.000%	-0.001%	-0.006%	-0.016%	-0.032%	-0.049%	-0.063%	-0.069%	-0.069%
1.3	0.000%	-0.002%	-0.006%	-0.017%	-0.033%	-0.050%	-0.065%	-0.074%	-0.077%
1.4	0.000%	-0.002%	-0.007%	-0.017%	-0.033%	-0.050%	-0.066%	-0.077%	-0.083%
1.5	0.000%	-0.002%	-0.007%	-0.018%	-0.033%	-0.050%	-0.066%	-0.079%	-0.088%
1.6	0.000%	-0.002%	-0.007%	-0.018%	-0.032%	-0.049%	-0.066%	-0.080%	-0.090%
1.7	0.000%	-0.002%	-0.008%	-0.017%	-0.031%	-0.048%	-0.065%	-0.080%	-0.092%
1.8	-0.001%	-0.003%	-0.008%	-0.017%	-0.031%	-0.047%	-0.063%	-0.079%	-0.092%
1.9	-0.001%	-0.003%	-0.008%	-0.017%	-0.030%	-0.045%	-0.062%	-0.077%	-0.091%
2.0	-0.001%	-0.003%	-0.008%	-0.017%	-0.029%	-0.044%	-0.059%	-0.075%	-0.090%

Table 5.6: Heston Model: Relative errors of the Lieberman method

	60	70	80	90	100	110	120	130	140
0.1	0.00%	0.00%	0.12%	0.23%	0.00%	-19.46%	-602.61%	-2.1E+3%	-1.9E+4%
0.2	0.00%	0.23%	1.13%	0.27%	0.00%	-4.29%	-116.01%	-428.92%	-1.3E+3%
0.3	0.05%	1.65%	1.79%	0.28%	0.00%	-2.04%	-45.91%	-200.64%	-466.73%
0.4	0.55%	3.58%	1.90%	0.28%	0.00%	-1.26%	-24.07%	-117.31%	-265.96%
0.5	2.03%	4.93%	1.78%	0.28%	0.00%	-0.88%	-14.60%	-75.62%	-181.11%
0.6	4.32%	5.46%	1.58%	0.27%	0.00%	-0.66%	-9.67%	-51.56%	-132.77%
0.7	6.66%	5.38%	1.38%	0.26%	0.00%	-0.51%	-6.78%	-36.52%	-100.62%
0.8	8.37%	4.98%	1.20%	0.24%	0.00%	-0.40%	-4.96%	-26.63%	-77.58%
0.9	9.21%	4.44%	1.04%	0.23%	0.00%	-0.33%	-3.73%	-19.87%	-60.47%
1.0	9.29%	3.89%	0.91%	0.22%	0.00%	-0.27%	-2.88%	-15.11%	-47.55%
1.1	8.84%	3.37%	0.80%	0.20%	0.00%	-0.23%	-2.27%	-11.69%	-37.68%
1.2	8.10%	2.91%	0.70%	0.19%	0.00%	-0.20%	-1.81%	-9.17%	-30.08%
1.3	7.26%	2.51%	0.62%	0.18%	0.00%	-0.17%	-1.47%	-7.29%	-24.19%
1.4	6.40%	2.17%	0.56%	0.17%	0.00%	-0.15%	-1.21%	-5.86%	-19.59%
1.5	5.61%	1.89%	0.50%	0.16%	0.00%	-0.13%	-1.00%	-4.76%	-15.97%
1.6	4.90%	1.65%	0.45%	0.15%	0.00%	-0.12%	-0.84%	-3.91%	-13.11%
1.7	4.27%	1.45%	0.41%	0.14%	0.01%	-0.10%	-0.71%	-3.23%	-10.83%
1.8	3.74%	1.28%	0.37%	0.13%	0.01%	-0.09%	-0.60%	-2.70%	-9.01%
1.9	3.28%	1.13%	0.34%	0.12%	0.01%	-0.08%	-0.52%	-2.27%	-7.53%
2.0	2.88%	1.01%	0.31%	0.12%	0.01%	-0.08%	-0.44%	-1.92%	-6.34%

Table 5.7: Heston Model: Relative errors of the L-LR method

	60	70	80	90	100	110	120	130	140
0.1	0.00%	0.00%	0.00%	0.04%	0.02%	-2.61%	-25.92%	-337.50%	-1.2E+4%
0.2	0.00%	0.00%	0.02%	0.19%	0.05%	-2.66%	-17.37%	-121.26%	-1.2E+3%
0.3	0.00%	0.01%	0.09%	0.34%	0.09%	-2.52%	-13.55%	-67.76%	-422.93%
0.4	0.00%	0.02%	0.18%	0.45%	0.13%	-2.29%	-11.03%	-44.65%	-207.63%
0.5	0.01%	0.05%	0.27%	0.54%	0.16%	-2.05%	-9.15%	-32.08%	-121.32%
0.6	0.01%	0.09%	0.35%	0.59%	0.19%	-1.81%	-7.67%	-24.30%	-78.78%
0.7	0.02%	0.13%	0.42%	0.62%	0.21%	-1.58%	-6.48%	-19.06%	-54.94%
0.8	0.03%	0.17%	0.46%	0.63%	0.22%	-1.38%	-5.52%	-15.33%	-40.34%
0.9	0.04%	0.20%	0.49%	0.63%	0.24%	-1.21%	-4.72%	-12.56%	-30.78%
1.0	0.05%	0.23%	0.51%	0.62%	0.24%	-1.05%	-4.07%	-10.44%	-24.19%
1.1	0.07%	0.25%	0.51%	0.60%	0.25%	-0.91%	-3.52%	-8.78%	-19.46%
1.2	0.08%	0.27%	0.51%	0.59%	0.25%	-0.79%	-3.06%	-7.46%	-15.95%
1.3	0.09%	0.28%	0.50%	0.56%	0.25%	-0.69%	-2.67%	-6.40%	-13.27%
1.4	0.10%	0.28%	0.49%	0.54%	0.25%	-0.60%	-2.34%	-5.52%	-11.18%
1.5	0.10%	0.29%	0.48%	0.52%	0.25%	-0.52%	-2.06%	-4.80%	-9.53%
1.6	0.11%	0.29%	0.47%	0.50%	0.25%	-0.45%	-1.81%	-4.20%	-8.19%
1.7	0.12%	0.29%	0.45%	0.48%	0.24%	-0.39%	-1.61%	-3.69%	-7.10%
1.8	0.12%	0.28%	0.44%	0.46%	0.24%	-0.33%	-1.43%	-3.26%	-6.20%
1.9	0.12%	0.28%	0.42%	0.44%	0.23%	-0.29%	-1.27%	-2.89%	-5.44%
2.0	0.13%	0.28%	0.41%	0.42%	0.23%	-0.25%	-1.13%	-2.58%	-4.81%

Table 5.8: Heston Model: Relative errors of the App-LR method

	60	70	80	90	100	110	120	130	140
0.1	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	-0.042%	-4.966%
0.2	0.000%	0.000%	0.000%	0.000%	-0.001%	-0.001%	0.000%	-0.015%	-1.091%
0.3	0.000%	0.000%	0.000%	-0.001%	-0.004%	-0.005%	-0.003%	-0.005%	-0.280%
0.4	0.000%	0.000%	0.000%	-0.003%	-0.008%	-0.011%	-0.008%	-0.004%	-0.087%
0.5	0.000%	0.000%	-0.001%	-0.005%	-0.013%	-0.018%	-0.015%	-0.008%	-0.033%
0.6	0.000%	0.000%	-0.001%	-0.007%	-0.017%	-0.025%	-0.023%	-0.015%	-0.018%
0.7	0.000%	0.000%	-0.002%	-0.009%	-0.022%	-0.031%	-0.032%	-0.025%	-0.019%
0.8	0.000%	0.000%	-0.003%	-0.011%	-0.025%	-0.037%	-0.040%	-0.035%	-0.027%
0.9	0.000%	-0.001%	-0.004%	-0.013%	-0.028%	-0.042%	-0.048%	-0.046%	-0.038%
1.0	0.000%	-0.001%	-0.004%	-0.014%	-0.030%	-0.045%	-0.054%	-0.055%	-0.049%
1.1	0.000%	-0.001%	-0.005%	-0.016%	-0.031%	-0.048%	-0.059%	-0.063%	-0.060%
1.2	0.000%	-0.001%	-0.006%	-0.016%	-0.032%	-0.049%	-0.063%	-0.069%	-0.069%
1.3	0.000%	-0.002%	-0.006%	-0.017%	-0.033%	-0.050%	-0.065%	-0.074%	-0.077%
1.4	0.000%	-0.002%	-0.007%	-0.017%	-0.033%	-0.050%	-0.066%	-0.077%	-0.083%
1.5	0.000%	-0.002%	-0.007%	-0.018%	-0.033%	-0.050%	-0.066%	-0.079%	-0.088%
1.6	0.000%	-0.002%	-0.007%	-0.018%	-0.032%	-0.049%	-0.066%	-0.080%	-0.090%
1.7	0.000%	-0.002%	-0.008%	-0.017%	-0.031%	-0.048%	-0.065%	-0.080%	-0.092%
1.8	-0.001%	-0.003%	-0.008%	-0.017%	-0.031%	-0.047%	-0.063%	-0.079%	-0.092%
1.9	-0.001%	-0.003%	-0.008%	-0.017%	-0.030%	-0.045%	-0.062%	-0.077%	-0.091%
2.0	-0.001%	-0.003%	-0.008%	-0.017%	-0.029%	-0.044%	-0.059%	-0.075%	-0.090%

Table 5.9: Heston Model: Cubic spline interpolation to approximate $\mathcal{K}'(z)$. $z \in [-15, 15]$ and step size 1

	60	70	80	90	100	110	120	130	140
0.1	0.000%	0.000%	0.000%	0.002%	-0.007%	-0.276%	-8.979%	-68.560%	-528.702%
0.2	0.000%	0.000%	-0.002%	0.000%	-0.023%	-0.001%	-0.520%	-7.643%	-33.650%
0.3	0.000%	-0.001%	-0.003%	-0.001%	-0.043%	-0.005%	-0.003%	-0.174%	-2.398%
0.4	0.000%	-0.002%	0.000%	-0.003%	-0.063%	-0.011%	-0.008%	-0.001%	0.013%
0.5	-0.001%	-0.001%	-0.001%	-0.005%	-0.081%	-0.017%	-0.015%	-0.006%	0.003%
0.6	0.000%	0.000%	-0.001%	-0.007%	-0.097%	-0.025%	-0.023%	-0.014%	-0.003%
0.7	0.000%	0.000%	-0.002%	-0.009%	-0.110%	-0.027%	-0.032%	-0.025%	-0.012%
0.8	0.000%	0.000%	-0.003%	-0.011%	-0.120%	-0.036%	-0.039%	-0.035%	-0.024%
0.9	0.000%	-0.001%	-0.004%	-0.013%	-0.127%	-0.037%	-0.047%	-0.044%	-0.036%
1.0	0.000%	-0.001%	-0.004%	-0.014%	-0.132%	-0.025%	-0.051%	-0.055%	-0.048%
1.1	0.000%	-0.001%	-0.005%	-0.015%	-0.136%	-0.019%	-0.059%	-0.061%	-0.058%
1.2	0.000%	-0.001%	-0.006%	-0.016%	-0.137%	-0.028%	-0.060%	-0.069%	-0.069%
1.3	0.000%	-0.002%	-0.006%	-0.017%	-0.138%	-0.047%	-0.057%	-0.072%	-0.076%
1.4	0.000%	-0.002%	-0.007%	-0.017%	-0.137%	-0.042%	-0.059%	-0.073%	-0.081%
1.5	0.000%	-0.002%	-0.007%	-0.018%	-0.135%	0.030%	-0.064%	-0.077%	-0.087%
1.6	0.000%	-0.002%	-0.007%	-0.018%	-0.133%	0.249%	-0.066%	-0.080%	-0.089%
1.7	0.000%	-0.002%	-0.008%	-0.017%	-0.131%	0.973%	-0.058%	-0.077%	-0.088%
1.8	-0.001%	-0.003%	-0.008%	-0.017%	-0.128%	5.383%	-0.046%	-0.071%	-0.087%
1.9	-0.001%	-0.003%	-0.008%	-0.017%	-0.125%	2034.383%	-0.034%	-0.068%	-0.089%
2.0	-0.001%	-0.003%	-0.008%	-0.016%	-0.122%	5.170%	-0.026%	-0.067%	-0.090%

Table 5.10: Relative errors of the PDE method with $\triangle X = \triangle v = \triangle t = 0.001$, $\epsilon = T - 0.02$ and 20 time steps.

	60	70	80	90	100
1	0.01%	0.09%	0.78%	5.70%	104.69%
1.1	0.01%	0.11%	0.89%	6.05%	98.80%
1.2	0.02%	0.13%	1.00%	6.36%	93.42%
1.3	0.02%	0.15%	1.12%	6.64%	88.50%
1.4	0.02%	0.18%	1.23%	6.88%	83.99%
1.5	0.03%	0.21%	1.34%	7.09%	79.85%
1.6	0.03%	0.24%	1.45%	7.28%	76.04%
1.7	0.04%	0.28%	1.56%	7.43%	72.52%
1.8	0.04%	0.31%	1.66%	7.57%	69.27%
1.9	0.05%	0.35%	1.76%	7.69%	66.27%
2	0.06%	0.39%	1.85%	7.78%	63.47%

Table 5.11: SVJ Model: Analytic option prices

	60	70	80	90	100	110	120	130	140
0.1	40.180	30.210	20.241	10.406	2.703	0.207	0.004	3.3E-05	1.5E-07
0.2	40.359	30.419	20.506	11.041	3.901	0.766	0.084	0.006	3.3E-04
0.3	40.538	30.632	20.816	11.693	4.851	1.379	0.274	0.042	0.005
0.4	40.716	30.853	21.159	12.322	5.672	1.985	0.539	0.121	0.024
0.5	40.896	31.083	21.521	12.923	6.412	2.573	0.851	0.243	0.063
0.6	41.077	31.321	21.893	13.497	7.094	3.142	1.192	0.402	0.125
0.7	41.260	31.567	22.268	14.047	7.733	3.692	1.551	0.591	0.210
0.8	41.444	31.819	22.644	14.576	8.337	4.225	1.922	0.804	0.317
0.9	41.631	32.075	23.019	15.086	8.913	4.744	2.301	1.037	0.445
1.0	41.820	32.335	23.392	15.581	9.465	5.248	2.684	1.287	0.590
1.1	42.010	32.596	23.761	16.060	9.997	5.740	3.069	1.550	0.752
1.2	42.202	32.859	24.126	16.527	10.511	6.220	3.455	1.824	0.928
1.3	42.395	33.123	24.488	16.982	11.009	6.690	3.841	2.107	1.117
1.4	42.589	33.388	24.845	17.425	11.493	7.150	4.226	2.397	1.318
1.5	42.785	33.653	25.198	17.859	11.965	7.601	4.610	2.693	1.529
1.6	42.981	33.917	25.547	18.283	12.425	8.044	4.992	2.993	1.748
1.7	43.177	34.181	25.893	18.699	12.875	8.479	5.372	3.298	1.976
1.8	43.374	34.445	26.234	19.107	13.315	8.906	5.749	3.605	2.211
1.9	43.572	34.708	26.571	19.508	13.746	9.326	6.124	3.915	2.452
2.0	43.769	34.969	26.904	19.901	14.168	9.740	6.496	4.227	2.698

Table 5.12: SVJ Model: Relative errors of the LR method

	60	70	80	90	100	110	120	130	140
0.1	0.000%	0.000%	0.000%	-0.004%	-0.043%	-0.136%	-0.241%	-0.323%	-0.376%
0.2	0.000%	0.000%	-0.001%	-0.008%	-0.043%	-0.116%	-0.209%	-0.296%	-0.362%
0.3	0.000%	0.000%	-0.002%	-0.012%	-0.046%	-0.113%	-0.197%	-0.275%	-0.338%
0.4	0.000%	0.000%	-0.003%	-0.015%	-0.050%	-0.115%	-0.194%	-0.268%	-0.325%
0.5	0.000%	0.000%	-0.004%	-0.017%	-0.054%	-0.119%	-0.196%	-0.267%	-0.322%
0.6	0.000%	-0.001%	-0.005%	-0.020%	-0.058%	-0.122%	-0.199%	-0.269%	-0.324%
0.7	0.000%	-0.001%	-0.006%	-0.022%	-0.061%	-0.124%	-0.200%	-0.271%	-0.327%
0.8	0.000%	-0.001%	-0.007%	-0.025%	-0.064%	-0.125%	-0.200%	-0.272%	-0.330%
0.9	0.000%	-0.002%	-0.008%	-0.027%	-0.066%	-0.126%	-0.198%	-0.270%	-0.331%
1.0	0.000%	-0.002%	-0.010%	-0.029%	-0.067%	-0.125%	-0.195%	-0.266%	-0.329%
1.1	0.000%	-0.003%	-0.011%	-0.031%	-0.068%	-0.123%	-0.191%	-0.260%	-0.325%
1.2	-0.001%	-0.003%	-0.012%	-0.032%	-0.069%	-0.121%	-0.185%	-0.253%	-0.318%
1.3	-0.001%	-0.004%	-0.013%	-0.033%	-0.069%	-0.119%	-0.180%	-0.245%	-0.310%
1.4	-0.001%	-0.004%	-0.014%	-0.034%	-0.069%	-0.116%	-0.174%	-0.237%	-0.300%
1.5	-0.001%	-0.005%	-0.015%	-0.035%	-0.068%	-0.113%	-0.168%	-0.228%	-0.289%
1.6	-0.001%	-0.005%	-0.016%	-0.036%	-0.067%	-0.110%	-0.162%	-0.219%	-0.278%
1.7	-0.002%	-0.006%	-0.016%	-0.036%	-0.067%	-0.107%	-0.156%	-0.211%	-0.267%
1.8	-0.002%	-0.006%	-0.017%	-0.036%	-0.066%	-0.104%	-0.151%	-0.202%	-0.256%
1.9	-0.002%	-0.007%	-0.018%	-0.037%	-0.065%	-0.101%	-0.145%	-0.194%	-0.246%
2.0	-0.002%	-0.007%	-0.018%	-0.037%	-0.063%	-0.098%	-0.140%	-0.186%	-0.236%

Table 5.13: SVJ Model: Relative errors of the App-LR method

	60	70	80	90	100	110	120	130	140
0.1	N/A	0.000%	0.000%	-0.004%	-0.043%	-0.136%	-0.259%	-1.084%	-0.456%
0.2	0.252%	0.004%	0.000%	-0.008%	-0.043%	-0.116%	-0.212%	-0.559%	-1.282%
0.3	0.012%	0.007%	-0.001%	-0.012%	-0.046%	-0.113%	-0.197%	-0.346%	-0.940%
0.4	0.040%	0.007%	-0.002%	-0.015%	-0.050%	-0.115%	-0.194%	-0.288%	-0.626%
0.5	0.103%	0.005%	-0.003%	-0.017%	-0.054%	-0.119%	-0.196%	-0.273%	-0.458%
0.6	0.133%	0.004%	-0.005%	-0.020%	-0.058%	-0.122%	-0.199%	-0.271%	-0.381%
0.7	0.106%	0.002%	-0.006%	-0.022%	-0.061%	-0.124%	-0.200%	-0.272%	-0.351%
0.8	0.067%	0.001%	-0.007%	-0.025%	-0.064%	-0.125%	-0.200%	-0.272%	-0.339%
0.9	0.039%	0.000%	-0.008%	-0.027%	-0.066%	-0.126%	-0.198%	-0.270%	-0.335%
1.0	0.022%	-0.001%	-0.009%	-0.029%	-0.067%	-0.125%	-0.195%	-0.266%	-0.331%
1.1	0.013%	-0.002%	-0.011%	-0.031%	-0.068%	-0.123%	-0.191%	-0.260%	-0.325%
1.2	0.007%	-0.003%	-0.012%	-0.032%	-0.069%	-0.121%	-0.185%	-0.253%	-0.318%
1.3	0.004%	-0.004%	-0.013%	-0.033%	-0.069%	-0.119%	-0.180%	-0.245%	-0.310%
1.4	0.002%	-0.004%	-0.014%	-0.034%	-0.069%	-0.116%	-0.174%	-0.237%	-0.300%
1.5	0.001%	-0.005%	-0.015%	-0.035%	-0.068%	-0.113%	-0.168%	-0.228%	-0.289%
1.6	0.000%	-0.005%	-0.016%	-0.036%	-0.067%	-0.110%	-0.162%	-0.219%	-0.278%
1.7	-0.001%	-0.006%	-0.016%	-0.036%	-0.067%	-0.107%	-0.156%	-0.211%	-0.267%
1.8	-0.001%	-0.006%	-0.017%	-0.036%	-0.066%	-0.104%	-0.151%	-0.202%	-0.256%
1.9	-0.002%	-0.007%	-0.018%	-0.037%	-0.065%	-0.111%	-0.145%	-0.194%	-0.246%
2.0	-0.002%	-0.007%	-0.018%	-0.037%	-0.063%	-0.098%	-0.140%	-0.186%	-0.236%

Table 5.14: SVJ Model: the LR method and gamma-based approximation with $\nu = 4$ for strikes 90, 100

	LR(90)	Gamma(90)	LR(100)	Gamma(100)
0.1	-0.004%	0.006%	-0.043%	-0.074%
0.2	-0.008%	0.006%	-0.043%	-0.065%
0.3	-0.012%	0.004%	-0.046%	-0.060%
0.4	-0.015%	0.003%	-0.050%	-0.059%
0.5	-0.017%	0.001%	-0.054%	-0.059%
0.6	-0.020%	-0.002%	-0.058%	-0.060%
0.7	-0.022%	-0.004%	-0.061%	-0.061%
0.8	-0.025%	-0.006%	-0.064%	-0.062%
0.9	-0.027%	-0.008%	-0.066%	-0.064%
1.0	-0.029%	-0.010%	-0.067%	-0.065%
1.1	-0.031%	-0.012%	-0.068%	-0.065%
1.2	-0.032%	-0.014%	-0.069%	-0.066%
1.3	-0.033%	-0.016%	-0.069%	-0.066%
1.4	-0.034%	-0.017%	-0.069%	-0.066%
1.5	-0.035%	-0.018%	-0.068%	-0.066%
1.6	-0.036%	-0.019%	-0.067%	-0.065%
1.7	-0.036%	-0.020%	-0.067%	-0.065%
1.8	-0.036%	-0.021%	-0.066%	-0.064%
1.9	-0.037%	-0.021%	-0.065%	-0.063%
2.0	-0.037%	-0.022%	-0.063%	-0.063%

Table 5.15: Scott Model: Analytic option prices

	60	70	80	90	100	110	120	130	140
0.1	40.299	30.349	20.399	10.519	2.484	0.111	0.001	2.0E-06	2.8E-07
0.2	40.597	30.697	20.806	11.186	3.654	0.536	0.034	0.001	3.1E-05
0.3	40.893	31.043	21.230	11.872	4.611	1.072	0.148	0.014	0.001
0.4	41.188	31.390	21.669	12.542	5.459	1.640	0.339	0.052	0.007
0.5	41.481	31.738	22.117	13.190	6.237	2.216	0.592	0.125	0.022
0.6	41.773	32.088	22.570	13.818	6.965	2.790	0.890	0.234	0.053
0.7	42.064	32.439	23.023	14.426	7.655	3.359	1.221	0.378	0.103
0.8	42.354	32.790	23.476	15.018	8.316	3.919	1.578	0.553	0.174
0.9	42.644	33.143	23.928	15.594	8.951	4.472	1.952	0.757	0.266
1.0	42.932	33.495	24.377	16.156	9.565	5.017	2.341	0.984	0.380
1.1	43.219	33.847	24.822	16.705	10.161	5.554	2.740	1.233	0.514
1.2	43.506	34.198	25.264	17.243	10.741	6.084	3.146	1.500	0.668
1.3	43.792	34.549	25.702	17.769	11.307	6.606	3.558	1.782	0.839
1.4	44.077	34.899	26.137	18.286	11.861	7.121	3.974	2.078	1.027
1.5	44.360	35.248	26.567	18.794	12.403	7.630	4.393	2.385	1.231
1.6	44.643	35.595	26.993	19.293	12.934	8.133	4.815	2.701	1.447
1.7	44.925	35.941	27.415	19.785	13.457	8.629	5.238	3.027	1.677
1.8	45.206	36.285	27.833	20.268	13.970	9.120	5.661	3.359	1.918
1.9	45.486	36.627	28.247	20.745	14.475	9.606	6.085	3.698	2.169
2.0	45.765	36.968	28.657	21.215	14.972	10.087	6.509	4.042	2.429

Table 5.16: Scott Model: Relative errors of the LR method

	60	70	80	90	100	110	120	130	140
0.1	0.000%	0.000%	0.000%	0.000%	0.002%	-0.025%	-4.501%	-12.167%	-99.217%
0.2	0.000%	0.000%	0.000%	0.000%	-0.002%	-0.017%	-0.045%	-0.063%	-0.459%
0.3	0.000%	0.000%	0.000%	-0.001%	-0.008%	-0.029%	-0.058%	-0.076%	-0.109%
0.4	0.000%	0.000%	0.000%	-0.002%	-0.011%	-0.034%	-0.066%	-0.091%	-0.108%
0.5	0.000%	0.000%	0.000%	-0.003%	-0.013%	-0.034%	-0.064%	-0.095%	-0.118%
0.6	0.000%	0.000%	-0.001%	-0.004%	-0.013%	-0.032%	-0.059%	-0.089%	-0.117%
0.7	0.000%	0.000%	-0.001%	-0.004%	-0.013%	-0.029%	-0.052%	-0.079%	-0.108%
0.8	0.000%	0.000%	-0.001%	-0.004%	-0.012%	-0.025%	-0.044%	-0.068%	-0.094%
0.9	0.000%	0.000%	-0.001%	-0.004%	-0.011%	-0.022%	-0.038%	-0.058%	-0.080%
1.0	0.000%	0.000%	-0.001%	-0.004%	-0.010%	-0.019%	-0.032%	-0.049%	-0.068%
1.1	0.000%	0.000%	-0.001%	-0.004%	-0.009%	-0.017%	-0.028%	-0.041%	-0.057%
1.2	0.000%	0.000%	-0.001%	-0.004%	-0.008%	-0.015%	-0.024%	-0.035%	-0.048%
1.3	0.000%	0.000%	-0.001%	-0.003%	-0.007%	-0.013%	-0.021%	-0.030%	-0.041%
1.4	0.000%	0.000%	-0.001%	-0.003%	-0.007%	-0.012%	-0.018%	-0.026%	-0.035%
1.5	0.000%	0.000%	-0.001%	-0.003%	-0.006%	-0.010%	-0.016%	-0.022%	-0.030%
1.6	0.000%	0.000%	-0.001%	-0.003%	-0.005%	-0.009%	-0.014%	-0.020%	-0.026%
1.7	0.000%	0.000%	-0.001%	-0.003%	-0.005%	-0.008%	-0.012%	-0.017%	-0.022%
1.8	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.007%	-0.011%	-0.015%	-0.020%
1.9	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.006%	-0.010%	-0.013%	-0.017%
2.0	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.006%	-0.009%	-0.012%	-0.015%

Table 5.17: Scott Model: Relative errors of the App-LR method using separate approximations for $\hat{z}$ and $\tilde{z}$

	60	70	80	90	100	110	120	130	140
0.1	-246.923%	N/A	-0.001%	0.000%	0.002%	-0.025%	-9.641%	246.595%	-2.956%
0.2	N/A	0.011%	-0.001%	0.000%	-0.002%	-0.017%	-0.204%	-9.696%	363.989%
0.3	N/A	0.033%	0.000%	-0.001%	-0.008%	-0.029%	-0.063%	-1.502%	-4.076%
0.4	0.015%	0.010%	0.000%	-0.002%	-0.011%	-0.034%	-0.066%	-0.219%	-3.503%
0.5	0.020%	0.003%	0.000%	-0.003%	-0.013%	-0.034%	-0.064%	-0.105%	-0.698%
0.6	0.018%	0.001%	0.000%	-0.004%	-0.013%	-0.032%	-0.059%	-0.090%	-0.198%
0.7	0.007%	0.001%	-0.001%	-0.004%	-0.013%	-0.029%	-0.052%	-0.079%	-0.119%
0.8	0.002%	0.000%	-0.001%	-0.004%	-0.012%	-0.025%	-0.044%	-0.068%	-0.096%
0.9	0.001%	0.000%	-0.001%	-0.004%	-0.011%	-0.022%	-0.038%	-0.058%	-0.081%
1.0	0.000%	0.000%	-0.001%	-0.004%	-0.010%	-0.019%	-0.032%	-0.049%	-0.068%
1.1	0.000%	0.000%	-0.001%	-0.004%	-0.009%	-0.017%	-0.028%	-0.041%	-0.057%
1.2	0.000%	0.000%	-0.001%	-0.004%	-0.008%	-0.015%	-0.024%	-0.035%	-0.048%
1.3	0.000%	0.000%	-0.001%	-0.003%	-0.007%	-0.013%	-0.021%	-0.030%	-0.041%
1.4	0.000%	0.000%	-0.001%	-0.003%	-0.007%	-0.012%	-0.018%	-0.026%	-0.035%
1.5	0.000%	0.000%	-0.001%	-0.003%	-0.006%	-0.010%	-0.016%	-0.022%	-0.030%
1.6	0.000%	0.000%	-0.001%	-0.003%	-0.005%	-0.009%	-0.014%	-0.020%	-0.026%
1.7	0.000%	0.000%	-0.001%	-0.003%	-0.005%	-0.008%	-0.012%	-0.017%	-0.022%
1.8	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.007%	-0.011%	-0.015%	-0.020%
1.9	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.006%	-0.010%	-0.013%	-0.017%
2.0	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.006%	-0.009%	-0.012%	-0.015%

Table 5.18: Scott Model: Relative errors of the App-LR method using $\hat{z}$ and the identity $\tilde{z} = \hat{z} - 1$.

	60	70	80	90	100	110	120	130	140
0.1	0.121%	N/A	0.000%	0.000%	0.002%	-0.025%	-4.499%	-12.211%	-99.224%
0.2	N/A	0.000%	0.000%	0.000%	-0.002%	-0.017%	-0.045%	-0.056%	-0.589%
0.3	N/A	0.000%	0.000%	-0.001%	-0.008%	-0.029%	-0.058%	-0.075%	-0.105%
0.4	0.001%	0.000%	0.000%	-0.002%	-0.011%	-0.034%	-0.066%	-0.090%	-0.103%
0.5	0.000%	0.000%	0.000%	-0.003%	-0.013%	-0.034%	-0.064%	-0.095%	-0.117%
0.6	0.000%	0.000%	-0.001%	-0.004%	-0.013%	-0.032%	-0.059%	-0.089%	-0.117%
0.7	0.000%	0.000%	-0.001%	-0.004%	-0.013%	-0.029%	-0.052%	-0.079%	-0.108%
0.8	0.000%	0.000%	-0.001%	-0.004%	-0.012%	-0.025%	-0.044%	-0.068%	-0.094%
0.9	0.000%	0.000%	-0.001%	-0.004%	-0.011%	-0.022%	-0.038%	-0.058%	-0.080%
1.0	0.000%	0.000%	-0.001%	-0.004%	-0.010%	-0.019%	-0.032%	-0.049%	-0.068%
1.1	0.000%	0.000%	-0.001%	-0.004%	-0.009%	-0.017%	-0.028%	-0.041%	-0.057%
1.2	0.000%	0.000%	-0.001%	-0.004%	-0.008%	-0.015%	-0.024%	-0.035%	-0.048%
1.3	0.000%	0.000%	-0.001%	-0.003%	-0.007%	-0.013%	-0.021%	-0.030%	-0.041%
1.4	0.000%	0.000%	-0.001%	-0.003%	-0.007%	-0.012%	-0.018%	-0.026%	-0.035%
1.5	0.000%	0.000%	-0.001%	-0.003%	-0.006%	-0.010%	-0.016%	-0.022%	-0.030%
1.6	0.000%	0.000%	-0.001%	-0.003%	-0.005%	-0.009%	-0.014%	-0.020%	-0.026%
1.7	0.000%	0.000%	-0.001%	-0.003%	-0.005%	-0.008%	-0.012%	-0.017%	-0.022%
1.8	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.007%	-0.011%	-0.015%	-0.020%
1.9	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.006%	-0.010%	-0.013%	-0.017%
2.0	0.000%	0.000%	-0.001%	-0.002%	-0.004%	-0.006%	-0.009%	-0.012%	-0.015%

Appendix A

A.1 Proofs for Chapter 3

Proof of Lemma 3.4.1 Define $x(t) = \Phi_t(u)$ and $y(t) = \Phi_t(\theta u)/\theta$; then

$$\begin{aligned} \dot{x} &= Ax + B(x_1^2, \ldots, x_n^2), \\ \dot{y} &= Ay + \theta B(y_1^2, \ldots, y_n^2) \end{aligned}$$

with $x(0) = y(0) = u$. It is immediate that $x^d \equiv y^d$ because they satisfy the same linear ODE with the same initial condition. So, we concentrate on x^v and y^v, for which the corresponding ODEs are

$$\begin{aligned} \dot{x}^v &= A^v x^v + (x_1^2, \ldots, x_m^2) + c(t) + d(t), \\ \dot{y}^v &= A^v y^v + \theta(y_1^2, \ldots, y_m^2) + c(t) + \theta d(t) \end{aligned}$$

where $c(t) = A^c x^d(t)$ and $d(t) = B^c(x_{m+1}^2, \ldots, x_n^2)$. Now define

$$f(x^v) = A^v x^v + (x_1^2, \ldots, x_m^2).$$

By condition **(C2)** (see the discussion preceding Lemma 3.4.1), the mapping $x^v \mapsto A^v x^v$ is quasi-monotone increasing, as is the mapping $x^v \mapsto (x_1^2, \ldots, x_m^2)$, and thus also f. Recalling

that B^c has nonnegative entries and $\theta > 1$, we get

$$\begin{aligned} \dot{x}^v - f(x^v) &= c(t) + d(t) \\ &\leq (\theta - 1)(y_1^2, \ldots, y_m^2) + c(t) + \theta d(t) \\ &= \dot{y}^v - f(y^v). \end{aligned}$$

It now follows from the comparison result (3.25) that $x(t) \leq y(t)$. ∎

For the proof of Lemma 3.4.2, we need a preliminary result that limits the crossing of coordinates of the solution to (3.9).

Lemma A.1.1 *For the system (3.9), suppose $(t, u) \in \Omega$ and let $x(t) = \Phi_t(u)$. For $i, j \in \{1, \ldots, n\}$, the set $\{s \in [0, t] : x_i(s) = x_j(s)\}$ has only finitely many isolated points.*

Proof As noted in Section 3.4.1, $\Phi_t(u)$ is analytic in (t, u) so long as it lies within the domain of analyticity of f_o in (3.9); but this function is analytic in the entire domain. It follows that $x_i(s) - x_j(s)$ is analytic in s. An analytic function can have only a finite number of isolated zeros on a compact interval. ∎

Proof of Lemma 3.4.2 Fix $u \in \mathbb{R}^n$ and let us denote $\Phi_t(u)$ by $x(t)$ to simplify notation. We define a piecewise differentiable function $\gamma(t) = \min_{i=1,\ldots,m} x_i(t)$. Since $x^d(t)$ converges to zero as (implied by (3.27)), we can find $M > 0$ such that $\sup_t |x^d(t)| < M$. The value of M depends on $x^d(0)$. Lemma A.1.1 implies that in any bounded interval $[0, t]$ with $x(t)$ finite, the set of s at which $x_i(s) = x_j(s)$ is either finite or an interval. Therefore, we can define a sequence of closed intervals of $\mathbb{R}_+$, $\{I_j\}$, such that $I(u) \cap \mathbb{R}_+ = \bigcup_{j=1}^{\infty} I_j$ and

$$\gamma(t) = x_{i(j)}, \quad \forall t \in I_j^o,$$

for some $i(j) \in \{1, \ldots, m\}$, where I_j^o denotes the interior of I_j.

In an interval I^o throughout which $\gamma(t) = x_i(t)$, we have

$$\begin{aligned} \dot{\gamma}(t) &= \gamma^2 + \sum_{k=1}^{n} A_{ik}x_k + \sum_{k=m+1}^{n} B_{ik}x_k^2 \\ &\geq \gamma^2 + \sum_{k=1}^{m} A_{ik}\gamma + \sum_{k=m+1}^{n} A_{ik}x_k \\ &\geq \gamma^2 + \sum_{k=1}^{m} A_{ik}\gamma - M \max_{j=1,\ldots,m} \sum_{k=m+1}^{n} |A_{jk}|. \end{aligned} \tag{A.1}$$

In the first inequality, we used the assumption that A^v has non-negative off-diagonal entries and $B^c \geq 0$. Next, we define a continuous, piecewise differentiable function v by

$$\dot{v} = L(v), \quad L(v) := v^2 + \sum_{k=1}^{m} A_{ik}v - K$$

whenever $\gamma(t) = x_i(t)$, with $K = M \max_j \sum_{k=m+1}^{n} |A_{jk}|$ and $\gamma(0) \geq v(0)$. Then, since γ and v satisfy

$$\dot{\gamma} - L(\gamma) \geq \dot{v} - L(v),$$

we get $\gamma \geq v$ by applying the standard comparison result repeatedly on each interval I_j. If we show that v is bounded below, then γ is also bounded below and the statement follows.

To see that v is indeed bounded below, we observe that M can be set large enough to make $L(x) = 0$ have two real solutions, $\eta_1^i < \eta_2^i$, for each i; in this case $\dot{v}(t) \geq 0$ or $v(t) \geq \eta_1^i$ (as is evident in Figure 3.1) when $\gamma(t) = x_i(t)$. ∎

Proof of Lemma 3.4.3 We write $x(t)$ for $\Phi_t(u)$ to simplify notation. Define a piecewise differentiable function $\gamma = \max_{i=1,\ldots,m} x_i$, similarly as in the proof of Lemma 3.4.2. We saw there that we can define a sequence of intervals $\{I_j\}$ until τ with $\gamma(t) = x_{i(j)}(t)$ in I_j^o. In an interval I on which $\gamma = x_i$, γ satisfies (A.1). Since the trajectory of $x(t)$ is bounded below (by Lemma 3.4.2), $\gamma \to \infty$. So, at some time $t_0 < \tau$, $\gamma(t_0)$ is sufficiently large that the right side of (A.1) becomes positive for all $i = 1, \ldots, m$, and then γ never decreases. We can then

divide both sides of (A.1) by γ to get

$$\begin{aligned}\frac{\dot{\gamma}}{\gamma} &= \sum_{k=1}^{m} A_{ik}\frac{x_k}{\gamma} + \gamma + \frac{1}{\gamma}\sum_{k=m+1}^{n} A_{ik}x_k + \frac{1}{\gamma}\sum_{k=m+1}^{n} B_{ik}x_k^2 \\ &\leq \sum_{k=1}^{m} A_{ik} + \gamma + M\end{aligned}$$

on (t_0, τ), for some sufficiently large M. Here we have used the fact that $x_k \leq \gamma, k \in \{1, \ldots, m\}$, and the nonnegativity of the off-diagonal entries of A^v. The existence of M is guaranteed by the fact that $|x^d|$ is bounded and γ never decreases after t_0. Then,

$$\int_{t_0}^{\tau} \frac{\dot{\gamma}}{\gamma} dt \leq \int_{t_0}^{\tau} \gamma dt + \Big(\max_i \sum_{k=1}^{m} A_{ik} + M\Big)(\tau - t_0).$$

However, the left side is infinite, so $\int_{t_0}^{\tau} \gamma dt = \infty$ as well. We can pick a constant C such that $|\Lambda^d x^d(t)| \leq C$ for all $t \geq 0$; then, since $\Lambda^v \gg 0$, we have

$$\int_0^{\tau} \Lambda \cdot x(t) dt \geq (\min_i \Lambda_i^v) \int_0^{\tau} \gamma dt - C\tau = \infty.$$

∎

The system (3.9) can be thought of a system of equations defined in $\mathbb{C}^n$ by setting $x(t) = \mathrm{Re}\, x(t) + i\mathrm{Im}\, x(t)$. Based on the analyticity of f_o, the solution $x(t)$ also has a nice analytic property which is used in the proof of Theorem 3.2.1.

Lemma A.1.2 *For the system (3.9), suppose $(t, u) \in \Omega$. Then we can find an open convex subset of $\mathbb{C}^n$, containing the line segment $L = \{\lambda u \in \mathbb{R}^n : \lambda \in [0, 1]\}$, in which $\Phi_t(\cdot)$ is analytic.*

Proof Since $\Phi_t(u)$ is finite, $\Phi_t(\lambda u)$ is finite for all $\lambda \in [0, 1]$. This is because, first, $\Phi_t(\lambda u) \leq \lambda \Phi_t(u)$ by Lemma 3.4.1 (take $\theta = 1/\lambda$, for $\lambda \in (0, 1]$) and, second, it is bounded below by Lemma 3.4.2.

For each $\lambda u \in L$, there is an open ball B_λ in $\mathbb{C}^n$ centered at λu in which $\Phi_t(\cdot)$ is analytic, because of the analyticity of f_o. Since L is compact, we can cover L by a finite number of such balls. We can then find an open convex set U that contains L and is contained within

the cover; for example, we can U to be the set of points less than a distance ϵ from L, for sufficiently small $\epsilon > 0$. Then $\Phi_t(\cdot)$ is analytic in U. ∎

Proof of Lemma 3.6.1 The proof uses an approach of Getz and Jacobson (1977). We write the ODE for x^v in (3.9) as

$$\dot{x}^v = \begin{pmatrix} x_1^2 \\ \vdots \\ x_m^2 \end{pmatrix} + A^v x^v + A^c x^d + B^c \begin{pmatrix} x_{m+1}^2 \\ \vdots \\ x_n^2 \end{pmatrix}, \quad x^v(0) = u^v.$$

Choose any $w \in \mathbb{R}^m_{++}$ and let $\rho = \min_i w_i$. Multiplying both sides of the ODE by $w^\top$, we get

$$w^\top \dot{x}^v = x^{v\top} diag(w) x^v + (w^\top A^v) x^v + w^\top A^c x^d + x^d diag(w^\top B^c) x^d.$$

Define $b = A^{v\top} w/2$ and $\tilde{x} = x^v + diag(w)^{-1} b$. Then,

$$\begin{aligned} w^\top \dot{\tilde{x}} &= \tilde{x}^\top diag(w)\tilde{x} - b^\top diag(w)^{-1} b + w^\top A^c x^d + x^d diag(w^\top B^c) x^d \\ &\geq \rho \tilde{x}^\top \tilde{x} - b^\top diag(w)^{-1} b + w^\top A^c x^d + x^d diag(w^\top B^c) x^d \\ &\geq \frac{\rho}{|w|^2} (w^\top \tilde{x})^2 - b^\top diag(w)^{-1} b + w^\top A^c x^d + x^d diag(w^\top B^c) x^d. \end{aligned} \tag{A.2}$$

Let $g(w) = b^\top diag(w)^{-1} b$, $y = w^\top \tilde{x}$ and $y(0) = w^\top \tilde{x}(0)$.

We want to determine whether there is a real number θ such that $x(s)$ blows up as $s \to t$ for the scaled initial condition $x(0) = \theta u$. We divide the rest of the proof into four cases.

Case (i): Suppose $u^v \neq 0$. From(A.2) we get

$$\begin{aligned} \dot{y} &\geq \frac{\rho}{|w|^2} y^2 - g(w) + w^\top A^c x^d \\ &\geq \frac{\rho}{|w|^2} y^2 - g(w) - C|\theta| \cdot |w^\top A^c| \cdot |u^d| \end{aligned} \tag{A.3}$$

with $y(0) = \theta w^\top u^v + e^\top b$, using (3.27) in the second inequality. Now choose w so that $w^\top u^v \neq 0$. Define a new function z by setting

$$\dot{z} = \frac{\rho}{|w|^2} z^2 - g(w) - |\theta| M, \tag{A.4}$$

with $z(0) = y(0)$ and $M = C|w^\top A^c| \cdot |u^d|$; then $y \geq z$ on their common interval of existence. Let $\eta_2 = \sqrt{(g(w) + |\theta|M)|w|^2/\rho}$ and $\eta_1 = -\eta_2$, the two equilibria of the ODE (A.4). Because $w \in \mathbb{R}_{++}$, $g(w) > 0$ so $\eta_2 \neq 0$. By increasing θ (if $w^\top u^v > 0$) or increasing $-\theta$ (if $w^\top u^v < 0$), we can make $z(0) > \eta_2$. Then, as in (3.15), z has a finite blow-up time

$$\tau = \frac{|w|^2}{\rho(\eta_2 - \eta_1)} \log \frac{z(0) - \eta_1}{z(0) - \eta_2}.$$

Since we always have $y \geq z$, τ is an upper bound on the blow-up time of y. Moreover, this upper bound can be made arbitrarily small because $\tau \downarrow 0$ as $\theta \to \infty$ or $\theta \to -\infty$, depending on the sign of $w^\top u^v$. Thus, by taking θ of sufficiently large magnitude and with the sign of $w^\top u^v$, we ensure that x blows up by time t.

Case (ii): Next, suppose $u^v = 0$ but $A^c x^d(s)$ is not identically zero, x^d having initial condition $x^d(0) = u^d$. The solution $x^d(t)$ is given by $\exp(A^d t)u^d$. So, there is some $t_0 < t$ for which $\int_0^{t_0} A^c \exp(A^d t)u^d ds \neq 0$; otherwise, $A^c x^d(s) = 0$ for all $s \in [0, t)$ and this implies $A^c x^d \equiv 0$ because $A^c x^d$ is analytic. Now consider the scaled initial condition $x(0) = \theta u$, and let y be the function defined above by $y = w^\top \tilde{x}$. Then, the initial condition becomes $y(0) = e^\top b$. For $s \leq t_0$, (A.3) yields

$$\dot{y} \geq \frac{\rho}{|w|^2} y^2 - g(w) + w^\top A^c x^d \geq -g(w) + w^\top A^c x^d,$$

and so

$$y(t_0) \geq e^\top b - g(w)t_0 + \theta w^\top \int_0^{t_0} A^c \exp(A^d s)u^d ds.$$

The integral in this expression is nonzero, so the last term is nonzero for some $w \in \mathbb{R}^m_{++}$. On the other hand, for $t \geq t_0$, we use

$$\dot{y} \geq \frac{\rho}{|w|^2} y^2 - g(w) - |\theta|M,$$

with M as before. We can make $y(t_0)$ greater than η_2 by increasing θ or $-\theta$. Applying the same argument we applied to z following (A.4), we conclude that y blows up in time t, and then x does too.

Case (iii): Suppose that $u^v = 0$ and $A^c x^d \equiv 0$, but $B^c(x^2_{m+1}(s), \ldots, x^2_n(s))$ is not identically zero. We can pick $t_0 < t$ such that

$$N \equiv \int_0^{t_0} (\exp(A^d s)u^d)^\top diag(w^\top B^c) \exp(A^d s)u^d ds \neq 0.$$

Now consider x with $x(0) = \theta u$ and take $y = w^\top \tilde{x}$. Then (A.2) yields

$$\begin{aligned} \dot{y} &\geq \frac{\rho}{|w|^2} y^2 - g(w) + w^\top A^c x^d + x^d diag(w^\top B^c) x^d \\ &\geq -g(w) + x^d diag(w^\top B^c) x^d \end{aligned}$$

and so $y(t_0) \geq e^\top b - g(w)t_0 + \theta^2 N$. And we use the following inequality for $t \geq t_0$ (by (A.2)):

$$\dot{y} \geq \frac{\rho}{|w|^2} y^2 - g(w).$$

By the argument in Cases (i)–(ii), we conclude that x blows up by time t for sufficiently large $|\theta|$.

Case (iv): Suppose $u^v = 0$, $A^c x^d \equiv 0$ and $B^c(x^2_{m+1}, \ldots, x^2_n) \equiv 0$. This means that x^v is a solution of

$$\dot{x^v} = \begin{pmatrix} x_1^2 \\ \vdots \\ x_m^2 \end{pmatrix} + A^v x^v, \quad x^v(0) = 0.$$

This makes $x^v \equiv 0$ and thus

$$\mathbb{E} \exp(2\theta u \cdot Y_t) = \exp\left(2\theta^2 \int_0^t |x^d(s)|^2 ds + 2\theta \left(\int_0^t \Lambda^d \cdot x^d(s) ds + x^d(t) \cdot Y_0^d \right) \right)$$

where x^d is the solution from the original (unscaled) initial condition, $x^d(0) = u^d$. Because the moment generating function of $u \cdot Y_t$ is the exponential of a quadratic function of θ, we conclude that $u \cdot Y_t$ is Gaussian. ∎

A.2 Useful Result for Chapter 4

Lemma A.2.1 *Suppose that two real numbers M, c are given satisfying $|u^d| \le M, c \le \min_{j=1,\dots,m} u_j$. Then, there exists a function $v(t)$ such that its dynamics only depends on A, M and $v(0) = c$, and $v(t)$ is bounded below while $\Phi_{t,j}(u) \ge v(t)$ for all $t \in [0, \tau(u))$ and any $j \in \{1, \dots, m\}$.*

Proof The proof is similar to that of Lemma 3.4.2. Since $\Phi_t(u)$ is analytic in t (also in x; see p.44 of Lefschetz 1957), a set $\{s \in [0,t] : \Phi_{s,i}(u) = \Phi_{s,j}(u)\}$ for fixed i, j has finitely many isolated points. Therefore, $\gamma(t) := \min_{k=1,\dots,m} \Phi_{t,k}(u)$ is well-defined in $[0, \tau(u))$ and we can define a sequence of closed intervals $\{I_j\}$ such that $[0, \tau(u)) = \bigcup_j I_j$ and $\gamma(t) = x_{i(j)}$ in I_j^o, interior of I_j, for some index $i(j)$.

If $\gamma(t) = x_i(t)$ in I^o, we have

$$\dot\gamma = \gamma^2 + \sum_{k=1}^{n} A_{ik}x_k + \sum_{k=m+1}^{n} B_{ik}x_k^2 \ge \gamma^2 + \sum_{k=1}^{m} A_{ik}\gamma - C|u^d| \max_{j=1,\dots,m} \sum_{k=m+1}^{n} |A_{jk}|$$

where we use (4.2) and the assumptions that A^v has non-negative off-diagonal entries and that $B \ge 0$. Now define a function v in I^o by

$$\dot v = v^2 + \sum_{k=1}^{m} A_{ik}v - K, \quad K = CM \max_{j=1,\dots,m} \sum_{k=m+1}^{n} |A_{jk}|. \tag{A.5}$$

Starting from $v(0) = c \le \gamma(0)$, $v(t)$ becomes a well-defined piecewise continuous and differentiable function such that $v(t) \le \gamma(t)$. Let us write $I_j^o = (a_j, b_j)$. A simple stability analysis of (A.5) reveals that

— if (A.5) has one or no equilibria, then $v(t)$ increases in I_j,

— if (A.5) has two equilibria, say η_1 and η_2 but $v(a_j) \notin [\eta_1, \eta_2]$, then $v(t)$ does not decrease,

— $v(t)$ decreases only if $v(a_j) \in (\eta_1, \eta_2)$, but then it is bounded below by η_1.

Therefore, $v(t)$ is bounded below and this bound is a function of A, M and c. ∎

A.3 Appendix for Chapter 5

Proof of Theorem 5.2.2. We follow the approach used in Theorem 1 in DPS. Throughout the proof, let us denote $\alpha(T-t)$, $\beta(T-t)$, $A(T-t)$, $B(T-t)$, $C(T-t)$ and $D(T-t)$ by α_t, β_t, A_t, B_t, C_t, D_t, respectively, for notational convenience. We also write $r(X_t)$, $\mu(X_t)$, $\sigma(X_t)$, $\lambda(X_t)$ as r_t, μ_t, σ_t and λ_t. We use a dot, as in $\dot{f}$, to denote a time derivative df/dt. Next we define

$$\Psi_t = \exp\left(-\int_0^t r(X_s)ds\right)e^{\alpha_t+\beta_t\cdot X_t}$$

and $\Phi_t = \Psi_t(A_t + B_t \cdot X_t)$. In addition, we set $\Phi'_t = \Psi_t(A_t + B_t \cdot X_t)^2$ and $\Psi'_t = \Psi_t(C_t + D_t \cdot X_t)$. If we show that $\Phi'_t + \Psi'_t$ is a martingale, then $\Phi'_t + \Psi'_t = \mathbb{E}[\Phi'_T + \Psi'_T \mid \mathcal{F}_t]$ leads to the desired result.

Itô's formula for jump-diffusion processes (as in Cont and Tankov 2003) yields

$$\begin{aligned} d\Phi'_t &= \Phi'_t\Big((-r_t + \dot{\alpha}_t + \dot{\beta}_t \cdot X_t)dt + (\beta_t \cdot \mu_t dt + \beta_t^\top \sigma_t dW_t) + \frac{1}{2}\beta^\top(\sigma_t\sigma_t^\top)\beta_t dt\Big) \\ &\quad +2\Phi_t\Big((\dot{A}_t + \dot{B}_t \cdot X_t)dt + (B_t \cdot \mu_t dt + B_t^\top \sigma_t dW_t) + \beta_t^\top(\sigma_t\sigma_t^\top)B_t dt\Big) \\ &\quad +\Psi_t B_t^\top(\sigma_t\sigma_t^\top)B_t dt + dJ_t \\ &= \Pi_t dt + \Upsilon_t dW_t + dJ_t \end{aligned}$$

for appropriate drift and volatility coefficients Π_t, Υ_t and $J_t = \sum_{0<\tau(i)\le t}(\Phi'_{\tau(i)} - \Phi'_{\tau(i)-})$ with $\tau(i) = \inf\{t : N_t = i\}$. Here N_t is the counting process with intensity λ_t. Letting $\mathbb{E}_t$ be the $\mathcal{F}_t$-conditional expectation under $\mathbb{P}$ for $0 \le t \le s \le T$, and writing $\triangle X_i$ for the increment in X at $\tau(i)$, we have

$$\begin{aligned} &\mathbb{E}_t\Big[\sum_{t<\tau(i)\le s}(\Phi'_{\tau(i)} - \Phi'_{\tau(i)-})\Big] \\ =\ &\mathbb{E}_t\Big[\sum_{t<\tau(i)\le s}\mathbb{E}[\Phi'_{\tau(i)} - \Phi'_{\tau(i)-} \mid X_{\tau(i)-}, \tau(i)]\Big] \\ =\ &\mathbb{E}_t\Big[\sum_{t<\tau(i)\le s}\Big\{\Phi'_{\tau(i)-}\Big(\mathbb{E}_{\tau(i)-}e^{\beta_{\tau(i)}\cdot\triangle X_i} - 1\Big) + 2\Phi_{\tau(i)-}\mathbb{E}_{\tau(i)-}[e^{\beta_{\tau(i)}\cdot\triangle X_i}B_{\tau(i)} \cdot \triangle X_i] \\ &\quad +\Psi_{\tau(i)-}\mathbb{E}_{\tau(i)-}[e^{\beta_{\tau(i)}\cdot\triangle X_i}(B_{\tau(i)} \cdot \triangle X_i)^2]\Big\}\Big] \\ =\ &\mathbb{E}_t\Big[\int_{t+}^{s}\Big\{\Phi'_{u-}(\theta(\beta_u) - 1) + 2\Phi_{u-}\nabla\theta(\beta_u)B_u + \Psi_{u-}B_u^\top\nabla^2\theta(\beta_u)B_u\Big\}dN_u\Big]. \end{aligned}$$

Proceeding similarly,

$$d\Psi'_t = \tilde{\Pi}_t dt + \tilde{\Upsilon}_t dW_t + d\tilde{J}_t$$

for suitable coefficients $\tilde{\Pi}_t$, $\tilde{\Upsilon}_t$ (they are straightforward to compute, but omitted to save some space) and $\tilde{J}_t = \sum_{0<\tau(i)\leq t}(\Psi'_{\tau(i)} - \Psi'_{\tau(i)-})$. The last term satisfies

$$\mathbb{E}_t\Big[\sum_{t<\tau(i)\leq s}(\Psi'_{\tau(i)} - \Psi'_{\tau(i)-})\Big] = \mathbb{E}_t\Big[\int_{t+}^{s}\big\{\Psi'_{u-}(\theta(\beta_u)-1) + \Psi_{u-}\nabla\theta(\beta_u)D_u\big\}dN_u\Big].$$

Now, we observe that if the condition (i) of Definition 5.2.1 is satisfied, then

$$\mathbb{E}_t\Big[J_s + \tilde{J}_s - J_t - \tilde{J}_t\Big] = \mathbb{E}_t\Big[\int_{t+}^{s}\gamma(u-)dN_u\Big] = \mathbb{E}_t\Big[\int_t^s \gamma(u)\lambda_u du\Big]$$

and $J_t + \tilde{J}_t - \int_0^t \gamma(u)\lambda_u du$ becomes a martingale thanks to the Integration theorem in p.27 of Brémaud (1981).

From these observations, by adding and subtracting $\gamma(t)\lambda_t dt$ we get

$$\begin{aligned}
d(\Phi'_t + \Psi'_t) &= d(J_t + \tilde{J}_t) - \gamma(t)\lambda_t dt + (\Upsilon_t + \tilde{\Upsilon}_t)dW_t \\
&\quad + \Phi'_t\Big(-r_t + \dot{\alpha}_t + \dot{\beta}_t \cdot X_t + \beta_t \cdot \mu_t + \frac{1}{2}\beta^\top(\sigma_t\sigma_t^\top)\beta_t + (\theta(\beta_t)-1)\lambda_t\Big)dt \\
&\quad + 2\Phi_t\Big(\dot{A}_t + \dot{B}_t \cdot X_t + B_t \cdot \mu_t + \beta_t^\top(\sigma_t\sigma_t^\top)B_t + \nabla\theta(\beta_t)B_t\lambda_t\Big)dt \\
&\quad + \Psi_t\Big(B_t^\top(\sigma_t\sigma_t^\top)B_t + B_t^\top\nabla^2\theta(\beta_t)B_t\lambda_t\Big)dt \\
&\quad + \Psi'_t\Big(-r_t + \dot{\alpha}_t + \dot{\beta}_t \cdot X_t + \beta_t \cdot \mu_t + \frac{1}{2}\beta^\top(\sigma_t\sigma_t^\top)\beta_t + (\theta(\beta_t)-1)\lambda_t\Big)dt \\
&\quad + \Psi_t\Big(\dot{C}_t + \dot{D}_t \cdot X_t + D_t \cdot \mu_t + \beta_t^\top(\sigma_t\sigma_t^\top)D_t + \nabla\theta(\beta_t)D_t\lambda_t\Big)dt \\
&= d(J_t + \tilde{J}_t) - \gamma(t)\lambda_t dt + (\Upsilon_t + \tilde{\Upsilon}_t)dW_t \qquad \text{(A.6)}
\end{aligned}$$

as α_t, β_t, A_t, B_t, C_t and D_t are solutions to (5.1)–(5.6). The condition (ii) of Definition 5.2.1 ensures that $\int_0^t(\Upsilon_u + \tilde{\Upsilon}_u)dW_u$ is a martingale. Therefore, $\Phi'_t + \Psi'_t$ is a martingale and the proof is complete. ∎

Theorem 5.2.2 can also be established as a consequence of Proposition 2 in Cheng and Scaillet (2002); for higher-order derivatives we need to consider higher powers of $b \cdot X_T$, and these require separate treatment.

Conditions for Theorem 5.2.3. The characteristics (K, H, l, θ, ρ) are well-behaved at (v, u, T), if all ODEs in Theorems 5.2.1, 5.2.2, 5.2.3 are solved uniquely, if θ is three times differentiable at $\beta(t)$ for all $t \leq T$, and if the following conditions are satisfied:

$$
\begin{aligned}
(i)\quad & \mathbb{E}\Big[\int_0^T |\gamma(t)\lambda(X_t)|dt\Big] < \infty, \\
& \text{where } \gamma(t) = f_1(t) + f_2(t) + f_3(t), \\
& f_1(t) := \Phi_t^1(\theta(\beta_t) - 1) + 3\Psi_t\Big\{(A_t + B_t \cdot X_t)^2 \nabla\theta(\beta_t)B_t \\
& \quad +(A_t + B_t \cdot X_t)B_t^\top \nabla^2\theta(\beta_t)B_t\Big\} + \Psi_t \int_{\mathbb{R}^n} e^{z\cdot\beta_t}(z \cdot B_t)^3 d\nu(z) \\
& f_2(t) := \Phi_t^2(\theta(\beta_t) - 1) + 3\Psi_t\Big\{(A_t + B_t \cdot X_t)\nabla\theta(\beta_t)D_t \\
& \quad +(C_t + D_t \cdot X_t)\nabla\theta(\beta_t)B_t\Big\} + \Psi_t B_t^\top \nabla^2\theta(\beta_t)D_t \\
& f_3(t) := \Phi_t^3(\theta(\beta_t) - 1) + \Psi_t \nabla\theta(\beta_t)F_t \\
(ii)\quad & \mathbb{E}\Big[\Big(\int_0^T \eta(t) \cdot \eta(t)dt\Big)^{1/2}\Big] < \infty, \text{ where } \eta(t) = (g_1(t) + g_2(t) + g_3(t))\,\sigma(X_t) \\
& g_1(t) := \Phi_t^1\beta_t^\top + 3\Psi_t(A_t + B_t \cdot X_t)^2 B_t^\top \\
& g_2(t) := \Phi_t^2\beta_t^\top + 3\Psi_t\Big\{(C_t + D_t \cdot X_t)B_t^\top + (A_t + B_t \cdot X_t)D_t^\top\Big\} \\
& g_3(t) := \Phi_t^3\beta_t^\top + \Psi_t F_t^\top \\
(iii)\quad & \mathbb{E}\Big[\Big|\Phi_T^1 + \Phi_T^2 + \Phi_T^3\Big|\Big] < \infty
\end{aligned}
$$

where Ψ_t, Φ_t^i for $i = 1, 2, 3$ are defined in the proof of Theorem 5.2.3 and $\alpha_t, \ldots, F_t$ stand for $\alpha(T-t), \ldots, F(T-t)$ which are the solutions to (5.1)–(5.8). ∎

Proof of Theorem 5.2.3. This can be proved by defining appropriate functions, as in the previous theorems. We set $\Psi_t = \exp(-\int_0^t r(X_s)ds)e^{\alpha(T-t)+\beta(T-t)\cdot X_t}$ as before and

$$
\begin{aligned}
\Phi_t^1 &= (A(T-t) + B(T-t) \cdot X_t)^3 \Psi_t \\
\Phi_t^2 &= 3(A(T-t) + B(T-t) \cdot X_t)(C(T-t) + D(T-t) \cdot X_t)\Psi_t \\
\Phi_t^3 &= (E(T-t) + F(T-t) \cdot X_t)\Psi_t
\end{aligned}
$$

and apply Itô's formula. Under the assumed conditions, $\Phi_t^1 + \Phi_t^2 + \Phi_t^3$ becomes a martingale.

∎

Characteristic of the model dynamics in the Heston model:

$$K_0 = \begin{pmatrix} r \\ \kappa\theta \end{pmatrix}, \; K_1 = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & -\kappa \end{pmatrix}, \; H_0 = 0,$$

$$H_{1,11} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; H_{1,12} = H_{1,21} = \begin{pmatrix} 0 \\ \sigma\rho \end{pmatrix}, \; H_{1,22} = \begin{pmatrix} 0 \\ \sigma^2 \end{pmatrix}$$

Characteristic of the model dynamics in the SVJ model:

$$K_0 = \begin{pmatrix} r - \lambda k \\ \kappa\theta \end{pmatrix}, \; K_1 = \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & -\kappa \end{pmatrix}, \; H_0 = 0,$$

$$H_{1,11} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; H_{1,12} = H_{1,21} = \begin{pmatrix} 0 \\ \sigma\rho \end{pmatrix}, \; H_{1,22} = \begin{pmatrix} 0 \\ \sigma^2 \end{pmatrix},$$

$$\theta(c) = \int_{\mathbb{R}^2} \exp(c \cdot z) d\nu(z) = \exp(c_1\mu_J + c_1^2\sigma_J^2/2), \; l_0 = \lambda, \; l_1 = 0$$

Characteristic of the model dynamics in the Scott model:

$$K_0 = \begin{pmatrix} -\lambda k \\ \kappa_1\theta_1 \\ \kappa_2\theta_2 \end{pmatrix}, \; K_1 = \begin{pmatrix} 0 & 1 - \frac{1}{2}\sigma^2 & 1 \\ 0 & -\kappa_1 & 0 \\ 0 & 0 & -\kappa_2 \end{pmatrix}, \; H_0 = 0,$$

$$H_{1,11} = \begin{pmatrix} 0 \\ \sigma^2 \\ 0 \end{pmatrix}, \; H_{1,12} = H_{1,21} = \begin{pmatrix} 0 \\ \rho\sigma\sigma_1 \\ 0 \end{pmatrix}, \; H_{1,22} = \begin{pmatrix} 0 \\ \sigma_1^2 \\ 0 \end{pmatrix}, \; H_{1,33} = \begin{pmatrix} 0 \\ 0 \\ \sigma_2^2 \end{pmatrix},$$

$$H_{1,13} = H_{1,31} = H_{1,23} = H_{1,32} = 0,$$

$$\theta(c) = \int_{\mathbb{R}^2} \exp(c \cdot z) d\nu(z) = \exp(c_1\mu_J + c_1^2\sigma_J^2/2), \; l_0 = \lambda, \; l_1 = 0$$

Bibliography

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

L. V. Ahlfors. *Complex Analysis.* McGRAW-HILL, 3rd and international edition, 1979.

Y. Aït-Sahalia and J. Yu. Saddlepoint approximations for continuous-time Markov processes. *Journal of Econometrics*, 134:507–551, 2006.

L. Andersen. Efficient simulation of the Heston stochastic volatility model. Working Paper, Banc of America Securities, 2005.

L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11:29–50, 2007.

T. G. Andersen, L. Benzoni, and J. Lund. An empirical investigation of continuous-time equity return models. *Journal of Finance*, 57:1239–1284, 2002.

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

J. S. Baris, P. J. Baris, and B. Ruchlewicz. On blow-up solutions of nonautonomous quadratic differential systems. *Differential Equations*, 42:320–326, 2006.

O. E. Barndorff-Nielsen. *Information and Exponential Families in Statistical Theory.* Wiley, 1978.

B. Basrak, R. A. Davis, and T. Mikosch. Regular variation of GARCH processes. *Stochastic Processes and Their Applications*, 99:95–115, 2002.

D. S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies*, 9:69–107, 1996.

D. S. Bates. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94: 181–238, 2000.

D. S. Bates. Empirical option pricing: A retrospection. *Journal of Econometrics*, 116:387–404, 2003.

S. Benaim and P. Friz. Regular variation and smile asymptotics. *Mathematical Finance*, 2006. Forthcoming.

A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences.* SIAM, Philadelphia, 1994.

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

N. Bleistein and R. A. Handelsman. *Asymptotic Expansion of Integrals*. Holt, Rinehart and Winston, New York, 1975.

L. Bondesson. On simulation from inifinitely divisible distributions. *Advances in Applied Probability*, 14:855–869, 1982.

Brémaud. *Point Processes and Queues*. Springer-Verlag, New York, 1981.

M. Broadie, M. Chernov, and M. Johannes. Model specification and risk premia: Evidence from futures options. *Journal of Finance*, 62:1453–1490, 2007.

M. Broadie and Ö. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54:217–231, 2006.

R. H. Brown and S. M. Schaefer. Interest rate volatility and the shape of the term structure. *Philosophical Transactions of the Royal Society of London, Series A*, 347:563–576, 1994.

R. W. Butler and A. T. A. Wood. Saddlepoint approximation for moment generating functions of truncated random variables. *Annals of Statistics*, 32:2712–2730, 2004.

R. Chen and L. O. Scott. Stochastic volatility and jumps in interest rates: An empirical analysis. Working Paper, Rutgers University, 2002.

D. Cheng, J. Ma, Q. Lu, and S. Mei. Quadratic form of stable sub-manifold for power systems. *International Journal of Robust and Nonlinear Control*, 14:773–788, 2004.

P. Cheng and O. Scaillet. Linear-quadratic jump-diffusion modelling with application to stochastic volatility. FAME Research Paper No.67, University of Geneva, Switzerland, 2002.

P. Cheridito, D. Filipović, and R.L. Kimmel. A note on the Dai-Singleton canonical representation of affine term structure models. Working Paper, Ohio State University, 2006.

M. Chernov. Empirical reverse engineering of the pricing kernel. *Journal of Econometrics*, 116: 329–364, 2003.

M. Chernov and E. Ghysels. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of option valuation. *Journal of Financial Economics*, 56:407–458, 2000.

M. Chernov, E. Ghysels, A. R. Gallant, and G. Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics*, 116:225–257, 2003.

H.-D. Chiang and L. Fekih-Ahmed. Quasi-stability regions of nonlinear dynamical systems: Optimal estimations. *IEEE Transactions on Circuits and Systems–I: Fundamental Theory and Applications*, 43:636–643, 1996.

H.-D. Chiang, M. W. Hirsch, and F. F. Wu. Stability regions of nonlinear autonomous dynamical systems. *IEEE Transactions on Automatic Control*, 33:16–27, 1988.

C. Chicone and D. S. Shafer. Separatrix and limit cycles of quadratic systems and Dulac's Theorem. *Transactions of the American Mathematical Society*, 278:585–612, 1983.

K. L. Chung. *A Course in Probability Theory*. Academic Press, 3rd edition, 2001.

P. Collin-Dufresne and R. Goldstein. Pricing swaptions within an affine framework. *Journal of Derivatives*, 10:9–26, 2002.

R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall / CRC Press, Boca Raton, Florida, 2003.

J. C. Cox. Notes on options pricing I: Constant elasticity of variance diffusions. Working Paper, Stanford University, 1975.

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

P. E. Crouch and M. Pavon. On the existence of solutions of the Riccati differential equation. *Systems and Control Letters*, 9:203–206, 1987.

J.-P. Crouziex. A review of continuity and differentiability properties of quasiconvex functions on $\mathbb{R}^n$. In J.-P. Aubin and R. Vinter, editors, *Convex Analysis and Optimization*, pages 18–34. Pitman Advanced Publishing Programs, 1982. Research Notes in Mathematics, vol.57.

Q. Dai and K. Singleton. Specification analysis of affine term structure models. *Journal of Finance*, 55:1943–1978, 2000.

H. E. Daniels. Saddlepoint approximations in statistics. *Annals of Mathematical Statistics*, 25:631–650, 1954.

H. E. Daniels. Tail probability approximations. *International Statistical Review*, 55:37–48, 1987.

A. Dembo, J.-D. Deuschel, and D. Duffie. Large portfolio losses. *Finance and Stochastics*, 8:3–16, 2004.

A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer-Verlag, New York, 1998.

L. Devroye. Simulating Bessel random variables. *Statistics & Probability Letters*, 57:249–257, 2002.

D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, 3rd edition, 2001.

D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053, 2003.

D. Duffie and P. Glynn. Efficient Monte Carlo simulation of security prices. *Annals of Applied Probability*, 5:897–905, 1995.

D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.

D. Duffie and J. Pan. Analytical value-at-risk with jumps and credit risk. *Finance and Stochastics*, 5: 155–180, 2001.

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

D. Duffie, L. H. Pedersen, and K. Singleton. Modeling sovereign yields spreads: A case study of Russian debt. *Journal of Finance*, 58:119–160, 1997.

D. Duffie and K. Singleton. An econometric model of the term structure of interest-rate swap yields. *Journal of Finance*, 52:1287–1321, 1997.

G. S. Easton and E. Ronchetti. General saddlepoint approximations with applications to L statistics. *Journal of the American Statistical Association*, 81:420–430, 1986.

B. Eraker. Do stock prices and volatility jump? reconciling evidence from spot and option prices. *Journal of Finance*, 59:1367–1403, 2004.

B. Eraker, M. Johannes, and N. Polson. The impact of jumps in equity index volatility and returns. *Journal of Finance*, 58:1269–1300, 2003.

R. Genesio, M. Tartaglia, and A. Vicino. On the estimation of asymptotic stability regions: State of the art and new proposals. *IEEE Transactions on Automatic Control*, 8:747–755, 1985.

W. M. Getz and D. H. Jacobson. Sufficiency conditions for finite escape times in systems of quadratic differential equations. *IMA Journal of Applied Mathematics*, 19:377–383, 1977.

P. Glasserman and K.-K. Kim. Saddlepoint approximations for affine jump-diffusion models. Working Paper, Columbia Business School, 2008.

M. B. Gordy. Saddlepoint approximation of CreditRisk$^+$. *Journal of Banking & Finance*, 26:1335–1353, 2002.

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

M. W. Hirsch and S. Smale. *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press, 1974.

R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

G. Iliopoulos and D. Karlis. Simulation from the Bessel distribution with applications. *Journal of Statistical Computation and Simulation*, 73:491–506, 2003.

J. L. Jensen. *Saddlepoint Approximations*. Oxford University Press, Oxford, United Kingdom, 1995.

M. Jonsson and K. R. Sircar. Partial hedging in a stochastic volatility environment. *Mathematical Finance*, 12:375–409, 2002.

C. Kahl and P. Jäckel. Fast strong approximation Monte Carlo schemes for stochastic volatility models. *Quantitative Finance*, 6:513–536, 2006.

O. Kallenberg. *Foundations of Modern Probability*. Springer, 2nd edition, 2002.

P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 3rd edition, 1999.

S. G. Kou. A jump diffusion model for option pricing. *Management Science*, 48:1086–1101, 2002.

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

S. Lefschetz. *Differential Equations: Geometric Theory*. Interscience Publishers, Inc., New York, 1957.

M. Leippold and L. Wu. Asset pricing under the quadratic class. *Journal of Financial and Quantitative Analysis*, 37:271–295, 2002.

A. Levin. An analytical method of estimating the domain of attraction for polynomial differential equations. *IEEE Transactions on Automatic Control*, 39:2471–2475, 1994.

O. Lieberman. On the approximation of saddlepoint expansions in statistics. *Econometric Theory*, 10:900–916, 1994.

F. A. Longstaff and E. S. Schwartz. Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance*, 47:1259–1282, 1992.

R. Lugannani and S. Rice. Saddle point approximation for the distribution of the sum of independent random variables. *Advances in Applied Probability*, 12:475–490, 1980.

C. Martin. Finite escape time for Riccati differential equations. *Systems and Control Letters*, 1:127–131, 1981.

R. Martin, K. Thompson, and C. Browne. Taking to the saddle. *Risk*, 14, June:91–94, 2001.

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

M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling*. Springer, 2nd edition, 2005.

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

J. Pitman and M. Yor. A decomposition of Bessel bridges. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 59:425–457, 1982.

J. Pitman and M. Yor. Infinitely divisible laws associated with hyperbolic functions. Technical report, University of California, Berkeley, 2000. No.581.

A. D. Polyanin and V. F. Zaitsev. *Handbook of Exact Solutions for Ordinary Differential Equations*. Chapman & Hall / CRC, 2nd edition, 2003.

D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, 3rd edition, 1999.

L. C. G. Rogers and O. Zane. Saddlepoint approximations to option prices. *Annals of Applied Probability*, 9:493–503, 1999.

S. Saha, A. Fouad, W. Kliemann, and V. Vittal. Stability boundary approximation of a power system using the real normal form of vector fields. *IEEE Transactions on Power Systems*, 12:797–802, 1997.

T. Sasagawa. On the finite escape phenomena for matrix Riccati equations. *IEEE Transactions on Automatic Control*, AC-27:977–979, 1982.

L. O. Scott. Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Applications of fourier inversion methods. *Mathematical Finance*, 7:413–424, 1997.

S. E. Shreve. *Stochastic Calculus for Finance II, Continuous-Time Models*. Springer-Verlag, New York, 2004.

K. Singleton. *Empirical Dynamic Asset Pricing*. Princeton University Press, Princeton, New Jersey, 2006.

S. Smale. Differentiable dynamical systems. *Bulletin of American Mathematical Society*, 73:747–817, 1967.

A. J. Sommese and C. W. Wampler. *The Numerical Solution of Systems of Polynomials*. World Scientific, 2005.

M. D. Springer. *The Algebra of Random Variables*. John Wily & Sons, 1979.

F. W. Steutel and K. van Harn. *Infinite Divisibility of Probability Distributions on the Real Line*. Marcel Dekker, 2004.

B. Tibken. Estimation of the domain of attraction for polynomial systems via LMI's. *Proceedings of the 39th IEEE Conference on Decision and Control, Sydney, Australia*, pages 3860–3864, 2000.

A. Vannelli and M. Vidyasagar. Maximal Lyapunov functions and domains of attraction for autonomous nonlinear systems. *Automatica*, 21:69–80, 1985.

O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5: 177–188, 1977.

F. Verhulst. *Nonlinear Differential Equations and Dynamical Systems*. Springer, 1996.

P. Volkmann. Gewöhnliche differentialungleichungen mit quasimonoton wachsenden funktionen in topologischen vektorräumen. *Mathematische Zeitschrift*, 127:157–164, 1972.

S. Wang. General saddlepoint approximations in the bootstrap. *Statistics & Probability Letters*, 13: 61–66, 1992.

A. T. A. Wood, J. G. Booth, and R. W. Butler. Saddlepoint approximations to the CDF of some statistics with nonnormal limit distribution. *Journal of the American Statistical Association*, 88: 680–686, 1993.

J. Yang, T. Hurd, and X. Zhang. Saddlepoint approximation method for pricing CDOs. *Journal of Computational Finance*, 10:1–20, 2006.

L. Yuan and J. D. Kalbfleisch. On the Bessel distribution and related problems. *Annals of the Institute of Statistical Mathematics*, 52:438–447, 2000.