The Estimation of Stochastic Models in Finance with Volatility and Jump Intensity

by

David Edward A. Wilson

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Statistics

Waterloo, Ontario, Canada, 2018

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	NAME: Liuren Wu Title: Full Professor
Internal-External Examiner	NAME: Kenneth Vetzal Title: Associate Professor
Committee Members	NAME: Adam Kolkiewicz Title: Associate Professor NAME: Martin Lysy Title: Associate Professor
Supervisors	NAME: Don McLeish Title: Professor NAME: Tony Wirjanto Title: Professor

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

David Edward A. Wilson

Abstract

This thesis covers the parametric estimation of models with stochastic volatility, jumps, and stochastic jump intensity, by FFT. The first primary contribution is a parametric minimum relative entropy optimal Q-measure for affine stochastic volatility jump-diffusion (ASVJD). Other attempts in the literature have minimized the relative entropy of Q given P either by nonparametric methods, or by numerical PDEs. These methods are often difficult to implement. We construct the relative entropy of Q given P from the Lebesgue densities under P and Q, respectively, where these can be retrieved by FFT from the closed form log-price characteristic function of any ASVJD model. We proceed by first estimating the fixed parameters of the P-measure by the Approximate Maximum Likelihood (AML) method of Bates (2006), and prove that the integrability conditions required for each Fourier inversion are satisfied. Then by using a structure preserving parametric model under the Q-measure, we minimize the relative entropy of Q given P with respect to the model parameters under Q. AML can be used to estimate P within the ASVJD class. Since, AML is much faster than MCMC, our main supporting contributions are to the theory of AML.

The second main contribution of this thesis is a non-affine model for time changed jumps with stochastic jump intensity called the Leveraged Jump Intensity (LJI) model. The jump intensity in the LJI model is modeled by the CIR process. Leverage occurs in the LJI model, since the Brownian motion driving the CIR process also appears in the log-price with a negative coefficient. Models with a leverage effect of this type are usually affine, but model the intensity with an Ornstein-Uhlenbeck process. The conditional characteristic function of the LJI log-price given the intensity is known in closed form. Thus, we price LJI call options by conditional Monte Carlo, using the Carr and Madan (1999) FFT formula for conditional pricing.

Acknowledgements

I would like to offer my sincere thanks most of all to each of my two thesis supervisors, Dr. Don McLeish and Dr. Tony S. Wirjanto, for their patience and support during this thesis. I am also greatly indebted to my committee members, Dr. Adam Kolkiewicz and Dr. Martin Lysy, for their helpful suggestions and encouragement.

My graduate studies professors, Drs. Bovas Abraham, Greg Bennett, Christian Boudreaux, Peter Forsyth, Andrew Heunis, Jeanette O'Hara-Hines, Chris Small, and Changbao Wu, have each taught me well.

I would also like to thank department chairs Paul Marriott, Christiane Lemieux, and Stefan Steiner, for making this degree possible, as well as Mary Lou Dufton, Mary Hardy, Laurent Marcoux, Cyntha Struthers, John Wainwright, Ross Willard, Henry Wolkowicz, and the many others who have given me the strength to undertake this degree.

In addition, I would like to thank David S. Bates, Stephen Joe, and Francis Kuo for their correspondence with me during my PhD research.

For financial support, I am very grateful to the Department of Statistics and Actuarial Science at the University of Waterloo, the Queen Elizabeth II Graduate Scholarship in Science and Technology, the WatRISQ TD Meloche Monnex Scholarship in Quantitative Finance and Insurance, and the Royal Bank of Canada Graduate Student Loans Program.

Dedication

This thesis is dedicated to my mother and father who have stood by me for many years during my lengthy education.

To

Sheila Maynard Grange

and

John MacKenzie Wilson

Contents

List of Figures

List of Tables

Chapter 1

Introduction to the Thesis

Motivation

At least since Mandelbrot (1963) it has been known that the log-return distribution of financial assets exhibits thicker than normal tails. This is one criticism of the Black-Scholes (1973) paradigm. Merton (1976) answered this by modeling the return process with jump-diffusion. Moreover, Black (1976) finds that the latent volatility should follow its own stochastic process. Indeed, Black (1976), see p. 179, argues that if the stock price falls then the relative level of the debt of the firm rises, and this causes the volatility for the firm to rise. Consequently, the negative correlation between the latent volatility and the stock price is known as the leverage effect. The Heston (1993) stochastic volatility model with $\rho < 0$ has become synonymous with the leverage effect, although Heston (1993) never uses the word leverage.

Apart from the Paretian stable processes of Mandelbrot (1963) there have been many attempts to find an alternative to normally distributed log-returns. For example McLeish (1982), the Variance Gamma model of Madan, Carr, and Chang (1998), the NIG of Barndorff-Nielsen (1998), and the CGMY model of Carr, Geman, Madan, and Yor (2002). All of these

models are for jumps with infinite-activity, and each was originally seen as a replacement for, rather than a complement to, Brownian motion. This was particularly the case in Carr, Geman, Madan, and Yor (2002). However, Brownian stochastic volatility models that include jumps in the log-price of either finite-activity or infinite-activity have been successfully estimated, for example in Huang and Wu (2004) under the Q-measure using historical daily options price data, and by Li, Wells, and Yu (2008) under the P-measure using historical daily log-return data. Furthermore, Huang and Wu (2004) considers the possibility introduced in Bates (2000) that the jumps have a stochastic jump intensity modeled by the integrated variance, and Bates (2006) successfully estimates this model under the P-measure. Thus, the estimation of stock price models with a diffusion, stochastic volatility, jumps, and stochastic jump intensity is a reasonable proposition. Moreover, in this thesis we do so entirely by fast Fourier transform techniques.

This thesis considers two main problems, and their solutions. First, if there is an additional latent factor such as volatility or jump intensity, then there is an additional source of randomness and the model for the market is incomplete. Thus, an optimal martingale measure needs to be selected. We provide an easily computable solution for this problem called the Parametric Minimum Entropy Martingale Measure (PMEMM). Secondly, consider the stochastic exponential time changed Lévy process model of Carr, Geman, Madan, and Yor (2003), with continuous stochastic jump intensity modeled by the CIR process of Cox, Ingersoll, and Ross (1985). By relaxing the assumption of an affine log-price characteristic function, we generalize this model so that it has a leverage effect. Our solution is called the Leveraged Jump Intensity model (LJI). We are able to price by a specialized method of conditional Monte Carlo. Furthermore, as we introduce the key themes of this thesis below, further supporting contributions will be revealed.

Key Themes

The six central themes of this thesis are as follows:

1. The Fast Fourier Transform
2. Estimation for the Affine Markov Models
3. Infinite-Activity Jumps
4. Incomplete Markets
5. Estimation for Stochastic Jump Intensity with Leverage
6. Joint Characteristic Functions

These six key themes will draw a complete picture of the thesis below. The two main contributions are highlighted again following the key themes.

The Fast Fourier Transform

We apply Fourier inversion via FFT in all estimation methods of this thesis. A detailed description of Fourier inversion, and our FFT contribution is provided in Chapter 3. Our method is based on the non-centred and shifted FFT implementation of Carr and Madan (1999), p. 68, and Carr, Geman, Madan, and Yor (2002), p. 320, with one important difference. We use the trapezoidal rule for quadrature, see Briggs and Henson (1995), p. 360. We find in Chapter 5 that Simpson's rule, as advocated in Carr and Madan (1999), p. 68, is what leads to the problem, identified in Carr and Madan (2009), Table 2, p. 60, of a negative call price deep out-of-the-money.

Estimation for the Affine Markov Models

There are five specific affine Markov models estimated in this thesis. These are the Heston model, the Heston model with three respectively different

independent jump types in the log-price (SVJ), and the Heston model with compound Poisson Merton (1976) jumps in the log-price, time changed by the integrated variance (SVSJ). The three SVJ independent jump types are Merton (1976), Huang and Wu (2004) type Variance Gamma, and the Meixner jumps of Schoutens and Teugels (1998). An introduction to these five models and their properties is given in Chapter 2. The P-measure and the Q-measure of Apple stock are estimated separately under each of the five the affine Markov models in Chapter 4 and Chapter 5 respectively.

In Chapter 4 we estimate the P-measure by the Bates (2006) method of Approximate Maximum Likelihood (AML). This method is a subsequence of three Fourier inversions evaluated iteratively through the data. The first inversion recovers the current point of the likelihood, and then the filtered volatility is updated by the second and third inversions. We contribute proof in subsection 4.3.1 that for the affine Markov models of this thesis, the L^1 integrability conditions sufficient for inversion of all three Fourier transforms of the AML method are met. In Appendix D of this thesis, we also contribute proof that for the five affine Markov models, the primary Fourier transform of the AML method may be twice differentiated under the integral with respect to the second transform variable, $v \in (u, v)$, as required by AML. We perform each Fourier inversion by FFT and approximate each inverse transform with a spline. There are two main advantages of AML over the Markov Chain Monte Carlo (MCMC), see for example Li, Wells, and Yu (2008) for MCMC. These are the speed of estimation and the authenticity of the filtered volatility. We also perform simulated estimation under AML on both the Apple stock daily log-returns (1991-2011), and the Dow-Jones Industrial Average (DJIA) daily log-returns (1988-2007).

In Chapter 5 we take a first step towards our ultimate goal of attaining a set of optimal Q-measures for Apple stock under the five affine Markov

models. We estimate the option implied Q-measures under least squares normal likelihood using our version of the Carr and Madan (1999) FFT formula with the trapezoidal rule for quadrature. We estimate based on a selection spot call options from the daily closing book on Apple stock from January 19th, 2011, the last day in the Apple stock (1991-2011) log-return data set. While these spot option results show the SVSJ Merton model to be slightly better than the others, the evidence from Chapter 4 under the P-measure is that Apple stock clearly has infinite-activity jumps.

Infinite-Activity Jumps

In addition to the P-measures for the DJIA (1988-2007) and Apple stock (1991-2011), in Chapter 4 we also use AML to estimate the P-measures of the S&P 500 (1988-2007), and six other major stock symbols for individual stocks (1988-2007). Thus, we make a contribution to the debate over jump types in models with continuous stochastic volatility and jumps. For a summary, see subsection 4.6.3. We find, consistent with Aït-Sahalia and Jacod (2012), that the DJIA has finite-activity jumps. But, we also find, consistent with Huang and Wu (2004), that the S&P 500 appears to contain an infinite variation pure jump process. Moreover, we find that banking and technology stocks are well suited to an infinite-activity jump component, and this is particularly the case for Apple stock (1991-2011). This is consistent with Aït-Sahalia and Jacod (2012), where it is suggested that individual DJIA components have infinite-activity jumps.

In Chapter 8 we consider the problem of simulating daily increments for the infinite variation Meixner process of Schoutens and Teugels (1998). We considered both Grigoletto and Provasi (2009), and Madan and Yor (2008), but found their methods to only be suitable for large Meixner increments. The Meixner scale parameter does not scale time, so it is not possible to

scale larger Meixner increments into smaller Meixner increments. Thus, in Chapter 8 we contribute a new method called least squares rejection, with an NIG rejection density, suitable for simulating daily Meixner increments. We use this method for the simulated AML estimations of Chapter 4.

Incomplete Markets

Estimation for incomplete market models is often handled by the simultaneous estimation jointly under P and Q of the log-return data and a time series of nearest to the money historical options prices, see for example Li, Wells, and Yu (2011). The measure change is accounted for by estimating the market prices of risk. The key problem with this approach is that the market prices of risk are generally not estimable, even for the Heston model, see Aït-Sahalia and Kimmel (2007), p. 442. Hence, in Chapter 6 we pursue an optimal martingale measure, similar in nature to the Minimal Entropy Martingale Measure (MEMM) of Frittelli (2000), only we propose something more easily computable.

For Lévy processes, and each of the five respective affine Markov models of this thesis, a description of the incomplete markets phenomenon is given in Chapter 2. For each model, we choose a structure preserving martingale measure such that the respective models under P and Q have the same characteristic function when the drift $\mu_0 = r$, but under the assumption that the parameters $\theta_P \in \Omega_P$ and $\theta_Q \in \Omega_Q$ may differ. We prove in Chapter 3 that the log-price characteristic functions for the affine Markov models are L^1. This implies that the respective measures P and Q are each absolutely continuous, so that for fixed $\theta_P \in \Omega_P$ we can compute the relative entropy of Q given P as a function of $\theta_Q \in \Omega_Q$, in terms of the Lebesgue densities retrieved via FFT. More specifically, we can compute the symmetric relative entropy, which is simply the sum of two relative

entropies with reversed arguments. The quantity that we minimize is the average of symmetric relative entropies, $AS(\theta_Q)$, between P_t and Q_t on a uniform grid for $t \in (0, T]$. We stabilize the minimization of $AS(\theta_Q)$ with a with a quadratic penalty function $\gamma \left\| \theta_Q - \theta_Q^* \right\|^2$, for some $\gamma > 0$. This is also known as Tikhonov regularization, see Engl, Hanke, and Neubauer (1996), pp. 241-43. The idea is that the fixed penalty parameter θ_Q^* is formed by the least squares option implied parameters under each of the five respective affine Markov models. These estimates were made in Chapter 5 for Apple stock. Correspondingly, θ_P was estimated in Chapter 4 for Apple stock, and is fixed. The stabilized minimizer,

$$\widetilde{\theta}_Q = \underset{\theta_Q \in \Omega_Q}{\arg\min} \left[AS(\theta_Q) + \gamma \left\| \theta_Q - \theta_Q^* \right\|^2 \right],$$

is the Parametric Minimum Entropy Martingale Measure (PMEMM).

Estimation for Stochastic Jump Intensity with Leverage

The Leveraged Jump Intensity model (LJI) is introduced in Chapter 2, along with the Barndorff-Nielsen and Shephard (2001) stochastic volatility model (BN-S). The proposed LJI model is compared on the basis of calibration performance to the BN-S and SVSJ Merton models in Chapter 7. The LJI model consists of time changed Lévy jumps, with a Brownian motion in the log-price perfectly negatively correlated with the driver of the continuous stochastic jump intensity, modeled by the CIR process. The LJI model is non-affine, but we show in this thesis that the LJI model appears to provide a better model for the leverage effect than the BN-S model does, due in part to the inherent skewness of the time changed jumps. There are two other well known alternative leverage models from the literature that we mention here, that are not included in this thesis. First, we mention the α-stable Lévy motion with $\beta = -1$, time changed by a an integrated CBI type process

for the intensity, driven by the same time changed α-stable Lévy motion, only with $\beta = 1$, see Carr and Wu (2004), pp. 134-35. The advantage of this model over the LJI model is that it is affine. But, the intensity of this model only takes the CIR process as a special case. Secondly, we mention the "double-jump" model of Duffie, Pan, and Singleton (2000), see pp. 1360-62, with simultaneous Poisson jumps in the log-price and the jump intensity. In the conclusion to Chapter 7 we propose a modified LJI model with simultaneous Poisson jumps similar to the "double-jump" model.

As the LJI model is non-affine, it requires a specialized pricing technique. For this purpose, in Chapter 7 we contribute a new conditional Monte Carlo pricing method called conditional FFT. The method allows us to condition on the jump intensity, yielding an affine conditional log-price characteristic function. Hence, we apply a conditional version of the Carr and Madan (1999) formula. However, instead of taking the mean of the conditional call prices, we take the mean of the respective Fourier transforms of the conditional damped call prices. Thus, conditional FFT only requires one fast Fourier transform to execute. We also contribute proof that the mean in the transform domain converges almost surely to the Fourier transform of the damped call price. Hence, we are able to calibrate the LJI model to the Apple stock spot call options data set from Chapter 5.

Joint Characteristic Functions

The AML method of Bates (2006), which we cover in Chapter 4, is atypical in the sense that for the affine Markov models, it requires the joint CF of the log-price and the latent variance given by

$$\phi_T(u,v) = \exp\left[iuY_{t_0} + \underline{C}(u,v;\tau) + \underline{D}(u,v;\tau)\,\sigma_{t_0}^2\right],$$

rather than simply the marginal log-price CF given by $\phi_T(u,0)$, as we use in Chapter 5 and Chapter 6. In Appendix A we present a set of joint affine

coefficients, $\underline{C}(u, v; \tau)$ and $\underline{D}(u, v; \tau)$, for the five affine Markov models, that are shown to be equivalent to Bates (2006), pp. 953-55, in the SVJ Merton and SVSJ Merton model cases. Then in Appendix B and Appendix C, by specializing and extending the arguments of Lord and Kahl (2010), and Kahl and Jäckel (2005), we contribute proof that our joint affine coefficients for the affine Markov models of this thesis are continuous on the principal branch. To the best of our knowledge, this argument has never before been made for the joint affine coefficients, and in particular this argument has never been made for the SVSJ Merton model. The above three appendices are cited in Section 3.3, and Section 4.3, both on L^1 integrability. They are also cited in Appendix D covering our contributed proof that the primary Fourier transform of AML is differentiable twice under the integral.

Summary of Main Contributions

There are two main contributions of this thesis that we wish to focus on, among the many supporting contributions highlighted above. These are:

1. The Parametric Minimum Entropy Martingale Measure of Chapter 6.

2. The Leveraged Jump Intensity Model (LJI) of Chapter 7.

The PMEMM of Chapter 6 is our original contribution to the solution of the incomplete markets problem for the five affine Markov models of this thesis. Note that the PMEMM requires an estimate of the P-measure, and an estimate of the option implied Q-measure. These are provided by results from Chapter 4 and Chapter 5. Though not equivalent, the PMEMM is similar in nature to the Minimal Entropy Martingale Measure (MEMM) of Frittelli (2000). The practical difference is that the PMEMM is easily computable by fast Fourier transform techniques.

The LJI model of Chapter 7 is our original contribution to the problem of modeling a leverage effect when the log-price contains time changed jumps, and the jump intensity is modeled by the purely continuous CIR process. We have provided a specialized conditional Monte Carlo pricing technique for the LJI model, and we successfully argue that the LJI model outperforms the Barndorff-Nielsen and Shephard (2001) stochastic volatility model on the basis of calibration to spot options prices on Apple stock.

Thesis Outline

The rest of the thesis is structured as follows. Part I: In Chapter 2 we construct, and give properties for, the five affine Markov models, and the LJI model. In Chapter 3 we describe how Fourier inversion via FFT will be carried out in this thesis. Part II: In Chapter 4, we make theoretical contributions to the Bates (2006) method of AML. Then, based on daily log-return data, we estimate the P-measures for Apple stock (1991-2011), and several other assets, under the five affine Markov models, by the AML method. In Chapter 5 we calibrate the option implied Q-measures, under the five affine Markov models, to spot call options prices selected from the daily close on Apple stock, January 19th, 2011, by our modified version of the Carr and Madan (1999) formula. Then in Chapter 6 we combine the results from Chapter 4 and Chapter 5 to construct optimal Q-measures, for each of the five affine Markov models, under our newly proposed Parametric Minimum Entropy Martingale Measure (PMEMM). Part III: In Chapter 7 we calibrate the newly proposed LJI model to the Apple stock spot call options data from Chapter 5 by a specialized method of conditional Monte Carlo, that we refer to as conditional FFT. Lastly, in Chapter 8 we present our least squares rejection technique for the simulation of Meixner process increments. Part IV is the conclusion of the thesis.

Part I

Model Building and Applied Fourier Analysis

Preface to Part I

The goal for Chapter 2 of this part is to lay down the technical foundations for the financial log-price models to be estimated in this thesis. Models from both the affine class, and the non-affine class, are proposed. From the affine class we have adapted two stochastic volatility jump models from Bates (2006), one featuring independent jumps of arbitrary type, and the other having Merton jumps time changed by the integrated variance. From our three proposed independent jump types, the Meixner jumps of Schoutens and Teugels (1998) have, to the best of our knowledge, not yet been seen in the stochastic volatility jump literature. We further propose a new non-affine model for the log-price featuring stochastic jump intensity with a leverage effect. This model uses the CIR process for the jump intensity, and generalizes the stochastic exponential time changed Lévy model of Carr, Geman, Madan, and Yor (2003). The leverage technique was inspired by the Barndorff-Nielsen and Shephard (2001) stochastic volatility model. The goal for Chapter 3 of this part is to establish the fast Fourier transform (FFT) as a refined technique. Fourier inversion plays the defining role in every estimation method proposed in this thesis. Hence, we have devoted a foundational chapter to this topic. Estimation of the affine models is the subject of Part II. Then, since it applies Monte Carlo methods, estimation of our non-affine model is treated in Part III.

Chapter 2

Construction of the Models

2.1 Introduction

In this chapter we develop the models of this thesis under the assumption that the log stock price has both a continuous component, and jumps. This is consistent with the S&P 500 index options price study of Carr and Wu (2003), see p. 2602, as well as the ultra-high frequency log-return study of the DJIA and each of its components found in Aït-Sahalia and Jacod (2012), see p. 1040. Carr and Wu (2003), see p. 2606-2608, also finds that both stochastic volatility and stochastic jump intensity can model the change in the relative sizes of the continuous and jump components over time. We will denote the three main models of this thesis as follows:

- **SVJ**: The Heston stochastic volatility model with independent jumps.
- **SVSJ**: The Heston stochastic volatility model with Merton jumps time changed by the integrated variance.
- **LJI**: Time changed Lévy jumps with a Brownian motion in the log-price perfectly negatively correlated with the driver of the continuous stochastic jump intensity process.

The SVJ model originated in Bates (1996) with Merton jumps. Later, the SVJ model appeared in Huang and Wu (2004), as well as Li, Wells, and Yu (2008), both under Merton, Variance Gamma, and log-stable jumps. Here we will consider the SVJ model under Merton, Variance Gamma, and the Meixner jumps of Schoutens and Teugels (1998). These choices were motivated by the observation in Aït-Sahalia and Jacod (2012), see p. 1037, that individual stock components of the DJIA appear to exhibit a diffusion with infinite-activity jumps. However, Carr and Wu (2003), see p. 2597, warns that infinite-activity jumps with infinite variation can be difficult to distinguish from the continuous component of the model.

The SVSJ Merton model first appeared in Bates (2000), and was the main model introducing Approximate Maximum Likelihood (AML) in Bates (2006). It also appeared in Pan (2002) with Merton jumps, and in Huang and Wu (2004) with Merton, Variance Gamma, and log-stable jumps. We treat the SVSJ model with time changed Merton jumps, and show that this coincides with the affine jump-diffusion representation from Bates (2006), see p. 953, due to the nature of finite-activity compound Poisson jumps.

The Leveraged Jump Intensity (LJI) model is new to the literature and represents one of the main contributions of this thesis. The LJI model is a combination of the time changed stochastic exponential Lévy model of Carr, Geman, Madan, and Yor (2003), see pp. 358-59, with the CIR process as the jump intensity, and the method of modeling a leverage effect taken from the Barndorff-Nielsen and Shephard (2001) stochastic volatility model.

The Huang and Wu (2004) study was based on S&P 500 index options data. They find that finite-activity jumps are inferior to infinite-activity jumps in the SVJ model for index data, see Table III, p. 1425. Under log-return data, we find that this is not true for the DJIA, but that it is true for the S&P 500. However, their results suggest that the SVSJ Merton

model is slightly better than the SVJ Merton model for index data, see Table V, p. 1427, and we agree. The main conclusion of interest to this thesis in the Huang and Wu (2004) study is that stochastic jump intensity, also called jump volatility, has little explanatory power on its own in the absence of a leverage effect for the jump intensity, see p. 1429. As mentioned in Chapter 1 above, Carr and Wu (2004), pp. 134-35, offers a solution to this problem in terms of time changed α-stable processes. But, in this chapter we introduce LJI model which features a more flexible choice of jumps, and models the intensity by the CIR process.

The remainder of this chapter is structured as follows. Section 2.2 covers Lévy processes. In Section 2.3 we present details of the three main jump types of this thesis. Section 2.4 briefly covers affine Markov processes. Then in Section 2.5 we cover the CIR process, the Heston stochastic volatility model, and time changed Lévy processes. In Section 2.6 we formally define the SVJ and SVSJ models and their properties. In Section 2.7 we present the Barndorff-Nielsen and Shephard stochastic volatility model, and the new Leveraged Jump Intensity (LJI) model. Section 2.8 concludes.

2.2 Lévy Processes

2.2.1 Basic Properties

In this subsection we present the main definitions and properties of Lévy processes, as applied in this thesis.

Definition 2.1 (Lévy Process) *A càdlàg stochastic process Y_t is called a Lévy process when*

i) the law of $Y_{t+h} - Y_t$ does not depend on t (stationary increments)

ii) the increments are independent.

(Cont and Tankov (2004a), §3.1, p. 68)

Remark 2.2 (Additive Process) *A stochastic process Y_t which has independent increments, but is not necessarily càdlàg, and does not necessarily have stationary increments, is called an additive process. Thus, every Lévy process is an additive process (see Sato (1999), p. 3).*

In the finance literature, when the assumptions of Definition 2.1 are appropriate, a Lévy process

$$Y_t = Y_0 + \mu t + \sigma W_t + X_t, \tag{2.1}$$

where W_t is Brownian motion, and X_t is a pure jump Lévy process, is used to model the log of a stock price. By Applebaum (2009), p. 123, the continuous part of a Lévy process must be Brownian motion.

Definition 2.3 (Lévy Measure) *Let $\mathbb{R}_0$ denote $\mathbb{R}\setminus\{0\}$. A Borel measure $\nu(dx)$ on $\mathbb{R}_0$ is called a Lévy measure if*

$$\int_{\mathbb{R}_0} \left(1 \wedge x^2\right) \nu(dx) < \infty. \tag{2.2}$$

(Applebaum (2009), §1.2, p. 29)

Remark 2.4 (Lévy Triplet) *For a Lévy process Y_t as in equation (2.1), a result called the Lévy-Itô decomposition, see Cont and Tankov (2004a), pp. 79-82, guarantees the existence of a Lévy triplet (μ, σ, ν) where $\mu \in \mathbb{R}$, $\sigma > 0$, and ν is the Lévy measure for the jumps in Y_t.*

Theorem 2.5 (Lévy-Khinchin Formula) *Let Y_t be a Lévy process with $Y_0 = 0$, and Lévy triplet (μ, σ, ν). Then the characteristic function Y_t, where $\phi_{Y_t}(u) = E\left[e^{iuY_t}\right]$, is given by*

$$\phi_{Y_t}(u) = e^{t\psi_Y(u)}, \text{ with} \tag{2.3}$$

$$\psi_Y(u) = iu\mu - \frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{\infty} \left(e^{iux} - 1 - iux\mathbf{1}_{|x|\leq 1}\right) \nu(dx), \tag{2.4}$$

where $h(x) = \mathbf{1}_{|x|\leq 1}$ is referred to as a truncation function for the small jumps, see Cont and Tankov (2004a), §3.4, p. 83.

Remark 2.6 *The truncation function $h(x)$ in Theorem 2.5 above allows the Lévy measure to have a singularity at $x = 0$. Moreover, any bounded measurable function $h : \mathbb{R} \to \mathbb{R}$ that satisfies $h(x) = 1 + o(|x|)$ as $x \to 0$ and $h(x) = 1 + O\left(\frac{1}{|x|}\right)$ as $x \to \infty$ is admissible as a truncation function, see Cont and Tankov (2004a), §3.4, p. 83. Furthermore, one may replace equation* (2.4) *in the Lévy-Khinchin formula above with*

$$\psi_Y(u) = iu\mu^h - \frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{\infty} \left(e^{iux} - 1 - iuxh(x)\right) \nu(dx)\, h(x), \qquad (2.5)$$

where μ^h is defined relative to the choice of truncation function $h(x)$ and is thus not necessarily unique. However, the other two parameters, σ and ν, of the Lévy triplet $\left(\mu^h, \sigma, \nu\right)$ are uniquely defined for all choices of the truncation function $h(x)$, see Cont and Tankov (2004a), §3.4, p. 83.

Pursuant to the above remark, if the choice $h(x)$ of a suitable truncation function is not otherwise clear, then the P-dynamics of the log-price Y_t under a Lévy process model may be written simply as

$$dY_t = \mu^h dt + \sigma dW_t + dX_t. \qquad (2.6)$$

Proposition 2.7 *Let X_t be a Lévy process. Then*

1. *X_t is Markov with respect to the natural filtration.*

2. *X_t is a semimartingale.*

Proof. *1.* Every additive process is Markov, see Sato (1999), Thm. 10.4, pp. 55-57. Thus, the result follows by Remark 2.2.

2. Every Lévy triplet $\left(\mu^h, \sigma, \nu\right)$ corresponds to a non-stochastic set of characteristics $\left(B^h, C, \nu\right)$ for a semimartingale process with stationary and independent increments. This is given in Barndorff-Nielsen and Shiryaev (2010), see p. 108. ■

2.2.2 Classification of Pure Jump Lévy Processes

For pure jump Lévy processes $\sigma^2 = 0$, and there are three main cases to consider. These are finite-activity jumps, infinite-activity jumps of finite variation, and infinite-activity jumps of infinite variation.

Finite-Activity Jumps

If $\int_{-1}^{1} \nu(dx) < \infty$, then none of the jumps need to be compensated and no truncation of the small jumps in the neighbourhood of zero is required, see Schoutens (2003), p. 45. This means that no truncation function $h(x)$ is required, and the parameter μ is uniquely determined. Moreover, in this case the jump sizes Y_j have a density $g(x)$ that exists. Thus, a finite-activity jump process is given by a compound Poisson process

$$X_t = \sum_{j=1}^{N_t} Y_j, \tag{2.7}$$

where N_t is a Poisson process with constant intensity λ. Also, the Lévy measure for finite-activity jumps is given by

$$\nu(dx) = \lambda g(x)\, dx. \tag{2.8}$$

Aït-Sahalia and Jacod (2012) classify all finite-activity jump processes as having a stability index of zero, see p. 1022.

Infinite-Activity Jumps of Finite Variation

If $\int_{-1}^{1} |x|\, \nu(dx) < \infty$, but we find that $\int_{-1}^{1} \nu(dx) = \infty$, then some small jumps are compensated and truncation of the small jumps is required, see Schoutens (2003), p. 45. In this case the Lévy measure $\nu(dx)$ is not defined at zero, thus there are finite intervals about zero which contain infinitely many small jumps. Therefore, we say that these jumps have infinite-activity.

In the finite variation case the rate at which the density corresponding to the Lévy measure diverges from zero, assuming that this density exists, is slower. Hence, the singularity at zero in $\nu(dx)$ is removable. Moreover, on any interval of finite length, the process can have at most countably many jumps. Consequently, this case is less difficult to distinguish from Brownian motion. Aït-Sahalia and Jacod (2012) classifies the finite variation case as having a stability index strictly less than one, see p. 1027.

Infinite-Activity Jumps of Infinite Variation

If the singularity at zero in the Lévy measure $\nu(dx)$ is not removable, such that we have

$$\int_{|x|\leq 1} |x| \, \nu(dx) = \infty, \tag{2.9}$$

then the jumps have infinite variation, see Schoutens (2003), p. 45. This means that over any interval of finite length, the number of jumps of the process has cardinality of the real line. Hence, this case can be more difficult to distinguish from Brownian motion. Aït-Sahalia and Jacod (2012) classifies the infinite variation case as having a stability index greater than or equal to one, but strictly less than two, see p. 1027.

2.2.3 Exponential Lévy Martingales

The distribution of the log-price $Y_t = \log(S_t)$ is used for estimation of models for the stock price S_t under the objective measure P in Chapter 4. In Chapter 5, Chapter 6, and also Chapter 7, the focus is on the risk-neutral measure Q. For the moment assume that the model under P is given by

$$Y_t = Y_0 + \mu dt + \sigma W_t + X_t, \tag{2.10}$$

where W_t is Brownian motion, and X_t is a pure jump Lévy process. The problem we now face in specifying a model under the Q-measure is that due

to the presence of the jumps under P in equation (2.10) there are infinitely many possible equivalent martingale measures. Indeed the risk-neutral model under Q need not even be a Lévy process, see Cont and Tankov (2004a), pp. 306-307. Therefore, we must choose an equivalent martingale measure Q such that the discounted stock price,

$$e^{-rt}S_t = e^{-rt}\exp\left(Y_t\right), \tag{2.11}$$

is a martingale where r is the risk-free rate, from among the infinite collection of possibilities. In this thesis we choose a structure preserving martingale measure similar to the choice made in Merton (1976). That is, we choose to model the log-price Y_t under the Q-measure as

$$Y_t = Y_0 + \left(r - \frac{1}{2}\sigma^2 - \psi_X\left(-i\right)\right)dt + \sigma W_t + X_t. \tag{2.12}$$

Note that in this thesis we estimate the P and Q measures separately. Hence, we also make the assumption that the respective models under P and Q may have different parameters, outside of any constraints imposed by equivalence. The following three well known results lead to verification that under the log-price model for the Q-measure in equation (2.12) above, the discounted stock price is a martingale.

Proposition 2.8 *Let Y_t be a Lévy process with $Y_0 = 0$ and $\psi_Y\left(-i\right) < \infty$. Then the process*

$$M_t = e^{Y_t - t\psi_Y(-i)} \tag{2.13}$$

is a martingale.

Proof. If $\psi_Y\left(-i\right) < \infty$, then by Theorem 2.5,

$$E\left[M_t|M_0\right] = \frac{E\left[e^{Y_t}\right]}{e^{t\psi_Y(-i)}} = \frac{e^{t\psi_Y(-i)}}{e^{t\psi_Y(-i)}} = 1.$$

Hence, since Y_t has independent increments, the result follows. ■

Theorem 2.9 (Exponential Lévy Martingale) *Consider a Lévy process* $Y_t = \mu t + \sigma W_t + X_t$ *with* $Y_0 = 0$. *Then the process* $\exp(Y_t)$ *is a martingale if and only if* $\psi_Y(-i) < \infty$, *and*

$$\mu = -\frac{1}{2}\sigma^2 - \int_{-\infty}^{\infty} \left(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\right) \nu(dx) \tag{2.14}$$

with $\psi_Y(-i)$ *defined from equation* (2.4) *in the Lévy-Khinchin formula.*

Proof. Let $Y_t = \mu t + \sigma W_t + X_t$ be a Lévy process, and assume that $\exp(Y_t)$ is a martingale. Then by Theorem 2.5, $E\left[e^{Y_t}\right] = \phi_{Y_t}(-i) = e^{t\psi_Y(-i)} = 1$ implies that $\psi_Y(-i) = 0 < \infty$. Moreover, by equation (2.4)

$$\psi_Y(-i) = \mu + \frac{1}{2}\sigma^2 + \int_{-\infty}^{\infty} \left(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\right) \nu(dx) = 0.$$

This implies $\mu = -\frac{1}{2}\sigma^2 - \int_{-\infty}^{\infty} \left(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\right) \nu(dx)$. Now assume that $\psi_Y(-i) < \infty$, and $\mu = -\frac{1}{2}\sigma^2 - \int_{-\infty}^{\infty} \left(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\right) \nu(dx)$. Since Y_t is a Lévy process with $\psi_Y(-i) < \infty$, Proposition 2.8 implies that

$$M_t = e^{Y_t - t\psi_Y(-i)}$$

is a martingale. But then $\mu = -\frac{1}{2}\sigma^2 - \int_{-\infty}^{\infty} \left(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\right) \nu(dx)$ implies that $\psi_Y(-i) = 0$ by equation (2.4), and $M_t = \exp(Y_t)$ is a martingale. ■

Corollary 2.10 *Let the log-price* Y_t *be a Lévy process with volatility* σ, *pure jump component* X_t, *and drift* μ *given by*

$$\mu = r - \frac{1}{2}\sigma^2 - \psi_X(-i), \text{ where} \tag{2.15}$$

$$\psi_X(-i) = \int_{-\infty}^{\infty} \left(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\right) \nu(dx). \tag{2.16}$$

Then if $\psi_Y(-i) < \infty$, *the discounted stock price process under* Q,

$$e^{-rt} S_t = S_0 \exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t + X_t - \psi_X(-i)\, t\right), \tag{2.17}$$

is a martingale. Moreover, the Q*-dynamics of* Y_t *are given by*

$$dY_t = \left(r - \frac{1}{2}\sigma^2 - \psi_X(-i)\right) dt + \sigma dW_t + dX_t. \tag{2.18}$$

Remark 2.11 (Exponential Compensators) *The processes given by* $\frac{1}{2}\sigma^2 t$ *and* $\psi_X(-i)\,t$ *respectively are referred to as the exponential compensators for Brownian motion and the pure jump Lévy process* X_t*.*

(see Jacod and Shiryaev (2003), II.8, p.141)

2.3 Specific Types of Jumps

In this section we consider both the strengths and the weaknesses, and give technical details, of the three jump types that are to be analyzed in this thesis. The finite-activity, infinite-activity with finite variation, and infinite-activity with infinite variation cases will respectively be represented by Merton (1976) jumps in subsection 2.3.1, the Huang and Wu (2004) version of Variance Gamma jumps in subsection 2.3.2, and the Meixner jumps of Schoutens and Teugels (1998) in subsection 2.3.3.

The marginal density for each type of jump process in this section is derived with drift parameter μ. To obtain the marginal density for an exponential Lévy martingale process as in Corollary 2.10, simply set

$$\mu = \mu_0 - \psi_X(-i)\,. \tag{2.19}$$

In brief, the Merton jumps are of finite-activity, thus having a stability index of zero, see Aït-Sahalia and Jacod (2012), p. 1022, the Variance Gamma jumps also happen to have a stability index of zero, see Aït-Sahalia and Jacod (2012), p. 1025, and the Meixner jumps have a stability index of one, see Madan and Yor (2008), p. 43.

Theorem 2.12 (Continuous Density) *Let* $Y_t = \sigma W_t + X_t$ *be a Lévy process. If* $\sigma > 0$ *or* X_t *has infinite-activity, then* Y_t *has a continuous marginal density* $f_t(y)$*.*

(Cont and Tankov (2004a), Prop. 3.12, p. 90)

2.3.1 Merton Jumps

Merton jumps are large and occur rarely. Along with a diffusion component, they are often used for modeling stock indices, in particular market crashes and other wide tailed phenomena. But, they are less good at modeling day to day events such as company announcements that lead to smaller jumps, and Merton jumps must be accompanied by a diffusion to be realistic.

The finite-activity compound Poisson jumps of Merton (1976) are constructed as a random sum given by

$$X_t = \sum_{j=1}^{N_t} \epsilon_j. \tag{2.20}$$

The process N_t is Poisson with intensity $\lambda > 0$, where each ϵ_j is $i.i.d.$ normal $N\left(\beta, \alpha^2\right)$, so the jump size density is given by

$$g(x) = \frac{1}{\alpha\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\beta}{\alpha}\right)^2}. \tag{2.21}$$

Thus, from equation (2.8) the Lévy measure is given by

$$\nu\left(dx\right) = \lambda g\left(x\right) dx = \frac{\lambda}{\alpha\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\beta}{\alpha}\right)^2} dx. \tag{2.22}$$

From Cont and Tankov (2004a), p. 112, the characteristic exponent is

$$\psi_X\left(u\right) = \lambda\left(\exp\left(i\beta u - \frac{1}{2}\alpha^2 u^2\right) - 1\right). \tag{2.23}$$

Hence, the compensator is given by

$$\psi_X\left(-i\right) = \lambda\left(\exp\left(\beta + \frac{1}{2}\alpha^2\right) - 1\right). \tag{2.24}$$

Notice in (2.24) that for Merton jumps the compensator $\psi_X\left(-i\right)$ is finite for all $\lambda > 0$, $\alpha > 0$, and $\beta \in \mathbb{R}$. However, Merton jumps do not have a continuous marginal density (see Cont and Tankov (2004a), p. 90).

The following lemma is used to obtain the representation of a Merton jump process by a Poisson driven integral.

Lemma 2.13 (Compound Poisson Integration Formula) *Let X_t be a compound Poisson process with zero drift and $X_0 = 0$. Then if $f \in C^1$ and $f(0) = 0$ we have*

$$f(X_t) = \int_0^t (f(X_s) - f(X_{s-}))\, dN_s, \tag{2.25}$$

where N_s is a homogeneous Poisson process.

Proof. See Jeanblanc, Yor, and Chesney (2009), p. 484. ■

So consider X_t to be a compound Poisson Merton process as in equation (2.20) above. Then by equation (2.25) of Lemma 2.13 above with f equal to the identity function we have

$$X_t = \int_0^t \epsilon_s dN_s, \tag{2.26}$$

where $\epsilon_s \sim Normal\left(\beta, \alpha^2\right)$, for all $s \in [0,t]$, and N_s is a homogeneous Poisson process with intensity $\lambda > 0$.

2.3.2 Variance Gamma Jumps

Variance Gamma jumps are of infinite-activity, and can model small jumps. But, they have a low activity rate so they can also model less frequent medium to large size jumps. They can be used with or without a diffusion, but have a rough density when Δt is small, see Proposition 3.13, item 2. VG jumps are of finite variation, and decompose into the difference of two independent gamma processes, see Madan, Carr, and Chang (1998), p. 81.

The Variance Gamma jumps may be constructed as the following time changed Brownian motion

$$X_t = \beta G_t + \alpha W_{G_t}, \tag{2.27}$$

where the time change G_t is a gamma process, $\alpha > 0$, and $\beta \in \mathbb{R}$.

Remark 2.14 (Gamma Distribution) *Throughout this thesis we assume that the* $\Gamma(a,b)$ *distribution is parameterized such that it has a characteristic function (C.F.) given by*

$$\phi(u) = \left(\frac{1}{1-ibu}\right)^a. \tag{2.28}$$

This implies that the $\Gamma(a,b)$ *distribution has mean* ab *and variance* ab^2*.*

In the Huang and Wu (2004) version of Variance Gamma that we use in this thesis, the distribution of the gamma process G_t in equation (2.27) is given by

$$G_t \sim \Gamma(\lambda t, 1), \text{ with } \lambda > 0. \tag{2.29}$$

Note that the above implies $E[G_t] = \lambda t$ for Huang and Wu (2004) type Variance Gamma jumps. Moreover, as given in Huang and Wu (2004), see p. 1413, the Lévy measure for this version of the Variance Gamma process may be written as

$$\begin{aligned} \nu(dx) &= \lambda \frac{\exp\left(\frac{-|x|}{A_\pm}\right)}{|x|} dx, \text{ where} \\ A_\pm &= \frac{1}{2}\left(\sqrt{\beta^2 + 2\alpha^2} \pm \beta\right). \end{aligned} \tag{2.30}$$

Also from Huang and Wu (2004), p. 1414, the characteristic exponent is

$$\psi_X(u) = -\lambda \log\left(1 - i\beta u + \frac{1}{2}\alpha^2 u^2\right). \tag{2.31}$$

Hence, the compensator is given by

$$\psi_X(-i) = -\lambda \log\left(1 - \beta - \frac{1}{2}\alpha^2\right). \tag{2.32}$$

Notice in (2.32) that for compensated VG jumps of Huang and Wu type,

$$\psi_X(-i) < \infty \text{ only if } \frac{1}{2}\alpha^2 + \beta < 1, \text{ and } \lambda > 0. \tag{2.33}$$

However, in estimation with financial data the main constraint in (2.33) is almost always met during estimation, because the estimated value of α will be small and β will be either small or negative.

Results for the marginal density of a Brownian motion time changed by a gamma process were given independently in McLeish (1982) *(pp. 90-91)* and Madan, Carr, and Chang (1998) *(pp. 87 & 98)*. Here we prove the result for the Huang and Wu (2004) version of Variance Gamma.

Consider the log-price process Y_t with $\mu > 0$, given by

$$dY_t = \mu dt + dX_t, \tag{2.34}$$

where X_t is from equation (2.27). The following lemma is needed to prove the main result in Theorem 2.16 for the VG density.

Lemma 2.15 *Let $K_\nu(z)$ denote the modified Bessel function of the second kind with order ν. Assume that $\operatorname{Re} a > 0$ and $\operatorname{Re} b > 0$. Then*

$$\int_0^\infty x^{\nu-1} e^{-\frac{a}{x} - bx} dx = 2\left(\frac{a}{b}\right)^{\frac{\nu}{2}} K_\nu\left(2\sqrt{ab}\right). \tag{2.35}$$

(Gradshteyn and Ryzhik (1965), 3.47.9, p. 340)

Theorem 2.16 (VG Density) *The log-price process Y_t for the model (2.34) has a continuous marginal density given by*

$$f_t(y) = \frac{2e^{\frac{\beta(y-\mu t)}{\alpha^2}}}{\alpha\sqrt{2\pi}\Gamma(\lambda t)} \left(\frac{|y-\mu t|}{\delta}\right)^{\lambda t - \frac{1}{2}} K_{\lambda t - \frac{1}{2}}\left(\frac{\delta|y-\mu t|}{\alpha^2}\right), \tag{2.36}$$
$$\delta = \sqrt{2\alpha^2 + \beta^2},\ \alpha > 0,\ \beta \in \mathbb{R},\ \lambda > 0.$$

Proof. By the model (2.34) with jumps defined by (2.27) and (2.29)

$$Y_t | (G_t = \gamma) \sim N\left(\mu t + \beta\gamma, \alpha^2\gamma\right), \text{ with}$$
$$G_t \sim \Gamma(\lambda t, 1). \text{ Therefore,}$$
$$f_t(y) = \frac{1}{\alpha\sqrt{2\pi}\Gamma(\lambda t)} \int_0^\infty \gamma^{\left(\lambda t - \frac{1}{2}\right)-1} e^{-\frac{(y-\mu t-\beta\gamma)^2}{2\alpha^2\gamma} - \gamma} d\gamma. \tag{2.37}$$

Equation (2.36) follows from equation (2.37) by Lemma 2.15 with

$$\nu = \lambda t - \frac{1}{2},\ a = \frac{(y-\mu t)^2}{2\alpha^2},\ \text{and}\ b = \frac{\delta^2}{2\alpha^2}.$$

The marginal density is continuous by Theorem 2.12. ■

2.3.3 Meixner Jumps

Meixner jumps are of infinite-activity with infinite variation, hence they can model very small and rapid movements. In this thesis we take the position that Meixner jumps can be modeled with or without a diffusion. But, since these jumps have infinite variation, they can create identification problems in the presence of a diffusion, see Carr and Wu (2003). For log-return data under a diffusion with stochastic volatility and independent jumps, Meixner jumps outperform Variance Gamma jumps only in unusual circumstances.

The Meixner jumps were first seen in Schoutens and Teugels (1998). They have a closed form representation in terms their Lévy measure which is given by

$$\nu(dx) = \lambda \frac{\exp\left(\frac{\beta x}{\alpha}\right)}{x \sinh\left(\frac{\pi x}{\alpha}\right)} dx, \tag{2.38}$$

where $\lambda > 0$, $|\beta| < \pi$, and $\alpha > 0$ is the scale parameter of the marginal density. The Meixner jumps may also be expressed as a time changed Brownian motion, see Madan and Yor (2008). The Meixner characteristic exponent is given by

$$\psi_X(u) = 2\lambda \log\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cosh\left(\frac{1}{2}(\alpha u - i\beta)\right)}\right), \tag{2.39}$$

see Schoutens and Teugels (1998), p. 346. It follows that the compensator is given by

$$\psi_X(-i) = 2\lambda \log\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cos\left(\frac{\alpha+\beta}{2}\right)}\right). \tag{2.40}$$

Compensated jumps require $\psi_X(-i) < \infty$ in (2.40). The constraints

$$\log\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cos\left(\frac{\alpha+\beta}{2}\right)}\right) < \infty, \text{ for all } \alpha > 0 \text{ and all } -\pi < \beta < \pi, \tag{2.41}$$

further imply that $|\alpha + \beta| < \pi$. However, unlike the Merton and Variance Gamma cases, the Meixner parameter $\beta \in (-\pi, \pi)$ has finite support and the additional constraint

$$|\alpha + \beta| < \pi, \tag{2.42}$$

imposed by the condition $\psi_X(-i) < \infty$, can occasionally be breached during estimation.

The closed form solution for the marginal density of the Meixner process is given in Grigelionis (1999) *(Eq. 4 of Theorem 1, p. 34)*. Below we provide a more detailed version of the proof of this result for completeness.

Consider the Meixner log-price process Y_t with $\mu > 0$, given by

$$dY_t = \mu dt + dX_t, \tag{2.43}$$

where X_t is a Meixner jump process. Lemma 2.17 below is needed to prove Lemma 2.18 that follows which in turn proves the main result Theorem 2.19 for the closed form Meixner density.

Lemma 2.17 *For the complex-valued gamma function*

$$\int_{-\infty}^{+\infty} e^{i\theta z} \left|\Gamma(\lambda + iz)\right|^2 dz = 2\pi\Gamma(2\lambda)\left(\frac{1}{2\cosh\left(\frac{\theta}{2}\right)}\right)^{2\lambda}. \tag{2.44}$$

(Grigelionis (1999), Eq. 5, p. 34)

Lemma 2.18 *Let Z be a Meixner Y_1 random variable with drift $\mu = 0$ and scale parameter $\alpha = 1$. The density of Z is given by*

$$f(z) = \frac{\left(2\cos\left(\frac{\beta}{2}\right)\right)^{2\lambda} e^{\beta z}}{2\pi\Gamma(2\lambda)} \left|\Gamma[\lambda + iz]\right|^2. \tag{2.45}$$

Proof. The C.F. of Z obtained from (2.39) with $\alpha = 1$ is

$$\phi_Z(u) = \left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cosh\left(\frac{1}{2}(u - i\beta)\right)} \right)^{2\lambda}.$$

Thus, by Lemma 2.17 with $\theta = u - i\beta$

$$\begin{aligned}\phi_Z(u) &= \frac{\left(2\cos\left(\frac{\beta}{2}\right)\right)^{2\lambda}}{2\pi\Gamma(2\lambda)} \int_{-\infty}^{+\infty} e^{i\theta z} \left|\Gamma(\lambda + iz)\right|^2 dz \\ &= \int_{-\infty}^{+\infty} e^{iuz} \frac{\left(2\cos\left(\frac{\beta}{2}\right)\right)^{2\lambda} e^{\beta z}}{2\pi\Gamma(2\lambda)} \left|\Gamma(\lambda + iz)\right|^2 dz.\end{aligned}$$

The result for $f(z)$ follows by the definition of a characteristic function. ■

Theorem 2.19 (Meixner Density) *The log-price process Y_t for the model* (2.43) *has a continuous marginal density given by*

$$f_t(y) = \frac{\left(2\cos\left(\frac{\beta}{2}\right)\right)^{2\lambda t} e^{\beta\left(\frac{y-\mu t}{\alpha}\right)}}{2\pi\Gamma(2\lambda t)} \left|\Gamma\left[\lambda t + i\left(\frac{y-\mu t}{\alpha}\right)\right]\right|^2. \tag{2.46}$$

Proof. The density of Y_1 follows from Lemma 2.18 where (μ, α) is a location-scale parameter pair. Secondly, the model (2.43) has C.F.

$$\phi_{Y_t}(u) = e^{iu\mu t}\phi_{X_t}(u)$$

where $\phi_{X_t}(u)$ is obtain from equation (2.39). The main result (2.46) follows. The marginal density is continuous by Theorem 2.12. ■

2.3.4 Jump Moments

In this subsection we summarize our introduction to the three main jump processes of this thesis. In Table 2.1 below, we give the theoretical moments, in the zero drift case, for the Merton, Huang and Wu (2004) type Variance Gamma, and Meixner processes. The moments in Table 2.1 are derived from

the respective moment generating functions of the three Lévy processes. By Theorem 2.5, each MGF is given by

$$M_t(u) = \exp\left(t\psi_X(-iu)\right), \tag{2.47}$$

where the respective characteristic exponents $\psi_X(u)$ were given in equations (2.23), (2.31), and (2.39), above. The algebraic results in Table 2.1 above have been verified with mathematical software, and the Meixner moments are also given in Schoutens (2003), p. 63.

	Merton	**Variance Gamma**	**Meixner**
Mean	$\lambda t\beta$	$\lambda t\beta$	$\lambda t\alpha \tan\left(\frac{\beta}{2}\right)$
Variance	$\lambda t\left(\alpha^2+\beta^2\right)$	$\lambda t\left(\alpha^2+\beta^2\right)$	$\lambda t\frac{\alpha^2}{2}\cos^{-2}\left(\frac{\beta}{2}\right)$
Skewness	$\frac{\beta}{\sqrt{\lambda t}}\frac{\left(3\alpha^2+\beta^2\right)}{\left(\alpha^2+\beta^2\right)^{\frac{3}{2}}}$	$\frac{\beta}{\sqrt{\lambda t}}\frac{\left(3\alpha^2+2\beta^2\right)}{\left(\alpha^2+\beta^2\right)^{\frac{3}{2}}}$	$\sqrt{\frac{2}{\lambda t}}\sin\left(\frac{\beta}{2}\right)$
Excess Kurtosis	$\frac{1}{\lambda t}\left(1+\frac{2\alpha^2\left(\alpha^2+2\beta^2\right)}{\left(\alpha^2+\beta^2\right)^2}\right)$	$\frac{3}{\lambda t}\left(1+\frac{\beta^2\left(2\alpha^2+\beta^2\right)}{\left(\alpha^2+\beta^2\right)^2}\right)$	$\frac{1}{\lambda t}\left(2-\cos\left(\beta\right)\right)$

Table 2.1: Jump Process Moments: Zero Drift.

Notice in Table 2.1 above that for all three jump processes, λt scales the mean and the variance, the skewness has a coefficient proportional to $\frac{1}{\sqrt{\lambda t}}$, and the excess kurtosis has a coefficient proportional to $\frac{1}{\lambda t}$. That is, as λ gets large, the skewness and excess kurtosis become smaller in magnitude, hence the jumps will tend to become more balanced, with less kurtosis.

2.4 Affine Markov Processes

The Heston model to be defined in subsection 2.5.2 below, the SVJ model to be defined in subsection 2.6.1 below, and the SVSJ Merton model with time changed compound Poisson jumps to be treated in subsection 2.6.2 below, are all special cases of the affine Markov processes. This section takes a brief moment to explain precisely what an affine Markov process is.

In this thesis we will view the class of affine processes as a subset of the class of all Markov processes, consistent with Duffie, Filipović, and Schachermayer (2003). We only require two main definitions, thus this section will be brief. The advantage of financial modeling with affine Markov processes is that an affine Markov process has an exponentially affine conditional characteristic function for the terminal state given the initial state that can be known in closed form, subject to a set of regularity conditions. The following definition of an affine Markov process is adapted from Duffie, Filipović, and Schachermayer (2003), see Definition 12.1, p. 1038. We have added the notion of a terminal state for clarity of exposition, and consider only the two dimensional case.

Definition 2.20 (Affine Markov Process) *Let $Z_t = \left(Y_t, \sigma_t^2\right)$ be a time homogeneous Markov process with state space $H = \mathbb{R} \times (0, \infty)$, and initial state $Z_{t_0} = \left(Y_{t_0}, \sigma_{t_0}^2\right) \in H$. For some $T \geq t_0$, let $Z_T \in H$ be the terminal state where we define $\tau = T - t_0$ as the gap time. We say that Z_t is affine if for each $\tau \geq 0$ and $(u, v) \in \mathbb{R}^2$ there exists a set of joint affine coefficients $\underline{C}(u, v; \tau)$, $\underline{D}_0(u, v; \tau)$, and $\underline{D}(u, v; \tau)$ such that*

$$\begin{aligned} \phi_{Z_T|Z_{t_0}}(u, v) &= E\left[e^{iuY_T + iv\sigma_T^2} | \left(\left(Y_{t_0}, \sigma_{t_0}^2\right)\right)\right] \\ &= \exp\left(\underline{C}(u, v; \tau) + \underline{D}_0(u, v; \tau) Y_{t_0} + \underline{D}(u, v; \tau) \sigma_{t_0}^2\right) \end{aligned} \tag{2.48}$$

is exponentially affine in $\left(Y_{t_0}, \sigma_{t_0}^2\right)$.

Remark 2.21 *For the affine models of this thesis where $Z_t = \left(Y_t, \sigma_t^2\right)$ and Y_t models the log stock price, the affine coefficient $\underline{D}_0(u, v; \tau)$ in equation (2.48) of Definition 2.20 above turns out to be trivial. That is, we typically have $\underline{D}_0(u, v; \tau) = iu$. Thus, we write*

$$\phi_{Z_T|Z_{t_0}}(u, v) = \exp\left(iuY_{t_0} + \underline{C}(u, v; \tau) + \underline{D}(u, v; \tau) \sigma_{t_0}^2\right). \tag{2.49}$$

When an affine Markov process Z_t, affine in the sense of Definition 2.20 above, is stochastically continuous, and the right-hand derivative of $\phi_{Z_T|Z_{t_0}}(u,v)$ in equation (2.48) above exists at $(u,v)=(0,0)$, then Z_t is said to be *regular affine*, see Duffie, Filipović, and Schachermayer (2003), p. 990. In this case the affine coefficients $\underline{C}(u,v;\tau)$, $\underline{D}_0(u,v;\tau)$, and $\underline{D}(u,v;\tau)$ are the solution to an affine system of generalized Riccati equations, provided this system has admissible parameters, see Duffie, Filipović, and Schachermayer (2003), pp. 991-92. This is a formalization of the method used to solve for the affine coefficients in Heston (1993), see pp. 340-42.

The second main definition of this section is for the Affine Jump-Diffusion class as adapted from Duffie, Pan, and Singleton (2000), see pp. 1349-50. This class contains many affine stochastic volatility jump models, but only treats Poisson type jumps. However, it treats stochastic jump intensity, and shows that a particular representation of the SVSJ Merton model which we propose in subsection 2.6.2 is an affine Markov process. Once again, we will stick to only two dimensional processes.

Definition 2.22 (Affine Jump-Diffusion) *Fix $\mathcal{F}_t = \sigma\left\{Y_t, \sigma_t^2\right\}$ as the filtration, and let $Z_t = \left(Y_t, \sigma_t^2\right)$ be a Markov process with respect to $\mathcal{F}_t$ on state space $H = \mathbb{R}\times(0,\infty)$ solving the SDE given by*

$$dZ_t = \mu(Z_t)\,dt + \sigma(Z_t)\,dW_t + \nu_s d\widetilde{N}_t \tag{2.50}$$

where W_t is a two dimensional standard Brownian motion, $\mu : H \to \mathbb{R}^2$, $\sigma : H \to \mathbb{R}^{2\times 2}$, ν_s is the random jump size independent of $\widetilde{N}_t$, and $\widetilde{N}_t$ is a non-homogeneous Poisson process with stochastic jump intensity given by

$$\lambda_t = b(Z_t), \tag{2.51}$$

with $a : H \to (0,\infty)$. When the functions μ, $\sigma\sigma^T$, and b are all affine on the state space H, we say that with respect to $\mathcal{F}_t$, the Markov process Z_t is an affine jump-diffusion.

2.5 Latent Factor Models

A *latent factor* refers to a secondary process of the model for the stock price that cannot be directly observed from stock price data. In this thesis we will consider stochastic volatility, and stochastic jump intensity as latent factors. Our main model for the latent factor is the CIR process from Cox, Ingersoll, and Ross (1985) which originated in Feller (1951).

2.5.1 The CIR Process

Definition 2.23 *The CIR process is a diffusion solving the SDE*

$$dv_t = \kappa\left(\eta - v_t\right)dt + \omega\sqrt{v_t}dW_t, \tag{2.52}$$

where W_t is Brownian motion. The solution to the SDE (2.52) *is a deterministically time changed squared Bessel process*

$$v_t = e^{-\kappa t}BESQ^{\delta}\left(\frac{\omega^2}{4\kappa}\left(e^{\kappa t}-1\right)\right) \tag{2.53}$$

with order $\delta = \frac{4\kappa\eta}{\omega^2}$ (see Jeanblanc, Yor, and Chesney (2009), p. 357). The integrated CIR process (ICIR) is given by

$$V_t = \int_0^t v_s ds. \tag{2.54}$$

The CIR parameter vector is $\theta = (v_0, \kappa, \eta, \omega)$. In this thesis we assume that the parameter space for the CIR process is given by

$$\Omega = \left\{\theta \in \mathbb{R}^4 | v_0 > 0,\, \kappa > 0,\, \eta > 0,\, \omega > 0\right\}. \tag{2.55}$$

Proposition 2.24 (CIR Properties) *Assume that v_t is given by* (2.53) *from Definition 2.23.*

1. *If $2\kappa\eta > \omega^2$ with $v_0 > 0$, then for all $t \geq 0$ $\Pr\left[v_t > 0\right] = 1$. This property is known as the Feller condition.*

2. $E[v_t] = \eta + (v_0 - \eta)\, e^{-\kappa t}$.

3. If $\kappa > 0$ *then the process* $\{v_t,\, t \geq 0\}$ *is mean-reverting.*

4. The process $\{v_t,\, t \geq 0\}$ *is a Markov process.*

5. For $t > s$, $v_t | v_s$ *follows a scaled non-central chi-squared distribution.*

6. The limiting distribution is $\overline{v} \sim \Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$ *where* $\overline{v} = \lim_{t\to\infty} v_t$.

Proof. *1.* This follows from equation (2.53) in Definition 2.23 (see Jeanblanc, Yor, and Chesney (2009), p.357). This property was originally proven in Feller (1951), p. 178. Hence, it is called the Feller condition.

2. See Jeanblanc, Yor, and Chesney (2009), p. 360.

3. See Jeanblanc, Yor, and Chesney (2009), p. 361.

4. Since it solves the SDE (2.52), the process $\{v_t,\, t \geq 0\}$ is a Markov diffusion. This class of processes is a particular extension of continuous-time Markov chains to continuous-time Markov processes with a continuous state-space, see McLeish (2005), p.148.

5. Let $F_{\chi'^2_d(\lambda)}(x)$ be the distribution function of a noncentral chi-squared random variable with degrees of freedom d and noncentrality parameter λ. Consider the conditional distribution given by

$$d = \frac{4\kappa\eta}{\omega^2}, \tag{2.56}$$

$$\lambda = \frac{4\kappa e^{-\kappa(t-s)}}{\omega^2\left(1 - e^{-\kappa(t-s)}\right)} v_s, \text{ and} \tag{2.57}$$

$$\Pr[v_t \leq y | v_s] = F_{\chi'^2_d(\lambda)}\left(\frac{4\kappa}{\omega^2\left(1 - e^{-\kappa(t-s)}\right)} y\right). \tag{2.58}$$

Equation (2.58) is the distribution function corresponding to the transition density of the Markov process $\{v_t,\, t \geq 0\}$. Equation (2.58) above, and the transition law itself, are both given in Glasserman (2004), see p. 122.

6. For all $s \geq 0$, as $t \to \infty$, $\lambda \to 0$ in equation (2.57), while the scalar $\frac{4\kappa}{\omega^2\left(1-e^{-\kappa(t-s)}\right)} \to \frac{4\kappa}{\omega^2}$ in equation (2.58). This implies that the scaled non-central chi-squared distribution becomes a scaled chi-squared distribution which is a gamma distribution, see Glasserman (2004), p. 122. The limiting distribution is $\Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$, see McLeish (2005), p.149.

This ends the proof of Proposition 2.24. ■

Proposition 2.25 (Integrated CIR Properties) *Assume that V_t is given by* (2.54) *from Definition 2.23.*

1. $E[V_t] = \eta t + (v_0 - \eta)\frac{\left(1-e^{-\kappa t}\right)}{\kappa}$.
2. $V_t = \int_0^t v_s ds$ *remains finite for all finite* $t \geq 0$.

Proof. *1.* This result follows by integrating the first part of Property *2.* from Proposition 2.24 over $[0, t]$.

2. By Property *1.* clearly $E[V_t] < \infty$ for all finite $t \geq 0$. This implies that $\Pr[V_t = \infty] = 0$ for all finite $t \geq 0$. ■

2.5.2 Stochastic Volatility

The main stochastic volatility model of this thesis is the Heston (1993) model which forms the continuous part of both the SVJ model introduced in subsection 2.6.1 and the SVSJ model introduced in subsection 2.6.2. We define the Heston model under the P-measure in log-price form as

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2\right)dt + \sigma_t dW_t^{(S)} \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)} \\ \rho dt &= E\left[dW_t^{(S)} dW_t^{(V)}\right]. \end{aligned} \tag{2.59}$$

In the Heston model the squared volatility, denoted by σ_t^2 , follows a CIR process, and the correlation coefficient $\rho \in [-1, 1]$ is the leverage parameter. The Heston parameter vector is $\theta = \left(\sigma_0^2, \kappa, \eta, \omega, \rho\right)$. In this thesis we assume that the parameter space for the Heston model is given by

$$\Omega = \left\{\theta \in \mathbb{R}^5 | \sigma_0^2 > 0,\, \kappa > 0,\, \eta > 0,\, \omega > 0,\, -1 \leq \rho \leq 1\right\}. \tag{2.60}$$

Notice also in equation (2.59) above for the Heston P-measure that the drift is $\mu_0 - \frac{1}{2}\sigma_t^2$ rather than the usual $\mu_0 + b\sigma_t^2$ for some additional parameter $b \in \mathbb{R}$. We do this for two reasons. First, according to Aït-Sahalia and Kimmel (2007), see p. 442, the parameter b is not identifiable from log-return and options data combined. Secondly, the log-price in equation (2.59) above has the property that $S_t = S_0 e^{Y_t}$ under the ordinary exponential.

But, the Q-measure is not unique, and this is precisely because of the additional source of randomness under P in equation (2.59) above that is induced by the presence of stochastic volatility as a non-traded asset, see Björk (2009), p. 122. Infinitely many equivalent martingale measures exist. Moreover, the risk-neutral model under Q need not be a Heston process, see for example the Heston minimal entropy martingale measure (MEMM) in Hobson (2004), p. 554. Thus, we must choose arbitrarily an equivalent martingale measure Q such that the discounted stock price, $e^{-rt}S_t = e^{-rt}\exp\left(Y_t\right)$, is a martingale where r is the risk-free rate. In this thesis we choose a structure preserving Heston martingale measure. That is, we choose to model the Heston Q-measure of the form

$$\begin{aligned} dY_t &= \left(r - \frac{1}{2}\sigma_t^2\right)dt + \sigma_t dW_t^{(S)} \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)} \\ \rho dt &= E\left[dW_t^{(S)} dW_t^{(V)}\right]. \end{aligned} \tag{2.61}$$

Note that in this thesis we estimate the P and Q measures separately. Hence, we also make the assumption that the respective models under P and Q may have different parameters. The following result offers direct verification that under the Heston model for the Q-measure in equation (2.61) above, the discounted stock price is a martingale.

Theorem 2.26 *If $\omega^2 < 2\kappa\eta$ then the discounted stock price $e^{-rt}S_t$ for the Heston model in equation* (2.61) *above is a martingale for all $\rho \in [-1, 1]$.*

Proof. See Proposition 5.1, pp. 18, of Bernard, Cui, and McLeish (2014). In the special case $\omega^2 < 2\kappa\eta$ the state space is $(0, \infty)$ and the process does not need to be stopped. Prop. 5.1 of Bernard, Cui, and McLeish (2014) is consistent with Andersen and Piterbarg (2007) Proposition 2.5, p. 34. ■

Proposition 2.27 *The Heston process is affine Markov w.r.t. $\sigma\left\{Y_t, \sigma_t^2\right\}$.*

Proof. The Heston model is a special case, with no jumps, of the Affine Jump-Diffusion model in Definition 2.22 of Section 2.4. ■

Remark 2.28 *Due to our choices for the Heston P and Q measures in equations* (2.59) *and* (2.61) *respectively above, these models have the same characteristic function when $\mu_0 = r$, albeit with different parameters.*

2.5.3 Stochastic Jump Intensity

In this subsection we construct a general notion of stochastic jump intensity, suitable for all Lévy processes, around the stochastic exponential model for time changed Lévy processes from §4.3 of Carr, Geman, Madan, and Yor (2003). Under this model the intensity is uncorrelated with the jumps, hence there is no direct provision for a leverage effect. But, the stochastic jump intensity is assumed to follow a CIR process. In particular, we treat

the martingale aspects of time changed jumps. We show separately in subsection 2.6.2 below that the SVSJ Merton model with finite-activity time changed compound Poisson constitutes a Markov process.

As in section 3, pp. 351-352, of Carr, Geman, Madan, and Yor (2003), let the latent factor λ_t follow a CIR process,

$$d\lambda_t = \kappa\left(\eta - \lambda_t\right)dt + \omega\sqrt{\lambda_t}dW_t. \tag{2.62}$$

The process λ_t is referred to as the instantaneous rate of time change, but for reasons that will become clear later on we call it the stochastic jump intensity. The time change is the integrated CIR process,

$$\Lambda_t = \int_0^t \lambda_s ds. \tag{2.63}$$

The time changed Lévy stochastic exponential model is developed in Carr, Geman, Madan, and Yor (2003) §4.3 as follows. Assume for simplicity that the jumps X_t have a density $g\left(x\right)$, and that no truncation is required. Then we may write

$$X\left(\Lambda_t\right) = \int_0^t \int_{-\infty}^{+\infty} x\mu_\lambda\left(ds, dx\right), \tag{2.64}$$

where $\mu_\lambda\left(ds, dx\right)$ is a Poisson random measure with a random characteristic given by the stochastic jump intensity process λ_s. A predictable compensator is given by

$$\nu_\lambda\left(ds, dx\right) = \lambda_s v\left(dx\right)ds = \lambda g\left(x\right)dx\lambda_s ds. \tag{2.65}$$

Evidently the pure jump process

$$m_t = X\left(\Lambda_t\right) - \int_0^t \int_{-\infty}^{+\infty} x\nu_\lambda\left(ds, dx\right) = x * \left(\mu_\lambda - \nu_\lambda\right) \tag{2.66}$$

is a martingale. Thus, we may define the modified pure jump martingale

$$M_t = m_t + \left(e^x - 1 - x\right) * \left(\mu_\lambda - \nu_\lambda\right) \tag{2.67}$$

$$= \left(e^x - 1\right) * \left(\mu_\lambda - \nu_\lambda\right). \tag{2.68}$$

It is shown in Carr, Geman, Madan, and Yor (2003) that the stochastic exponential of the pure jump martingale M_t in equation (2.67) is given by

$$\mathcal{E}\left(M_t\right) = e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)}. \tag{2.69}$$

On the basis of equation (2.69) we choose the log-price Q-dynamics of the unleveraged stochastic jump intensity model to be

$$\begin{aligned} dY_t &= \left(r - \lambda_t \psi_X\left(-i\right)\right) dt + dX\left(\Lambda_t\right), \\ d\lambda_t &= \kappa\left(\eta - \lambda_t\right) dt + \omega \sigma_t dW_t, \\ d\Lambda_t &= \lambda_t dt,\ \lambda_0 = 1. \end{aligned} \tag{2.70}$$

Note that to make the parameter λ identifiable in the above model, we assume that $\lambda_0 = 1$, see Schoutens (2003), p. 92.

Lemma 2.29 *Let Λ_t be an integrated CIR process, and let $X\left(\Lambda_t\right)$ be a time changed Lévy process null at the origin and without drift. Then*

$$E\left[e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)} | \Lambda_t\right] = 1, \text{ for all finite } t \geq 0. \tag{2.71}$$

Proof.

$$\begin{aligned} E\left[e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)} | \Lambda_t\right] &= \frac{E\left[e^{X(\Lambda_t)} | \Lambda_t\right]}{e^{\Lambda_t \psi_X(-i)}} \\ &= \frac{e^{\Lambda_t \psi_X(-i)}}{e^{\Lambda_t \psi_X(-i)}} \text{ by Theorem 2.5} \\ &= 1 \text{ for all finite } t \geq 0, \end{aligned}$$

by Property 2. of Proposition 2.25. ■

Corollary 2.30 *If $\psi_X\left(-i\right) < \infty$, then the discounted price process $e^{-rt} S_t$ in the time changed Lévy model with independent intensity is a martingale.*

Proof. $E\left[S_0 e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)}\right] = E\left[E\left[S_0 e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)} | \Lambda_t\right]\right] = E\left[S_0\right]$ by Lemma 2.29, and the result is equal to S_0 for all finite $t \geq 0$. ■

2.6 Stochastic Volatility Jump Models

This section presents the SVJ model in subsection 2.6.1, and the SVJ model is treated under all three jump types from Section 2.3. The Bates (2000) SVSJ Merton model is covered in subsection 2.6.2.

2.6.1 The Heston Model with Independent Jumps (SVJ)

By combining the Heston P-measure from equation (2.59) in subsection 2.5.2 independently with compensated Lévy process jumps, we obtain a P-measure for the SVJ model given by

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \psi_X(-i)\right)dt + \sigma_t dW_t^{(S)} + dX_t, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ \rho dt = E\left[dW_t^{(S)} dW_t^{(V)}\right]. \end{aligned} \tag{2.72}$$

However, due to both the presence of stochastic volatility as a latent factor, and the jumps, the Q-measure for the SVJ model is not unique. In fact, like both the Heston model and the Lévy model, the SVJ Q-measure may not even have the same structural form as the SVJ model itself. Since there are infinitely many possible Q-measures for the SVJ model, we must choose arbitrarily an equivalent martingale measure Q such that the discounted stock price, $e^{-rt}S_t = e^{-rt}\exp(Y_t)$, is a martingale where r is the risk-free rate. We choose a structure preserving SVJ Q-measure of the form

$$\begin{aligned} dY_t &= \left(r - \frac{1}{2}\sigma_t^2 - \psi_X(-i)\right)dt + \sigma_t dW_t^{(S)} + dX_t, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ \rho dt = E\left[dW_t^{(S)} dW_t^{(V)}\right]. \end{aligned} \tag{2.73}$$

Recall that in this thesis we estimate the P and Q measures separately. Thus, we make the natural assumption that the respective models under P and Q may have different parameters. The following result verifies that

under the SVJ Q-measure in equation (2.73) above, the discounted stock price is a martingale.

Theorem 2.31 (SVJ Martingale) *Assume that* $\omega^2 < 2\kappa\eta$*, and also that* $\psi_X(-i) < \infty$*. Then the discounted price process* $e^{-rt}S_t$*,* $t \geq 0$*, defined by equation* (2.73) *for SVJ model under the Q-measure, is a martingale.*

Proof. Let $H_t = \sigma\left\{W_t^{(S)}, W_t^{(V)}, X_t\right\}$, and let $s < t$. Then

$$E^Q\left[e^{-rt}S_t|H_s\right] = S_0E^Q\left[e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}\int_0^t \sigma_u^2 du + X_t - t\psi_X(-i)}|H_s\right]$$

$$= S_0E^Q\left[e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}\int_0^t \sigma_u^2 du}|H_s\right] E^Q\left[e^{X_t - t\psi_X(-i)}|H_s\right],$$

by independent components,

$$= S_0e^{\int_0^s \sigma_u dW_u^{(S)} - \frac{1}{2}\int_0^s \sigma_u^2 du}E^Q\left[e^{X_t - t\psi_X(-i)}|H_s\right], \text{ by Theorem 2.26,}$$

$$= S_0e^{\int_0^s \sigma_u dW_u^{(S)} - \frac{1}{2}\int_0^s \sigma_u^2 du}e^{X_s - s\psi_X(-i)}, \text{ by Proposition 2.8,}$$

$$= e^{-rs}S_s, \text{ for } s < t.$$

Hence, $e^{-rt}S_t$, $t \geq 0$ is a martingale under the Q-measure, as required. ■

We estimate the SVJ model with Merton (MJ), Variance Gamma (VG), and Meixner (MX) jumps. These SVJ models will be denoted by SVMJ, SVVG, and SVMX, respectively in what follows. Note that if the SVJ model has infinite-activity jumps, then it does not satisfy Definition 2.22 of Section 2.4 for Affine Jump-Diffusions.

Proposition 2.32 *The SVJ model is affine Markov with respect to* $\sigma\left\{Y_t, \sigma_t^2\right\}$.

Proof. The SVJ model is a Heston model plus an independent Lévy process. The Heston part is Markov w.r.t. $\sigma\left\{Y_t, \sigma_t^2\right\}$ by Proposition 2.27, and the Lévy process in Y_t is Markov by Proposition 2.7, part *1.* Hence, the SVJ model is Markov w.r.t. $\sigma\left\{Y_t, \sigma_t^2\right\}$. But, the joint C.F. for the SVJ model as given in Appendix A is exponentially affine in $\left(Y_{t_0}, \sigma_{t_0}^2\right)$. Hence, by Definition 2.20 of Section 2.4, the SVJ model is affine Markov. ■

Remark 2.33 *Due to our choices for the SVJ model P and Q measures in equations* (2.72) *and* (2.73) *respectively above, these models have the same characteristic function when* $\mu_0 = r$*, albeit with different parameters.*

2.6.2 The Bates Stochastic Jump Intensity Model (SVSJ)

By combining the Heston P-measure from equation (2.59) in subsection 2.5.2 with time changed Merton jumps and electing to compensate the drift in a suitable manner, we obtain a P-measure for the SVSJ Merton model, as a special case of Huang and Wu (2004), see p. 1409, given as follows by

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \sigma_t^2\psi_X\left(-i\right)\right)dt + \sigma_t dW_t^{(S)} + dX\left(V_t\right), \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right] = \rho dt, \\ dV_t &= \sigma_t^2 dt. \end{aligned} \tag{2.74}$$

Once again, due to both the latent factor and the jumps, the equivalent martingale measure for the SVSJ Merton model is not unique, and we must arbitrarily choose a Q-measure for the model such that the discounted stock price, $e^{-rt}S_t = e^{-rt}\exp\left(Y_t\right)$, is a martingale where r is the risk-free rate. We choose a structure preserving Q-measure of the form

$$\begin{aligned} dY_t &= \left(r - \frac{1}{2}\sigma_t^2 - \sigma_t^2\psi_X\left(-i\right)\right)dt + \sigma_t dW_t^{(S)} + dX\left(V_t\right), \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right] = \rho dt, \\ dV_t &= \sigma_t^2 dt. \end{aligned} \tag{2.75}$$

Since we estimate the P and Q measures separately in this thesis, we assume that the respective models under P and Q may have different parameters. The following lemma is used to prove that under the SVSJ Q-measure in equation (2.75) above, the discounted stock price is a martingale.

Lemma 2.34 *Let M_t be the pure jump martingale process of equation* (2.67) *in subsection 2.5.3 such that by equation* (2.69) *in subsection 2.5.3,*

$$\mathcal{E}\left(M_t\right) = e^{X(V_t) - V_t \psi_X(-i)}. \tag{2.76}$$

Assume that $\omega^2 < 2\kappa\eta$, and that $\psi_X(-i) < \infty$, in the SVSJ Merton model of equation (2.75) *above, under the Q-measure. Then*

$$E^Q\left[\mathcal{E}\left(\sigma_t \bullet W_t^{(S)} + M_t\right)\right] = 1. \tag{2.77}$$

Proof. First of all,

$$\begin{aligned}
\mathcal{E}\left(\sigma_t \bullet W_t^{(S)} + M_t\right) &= \mathcal{E}\left(\sigma_t \bullet W_t^{(S)}\right)\mathcal{E}\left(M_t\right), \text{ by orthogonality,} \\
&= e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}V_t} e^{X(V_t) - V_t\psi_X(-i)}, \text{ by equation } (2.76).
\end{aligned}$$

But, by conditioning on V_t under the Q-measure, we obtain

$$\begin{aligned}
&E^Q\left[e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}V_t} e^{X(V_t) - V_t\psi_X(-i)}\right] \\
&= E^Q\left[E^Q\left[e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}V_t} e^{X(V_t) - V_t\psi_X(-i)} | V_t\right]\right] \\
&= E^Q\left[E^Q\left[e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}V_t} | V_t\right] E^Q\left[e^{X(V_t) - V_t\psi_X(-i)} | V_t\right]\right], \\
&\text{by conditional independence,} \\
&= E^Q\left[E^Q\left[e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}V_t} | V_t\right](1)\right], \text{ by Lemma 2.29,} \\
&= E^Q\left[e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}V_t}\right] \\
&= 1, \text{ by Theorem 2.26.}
\end{aligned}$$

Hence, equation (2.77) follows, as required. ■

Theorem 2.35 (SVSJ Martingale) *Assume that $\omega^2 < 2\kappa\eta$, and that $\psi_X(-i) < \infty$. Also, let the process M_t be defined as in Lemma 2.34 above. Then the discounted price process $e^{-rt}S_t$, $t \geq 0$, defined by equation* (2.75) *for the SVSJ Merton model under the Q-measure, is a martingale.*

Proof. Similar to the first part of the proof of Lemma 2.34 above,

$$\begin{aligned}
&e^{-rt}S_t = e^{-rt}S_0e^{Y_t} \\
&= S_0e^{\int_0^t \sigma_u dW_u^{(S)} - \frac{1}{2}V_t} e^{X(V_t) - V_t\psi_X(-i)}, \text{ by the ordinary exponential,} \\
&= S_0\mathcal{E}\left(\sigma_t \bullet W_t^{(S)}\right)\mathcal{E}\left(M_t\right), \text{ by equation (2.76) of Lemma 2.34,} \\
&= S_0\mathcal{E}\left(\sigma_t \bullet W_t^{(S)} + M_t\right), \text{ by orthogonality.}
\end{aligned}$$

But, upon letting $H_t = \sigma\left\{W_t^{(S)}, W_t^{(V)}, X_t\right\}$, it may be observed that the processes $\sigma_t \bullet W_t^{(S)}$ and M_t are both Q-martingales with respect to H_t. Hence, letting $s < t$,

$$\begin{aligned}
E^Q\left[\sigma_t \bullet W_t^{(S)} | H_s\right] &= \sigma_s \bullet W_s^{(S)}, \text{ and} \\
E^Q\left[M_t | H_s\right] &= M_s, \text{ therefore} \\
E^Q\left[\sigma_t \bullet W_t^{(S)} + M_t | H_s\right] &= \sigma_s \bullet W_s^{(S)} + M_s.
\end{aligned}$$

That is, the process $\sigma_t \bullet W_t^{(S)} + M_t$ is a Q-martingale with respect to H_t. But, by Lemma 2.34,

$$E^Q\left[\mathcal{E}\left(\sigma_t \bullet W_t^{(S)} + M_t\right)\right] = 1.$$

Hence, $e^{-rt}S_t$, $t \geq 0$ is a martingale under the Q-measure, as required. ■

To show that the SVSJ Merton model is affine Markov, we will show that it is an Affine Jump-Diffusion by Definition 2.22 of Section 2.4. Recall that by equation (2.26) from subsection 2.3.1 a compound Poisson Merton process X_t may be written as

$$X_t = \int_0^t \epsilon_s dN_s, \tag{2.78}$$

where $\epsilon_s \sim Normal\left(\beta, \alpha^2\right)$, for all $s \in [0, t]$, and N_s is a homogeneous Poisson process with intensity $\lambda > 0$. So let

$$\Lambda_t = \int_0^t \lambda_s ds, \tag{2.79}$$

be a random change of time where λ_s, $s \in [0,t]$, is some nonnegative random process. Then by equation (2.78) above, since on the right hand side only the Poisson process N_s itself is time changed, we obtain

$$X\left(\Lambda_t\right) = \int_0^t \epsilon_s dN\left(\Lambda_s\right), \tag{2.80}$$

where N is a homogeneous Poisson process, and $\epsilon_s \sim Normal\left(\beta, \alpha^2\right)$. But, a time changed homogeneous Poisson process $N\left(\Lambda_s\right)$ may be written as a non-homogenous Poisson process $\widetilde{N}$ with stochastic jump intensity λ_s, see Jeanblanc, Yor, and Chesney (2009), p. 476. Therefore, equation (2.80) above becomes

$$X\left(\Lambda_t\right) = \int_0^t \epsilon_s d\widetilde{N}_s, \tag{2.81}$$

where $\widetilde{N}$ is a non-homogenous Poisson process with stochastic jump intensity λ_s, and $\epsilon_s \sim Normal\left(\beta, \alpha^2\right)$. Based on equation (2.81) the SVSJ model from equation (2.74) above may equivalently be written as

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \sigma_t^2 \psi_X\left(-i\right)\right) dt + \sigma_t dW_t^{(S)} + \epsilon_s d\widetilde{N}_t, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right) dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)} dW_t^{(V)}\right] = \rho dt, \\ \lambda_t &= \lambda\sigma_t^2, \text{ as in Bates (2006), see p. 953.} \end{aligned} \tag{2.82}$$

Proposition 2.36 *The SVSJ model is affine Markov with respect to* $\sigma\left\{Y_t, \sigma_t^2\right\}$.

Proof. The SVSJ model in equation (2.74) is equivalent to equation (2.82) which satisfies Definition 2.22 for an Affine Jump-Diffusion. Thus, the SVSJ model is affine Markov with respect to $\sigma\left\{Y_t, \sigma_t^2\right\}$. Also, the SVSJ joint C.F. as given in Appendix A is exponentially affine in $\left(Y_{t_0}, \sigma_{t_0}^2\right)$. ■

Remark 2.37 *Due to our choices for the SVSJ model P and Q measures in equations* (2.74) *and* (2.75) *respectively above, these models have the same characteristic function when* $\mu_0 = r$*, albeit with different parameters.*

2.7 Alternative Leverage Models

The classical leverage effect is best embraced by the Heston model with $\rho < 0$ in equation (2.59) of subsection 2.5.2 above. When the volatility σ_t rises in the Heston model with $\rho < 0$, the log-price Y_t tends to fall, and this fall is amplified by the larger value of σ_t. Conversely, if σ_t falls then Y_t tends to rise, but this rise is damped by the smaller value of σ_t. The result is a distribution for the log-price that is skewed to the left, see Heston (1993), p. 337. This section covers models with alternative forms of leverage, in particular, the Barndorff-Nielsen and Shephard (2001) stochastic volatility model (BN-S), see subsection 2.7.1. The newly proposed Leveraged Jump Intensity model (LJI) appears in subsection 2.7.2.

2.7.1 The Barndorff-Nielsen and Shephard Model

As adapted from Nicolato and Venardos (2003), p. 454, the structure preserving Q-dynamics of the Barndorff-Nielsen and Shephard (2001) stochastic volatility model are given by

$$dY_t = \left(r - \kappa\psi_U\left(-i\rho;\nu,\delta\right) - \frac{1}{2}\sigma_t^2 \right) dt + \sigma_t dW_t + \rho dU\left(\nu,\delta;\kappa t\right), \quad (2.83)$$

$$d\sigma_t^2 = -\kappa\sigma_t^2 dt + dU\left(\nu,\delta;\kappa t\right), \text{ with } \rho < 0, \quad (2.84)$$

where $\psi_U\left(-i\rho;\nu,\delta\right) = \log E\left[e^{\rho U(\nu,\delta:1)}\right]$ is the compensator for the positive Lévy process $U_{\kappa t} = U\left(\nu,\delta;\kappa t\right)$, and the leverage parameter $\rho < 0$ gives the jumps a negative coefficient in the log-price equation (2.83). The variance in equation (2.84) is a non-Gaussian Ornstein-Uhlenbeck (OU) process. The joint model in equations (2.83) and (2.84) is commonly referred to as the BN-S model. In this thesis we examine the stationary gamma version of the BN-S. In this case there is a positive Lévy process $U_{\kappa t}$ such that the OU equation (2.84) has a $\Gamma\left(\nu,\delta\right)$ stationary distribution with shape parameter $\nu > 0$, see Barndorff-Nielsen and Shephard (2001), pp. 171-172.

The positive Lévy process $U_{\kappa t}$ driving the variance in the BN-S model is uncorrelated with the Brownian motion W_t in the log-price Y_t. Thus, the leverage effect depends entirely on $\rho U_{\kappa t}$ in the log-price Y_t with $\rho < 0$. This is effective for large values of σ_t, where $U_{\kappa t}$ will be large, and any fall in Y_t will be amplified. But on the opposite side, when σ_t is small, $U_{\kappa t}$ will also be small, so that the remainder of the log-price Y_t is determined by an uncorrelated diffusion which has a symmetric effect. The net effect is not necessarily a damped rise in Y_t. Thus, while the non-Gaussian OU process for the BN-S variance is good for modeling volatility spikes, the BN-S model is less than ideal for modeling the leverage effect.

2.7.2 The Leveraged Jump Intensity Model

In this subsection we introduce the Leveraged Jump Intensity model which improves upon the leverage structure of the Barndorff-Nielsen and Shephard (2001) stochastic volatility model. In the SVJ and SVSJ models, leverage comes from the Heston component. Moreover, Carr and Wu (2004), pp. 130-131 states that if the jump intensity is continuous and the log-price is purely discontinuous then Heston style correlation and leverage is not possible. The Leveraged Jump Intensity model (LJI) that we propose has time changed jumps in the log-price, and uses a continuous CIR process for the jump intensity. Similar to the BN-S model, we obtain leverage in the LJI model by placing the driver of the jump intensity process in the log-price with a negative coefficient. The proposed Q-dynamics of the LJI model are given by

$$dY_t = \left(r - \lambda_t \psi_X(-i) - \frac{1}{2}\gamma^2\right) dt + dX(\Lambda_t) - \gamma dW_t, \tag{2.85}$$

$$d\lambda_t = \kappa(\eta - \lambda_t)\, dt + \omega\sqrt{\lambda_t} dW_t, \tag{2.86}$$

$$d\Lambda_t = \lambda_t dt \ , \ \lambda_0 = 1, \text{ with } \gamma > 0. \tag{2.87}$$

The process X_t in the LJI model log-price equation (2.85) above is any pure jump Lévy process whose characteristic exponent $\psi_X(u)$ is known in closed form. The Lévy process X_t in the LJI log-price is also time changed by the integrated intensity process Λ_t. The intensity process λ_t in equation (2.86) above is modeled by the CIR process. The driver of the CIR intensity is taken to be the Brownian motion W_t, and the reflection of W_t also appears in the log-price equation. This is the basis of the LJI leverage effect.

The LJI leverage effect improves on the BN-S model for two main reasons. First, W_t in the log-price of the LJI takes on both positive and negative values, whereas $U_{\kappa t}$ in the log-price of the BN-S takes positive values only. To make the second point more clearly, we define the discrete increment of the time changed jumps to be

$$\begin{aligned}\Delta X(\Lambda_t) &\triangleq X(\Lambda_{t+1}) - X(\Lambda_t) \\ &\overset{law}{=} X(\Lambda_{t+1} - \Lambda_t) \\ &\approx X\left(\sum\nolimits_{j=0}^{t} \lambda_j \Delta t - \sum\nolimits_{j=0}^{t-1} \lambda_j \Delta t\right) \\ &\overset{law}{=} X(\lambda_t \Delta t), \end{aligned} \tag{2.88}$$

leading to the above approximation based on the discrete time intensity process. Note that the underlying jumps X_t in the LJI model have inherent skewness which will be negative. Hence, as the intensity λ_t gets large this inherent skewness will dissipate, and as λ_t gets small this inherent skewness will grow in magnitude, see Table 2.1 in subsection 2.3.4. Meanwhile the diffusion in the BN-S is always symmetric for all values of σ_t^2. Therefore, for the LJI model when λ_t is small, $-\gamma W_t$ is clearly positive and the log-price Y_t tends to rise. But, this rise is damped by the time changed jump increment with strongly negative skewness. Conversely, when λ_t is large, $-\gamma W_t$ is clearly negative and the log-price Y_t will tend to fall. Moreover, the jumps are more neutral in this case. Thus, they tend to leave the fall in

the log-price as is. Altogether, the LJI model appears to be slightly better than the BN-S model when it comes to the leverage effect.

The LJI model C.F. depends on the joint distribution of (Λ_t, W_t) about which nothing is known. Thus, we say it is non-affine. The following lemma is used to prove that the discounted stock price is a martingale under the Q-measure in the LJI model.

Lemma 2.38 *Let M_t be the pure jump martingale process of equation* (2.67) *in subsection 2.5.3 such that by equation* (2.69) *in subsection 2.5.3,*

$$\mathcal{E}(M_t) = e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)}. \tag{2.89}$$

Assume that $\omega^2 < 2\kappa\eta$, and that $\psi_X(-i) < \infty$, in the LJI model of equations (2.85) *to* (2.87) *above, under the Q-measure. Then*

$$E^Q[\mathcal{E}(M_t - \gamma W_t)] = 1. \tag{2.90}$$

Proof. First note that,

$$\begin{aligned}\mathcal{E}(M_t - \gamma W_t) &= \mathcal{E}(M_t)\,\mathcal{E}(-\gamma W_t), \text{ by orthogonality,} \\ &= e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)} e^{-\gamma W_t - \frac{1}{2}\gamma^2 t}, \text{ by equation } (2.89).\end{aligned}$$

Then, by conditioning on W_t under the Q-measure, we obtain

$$\begin{aligned}&E^Q\left[e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)} e^{-\gamma W_t - \frac{1}{2}\gamma^2 t}\right] \\ &= E^Q\left[E^Q\left[e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)} e^{-\gamma W_t - \frac{1}{2}\gamma^2 t} | W_t\right]\right] \\ &E^Q\left[E^Q\left[e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)} | W_t\right] e^{-\gamma W_t - \frac{1}{2}\gamma^2 t}\right] \\ &= E^Q\left[(1)\, e^{-\gamma W_t - \frac{1}{2}\gamma^2 t}\right], \text{ by Lemma 2.29 since } \Lambda_t \in \sigma(W_t), \\ &= 1, \text{ since } e^{-\gamma W_t} \text{ is log-normal with mean } e^{\frac{1}{2}\gamma^2 t}.\end{aligned}$$

Hence, equation (2.90) follows, as required. ■

Theorem 2.39 (LJI Martingale) *Assume that $\omega^2 < 2\kappa\eta$, and also that $\psi_X(-i) < \infty$. Moreover, let the process M_t be defined as in Lemma 2.38 above. Then the discounted price process $e^{-rt}S_t$, $t \geq 0$, defined by equations* (2.85) *to* (2.87) *for the LJI model under the Q-measure, is a martingale.*

Proof. By reversing the first part the proof of Lemma 2.38,

$$\begin{aligned}
&e^{-rt}S_t = e^{-rt}S_0 e^{Y_t} \\
&= S_0 e^{X(\Lambda_t) - \Lambda_t \psi_X(-i)} e^{-\gamma W_t - \frac{1}{2}\gamma^2 t}, \text{ by the ordinary exponential,} \\
&= S_0 \mathcal{E}(M_t)\, \mathcal{E}(-\gamma W_t), \text{ by equation (2.89) of Lemma 2.38,} \\
&= S_0 \mathcal{E}(M_t - \gamma W_t), \text{ by orthogonality.}
\end{aligned}$$

However, if we let $H_t = \sigma\{X_t, W_t\}$, then it may be observed that the two processes M_t and $-\gamma W_t$ are both martingales with respect to H_t, under the Q-measure . Thus, if we let $s < t$, then

$$\begin{aligned}
E^Q[M_t | H_s] &= M_s, \text{ and} \\
E^Q[-\gamma W_t | H_s] &= -\gamma W_s, \text{ therefore} \\
E^Q[M_t - \gamma W_t] &= M_s - \gamma W_s.
\end{aligned}$$

That is, the process $M_t - \gamma W_t$ is a martingale with respect to H_t, under the Q-measure. But, by Lemma 2.38,

$$E^Q[\mathcal{E}(M_t - \gamma W_t)] = 1.$$

Hence, $e^{-rt}S_t$, $t \geq 0$ is a martingale under the Q-measure, as required. ■

2.8 Conclusion

1. The Heston, SVJ, and SVSJ Merton models are all *affine Markov* under both the P-measure, and also the structure preserving choice

of Q-measure, as outlined above in this chapter. This means that each of these models may be estimated by the FFT based methods of Chapter 4, Chapter 5, and Chapter 6, respectively in Part II.

2. The discounted price process is a martingale under the Q-measure for each of the Heston, SVJ, SVSJ Merton, and LJI models. This is of particular importance for the Heston, SVJ, and SVSJ Merton models in Chapter 6 of Part II where the parametric minimum entropy martingale measure (PMEMM) is introduced.

3. The time changed stochastic exponential Lévy model of Carr, Geman, Madan, and Yor (2003), the Barndorff-Nielsen and Shephard (2001) stochastic volatility (BN-S) model, and the newly proposed Leveraged Jump Intensity (LJI) model have all been introduced. The LJI model appears to provide a better treatment of the leverage effect than the BN-S model does. These models will be compared on the basis of options price calibration performance in Chapter 7 of Part III where a new FFT based conditional Monte Carlo options price method will be introduced to treat the non-affine LJI model.

4. The infinite-activity with infinite variation Meixner Lévy process of Schoutens and Teugels (1998) has been introduced. A new acceptance-rejection method for simulating the increments of this process will be presented in Chapter 8 of Part III.

Chapter 3

Fourier Analysis and the Fast Fourier Transform

3.1 Introduction

Every estimation method in this thesis uses Fourier inversion via FFT as a means to an end, and this chapter provides the technical background. We insist that both the Fourier transform and its inverse transform be L^1 to match the hypotheses of the Fourier inversion theorem. While these conditions are sometimes met by damping, in the absence of any special techniques, L^1 integrability must be proven. An exception occurs when the inverse transform is a density. Densities are L^1 by definition. In this case the Fourier transform is a characteristic function, and it must be proven that the C.F. is L^1. When the C.F. is L^1, the corresponding density is continuous on the real line. Regarding quadrature in the FFT, main line theory from Carr and Madan (1999), see p. 68, advocates Simpson's rule. We diverge here, and propose the trapezoidal rule in place of Simpson's rule, see for example Briggs and Henson (1995), pp. 358-360. We find that Simpson's

rule leads to sudden negativity in the deep tails of the inverse transform, specifically in the non-centred and shifted version of FFT as proposed by Carr and Madan (1999), see pp. 67-68. We propose to keep the non-centred and shifted version of FFT, but use it with the trapezoidal rule instead of Simpson's rule. We find, by way of example, that if the inverse transform has only semi-heavy tails, then this change does not appear to actually lose much accuracy on the interior of the domain of the inverse transform.

The remainder of this chapter is organized as follows. In Section 3.2 we cover Fourier inversion and its approximation via FFT. In Section 3.3 we prove that the marginal log-price C.F. is L^1 integrable in both the SVJ and SVSJ models. Then in Section 3.4 we advocate in favour of the trapezoidal rule as the choice of quadrature method in the non-centred and shifted FFT integral approximation. Section 3.5 briefly concludes.

3.2 Transform Theory

3.2.1 Fourier Transforms

In this thesis we use the Fourier transform on $L^1 = L^1(\mathbb{R})$ where

$$L^1 = \left\{ f | \int_{-\infty}^{+\infty} |f(x)| \, dx < \infty \right\}. \tag{3.1}$$

We will denote the Fourier transform by $\widehat{f}(u)$, under the following definition.

Definition 3.1 (Fourier Transform) *Let $f : \mathbb{R} \to \mathbb{C}$ be a function. If $f \in L^1$ then the improper integral*

$$\widehat{f}(u) = \int_{-\infty}^{+\infty} e^{iux} f(x) \, dx < \infty \tag{3.2}$$

is well defined for all $u \in \mathbb{R}$. We refer to $\widehat{f}(u)$ as the Fourier transform of the function $f(x)$, see Rudin (1987), §9.1, p. 178.

Remark 3.2 *If $f \in L^1$, then equation* (3.2) *in Definition 3.1 above defines $\widehat{f}(u)$ unambiguously for all $u \in \mathbb{R}$. However, if $f \in L^2$ and not L^1, then equality holds in equation* (3.2) *of Definition 3.1 above only in the almost everywhere sense, see Rudin (1987), p. 188.*

Lemma 3.3 *If $f \in L^1$, then the Fourier transform $\widehat{f}(u)$ in equation* (3.2) *is uniformly continuous on $\mathbb{R}$.*

Proof. From equation (3.2)

$$\left|\widehat{f}(u+h) - \widehat{f}(u)\right| \leq \int_{-\infty}^{+\infty} \left|\left(e^{ihx} - 1\right) f(x)\right| dx. \tag{3.3}$$

But, $\left|\left(e^{ihx} - 1\right) f(x)\right| \leq 2\left|f(x)\right| < \infty$ since $f \in L^1$. Moreover,

$$\left|\left(e^{ihx} - 1\right) f(x)\right| \to 0 \text{ as } h \to 0 \text{ for all } x \in \mathbb{R}.$$

Therefore, by the dominated convergence theorem

$$\int_{-\infty}^{+\infty} \left|\left(e^{ihx} - 1\right) f(x)\right| dx \to 0 \text{ as } h \to 0. \tag{3.4}$$

The inequality (3.3) and the limit (3.4) imply that $\widehat{f}(u)$ is continuous. Moreover, the right hand side of (3.3) does not depend on u. Thus, $\widehat{f}(u)$ is uniformly continuous, see Kawata (1972), §2.7, pp. 62-63. ■

Remark 3.4 (Rudin (1987), §9.4, p. 181) *Similar to Definition 3.1 above, if $\widehat{f} \in L^1$, we may assert that*

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux} \widehat{f}(u)\, du, \tag{3.5}$$

for some well defined function $g(x)$. The fact that $f(x) = g(x)$ precisely when f and $\widehat{f}$ are both L^1 is the Fourier inversion theorem which we will come to momentarily.

Selected properties of the Fourier transform on L^1 that are used in this thesis are summarized as follows.

Proposition 3.5 (Fourier Transform Properties) *Let $f \in L^1$. Then*

1. $\widehat{f}(0) = \int_{-\infty}^{+\infty} f(x)\,dx \leq \int_{-\infty}^{+\infty} |f(x)|\,dx < \infty$.
2. $\left|\widehat{f}(u)\right| \leq \int_{-\infty}^{+\infty} |f(x)|\,dx < \infty$ *for all* $u \in \mathbb{R}$.
3. $\widehat{f}(u)$ *is uniformly continuous on* $\mathbb{R}$.
4. *If* f *is also continuous, then* $\widehat{f}(u) \to 0$ *as* $|u| \to \infty$.
5. *If* f *is purely real then* $\overline{\widehat{f}(u)} = \widehat{f}(-u)$.
6. *If* f *is purely complex then* $\overline{\widehat{f}(u)} = -\widehat{f}(-u)$.

Proof. *1.* and *2.* follow from equation (3.2) since $f \in L^1$.

3. This is Lemma 3.3 above.

4. This is the Riemann-Lebesgue lemma (see Körner (1989), p. 263).

5. By Definition 3.1 of the Fourier transform,

$$\overline{\widehat{f}(u)} = \int_{-\infty}^{+\infty} \overline{e^{iux} f(x)} dx = \int_{-\infty}^{+\infty} e^{-iux} \overline{f(x)} dx = \widehat{f}(-u),$$

since f is purely real.

6. By Definition 3.1 again,

$$\overline{\widehat{f}(u)} = \int_{-\infty}^{+\infty} \overline{e^{iux} f(x)} dx = \int_{-\infty}^{+\infty} e^{-iux} \overline{f(x)} dx = -\widehat{f}(-u),$$

since f is purely complex. ■

The Fourier inversion theorem is one of the most important results in classical analysis. The version of the theorem that we have selected, Rudin (1987), Theorem 9.11, p. 185, takes the continuity of the underlying function f as a conclusion. Other versions, for example Körner (1989), Theorem 60.1, p. 296, take the continuity of f as a hypothesis. The latter is often impractical when f is known only through its Fourier transform. This has typically been the case in mathematical finance since Heston (1993).

Theorem 3.6 (Fourier Inversion Theorem) *Let the functions f and $\widehat{f}$ both be L^1. Then*

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iux}\widehat{f}(u)\,du \quad a.e. \tag{3.6}$$

Moreover, $f(x)$ is continuous for all $x \in \mathbb{R}$, see Rudin (1987), §9.4, p.185.

Corollary 3.7 *1. If f is purely real then for all $x \in \mathbb{R}$*

$$f(x) = \frac{1}{\pi}\operatorname{Re}\int_0^{\infty} e^{-iux}\widehat{f}(u)\,du \tag{3.7}$$

$$= \frac{1}{\pi}\int_0^{\infty}\left[\cos(ux)\operatorname{Re}\widehat{f}(u) + \sin(ux)\operatorname{Im}\widehat{f}(u)\right]du. \tag{3.8}$$

2. If f is purely complex then for all $x \in \mathbb{R}$

$$f(x) = \frac{i}{\pi}\operatorname{Im}\int_0^{\infty} e^{-iux}\widehat{f}(u)\,du \tag{3.9}$$

$$= \frac{i}{\pi}\int_0^{\infty}\left[\cos(ux)\operatorname{Im}\widehat{f}(u) - \sin(ux)\operatorname{Re}\widehat{f}(u)\right]du. \tag{3.10}$$

Proof. *1.* Equation (3.7) is obtained from Theorem 3.6 as follows.

$$f(x) = \frac{1}{2\pi}\left[\int_{-\infty}^{0} e^{-iux}\widehat{f}(u)\,du + \int_0^{\infty} e^{-iux}\widehat{f}(u)\,du\right]$$

$$= \frac{1}{2\pi}\left[\int_0^{\infty} e^{iux}\widehat{f}(-u)\,du + \int_0^{\infty} e^{-iux}\widehat{f}(u)\,du\right]$$

$$= \frac{1}{2\pi}\left[\int_0^{\infty}\overline{e^{-iux}\widehat{f}(u)}du + \int_0^{\infty} e^{-iux}\widehat{f}(u)\,du\right], \text{ by } 5. \text{ in Proposition 3.5,}$$

$$= \frac{1}{\pi}\operatorname{Re}\int_0^{\infty} e^{-iux}\widehat{f}(u)\,du\ , \text{ by the identity } z + \overline{z} = 2\operatorname{Re}(z).$$

Equation (3.8) follows from $\operatorname{Re}(z_1 z_2) = \operatorname{Re} z_1 \operatorname{Re} z_2 - \operatorname{Im} z_1 \operatorname{Im} z_2$.

2. Is similar to the above using *6.* from Proposition 3.5, $z - \overline{z} = 2i\operatorname{Im}(z)$, and the property $\operatorname{Im}(z_1 z_2) = \operatorname{Re} z_1 \operatorname{Im} z_2 + \operatorname{Im} z_1 \operatorname{Re} z_2$. ■

The next result is the last of this subsection on Fourier transforms. It provides the key to the proof, see Appendix D of this thesis, that we may differentiate twice under the Fourier integral in the Bates (2006) Approximate Maximum Likelihood method of Chapter 4.

Lemma 3.8 (Uniform Convergence) *Assume that* $f\ , \widehat{f} \in L^1$. *Then the improper integral in equation* (3.6) *of Theorem 3.6 converges uniformly and absolutely for each* $x \in \mathbb{R}$.

Proof. Since $f \in L^1$, Lemma 3.3 implies that $\widehat{f}(u)$ is uniformly continuous. Moreover, since both f and $\widehat{f}$ are L^1, Theorem 3.6 implies that $f(x)$ is continuous. Hence, from Körner (1989), pp. 296-97,

$$\begin{aligned}
&\left|\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iux}\widehat{f}(u)\,du - \frac{1}{2\pi}\int_{-R}^{S} e^{-iux}\widehat{f}(u)\,du\right| \\
&= \frac{1}{2\pi}\left|\int_{u>S \text{ or } u<-R} e^{-iux}\widehat{f}(u)\,du\right| \leq \frac{1}{2\pi}\int_{u>S \text{ or } u<-R}\left|e^{-iux}\widehat{f}(u)\right| du \\
&= \frac{1}{2\pi}\int_{u>S \text{ or } u<-R}\left|\widehat{f}(u)\right| du \to 0 \text{ as } R,\, S \to +\infty,
\end{aligned}$$

and, since $\widehat{f}(u)$ is uniformly continuous, this implies that the improper integral

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iux}\widehat{f}(u)\,du,$$

converges uniformly and absolutely for each $x \in \mathbb{R}$, as required. ■

3.2.2 Characteristic Functions

Definition 3.9 (Characteristic Function) *Let* X *be a random variable with distribution function* F. *Then*

$$\phi(u) = E\left[e^{iuX}\right] = \int_{-\infty}^{+\infty} e^{iux}dF\ ,\ \ u \in \mathbb{R} \tag{3.11}$$

is the characteristic function of X. *(Stuart and Ord (1994), Vol. 1, p. 80)*

The characteristic function exists and inverts for any random variable.

Theorem 3.10 (Characteristic Inversion) *If* $\phi(u)$ *is given by* (3.11) *then*

$$F(x) - F(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1 - e^{-iux}}{iu}\phi(u)\,du, \text{ and}$$

if F is everywhere continuous with $dF = f(x)\,dx$ then

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iux}\phi(u)\,du. \tag{3.12}$$

(Stuart and Ord (1994), Vol. 1, §4.1, pp. 125-126)

When X is a continuous random variable on the real line, F is everywhere continuous with $dF = f(x)\,dx$. Thus, by equation (3.11) of Definition 3.9, the characteristic function of a continuous random variable on the real line is a special case of the Fourier transform, and equation (3.12) of Theorem 3.10 may be viewed as the inverse Fourier transform.

Proposition 3.11 (Characteristic Function Properties) *Let X be any random variable. Then*

1. $\phi(0) = 1$.
2. $|\phi(u)| \leq 1$ *for all* $u \in \mathbb{R}$.
3. $\phi(u)$ *is uniformly continuous on* $\mathbb{R}$.
4. *If X is continuous on the real line, then* $\phi(u) \to 0$ *as* $|u| \to \infty$.
5. $\overline{\phi(u)} = \phi(-u)$.

Proof. The proof is similar to the proof of Proposition 3.5. ■

3.2.3 The Discrete Fourier Transform

Definition 3.12 (Standard DFT) *Consider a pair of sequences $\{f_j\}_{j=1}^N$ and $\{F_k\}_{k=1}^N$ for positive integer N. The standard N-point discrete Fourier transform is given by*

$$F_k = \sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} f_j. \tag{3.13}$$

(Walker (1996), §2.1, p. 36)

In this thesis we use the standard discrete Fourier transform (DFT) to approximate the integral function

$$g(x) = \int_0^\infty e^{-iux} \widehat{f}(u)\, du \tag{3.14}$$

along the left-endpoint discretization of $x \in [-b, b]$ given by

$$x_k = -b + (k-1)\Delta x \ , \ \ k = 1, 2, ..., N \ , \ \text{ where } b = \frac{N}{2}\Delta x. \tag{3.15}$$

This leads to an approximation for the integral in the inverse transform given by either equation (3.7) or (3.9).

Under the standard DFT, this choice of x_k leads to a non-centred and shifted discretization scheme as in Carr and Madan (1999) pp. 67-69. The domain of integration is discretized by

$$u_j = (j-1)\Delta u \ , \ \ j = 1, 2, ..., N \tag{3.16}$$

intending a left-endpoint Riemann approximation to (3.14) for each x_k. The approximate domain of integration becomes

$$[0, A], \ \text{ where } \ A = N\Delta u. \tag{3.17}$$

Nyquist optimal sampling (see Fang and Oosterlee (2008), p. 828) specifies

$$\Delta x \Delta u = \frac{2\pi}{N}. \tag{3.18}$$

The DFT also requires the sampling width Δu to be fixed. Carr and Madan (1999) p. 69 recommends $\Delta u = \frac{1}{4}$. We use this value throughout this thesis. Under Nyquist sampling this choice of Δu implies

$$\Delta x = \frac{2\pi}{A} = \frac{8\pi}{N} \ , \tag{3.19}$$

$$\text{and } \ b = \frac{N}{2}\Delta x = \frac{\pi}{\Delta u} = 4\pi. \tag{3.20}$$

The two main choices of N are $N = 2^{14}$ for daily log-return precision (see Carr, Geman, Madan, and Yor (2002), p. 320) and $N = 2^{12}$ for high-speed options price calibration (see Carr and Madan (1999), p. 69).

For each x_k the Riemann approximation with quadrature weights w_j for the integral in equation (3.14) is given by

$$\begin{aligned} g(x_k) &\approx \int_0^A e^{-iux_k} \widehat{f}(u)\, du \\ &\approx \sum\nolimits_{j=1}^{N} e^{-iu_j x_k} \widehat{f}(u_j)\, w_j \Delta u, \quad k = 1, 2, ..., N. \end{aligned} \tag{3.21}$$

To suit the standard DFT, Carr and Madan (1999) p. 68 shows that

$$\begin{aligned} g(x_k) &\approx \sum\nolimits_{j=1}^{N} e^{-iu_j x_k} \widehat{f}(u_j)\, w_j \Delta u \\ &= \sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \widehat{f}(u_j)\, w_j \Delta u, \\ &\quad \text{for each } x_k,\ k = 1, 2, ..., N. \end{aligned} \tag{3.22}$$

For the w_j Carr and Madan (1999) p. 68 proposes Simpson's rule. We propose a different quadrature rule in subsection 3.4.1.

3.2.4 The Fast Fourier Transform

The fast Fourier transform (FFT) is an algorithm for the DFT that was proposed in Cooley and Tukey (1965). The complexity of the FFT algorithm is $O(N \log_2 N)$ when $N = 2^p$, whereas the DFT has complexity $O\left(N^2\right)$. The FFT is unbeatable when every element of the transform is in use.

For simplicity, first assume that the inverse transform $f(x)$ is purely real. This is the case when $\widehat{f}(u)$ is a characteristic function, or also if $\widehat{f}(u)$ is a damped call price. The latter leads to the Carr and Madan (1999) formula, which will be introduced in Chapter 5. If $f(x)$ is purely real, then by equation (3.7) of Corollary 3.7,

$$f(x) = \frac{1}{\pi} \operatorname{Re} \int_0^\infty e^{-iux} \widehat{f}(u)\, du, \text{ for all } x \in \mathbb{R}.$$

Hence, similar to the argument in subsection 3.2.3 on the DFT, leading to the approximation in (3.22),

$$f(x_k) \approx \frac{1}{\pi} \operatorname{Re} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \widehat{f}(u_j)\, w_j \Delta u, \tag{3.23}$$
$$\text{for each } x_k,\ k = 1, 2, ..., N.$$

If $\widehat{f}(u) = \phi(u)$ is a C.F. then (3.23) is the FFT approximation to the density. This will be applied to the conditional C.F. of the log-return density given the price history in Chapter 4, and to both the risk-neutral and objective characteristic functions for the log-price density in Chapter 6. Also, the Carr and Madan (1999) FFT formula is used in Chapter 5, and again in Chapter 7 for conditional call pricing under the LJI model of subsection 2.7.2.

The other simple case of the FFT approximation occurs when the inverse transform $f(x)$ is purely imaginary. In this case equation (3.8) of Corollary 3.7 gives

$$f(x) = \frac{i}{\pi} \operatorname{Im} \int_0^{\infty} e^{-iux} \widehat{f}(u)\, du, \text{ for all } x \in \mathbb{R}.$$

Hence, similar to (3.23),

$$f(x_k) \approx \frac{i}{\pi} \operatorname{Im} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \widehat{f}(u_j)\, w_j \Delta u, \tag{3.24}$$
$$\text{for each } x_k,\ k = 1, 2, ..., N.$$

The approximation in (3.24) arises in the Bates (2006) volatility filter in Chapter 4, since we use the C.F. in place of the M.G.F. to extract moments.

3.3 L1 Characteristic Functions

In this section we consider the L^1 criterion for the pure jump models of Section 2.3, as well as the marginal log-price characteristic functions for the SVJ and SVSJ models used to be applied in Chapter 6 of this thesis. The conditional characteristic functions and joint transforms of the Bates (2006) AML method will be treated in Chapter 4.

3.3.1 Jump Characteristic Functions

Let $\phi_{X_t}(u)$ be the characteristic function of a pure jump process. As the following proposition illustrates, the condition that $\phi_{X_t}(u) \in L^1$ depends on the degree of activity of the jumps.

Proposition 3.13 *Consider the pure jump processes analyzed in Section 2.3.*

1. *For Merton jumps* $\phi_{X_t}(u) \notin L^1$.
2. *For Variance Gamma jumps* $\phi_{X_t}(u) \in L^1$ *only if* $\lambda t > \frac{1}{2}$.
3. *For Meixner jumps* $|\phi_{X_t}(u)|$ *has exponential tails; hence,* $\phi_{X_t}(u) \in L^1$.

Proof. *1.* For Merton jumps, equation (2.23) in subsection 2.3.1 yields

$$\phi_{X_t}(u) = \exp\left(\lambda t\left(e^{\left(i\beta u - \frac{1}{2}\alpha^2 u^2\right)} - 1\right)\right).$$

However, the complex exponential is bounded such that

$$\begin{aligned} \left|\operatorname{Re} e^{\left(i\beta u - \frac{1}{2}\alpha^2 u^2\right)}\right| &= e^{-\frac{1}{2}\alpha^2 u^2}\left|\cos(\beta u)\right| \le e^{-\frac{1}{2}\alpha^2 u^2}, \\ \text{and } \left|\operatorname{Im} e^{\left(i\beta u - \frac{1}{2}\alpha^2 u^2\right)}\right| &= e^{-\frac{1}{2}\alpha^2 u^2}\left|\sin(\beta u)\right| \le e^{-\frac{1}{2}\alpha^2 u^2}. \\ \text{Hence, } e^{\left(i\beta u - \frac{1}{2}\alpha^2 u^2\right)} &\to 0 \text{ as } |u| \to \infty. \end{aligned}$$

This implies $\phi_{X_t}(u) \to e^{-\lambda t}$ as $|u| \to \infty$. Thus, $\int_{-\infty}^{+\infty} |\phi_{X_t}(u)|\, du = \infty$.

2. For VG jumps, equation (2.31) in subsection 2.3.2 yields

$$\begin{aligned} \phi_{X_t}(u) &= \left(1 - iu\beta + \frac{1}{2}\alpha^2 u^2\right)^{-\lambda t}, \\ \text{so that } |\phi_{X_t}(u)| &= \left(1 + \left(\alpha^2 + \beta^2\right)u^2 + \frac{1}{4}\alpha^4 u^4\right)^{-\frac{\lambda t}{2}}. \end{aligned} \tag{3.25}$$

Note that by Property *5.* of Proposition 3.11, $|\phi_{X_t}(u)|$ is even on $\mathbb{R}$. Thus, for all characteristic functions

$$\int_{-\infty}^{+\infty} |\phi_{X_t}(u)|\, du = 2\int_0^1 |\phi_{X_t}(u)|\, du + 2\int_1^{\infty} |\phi_{X_t}(u)|\, du. \tag{3.26}$$

Moreover, since $|\phi_{X_t}(u)| \leq 1$ for all $u \in \mathbb{R}$, $\int_0^1 |\phi_{X_t}(u)|\, du < \infty$ for any C.F. However, in the Variance Gamma case

$$|\phi_{X_t}(u)| = \left(1 + \left(\alpha^2 + \beta^2\right) u^2 + \frac{1}{4}\alpha^4 u^4\right)^{-\frac{\lambda t}{2}} \sim K \frac{1}{u^{2\lambda t}}$$

for some $K > 0$, as $u \to +\infty$.

$$\text{But,} \quad \int_1^\infty \frac{1}{u^{2\lambda t}} du$$

diverges for $2\lambda t \leq 1$, and converges for $2\lambda t > 1$.

Hence, $\phi_{X_t}(u) \in L^1$ only if $\lambda t > \frac{1}{2}$.

3. For Meixner jumps, equation (2.39) in subsection 2.3.3 yields

$$\phi_{X_t}(u) = \left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cosh\left(\frac{1}{2}(\alpha u - i\beta)\right)}\right)^{2\lambda t}. \tag{3.27}$$

From Brown and Churchill (2007), p. 106, for $z = x + iy$

$$|\cosh(z)|^2 = \sinh^2(x) + \cos^2(y). \tag{3.28}$$

Hence, we obtain for the Meixner

$$|\phi_{X_t}(u)| = \left(\frac{\cos^2\left(\frac{\beta}{2}\right)}{\sinh^2\left(\frac{\alpha u}{2}\right) + \cos^2\left(\frac{\beta}{2}\right)}\right)^{\lambda t}, \tag{3.29}$$

and indeed, $|\phi_{X_t}(0)| = 1 < \infty$. Again by Property *5.* of Proposition 3.11, $|\phi_{X_t}(u)|$ is even on $\mathbb{R}$. Thus, we may focus on the right tail. Now since

$$\sinh^2\left(\frac{\alpha u}{2}\right) = \frac{e^{\alpha u} + e^{-\alpha u} - 1}{4} \sim \frac{1}{4} e^{\alpha u} \text{ for large } u > 0,$$

we obtain $|\phi_{X_t}(u)| \leq K e^{-\alpha \lambda t u}$, as $u \to \infty$, for some $K > 0$.

That is, for the Meixner, the function $|\phi_{X_t}(u)|$ has exponential tails with decay rate $\alpha\lambda t > 0$. And since $|\phi_{X_t}(u)| \leq 1$, this implies $\phi_{X_t}(u) \in L^1$. ■

The next result is a formality to be used in L^1 proofs concerning the SVJ model where the jumps are an independent process. First, some definitions. For a pure jump process X_t, equation $(A.4)$ in Appendix A defines the compensated characteristic exponent to be

$$\widehat{\psi}_X(u) = \psi_X(u) - iu\psi_X(-i). \tag{3.30}$$

Hence, by Theorem 2.5 the compensated characteristic function of the jumps is given by

$$\widehat{\phi}_{X_t}(u) = e^{t\psi_X(u)-iut\psi_X(-i)} = \phi_{X_t}(u)\, e^{-iut\psi_X(-i)}. \tag{3.31}$$

Lemma 3.14 *For any pure jump Lévy process X_t*

$$\left\|\widehat{\phi}_{X_t}(u)\right\|_\infty < M < \infty. \tag{3.32}$$

Proof. By equation (3.31) and Proposition 3.11 *2.*

$$\left|\widehat{\phi}_{X_t}(u)\right| = \left|\phi_{X_t}(u)\, e^{-iut\psi_X(-i)}\right| = |\phi_{X_t}(u)| \leq 1.$$

The result follows by definition of the essential supremum. ■

3.3.2 Exponentially Affine Characteristic Functions

Let $\tau = T - t_0 > 0$. As given in equation $(A.3)$ of Appendix A, for both the SVJ and SVSJ models, the joint conditional C.F. of the log-price and the latent factor has an exponentially affine form given by

$$\phi_{Y_T,\sigma_T^2|Y_{t_0},\sigma_{t_0}^2}(u,v) = \exp\left[iuY_{t_0} + \underline{C}(u,v;\tau) + \underline{D}(u,v;\tau)\,\sigma_{t_0}^2\right]. \tag{3.33}$$

In this thesis we denote the joint affine coefficients with an underscore,

$$\underline{C} = \underline{C}(u,v;\tau) \text{ and } \underline{D} = \underline{D}(u,v;\tau), \tag{3.34}$$

and the marginal log-price affine coefficients with no underscore,

$$C(u;\tau) = \underline{C}(u,0;\tau) \text{ and } D(u;\tau) = \underline{D}(u,0;\tau). \tag{3.35}$$

The affine coefficients $\underline{C}(u,v;\tau)$ and $\underline{D}(u,v;\tau)$ are derived in Bates (2006), see pp. 953-955. Our version is given in Appendix A of this thesis. The Bates (2006) AML method of Chapter 4 uses the affine joint conditional C.F. given in equation (3.33) above. But, when the initial point $\left(Y_{t_0},\sigma_{t_0}^2\right)$ is a known quantity, we simply denote the affine joint C.F. by

$$\phi_{Y_T,\sigma_T^2}(u,v)=\exp\left[iuY_{t_0}+\underline{C}(u,v;\tau)+\underline{D}(u,v;\tau)\,\sigma_{t_0}^2\right]. \tag{3.36}$$

In the analysis in Chapter 5 and Chapter 6, $\left(Y_{t_0},\sigma_{t_0}^2\right)$ is a known, hence we denote the marginal log-price C.F. by

$$\phi_{Y_T}(u)=\exp\left[iuY_{t_0}+C(u,\tau)+D(u,\tau)\,\sigma_{t_0}^2\right]. \tag{3.37}$$

Notice that for the each of the above SVJ and SVSJ model affine CFs, the coefficient of Y_{t_0} is triviallly equal to the purely complex number iu. This leads to the following very useful theoretical result.

Lemma 3.15 *Let $\tau>0$. For the joint C.F. in equation (3.33), assume that the coefficients $\underline{C}(u,v;\tau)$ and $\underline{D}(u,v;\tau)$ do not depend on either Y_{t_0} or $\sigma_{t_0}^2$, and that $\rho<0$. Then, for all $(u,v)\in\mathbb{R}\times(-\delta,\delta)$, and for some $\delta>0$,*

$$\begin{aligned}&\operatorname{Re}\underline{C}(u,v;\tau)\leq 0 \text{ and } \operatorname{Re}\underline{D}(u,v;\tau)\leq 0. \text{ Moreover, when } v=0,\\ &\operatorname{Re}C(u,\tau)\leq 0 \quad \text{and} \quad \operatorname{Re}D(u;\tau)\leq 0, \text{ for all } u\in\mathbb{R}.\end{aligned}$$

Proof. First, by Proposition C.12 of Appendix C.3, since $\rho<0$ in the SVJ and SVSJ models, and $\tau>0$, we have that $\underline{C}(u,v;\tau)$ and $\underline{D}(u,v;\tau)$ are both continuous on the principal branch, for all $(u,v)\in\mathbb{R}\times(-\delta,\delta)$, and for some $\delta>0$. Second, for any $\sigma_{t_0}^2>0$, we must have $\left|\phi_{Y_T,\sigma_T^2|Y_{t_0},\sigma_{t_0}^2}(u,v)\right|\leq 1$, since equation (3.33) defines a C.F. Thus, by taking the logarithm of the norm of the exponential on the right hand side of equation (3.33),

$$\operatorname{Re}\underline{C}(u,v;\tau)+\operatorname{Re}\underline{D}(u,v;\tau)\,\sigma_{t_0}^2\leq 0. \tag{3.38}$$

However, (3.38) implies that at most one of $\operatorname{Re}\underline{C}(u,v;\tau)$ and $\operatorname{Re}\underline{D}(u,v;\tau)$ can be positive. Thus, suppose that $\operatorname{Re}\underline{C}(u,v;\tau) > 0$. Then there exists $\sigma_{t_0}^2 > 0$ such that (3.38) is false. Similarly, if $\operatorname{Re}\underline{D}(u,v;\tau) > 0$ then there exists $\sigma_{t_0}^2 > 0$ such that (3.38) is false. It follows that

$$\operatorname{Re}\underline{C}(u,v;\tau) \leq 0 \text{ and } \operatorname{Re}\underline{D}(u,v;\tau) \leq 0, \tag{3.39}$$

for all $(u,v) \in \mathbb{R} \times (-\delta,\delta)$, and for some $\delta > 0$, as required. ■

3.3.3 SVJ and SVSJ Log-Price Characteristic Functions

In this subsection we prove that the log-price C.F.s for the SVJ and SVSJ models to be used in Chapter 6 are L^1. In Chapter 5 damping will satisfy integrability. Consider a maturity date $T = \tau > 0$, and initial time $t_0 = 0$. Let S_0 be the initial stock price, and define

$$Z_T = \log\left(\frac{S_T}{S_0}\right). \tag{3.40}$$

Then by equation (3.37) we obtain the normalized log-price C.F. as

$$\phi_{Z_T}(u) = \exp\left[C(u;\tau) + D(u;\tau)\,\sigma_0^2\right]. \tag{3.41}$$

The following are conventions of this thesis. A superscript $J0$ indicates the SVJ case, a superscript $J1$ indicates the SVSJ case, and either a superscript ψ or no superscript indicates both. A superscript SV denotes the Heston model. The affine coefficients of the log-price C.F.,

$$C(u) = \left\{ \begin{array}{c} C^{J0}(u) \\ C^{J1}(u) \end{array} \right\} = \left\{ \begin{array}{c} C^{SV}(u) \\ C^{J1}(u) \end{array} \right\} + \mathbf{1}_{SVJ}\widehat{\psi}_X(u)\,\tau, \tag{3.42}$$

$$\text{and } \; D(u) = D^{\psi}(u), \tag{3.43}$$

for the SVJ and SVSJ models, are fully defined in Appendix A, in terms of their constituent auxiliary variables. By Proposition B.16, since $\tau > 0$, the

hypothesis $\rho < 0$ is sufficient for principal branch continuity of the log-price coefficients $C(u)$ and $D(u)$, see Appendix B.2. In the proof of the next result, we refer to the modified $C(u)$ coefficient given by

$$\widetilde{C}(u) = C(u) - \mathbf{1}_{SVJ}\widehat{\psi}_X(u)\tau = \left\{ \begin{array}{l} C^{SV}(u;\tau) \\ C^{J1}(u;\tau) \end{array} \right\}, \tag{3.44}$$

from equation $(C.18)$ in Appendix C. This is done because the jumps in the SVJ model are an external independent process, whereas in the SVSJ model the jumps are internal, see equations $(A.4)$ to $(A.11)$ of Appendix A.1. The following result is also used in Chapter 4.

Lemma 3.16 *Let $\tau > 0$. For the SVJ and SVSJ models, if $\rho < 0$, then the affine coefficient $C(u)$ has the property*

$$e^{C(u)} \in L^1.$$

Proof. First we prove that $e^{\widetilde{C}(u)} \in L^1$. Since $\tau > 0$ and $\rho < 0$, Proposition B.16 *1.* in Appendix B.2 gives that $C(u)$ is continuous on the principal branch. But, by Lemma B.1, $\widehat{\psi}_X(u)$ is continuous. Hence, by equation (3.44) above, $\widetilde{C}(u)$ is continuous on the principal branch. It follows that on the principal branch,

$$\int_{-M}^{M} \left| e^{\widetilde{C}(u)} \right| du < \infty, \text{ for all finite } M > 0.$$

However, since $\tau > 0$ and $\rho < 0$, Lemma C.6 in Appendix C.1 implies that

$$\widetilde{C}(u) \sim \mp \frac{\kappa\eta\tau}{\omega}\sqrt{1-\rho^2}u + i\tau\left(\mu_0 - \frac{\kappa\eta\rho}{\omega}\right)u, \text{ as } u \to \pm\infty. \tag{3.45}$$

Therefore,

$$\left| e^{\widetilde{C}(u)} \right| = e^{\operatorname{Re}\widetilde{C}(u)} \sim e^{\mp\frac{\kappa\eta\tau}{\omega}\sqrt{1-\rho^2}u}, \text{ as } u \to \pm\infty. \tag{3.46}$$

This implies that

$$\int_{-\infty}^{+\infty} \left| e^{\widetilde{C}(u)} \right| du < \infty. \tag{3.47}$$

Note that (3.47) is sufficient for the SVSJ case by equation (3.44). Thus,

$$e^{C^{J1}(u)} \in L^1. \tag{3.48}$$

However, (3.47) also implies that

$$e^{C^{SV}(u)} \in L^1, \text{ where} \tag{3.49}$$

$$C^{J0}(u) = C^{SV}(u) + \widehat{\psi}_X(u)\,\tau. \tag{3.50}$$

Moreover, by equation (3.41)

$$e^{\widehat{\psi}_X(u)\tau} = \widehat{\phi}_{X_\tau}(u), \tag{3.51}$$

and, by Lemma 3.14, for any jump type,

$$\left\|\widehat{\phi}_{X_t}(u)\right\|_\infty < \infty. \tag{3.52}$$

Therefore, by Hölder's inequality

$$\begin{aligned}\int_{-\infty}^{+\infty} \left|e^{C^{J0}(u)}\right| du &= \int_{-\infty}^{+\infty} \left|e^{C^{SV}(u)}\right| \left|e^{\widehat{\psi}_X(u)\tau}\right| du \\ &\leq \left\|e^{C^{SV}(u)}\right\|_1 \left\|\widehat{\phi}_{X\tau}(u)\right\|_\infty < \infty,\end{aligned}$$

by (3.49) and (3.52), as required. ■

Theorem 3.17 *Let $T = \tau > 0$ be a maturity date, with initial time $t_0 = 0$, and let S_0 be the initial stock price. Define for the SVJ and SVSJ models*

$$Z_T = \log\left(\frac{S_T}{S_0}\right).$$

Then, under the assumption that $\rho < 0$,

$$\phi_{Z_T}(u) = e^{C(u)+D(u)\sigma_0^2} \in L^1. \tag{3.53}$$

Proof. By Lemma 3.16, if $\tau > 0$ and $\rho < 0$, then

$$\left\|e^{C(u)}\right\|_1 = \int_{-\infty}^{+\infty} \left|e^{C(u)}\right| du < \infty. \tag{3.54}$$

Moreover, by Lemma 3.15, for the SVJ and SVSJ models, $\tau > 0$ with $\rho < 0$ implies

$$\operatorname{Re} D(u) \leq 0, \text{ for all } u \in \mathbb{R}. \tag{3.55}$$

Since we assume that $\sigma_0^2 > 0$, this implies that

$$\left\| e^{D(u)\sigma_0^2} \right\|_\infty = \sup_{u \in \mathbb{R}} \left| e^{D(u)\sigma_0^2} \right| = \sup_{u \in \mathbb{R}} \left[e^{\operatorname{Re} D(u)\sigma_0^2} \right] < \infty. \tag{3.56}$$

Therefore, by Hölder's inequality

$$\begin{aligned} \int_{-\infty}^{+\infty} \left| e^{C(u)+D(u)\sigma_0^2} \right| du &= \int_{-\infty}^{+\infty} \left| e^{C(u)} \right| \left| e^{D(u)\sigma_0^2} \right| du \\ &\leq \left\| e^{C(u)} \right\|_1 \left\| e^{D(u)\sigma_0^2} \right\|_\infty < \infty, \end{aligned}$$

by (3.54) and (3.56), as required. ■

3.4 The FFT Approximation in Practice

3.4.1 The Trapezoidal Rule

For simplicity, consider the integral on the half real line given by

$$g(x) = \int_0^\infty e^{-iux} \widehat{f}(u)\, du, \tag{3.57}$$

from equation (3.14). Note that $g(x)$ in equation (3.57) may have both real and imaginary parts. By following the argument in equations (3.15) to (3.22) of subsection 3.2.3, and setting $N = 2^p$, we obtain for the integral in (3.57), the FFT approximation with quadrature weights w_j given by

$$\begin{aligned} g(x_k) \approx \sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \widehat{f}(u_j)\, w_j \Delta u, \\ \text{for each } x_k,\ k = 1, 2, ..., N. \end{aligned} \tag{3.58}$$

The FFT approximations in (3.23) and (3.24) of subsection 3.2.4 are merely special cases of (3.58). In this subsection we give the standard theoretical

argument, from Briggs and Henson (1995), in favour of using the trapezoidal rule for the quadrature weights w_j in the FFT approximation (3.58). The main idea here is that the trapezoidal rule properly weights the sampled endpoints, $\widehat{f}(0)$ and $\widehat{f}(A)$. That is, we show in this subsection that the sampled endpoints should each be weighted by one $\frac{1}{2}$.

Definition 3.18 (Periodic Extension) *Let $[0, A]$ be the domain of h. Then the function h_A defined by*

$$h_A(x + mA) = h(x) \ , \ x \in [0, A] \ , \ m = 0, \pm 1, \pm 2, ... \tag{3.59}$$

is periodic with period A, and is called the A-periodic extension of h.

(Briggs and Henson (1995), §2.4, p. 38)

Definition 3.19 (Fourier Series) *Let f be an A-periodic function. The non-centred Fourier coefficients of f are given by*

$$c_k = \frac{1}{A} \int_0^A e^{-i\frac{2\pi}{A}kx} f(x)\, dx, \ \ k = 0, \pm 1, \pm 2, ... \ , \text{ and} \tag{3.60}$$

$$S_M(x) = \sum_{k=-M}^{M} c_k e^{i\frac{2\pi}{A}kx} \tag{3.61}$$

is the M^{th} partial sum of the Fourier series for $f(x)$. (Walker (1996), p.4)

Theorem 3.20 (Pointwise Convergence) *Let f be piecewise smooth and A-periodic. Then*

$$\sum_{k=-M}^{M} c_k e^{i\frac{2\pi}{A}kx} \to \frac{f(x+) + f(x-)}{2} \tag{3.62}$$

pointwise as $M \to \infty$, for all x. (Briggs and Henson (1995), pp. 37-38)

The key point raised in §3.4 pp. 93-95 of Briggs and Henson (1995) is that the periodic extension $\widehat{f}_A(u)$ should be sampled from in place of $\widehat{f}(u)$

in the FFT approximation. In so doing we can apply Theorem 3.20 to obtain the correct weights for the sampled endpoints, $\widehat{f}(0)$ and $\widehat{f}(A)$.

By Theorem 3.20 in the limit as $M \to \infty$

$$\widehat{f}_A(0) = \widehat{f}_A(A) = \frac{\widehat{f}_A(0) + \widehat{f}_A(A)}{2} = \frac{\widehat{f}(0) + \widehat{f}(A)}{2}. \tag{3.63}$$

However, since $\widehat{f}_A(u)$ is A-periodic, the interval $[0, A]$ should be sampled only at one endpoint. Consequently, in order to sample correctly from $\widehat{f}(u)$ on $[0, A]$ we should sample from the auxiliary function

$$f^*(u) = \left\{ \begin{array}{cc} \widehat{f}(u) & \text{if } u \in (0, A) \\ \frac{\widehat{f}(0)+\widehat{f}(A)}{2} & \text{if } u = A \end{array} \right\}, \text{ on } (0, A], \tag{3.64}$$

see Briggs and Henson (1995) §3.4 p. 95. Observe that this prescription is met by sampling uniformly from $\widehat{f}(u)$ on $[0, A]$ and then weighting each endpoint by one $\frac{1}{2}$. The most direct way to apply this in practice is to simply use the trapezoidal rule. This choice is critical in the non-centred and shifted scheme of Carr and Madan (1999), pp. 67-68, where $\widehat{f}(0)$ and $\widehat{f}(A)$ do not match, see Briggs and Henson (1995), p. 95 and p. 360.

3.4.2 Comparison to Simpson's Rule

Consider the standard normal distribution which has a characteristic function and a probability density function, each respectively given by

$$\phi(u) = e^{-\frac{1}{2}u^2}, \text{ and} \tag{3.65}$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{3.66}$$

Since the C.F. in equation (3.65) is obviously L^1, Fourier inversion applies, and we have

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux} \phi(u)\, du. \tag{3.67}$$

By equation (3.23) from subsection 3.2.4 the FFT approximation is

$$f(x_k) \approx \frac{1}{\pi} \operatorname{Re} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \phi(u_j) w_j \Delta u, \tag{3.68}$$

$$\text{for each } x_k,\ k = 1, 2, ..., N.$$

For the FFT approximation in (3.68) we consider two quadrature rules:

1. Simpson's rule $w_1 = \frac{1}{3}$, and $w_j = \left\{ \begin{array}{c} \frac{2}{3} \text{ if } j \text{ is odd} \\ \frac{4}{3} \text{ if } j \text{ is even} \end{array} \right\}$, $j = 2, 3, ..., N$
   (cf. Carr and Madan (1999), p. 68)

2. The trapezoidal rule $w_1 = w_N = \frac{1}{2}$, and $w_j = 1$, for $j = 2, 3, ..., N-1$
   (recommended in Briggs and Henson (1995), pp. 19-20).

Table 3.1 below confirms that the trapezoidal rule performs at high precision, for each x_k, $k = 1, 2, ..., N$. However, Simpson's rule does not.

$N = 2^{14}$	**Total Abs Err**	**Mean Abs Err**
Trapezoidal Rule	7.23E-13	4.41E-17
Simpson's Rule	217.29	0.01326

Table 3.1: Errors in the FFT Approximation: Standard Normal.

In Figure 3.1 below, the error in Simpson's rule is clearly illustrated for the standard normal FFT approximation. Simpson's rule goes strongly negative in the deep tails. Moreover, as will be shown in Chapter 5, this is also why the Carr and Madan (1999) formula gives a negative European call option price deep out-of-the-money. According to Briggs and Henson (1995) §3.4, pp. 93-95, this type of error is expected whenever the sampled endpoints, in this case $\widehat{f}(0)$ and $\widehat{f}(A)$, are not matched. Briggs and Henson (1995), p. 360, recommends the trapezoidal rule in the event of such a mismatch. Moreover, such a mismatch is inherent in the non-centred and shifted FFT scheme of Carr and Madan (1999).

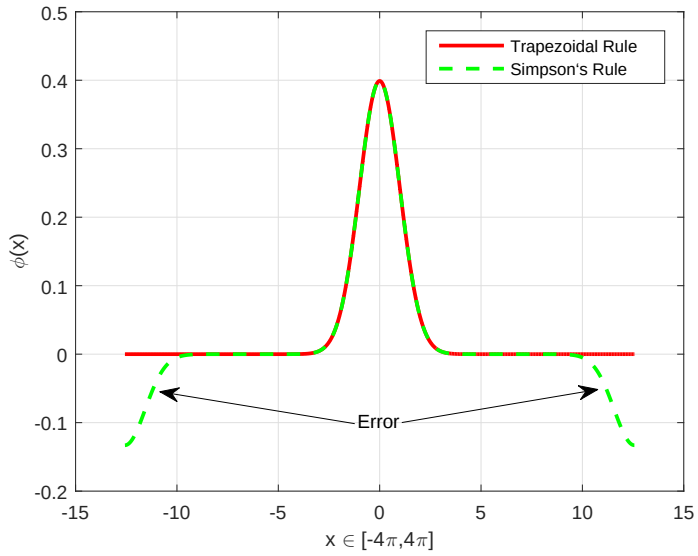

Figure 3.1: FFT Quadrature Rules: Standard Normal Density.

We turn now to the accuracy of the trapezoidal versus Simpson's rule on the interior of the inverse transform domain. Standard quadrature theory, see for example Burden and Faires (1997), pp. 201-203, states that the order of the absolute error for Simpson's rule in the general case is $O\left(\Delta u^4\right)$, whereas the same measure is only $O\left(\Delta u^2\right)$ for the trapezoidal rule. But, as we will show, the advantage to Simpson's rule can pass by unfulfilled in certain circumstances. For example, if the integrand is periodic, then the absolute error in the trapezoidal rule decays geometrically with N, see Davis and Rabinowitz (1984), pp. 315-16. Recall from equation (3.57) in subsection 3.4.1 that the integral that we intend to approximate using FFT is given by

$$g(x) = \int_0^\infty e^{-iux} \widehat{f}(u)\, du \tag{3.69}$$

$$= \int_0^\infty \left[\cos(ux)\, \widehat{f}(u) - i \sin(ux)\, \widehat{f}(u)\right] du. \tag{3.70}$$

Since the integrand in equation (3.70) above exhibits damped oscillation on $(0, \infty)$, the trapezoidal rule may outperform the error rate of $O\left(\Delta u^2\right)$ for the FFT approximation. We now consider two log-return distributions, for which the exact density is known in closed form, and examine the error in the FFT approximation. The two distributions are the Meixner which has semi-heavy density tails, see Schoutens (2003), p. 63, and the Cauchy distribution which is known to have very heavy density tails. From subsection 2.3.3, the Meixner log-return distribution has a C.F. and a density given by

$$\phi\left(u;\Delta t\right) = e^{iu\mu\Delta t}\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cosh\left(\frac{1}{2}\left(\alpha u - i\beta\right)\right)}\right)^{2\lambda\Delta t}, \text{ and} \tag{3.71}$$

$$f\left(y;\Delta t\right) = \frac{\left(2\cos\left(\frac{\beta}{2}\right)\right)^{2\lambda\Delta t} e^{\beta\left(\frac{y-\mu\Delta t}{\alpha}\right)}}{2\pi\Gamma\left(2\lambda\Delta t\right)}\left|\Gamma\left[\lambda\Delta t + i\left(\frac{y-\mu\Delta t}{\alpha}\right)\right]\right|^2. \tag{3.72}$$

By Proposition 3.13, *3.*, $|\phi\left(u;\Delta t\right)|$ has exponential tails, thus $\phi\left(u;\Delta t\right) \in L^1$. Hence, Fourier inversion applies. We assume the non-centred and shifted FFT setup from subsection 3.2.3, and set $N = 2^{14}$. But, we truncate the density range down to $\left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$, so that the contaminated part of Simpson's rule is discarded. That is, we apply the FFT approximation from (3.23) in subsection 3.2.4 for $y_k \in [-4\pi, 4\pi]$, then discard values to obtain

$$f\left(y_k;\Delta t\right) \approx \frac{1}{\pi}\operatorname{Re}\sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j}\phi\left(u_j;\Delta t\right) w_j \Delta u, \tag{3.73}$$
$$\text{for each } y_k \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right].$$

$\Delta u = 2.5 \times 10^{-1}$	Exact Points y_k	Midpoints $\frac{y_k + y_{k+1}}{2}$
Trapezoidal Rule	2.79E-14	2.28E-07
Simpson's Rule	2.78E-14	2.28E-07

Table 3.2: Average FFT Approximation Errors: Meixner Density.

In Table 3.2 above, the average absolute errors are computed for the exact $y_k \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$, and a spline is used for the midpoints. Table 3.2 shows that for the Meixner density there is essentially no difference between the trapezoidal rule and Simpson's rule on the interior points of the FFT approximation. The parameters of the Meixner density in Table 3.2 are $(\mu, \alpha, \beta, \lambda) = (-.0485, .0702, .0790, 92.48)$, estimated by ML on an annual basis with $\Delta t = \frac{1}{252}$ from the daily log returns on Apple stock (1991-2011), see subsection 4.4.2. Next we pursue a similar treatment for the Cauchy distribution which has a C.F. and a density function given by

$$\phi(u; \Delta t) = e^{iu\mu\Delta t - \beta\Delta t|u|}, \text{ and} \tag{3.74}$$

$$f(y; \Delta t) = \frac{1}{\beta\Delta t\pi\left[1 + \left(\frac{y-\mu\Delta t}{\beta\Delta t}\right)^2\right]}. \tag{3.75}$$

By equation (3.74), $|\phi(u; \Delta t)| = \left|e^{iu\mu\Delta t - \beta\Delta t|u|}\right| = e^{-\beta\Delta t|u|}$, so that clearly $\phi(u; \Delta t) \in L^1$. Thus, Fourier inversion applies. We now proceed exactly as we did for the Meixner density, with the FFT approximation restricted to $y_k \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$ and a spline interpolant within this range. The annual parameters $(\mu, \beta) = (.1173, 3.790)$ were estimated by ML from the daily log returns on Apple stock (1991-2011), see subsection 4.4.2.

$\Delta u = 2.5 \times 10^{-1}$	Exact Points y_k	Midpoints $\frac{y_k + y_{k+1}}{2}$
Trapezoidal Rule	2.49E-05	3.33E-05
Simpson's Rule	6.45E-08	8.84E-06

Table 3.3: Average FFT Approximation Errors: Cauchy Density.

In Table 3.3, Simpson's rule is better for the Cauchy. But, the Cauchy trapezoidal error is better than the rate of $O\left(\Delta u^4\right)$ typifying Simpson's rule. The difference here appears to be that the Meixner density has semi-heavy tails, while the Cauchy density has heavy tails.

Lastly, let $\Delta t = \frac{1}{252}$, and let $Y_1 = \log\left(\frac{S_{\Delta t}}{S_0}\right)$ be the first log return. Then by equation (3.41) the SVSJ model with $\sigma_0^2 = \eta$ has a C.F. given by

$$\phi_{Y_1}(u;\Delta t) = \exp\left[C(u;\Delta t) + D(u;\Delta t)\,\eta\right], \tag{3.76}$$

for a suitable set of affine coefficients $C(u)$ and $D(u)$ from Appendix A. By Theorem 3.17, $\phi_{Y_1}(u;\Delta t) \in L^1$. Hence, by Fourier inversion we are able to obtain the SVSJ Merton first log-return density. Approximate Maximum Likelihood parameter estimates for Apple stock daily log returns (1991-2011) under the SVSJ Merton model are taken from Table 4.2 in subsection 4.4.2. In Figure 3.2 below, we use FFT to graph log-density of the SVSJ Merton first log-return under these parameters, and compare with the Meixner and Cauchy log-return densities. Observe that the SVSJ Merton first log-return log-density is much closer to the Meixner log-density than it is to the Cauchy. This suggests that the SVSJ Merton model has semi-heavy tails, and that the trapezoidal rule will perform well for FFT in this case. We expect the same for the SVJ model.

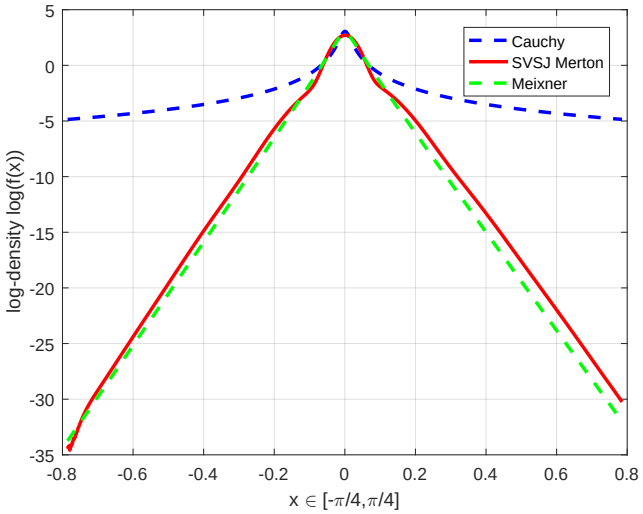

Figure 3.2: Log-Densities: SVSJ Merton vs. Cauchy and Meixner.

3.5 Conclusion

1. The marginal log-price characteristic function for the SVJ and SVSJ Merton models is L^1 integrable. Moreover, the Heston model is a special case of each of these models with no jumps. Furthermore, by Remarks 2.28, 2.33, and 2.37, this holds for the SVJ and SVSJ Merton models under both the P and Q measures. This information will be used in Chapter 6 of Part II for Fourier inversion via FFT, and also to show that both the P and Q measures of these respective models are absolutely continuous with respect to the Lebesgue measure.

2. The trapezoidal rule is a much safer alternative to the usual Simpson's rule for quadrature in the FFT approximation. Moreover, we find that in practice the trapezoidal rule should be suitably accurate on the interior of the inverse transform domain, provided that the tails of the inverse transform are not too heavy. This is likely because the inverse Fourier integrand exhibits damped oscillation which is similar to periodicity. The trapezoidal rule is known to converge geometrically for periodic integrands. For the remainder of this thesis, we use the trapezoidal rule for quadrature in FFT, outside of further illustration of the drawbacks to Simpson's rule in Chapter 5 of Part II.

Part II

Estimation and Relative Entropy

Preface to Part II

The focus of this part is on estimation and the change of measure for the five affine Markov models defined in Chapter 2 above. Recall that these are the Heston model, the SVJ model with independent Merton, Variance Gamma, and Meixner jumps, respectively, as well as the SVSJ model with time changed compound Poisson Merton jumps. Each chapter in this part uses the FFT method established in Chapter 3 above for Fourier inversion in respectively different ways. In Chapter 4 we estimate the P-measures, under the five affine Markov models, from the log-returns on two indices, and seven individual stocks, using the Bates (2006) method of Approximate Maximum Likelihood (AML). We use FFT to invert each of the three main Fourier transforms of the AML method. Also, we prove that the requisite L^1 integrability conditions for Fourier inversion hold for each of the three transforms of the five affine Markov models, under the AML method. Of the nine assets estimated under the P-measure in Chapter 4, the best fit of the log-returns for the SVJ model with infinite-activity jumps is Apple stock (1991-2011). Moreover, the rest of the thesis considers only Apple stock. In Chapter 5 we estimate the option implied Q-measures, by calibration using least squares normal likelihood, under the five affine Markov models, from spot options prices on Apple stock, as selected from the daily closing options price book of January 19th, 2011. This is the last day in the Apple

stock (1991-2011) log-return data set. For pricing, we employ the Carr and Madan (1999) formula, thus inverting the Fourier transform of the damped call price by FFT. We also propose a modest improvement that appears to solve the known problem of negative call prices deep out-of-the-money in the Carr and Madan (1999) formula. Moreover, we find that for spot options on the daily closing book of Apple stock, the SVSJ Merton model with stochastic jump intensity has a slightly better fit than the other affine Markov models. Lastly, in Chapter 6 we present our contribution to the change of measure problem posed by the presence of incomplete markets in the five affine Markov models. By combining the estimated Apple stock P-measures from Chapter 4, in a unique way with the estimated Apple stock implied option Q-measures from Chapter 5, we construct what we refer to as the parametric minimum entropy martingale measure (PMEMM). The PMEMM is similar to the minimal entropy martingale measure (MEMM) of Frittelli (2000), only the PMEMM is much more easily computable. Since we assume a structure preserving Q-measure, the risk-neutral log-price CF is known in closed form. Thus, we can compute the relative entropy from the densities corresponding to the absolutely continuous measures P and Q, respectively, as obtained by Fourier inversion from their log-price CFs via FFT. These marginal log-price CFs were shown to be L^1 in Chapter 3. We also make forecasts of at-the-money discounted call option payoffs on Apple stock, thus comparing the PMEMM to the option implied Q-measure. Relative to actual discounted call option payoffs computed from the future Apple stock price, the PMEMM forecast outperforms for long dated options, but the option implied Q-measure is better for short term options.

Chapter 4

Approximate Maximum Likelihood

4.1 Introduction

This chapter covers estimation under the P-measure, using daily log-return data, for the five affine Markov models defined in Chapter 2 above. These are the Heston model, the SVJ model under Merton, Variance Gamma, and Meixner jumps, and the SVSJ Merton model. We often refer to these five models simply as the SVJ and SVSJ models, where the Heston model is a special case of either with no jumps. In the literature, MCMC based methods dominate for estimation of the P-measure under these models, see for example Eraker, Johannes, and Polson (2003). But, we have chosen the Fourier based method of Approximate Maximum Likelihood (AML) from Bates (2006) to estimate the P-measures required in this thesis. Since they are affine, and also as required by the AML method, the SVJ and SVSJ models have a known joint characteristic function for the log-price and the latent factor, see Appendix A of this thesis.

In this thesis we have made substantial contributions to the theory of AML. First, in Appendix A.2 we give an equivalent set of affine coefficients for the joint characteristic function of the SVJ and SVSJ models. Our affine coefficients are shown to be continuous on the principal branch, as proven in Appendix B and Appendix C.3. Equivalence to the Bates (2006) version is proven in Theorem A.1 of Appendix A.3. Secondly, in Section 4.3 below, we prove that for the SVJ and SVSJ models all three Fourier transforms of the AML method are both continuous on the principal branch, and L^1 integrable. Furthermore, we prove in Section 4.3 that the three respective inverse transforms are also L^1. A consequence of these results is that by Lemma 3.8 of subsection 3.2.1, the primary Fourier integral of the AML method converges uniformly. This last result leads, in Appendix D of this thesis, to proof that the right hand side of the conditional transform for the latent factor in the AML volatility filter can be differentiated twice under the integral. Based on these two key differentiability results, the theory of AML for the SVJ and SVSJ models is complete.

We also make a number of empirical contributions in this chapter using AML for estimation. Consistent with Eraker, Johannes, and Polson (2003), our standard sample size is about 5,000 daily log-returns covering twenty years of daily data. Table 4.2 and Table 4.3 in subsection 4.4.2 cover AML estimation, under each of the five SVJ and SVSJ models, for Apple stock (1991-2011) and the DJIA (1988-2007), respectively. These estimates are used for the true values in the corresponding pair of simulated AML estimation studies of subsection 4.4.3. Each study covers all five of the SVJ and SVSJ models, and the simulation results are based on 100 model data sets of 5,000 simulated log-returns. The SVJ and SVSJ model path simulation methods are given in Appendix E. These are based on Bates (2006), pp. 957-958, Broadie and Kaya (2006), pp. 221-22, and Glasserman (2004), pp.

138-144. For simulation of the Meixner process, see Chapter 8 of this thesis. In Table 4.15 of subsection 4.5.1 we cover AML estimation of the S&P 500 (1988-2007) for the five SVJ and SVSJ models. The immediate purpose of these estimates is to present the AML filtered volatility as compared to the VIX index in subsection 4.5.2. Furthermore, in Table 4.16 to Table 4.21 of subsection 4.6.1 we provide the AML estimates for each of the five SVJ and SVSJ models, under the daily log-returns of six additional individual stocks, over the 1988-2007 period. Five of these six stocks have been components of the Dow-Jones Industrial Average since at least 2001. British Petroleum (BP) is the other stock. Based on the analysis of Aït-Sahalia and Jacod (2012), see pp. 1036-37, we expect to find that at least some of these DJIA components will fit the SVJ model with infinite-activity jumps well. We find this to be the case for banking and technology stocks. In particular, the main asset of this thesis, Apple stock (1991-2011), fits the SVJ model with infinite-activity jumps exceptionally well. We also find that the model for the S&P 500 (1988-2007) daily log-returns with the highest log-likelihood, including a constraint for the Feller condition, is the infinite variation SVJ Meixner model, see Table 4.24 in subsection 4.6.2.

The remainder of the chapter is structured as follows. In Section 4.2 we cover AML theory from Bates (2006). In Section 4.3 we present our improvements to AML theory, and give our FFT based AML implementation in subsection 4.3.2. Section 4.4 covers the two AML simulated estimation studies, and also our comparisons to results from the Heston, SVJ, and SVSJ model literature. Section 4.5 treats the AML filtered volatility. In particular, we estimate the Heston and SVSJ models from S&P 500 index daily log-returns by AML, and compare AML filtered volatilities to the VIX. In Section 4.6 we table the additional AML estimates, and provide an analysis of all AML estimates from this chapter. Section 4.7 concludes.

4.2 The AML Method

4.2.1 Introduction

The method of Approximate Maximum Likelihood (AML) was pioneered in Bates (2006) where it was used to estimate affine models similar to our SVJ and SVSJ models defined in equations (4.1) and (4.2) below. The AML method succeeds in filtering the latent factor from the history of the log-price, so that the likelihood may be computed from the log-price data alone. The method employs conditional characteristic functions, and consequently Fourier analysis, for both the likelihood and the filtration of the latent factor. Recall from equation (2.72) in subsection 2.6.1 that for an arbitrary Lévy process X_t the continuous-time SVJ model under the P-measure is given by

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \psi_X(-i)\right)dt + \sigma_t dW_t^{(S)} + dX_t, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ \rho dt = E\left[dW_t^{(S)}dW_t^{(V)}\right], \end{aligned} \tag{4.1}$$

and that by Proposition 2.32 from subsection 2.6.1 the SVJ model is affine Markov with respect to $\sigma\left\{Y_t, \sigma_t^2\right\}$. Moreover, we have from equation (2.74) in subsection 2.6.2 that the continuous-time SVSJ time changed compound Poisson Merton model under the P-measure is given by

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \sigma_t^2\psi_X(-i)\right)dt + \sigma_t dW_t^{(S)} + dX(V_t), \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right] = \rho dt, \\ dV_t &= \sigma_t^2 dt, \end{aligned} \tag{4.2}$$

where in subsection 2.6.2 we argue that the SVSJ Merton model may be written as an Affine Jump-Diffusion. Thus, by Proposition 2.36 it is affine Markov with respect to $\sigma\left\{Y_t, \sigma_t^2\right\}$, see also Bates (2006), pp. 913 & 926.

Note that the Heston model is a special case of either the SVJ model or the SVSJ model, with no jumps. For the SVJ and SVSJ models, define the continuous-time process given by

$$Z_t = \left(Y_t, \sigma_t^2\right), \, t \geq 0. \tag{4.3}$$

From equation $(A.3)$ in Appendix A, the continuous-time joint affine conditional characteristic function for the SVJ and SVSJ models is given by

$$\phi_{Z_T|Z_{t_0}}(u,v) = E\left[e^{iuY_T + iv\sigma_T^2} | \left(Y_{t_0}, \sigma_{t_0}^2\right)\right] \tag{4.4}$$

$$= \exp\left[iuY_{t_0} + \underline{C}(u,v;\tau) + \underline{D}(u,v;\tau)\,\sigma_{t_0}^2\right], \tag{4.5}$$

where $T > 0$ is the terminal time, $t_0 \geq 0$ is the initial time, and $\tau = T - t_0 > 0$ is the gap time. The functional forms

$$\underline{C} = \underline{C}(u,v;\tau), \text{ and } \underline{D} = \underline{D}(u,v;\tau), \tag{4.6}$$

are referred to as the joint affine coefficients.

In subsection 4.2.2 below we treat the transition from the continuous-time models in equations (4.1) and (4.2) above to the discrete case, and summarize the main assumptions of the AML method. The transition to discrete-time is facilitated by the joint affine conditional characteristic function in equation (4.5) above. In subsection 4.2.3 we show how the AML method provides an approximation for the conditional distribution of the latent factor given the log-price history, yielding a joint C.F. conditional on the log-price history alone. Since the latent factor in the SVJ and SVSJ models is positive, and follows a CIR process, the gamma approximation is the most appropriate choice, see Bates (2006), p. 920. Then, since the log-return history augmented by the initial log-price, and the log-price history itself, each contain the same information, we switch to the augmented log-return history for convenience. In subsection 4.2.4 we present a result

from Bartlett (1938), p. 63, that the Bates (2006) AML method uses to obtain an expression for the partial inverse Fourier transform of the joint C.F. conditional on the log-price history. This allows the AML method to obtain the first two noncentral moments of the conditional distribution of the latent factor by the gamma approximation from subsection 4.2.3. The first moment is the filtered volatility squared. Furthermore, the parameters of the gamma approximation for the conditional distribution are sequentially updated by the method of moments, as the method proceeds through the log-return history. In subsection 4.2.5 we present the theoretical version of the Approximate Maximum Likelihood algorithm. A technical version of this algorithm, with FFT, is given in subsection 4.3.2.

4.2.2 The Discrete Time Joint Conditional C.F.

We wish to estimate the continuous-time latent factor SVJ and SVSJ models in equations (4.1) and (4.2) above. However, the continuous-time log-price Y_t is discretely observed, and the continuous-time latent factor σ_t^2 is also filtered discretely from the log-price data. Thus, for the SVJ and SVSJ models, where the joint continuous-time process is $Z_t = \left(Y_t, \sigma_t^2\right)$, $t \geq 0$, we consider a collection of observation points in continuous time given by

$$t_n = n\Delta t,\ n = 0, 1, 2, ..., \tag{4.7}$$

where $\Delta t > 0$. Then we define the process given by

$$Z_n = \left(Y_n, \sigma_n^2\right),\ n = 0, 1, 2, ..., \tag{4.8}$$

$$\text{such that } Z_n = Z_{t_n},\ n = 0, 1, 2, ..., \tag{4.9}$$

where $Z_n = \left(Y_n, \sigma_n^2\right)$, $n = 0, 1, 2, ...$, is the desired discrete-time process for the SVJ and SVSJ models. Observe that a step from time n to time $n+1$ under the discrete-time process Z_n, $n = 0, 1, 2, ...$, corresponds exactly

to a time step of length Δt under the continuous-time process Z_t, $t \geq 0$. This implies that for the SVJ and SVSJ models the process $Z_n = \left(Y_n, \sigma_n^2\right)$ inherits the Markov property from the process $Z_t = \left(Y_t, \sigma_t^2\right)$.

The key point of the AML method is that we only need to know the model for the discrete-time process $Z_n = \left(Y_n, \sigma_n^2\right)$, $n = 0, 1, 2, ...$, through its joint conditional characteristic function. Moreover, since we assume daily data, the gap time τ is a constant such that

$$\tau = \Delta t = \frac{1}{252}, \text{ per annum.} \tag{4.10}$$

Hence, the continuous-time joint affine conditional characteristic function from equation (4.5) above takes the special case

$$\phi_{Z_{n+1}|Z_n}(u, v) = E\left[e^{iuY_{n+1}+iv\sigma_{n+1}^2} | \left(Y_n, \sigma_n^2\right)\right] \tag{4.11}$$

$$= \exp\left[iuY_n + \underline{C}(u, v; \Delta t) + \underline{D}(u, v; \Delta t)\sigma_n^2\right], \tag{4.12}$$

which is precisely the discrete-time joint conditional characteristic function of the process $Z_n = \left(Y_n, \sigma_n^2\right)$.

The final assumption of the AML method is that the process σ_n^2, where by equation (4.7) above

$$\sigma_n^2 = \sigma_{t_n}^2, \; n = 0, 1, 2, ..., \tag{4.13}$$

is strictly stationary and strongly ergodic. Since the continuous-time process σ_t^2 is a CIR process in both the SVJ and the SVSJ models, the process σ_n^2 inherits the Markov property, similar to Z_n. Moreover, equations (4.7) and (4.9) above imply that the Markov process σ_n^2, $n = 0, 1, 2, ...$, has a j-step transition law given by

$$\sigma_{n+j}^2 = \frac{\omega^2\left(1 - e^{-\kappa j \Delta t}\right)}{4\kappa} \chi_d'^2\left(\frac{4\kappa e^{-\kappa j \Delta t}}{\omega^2\left(1 - e^{-\kappa j \Delta t}\right)}\sigma_n^2\right), \; j = 0, 1, 2, ..., \tag{4.14}$$

where $d = \frac{4\kappa\eta}{\omega^2}$, and $\chi'^2_d(\xi)$ is a noncentral chi-squared random variable with noncentrality parameter

$$\xi = \frac{4\kappa e^{-\kappa j \Delta t}}{\omega^2 \left(1 - e^{-\kappa j \Delta t}\right)} \sigma_n^2, \tag{4.15}$$

as adapted from Glasserman (2004), see p. 122. Define

$$\overline{\sigma}_n^2 = \lim_{j \to \infty} \sigma_{n+j}^2. \tag{4.16}$$

Similar to the proof of Proposition 2.24, item 6, from subsection 2.5.1, it can be shown by equation (4.14) above that as $j \to \infty$, $\xi \to 0$, such that $\overline{\sigma}_n^2$ is chi-squared random variable with degrees of freedom $d = \frac{4\kappa\eta}{\omega^2}$, and scaled by a factor of $\frac{\omega^2}{4\kappa}$. This implies that for $\overline{\sigma}_n^2$ in equation (4.16) above

$$\overline{\sigma}_n^2 \sim \Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right). \tag{4.17}$$

The following lemma shows that the final AML assumption holds for the SVJ and SVSJ Merton models, as discussed in Bates (2006), see p. 917.

Lemma 4.1 *For the SVJ and SVSJ Merton models, consider the discrete-time variance process σ_n^2 in equation* (4.13) *above, and let $\sigma_{n_0}^2 \sim \Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$ have the limiting distribution from equation* (4.17)*. If the Feller condition, $2\kappa\eta > \omega^2$, holds for the underlying continuous-time CIR process σ_t^2, $t \geq 0$, then σ_n^2, $n = 0, 1, 2, ...$, is strictly stationary and strongly ergodic.*

Proof. The process σ_n^2, $n = 0, 1, 2, ...$, inherits the boundaries of σ_t^2, $t \geq 0$. Thus, since the Feller condition holds for σ_t^2, $t \geq 0$, we have $\frac{2\kappa\eta}{\omega^2} > 1$, and no mass escapes at zero from the limiting gamma density of σ_n^2. Hence, σ_n^2 is positive ergodic, see Karlin and Taylor (1981), p. 221. Also, the initial points $\sigma_{n_0}^2$ and $\sigma_{t_0}^2$are the same, and have the same distribution. Thus, since $\sigma_{n_0}^2 \sim \Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$, and the Feller condition holds for σ_t^2, $t \geq 0$, the process σ_n^2 never touches zero. Moreover, since ∞ is a natural boundary

for both processes, the limiting distribution of σ_n^2 is a unique stationary distribution, and the process σ_n^2, $n = 0, 1, 2, ...$, is strongly ergodic, see Karlin and Taylor (1981), pp. 240-241. Furthermore, with $j = 1$, equation (4.14) is the one-step transition law of σ_n^2, $n = 0, 1, 2, ...$, so the process is Markov. Also, we proved above that σ_n^2, $n = 0, 1, 2, ...$, has a unique stationary distribution, and this is given by $\overline{\sigma}_n^2 \sim \Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$. Thus, since $\sigma_{n_0}^2 \sim \Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$ has this unique stationary distribution, every pair of finite dimensional distributions must be the same, and σ_n^2, $n = 0, 1, 2, ...$, is strictly stationary, see Grimmett and Stirzaker (2001), p. 362. ■

Remark 4.2 (Summary of AML Assumptions) *Let $Z_n = \left(Y_n, \sigma_n^2\right)$ be the discrete-time estimation model under AML. The following assumptions are made, see Bates (2006) pp. 912-13.*

1. *The joint process $Z_n = \left(Y_n, \sigma_n^2\right)$ is Markov.*

2. *The conditional joint C.F. $\phi_{Z_{n+1}|Z_n}(u, v)$, with gap time $\tau = \Delta t$, is exponentially affine in Y_n, and σ_n^2 such that*

$$\begin{aligned}\phi_{Z_{n+1}|Z_n}(u, v) &= E\left[e^{iuY_{n+1}+iv\sigma_{n+1}^2}|\left(Y_n, \sigma_n^2\right)\right] \\ &= \exp\left[iuY_n + \underline{C}(u, v; \Delta t) + \underline{D}(u, v; \Delta t)\sigma_n^2\right]. \end{aligned} \tag{4.18}$$

3. *The latent factor σ_n^2 is strictly stationary and strongly ergodic.*

4.2.3 The Gamma Approximation

The Bates (2006) approximation to the joint conditional C.F. of equations (4.11) and (4.12) requires the following. Let

$$Y_n^H = \{Y_1, Y_2, ..., Y_n\} \tag{4.19}$$

be the history of the log-price Y_n at time n. Then for $n \geq m$, let

$$G_{\sigma_n^2|Y_m^H}(v) = E\left[e^{v\sigma_n^2}|Y_m^H\right] \tag{4.20}$$

be the conditional moment generating function of the latent factor σ_n^2 given the log-price history Y_m^H at time $m \leq n$. Hence, we find in the both the SVJ and SVSJ cases, since $Z_n = \left(Y_n, \sigma_n^2\right)$ has the Markov property, that the affine structure in equation (4.12), combined with the definition of the conditional MGF in (4.20), leads by iterated expectations to

$$\begin{aligned}
&\phi_{Z_{n+1}|Y_n^H}(u,v) = E\left[e^{iuY_{n+1}+iv\sigma_{n+1}^2}|Y_n^H\right] \\
&= E\left[E\left[e^{iuY_{n+1}+iv\sigma_{n+1}^2}|\left(Y_n,\sigma_n^2\right)\right]|Y_n^H\right] \\
&= E\left[e^{iuY_n+\underline{C}(u,v;\Delta t)+\underline{D}(u,v;\Delta t)\sigma_n^2}|Y_n^H\right] \\
&= e^{iuY_n+\underline{C}(u,v;\Delta t)}G_{\sigma_n^2|Y_n^H}\left[\underline{D}(u,v;\Delta t)\right]. \qquad (4.21)
\end{aligned}$$

Thus, in both the SVJ and SVSJ cases, equation (4.21) provides a joint conditional C.F. for $Z_{n+1} = \left(Y_{n+1}, \sigma_{n+1}^2\right)$ at time $n+1$, given only the log-price history Y_n^H at time n.

We now define

$$y_{n+1} = \Delta Y_{n+1} = Y_{n+1} - Y_n \qquad (4.22)$$

to be the log-return process at time $n+1$, and

$$y_n^H = \{Y_0, y_1, y_2, ..., y_n\} \qquad (4.23)$$

to be the log-return history at time n, augmented by the initial log-price. Observe that the log-price history Y_n^H and the augmented log-return history y_n^H each contain the same information. This implies that

$$\sigma_n^2|Y_n^H = \sigma_n^2|y_n^H. \qquad (4.24)$$

Thus, for the rest of this thesis we will condition on the augmented log-return history instead, and for simplicity we will refer to y_n^H in equation (4.23) simply as the log-return history. Furthermore, for convenience we will denote the conditional moment generating function of the latent factor

σ_n^2 given the log-return history y_n^H at time n by

$$G_{n|n}(v) = E\left[e^{v\sigma_n^2}|y_n^H\right]. \tag{4.25}$$

Hence, by equation (4.21) for the SVJ and SVSJ models, using equations (4.22) to (4.25) above, we obtain

$$E\left[e^{iuy_{n+1}+iv\sigma_{n+1}^2}|y_n^H\right] = e^{\underline{C}(u,v;\Delta t)}G_{n|n}\left[\underline{D}(u,v;\Delta t)\right] \tag{4.26}$$

as the joint conditional C.F. of the log-return and the latent factor at time $n+1$, given only the log-return history at time n. Thus, we define

$$\begin{aligned}\underline{F}\left(u,v|y_n^H\right) &= E\left[e^{iuy_{n+1}+iv\sigma_{n+1}^2}|y_n^H\right] \\ &= e^{\underline{C}(u,v;\Delta t)}G_{n|n}\left[\underline{D}(u,v;\Delta t)\right].\end{aligned} \tag{4.27}$$

For the SVJ and SVSJ models, the distribution of $\sigma_n^2|y_n^H$ defining $G_{n|n}(v)$ in equation (4.25), and $\underline{F}\left(u,v|y_n^H\right)$ in equation (4.27) is not suitably tractable for practical purposes. Thus, as an approximation, we assume that $\sigma_n^2|y_n^H$ has a gamma distribution, see Bates (2006), pp. 919-921. In subsection 4.2.4 we reveal that this choice has a Bayesian interpretation. The distribution of $\sigma_n^2|y_n^H$ can be viewed as a prior, with the distribution of $\sigma_{n+1}^2|y_{n+1}^H$ being the posterior. Thus, it is practical to choose a conjugate prior, for example a gamma prior, see Bates (2006), p. 920. In view of this, recall also that the latent factor is modeled by the CIR process in both the SVJ and SVSJ models. Moreover, by equation (4.16) above, the limiting distribution of σ_n^2, $n = 0, 1, 2, ...$, is $\Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$. This further suggests that the gamma distribution with C.F. given by

$$G_{n|n}(v) = \left(\frac{1}{1-b_n v}\right)^{a_n}, \tag{4.28}$$

for some prior gamma parameters (a_n, b_n), is a reasonable approximation for the conditional distribution of $\sigma_n^2|y_n^H$. However, note that the affine

coefficient $\underline{D}(u,v;\Delta t)$ is a complex-valued function. This means that $G_{n|n}[\underline{D}(u,v;\Delta t)]$, as it appears in equation (4.27), is a generalized transform in the sense that it is an MGF with a parameter that has both real and imaginary parts. Hence, we need to show the values of u and v for which $G_{n|n}[\underline{D}(u,v;\Delta t)]$ is well defined. We show in Lemma 4.6 of subsection 4.3.1 on L^1 transforms below, that $G_{n|n}[\underline{D}(u,v;\Delta t)]$ is continuous and bounded for all $(u,v) \in \mathbb{R} \times (-\delta,\delta)$, for some $\delta > 0$.

4.2.4 Bayesian Updating and Volatility Filtering

The following theorem and its proof are adapted from Bates (2006), pp. 915-916. This is regarded as an extension of Bayes' rule. We will let $p_Y(y)$ denote the density of a continuous random variable Y.

Theorem 4.3 (Bartlett (1938), p. 63) *If it exists, the partial inverse Fourier transform in the u-variable of the joint characteristic function*

$$\phi_{Y,X}(u,v) = E\left[e^{iuY+ivX}\right] \tag{4.29}$$

may be written as

$$\left[\phi_{X|Y}(v)\, p_Y(y)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy}\phi_{Y,X}(u,v)\,du, \tag{4.30}$$

where $\phi_{X|Y}(v)$ is the conditional characteristic function of X given Y, and $p_Y(y)$ is the density of Y.

Proof. Since the conditional C.F. is defined by

$$\begin{aligned}\phi_{X|Y}(v) &= E\left[e^{ivX}|Y\right] \\ &= \int_{-\infty}^{+\infty} e^{ivx} p_{X|Y}(x|y)\,dx,\end{aligned} \tag{4.31}$$

we have by Bayes' rule that the joint C.F. may be written as

$$\phi_{Y,X}(u,v) = E\left[e^{iuY+ivX}\right]$$

$$
\begin{aligned}
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{iuy+ivx} p_{Y,X}(y,x)\, dydx \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{iuy+ivx} p_{X|Y}(x|y)\, p_Y(y)\, dxdy \\
&= \int_{-\infty}^{+\infty} e^{iuy} \left[\int_{-\infty}^{+\infty} e^{ivx} p_{X|Y}(x|y)\, dx \right] p_Y(y)\, dy \\
&= \int_{-\infty}^{+\infty} e^{iuy} \phi_{X|Y}(v)\, p_Y(y)\, dy. \qquad (4.32)
\end{aligned}
$$

Thus, since the Fourier inversion theorem holds by hypothesis, we have

$$
\left[\phi_{X|Y}(v)\, p_Y(y)\right] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuy} \phi_{Y,X}(u,v)\, du,
$$

as required. ■

In this subsection we introduce the three inverse Fourier transform pairs of the AML method. Recall from equation (4.27) in subsection 4.2.3 that under the gamma approximation from equation (4.28),

$$
\begin{aligned}
\underline{F}\left(u,v|y_n^H\right) &= E\left[e^{iuy_{n+1}+iv\sigma_{n+1}^2}|y_n^H\right] \qquad (4.33) \\
&= e^{\underline{C}(u,v;\Delta t)} G_{n|n}\left[\underline{D}(u,v;\Delta t)\right] \qquad (4.34) \\
&= e^{\underline{C}(u,v;\Delta t)} \left[1 - b_n \underline{D}(u,v;\Delta t)\right]^{-a_n}, \qquad (4.35)
\end{aligned}
$$

is the joint characteristic function of the time $n+1$ log-return and latent factor, conditional on the time n log-return history. This leads to our first inverse Fourier transform pair. As in Bates (2006), p. 917, assuming that Fourier inversion holds, we obtain

$$
p\left(y_{n+1}|y_n^H\right) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u,0|y_n^H\right) du, \qquad (4.36)
$$

as the time $n+1$ log-return density, conditional on the history y_n^H. Equation (4.36) only partially defines the likelihood. The sequence of gamma parameters (a_n, b_n), defining $\underline{F}\left(u,v|y_n^H\right)$ for $n = 1, 2, ..., N_y$ needs to be updated. We proceed as follows. By shifting time n to $n+1$ in equation (4.25) we

obtain

$$\phi_{n+1|n+1}(v) = G_{n+1|n+1}(iv) = E\left[e^{iv\sigma^2_{n+1}}|y^H_{n+1}\right], \tag{4.37}$$

as the time $n+1$ characteristic function of the latent factor, conditional on the history y^H_n. Hence, using equation (4.33) for $\underline{F}\left(u, v|y^H_n\right)$, equation (4.36) for $p\left(y_{n+1}|y^H_n\right)$, and equation (4.37) above, if Fourier inversion holds, then after a minor re-arrangement Theorem 4.3 implies

$$\phi_{n+1|n+1}(v) = \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u, v|y^H_n\right)du}{p\left(y_{n+1}|y^H_n\right)} \tag{4.38}$$

$$= \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}+\underline{C}(u,v;\Delta t)}G_{n|n}\left[\underline{D}(u, v; \Delta t)\right]du}{p\left(y_{n+1}|y^H_n\right)}, \tag{4.39}$$

see Bates (2006), p. 918. Note, equation (4.39) follows from the definition of $\underline{F}\left(u, v|y^H_n\right)$ in equation (4.34). From Bates (2006), p. 918, equations (4.38) and (4.39) form a Bayesian updating mechanism. At time n under the gamma approximation, $\sigma^2_n|y^H_n$ has a prior $\Gamma(a_n, b_n)$ distribution, with a conditional gamma MGF $G_{n|n}(v) = E\left[e^{v\sigma^2_n}|y^H_n\right]$. Then, at time $n+1$, $\sigma^2_{n+1}|y^H_{n+1}$ has a posterior $\Gamma(a_{n+1}, b_{n+1})$ distribution, with a conditional MGF $G_{n+1|n+1}(v) = E\left[e^{v\sigma^2_{n+1}}|y^H_{n+1}\right]$, or equivalently a conditional C.F. $\phi_{n+1|n+1}(v) = E\left[e^{iv\sigma^2_{n+1}}|y^H_{n+1}\right]$. The posterior parameters may be obtained through equations (4.38) and (4.39) given the prior, and these equations are a realization of Theorem 4.3 above, which is proven by Bayes' rule. The filtered volatility squared is the posterior mean given by

$$\widetilde{\sigma}^2_{n+1} = E\left[\sigma^2_{n+1}|y^H_{n+1}\right] = a_{n+1}b_{n+1}. \tag{4.40}$$

Since the prior and the posterior are both gamma, it is recommended in Bates (2006), pp. 919-920 and pp. 955-956, that the updated posterior parameters (a_{n+1}, b_{n+1}) be obtained by the method of moments using an MGF version of equations (4.38) and (4.39). However, in this thesis we obtain the requisite moments from equations (4.38) and (4.39) using the

conditional C.F. instead of the conditional MGF. This choice implies that $\underline{F}\left(u, v|y_n^H\right)$, as it appears in the right hand side of equation (4.38), is a joint conditional C.F. rather than a mixed transform. This has some theoretical advantages, although the two methods are equivalent in practice. Since all moments exist for the gamma distribution, the noncentral moments obtained from the gamma C.F. in equation (4.38) are governed by

$$i^k m'_{k,n+1} = \frac{\partial^k}{\partial v^k}\phi_{n+1|n+1}\left(v\right)|_{v=0}, \qquad (4.41)$$
$$k = 1, 2, ...,$$

see Grimmett and Stirzaker (2001), p. 183. We only need the first two noncentral moments. However, we will need to differentiate under the Fourier integral twice on the right hand side of equation (4.38) to obtain

$$\frac{\partial}{\partial v}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u, v|y_n^H\right) du \qquad (4.42)$$
$$= \int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u, v|y_n^H\right) du$$

and

$$\frac{\partial^2}{\partial v^2}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u, v|y_n^H\right) du \qquad (4.43)$$
$$= \int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u, v|y_n^H\right) du.$$

We prove in Appendix D of this thesis that for both the SVJ model and the SVSJ model, differentiation under the integral holds in equations (4.42) and (4.43) above. Hence, the remaining two inverse Fourier transform pairs are both simply re-arrangements of equation (4.38), differentiated on both sides with respect to v. The first case is the first derivative, and the second case is the second derivative. However, to show that Fourier inversion holds in each case, one needs to arrange the equations such that it is clear what needs to be L^1 integrable. A detailed treatment of L^1 integrability for each

inverse Fourier transform pair is deferred until subsection 4.3.1. For the rest of this thesis we denote the following derivatives by

$$\phi^{(1)}_{n+1|n+1}(v) = \frac{\partial}{\partial v}\phi_{n+1|n+1}(v), \tag{4.44}$$

$$\text{and } \phi^{(2)}_{n+1|n+1}(v) = \frac{\partial^2}{\partial v^2}\phi_{n+1|n+1}(v). \tag{4.45}$$

Again, since $\underline{F}\left(u, v|y_n^H\right) = E\left[e^{iuy_{n+1}+iv\sigma^2_{n+1}}|y_n^H\right]$ is a joint characteristic function, if Fourier inversion holds, then similar to equation (4.38), we have by Theorem 4.3 that

$$\left[\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u, v|y_n^H\right) du. \tag{4.46}$$

But, by equation (4.36), $p\left(y_{n+1}|y_n^H\right)$ does not depend on v. Hence, assuming we can differentiate twice under the integral as in equations (4.42) and (4.43), and that Fourier inversion holds in each case, equation (4.46) implies

$$\left[\phi^{(1)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u, v|y_n^H\right) du, \text{ and} \tag{4.47}$$

$$\left[\phi^{(2)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u, v|y_n^H\right) du. \tag{4.48}$$

Equations (4.47) and (4.48) are, respectively, the second and third inverse transform pairs of the AML method. Moreover, by equation (4.41), estimates for the first two noncentral gamma posterior moments, $\widehat{m}'_{1,n+1}$ and $\widehat{m}'_{2,n+1}$, of the AML filtration process may be obtained from

$$i\widehat{m}'_{1,n+1} = \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u, 0|y_n^H\right) du}{p\left(y_{n+1}|y_n^H\right)}, \text{ and} \tag{4.49}$$

$$-\widehat{m}'_{2,n+1} = \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u, 0|y_n^H\right) du}{p\left(y_{n+1}|y_n^H\right)}. \tag{4.50}$$

We are now in a position to finalize the updating mechanism. Recall that by equation (4.34), under the gamma approximation of equation (4.28),

$$\begin{aligned}\underline{F}\left(u,v|y_n^H\right) &= E\left[e^{iuy_{n+1}+iv\sigma_{n+1}^2}|y_n^H\right] \\ &= e^{\underline{C}(u,v;\Delta t)}G_{n|n}\left[\underline{D}\left(u,v;\Delta t\right)\right] \\ &= e^{\underline{C}(u,v;\Delta t)}\left(\frac{1}{1-b_n\underline{D}\left(u,v;\Delta t\right)}\right)^{a_n}. \end{aligned} \tag{4.51}$$

Expressions for $\underline{F}_v\left(u,v|y_n^H\right)$ and $\underline{F}_{vv}\left(u,v|y_n^H\right)$ in terms of the prior gamma parameters (a_n,b_n), the affine coefficients, and their v-derivatives, will be given in subsection 4.3.1 below. The update to the posterior gamma parameters (a_{n+1},b_{n+1}) is made from equations (4.49) and (4.50), by a moment match involving the first two noncentral gamma moments. We set

$$\begin{aligned}\widehat{m}'_{1,n+1} &= a_{n+1}b_{n+1} \\ \widehat{m}'_{2,n+1} &= a_{n+1}b_{n+1}^2\left(1+a_{n+1}\right).\end{aligned}$$

Thus, we obtain

$$b_{n+1} = \frac{\widehat{m}'_{2,n+1}-\left(\widehat{m}'_{1,n+1}\right)^2}{\widehat{m}'_{1,n+1}}, \text{ and } a_{n+1} = \frac{\widehat{m}'_{1,n+1}}{b_{n+1}}. \tag{4.52}$$

4.2.5 Log-Likelihood I

The AML algorithm for the joint likelihood given θ is a subsequence of three inverse Fourier transforms, evaluated sequentially through the data. The first inverse transform is for the current likelihood. The second inverse transform filters the volatility squared, and the third inverse transform allows the parameters (a_n,b_n) of the gamma approximation to be updated. For the SVJ and SVSJ models, $\theta=(\mu,\kappa,\eta,\omega,\rho,\lambda,\alpha,\beta)$, and under AML the initial variance $\sigma_{n_0}^2$ is random. But, given θ, we can specify the distribution of $\sigma_{n_0}^2$. Recall from Remark 4.2 *3.*, that under AML, σ_n^2, $n=0,1,2,...$, must be strictly stationary and strongly ergodic. Moreover, by Lemma 4.1 from subsection 4.2.2, these conditions are met when the Feller condition,

$2\kappa\eta > \omega^2$, holds for the underlying CIR process σ_t^2, $t \geq 0$, and when for σ_n^2 we have that $\sigma_{n_0}^2 \sim \Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$ has the limiting distribution from equation (4.16) in subsection 4.2.2. Given θ, this implies that we should choose

$$(a_0, b_0) = \left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right), \tag{4.53}$$

to initialize the sequence of gamma parameters for the gamma approximation, see Bates (2006), p. 916 and p. 955. Algorithm 4.4 below summarizes the results from this section.

Algorithm 4.4 (Log-Likelihood I) *Given* $\theta = (\mu, \kappa, \eta, \omega, \rho, \lambda, \alpha, \beta)$

$(a_0, b_0) = \left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$

$p\left(y_1|\theta\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_1 + \underline{C}(u,0;\Delta t)}\left[1 - b_0\underline{D}\left(u, 0; \Delta t\right)\right]^{-a_0} du$

BEGIN FOR $n = 1$ *to* $N_y - 1$

$p\left(y_{n+1}|y_n^H, \theta\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u, 0|y_n^H\right) du$

$\widehat{m}'_{1,n+1} = \frac{\frac{-i}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u,0|y_n^H\right)du}{p(y_{n+1}|y_n^H)}$

$\widehat{m}'_{2,n+1} = \frac{\frac{-1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u,0|y_n^H\right)du}{p(y_{n+1}|y_n^H)}$

$b_{n+1} = \frac{\widehat{m}'_{2,n+1} - \left(\widehat{m}'_{1,n+1}\right)^2}{\widehat{m}'_{1,n+1}}$, $a_{n+1} = \frac{\widehat{m}'_{1,n+1}}{b_{n+1}}$

END FOR

$\ell\left(\theta\right) = \ell_0 + \log p\left(y_1|\theta\right) + \sum_{n=1}^{N_y-1}\log p\left(y_{n+1}|y_n^H, \theta\right)$

4.3 L1 Transforms and FFT

For AML to be stable, the Fourier inversion theorem must hold for each transform. Thus, in subsection 4.3.1 we prove that for the SVJ and SVSJ models, the three Fourier transforms of the AML method, and their inverse transforms, are each L^1. Since Appendix D proves that we can differentiate twice under the integral on the right hand side of equation (4.46), equations (4.47) and (4.48) of subsection 4.2.4 hold, and in subsection 4.3.2 we apply FFT. In this section we take the hypothesis $\tau = \Delta t > 0$ to be given.

4.3.1 L1 Transforms in the AML Method

Definitions and results from this subsection are used again in subsection 4.3.2 below, and also for the two proofs that we may differentiate under the integral as in equations (4.42) and (4.43). The proofs for each of the first and second derivatives both appear in Appendix D of this thesis. From equations $(A.12)$ and $(A.13)$ in Appendix A.1 we have

$$\underline{C} = \underline{C}(u,v) = C(u) - \frac{2\kappa\eta}{\omega^2}\log\left(1 - K(u)\,iv\right), \text{ and} \tag{4.54}$$

$$\underline{D} = \underline{D}(u,v) = D(u) + \frac{R(u)\,iv}{1 - K(u)\,iv}, \tag{4.55}$$

where the log-price coefficients $C(u)$ and $D(u)$, given respectively by equations $(A.22)$ and $(A.23)$ of Appendix A.2, do not depend on v. In what follows, the assumption $\rho < 0$ primarily ensures the principal branch continuity of $C(u)$ and $D(u)$, see Proposition B.16. It also arises in Lemma C.11, and Proposition C.12 on the continuity of the joint affine coefficients. The auxiliary variables $K(u)$ and $R(u)$, given respectively by equations $(A.24)$ and $(A.25)$ of Appendix A.2, also do not depend on v. The coefficients $\underline{C} = \underline{C}(u,v)$ and $\underline{D} = \underline{D}(u,v)$, in equations (4.54) and (4.55) above, are the joint affine coefficients for the SVJ and SVSJ models. Moreover, from equations (4.54) and (4.55) above, the first and second derivatives with respect to v of the joint affine coefficients are given by

$$\underline{C}_v = \frac{\partial}{\partial v}\underline{C}(u,v) = \frac{\frac{2\kappa\eta}{\omega^2}K(u)\,i}{1 - K(u)\,iv}, \tag{4.56}$$

$$\underline{C}_{vv} = \frac{\partial^2}{\partial v^2}\underline{C}(u,v) = \frac{-\frac{2\kappa\eta}{\omega^2}K(u)^2}{\left(1 - K(u)\,iv\right)^2}, \tag{4.57}$$

$$\underline{D}_v = \frac{\partial}{\partial v}\underline{D}(u,v) = \frac{R(u)\,i}{\left(1 - K(u)\,iv\right)^2}, \text{ and} \tag{4.58}$$

$$\underline{D}_{vv} = \frac{\partial^2}{\partial v^2}\underline{D}(u,v) = \frac{-2K(u)\,R(u)}{\left(1 - K(u)\,iv\right)^3}. \tag{4.59}$$

Proposition 4.5 *Let $\rho < 0$. Then the first and second derivatives with respect to v of the joint affine coefficients $\underline{C} = \underline{C}(u,v)$ and $\underline{D} = \underline{D}(u,v)$, denoted by $\underline{C}_v$, $\underline{C}_{vv}$, $\underline{D}_v$, and $\underline{D}_{vv}$, in equations (4.56) to (4.59) above, are each continuous and bounded for all $(u,v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$.*

Proof. Under the hypothesis $\rho < 0$, by Lemma C.11, $\operatorname{Re}(1 - K(u)\, iv) > 0$, for all $(u,v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$. Hence, the denominator is never zero in any of the four derivatives, on the domain in question. Moreover, if $\rho < 0$, then by Proposition B.14, items *2.* and *3.*, the auxiliary variables $K(u)$ and $R(u)$ are both continuous for all $u \in \mathbb{R}$. Thus, each of the four derivatives is continuous, for all $(u,v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$. Furthermore, by Lemmas C.9 and C.10, respectively,

$$\lim_{u \to \pm\infty} K(u) = 0 \quad \text{and} \quad \lim_{u \to \pm\infty} R(u) = 0.$$

Thus, given that each derivative is continuous, each of the four derivatives defined in equations (4.56) to (4.59) is also bounded for all $u \in \mathbb{R}$, and for all $|v| < \delta$, for some $\delta > 0$, as required. ■

We now turn our attention to the derivatives with respect to v of the main conditional characteristic function $\underline{F}\left(u, v | y_n^H\right)$. Recall from equation (4.51) that under the gamma approximation in equation (4.28),

$$\underline{F} = \underline{F}\left(u, v | y_n^H\right) = E\left[e^{iuy_{n+1} + iv\sigma_{n+1}^2} | y_n^H\right] \tag{4.60}$$

$$= e^{\underline{C}(u,v)} G_{n|n}\left[\underline{D}(u,v)\right] \tag{4.61}$$

$$= e^{\underline{C}(u,v)} \left(\frac{1}{1 - b_n \underline{D}(u,v)}\right)^{a_n} \tag{4.62}$$

$$= e^{\underline{C}(u,v) - a_n \log[1 - b_n \underline{D}(u,v)]} \tag{4.63}$$

Based on equation (4.63) we define

$$\underline{f} = \underline{f}\left(u, v | y_n^H\right) = \log \underline{F}\left(u, v | y_n^H\right) \tag{4.64}$$

$$= \underline{C}(u,v) - a_n \log\left[1 - b_n \underline{D}(u,v)\right]. \tag{4.65}$$

Taking the derivative with respect to v in equation (4.65) we obtain

$$\underline{f}_{v} = \frac{\partial}{\partial v}\underline{f}\left(u, v|y_n^H\right) = \underline{C}_{v} + \frac{a_n b_n \underline{D}_{v}}{1 - b_n \underline{D}}, \tag{4.66}$$

and for the second derivative we obtain,

$$\underline{f}_{vv} = \frac{\partial^2}{\partial v^2}\underline{f}\left(u, v|y_n^H\right) = \underline{C}_{vv} + \frac{a_n b_n \underline{D}_{v} + a_n b_n^2 \left(\underline{D}_{v}^2 - \underline{D}\underline{D}_{vv}\right)}{\left(1 - b_n \underline{D}\right)^2}. \tag{4.67}$$

It now follows from the definition of $\underline{f}$ in equation (4.64) that

$$\underline{F} = \underline{F}\left(u, v|y_n^H\right) = e^{\underline{f}\left(u,v|y_n^H\right)}, \text{ leading directly to} \tag{4.68}$$

$$\underline{F}_{v} = \frac{\partial}{\partial v}\underline{F}\left(u, v|y_n^H\right) = \underline{f}_{v}\underline{F}, \text{ and} \tag{4.69}$$

$$\underline{F}_{vv} = \frac{\partial^2}{\partial v^2}\underline{F}\left(u, v|y_n^H\right) = \left(\underline{f}_{vv} + \underline{f}_{v}^2\right)\underline{F}. \tag{4.70}$$

Lemma 4.6 *Assume that $\rho < 0$ in the SVJ and SVSJ models. Let $u \in \mathbb{R}$, and let $\delta > 0$. Then for each fixed $v \in (-\delta, \delta)$,*

$$\sup_{u\in\mathbb{R}}\left|G_{n|n}\left[\underline{D}\left(u, v\right)\right]\right| = \left\|G_{n|n}\left[\underline{D}\left(u, v\right)\right]\right\|_{\infty} < \infty, \tag{4.71}$$

where $G_{n|n}\left[\bullet\right]$ has the form of a gamma MGF as in equation (4.28), and $\underline{D}\left(u, v\right)$ is defined in equation (4.55) above. Moreover, $G_{n|n}\left[\underline{D}\left(u, v\right)\right]$ is continuous for all $(u, v) \in \mathbb{R} \times (-\delta, \delta)$.

Proof. By equation (4.28) we have

$$G_{n|n}\left[\underline{D}\left(u, v\right)\right] = \left(\frac{1}{1 - b_n \underline{D}\left(u, v\right)}\right)^{a_n}. \tag{4.72}$$

But, $\rho < 0$. Thus, for all $u \in \mathbb{R}$, and for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$, we have both that $\underline{D}\left(u, v\right)$ is continuous by Proposition C.12 *2.*, and $\operatorname{Re}\underline{D}\left(u, v\right) \leq 0$ by Lemma 3.15. Hence, since $b_n > 0$,

$$\operatorname{Re}\left[1 - b_n \underline{D}\left(u, v\right)\right] > 0,$$

for all $u \in \mathbb{R}$, and for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$. Therefore, since $a_n > 0$, $G_{n|n}\left[\underline{D}\left(u, v; \Delta t\right)\right]$ is continuous for all $(u, v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$, and

$$\left|G_{n|n}\left[\underline{D}\left(u, v; \Delta t\right)\right]\right| = \left(\frac{1}{\left|1 - b_n \underline{D}\left(u, v; \Delta t\right)\right|}\right)^{a_n} < \infty,$$

for all $u \in \mathbb{R}$, and for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$, as required. ■

Lemma 4.7 *Assume that $\rho < 0$ in the SVJ and SVSJ models. Let $u \in \mathbb{R}$, and let $\delta > 0$. Then for each fixed $v \in (-\delta, \delta)$,*

$$\int_{-\infty}^{+\infty} \left|e^{\underline{C}(u,v)}\right| du < \infty. \tag{4.73}$$

Proof. By equation (4.54),

$$\underline{C}\left(u, v\right) = C\left(u\right) - \frac{2\kappa\eta}{\omega^2} \log\left(1 - K\left(u\right) iv\right), \tag{4.74}$$

and, since $\rho < 0$, Lemma C.11 implies that $\mathrm{Re}\left(1 - K\left(u\right) iv\right) > 0$, for all $(u, v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$. Thus,

$$e^{\underline{C}(u,v)} = e^{C(u)} \left(\frac{1}{1 - K\left(u\right) iv}\right)^{\frac{2\kappa\eta}{\omega^2}}. \tag{4.75}$$

But, $\mathrm{Re}\left(1 - K\left(u\right) iv\right) > 0$ on this domain, and $\frac{2\kappa\eta}{\omega^2} > 0$. Hence,

$$\left|\left(\frac{1}{1 - K\left(u\right) iv}\right)^{\frac{2\kappa\eta}{\omega^2}}\right| = \left(\frac{1}{\left|1 - K\left(u\right) iv\right|}\right)^{\frac{2\kappa\eta}{\omega^2}} < \infty, \tag{4.76}$$

for all $u \in \mathbb{R}$, and for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$. Moreover, recall from subsection 3.3.3 that if $\rho < 0$, then $e^{C(u)} \in L^1$ by Lemma 3.16. Therefore, by Hölder's inequality,

$$\begin{aligned}\int_{-\infty}^{+\infty} \left|e^{\underline{C}(u,v)}\right| du &= \int_{-\infty}^{+\infty} \left|e^{C(u)} \left(\frac{1}{1 - K\left(u\right) iv}\right)^{\frac{2\kappa\eta}{\omega^2}}\right| du \\ &\leq \left\|e^{C(u)}\right\|_1 \left\|\left(\frac{1}{1 - K\left(u\right) iv}\right)^{\frac{2\kappa\eta}{\omega^2}}\right\|_\infty < \infty,\end{aligned}$$

for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$, as required. ■

Lemma 4.8 *Assume that $\rho < 0$ in the SVJ and SVSJ models. From equation* (4.61)*, define*

$$\underline{F}\left(u, v|y_n^H\right) = e^{\underline{C}(u,v)} G_{n|n}\left[\underline{D}\left(u, v\right)\right],$$

where $G_{n|n}\left[\underline{D}\left(u, v\right)\right]$ is defined as in Lemma 4.6 above. Then, for some $\delta > 0$, we have both that $\underline{F}\left(u, v|y_n^H\right)$ is continuous on the principal branch for all $(u, v) \in \mathbb{R} \times (-\delta, \delta)$, and

$$\int_{-\infty}^{+\infty} \left|\underline{F}\left(u, v|y_n^H\right)\right| du < \infty, \text{ for each fixed } v \in (-\delta, \delta). \tag{4.77}$$

In the latter case we shall say that $\underline{F}\left(u, v|y_n^H\right)$ is L^1 in the u-variable.

Proof. Since $\rho < 0$, by Proposition C.12, for all $(u, v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$, $\underline{C}\left(u, v\right)$ is principal branch continuous, and $\underline{D}\left(u, v\right)$ is continuous. Therefore, on this domain, clearly $e^{\underline{C}(u,v)}$ is principal branch continuous, and, since $\rho < 0$, $G_{n|n}\left[\underline{D}\left(u, v\right)\right]$ is also continuous, by Lemma 4.6 above. Hence, $\underline{F}\left(u, v|y_n^H\right)$ is principal branch continuous, for all $(u, v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\dot{\delta} > 0$. Moreover, since $\rho < 0$, Lemma 4.7 implies that $e^{\underline{C}(u,v)} \in L^1$ in the u-variable for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$. Also, since $\rho < 0$, Lemma 4.6 implies that if $u \in \mathbb{R}$, then $\left\|G_{n|n}\left[\underline{D}\left(u, v\right)\right]\right\|_\infty < \infty$, for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$. Hence, by Hölder's inequality,

$$\begin{aligned}\int_{-\infty}^{+\infty} \left|\underline{F}\left(u, v|y_n^H\right)\right| du &= \int_{-\infty}^{+\infty} \left|e^{\underline{C}(u,v)} G_{n|n}\left[\underline{D}\left(u, v\right)\right]\right| du \\ &\leq \left\|e^{\underline{C}(u,v)}\right\|_1 \left\|G_{n|n}\left[\underline{D}\left(u, v\right)\right]\right\|_\infty < \infty,\end{aligned}$$

for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$, as required. ■

Proposition 4.9 *Assume that $\rho < 0$ in the SVJ and SVSJ models. As in equations* (4.66) *and* (4.67)*, define*

$$\begin{aligned}\underline{f}_v &= \frac{\partial}{\partial v}\underline{f}\left(u, v|y_n^H\right) = \underline{C}_v + \frac{a_n b_n \underline{D}_v}{1 - b_n \underline{D}}, \text{ and} \\ \underline{f}_{vv} &= \frac{\partial^2}{\partial v^2}\underline{f}\left(u, v|y_n^H\right) = \underline{C}_{vv} + \frac{a_n b_n \underline{D}_{vv} + a_n b_n^2\left(\underline{D}_v^2 - \underline{D}\underline{D}_{vv}\right)}{\left(1 - b_n \underline{D}\right)^2},\end{aligned}$$

where $\underline{f}\left(u,v|y_n^H\right) = \log \underline{F}\left(u,v|y_n^H\right)$. *Then, for some* $\delta > 0$, *each of* $\underline{f}_v\left(u,v|y_n^H\right)$ *and* $\underline{f}_{vv}\left(u,v|y_n^H\right)$ *is continuous for all* $(u,v) \in \mathbb{R} \times (-\delta,\delta)$. *Moreover, if we let* $u \in \mathbb{R}$, *then*

$$\left\|\underline{f}_v\left(u,v|y_n^H\right)\right\|_\infty < \infty, \text{ and}$$
$$\left\|\underline{f}_{vv}\left(u,v|y_n^H\right)\right\|_\infty < \infty, \text{ both for each fixed } v \in (-\delta,\delta).$$

Proof. By the gamma approximation, $b_n > 0$. Moreover, since $\rho < 0$, by Lemma 3.15 $\operatorname{Re}\underline{D} \leq 0$, for all $(u,v) \in \mathbb{R} \times (-\delta,\delta)$, for some $\delta > 0$. Therefore,

$$\operatorname{Re}\left(1 - b_n\underline{D}\right) > 0,$$

in the denominator of the second term for each of $\underline{f}_v$ and $\underline{f}_{vv}$. Furthermore, since $\rho < 0$, by Proposition 4.5, each derivative $\underline{C}_v$, $\underline{C}_{vv}$, $\underline{D}_v$, and $\underline{D}_{vv}$, used in the definitions of $\underline{f}_v$ and $\underline{f}_{vv}$ above, is continuous and bounded for all $(u,v) \in \mathbb{R} \times (-\delta,\delta)$, for some $\delta > 0$. Thus, $\underline{f}_v$ and $\underline{f}_{vv}$ are both continuous and bounded, for all $(u,v) \in \mathbb{R} \times (-\delta,\delta)$, for some $\delta > 0$, as required. ■

Proposition 4.10 *Assume that* $\rho < 0$ *in the SVJ and SVSJ models. As in equations* (4.69) *and* (4.70), *define*

$$\underline{F}_v = \frac{\partial}{\partial v}\underline{F}\left(u,v|y_n^H\right) = \underline{f}_v\underline{F}, \text{ and}$$
$$\underline{F}_{vv} = \frac{\partial^2}{\partial v^2}\underline{F}\left(u,v|y_n^H\right) = \left(\underline{f}_{vv} + \underline{f}_v^2\right)\underline{F},$$

where $\underline{F} = \underline{F}\left(u,v|y_n^H\right)$. *Then, each of* $\underline{F}_v\left(u,v|y_n^H\right)$ *and* $\underline{F}_{vv}\left(u,v|y_n^H\right)$ *is continuous on the principal branch for all* $(u,v) \in \mathbb{R} \times (-\delta,\delta)$, *for some* $\delta > 0$. *Moreover,*

$$\int_{-\infty}^{+\infty} \left|\underline{F}_v\left(u,v|y_n^H\right)\right| du < \infty, \text{ and}$$
$$\int_{-\infty}^{+\infty} \left|\underline{F}_{vv}\left(u,v|y_n^H\right)\right| du < \infty, \text{ both for each fixed } v \in (-\delta,\delta).$$

That is, $\underline{F}_v$ *and* $\underline{F}_{vv}$ *are both* L^1 *in the* u*-variable.*

Proof. Consider $(u,v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$. Since $\rho < 0$, we have by Lemma 4.8 that $\underline{F}$ is continuous on the principal branch, and by Proposition 4.9 both $\underline{f}_v$ and $\underline{f}_{vv}$ are continuous, on this domain. Therefore, clearly both $\underline{F}_v$ and $\underline{F}_{vv}$ are continuous on the principal branch, for all $(u,v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$. Moreover, since $\rho < 0$, on this domain $\underline{F}$ is L^1 in the u-variable, by Lemma 4.8, and both $\left\| \underline{f}_v \right\|_\infty$ and $\left\| \underline{f}_{vv} \right\|_\infty$ are finite by Proposition 4.9. Therefore, by Hölder's inequality,

$$\begin{aligned} \int_{-\infty}^{+\infty} \left| \underline{F}_v \left(u, v | y_n^H \right) \right| du &= \int_{-\infty}^{+\infty} \left| \underline{f}_v \underline{F} \right| du \\ &\leq \| \underline{F} \|_1 \left\| \underline{f}_v \right\|_\infty < \infty, \end{aligned}$$

and again by Hölder's inequality,

$$\begin{aligned} &\int_{-\infty}^{+\infty} \left| \underline{F}_{vv} \left(u, v | y_n^H \right) \right| du = \int_{-\infty}^{+\infty} \left| \left(\underline{f}_{vv} + \underline{f}_v^2 \right) \underline{F} \right| du \\ &\leq \| \underline{F} \|_1 \left\| \underline{f}_{vv} + \underline{f}_v^2 \right\|_\infty \\ &\leq \| \underline{F} \|_1 \left(\left\| \underline{f}_{vv} \right\|_\infty + \left\| \underline{f}_v \right\|_\infty^2 \right) < \infty, \end{aligned}$$

in both cases, for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$, as required. ■

We have shown that the three Fourier transforms of the AML method, $\underline{F}$, $\underline{F}_v$, and $\underline{F}_{vv}$, respectively are, for some $\delta > 0$, each continuous for all $(u,v) \in \mathbb{R} \times (-\delta, \delta)$, and each L^1 in the u-variable, for each fixed $v \in (-\delta, \delta)$. The respective inverse transforms will now be shown to be both continuous and L^1 within the context of the following two propositions. These two propositions complete the theory of the AML method.

Recall from equation (4.38) in subsection 4.2.4 that by Theorem 4.3, under the assumption that Fourier inversion holds,

$$\left[\phi_{n+1|n+1}(v)\, p\left(y_{n+1} | y_n^H \right) \right] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v | y_n^H \right) du. \tag{4.78}$$

Proposition 4.11 (Density Transform) *Assume that $\rho < 0$ in the SVJ and SVSJ models. Then, for some $\delta > 0$, the following hold.*

1. $\int_{-\infty}^{+\infty} \left|\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right| dy_{n+1} < \infty$, *for each fixed* $v \in (-\delta, \delta)$. *That is,* $\left[\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right]$ *is* L^1 *in* y_{n+1}.

2. $\left[\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du$ *a.e., and* $\left[\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right]$ *is continuous for all* $(v, y_{n+1}) \in (-\delta, \delta) \times \mathbb{R}$. *That is, Fourier inversion holds.*

3. $\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du$ *converges uniformly, for every* $v \in (-\delta, \delta)$.

Proof. Let $\delta > 0$, and for each fixed $v \in (-\delta, \delta)$ consider

$$\int_{-\infty}^{+\infty} \left|\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right| dy_{n+1}$$

$$\begin{aligned}
&= \int_{-\infty}^{+\infty} \left| \int_{-\infty}^{+\infty} e^{iv\sigma_{n+1}^2} p\left(\sigma_{n+1}^2|y_n^H\right) d\sigma_{n+1}^2 \right| p\left(y_{n+1}|y_n^H\right) dy_{n+1} \\
&\leq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left| e^{iv\sigma_{n+1}^2} p\left(\sigma_{n+1}^2|y_n^H\right) d\sigma_{n+1}^2 \right| p\left(y_{n+1}|y_n^H\right) dy_{n+1} \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p\left(\sigma_{n+1}^2|y_n^H\right) d\sigma_{n+1}^2 p\left(y_{n+1}|y_n^H\right) dy_{n+1} \\
&= \int_{-\infty}^{+\infty} p\left(\sigma_{n+1}^2|y_n^H\right) d\sigma_{n+1}^2 \int_{-\infty}^{+\infty} p\left(y_{n+1}|y_n^H\right) dy_{n+1} = 1 < \infty.
\end{aligned}$$

This proves *1.* But, since $\rho < 0$, by Lemma 4.8, $\underline{F}$ is L^1 in the u-variable, for each fixed $v \in (-\delta, \delta)$. Hence, since $\left[\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right]$ is also L^1 in y_{n+1} by *1.*, Theorem 3.6, the Fourier inversion theorem, implies *2.* Lastly, since we have established both that $\underline{F}$ is L^1 in the u-variable, for each fixed $v \in (-\delta, \delta)$, and $\left[\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right]$ is L^1 in y_{n+1}, Lemma 3.8 from subsection 3.2.1 implies that

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du \text{ converges uniformly,}$$

for every $v \in (-\delta, \delta)$, for some $\delta > 0$, as required. ■

Remark 4.12 *In Proposition 4.11 above,* $\left[\phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right]$ *is* L^1 *in* y_{n+1}*, and* $\underline{F}\left(u,v|y_n^H\right)$ *is the corresponding Fourier transform in* u*. Thus, by Lemma 3.3,* $\underline{F}\left(u,v|y_n^H\right)$ *is uniformly continuous in* u*, for all* $u \in \mathbb{R}$*, for each* $v \in (-\delta,\delta)$*, for some* $\delta > 0$*.*

Remark 4.13 *The results from Proposition 4.11 above hold when* $v = 0$*. That is, since* $p\left(y_{n+1}|y_n^H\right) \in L^1$*, when* $\underline{F}\left(u,0|y_n^H\right)$ *is also* L^1

$$p\left(y_{n+1}|y_n^H\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u,0|y_n^H\right)du \quad a.e., \tag{4.79}$$

and $p\left(y_{n+1}|y_n^H\right)$ *is continuous, by the Fourier inversion theorem.*

We now show that Fourier inversion holds for the two filtering transforms. Recall from equations (4.47) and (4.48) in subsection 4.2.4, that if the right hand side of equation (4.78) above can be differentiated twice under the integral, and the Fourier inversion theorem holds in each case, then we obtain the two Fourier transform pairs

$$\left[\phi^{(1)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u,v|y_n^H\right)du, \text{ and} \tag{4.80}$$

$$\left[\phi^{(2)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u,v|y_n^H\right)du. \tag{4.81}$$

Proposition 4.14 (Filtering Transforms) *Assume that* $\rho < 0$ *in the SVJ and SVSJ models. Further assume that from equation* (4.37)*,*

$$\phi_{n+1|n+1}(v) = \left(\frac{1}{1-ivb_{n+1}}\right)^{a_{n+1}},$$

by the gamma approximation of subsection 4.2.3. Then, for some $\delta > 0$*, the following hold.*

1. *Each of* $\left[\phi^{(1)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right]$ *and* $\left[\phi^{(2)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right]$ *is* L^1 *in the variable* y_{n+1} *for each fixed* $v \in (-\delta,\delta)$*.*

2. $\frac{\partial}{\partial v}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u,v|y_n^H\right)du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u,v|y_n^H\right)du$, *holds for every* $v\in(-\delta,\delta)$. *Moreover,*

$$\left[\phi^{(1)}_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u,v|y_n^H\right)du \text{ a.e.},$$

and the left hand side is continuous for all $(v,y_{n+1})\in(-\delta,\delta)\times\mathbb{R}$. *That is, Fourier inversion holds.*

3. $\frac{\partial^2}{\partial v^2}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u,v|y_n^H\right)du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u,v|y_n^H\right)du$, *holds for every* $v\in(-\delta,\delta)$. *Moreover,*

$$\left[\phi^{(2)}_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u,v|y_n^H\right)du \text{ a.e.},$$

and the left hand side is continuous for all $(v,y_{n+1})\in(-\delta,\delta)\times\mathbb{R}$. *That is, Fourier inversion holds.*

Proof. To prove *1.*, observe that

$$\begin{aligned}&\frac{\partial}{\partial v}\left[\phi_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right] = \left[\phi^{(1)}_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right]\\ &= \frac{ia_{n+1}b_{n+1}}{1-ivb_{n+1}}\left[\phi_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right], \text{ and}\end{aligned} \tag{4.82}$$

$$\begin{aligned}&\frac{\partial^2}{\partial v^2}\left[\phi_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right] = \left[\phi^{(2)}_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right]\\ =&\frac{-(1+a_{n+1})\,a_{n+1}b_{n+1}^2}{(1-ivb_{n+1})^2}\left[\phi_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right].\end{aligned} \tag{4.83}$$

But, $\mathrm{Re}\,(1-ivb_{n+1})>0$, for all $v\in(-\delta,\delta)$, and for any $\delta>0$. Hence, $\frac{ia_{n+1}b_{n+1}}{1-ivb_{n+1}}$ and $\frac{-(1+a_{n+1})a_{n+1}b_{n+1}^2}{(1-ivb_{n+1})^2}$ are continuous and bounded for all $v\in\mathbb{R}$, and neither of them depends on the variable y_{n+1}. However, since $\rho<0$, by Proposition 4.11, item *1.*, for some $\delta>0$, $\left[\phi_{n+1|n+1}(v)\,p\left(y_{n+1}|y_n^H\right)\right]$ is L^1 in the variable y_{n+1} for each fixed $v\in(-\delta,\delta)$. Therefore, by equations

(4.82) and (4.83) above, for each fixed $v \in (-\delta, \delta)$, for some $\delta > 0$,

$$\int_{-\infty}^{+\infty} \left| \phi^{(1)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right) \right| dy_{n+1}$$
$$= \frac{ia_{n+1}b_{n+1}}{1 - ivb_{n+1}} \int_{-\infty}^{+\infty} \left| \phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right) \right| dy_{n+1} < \infty,$$
$$\text{and } \int_{-\infty}^{+\infty} \left| \phi^{(2)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right) \right| dy_{n+1}$$
$$= \frac{-(1 + a_{n+1})\, a_{n+1}b_{n+1}^2}{(1 - ivb_{n+1})^2} \int_{-\infty}^{+\infty} \left| \phi_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right) \right| dy_{n+1} < \infty.$$

For *2.*, since $\rho < 0$, by Lemma 4.8, $\underline{F}$ is L^1 in u, by Remark 4.12, $\underline{F}\left(u, v|y_n^H\right)$ is uniformly continuous in u, for all $u \in \mathbb{R}$, and by Proposition 4.11 item *3.*,

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du \text{ converges uniformly,} \tag{4.84}$$

where each of the above holds for all $v \in (-\delta, \delta)$, for some $\delta > 0$. Therefore, by Theorem D.7 from Appendix D, we have that $\underline{F}_v\left(u, v|y_n^H\right)$ is uniformly continuous in u, for all $u \in \mathbb{R}$,

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u, v|y_n^H\right) du \text{ converges uniformly, and} \tag{4.85}$$

$$\frac{\partial}{\partial v} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u, v|y_n^H\right) du, \tag{4.86}$$

where each of the above holds for all $v \in (-\delta, \delta)$, for some $\delta > 0$. Moreover, since $\rho < 0$, by Proposition 4.10, $\underline{F}_v\left(u, v|y_n^H\right)$ is L^1 in u, for $v \in (-\delta, \delta)$. Hence, since $\left[\phi^{(1)}_{n+1|n+1}(v)\, p\left(y_{n+1}|y_n^H\right)\right]$ is also L^1 in y_{n+1} by *1.*, the rest of *2.* follows from Theorem 3.6, the Fourier inversion theorem. To prove *3.*, since $\rho < 0$, we may use the same hypotheses as for Theorem D.7 above, plus the result $\underline{F}_v$ is L^1 in u from Proposition 4.10. Then by Theorem D.13 from Appendix D we have that.

$$\frac{\partial^2}{\partial v^2} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_{vv}\left(u, v|y_n^H\right) du. \tag{4.87}$$

The rest is similar to the proof of *2.* This completes the proof. ■

4.3.2 Log-Likelihood II (FFT)

In this subsection we describe our FFT based implementation of AML. This is summarized in Algorithm 4.15 below. By letting $v = 0$ in equations (4.78), (4.80), and (4.81) of subsection 4.3.1 above, and by using the result (4.41) from subsection 4.2.4, we obtain

$$p\left(y_{n+1}|y_n^H\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u,0|y_n^H\right)du, \tag{4.88}$$

$$i\widehat{m}'_{1,n+1} = \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u,0|y_n^H\right)du}{p\left(y_{n+1}|y_n^H\right)}, \text{ and} \tag{4.89}$$

$$-\widehat{m}'_{2,n+1} = \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u,0|y_n^H\right)du}{p\left(y_{n+1}|y_n^H\right)}, \tag{4.90}$$

where $\widehat{m}'_{1,n+1}$ and $\widehat{m}'_{2,n+1}$ are the estimates of the first two noncentral gamma posterior moments of the AML filtration process, and $p\left(y_{n+1}|y_n^H\right)$ is simply the conditional log-return density. Alternatively, if we had used the MGF $\phi_{n+1|n+1}\left(-iv\right)$ in equation (4.41) from subsection 4.2.4 to obtain moments instead of the C.F. then we would have

$$\begin{aligned}
p\left(y_{n+1}|y_n^H\right) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u,0|y_n^H\right)du, \\
\widehat{m}'_{1,n+1} &= \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\frac{\partial}{\partial v}\underline{F}\left(u,-iv|y_n^H\right)|_{v=0}du}{p\left(y_{n+1}|y_n^H\right)}, \text{ and} \\
\widehat{m}'_{2,n+1} &= \frac{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\frac{\partial^2}{\partial v^2}\underline{F}\left(u,-iv|y_n^H\right)|_{v=0}du}{p\left(y_{n+1}|y_n^H\right)}.
\end{aligned}$$

While the latter system is somewhat simpler to implement, equations (4.88), (4.89), and (4.90) above maintain the assumption that $\underline{F}\left(u,v|y_n^H\right)$ in each integrand is a joint characteristic function. The system of equations (4.88), (4.89), and (4.90) above is no different from Algorithm 4.4 in subsection 4.2.5. But, we now know from Propositions 4.11 and 4.14, of subsection 4.3.1, that the hypotheses of the Fourier inversion theorem are met in each of equations (4.88), (4.89), and (4.90) above. These two propositions also tell us,

via Theorem D.7 and Theorem D.13 from Appendix D of this thesis, that differentiation under the integral holds in equations (4.42) and (4.43) from subsection 4.2.4, in some neighbourhood of $v = 0$. This defines the right hand side of each of equations (4.80) and (4.81), thus also the right hand sides of equations (4.89) and (4.90). These properties ensure evaluation of equations (4.88), (4.89), and (4.90) above by a numerical integration scheme, such as FFT or quadrature, will be stable.

The FFT scheme proposed in Chapter 3 assumes the domain of integration is $[0, \infty)$. The above three integrals can be reduced to this domain as follows. By the Fourier inversion theorem, Theorem 3.6,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux} \widehat{f}(u)\, du.$$

However, by Corollary 3.7 to Theorem 3.6, from subsection 3.2.1,

$$\text{If } f \text{ is purely real, then } f(x) = \frac{1}{\pi} \operatorname{Re} \int_0^{\infty} e^{-iux} \widehat{f}(u)\, du, \text{ and}$$
$$\text{if } f \text{ is purely imaginary, then } f(x) = \frac{i}{\pi} \operatorname{Im} \int_0^{\infty} e^{-iux} \widehat{f}(u)\, du.$$

Hence, since the left hand sides of (4.88) and (4.90) are purely real, and the left hand side of (4.89) is purely imaginary,

$$p\left(y_{n+1}|y_n^H\right) = \frac{1}{\pi} \operatorname{Re} \int_0^{\infty} e^{-iuy_{n+1}} \underline{F}\left(u, 0|y_n^H\right) du, \tag{4.91}$$

$$\widehat{m}'_{1,n+1} = \frac{\frac{1}{\pi} \operatorname{Im} \int_0^{\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u, 0|y_n^H\right) du}{p\left(y_{n+1}|y_n^H\right)}, \text{ and} \tag{4.92}$$

$$\widehat{m}'_{2,n+1} = \frac{\frac{-1}{\pi} \operatorname{Re} \int_0^{\infty} e^{-iuy_{n+1}} \underline{F}_{vv}\left(u, 0|y_n^H\right) du}{p\left(y_{n+1}|y_n^H\right)}. \tag{4.93}$$

Recall from subsection 3.2.3 that we approximate $g : \mathbb{R} \to \mathbb{C}$ given by

$$g(x) = \int_0^{\infty} e^{-iux} \widehat{f}(u)\, du, \tag{4.94}$$

along the left-endpoint discretization $x_k = -b + (k-1)\,\Delta x,\ k = 1, 2, ..., N$, by the DFT, or equivalently by FFT when $N = 2^p$, such that

$$\widetilde{g}\,(x_k) = \sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j}\,\widehat{f}\,(u_j)\,w_j \Delta u, \qquad (4.95)$$
$$k = 1, 2, ..., N,$$

as in equation (3.22), where $u_j = (j-1)\,\Delta u,\ j = 1, 2, ..., N$. Following Carr, Geman, Madan, and Yor (2002), p. 320, we choose $N = 2^{14}$ and $\Delta u = \frac{1}{4}$. By equations (3.19) and (3.20), of subsection 3.2.3 respectively, this implies $\Delta x = \frac{\pi}{2048}$, and $b = 4\pi$. For the w_j we proposed the trapezoidal rule, $w_1 = w_N = \frac{1}{2}$, $w_j = 1$ otherwise, in subsection 3.4.1.

In Algorithm 4.15 below, we use the FFT approximation in equation (4.95) to evaluate equations (4.91), (4.92), and (4.93) above, thus forming an FFT approximation for the AML likelihood. Since the output of FFT is a grid rather than the exact point of interest, in each step of Algorithm 4.15 below we use a spline interpolant based on a subset,

$$x_k \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right],$$

of the FFT discretized points. The end-points of the above sub-grid are implicitly given by $x_{7680} = -\frac{\pi}{4}$ and $x_{8704} = \frac{\pi}{4}$. These can be obtained, with $\Delta x = \frac{\pi}{2048}$, from the main FFT grid given by,

$$x_k = -4\pi + (k-1)\,\Delta x,\ k = 1, 2, ..., N = 2^{14}.$$

The interior points follow by the uniformity of the grid. Thus, each cubic spline provides a continuous approximation to the function of interest, for $x \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$. Moreover, at each step, our version of AML uses only 1,025 of the 16,384 values computed by FFT. Since 1,025 is a sufficiently small number of points, the cost of the cubic spline interpolant is negligible compared to the cost of FFT.

Algorithm 4.15 (Log-Likelihood II) *FFT Version*

Given $\theta = (\mu, \kappa, \eta, \omega, \rho, \lambda, \alpha, \beta)$

$(a_0, b_0) = \left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$, $\underline{F}_0\left(u, 0|y_n^H\right) = e^{\underline{C}(u,0;\Delta t)}\left[1 - b_0\underline{D}\left(u, 0; \Delta t\right)\right]^{-a_0}$

STEP 0:

$h_0\left(x_k\right) = \frac{1}{\pi}\,\mathrm{Re}\sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)}e^{ibu_j}\underline{F}_0\left(u_j, 0|y_n^H\right)w_j\Delta u,$

$k = 1, 2, ..., N$

$S_0\left(x\right) = Spline\left\{\left(x_k, h_0\left(x_k\right)\right)\right\}_{k=7680}^{8704},\ x \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$

$p\left(y_1|\theta\right) = S_0\left(y_1\right)$

BEGIN FOR $n = 1$ *TO* $N_y - 1$

STEP 1:

$h_1\left(x_k\right) = \frac{1}{\pi}\,\mathrm{Re}\sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)}e^{ibu_j}\underline{F}\left(u_j, 0|y_n^H\right)w_j\Delta u,$

$k = 1, 2, ..., N$

$S_1\left(x\right) = Spline\left\{\left(x_k, h_1\left(x_k\right)\right)\right\}_{k=7680}^{8704},\ x \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$

$p\left(y_{n+1}|y_n^H, \theta\right) = S_1\left(y_{n+1}\right)$

STEP 2:

$h_2\left(x_k\right) = \frac{1}{\pi}\,\mathrm{Im}\sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)}e^{ibu_j}\underline{F}_v\left(u_j, 0|y_n^H\right)w_j\Delta u,$

$k = 1, 2, ..., N$

$S_2\left(x\right) = Spline\left\{\left(x_k, h_2\left(x_k\right)\right)\right\}_{k=7680}^{8704},\ x \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$

$\widehat{m}'_{1,n+1} = \frac{S_2(y_{n+1})}{p(y_{n+1}|y_n^H,\theta)}$

STEP 3:

$h_3\left(x_k\right) = \frac{-1}{\pi}\,\mathrm{Re}\sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)}e^{ibu_j}\underline{F}_{vv}\left(u_j, 0|y_n^H\right)w_j\Delta u,$

$k = 1, 2, ..., N$

$S_3\left(x\right) = Spline\left\{\left(x_k, h_3\left(x_k\right)\right)\right\}_{k=7680}^{8704},\ x \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$

$\widehat{m}'_{2,n+1} = \frac{S_3(y_{n+1})}{p(y_{n+1}|y_n^H,\theta)}$

UPDATE:

$b_{n+1} = \frac{\widehat{m}'_{2,n+1} - \left(\widehat{m}'_{1,n+1}\right)^2}{\widehat{m}'_{1,n+1}},\ a_{n+1} = \frac{\widehat{m}'_{1,n+1}}{b_{n+1}}$

END FOR

$\ell\left(\theta\right) = \ell_0 + \log p\left(y_1|\theta\right) + \sum_{n=1}^{N_y-1}\log p\left(y_{n+1}|y_n^H, \theta\right)$

4.4 Simulation and Comparative Results

4.4.1 Implementation Details

The computations in this chapter were all completed using an Intel Xeon 2xE5-2643v3 dual processor at 3.4 GHz on 12 cores. The code was written and executed in Matlab R2016a running on Windows 10 Pro. For all AML estimations in this chapter, except those reported in subsection 4.4.5 where quadrature was used, the negative log-likelihood was computed by FFT using Algorithm 4.15. The negative log-likelihood was minimized, subject to bound constraints only, by the 'active-set' algorithm of the Matlab function *fmincon*. The 'active-set' algorithm uses sequential quadratic programming (SQP). Under suitable conditions it will behave like a Newton method, see Nocedal and Wright (2006), p. 533. We employed moderately loose bound constraints so that the final Lagrange multipliers were suitably close to zero. Therefore, the final Hessian of the Lagrangian served as close approximation to the final Hessian of the unconstrained problem. Moreover, these bounds were enlarged in the event of any breach, and in this sense, for simplicity, we will refer to the problem as being unconstrained.

Two types of parallel processing were employed. For individual estimation the elements of each gradient were computed in parallel. This allowed us to match, using FFT, a stated average estimation time from Bates (2006), p. 928. For group estimation of simulated data, the Monte Carlo loop was run in parallel. Within a parallel program, both methods cannot be used simultaneously. In subsection 4.4.2 below, and in the remainder of this chapter, exact estimation times, as well as number of iterations, are both tabled for all individual estimation results. Likewise, for group estimation, as in the simulation studies presented in subsection 4.4.3 and subsection 4.4.5 below, the total simulation time is reported. To illustrate the benefits of

parallel processing, we need to compare to the sequential case which does not use parallel processing at all. This case is considered, under the DJIA 1988-2007 daily log-return data, at the top of Table 4.1 below.

Estimation Type	**Model Type**	**Effective Time**	**No. of Obs.**
Sequential FFT	SVSJ Merton	68.95 min	5,042
Parallel Gradient FFT	SVSJ Merton	36.63 min	5,042
Parallel Simulation FFT	SVSJ Merton	14.96 min	5,042

Table 4.1: AML FFT Estimation Times: DJIA 1988-2007.

As shown in Table 4.1 above, for about 5,000 observations with FFT, the advantage of parallel gradients over a sequential program is about a factor of 2. For parallel simulation Table 4.1 reports the total simulation time in parallel divided by number of paths, and the benefit is almost a factor of 5 over a sequential program. However, we find in subsection 4.4.5 that the cost of FFT becomes too high if we double the number of observations. Bates (2006), p. 928, reports an average sequential AML estimation time of 29 minutes for 4,000 log-returns simulated under the SVJ Merton model estimates for the S&P 500, 1980-1999, from Eraker, Johannes, and Polson (2003). Adjusting for the model type, and the number of observations, this is similar to our parallel gradient FFT time reported in Table 4.1 above.

4.4.2 Estimation Results

In preparation for the two simulated AML estimation studies presented in subsection 4.4.3 below, this subsection provides the AML estimates for the true values of the parameters under the SVJ Merton, SVJ Variance Gamma, SVJ Meixner, SVSJ Merton, and Heston models, based on two respective real world data sets. These five models were described in Chapter 2 above. Our trading day data is from Bloomberg L.P. The first data set covers

Apple stock (AAPL) from January 2nd, 1991 to January 19th, 2011. The log-return for September 29th, 2000, is excluded. Thus, there are 5,052 log-returns, and this is the main data set of this thesis. On September 29th, 2000, Apple lost just over half its value in one day due to an unfavourable earnings report. The log-return for that day is -0.7312 which is more than 22 standard deviations away from the mean based on the complete data. We elected to remove this outlier from the data, and this is the only case in which this was done. The second data set covers the Dow-Jones Industrial Average (INDU) from January 4th, 1988 to December 31st, 2007, for a total of 5,042 log-returns. This is the main secondary time period used in this chapter. In order to be consistent with other P-measure estimation studies in the literature, of models similar to the SVJ and SVSJ models, we have chosen a sample size of approximately 5,000 daily log-returns, see for example Eraker, Johannes, and Polson (2003), p. 1279, and Li, Wells, and Yu (2008), p. 2364. For internal consistency we also chose 5,000 daily log-returns for the Heston model, although in the Heston case 2,500 might lead to a better fit, see Lysy (2012) as cited in subsection 4.4.4 below.

For Apple stock (1991-2011), Table 4.2 below reports the AML estimates for the true values of the parameters under each of the five respective models. Similarly, Table 4.3 below reports the same for the Dow-Jones Industrial Average (1988-2007). For all maximum likelihood estimates in this thesis, the Euclidean norm of the final gradient of the log-likelihood, denoted by

$$\left\| \nabla \ell \left(\theta \right) \right\|_{\theta = \widehat{\theta}}, \tag{4.96}$$

is reported. When $\left\| \nabla \ell \left(\theta \right) \right\|_{\theta = \widehat{\theta}}$ is not sufficiently close to zero, we say that the numerical estimate of θ is neither a local nor a global maximum. In this situation there is also no guarantee that the inverse of the final Hessian is a good approximation for the observed information. Thus, standard errors based on the final Hessian may be incorrect in this case.

AML Estimates of True Values: Apple Stock								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.361	3.990	0.175	0.726	-0.268	10.20	0.070	0.007
Std. Err.	0.098	1.031	0.018	0.105	0.082	2.53	0.008	0.006
Estimation Time			31.60 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.070	
Number of Iterations			23		$\max \ell(\theta)$		11,019	
SVJ Variance Gamma								
Estimate	0.367	3.420	0.161	0.691	-0.293	13.50	0.068	0.008
Std. Err.	0.101	1.101	0.019	0.120	0.090	4.71	0.012	0.006
Estimation Time			27.93 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.030	
Number of Iterations			19		$\max \ell(\theta)$		11,026	
SVJ Meixner								
Estimate	0.368	3.260	0.152	0.715	-0.299	5.23	0.166	0.277
Std. Err.	0.095	0.991	0.020	0.120	0.089	1.74	0.027	0.212
Estimation Time			28.08 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.015	
Number of Iterations			21		$\max \ell(\theta)$		11,027	
SVSJ Merton								
Estimate	0.360	3.850	0.167	0.643	-0.278	91.4	0.062	0.009
Std. Err.	0.089	1.095	0.016	0.103	0.076	29.19	0.008	0.005
Estimation Time			32.63 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.021	
Number of Iterations			20		$\max \ell(\theta)$		11,023	
Heston								
Estimate	0.254	22.000	0.226	2.430	-0.125	-		
Std. Err.	0.074	3.510	0.012	0.241	0.044	-		
Estimation Time			22.39 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.033	
Number of Iterations			16		$\max \ell(\theta)$		10,949	

Table 4.2: Apple Inc. (AAPL) 1991-2011: Estimation Results.

AML Estimates of True Values: Dow-Jones Index								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.059	3.390	0.024	0.282	-0.697	1.540	0.030	-0.015
Std. Err.	0.029	0.643	0.003	0.026	0.041	0.739	0.006	0.009
Estimation Time			27.35 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.017	
Number of Iterations			13		$\max\ell(\theta)$		16,851	
SVJ Variance Gamma								
Estimate	0.060	3.220	0.024	0.277	-0.705	2.220	0.028	-0.013
Std. Err.	0.030	0.697	0.003	0.028	0.045	2.005	0.010	0.010
Estimation Time			35.64 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.015	
Number of Iterations			22		$\max\ell(\theta)$		16,851	
SVJ Meixner								
Estimate	0.060	3.170	0.024	0.280	-0.719	1.330	0.057	-0.849
Std. Err.	0.029	0.674	0.003	0.027	0.044	1.165	0.021	0.416
Estimation Time			31.64 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.123	
Number of Iterations			16		$\max\ell(\theta)$		16,852	
SVSJ Merton								
Estimate	0.064	3.560	0.023	0.274	-0.703	134.0	0.028	-0.010
Std. Err.	0.027	0.644	0.002	0.026	0.038	51.78	0.005	0.006
Estimation Time			36.63 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.026	
Number of Iterations			23		$\max\ell(\theta)$		16,855	
Heston								
Estimate	0.046	5.460	0.027	0.373	-0.679	-		
Std. Err.	0.017	0.113	0.002	0.019	0.045	-		
Estimation Time			23.27 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		11.7	
Number of Iterations			10		$\max\ell(\theta)$		16,819	

Table 4.3: DJIA Index (INDU) 1988-2007: Estimation Results.

Estimation Remarks

Notice that the estimated maximum log-likelihood values for Apple stock (1991-2011), in Table 4.2 above, suggest that infinite-activity jumps are a good fit. On the other hand, for the DJIA index (1988-2007) in Table 4.3, there appears to be no advantage in choosing infinite-activity jumps. This will be discussed further in subsection 4.6.3 below.

4.4.3 Simulation Results

In this subsection we present two simulated AML estimation studies based on the AML estimated true values from subsection 4.4.2 above for the SVJ Merton, SVJ Variance Gamma, SVJ Meixner, SVSJ Merton, and Heston models, where these five models were described in Chapter 2 above. The first study uses the AML estimated true values for Apple stock (1991-2011) under each of the five models, where these estimates were reported above in Table 4.2. Similarly, the second study covers the Dow-Jones Industrial Average (1988-2007), and these estimates were reported in Table 4.3 above.

For each of the two studies, using the simulation methods described in Appendix E, and also Chapter 8 for simulating Meixner jumps, 100 paths, each with 5,000 log-returns, were simulated under the five models listed above, using the AML estimated true values from subsection 4.4.2 above, for Apple stock (1991-2011), and the DJIA (1988-2007), respectively. Then, the AML estimated true values from subsection 4.4.2 above were also used as the starting values for AML estimation along each path, where this was also done in Bates (2006), see p. 927. Below, for Apple stock (1991-2011), Table 4.4 reports the simulation results, and Table 4.5 considers the standard errors. For the Dow-Jones Industrial Average (1988-2007), these results are reported in Table 4.6, and Table 4.7, respectively below. A simulation analysis is given following the two sets of tables.

Simulation Results: Apple Stock								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
True Value	0.361	3.990	0.175	0.726	-0.268	10.20	0.070	0.007
Bias	0.001	0.229	0.001	-0.003	-0.023	0.359	-0.001	0.000
RMSE	0.101	1.119	0.018	0.072	0.077	2.027	0.006	0.007
Total Simulation Time			16.47 hours					
SVJ Variance Gamma								
True Value	0.367	3.420	0.161	0.691	-0.293	13.50	0.068	0.008
Bias	-0.012	0.345	-0.002	-0.008	-0.001	0.834	0.000	0.000
RMSE	0.099	1.140	0.019	0.086	0.081	4.86	0.011	0.006
Total Simulation Time			18.50 hours					
SVJ Meixner								
True Value	0.368	3.260	0.152	0.715	-0.299	5.23	0.166	0.277
Bias	-0.013	0.262	0.002	0.011	0.002	0.267	0.005	-0.014
RMSE	0.089	0.851	0.021	0.077	0.089	1.896	0.032	0.198
Total Simulation Time			18.32 hours					
SVSJ Merton								
True Value	0.360	3.850	0.167	0.643	-0.278	91.4	0.062	0.009
Bias	0.001	0.341	-0.003	-0.006	0.004	3.37	0.000	-0.001
RMSE	0.082	1.122	0.015	0.068	0.092	24.36	0.006	0.005
Total Simulation Time			21.64 hours					
Heston								
True Value	0.254	22.000	0.226	2.430	-0.125	-		
Bias	0.005	1.739	-0.002	0.029	0.001	-		
RMSE	0.087	3.413	0.013	0.180	0.050	-		
Total Simulation Time			8.24 hours		-			

Table 4.4: Apple Inc. (AAPL) 1991-2011: Simulation Results.

Standard Errors: Apple Stock								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Sim.SE	0.101	1.095	0.017	0.072	0.073	1.995	0.006	0.007
Asym.SE	0.098	1.031	0.018	0.105	0.082	2.530	0.008	0.006
$\overline{\textbf{Asym.SE}}$	0.099	1.036	0.018	0.083	0.079	2.122	0.006	0.007
SVJ Variance Gamma								
Sim.SE	0.098	1.087	0.019	0.085	0.081	4.783	0.011	0.006
Asym.SE	0.101	1.101	0.019	0.120	0.090	4.707	0.012	0.006
$\overline{\textbf{Asym.SE}}$	0.105	1.214	0.019	0.099	0.097	4.884	0.012	0.007
SVJ Meixner								
Sim.SE	0.089	0.810	0.020	0.076	0.089	1.877	0.031	0.197
Asym.SE	0.095	0.991	0.020	0.120	0.089	1.735	0.027	0.212
$\overline{\textbf{Asym.SE}}$	0.099	0.899	0.021	0.082	0.093	1.931	0.031	0.210
SVSJ Merton								
Sim.SE	0.082	1.069	0.014	0.068	0.092	24.12	0.006	0.005
Asym.SE	0.089	1.095	0.016	0.103	0.076	29.19	0.008	0.005
$\overline{\textbf{Asym.SE}}$	0.093	1.009	0.015	0.076	0.081	23.47	0.006	0.005
Heston								
Sim.SE	0.087	2.936	0.013	0.178	0.050	-		
Asym.SE	0.074	3.510	0.012	0.241	0.044	-		
$\overline{\textbf{Asym.SE}}$	0.092	3.467	0.012	0.206	0.049	-		
Notes								
Sim.SE $=\sqrt{RMSE^2 - Bias^2}$ (simulated SE)								
Asym.SE = estimated asymptotic standard error								
$\overline{\textbf{Asym.SE}}$ = mean asymptotic standard error over each simulation								

Table 4.5: Apple Inc. (AAPL) 1991-2011: Standard Errors.

Simulation Results: Dow-Jones Index								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
True Value	0.059	3.390	0.024	0.282	-0.697	1.54	0.030	-0.015
Bias	-0.003	0.156	0.000	-0.001	-0.003	0.004	-0.004	-0.005
RMSE	0.034	0.658	0.003	0.021	0.042	0.782	0.010	0.015
Total Simulation Time			18.72 hours					
SVJ Variance Gamma								
True Value	0.060	3.220	0.024	0.277	-0.705	2.22	0.028	-0.013
Bias	-0.005	0.264	0.000	0.001	-0.003	0.752	-0.001	-0.003
RMSE	0.031	0.777	0.003	0.023	0.048	2.23	0.011	0.011
Total Simulation Time			19.72 hours					
SVJ Meixner								
True Value	0.060	3.170	0.024	0.280	-0.719	1.330	0.057	-0.849
Bias	-0.003	0.026	0.000	-0.003	-0.002	0.428	-0.000	-0.033
RMSE	0.030	0.538	0.003	0.021	0.046	1.257	0.023	0.506
Total Simulation Time			19.18 hours					
SVSJ Merton								
True Value	0.064	3.560	0.023	0.274	-0.703	134.0	0.028	-0.010
Bias	-0.005	0.255	0.000	-0.002	-0.005	4.64	-0.001	-0.002
RMSE	0.034	0.758	0.003	0.023	0.049	57.02	0.006	0.009
Total Simulation Time			24.94 hours					
Heston								
True Value	0.046	5.460	0.027	0.373	-0.679	-		
Bias	-0.003	0.307	0.000	0.004	-0.001	-		
RMSE	0.028	0.874	0.002	0.027	0.040	-		
Total Simulation Time			8.25 hours		-			

Table 4.6: DJIA Index (INDU) 1988-2007: Simulation Results.

Standard Errors: Dow-Jones Index								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Sim.SE	0.034	0.640	0.003	0.021	0.042	0.782	0.009	0.014
Asym.SE	0.029	0.643	0.003	0.026	0.041	0.739	0.006	0.009
$\overline{\textbf{Asym.SE}}$	0.030	0.598	0.003	0.021	0.045	0.709	0.006	0.009
SVJ Variance Gamma								
Sim.SE	0.031	0.731	0.003	0.023	0.048	2.095	0.011	0.011
Asym.SE	0.030	0.697	0.003	0.028	0.045	2.005	0.010	0.010
$\overline{\textbf{Asym.SE}}$	0.031	0.631	0.003	0.022	0.048	2.253	0.015	0.012
SVJ Meixner								
Sim.SE	0.030	0.538	0.003	0.021	0.046	1.182	0.023	0.505
Asym.SE	0.029	0.674	0.003	0.027	0.044	1.165	0.021	0.416
$\overline{\textbf{Asym.SE}}$	0.032	0.732	0.003	0.024	0.050	1.208	0.023	0.559
SVSJ Merton								
Sim.SE	0.033	0.714	0.003	0.022	0.048	56.83	0.006	0.009
Asym.SE	0.027	0.644	0.002	0.026	0.038	51.78	0.005	0.006
$\overline{\textbf{Asym.SE}}$	0.030	0.650	0.003	0.022	0.045	57.54	0.005	0.006
Heston								
Sim.SE	0.028	0.819	0.002	0.026	0.040	-		
Asym.SE	0.017	0.113	0.002	0.019	0.045	$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}} = 11.7$		
$\overline{\textbf{Asym.SE}}$	0.030	0.852	0.002	0.027	0.040	-		
Notes								
Sim.SE $=\sqrt{RMSE^2 - Bias^2}$ (simulated SE)								
Asym.SE = estimated asymptotic standard error								
$\overline{\textbf{Asym.SE}}$ = mean asymptotic standard error over each simulation								

Table 4.7: DJIA Index (INDU) 1988-2007: Standard Errors.

Simulation Analysis

In Table 4.4 for Apple stock, and Table 4.6 for the Dow-Jones Industrial Average, only two parameters show any substantial degree of bias for any of the models considered. These are the rate of mean-reversion κ, and the jump intensity λ. From Bates (2006), p. 931, the upward bias to κ is explained by the noise in the variance process, even more so because we have filtered the variance from noisy log-return data. This particular bias is expected to exist even if the variance process is directly observed, although it is also expected to diminish in samples of log-return data larger than 5,000, see Bates (2006), p. 931. The noisy variance also gives κ a larger RMSE. The bias to λ depends on the jump activity-rate present in the data, the jump activity-rate specified by the model, and in our implementation to a mild extent on whether or not the jump intensity is stochastic. We will come back to complications of the λ analysis in a moment.

The following information is useful in explaining the bias and RMSE of certain parameters estimated by AML. Recall from equations (4.34) and (4.36) in subsection 4.2.4 above, that the conditional density used to compute the likelihood under AML is given by

$$p\left(y_{n+1}|y_n^H\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u,0|y_n^H\right)du, \tag{4.97}$$

$$\text{where } \underline{F}\left(u,0|y_n^H\right) = e^{C(u;\Delta t)}G_{n|n}\left[D\left(u;\Delta t\right)\right]. \tag{4.98}$$

That is, by equation (4.98), the log-price coefficient $D\left(u;\Delta t\right)$ is affected by the noise in the variance filtration process, whereas the $C\left(u;\Delta t\right)$ is not. Further note the following parameter dependencies from Remark A.2 at the end of Appendix A. For both the SVJ and SVSJ models,

$$C\left(u;\Delta t\right) = C\left(u;\mu_0,\kappa,\eta,\omega,\rho,\lambda,\alpha,\beta;\Delta t\right). \tag{4.99}$$

Whereas for the $D(u;\Delta t)$ coefficient,

$$D^{J0}(u;\Delta t) = D^{J0}(u;\kappa,\omega,\rho;\Delta t), \text{ for the SVJ model, and} \tag{4.100}$$

$$D^{J1}(u;\Delta t) = D^{J1}(u;\kappa,\omega,\rho,\lambda,\alpha,\beta;\Delta t), \text{ for the SVSJ model.} \tag{4.101}$$

Specifically, the drift μ_0 and long run variance η only appear in $C(u;\Delta t)$ for both models, and the three jump parameters λ, α, and β appear in $D^{J1}(u;\Delta t)$ but not $D^{J0}(u;\Delta t)$.

Again consider Table 4.4 for Apple stock, and Table 4.6 for the Dow-Jones Industrial Average. We claim that the sole drift parameter μ_0 has low bias because it is not subject to the variance filtration process, but it has a high RMSE since the drift is typically difficult to estimate. By contrast, the long run variance η has both low bias and low RMSE. We expect this is because it only appears in the $C(u;\Delta t)$ coefficient, thus is not subject to the variance filtration process. The volatility of the volatility ω has low bias, but high RMSE. This is consistent with Bates (2006), p. 931, where it is stated that this is because of the noise in the filtration process. Observe that the same is true for ρ, and both ω and ρ appear in the $D(u;\Delta t)$ coefficient in both the SVJ and SVSJ models. Hence, ω and ρ are both subject to the variance filtration process. Given the small size of some of the estimates, the jump parameters α, and β also perform quite well, and this is consistent with Bates (2006), p. 930.

We now return to the subject of λ. According to Aït-Sahalia and Jacod (2012), §9.2, pp. 1036-37, the DJIA index typically exhibits finite-activity jumps, whereas the individual stocks comprising the index typically exhibit infinite-activity jumps, both with a diffusion present, see §9.3, p. 1040 and §10, p. 1047, respectively. For the DJIA index (1988-2007) in Table 4.6 above, we found that under the SVJ Merton model, the DJIA appears to be unbiased for λ. This is consistent with Bates (2006), Table 2, p. 928, for the S&P 500. But, in Table 4.6 for the DJIA index under the SVSJ Merton

model, λ shows a 3.5% bias. While this is not large, it is larger than Bates (2006), Table 4, p. 930 predicts for 5,000 log-returns. We attribute this to the dependence of the $D(u;\Delta t)$ coefficient on λ in the SVSJ model, within the framework of our implementation. More importantly, Table 4.6 shows that models with infinite-activity jumps are upwardly biased for λ with data such as the DJIA index, containing only finite-activity jumps. This occurs since these models classify some part of the many diffusive movements in index data as small jumps. Consequently, λ is biased upwards. Next consider Apple Stock (1991-2011) in Table 4.4 above. For the SVSJ Merton model, with stochastic intensity given by $\lambda_t = \lambda\sigma_t^2$, we will define the long run intensity to be

$$\overline{\lambda} = \lambda\eta. \tag{4.102}$$

For Apple stock (1991-2011) in Table 4.2, $\overline{\lambda} = 15.26$, whereas for the DJIA index (1988-2007) in Table 4.3, $\overline{\lambda} = 3.08$. This, and other measurements of λ, all show that Apple stock has many more jumps than the DJIA index. Thus, since Apple is an individual stock, we will assume that its jumps have infinite-activity. From Table 4.4 for Apple Stock (1991-2011), λ shows an upward bias of 3.5% for both the SVJ Merton and SVSJ Merton models. However, similar to Honoré (1998), §5.1, estimating a diffusion model with finite-activity jumps can lead to an upward bias in λ when there are too many jumps, though this occurs here to a much lesser extent given the latent factor. What is surprising is that in Table 4.4 for Apple Stock a similar upward bias occurs with infinite-activity, and the relative bias to λ, though still small, is nearly double what it was for finite-activity. Lastly on the subject of λ, exposure to the noise from the filtered volatility gives λ a very high RMSE in the SVSJ Merton model for both Apple stock and the DJIA index. This is consistent with Bates (2006), p. 930. Moreover, the RMSE of λ appears to be relatively high in general, since jumps are typically

more difficult to estimate. We conclude that if the log-returns exhibit finite-activity jumps, then diffusion models providing for infinite-activity jumps will be upwardly biased for λ, but if the log-returns exhibit infinite-activity jumps, then diffusion models providing for infinite-activity jumps will have a smaller upward bias for λ, both in the presence of a latent factor.

In Table 4.5 for Apple stock, and Table 4.7 for the DJIA index, it is shown that the asymptotic standard errors from the original AML estimation, and the mean of the asymptotic standard errors from AML estimations based on the simulated paths, are both highly consistent with the simulated standard errors given by $Sim.SE = \sqrt{RMSE^2 - Bias^2}$. An exception occurs in Table 4.7 for the DJIA index under the Heston model. In this case the final gradient of the negative log-likelihood is quite far from zero. This means that the estimate is neither a local nor a global maximum. Therefore, we cannot conclude asymptotic normality of the AML estimates in this case, see for example Lehmann and Casella (1998), Theorem 3.10, pp. 449-50. Hence, the asymptotic standard errors are not necessarily reliable.

The following concerns exception handling in the simulated estimation of the SVJ Meixner model, due to numerical instability in the log-price CF that we describe below. The independent jump part of the SVJ Meixner log-price characteristic function is given by

$$\phi_X(u) = \left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cosh\left(\frac{\alpha u - i\beta}{2}\right)} \right)^{\lambda \Delta t}, \text{ with } |\beta| < \pi. \tag{4.103}$$

As $\beta \to \pm\pi$ in equation (4.103) above, $\cos\left(\frac{\beta}{2}\right) \to 0$. So, either $\phi_X(u) \to 0$, or $\phi_X(u)$ becomes numerically undefined as $\cosh\left(\frac{\alpha u - i\beta}{2}\right)$ also vanishes. It can be shown that the latter occurs as $\beta \to \pm\pi$, when $u = 0$. Either phenomenon creates instability in the SVJ Meixner density which we retrieve from its CF. Thus, the log-likelihood can become numerically unreliable.

Occasionally this leads to estimated values for β that are unrealistically large either positively or negatively. In these cases, the estimates of the other SVJ Meixner parameters tend to be unpredictable. We found this phenomenon to be relatively rare. For the SVJ Meixner, it occurred twice in the Apple stock simulated estimation in Table 4.4, and once for the DJIA index simulated estimation in Table 4.6. Estimates for real world data were not affected by this phenomenon. Our solution was to discard the tainted simulations and simulate again with new random seeds.

Last of all, we discuss natural constraints. By Proposition 2.24 *1.* in subsection 2.5.1, if the Feller condition, $2\kappa\eta > \omega^2$, holds then the latent factor in the SVJ and SVSJ models is strictly positive for all $t \geq 0$. Also, by Proposition B.16 in Appendix B.2, the condition $\rho < 0$ is sufficient for principal branch continuity of the marginal log-price coefficients, $C(u)$ and $D(u)$, as defined in Appendix A.2, for the SVJ and SVSJ models. While in general, we do not use constraints on the likelihood for either the Feller condition, or the condition $\rho < 0$, these two conditions were both met by the original estimates for Apple stock and the DJIA index in Table 4.2 and Table 4.3, respectively. Also, both conditions were met in nearly all of the estimates underlying the simulations in Table 4.4 for Apple stock, and Table 4.6 for the DJIA index. The only exceptions were the three SVJ Meixner cases discarded for estimated values of β near $\pm\pi$, as described above. The situation differs for the other data sets in Section 4.5 and Section 4.6 below. Here, the condition $\rho < 0$ is always met, but if λ is too large, then the Feller condition can fail for the SVJ model with infinite-activity jumps. The Feller condition does not usually fail in the AML estimations of subsection 4.5.1 and subsection 4.6.1 below. But, if it does we re-estimate the model with a Feller constraint. These re-estimation results are reported separately in Table 4.23 of subsection 4.6.2.

4.4.4 Comparisons to MCMC Literature

In this brief subsection we compare FFT based AML estimates and asymptotic standard errors with two sets of results from the MCMC literature. Our first comparison is to Eraker, Johannes, and Polson (2003) where the following alternative version of the SVJ Merton model is used.

$$\begin{aligned} dY_t &= \mu_0 dt + \sigma_t dW_t^{(S)} + dX_t^{MJ}, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right) dt + \omega\sigma_t dW_t^{(V)}, \\ \rho dt &= E\left[dW_t^{(S)} dW_t^{(V)}\right]. \end{aligned} \tag{4.104}$$

The characteristic function for the model (4.104) is a special case of the one presented in Appendix A. Simply change $\widehat{\psi}_W(u)$ in equation $(A.5)$ from $-\frac{1}{2}\left(u^2+iu\right)$ to $-\frac{1}{2}u^2$, and assuming Merton jumps, simply change $\widehat{\psi}_X(u)$ in equation $(A.4)$ from $\psi_X(u) - iu\psi_X(-i)$ to $\psi_X(u)$.

Parameter	μ_0	κ	η	ω	ρ	λ	α	β
Our Estimation Results: SVJ Merton Model (alternative)								
Estimate	0.119	3.450	0.021	0.236	-0.476	1.174	0.055	-0.024
Asym.SE	0.028	0.718	0.002	0.026	0.065	0.301	0.011	0.013
Estimation Time			30.02 min		$\left\|\nabla\ell(\theta)\right\|_{\theta=\widehat{\theta}}$		0.082	
Number of Iterations			15		$\max\ell(\theta)$		16,951	
Eraker et al. (2003), Table III, p. 1280								
Estimate	0.125	3.226	0.021	0.240	-0.467	1.512	0.041	-0.026
Std. Err.	0.027	0.983	0.003	0.026	0.058	0.529	0.017	0.013

Table 4.8: S&P 500 Index (SPX) 1980-1999: Comparing AML to MCMC I. The results from Eraker, Johannes, and Polson (2003), Table III, have been annualized, see Bates (2006), Table 2, p. 928.

Bates (2006), Table 2, p. 928, compares to the simulation results from Eraker et al. (2003). In Table 4.8 above, we compare to the actual estimates. The data used in Table 4.8 is the S&P 500 index from January 2nd, 1980 to December 31st, 1999 for a total of 5,054 trading day log-returns, see Eraker et al. (2003), Table II, p. 1280 for a summary. We used similar data from Bloomberg L.P. In Table 4.8, according to our AML estimates, the jump intensity λ is slightly smaller, but α, the standard deviation of the jumps size, is slightly larger. Otherwise, MCMC and AML are remarkably similar in this case for both the estimates and asymptotic standard errors.

Our next MCMC comparison is to Lysy (2012), p. 73. In Table 4.9 below, Lysy (2012) treats the Heston model with a drift given by $\mu_0 - \frac{1}{2}\sigma_t^2$, similar to our treatment in this thesis. The data used in Table 4.9 is the DJIA index from August 4th, 1997 to June 4th, 2007 for a total of 2,472 log-returns. Table 4.9 below also shows consistency between MCMC and AML. Notice that the Heston model has a good fit with only 2,500 log-returns.

Parameter	μ_0	κ	η	ω	ρ
Our Estimation Results: Heston Model					
Estimate	0.003	3.204	0.035	0.375	-0.848
Asym.SE	0.042	1.201	0.007	0.044	0.032
Time (min)	11.39	$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.015	
Iterations	18	$\max \ell(\theta)$		8,019.9	
Lysy (2012), Table 4.1, p. 73					
Estimate	3.0E-4	3.760	0.032	0.401	-0.823
Std. Err.	0.041	0.997	0.006	0.043	0.037

Table 4.9: DJIA Index (INDU) 1997-2007: Comparing AML to MCMC II. Standard errors were supplied by Lysy in correspondence, and the long run variance is reported here in place of the long run volatility.

4.4.5 Quadrature and the Bates (2006) Results

In this subsection we ultimately replicate the AML estimation results from Bates (2006), Table 7, p. 938, see Table 4.14 below, and the comments that follow. But, two preliminary steps are required here. First, we give the Bates (2006) model in equation (4.105), and introduce the required data set, the S&P 500 (1953-1996) daily log-returns, from Andersen, Benzoni, and Lund (2002), see pp. 1250-1252. Secondly, because the data set contains 11,076 log-returns, and is too large for AML by FFT, we modify Algorithm 4.15 above to use five point Newton-Cotes quadrature in place of FFT. Lastly, for the final replication results in Table 4.14 below, the gradient is not close to zero, thus we simulate in order to obtain reliable standard errors.

From Bates (2006), eq. A.10, p. 953, the original SVSJ Merton model in log-price form may be written in the notation from Chapter 2 as

$$\begin{aligned} dY_t &= \left(\mu_0 + \left(\mu_1 - \frac{1}{2}\right)\sigma_t^2 - \sigma_t^2\psi_X(-i)\right)dt + \sigma_t dW_t^{(S)} + dX(V_t), \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)}, \\ dV_t &= \sigma_t^2 dt, \;\; \rho dt = E\left[dW_t^{(S)} dW_t^{(V)}\right]. \end{aligned} \tag{4.105}$$

The characteristic function for the model (4.105) is similar to the one given in Appendix A. Simply change $\widehat{\psi}_W(u)$ in equation $(A.5)$ from

$$-\frac{1}{2}\left(u^2 + iu\right) \text{ to } -\frac{1}{2}\left(u^2 + iu\right) + iu\mu_1.$$

For the data we begin with the S&P 500 from January 2nd, 1953 to December 31st, 1996, yielding 11,076 log-returns from Bloomberg L.P. The sample moments of these log-returns are similar to Andersen, Benzoni, and Lund (2002), Table I, Panel A, p. 1250. Thus, we proceed. Andersen et al. (2002), Table I, Panel B, provides the autocorrelation function of the log-return data to six lags. We provide the same in Table 4.10 below.

Lags	1	2	3	4	5	6
Andersen et al. (2002)	.1240	-.0320	-.0084	-.0056	.0222	-.0131
Bloomberg L.P.	.1211	-.0321	-.0080	-.0076	.0226	-.0124

Table 4.10: S&P 500 Index 1953-1996: Autocorrelation Function.

It is believed that the significant high order lags are indicative of nonsynchronous trading in the stocks underlying the S&P 500 index, see Andersen et al. (2002), p. 1251. Their prescription is to fit an MA(1) model, and replace the original returns with the residuals, rescaled to match the sample mean and variance of the original returns, see Andersen et al. (2002), p. 1252. Under this prescription, we obtained the data set underlying Table 4.11 below. This is the data set used in Bates (2006), Table 7.

	Mean	Std. Dev.	Skewness	Kurtosis
Filtered Data	0.0301	0.8286	-1.8522	56.7595

Table 4.11: S&P 500 Index 1953-1996: Sample Moments based on MA(1) Filtered Percentage Daily Log-Return Data.

Quadrature Based AML

To tackle the 11,076 log-returns of the Andersen, Benzoni, and Lund (2002) S&P 500 data set, we introduce a simple form of quadrature that is faster and more accurate than FFT. The method is known as composite Boole's rule, see Davis and Rabinowitz (1984), p. 70 and p. 78.

$\ell(\theta)$ **Eval. Time**	**5,042 Obs.**	**11,076 Obs.**
Boole's Rule	6.23 sec	10.65 sec
FFT with Splines	14.87 sec	33.69 sec

Table 4.12: Speed of Boole's Rule vs. FFT: AML Likelihood Evaluation.

In Table 4.12 likelihood evaluation times were measured on the final iteration of the Bates SVSJ model (4.105) defined above. For Table 4.13 below, we estimated the Meixner model from subsection 2.3.3, which has an exact density, using the MA(1) filtered data underlying Table 4.11 above. The following errors were taken based on the MA(1) filtered log-returns.

Meixner	**Mean Abs Error vs. Exact**
Boole's Rule	2.21E-13
FFT with Splines	5.60E-11

Table 4.13: Accuracy of Boole's Rule vs. FFT: Exact Meixner Density.

FFT appears to be reasonable for 5,000 log-returns, but quadrature is needed for larger data sets, especially when tails are important. Recall from subsection 4.3.2 that we applied FFT with splines three times in Algorithm 4.15 to evaluate the Fourier integral (4.94) given by

$$g(x) = \int_0^\infty e^{-iux} \widehat{f}(u)\, du, \tag{4.106}$$

with the Fourier transform $\widehat{f}(u)$ set to each of $\underline{F}\left(u, 0|y_n^H\right)$, $\underline{F}_v\left(u, 0|y_n^H\right)$, and $\underline{F}_{vv}\left(u, 0|y_n^H\right)$, defined in equations (4.63), (4.69), and (4.70), respectively above. We proved in Section 4.3 that these Fourier transforms are L^1, thus (4.106) converges for all $x \in \mathbb{R}$. However, while FFT approximates (4.106) at a grid of points x surrounding y_{n+1}, quadrature specifies y_{n+1} precisely with no interpolation error. Similar to FFT, the composite Boole's rule approximates (4.106) with a finite upper bound $A = N\Delta u$, and a Riemann sum. But, Boole's rule places the weights at the $N+1$ end-points of the N equal subintervals,

$$g(y_{n+1}) \approx \int_0^A e^{-iuy_{n+1}} \widehat{f}(u)\, du, \tag{4.107}$$

$$\approx \sum\nolimits_{j=1}^{N+1} e^{-iu_j y_{n+1}} \widehat{f}(u_j)\, w_j \Delta u. \tag{4.108}$$

Our implementation of Boole's rule uses $N = 10,000$, and $\Delta u = \frac{1}{4}$. The weights for a single interval of Boole's rule are well known, see Davis and Rabinowitz (1984), p. 78. But, for a composite closed Newton-Cotes formula such as Boole's rule, the end-points of the intervals coincide. Both weights must be counted at the matching end-points, see Burden and Faires (1997), pp. 199-201, for the composite Simpson's rule. Thus, the weights for the composite Boole's rule are given by

$$w_1 = \frac{14}{45},\ w_2 = \frac{64}{45},\ w_3 = \frac{24}{45},\ w_4 = \frac{64}{45};$$

$$\text{and for } k = 1 \text{ to } \frac{N}{4} - 1\text{:}$$

$$w_{4k+1} = \frac{28}{45},\ w_{4k+2} = \frac{64}{45},\ w_{4k+3} = \frac{24}{45},\ w_{4k+4} = \frac{64}{45};$$

$$\text{then } w_{N+1} = \frac{14}{45}.$$

Algorithm 4.16 (Log-Likelihood III) *Given* $\theta = (\mu, \kappa, \eta, \omega, \rho, \lambda, \alpha, \beta)$

$(a_0, b_0) = \left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$

$\underline{F}_0\left(u, 0|y_n^H\right) = e^{\underline{C}(u,0;\Delta t)}\left[1 - b_0\underline{D}\left(u, 0; \Delta t\right)\right]^{-a_0}$

$p\left(y_1|\theta\right) = \frac{1}{\pi}\,\mathrm{Re}\sum_{j=1}^{N+1} e^{-iu_j y_{n+1}}\underline{F}_0\left(u_j, 0|y_n^H\right) w_j \Delta u$

BEGIN FOR $n = 1$ *TO* $N_y - 1$

$p\left(y_{n+1}|y_n^H, \theta\right) = \frac{1}{\pi}\,\mathrm{Re}\sum_{j=1}^{N+1} e^{-iu_j y_{n+1}}\underline{F}\left(u_j, 0|y_n^H\right) w_j \Delta u$

$\widehat{m}'_{1,n+1} = \frac{\frac{1}{\pi}\,\mathrm{Im}\sum_{j=1}^{N+1} e^{-iu_j y_{n+1}}\underline{F}_v\left(u_j, 0|y_n^H\right) w_j \Delta u}{p\left(y_{n+1}|y_n^H, \theta\right)}$

$\widehat{m}'_{2,n+1} = \frac{\frac{-1}{\pi}\,\mathrm{Re}\sum_{j=1}^{N+1} e^{-iu_j y_{n+1}}\underline{F}_{vv}\left(u_j, 0|y_n^H\right) w_j \Delta u}{p\left(y_{n+1}|y_n^H, \theta\right)}$

$b_{n+1} = \frac{\widehat{m}'_{2,n+1} - \left(\widehat{m}'_{1,n+1}\right)^2}{\widehat{m}'_{1,n+1}},\ a_{n+1} = \frac{\widehat{m}'_{1,n+1}}{b_{n+1}}$

END FOR

$\ell\left(\theta\right) = \ell_0 + \log p\left(y_1|\theta\right) + \sum_{n=1}^{N_y - 1}\log p\left(y_{n+1}|y_n^H, \theta\right)$

Algorithm 4.16 above executes the AML method using Boole's rule. It will only be used for comparison to the Bates (2006) results. In the remainder of this chapter, only data sets of 5,042 points are considered for AML estimation. For continuity, we will let FFT suffice.

Replication of the Bates (2006) Results

For the Bates (2006) version of the SVSJ Merton model, see model (4.105) above, Table 4.14 below compares our AML estimates for the Andersen, Benzoni, and Lund (2002) S&P 500 (1953-1996) data, see Table 4.11 above, to the estimates from Bates (2006), Table 7, p. 938.

Param	μ_0	μ_1	κ	η	ω	ρ	λ	α	β
Our Estimation Results: SVSJ Merton Model (modified drift)									
Est.	.045	3.14	4.43	.014	.232	-.596	93.4	.039	-.003
Sim.SE	.027	2.59	.620	.001	.012	.031	22.2	.004	.007
Estimation Time				49.14 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		18.95	
Number of Iterations				21		$\max \ell(\theta)$		39,359	
Bates (2006) Estimation Results, Table 7, p. 938									
Est.	.040	3.09	4.25	.014	.237	-.611	93.4	.039	-.002
Asy.SE	.025	2.16	0.59	2E-5	.015	.031	33.4	.008	.006
-						$\max \ell(\theta)$		39,310	
Param	μ_0	μ_1	κ	η	ω	ρ	λ	α	β
Our Simulation Results (11,076 log returns, M=100)									
True	.045	3.14	4.43	.014	.232	-.596	93.4	.039	-.003
Bias	-.008	.735	.184	.000	-.001	-.011	3.95	-.001	.000
RMSE	.028	2.69	.646	.001	.012	.033	22.6	.004	.007
Bates (2006) Table 4, p. 930 (12,000 log returns, M=100)									
True	.040	3.09	4.25	.014	.246	-.611	93.4	.039	-.024
Bias	.009	2.15	.000	.000	-.002	-.003	-2.7	-.001	-.002
RMSE	.027	3.23	.050	2E-5	.011	.038	21.0	.004	.007

Table 4.14: S&P 500 Index (SPX) 1953-1996: Comparison to Bates (2006) with MA(1) filtered daily data from Andersen, Benzoni, and Lund (2002).

Notice in the upper panel of Table 4.14 above, that the AML estimates from Bates (2006), Table 7, p. 938 are reasonably well replicated. Notice also in the upper panel of Table 4.14 above that our final gradient is not close to zero. Hence, we ran a simulation to obtain our standard errors. Details of the simulation are in the lower panel of Table 4.14 above, and are compared to Bates (2006), Table 4, p. 930. Our standard error for λ in Table 4.14 is smaller than Bates (2006), Table 7, p. 938. But, otherwise the estimates and standard errors are quite similar. There are three main differences in the biases in Table 4.14. First, our bias to κ fails to vanish at 11,076 observations, whereas it does in Bates (2006), Table 4, p. 930. Secondly, we still have the small but persistent upward bias to λ, whereas Bates (2006), Table 4, p. 930 has a negative bias to λ. Lastly, our estimate of μ_1 is surprisingly less biased than Bates (2006), Table 4, p. 930. The RMSEs in Table 4.14 above are similar to Bates (2006), Table 4, p. 930.

4.5 AML Filtered Volatility and the VIX Index

In subsection 4.5.2 below, we compare the AML filtered volatility for the S&P 500 daily to the VIX daily index (1988-2007). However, we naturally begin with the unconstrained AML parameter estimates themselves for the S&P 500 over this period, in subsection 4.5.1 immediately below.

4.5.1 AML Estimates for the S&P 500

One would expect the AML estimates for the S&P 500 daily log-returns (1988-2007), given in Table 4.15 below, to be highly similar to the DJIA estimates from Table 4.3 in subsection 4.4.2. While this is true for most of the five models that we have estimated, something unusual happens in the SVJ Meixner with infinite-activity jumps of infinite variation.

Unconstrained AML Estimates: S&P 500 (SPX)								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.046	3.248	0.026	0.300	-0.727	1.504	0.026	-0.014
Std. Err.	0.028	0.699	0.003	0.029	0.038	0.891	0.006	0.010
Estimation Time			28.96 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.013	
Number of Iterations			14		$\max\ell(\theta)$		16,869	
SVJ Variance Gamma								
Estimate	0.047	3.033	0.025	0.294	-0.745	4.662	0.019	-0.006
Std. Err.	0.029	0.626	0.003	0.029	0.029	1.736	0.004	0.004
Estimation Time			35.95 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.348	
Number of Iterations			20		$\max\ell(\theta)$		16,870	
SVJ Meixner (Feller condition fails)								
Estimate	0.072	2.710	0.022	0.505	-1.000	83.84	0.013	0.027
Std. Err.	0.029	0.472	0.005	0.055	0.039	19.15	0.002	0.109
Estimation Time			38.37 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		3.17	
Number of Iterations			27		$\max\ell(\theta)$		16,892	
SVSJ Merton								
Estimate	0.052	3.346	0.025	0.289	-0.731	134.5	0.024	-0.010
Std. Err.	0.027	0.579	0.003	0.025	0.035	48.51	0.004	0.006
Estimation Time			42.52 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.013	
Number of Iterations			28		$\max\ell(\theta)$		16,872	
Heston								
Estimate	0.025	5.019	0.028	0.384	-0.706	-		
Std. Err.	0.020	0.153	0.002	0.022	0.035	-		
Estimation Time			24.28 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		4.09	
Number of Iterations			15		$\max\ell(\theta)$		16,843	

Table 4.15: S&P 500 Index (SPX) 1988-2007: Estimation Results.

We estimated nine sets of log-returns, including the Apple Inc. stock in Table 4.2, the DJIA in Table 4.3, and the S&P 500 in Table 4.15 above, by unconstrained AML estimation, under the five models of this chapter. The remaining results are given in subsection 4.6.1 below. All forty-five of these estimates are given with asymptotic standard errors. The Feller condition fails for the SVJ Meixner model in four cases, including the S&P 500 in Table 4.15 above. The Feller condition also fails for VG jumps in only two of these four cases. Thus, the problem appears more likely in the infinite variation Meixner case. Since the problem occurs when λ is quite large, and the jumps are quite small, it appears that AML confuses the infinite-activity process with a second diffusion, see Carr and Wu (2003), p. 2602. In re-estimating these six cases with a Feller constraint, we found that λ is smaller, but still quite large. Also, the Feller constraint is binding in all cases, biasing the Hessian, so that asymptotic standard errors are unreliable. Hence, we obtained the Feller constrained standard errors by simulation. The six Feller constrained estimates and their standard errors are reported separately in Table 4.23 of subsection 4.6.2 below. For the other thirty-nine sets of log-returns that we estimated, the estimates of λ are lower, and the Feller condition holds in all five models. An objective for including jumps in a stochastic volatility model is to make the variance parameters more reasonable. Adding Merton jumps to the Heston model will make the volatility of the volatility ω smaller, see Bates (2000), p. 216. We find that when λ is low enough, adding jumps to the Heston model makes ω smaller, even for the two infinite-activity models. But, for infinite-activity, if λ is too large, then typically the Feller condition either fails or may be binding. Recall from Table 2.1 in subsection 2.3.4 that the theoretical skewness and excess kurtosis of the jumps both vanish as λ gets large. This may partly explain the problem with infinite-activity jumps when λ is large.

4.5.2 The AML Filtered Volatility

Throughout each of the figures in this subsection we follow, with markings, the five biggest S&P 500 daily losses over 1988-2007. In chronological order, these are (with log-returns in parentheses): the mini-crash of January 8th, 1988 (-.07008), the Friday October 13th, 1989 mini-crash (-.06321), the Asian crisis panic of October 27th, 1997 (-.07113), the Russian crisis fall-out of August 31st, 1998 (-.07044), and the initial burst of the tech-bubble on April 14th, 2000 (-.06005).

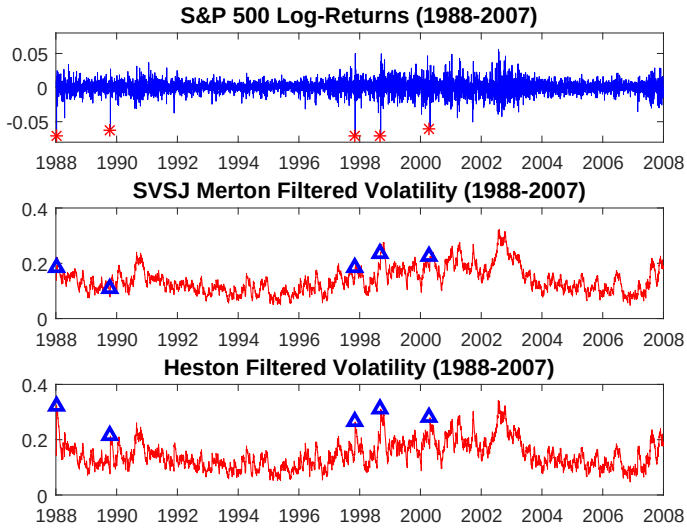

Figure 4.1: S&P 500 SVSJ Merton and Heston Filtered Volatilities.

As suggested in Figure 4.1 above, the SVSJ Merton filtered volatility is slightly conservative at the large loss dates. Thus, we will use the SVSJ Merton filtered volatility as the baseline in the comparisons below. It is argued in Bates (2006), see pp. 931-33, that the SVSJ Merton filtered volatility is the most suitable if we wish to attribute large losses to jumps.

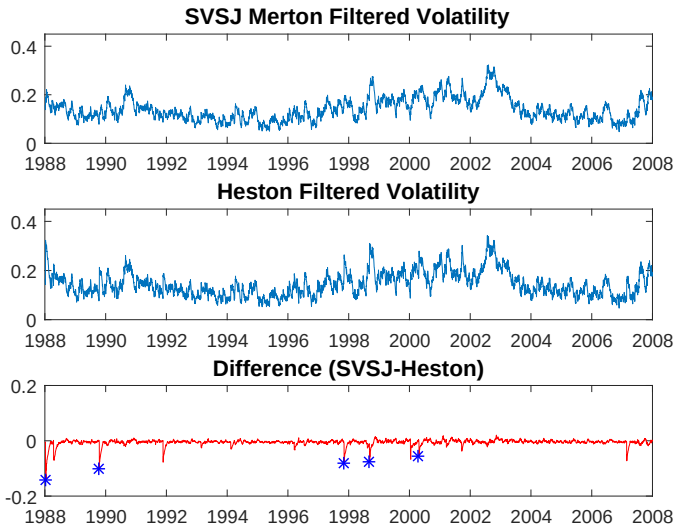

Figure 4.2: SPX Filtered Volatility Differences: SVSJ minus Heston.

As shown in Figure 4.2 above, there is essentially no difference between the Heston AML filtered volatility and the SVSJ Merton AML filtered volatility, except immediately following the large loss dates, where the five biggest daily losses for the S&P 500 over the 1988-2007 period are marked, as outlined at the start of this subsection. This is consistent with Bates (2006), see p. 942. Also, Figure 4.2 above shows that the Heston filtered volatility overshoots the SVSJ filtered volatility immediately following these dates. This is because for large observations the Heston model has a steeper volatility news impact curve than the SVSJ model does, see Bates (2006), pp. 931-32. Thus, the Heston filtered volatility resembles the VIX index the most. Moreover, for this reason, many people prefer the Heston filtered volatility to the more conservative SVSJ Merton filtered volatility. But, Bates (2006) prefers the SVSJ Merton filtered volatility, in the event that jumps are to be included in the model.

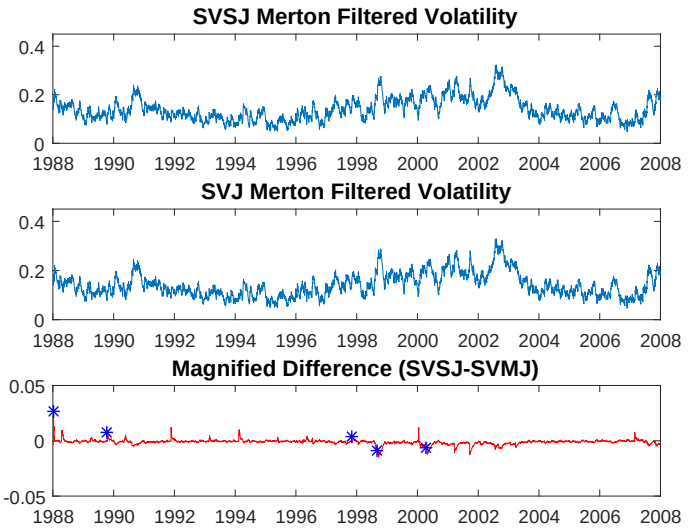

Figure 4.3: SPX Filtered Volatility Differences: SVSJ minus SVJ Merton.

The scale of the Magnified Difference in Figure 4.3 above is a quarter of the scale used to compare the Heston and SVSJ filtered volatilities above in Figure 4.2, and the difference between the SVSJ and the SVJ Merton filtered volatilities is smaller. However, Bates (2006) maintains that there is a difference. The SVJ Merton model initially classifies large observations as jumps. But, the SVJ Merton volatility news impact curve vanishes for large observations, see Bates (2006), p. 932. Alternatively, while the SVSJ volatility news impact curve is damped relative to the Heston news curve, it maintains a positive level for large observations, see Bates (2006), p. 932. These effects dominate when volatility is normal from 1988 to 1998. Hence, the SVJ Merton filtered volatility undershoots the SVSJ filtered volatility immediately following the earlier large loss dates, see Figure 4.3 above. However, from 1998 to 2003, the volatility is higher, see Figure 4.3 above. Thus, the Heston component of the SVJ Merton model dominates,

and due to the sharper Heston news impact curve, the SVJ Merton filtered volatility now overshoots the SVSJ filtered volatility for the last two large loss dates, see Figure 4.3 above. Notice that the SVSJ filtered volatility provides a dynamic adjustment for the volatility regime. Despite this, many people prefer the SVJ Merton filtered volatility, since, like the Heston filtered volatility, it more clearly attributes large losses to volatility. At the cost of being volatility conservative, the SVSJ Merton filtered volatility provides a more dynamic alternative that attributes large moves to jumps.

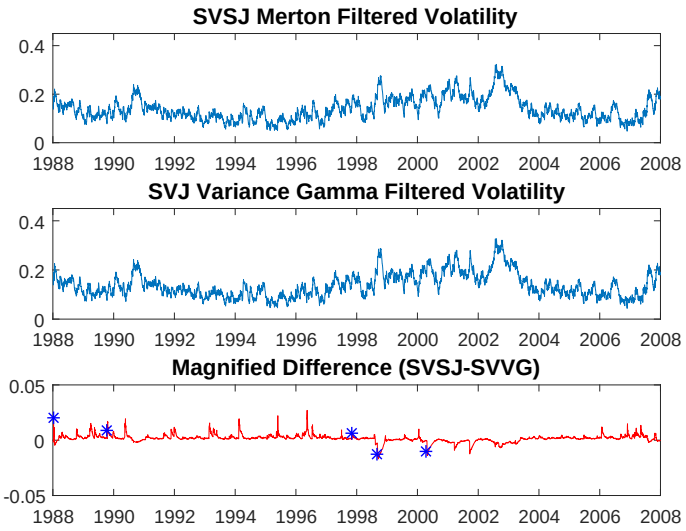

Figure 4.4: SPX Filtered Volatility Differences: SVSJ minus SVVG.

The analysis given with Figure 4.3 for the SVJ Merton case also applies to Figure 4.4 above for the SVVG versus SVSJ filtered volatilities. The main difference is that due to the smaller and more frequent infinite-activity jumps, there are more undershoots from 1988 to 1998, and more overshoots from 1998 to 2003. However, Figure 4.4 above continues to show that there is a noticeable impact on the difference at each of the large loss dates.

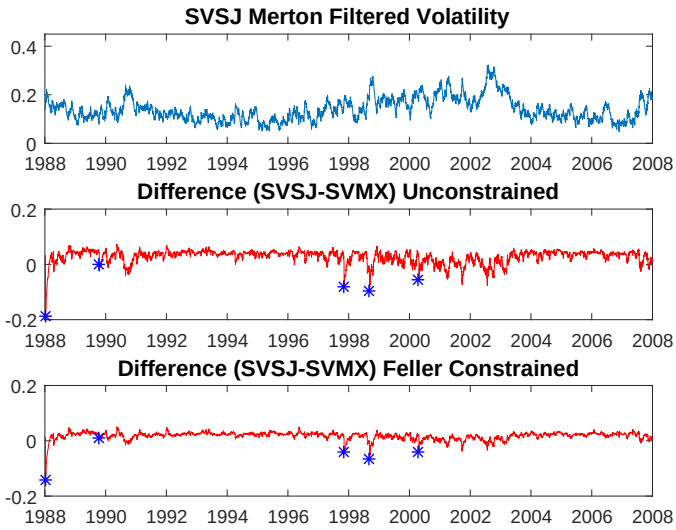

Figure 4.5: SPX Filtered Volatility Differences: SVSJ minus SVJ Meixner.

Figure 4.5 above shows the difference between the SVSJ Merton and SVJ Meixner S&P 500 filtered volatilities, for both the unconstrained SVJ Meixner estimates in Table 4.15 of subsection 4.5.1 above, and the corresponding Feller constrained estimates in Table 4.23 of subsection 4.6.2 below. In both cases, the estimate of λ is large, and the estimate of the jump scale parameter α is small. This suggests that the Meixner jumps are behaving like a second diffusion. Also, the volatility differences in Figure 4.5 above resemble both each other, and the Heston case from Figure 4.2 above.

4.5.3 Comparison to the VIX Volatility Index

The VIX market volatility index is a weighted average of Black-Scholes (1973) implied volatilities computed from a basket of one month near the money market options. It has been published in real time by the CBOE since 1993, and the CBOE has backdated it to 1986 using historical data.

The VIX was originally computed from S&P 100 index options, see Whaley (2000), pp. 13-15. The VIX underlying formally became the S&P 500 in 2003, see the CBOE website. The VIX became a publicly traded asset in 2004, see Hull (2009), p. 297. Our VIX (1988-2007) historical daily index data set was obtained directly from the CBOE website.

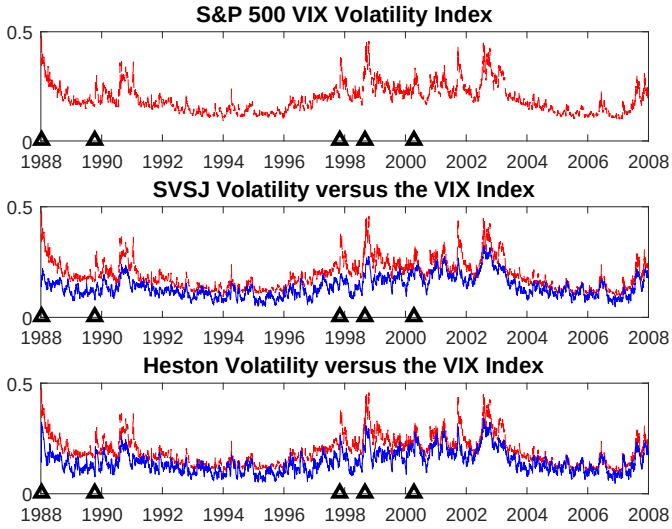

Figure 4.6: S&P 500 SVSJ & Heston Filtered Volatilities vs. the VIX.

The VIX volatility index is obtained from options prices under the risk-neutral measure Q, whereas AML filters the volatility from historical log-returns under the objective measure P. However, Figure 4.6 above shows that the AML filtered S&P 500 daily volatility is quite similar to the VIX. Also notice in Figure 4.6 that at the large loss dates, the VIX overshoots the SVSJ volatility, but less so for the Heston volatility. For the 1988-2007 period, the mean of the VIX is 19.27%, the mean of the SVSJ volatility is 13.96%, and the mean of the Heston volatility is 14.50%. This is consistent with the hypothesis that the volatility is higher under the Q-measure.

4.6 Results from Stock and Index Returns

In addition to the AML estimates for Apple stock (1991-2011) in Table 4.2 of subsection 4.4.2, the DJIA index (1988-2007) in Table 4.3 of subsection 4.4.2, and the S&P 500 index (1988-2007) in Table 4.15 of subsection 4.5.1, this section provides AML estimates for six additional well known individual stocks, each over the 1988-2007 period. The six individual stocks are American Express (AXP), British Petroleum (BP), Citigroup (C), Intel (INTC), JPMorgan Chase (JPM), and Microsoft (MSFT). In each case, estimation is carried out for the SVJ Merton, SVVG, SVJ Meixner, SVSJ Merton, and Heston models. In subsection 4.6.1 below, we table the AML estimates for the six additional stocks. In subsection 4.6.2, Table 4.22 provides the moments of the percentage log-returns for each of the nine estimated assets, and Table 4.24 summarizes the maximum log-likelihood values for convenience. In subsection 4.6.3 we give an analysis of the AML estimates for the nine assets, categorized into four groups, and featuring log-return plots. Apple stock (1991-2011) is the main example from the technology group.

4.6.1 Additional Estimation Results

The implementation details are identical to subsection 4.4.1 above in all respects. Algorithm 4.15 from subsection 4.3.2 is used for AML by FFT with parallel gradients. The AML estimates in Tables 4.16 to 4.21 below are unconstrained in the sense that unbreached bounds are the only constraints. In particular, the Feller condition is not initially imposed as a constraint. The five additional cases in which the Feller condition fails are clearly marked in Tables 4.16 to 4.21 below. All six cases of infinite-activity SVJ models, including the S&P 500 under the SVJ Meixner in Table 4.15 above, in which the Feller condition fails are re-estimated by AML with a Feller constraint. The results are given in Table 4.23 of subsection 4.6.2 below.

Unconstrained AML Estimates: American Express (AXP)								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.123	1.857	0.092	0.530	-0.505	114.9	0.011	0.004
Std. Err.	0.046	0.242	0.016	0.065	0.085	48.81	0.002	0.001
Estimation Time			49.59 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		1.94	
Number of Iterations			36		$\max \ell(\theta)$		13,210	
SVJ Variance Gamma (Feller condition fails)								
Estimate	0.131	1.405	0.089	0.517	-0.559	106.9	0.013	0.004
Std. Err.	0.069	0.226	0.015	0.041	0.109	33.9	0.001	0.002
Estimation Time			41.65 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.154	
Number of Iterations			26		$\max \ell(\theta)$		13,216	
SVJ Meixner (Feller condition fails)								
Estimate	0.132	1.277	0.089	0.517	-0.568	32.53	0.033	0.729
Std. Err.	0.057	0.414	0.023	0.062	0.103	12.45	0.006	0.190
Estimation Time			33.29 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.174	
Number of Iterations			23		$\max \ell(\theta)$		13,217	
SVSJ Merton								
Estimate	0.090	1.773	0.080	0.371	-0.471	858.4	0.019	0.005
Std. Err.	0.053	0.465	0.009	0.043	0.093	1.75	0.001	0.002
Estimation Time			30.12 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.063	
Number of Iterations			17		$\max \ell(\theta)$		13,213	
Heston								
Estimate	0.074	3.578	0.106	0.653	-0.298	-		
Std. Err.	0.057	0.921	0.014	0.070	0.074	-		
Estimation Time			15.81 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.023	
Number of Iterations			11		$\max \ell(\theta)$		13,182	

Table 4.16: American Express (AXP) 1988-2007: Estimation Results.

Unconstrained AML Estimates: British Petroleum (BP)								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.097	2.641	0.050	0.244	-0.421	3.233	0.034	-0.006
Std. Err.	0.049	1.254	0.005	0.052	0.091	2.006	0.009	0.012
Estimation Time			34.03 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.035	
Number of Iterations			21		$\max\ell(\theta)$		14,465	
SVJ Variance Gamma								
Estimate	0.098	2.672	0.047	0.253	-0.442	10.73	0.024	-0.002
Std. Err.	0.049	0.645	0.005	0.035	0.093	2.44	0.002	0.004
Estimation Time			37.91 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.725	
Number of Iterations			21		$\max\ell(\theta)$		14,466	
SVJ Meixner (Feller condition fails)								
Estimate	0.104	2.649	0.026	0.594	-0.759	230.4	0.015	0.077
Std. Err.	0.045	0.472	0.007	0.103	0.121	51.49	0.002	0.166
Estimation Time			39.56 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.083	
Number of Iterations			31		$\max\ell(\theta)$		14,479	
SVSJ Merton								
Estimate	0.101	2.464	0.048	0.221	-0.432	153.5	0.029	-0.002
Std. Err.	0.046	0.686	0.005	0.033	0.086	85.45	0.005	0.004
Estimation Time			38.65 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.152	
Number of Iterations			27		$\max\ell(\theta)$		14,471	
Heston								
Estimate	0.094	4.228	0.053	0.350	-0.378	-		
Std. Err.	0.048	1.244	0.005	0.051	0.079	-		
Estimation Time			19.08 min		$\|\nabla\ell(\theta)\|_{\theta=\widehat{\theta}}$		0.026	
Number of Iterations			15		$\max\ell(\theta)$		14,451	

Table 4.17: British Petroleum (BP) 1988-2007: Estimation Results.

Unconstrained AML Estimates: Citigroup Inc. (C)								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.033	1.585	0.121	0.492	-0.487	2.631	0.067	0.001
Std. Err.	0.057	0.408	0.024	0.049	0.070	1.166	0.012	0.011
Estimation Time			27.18 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.031	
Number of Iterations			16		$\max \ell(\theta)$		13,106	
SVJ Variance Gamma								
Estimate	0.031	1.434	0.122	0.484	-0.513	2.914	0.066	0.004
Std. Err.	0.056	0.375	0.025	0.046	0.060	0.817	0.013	0.012
Estimation Time			29.32 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.018	
Number of Iterations			18		$\max \ell(\theta)$		13,111	
SVJ Meixner								
Estimate	0.035	1.347	0.123	0.492	-0.525	1.117	0.163	0.202
Std. Err.	0.055	0.579	0.033	0.056	0.071	0.409	0.037	0.352
Estimation Time			30.29 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.008	
Number of Iterations			20		$\max \ell(\theta)$		13,112	
SVSJ Merton								
Estimate	0.028	1.561	0.113	0.459	-0.475	47.33	0.060	-0.008
Std. Err.	0.055	0.673	0.027	0.054	0.069	16.82	0.010	0.009
Estimation Time			31.95 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.013	
Number of Iterations			20		$\max \ell(\theta)$		13,107	
Heston								
Estimate	0.091	8.933	0.101	1.038	-0.360	-		
Std. Err.	0.002	0.017	0.000	0.000	0.001	-		
Estimation Time			33.84 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		4,969.2	
Number of Iterations			10		$\max \ell(\theta)$		13,040	

Table 4.18: Citigroup Inc. (C) 1988-2007: Estimation Results.

Unconstrained AML Estimates: Intel Corp. (INTC)								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.203	0.981	0.153	0.325	-0.337	5.477	0.063	-0.021
Std. Err.	0.080	0.445	0.028	0.051	0.098	1.538	0.008	0.010
Estimation Time			28.31 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.030	
Number of Iterations			17		$\max \ell(\theta)$		11,732	
SVJ Variance Gamma								
Estimate	0.202	0.935	0.150	0.323	-0.348	5.466	0.070	-0.024
Std. Err.	0.081	0.255	0.029	0.044	0.096	2.091	0.014	0.012
Estimation Time			27.56 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.032	
Number of Iterations			16		$\max \ell(\theta)$		11,733	
SVJ Meixner								
Estimate	0.202	0.939	0.147	0.332	-0.363	2.299	0.159	-0.746
Std. Err.	0.082	0.374	0.028	0.051	0.094	1.110	0.037	0.339
Estimation Time			29.93 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.011	
Number of Iterations			20		$\max \ell(\theta)$		11,733	
SVSJ Merton								
Estimate	0.219	0.953	0.144	0.294	-0.340	69.45	0.056	-0.016
Std. Err.	0.081	0.385	0.028	0.048	0.097	22.58	0.007	0.008
Estimation Time			32.89 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.026	
Number of Iterations			21		$\max \ell(\theta)$		11,737	
Heston								
Estimate	0.246	5.084	0.173	0.844	-0.257	-		
Std. Err.	0.054	0.513	0.016	0.044	0.168	-		
Estimation Time			26.54 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.336	
Number of Iterations			13		$\max \ell(\theta)$		11,655	

Table 4.19: Intel Corporation (INTC) 1988-2007: Estimation Results.

Unconstrained AML Estimates: JPMorgan Chase (JPM)								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.086	2.302	0.113	0.546	-0.410	3.696	0.042	0.011
Std. Err.	0.060	0.626	0.018	0.059	0.068	2.554	0.014	0.010
Estimation Time			27.18 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.010	
Number of Iterations			16		$\max \ell(\theta)$		13,033	
SVJ Variance Gamma (Feller condition fails)								
Estimate	0.105	1.657	0.103	0.685	-0.516	163.8	0.011	0.001
Std. Err.	0.058	0.594	0.027	0.077	0.082	56.84	0.002	0.001
Estimation Time			37.19 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.124	
Number of Iterations			24		$\max \ell(\theta)$		13,039	
SVJ Meixner (Feller condition fails)								
Estimate	0.106	1.619	0.103	0.669	-0.521	38.38	0.033	0.246
Std. Err.	0.055	0.744	0.031	0.070	0.093	11.06	0.004	0.167
Estimation Time			35.22 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.145	
Number of Iterations			27		$\max \ell(\theta)$		13,042	
SVSJ Merton								
Estimate	0.083	2.054	0.103	0.472	-0.417	137.8	0.035	0.006
Std. Err.	0.055	0.416	0.015	0.044	0.067	58.14	0.006	0.004
Estimation Time			40.91 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.044	
Number of Iterations			28		$\max \ell(\theta)$		13,039	
Heston								
Estimate	0.060	3.010	0.122	0.664	-0.357	-		
Std. Err.	0.062	0.823	0.017	0.067	0.069	-		
Estimation Time			16.36 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.742	
Number of Iterations			9		$\max \ell(\theta)$		13,011	

Table 4.20: JPMorgan Chase (JPM) 1988-2007: Estimation Results.

Unconstrained AML Estimates: Microsoft (MSFT)								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Estimate	0.162	2.112	0.105	0.475	-0.134	6.055	0.053	0.001
Std. Err.	0.067	1.040	0.019	0.085	0.082	1.859	0.008	0.007
Estimation Time			28.26 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.059	
Number of Iterations			17		$\max \ell(\theta)$		12,800	
SVJ Variance Gamma								
Estimate	0.164	1.872	0.102	0.461	-0.146	5.780	0.059	0.003
Std. Err.	0.057	0.586	0.018	0.059	0.087	0.826	0.006	0.008
Estimation Time			30.18 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.060	
Number of Iterations			17		$\max \ell(\theta)$		12,802	
SVJ Meixner								
Estimate	0.166	1.743	0.101	0.478	-0.167	2.308	0.140	0.167
Std. Err.	0.068	0.685	0.023	0.066	0.090	1.914	0.064	0.353
Estimation Time			28.09 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.006	
Number of Iterations			19		$\max \ell(\theta)$		12,804	
SVSJ Merton								
Estimate	0.151	2.039	0.098	0.428	-0.126	123.1	0.045	-0.004
Std. Err.	0.059	0.887	0.015	0.090	0.080	39.08	0.005	0.005
Estimation Time			38.32 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		0.066	
Number of Iterations			25		$\max \ell(\theta)$		12,805	
Heston								
Estimate	0.128	5.812	0.117	0.920	-0.051	-		
Std. Err.	0.037	0.411	0.022	0.064	0.096	-		
Estimation Time			22.51 min		$\|\nabla \ell(\theta)\|_{\theta=\widehat{\theta}}$		1.49	
Number of Iterations			12		$\max \ell(\theta)$		12,738	

Table 4.21: Microsoft Corp. (MSFT) 1988-2007: Estimation Results.

4.6.2 Summary of Exceptions and Log-Likelihoods

In Table 4.22 below we give the sample moments of the percentage log-returns for each of the nine assets estimated in this chapter. As discussed in subsection 4.5.1 above, the skewness and kurtosis, relative to the standard deviation, both give some indication as to whether or not λ will become too large in the SVJ models with infinite-activity jumps. We say that an exception occurs when the Feller constraint is binding.

	Mean	**Std. Dev.**	**Skewness**	**Kurtosis**	**Feller Constraint**	
Indices 1988-2007					**SVVG**	**SVMX**
DJIA	0.0374	0.983	-0.3731	8.247		
SPX	0.0346	0.991	-0.2417	7.326		binding
Stocks 1988-2007						
AXP	0.0424	2.018	-0.0227	5.985	binding	binding
BP	0.0320	1.463	-0.1133	5.680		binding
C	0.0550	2.101	-0.3236	10.943		
INTC	0.0679	2.655	-0.4490	8.498		
JPM	0.0345	2.162	0.0647	8.356	binding	binding
MSFT	0.0896	2.175	-0.1575	7.730		
Stocks 1991-2011						
AAPL	0.0825	3.0549	0.1059	9.358		

Table 4.22: Percentage Log-Return Moments: All Assets: AML.

Table 4.23 below gives the Feller constrained AML estimation results for the six exceptional infinite-activity SVJ models as shown in Table 4.22 above. In Table 4.23 below, both the unconstrained and Feller constrained maximum log-likelihoods values are given, and the Feller constrained values for the maximum log-likelihoods are only a few points less.

Feller Constrained AML Estimates								
Parameter	μ_0	κ	η	ω	ρ	λ	α	β
SVJ Meixner: S&P 500 (SPX)								
Estimate	0.036	2.98	0.025	0.383	-0.91	43.62	0.016	0.017
Std. Err.	0.033	0.481	0.004	0.025	0.049	12.39	0.003	0.187
$\max \ell(\theta)$	Constrained		16,884		Unconstrained		16,892	
SVJ Variance Gamma: American Express (AXP)								
Estimate	0.127	1.46	0.089	0.510	-0.56	98.89	0.013	0.004
Std. Err.	0.067	0.555	0.023	0.049	0.082	28.85	0.004	0.003
$\max \ell(\theta)$	Constrained		13,216		Unconstrained		13,216	
SVJ Meixner: American Express (AXP)								
Estimate	0.123	1.43	0.090	0.507	-0.56	27.84	0.034	0.758
Std. Err.	0.054	0.534	0.021	0.053	0.084	12.01	0.010	0.265
$\max \ell(\theta)$	Constrained		13,217		Unconstrained		13,217	
SVJ Meixner: British Petroleum (BP)								
Estimate	0.091	3.27	0.031	0.449	-0.65	155.7	0.017	0.060
Std. Err.	0.055	0.953	0.007	0.060	0.110	53.05	0.004	0.189
$\max \ell(\theta)$	Constrained		14,475		Unconstrained		14,479	
SVJ Variance Gamma: JPMorgan Chase (JPM)								
Estimate	0.092	1.97	0.104	0.642	-0.50	107.3	0.013	0.001
Std. Err.	0.053	0.620	0.020	0.052	0.072	28.85	0.003	0.002
$\max \ell(\theta)$	Constrained		13,039		Unconstrained		13,039	
SVJ Meixner: JPMorgan Chase (JPM)								
Estimate	0.095	1.92	0.104	0.631	-0.50	25.49	0.038	0.303
Std. Err.	0.064	0.574	0.020	0.055	0.075	9.966	0.007	0.219
$\max \ell(\theta)$	Constrained		13,041		Unconstrained		13,042	

Table 4.23: Exceptional Cases: Feller Constrained AML Estimates.

In Table 4.24 below we give a summary of the maximum log-likelihood values from each estimation in this chapter. This is the main criterion determining the quality of the fit for the AML estimated models. Where the Feller constraint is binding, the Feller constrained log-likelihood value, which is always lower, is used. These entries are marked with an asterisk.

	HES	**SVMJ**	**SVVG**	**SVMX**	**SVSJ**
Indices 1988-2007					
DJIA	16,819	16,851	16,851	16,852	16,855
SPX	16,843	16,869	16,870	16,884*	16,872
Stocks 1988-2007					
AXP	13,182	13,210	13,216*	13,217*	13,213
BP	14,451	14,465	14,466	14,475*	14,471
C	13,040	13,106	13,111	13,112	13,107
INTC	11,655	11,732	11,733	11,733	11,737
JPM	13,011	13,033	13,039*	13,041*	13,039
MSFT	12,738	12,800	12,802	12,804	12,805
Stocks 1991-2011					
AAPL	10,949	11,019	11,026	11,027	11,023

Table 4.24: Maximum Log-Likelihood Values: All Assets: AML.

4.6.3 Analysis of AML Estimates

Aït-Sahalia and Jacod (2012) considers a Brownian semimartingale model with stochastic volatility and independent jumps, see p. 1012. They find that the DJIA index has finite-activity jumps, but that its components may have infinite-activity jumps, see pp. 1036-37. Thus, the SVJ model of subsection 2.6.1, with infinite-activity jumps, might be advantageous for some individual stocks. In this analysis we identify two types of individual

stocks, and surprisingly one index, for which infinite-activity jumps lead to a superior fit. Of the five models that we have considered, the SVJ Meixner model is the best fit, purely on the basis of likelihood, for six of the nine cases covered. For the remaining three cases, the Bates (2006) SVSJ Merton model is the best overall fit.

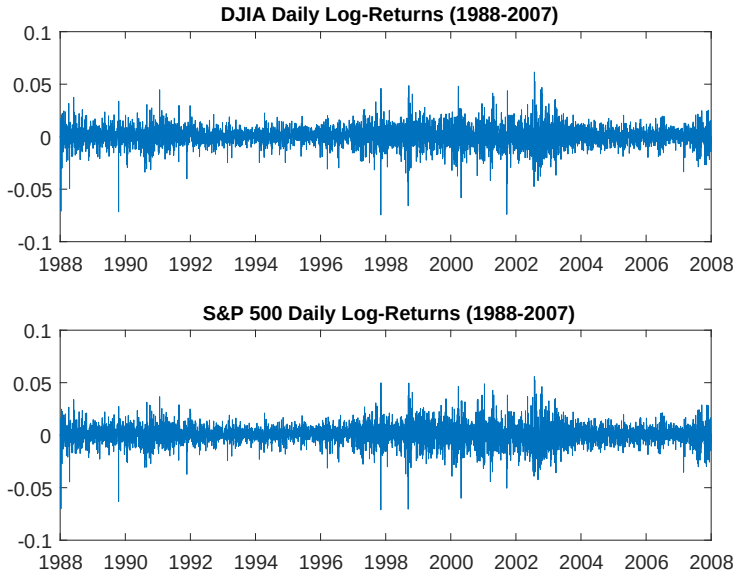

Figure 4.7: Index Returns: the DJIA and the S&P 500.

As illustrated in Figure 4.7 above, stock indices typically have relatively few, large systematic jumps. The likelihood summary in Table 4.24 above shows that for the DJIA index (1988-2007) there is no likelihood gain for the SVJ model with infinite-activity jumps, compared to finite-activity Merton jumps. But, the SVSJ Merton model is the best fit for the DJIA. The S&P 500 index is closely correlated with the DJIA. But, the SPX contains 500 stocks, while the DJIA contains only 30 stocks. Also, the likelihood summary in Table 4.24 above indicates that the SVJ Meixner model is the

best fit for the S&P 500. That is, the S&P 500 (1988-2007) appears to contain a pure jump process of infinite variation, consistent with the options price study of Huang and Wu (2004), see pp. 1425-26. We conclude that while the jumps in stock indices are traditionally best modeled with finite-activity, consistent with Aït-Sahalia and Jacod (2012), a broad index such as the S&P 500 may contain an infinite variation pure jump process, consistent with Carr and Wu (2003).

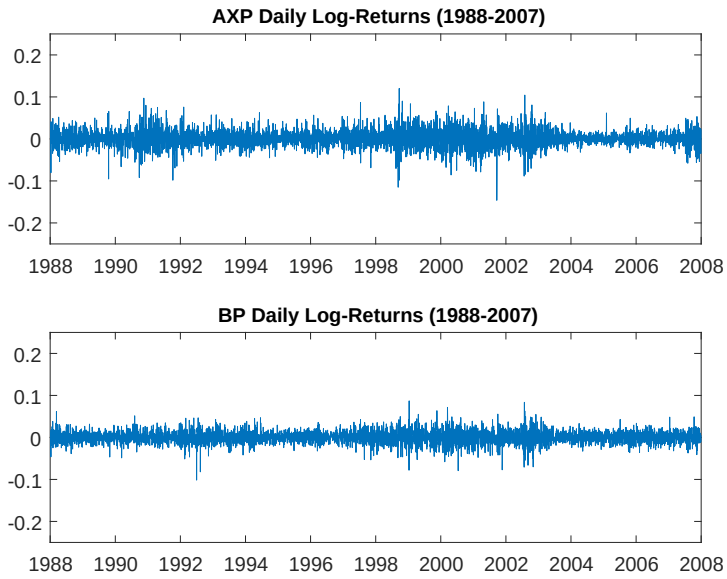

Figure 4.8: Diffusive Returns: American Express & British Petroleum.

In Figure 4.8 above, the returns for AXP and BP over the 1988-2007 period appear to exhibit at most only small jumps. Moreover, Table 4.22 indicates that the kurtosis relative to the standard deviation of the returns is quite small. The likelihood summary in Table 4.24 above states that the infinite variation SVJ Meixner model is the best fit for AXP and BP. However, based on the sparse moments for these two stocks in Table 4.22

above, we believe that the underlying second process is actually a diffusion in both cases. That is, we believe that there are essentially no jumps in either AXP or BP over the 1988-2007 period. This suggests that a model with two diffusions would be superior to a second component with jumps. For example, the model with two independent Heston processes, see Bates (2000), p. 202, is what we would advise here.

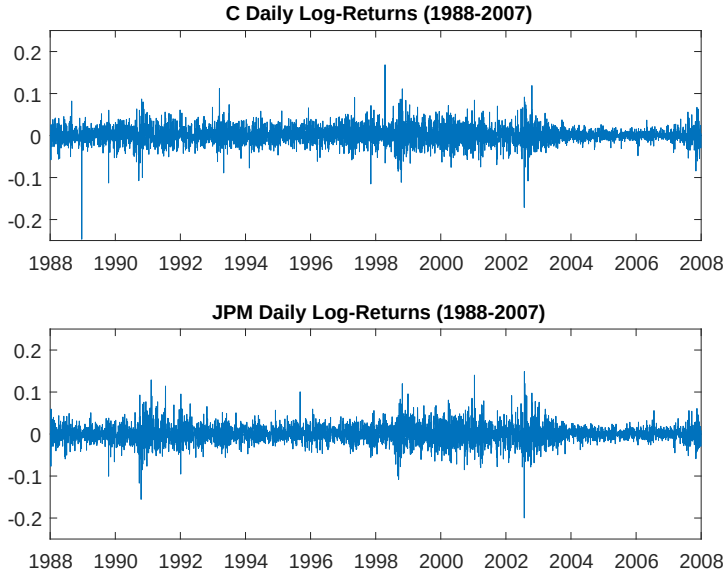

Figure 4.9: Bank Stock Returns: Citigroup (C) & JPMorgan Chase (JPM).

In Figure 4.9 above, it appears that the returns of C and JPM both have a high-activity jump component. Also, Table 4.22 indicates that the kurtosis relative to the standard deviation of the returns should be large enough for infinite-activity to succeed, for each of C and JPM. The likelihood summary, see Table 4.24 above, indicates that for both the returns on C, and the returns on JPM, the SVJ Meixner model is the best fit. However, it should be noted that in the case of the JPM returns, the likelihood summary

in Table 4.24 above indicates that the SVSJ Merton model is also a relatively good fit. We conclude that the log-returns of some large bank stocks can be well modeled with an infinite-activity jump component, of either finite variation or infinite variation.

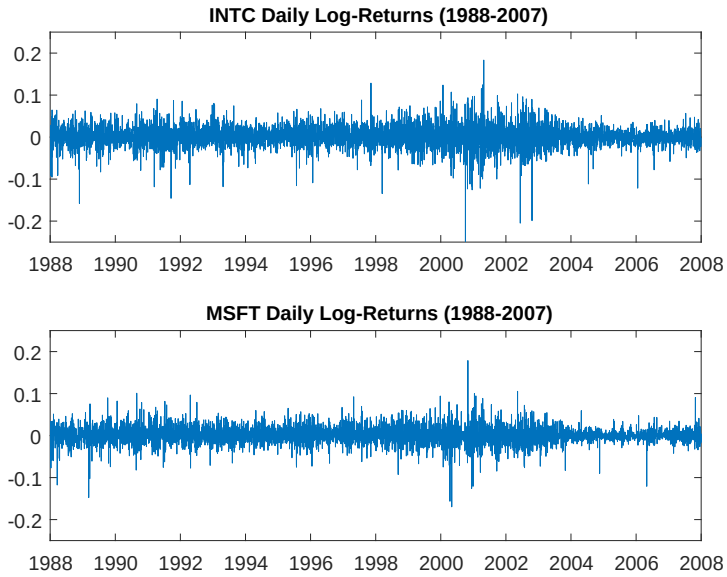

Figure 4.10: Tech Stock Returns: Intel (INTC) and Microsoft (MSFT).

The 1988-2007 period is marked by three regimes. This is particularly the case for INTC and MSFT returns. From 1988 to 1998 there is a bull market, and from 1998 to 2003 there is high volatility period known as the tech-bubble. Then from 2003 to 2007, returns are particularly flat for INTC and MSFT. These regimes make the dynamic jumps of the SVSJ Merton model more appealing. The likelihood summary, see Table 4.24 above, shows that the SVSJ Merton model is the best fit for both INTC and MSFT. But, Figure 4.10 also suggests that there might be an infinite-activity jump component for at least one of INTC and MSFT. According

to Table 4.24, there is no likelihood advantage to infinite-activity for INTC, as there are probably too many large jumps in the INTC returns during the tech-bubble, see Figure 4.10 above. However, Table 4.24 shows that the SVJ Meixner model is almost equal in likelihood to the SVSJ Merton model, for the MSFT returns.

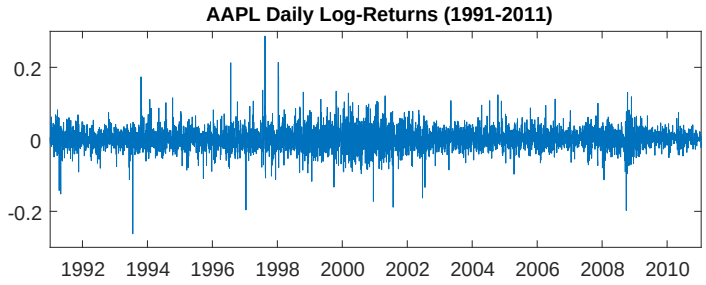

Figure 4.11: The Returns on Apple Stock (AAPL) 1991-2011.

Apple stock (1991-2011) is our leading example for stocks with infinite-activity jumps. Moreover, the crash of 2008 was just another large jump for AAPL, see Figure 4.11 above. But, recall from subsection 4.4.2 above, that the AAPL log-return of -0.7312 on September 29th, 2000 was censored from this sample. Thus, this version of AAPL (1991-2011) does not experience the burst of the tech-bubble. In Table 4.22 above, AAPL exhibits an abnormally large mean and standard deviation, as well as relatively ample skewness and kurtosis. Also, in Figure 4.11 above, AAPL exhibits many jumps, and few regime effects. Therefore, it is not surprising that in the likelihood summary, see Table 4.24 above, models with an infinite-activity jump component fit AAPL log-returns the best. We conclude that the returns of some technology stocks, especially Apple (1991-2011), may be well modeled with an infinite-activity jump component.

The Future of Apple Stock

In this chapter we estimated the P-measure for Apple stock, under five models, based on the log-returns from January 2nd, 1991 to January 19th, 2011. In Chapter 5 we will estimate the option implied Q-measure for Apple stock, under the same five models, based on the January 19th, 2011 closing book of spot options prices. In Chapter 6 we combine each P-measure with its respective option implied Q-measure to construct an optimal Q-measure based on relative entropy for each model. Based on the optimal Q-measure we construct a daily forecast for each model of the realized at-the-money discounted call payoffs for the two years following January 19th, 2011.

4.7 Conclusion

1. The method of Approximate Maximum Likelihood (AML) of Bates (2006) is a fast and accurate way to estimate affine models such as the SVJ and SVSJ models, provided that the joint characteristic function of the log-price and the latent factor is known in closed form, as in Bates (2006), or equivalently as in Appendix A of this thesis.

2. For the SVJ and SVSJ models, we have proven that all three Fourier transform pairs of the AML method satisfy both of the L^1 integrability conditions of the Fourier inversion theorem. In addition to this, we prove in Appendix D below that the first Fourier transform of the AML method may be differentiated twice under the integral with respect to the secondary variable v, as assumed in Bates (2006).

3. Consistent with Aït-Sahalia and Jacod (2012), we have shown that some individual stocks, in particular banking and technology stocks, appear to have infinite-activity jumps in addition to a diffusion with stochastic volatility. Also, consistent with Huang and Wu (2004), we

find that the S&P 500 may contain an infinite variation pure jump process in addition to a diffusion with stochastic volatility. However, the DJIA appears to have finite-activity jumps.

4. We have shown that in AML estimation of the SVJ and SVSJ models, most of the parameters are relatively unbiased, but there is some small bias to κ and λ. Nonetheless, we have shown that for AML estimation, the asymptotic standard errors are reliable.

5. For both the Heston and the SVSJ Merton models, the AML filtered volatility for the S&P 500 is highly similar to the VIX daily index of S&P 500 implied volatilities over the 1988-2007 period.

6. Consistent with Eraker, Johannes, and Polson (2003), we find that the SVJ and SVSJ models can handle 5,000 daily log-returns. Although, it is possible that the Heston model is better with only 2,500 data points, see Lysy (2012). However, based on Bates (2006), the SVJ and SVSJ models with jumps can actually handle 10,000 daily log-returns or more, with AML implemented by quadrature. For future research we propose to use AML by quadrature, and daily rather than annual parameters in the AML method. Daily parameters will allow data at five minute frequency. We propose three months of five minute data for a total of just under 5,000 data points, to preserve speed in AML estimation.

7. Carr, Geman, Madan, and Yor (2003), see p. 364, propose a version of the Barndorff-Nielsen and Shephard (2001) model with time changed jumps and an Ornstein-Uhlenbeck equation for the latent factor. This model is affine, and a joint characteristic function for the log-price and the latent factor can be found. Thus, for future research, this model is a candidate for estimation of the P-measure via AML.

Chapter 5

Least Squares Calibration

5.1 Introduction

Starting with this chapter, and for the remainder of the thesis, the only asset that we will treat is Apple stock. The main goal of this present chapter is to calibrate to spot options prices from a single daily close on Apple stock. This chapter will be the first step in the analysis of the Q-measure for Apple stock, under the Heston model, all three SVJ models, and the SVSJ Merton model. We will calibrate using a suitably modified version of the Carr and Madan (1999) formula, and this needs to be discussed first. It is well known, see Carr and Madan (2009), Table 2, p. 60, that the Carr and Madan (1999) formula produces a negative call price deep out-of-the-money. We show in Table 5.1 of subsection 5.2.3 below that this problem similarly occurs in the log-moneyness version of the Carr and Madan formula, the preferred version for this thesis, see subsection 5.2.2 below. That is, we find, in the Black-Scholes environment of Table 5.1 below, that the log-moneyness version of the Carr and Madan formula produces a negative call price deep out-of-the-money when Simpson's rule is used for quadrature in FFT. Moreover, we find in Table 5.1 of subsection 5.2.3 below, that similar

to subsection 3.4.2 above, this problem is corrected by simply using the trapezoidal rule in place of Simpson's rule, in the log-moneyness version of the Carr and Madan formula. That being established, we proceed, similar to Carr, Geman, Madan, and Yor (2003), see p. 371, to calibrate to a single daily closing book of call options prices. The main differences are that we consider Apple stock instead of the S&P 500, we consider the SVJ and SVSJ models instead of time changed Lévy processes, and we use the trapezoidal rule in the log-moneyness version of the Carr and Madan formula. Also, we only consider one book of daily call option closing prices on Apple stock. The closing date is January 19th, 2011, the last day in the Apple stock log-return data set from subsection 4.4.2 above.

In Section 5.2 we re-evaluate the Carr and Madan (1999) formula as discussed above. In Section 5.3 we introduce the Apple stock call options data set from January 19th, 2011, prepare the data, and then calibrate the Heston, SVJ, and SVSJ models under this data set. Section 5.4 concludes.

5.2 Accuracy in the Carr and Madan Formula

For theoretical results in this section we rely primarily on Carr and Madan (1999), as summarized in subsection 5.2.1 below. The log-moneyness version of the Carr and Madan formula, which appears to have been employed in Huang and Wu (2004), see p. 1419, is described in full in subsection 5.2.2 below. The only difference here from the original is that the damped call price domain is the log-moneyness $\chi = \log\left(\frac{K}{S_0}\right)$ in place of the log-strike $\log(K)$ used in the original Carr and Madan (1999) formula. This does not affect the theoretical results. However, we mention that Carr and Madan (1999), see p. 63, states that the damped call price is L^2, when it can be shown to be L^1 under a condition that they later assume. We pursue this point in Lemma 7.8 of subsection 7.4.1 below. In subsection 7.4.1 we give a

formal proof for the log-moneyness version of the Carr and Madan formula, in a manner suitable for pricing under conditional Monte Carlo, as pursued in Chapter 7 below. But, the only difference between subsection 5.2.2 below and subsection 7.4.1 is that in subsection 7.4.1 we condition $Z_T = \log\left(\frac{S_T}{S_0}\right)$ on some quantity H_T. However, when H_T is independent of Z_T, the two treatments coincide.

5.2.1 The Original Carr and Madan Formula

Let $Y_T = \log(S_T)$ be the terminal log-price with risk-neutral density $f_T^Q(y)$, and and let $k = \log(K)$ be the log-strike price. Then for some damping parameter $a > 0$, define the damped call price by

$$C_T^a(k) = e^{-rT} \int_{-\infty}^{+\infty} e^{ak} \left(e^y - e^k\right)^+ f_T^Q(y)\, dy. \tag{5.1}$$

The damped call price $C_T^a(k)$ in equation (5.1) above is at least L^2, see Carr and Madan (1999), p. 63. Now let $\phi_{Y_T}(u)$ be the characteristic function of the terminal log-price Y_T. Then the Fourier transform of the damped call price is given by

$$\widehat{C}_T^a(u) = \frac{e^{-rT}\phi_{Y_T}(u - i(a+1))}{a^2 + a - u^2 + iu(2a+1)}. \tag{5.2}$$

Moreover, under the mild assumption that $E\left[S_T^{a+1}\right] < \infty$, $\widehat{C}_T^a(u) \in L^1$, see Carr and Madan (1999), pp. 64-65. Thus, by an alternative version of the Fourier inversion theorem, see Rudin (1987), p. 187, we have that

$$C_T^a(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuk} \widehat{C}_T^a(u)\, du. \tag{5.3}$$

Hence, when $E\left[S_T^{a+1}\right] < \infty$, the undamped call price $C_T(k)$ is given by

$$C_T(k) = \frac{e^{-ak}}{\pi} \operatorname{Re} \int_0^{\infty} e^{-iuk} \widehat{C}_T^a(u)\, du, \tag{5.4}$$

see Carr and Madan (1999), p. 64. Thus, assuming that $E\left[S_T^{a+1}\right] < \infty$, the Carr and Madan FFT formula is given by

$$C_T(k_m) \approx \frac{e^{-ak_m}}{\pi} \operatorname{Re} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(m-1)} e^{ibu_j} \widehat{C}_T^a(u_j) w_j \Delta u, \qquad (5.5)$$
$$m = 1, 2, ..., N,$$

where the w_j are defined by Simpson's rule, $w_1 = \frac{1}{3}$, and $w_j = \left\{ \begin{array}{c} \frac{2}{3} \text{ if } j \text{ is odd} \\ \frac{4}{3} \text{ if } j \text{ is even} \end{array} \right\}$, $j = 2, 3, ..., N$, see Carr and Madan (1999), p. 68.

5.2.2 The Log-Moneyness Carr and Madan Formula

This version of the Carr and Madan formula is essentially the same as the original in subsection 5.2.1 above. However, it is precisely this version that we use in this chapter and this thesis. Let $Z_T = \log\left(\frac{S_T}{S_0}\right)$ have risk-neutral density $f_T^Q(z)$, and and let $\chi = \log\left(\frac{K}{S_0}\right)$ be the log-moneyness. Then for some damping parameter $a > 0$, define the damped call price by

$$C_T^a(\chi) = S_0 e^{-rT} \int_{-\infty}^{+\infty} e^{a\chi} \left(e^z - e^{\chi}\right)^+ f_T^Q(z)\, dz. \qquad (5.6)$$

If $E\left[S_T^{a+1}\right] < \infty$, then the damped call price $C_T^a(\chi)$ in equation (5.6) is L^1, see Lemma 7.8 of subsection 7.4.1 below with H_T independent of Z_T. Now let $\phi_{Z_T}(u)$ be the characteristic function of Z_T. Then the Fourier transform of the damped call price is given by

$$\widehat{C}_T^a(u) = \frac{S_0 e^{-rT} \phi_{Z_T}(u - i(a+1))}{a^2 + a - u^2 + iu(2a+1)}, \qquad (5.7)$$

where similar to Carr and Madan (1999), pp. 64-65, if $E\left[S_T^{a+1}\right] < \infty$, then $\widehat{C}_T^a(u)$ in equation (5.7) above is L^1. Thus, by the Fourier inversion theorem, Theorem 3.6, we have that

$$C_T^a(\chi) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iu\chi} \widehat{C}_T^a(u)\, du, \qquad (5.8)$$

and similar to Carr and Madan (1999), p. 64, if $E\left[S_T^{a+1}\right] < \infty$, then the undamped call price $C_T(\chi)$ is given by

$$C_T(\chi) = \frac{e^{-a\chi}}{\pi} \operatorname{Re} \int_0^{\infty} e^{-iu\chi} \widehat{C}_T^a(u)\, du. \tag{5.9}$$

Again, similar to Carr and Madan (1999), p. 68, when $E\left[S_T^{a+1}\right] < \infty$, the log-moneyness version of the Carr and Madan FFT formula is given by

$$C_T(\chi_k) \approx \frac{e^{-a\chi_k}}{\pi} \operatorname{Re} \sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \widehat{C}_T^a(u_j)\, w_j \Delta u, \tag{5.10}$$
$$k = 1, 2, ..., N.$$

For the log-moneyness version in equation (5.10) above we will allow the w_j to follow either

1. Simpson's rule: $w_1 = \frac{1}{3}$, and $w_j = \left\{ \begin{array}{c} \frac{2}{3} \text{ if } j \text{ is odd} \\ \frac{4}{3} \text{ if } j \text{ is even} \end{array} \right\}$, $j = 2, 3, ..., N$, see Carr and Madan (1999), p. 68, or

2. The trapezoidal rule: $w_1 = w_N = \frac{1}{2}$, and $w_j = 1$, for $j = 2, 3, ..., N-1$, see Briggs and Henson (1995), p. 360.

5.2.3 Accuracy Analysis

Table 5.1 below examines the problem of negative call prices computed by the log-moneyness version of the Carr and Madan formula deep out-of-the-money, with Simpson's rule used for the quadrature weights. By using the exact price from the Black-Scholes model, we have discovered that the problem in the log-moneyness version of the Carr and Madan formula is indeed the use of Simpson's rule, as was recommended in Carr and Madan (1999), see p. 68, for the original formula. We suspect that the issue is the same with the original Carr and Madan formula. In Table 5.1 below, $S_0 = 100$, the risk-free rate is $r = 0.03$, $T = 1$ year, and the Black-Scholes

volatility is $\sigma = 0.30$. The FFT parameters are $N = 2^{13}$, and $\Delta u = \frac{1}{4}$. In this thesis we use the value $a = 1.5$ for the damping parameter. Also in Table 5.1 below, when a spline interpolant is used to obtain the call price, it is based on the uniform discretization $\chi_k \in [-\pi, \pi]$ of the log-moneyness points. To ensure accuracy in this subsection, all computations were done on an Intel Xeon 2xE5-2643v3 dual processor at 3.4 GHz.

Call Prices	**Simpson's Rule**		
Strike Price	Exact Price	FFT Linear	FFT Spline
300	2.31144E-03	2.31137E-03	2.31123E-03
400	3.66372E-05	3.64252E-05	3.64201E-05
500	8.17711E-07	6.00932E-07	6.00634E-07
600	2.48365E-08	−1.92239E-07	−1.92240E-07
700	9.83587E-10	−2.16091E-07	−2.16092E-07
Call Prices	**Trapezoidal Rule**		
Strike Price	Exact Price	FFT Linear	FFT Spline
300	2.31144E-03	2.31158E-03	2.31144E-03
400	3.66372E-05	3.66422E-05	3.66372E-05
500	8.17711E-07	8.18009E-07	8.17711E-07
600	2.48365E-08	2.48375E-08	2.48365E-08
700	9.83587E-10	9.84007E-10	9.83591E-10

Table 5.1: FFT Call Prices: Black-Scholes DOTM: $S_0 = 100$.

It is clear in the upper panel of Table 5.1 above that Simpson's rule leads to a negative FFT call price deep out-of-the-money. However, the lower panel of Table 5.1 above shows that the trapezoidal based FFT price does not behave in this way. Moreover, Table 5.1 above also shows that when combined with spline interpolation the trapezoidal rule is accurate to four or five significant digits against the Black-Scholes formula, for the

pricing of call options deep out-of-the-money. Next we consider the pricing performance of near-the-money call options in the log-moneyness version of the Carr and Madan (1999) formula. We consider the relative pricing errors under linear and spline interpolation in the same setting as Table 5.1 above. Specifically, we compare Simpson's rule to the trapezoidal rule for the pricing of call options near-the-money, using the log-moneyness version of the Carr and Madan formula, see Table 5.2 below.

Rel. Errors	**Simpson's Rule**		**Trapezoidal Rule**	
Strike Price	Linear	Spline	Linear	Spline
80	8.1E-07	8.6E-09	8.2E-07	1.5E-12
90	3.2E-06	1.2E-08	3.2E-06	5.0E-12
100	1.6E-08	1.6E-08	6.7E-16	6.7E-16
110	3.0E-06	2.3E-08	3.0E-06	2.2E-12
120	1.7E-05	3.4E-08	1.7E-05	4.4E-11

Table 5.2: FFT Relative Call Pricing Errors: Black-Scholes: $S_0 = 100$.

For performance near-the-money Table 5.2 above shows that in the Black-Scholes case the trapezoidal rule is as good as Simpson's rule with linear interpolation, and substantially better with splines. It should be noted that in Table 5.2 above, the phenomenon that the at-the-money price is computed to near machine precision under the trapezoidal rule is due to a special feature of the log-moneyness version of the Carr and Madan formula, which uses the non-centred and shifted form of FFT. That is, if $K = S_0$, then $\chi = \log\left(\frac{K}{S_0}\right) = 0$. It turns out that zero corresponds to the highest frequency node in the non-centred and shifted version of FFT, see Briggs and Henson (1995), p. 69. Altogether, we recommend the trapezoidal rule with splines, and the log-moneyness version of the Carr and Madan formula, for accuracy in call option pricing by FFT.

5.3 Calibration to Apple Stock Call Options Data

The ultimate goal of this section is to provide least squares estimates for the parameters of the Heston model, all three SVJ models, and the SVSJ Merton model, each under the Q-measure, based on the closing call options prices of Apple stock on Wednesday, January 19th, 2011. First of all, in subsection 5.3.1 we provide details on the preparation of the Apple stock call options data set from the raw data obtained from Market Data Express at the CBOE. Also, in subsection 5.3.1, we describe the nonlinear least squares likelihood method used in this chapter. In subsection 5.3.2 we provide the least squares calibration results, see Table 5.3 below.

5.3.1 Data Preparation and Estimation Methodology

We obtained market closing American options quotes for Apple stock on the third Wednesday of January 2011, from Market Data Express at the CBOE. At this time in January 2011, Apple stock was not paying any dividends. The next dividend date turned out to be August 9, 2012, at which time dividends became regular. We assume that this policy was unknown to market participants on January 19, 2011. Hence, we treat Apple stock, on that date, as a stock with no dividends. American call options on a stock paying no dividends may be treated as European call options, see Björk (2009), p. 112. Moreover, estimation based solely on call options should not unduly bias the results, see Bakshi, Cao, and Chen (1997), p. 2015. Hence, we discarded the American puts.

The call options were selected as follows. Any contract without strictly positive open interest, volume, bid price, or ask price was discarded. Then, only contracts with either substantial volume, or substantial open interest were retained. Next, contracts were discarded until, at each maturity, both the bid and ask price came to follow a strictly decreasing sequence

with respect to an increasing strike price, with volume used as the deciding factor in the event of a conflict. At this point the call price was taken to be the mean of the bid price and the ask price. Two more selection criteria were applied. First, assuming that the call price C is European, the arbitrage-free condition given by

$$C \geq S_0 - Ke^{-rT}, \tag{5.11}$$

was enforced, see Björk (2009), p. 111. Lastly, we insisted that the bid-ask spread be less than 10% of the call price. The resulting data set includes 194 call options quotes across seven maturities, with exact maturity dates of 22, 41, 61, 123, 192, 253, and 503 days. Given a 21 day month, these maturity dates correspond roughly to 1, 2, 3, 6, 9, 12, and 24 months. The corresponding numbers of options quotes at each maturity are given by 17, 22, 29, 39, 28, 38, and 21. The moneyness range for this data set is

$$0.69 \leq \frac{S_0}{K} \leq 1.54, \tag{5.12}$$

with $S_0 = 338.84$ at the close on January 19, 2011. A few words are in order here. First, the underlying is Apple stock, and not an index such as the S&P 500. In general, Apple was widely expected to outperform following January 2011, and objectively there is some existing liquidity in the data, further out-of-the-money than usual. Secondly, January 19th, 2011 was only two days after Steve Jobs announced what turned out to be his final medical leave of absence due to cancer. Consequently, there is some existing liquidity in the data, further in-to-the-money than usual. Lastly, in order to estimate jumps, it is important to include quotes that are both far in-to-the-money, and far out-of-the-money, see Cont and Tankov (2004b), p. 29. The risk-free rate was taken to be the average three-month LIBOR rate for January 2011 minus one tenth of a basis point. This gives $r = 0.2934\%$, see the fedprimerate website.

Least Squares Likelihood Methodology

We apply the following estimation methodology. For the Heston, SVJ, and SVSJ models, the characteristic function of the log-price is known in the exponentially affine form

$$\phi_{Z_T}(u) = \exp\left(C(u) + D(u)\sigma_0^2\right), \text{ with } Z_T = \log\left(S_T/S_0\right), \tag{5.13}$$

where $\mu_0 = r$ in the coefficients $C(u)$ and $D(u)$ taken from Appendix A. Let $E^Q\left[S_T^{a+1}\right] < \infty$ at each maturity T be assumed. Then, European call prices can be obtained, as in subsection 5.2.2 above, from the log-moneyness version of the Carr and Madan formula. Also, though not required by the Carr and Madan (1999) formula, by Theorem 3.17 from subsection 3.3.3, assuming $\rho < 0$, each of these log-price CFs can be shown to be L^1. Our European call price data set is based on the average of the bid price and the ask price. We represent this data set as the vector C given by

$$C_h = C(T_h, K_h), \; h = 1, 2, ..., n, \tag{5.14}$$

where T_h is the exact time to maturity in years, assuming a 252 day year, and K_h is a strike price at that maturity. The parameter vector for both the SVJ and SVSJ models is given by

$$\theta = \left(\sigma_0^2, \kappa, \eta, \omega, \rho, \lambda, \alpha, \beta\right), \tag{5.15}$$

where the Heston model is a special case with no jumps. Note that similar to Bakshi, Cao, and Chen (1997), p. 2016, we intend to treat σ_0^2 in the same manner as the other parameters. This distinction was pointed out in Christoffersen, Heston, and Jacobs (2009), see p. 1925. We believe that this is the most suitable choice for a small data set. Let the corresponding vector of model prices be given by $C(\theta)$. Then, we may write the sum of squared pricing errors as

$$S(\theta) = \|C(\theta) - C\|^2. \tag{5.16}$$

Under the assumption that the pricing errors are *i.i.d.* $N\left(0,\sigma_e^2\right)$, the normal log-likelihood is given by

$$\ell\left(\theta,\sigma_e^2\right)=-\frac{n}{2}\log\sigma_e^2-\frac{1}{2\sigma_e^2}S\left(\theta\right), \tag{5.17}$$

see Seber and Wild (2003), p. 32. However, for fixed θ, the log-likelihood in (5.17) is maximized when $\sigma_e^2=\frac{S(\theta)}{n}$. Hence, we may use the concentrated log-likelihood given by

$$\ell_M\left(\theta\right)=\ell\left(\theta,\frac{S\left(\theta\right)}{n}\right)=-\frac{n}{2}\log S\left(\theta\right)+\frac{n}{2}\left(\log n-1\right). \tag{5.18}$$

Since the log-likelihood in (5.18) is maximized when $S\left(\theta\right)$ is minimized,

$$\widehat{\sigma}_e^2=\frac{S\left(\widehat{\theta}\right)}{n} \tag{5.19}$$

is the MLE of σ_e^2. Moreover, it can be shown that the information matrix $I_M\left(\theta\right)$, based on the concentrated log-likelihood $\ell_M\left(\theta\right)$, is given by

$$\begin{aligned} I_M\left(\theta\right) &= -\frac{\partial^2\ell_M\left(\theta\right)}{\partial\theta\partial\theta'}|_{\widehat{\theta}} \\ &= \frac{1}{2\widehat{\sigma}_e^2}\frac{\partial^2 S\left(\theta\right)}{\partial\theta\partial\theta'}|_{\widehat{\theta}}, \end{aligned} \tag{5.20}$$

see Seber and Wild (2003), pp. 38-39. On the basis of equation (5.20), we obtain asymptotic standard errors for the parameter estimates.

5.3.2 Estimates and Standard Errors

All computations in this section were done on an Intel 2xE5-2643v3 at 3.4GHz. The estimates in Table 5.3 below are based on spot options prices for Apple stock on January 19, 2011, whereas the estimates in Table 4.2 from subsection 4.4.2 are based solely on daily log-return data for Apple stock from January 2, 1991 to January 19, 2011. When we calibrate to options prices, the stock price is not observed directly. Moreover, by choosing

to estimate only from the current book of options prices, we sacrifice the historical structure of the volatility in favour of current market expectations for the near future. The most striking difference between Table 5.3 below and Table 4.2 is in Heston model. In Table 4.2, for the Heston model, both the rate of mean reversion κ and the volatility of the volatility ω are far too large, indicating a need for jumps in the model. However, in Table 5.3 below, the Heston model parameter estimates are broadly consistent with the literature on call options price based least squares estimation. For example, see the Heston model parameter estimates based on the S&P 500 in Christoffersen, Heston, and Jacobs (2009), p. 1922. In Table 5.3 below, the t-statistics for the jump parameters λ, α, and β in the SVJ and SVSJ models are weak relative to the t-statistics for the variance parameters. Also, the RMSE for the Heston model is nearly as low as the RMSE for any of the jump models. However, the same problem, with the t-statistics for the jump parameters, exists in the study by Huang and Wu (2004), see pp. 1425-26, where models similar to the SVJ and SVSJ models were calibrated. Moreover, in the Huang and Wu (2004) study approximately 16,402 options prices were included in the in-sample data, see p. 1421. Thus, adding more options prices to the data will not necessarily solve this problem. We conclude that jumps are difficult to estimate in this setting. Nonetheless, jumps should be included in options price models, as they provide a better accounting of the skewness and the kurtosis, see Bakshi, Cao, and Chen (1997), p. 2017 for the SVJ Merton model, and Bates (2000), p. 216 for the SVSJ Merton model. Regarding this matter of kurtosis, observe that in Table 5.3 below, the Heston estimate for ω is nearly twice the value of the long run volatility. Too much kurtosis has been left to be explained by the volatility of the volatility ω. However, notice that in Table 5.3 below, the Feller condition, $2\kappa\eta > \omega^2$, holds in all five models.

Apple Stock Call Options Calibration Results								
θ_Q	σ_0^2	κ	η	ω	ρ	λ	α	β
SVJ Merton								
Est	0.058	1.767	0.136	0.567	-0.662	1.637	0.095	-0.008
S.E.	0.005	0.117	0.004	0.035	0.040	1.346	0.027	0.008
t-stat	11.25	15.05	33.74	16.02	-16.7	1.217	3.468	-0.992
Iter	126	**Time**	12.1 s	$\|\ell_M(\theta)\|_{\theta=\widehat{\theta}}$		0.024	**RMSE**	0.1976
SVJ Variance Gamma								
Est	0.057	1.736	0.135	0.559	-0.668	1.337	0.111	-0.014
S.E.	0.004	0.124	0.004	0.034	0.041	0.595	0.022	0.011
t-stat	14.90	14.02	36.76	16.28	-16.4	2.248	5.067	-1.330
Iter	165	**Time**	16.5 s	$\|\ell_M(\theta)\|_{\theta=\widehat{\theta}}$		7E-04	**RMSE**	0.1978
SVJ Meixner								
Est	0.056	1.720	0.134	0.554	-0.670	0.396	0.293	-0.380
S.E.	0.003	0.121	0.003	0.041	0.039	0.196	0.069	0.217
t-stat	16.48	14.19	39.36	13.46	-17.4	2.019	4.255	-1.750
Iter	346	**Time**	35.4 s	$\|\ell_M(\theta)\|_{\theta=\widehat{\theta}}$		0.001	**RMSE**	0.1979
SVSJ Merton								
Est	0.051	1.805	0.103	0.441	-0.706	64.25	0.082	-0.003
S.E.	0.004	0.100	0.009	0.039	0.043	22.64	0.009	0.006
t-stat	12.55	17.98	11.53	11.36	-16.6	2.838	9.171	-0.538
Iter	120	**Time**	12.4 s	$\|\ell_M(\theta)\|_{\theta=\widehat{\theta}}$		0.116	**RMSE**	0.1960
Heston								
Est	0.071	1.973	0.148	0.604	-0.556	-		
S.E.	0.001	0.079	0.001	0.021	0.011	-		
t-stat	125.4	24.99	114.9	29.10	-49.3	-		
Iter	44	**Time**	2.9 s	$\|\ell_M(\theta)\|_{\theta=\widehat{\theta}}$		7E-04	**RMSE**	0.2035

Table 5.3: Apple Inc. Call Options January 19, 2011: Calibration Results.

Remark 5.1 (Asymptotic Distribution) *Let the log-likelihood* $\ell_M(\theta)$ *be given by equation* (5.18) *above. Then if* $\widehat{\theta}_n$ *is a consistent root of*

$$\frac{\partial \ell_M(\theta)}{\partial \theta} = 0, \tag{5.21}$$

an approximate distribution for θ *is given by*

$$\theta \sim_a N\left(\widehat{\theta}_n, I_M\left(\widehat{\theta}_n\right)^{-1}\right), \tag{5.22}$$

where $I_M(\theta)$ *is given by equation* (5.20) *above.*

See Lehmann and Casella (1998), p. 449.

Table 5.3 above further shows that by introducing jumps to assist with skewness and kurtosis, as in the SVJ and SVSJ models, the estimates of ω are made smaller, and thus more reasonable. However, regarding skewness, also notice in Table 5.3 above that the estimate of ρ makes a paradigm shift to the left whenever jumps are introduced. This shift in ρ is typically both smaller and to the right with S&P 500 options data, see Bates (2000), Table 2, p. 203, and Bakshi, Cao, and Chen (1997), Table III, p. 2018. This discrepancy in the behaviour of ρ is explained by the jump skewness parameter β. Notice in Table 5.3 above that the estimates of β are relatively small, indeed quite small compared to Bates (2000), Table 2, p. 203, and Bakshi, Cao, and Chen (1997), Table III, p. 2018, for the SVJ and SVSJ Merton models, respectively. Consequently, when jumps are introduced, ρ moves to the left in Table 5.3 above, for options on Apple stock from January 19, 2011.

RMSE and the Choice of Jumps

Notice in Table 5.3 above that the SVJ Meixner and SVJ Variance Gamma models have almost the same RMSE. This is similar to the results for spot options prices on the S&P 500, see Carr, Geman, Madan, and Yor (2003),

pp. 373-74, under the time changed NIG and time changed Variance Gamma models. The NIG and the Meixner processes both have a stability index of one, see Aït-Sahalia and Jacod (2012), p. 1020, and Madan and Yor (2008), p. 43, respectively. To obtain an advantage for infinite variation jumps, it appears to be necessary to introduce a jump model such as the log-stable process, where the stability index is a model parameter, see for example Huang and Wu (2004), p. 1425. What is surprising about the Apple stock call options data set of this thesis is that there appears to be no advantage in RMSE for infinite-activity jumps over finite-activity Merton jumps. We attribute this to the smallness of the spot options data set, as there is such an advantage in the much larger historical data set of Huang and Wu (2004), see p. 1425. However, we find, when the estimate of λ is suitably large in the SVSJ Merton model, as it is in Table 5.3 above, that the RMSE for the SVSJ Merton model is slightly lower than that of the other models.

5.4 Conclusion

1. The first stage of the analysis of the Q-measure for Apple stock is complete, and the Heston, SVJ, and SVSJ Merton model estimates from Table 5.3 in subsection 5.3.2 will be used for stabilization in the Parametric Minimum Entropy Martingale Measure of Chapter 6.

2. Under the Apple stock call options data set of this thesis, on the basis of RMSE, all of the models with jumps are slightly better than the Heston model. But, within the SVJ models, no particular model stands out. However, the SVSJ Merton model is slightly better than the others on the basis of RMSE.

3. Use of the trapezoidal rule in the log-moneyness version of the Carr and Madan formula prevents negative call prices deep out-of-the-money.

Chapter 6

An FFT Based Approach to Relative Entropy

6.1 Introduction

In this chapter we present a minimum relative entropy method for selecting an optimal equivalent martingale measure where the model for the market is incomplete due to the presence of stochastic volatility, jumps, or stochastic jump intensity. The method was designed to take advantage of the closed form characteristic function for the normalized log-price as an attribute of the affine SVJ and SVSJ models. But, the method can be adapted to any affine model. Our method is parametric, and we give results for the SVSJ Merton model, the three SVJ models, and the Heston model, based on the P-measures, denoted here by $\theta_P \in \Omega_P$, for Apple stock log-returns, from January 2nd 1991 to January 19th 2011, as estimated for Table 4.2 in subsection 4.4.2. We offer a computable approximation to an alternative version of the minimum relative entropy problem. Let T be finite, and for any $t \in (0, T]$, let t be an option maturity date. Consider the measures

Q_t and P_t corresponding to the respective densities $f^Q(z_t)$ and $f^P(z_t)$ of the normalized log-price process $Z_t = \log\left(\frac{S_t}{S_0}\right)$ at maturity t. We obtain these densities by Fourier inversion of their characteristic functions, and thus obtain the relative entropy of Q_t given P_t. More specifically, we use the symmetric relative entropy, see Kullback and Leibler (1951), p. 81. The symmetric relative entropy is the sum of two relative entropies with reversed arguments. As such, it has the property that if finite, then Q_t is equivalent to P_t. After showing how to compute the symmetric relative entropy for any $t \in (0, T]$, we sample the maturity dates uniformly, $m = 1, 2, ..., M$, for some finite M, with $t(m) = \frac{mT}{M}$, and compute the average of the sample symmetric relative entropies. The quantity that we minimize is this average value. Based on the January 19th, 2011 Apple stock call options data set, we choose $T = 2$ years, and also $M = 24$ months, implying monthly maturities. Alternative choices of M are considered in subsection 6.5.3 below. When the average symmetric relative entropy is finite, the symmetric relative entropy of $Q_{t(m)}$ given $P_{t(m)}$ is finite for each $m = 1, 2, ..., M$, suggesting that $Q \sim P$ on $(0, T]$ holds approximately. The P-measure is fixed by the parameters $\theta_P \in \Omega_P$ from Table 4.2 in subsection 4.4.2 for Apple stock (1991-2011). Therefore, we minimize the average value with respect to the risk-neutral parameters $\theta \in \Omega_Q$ of the normalized log-price under the Q-measure. We naturally assume that the parametric form of the model remains closed to the original form of the model during minimization. This assumption has a key benefit. For, if the martingale parameter constraints for the proposed models are known, see Theorem 2.26 in subsection 2.5.2, Theorem 2.31 in subsection 2.6.1, and Theorem 2.35 in subsection 2.6.2, then we can ensure that the solution is a martingale simply by imposing the martingale constraints during numerical minimization. However, the parametric approach has a drawback. Relative entropy is not necessarily

convex in the model parameters, see Cont and Tankov (2004b), p. 11. For this reason we propose a two stage optimization routine to minimize the average symmetric relative entropy, equivalent to a Tikhonov regularization scheme, see Engl, Hanke, and Neubauer (1996), p. 250. Within this routine, the fixed parameters, $\theta_Q^* \in \Omega_Q$, implied by call options based on the closing book of call options prices on Apple stock for the date January 19th 2011, taken from Table 5.3 in subsection 5.3.2, play the role of the prior, see Section 6.3 below. We refer to our result as the parametric minimum entropy martingale measure (PMEMM). The PMEMM results in Section 6.5 below include forecasts of the at-the-money discounted call option payoffs on Apple stock for each model. The evaluation day is January 19th, 2011, and the maturities run from 15 days to 504 days past this date.

The remainder of the chapter is structured as follows. In Section 6.2 we explain how to compute the average symmetric relative entropy. In Section 6.3 we present our two stage optimization routine, equivalent to a Tikhonov regularization of the average symmetric relative entropy. In Section 6.4 we review the main literature on minimum relative entropy. Section 6.5 gives the PMEMM results along with the forecasts of the at-the-money discounted call payoffs. Section 6.6 concludes.

6.2 Computing the Average Relative Entropy

6.2.1 Models and Martingale Constraints

Recall from subsection 2.6.1 that the P-dynamics of the SVJ model are given in log-price form by

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \psi_X\left(-i\right)\right)dt + \sigma_t dW_t^{(S)} + dX_t, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right] = \rho dt, \end{aligned} \tag{6.1}$$

where the jumps X_t are either Merton, Variance Gamma, or Meixner. Also, from subsection 2.6.2 the P-dynamics of the SVSJ model are

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \sigma_t^2\psi_X(-i)\right)dt + \sigma_t dW_t^{(S)} + dX(V_t), \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right] = \rho dt, \\ dV_t &= \sigma_t^2 dt, \end{aligned} \tag{6.2}$$

where the jumps of the time changed Lévy process $X(V_t)$ are Merton. The Heston model of equation (2.59) in subsection 2.5.2 may be viewed as a special case of either the SVJ or SVSJ model with no jumps. Moreover, since the stochastic jump intensity in the SVSJ model is simply the latent variance σ_t^2 scaled by the intensity λ, the SVJ and SVSJ Merton models share a common parameter space, given the same jumps.

Remark 6.1 (Initial Variance) *Recall from subsection 4.2.5 that under the AML estimated P-measure the initial variance $\sigma_0^2 = \sigma_{n_0}^2$ is random, and does not have an estimate. In this chapter we will assume for convenience that σ_0^2 is a parameter under the fixed P-measure with an estimate equal to the estimated value of the long run variance η.*

Based on Remark 6.1 above, we can write the common parameter space under the P-measure of the SVJ and SVSJ models, apart from the interval B which depends on the jump type, as

$$\Omega_P = \left\{ \begin{array}{c} \theta \in \mathbb{R}^9 | \mu_0 \in \mathbb{R}, \sigma_0^2 > 0, \kappa > 0, \eta > 0, \omega > 0, \rho \in [-1,1], \\ \lambda > 0, \alpha > 0, \beta \in B \end{array} \right\}. \tag{6.3}$$

For Merton and Variance Gamma jumps $B = (-\infty, +\infty)$, see subsection 2.3.1 and subsection 2.3.2 respectively. However, for Meixner jumps $B = (-\pi, \pi)$, see subsection 2.3.3. Recall Theorem 2.31, from subsection 2.6.1, for the

SVJ model under the structure preserving Q-measure given by equation (2.73) of subsection 2.6.1. Also, recall Theorem 2.35, from subsection 2.6.2, for the SVSJ model under the structure preserving Q-measure given by equation (2.75) of subsection 2.6.2. Both of these theorems imply that if the Feller condition holds, and the jumps satisfy $\psi_X(-i) < \infty$, then, where μ_0 is the risk-free rate of return r, the discounted stock price, $e^{-rt}S_t$, $t \geq 0$ is a martingale. Thus, we write the common risk-neutral parameter space for the SVJ and SVSJ models as

$$\Omega_Q = \left\{\theta \in \Omega_P | \mu_0 = r, 2\kappa\eta > \omega^2, \psi_X(-i) < \infty\right\}. \tag{6.4}$$

The condition $\psi_X(-i) < \infty$ is given in the discussion surrounding equations (2.24), (2.33), and (2.42) for Merton, Variance Gamma, and Meixner jumps, respectively. These results are summarized in Table 6.1 below.

Jump Type	**Parameter Constraints for** $\psi_X(-i) < \infty$
Merton	$\lambda > 0, \alpha > 0, \beta \in \mathbb{R}$
Variance Gamma	$\lambda > 0, \alpha > 0, \frac{1}{2}\alpha^2 + \beta < 1$
Meixner	$\lambda > 0, \alpha > 0, \lvert\alpha + \beta\rvert < \pi$

Table 6.1: Martingale Constraints for the Jump Processes.

Remark 6.2 (Normalized Log-Price Filtration) *Since the risk-free rate r is constant, clearly S_t, $t \geq 0$ and the discounted stock price $e^{-rt}S_t$, $t \geq 0$ have the same filtration $\mathcal{F}$. Moreover, by equations* (6.1) *and* (6.2)*, the property $S_t = S_0 e^{Y_t}$ holds for both the SVJ and SVSJ models, respectively. Thus, the log-price Y_t also has filtration $\mathcal{F}$, and the relative entropies under S_t and Y_t are equivalent, see Cont and Tankov (2004b), p. 9. Hence, when S_0 is constant, the relative entropy can be obtained under the normalized log-price $Z_t = \log\left(\frac{S_t}{S_0}\right)$, $t \geq 0$, and no information will be lost.*

6.2.2 Relative Entropy and Densities

Definition 6.3 (Relative Entropy) *Let Q and P be arbitrary measures. The relative entropy of Q given P is defined by*

$$I(Q|P) = \left\{ \begin{array}{cc} E^Q\left[\log\left(\frac{dQ}{dP}\right)\right] = E^P\left[\frac{dQ}{dP}\log\left(\frac{dQ}{dP}\right)\right] & \text{if } Q \ll P, \\ \infty & \text{otherwise} \end{array} \right\}, \quad (6.5)$$

see Cont and Tankov (2004b), p. 8, and Frittelli (2000), p. 41.

Proposition 6.4 (Properties of Relative Entropy)

1. $I(Q|P) \geq 0$ with $I(Q|P) = 0$ if and only if $Q = P$.
2. The functional $Q \to I(Q|P)$ is strictly convex.

Proof. See Frittelli (2000), p. 41. ■

Definition 6.5 (Symmetric Relative Entropy) *Again let Q and P be arbitrary measures. Then the symmetric relative entropy between Q and P is defined by*

$$IS(Q,P) = I(Q|P) + I(P|Q), \quad (6.6)$$

see Kullback and Leibler (1951), p. 81.

Proposition 6.6 (Properties of Symmetric Relative Entropy)

1. $IS(Q,P) \geq 0$ with $IS(Q,P) = 0$ if and only if $Q = P$.
2. $IS(Q,P) = IS(P,Q)$.
3. The functional $Q \to IS(Q,P)$ is strictly convex.
4. The functional $P \to IS(P,Q)$ is also strictly convex.
5. If $IS(Q,P) < \infty$ then $Q \sim P$.

Proof. For *1.* and *2.*, see Kullback and Leibler (1951), p. 85. For *3.*, $Q \to I(Q|P)$ is strictly convex by Proposition 6.4, item *2.* Then by equation (6.5), if finite $I(P|Q)$ may be written as $E^P\left[\log\left(\frac{dP}{dQ}\right)\right]$. But the function $f(x) = \log\left(\frac{1}{x}\right)$ is strictly convex since $\frac{\partial^2}{\partial x^2}\log\left(\frac{1}{x}\right) = \frac{1}{x^2}$. Thus, with $x = \frac{dQ}{dP}$, $\log\left(\frac{dP}{dQ}\right)$ is strictly convex in Q, and the result follows. Lastly, *4.* and *5.* both follow by symmetry. ■

We intend to make a discrete approximation to the continuous case of symmetric relative entropy. Hence, separate definitions are expedient. Referring to equation (6.6) above, in the continuous case, we will denote the symmetric relative entropy by $HS(Q,P)$. Observe that the quantity $HS(Q,P)$ may be written as a single integral with respect to P,

$$HS(Q,P) = \left\{ \begin{array}{cc} \int\limits_{\Omega} \log\left(\frac{dQ}{dP}\right)\left(\frac{dQ}{dP} - 1\right) dP & \text{if } Q \sim P, \\ \infty & \text{otherwise} \end{array} \right\}. \tag{6.7}$$

Let λ denote the Lebesgue measure. We will assume here both that $Q \ll \lambda$, and $P \ll \lambda$. This assumption will be justified within the context of this chapter by Remark 6.8 of subsection 6.2.3 below. Under this assumption, equation (6.7) may be expressed as

$$HS(Q,P) = \left\{ \begin{array}{cc} \int\limits_{-\infty}^{+\infty} \log\left(\frac{f^Q(z)}{f^P(z)}\right)\left(f^Q(z) - f^P(z)\right) dz & \text{if } Q \sim P, \\ \infty & \text{otherwise} \end{array} \right\}, \tag{6.8}$$

where the densities $f^Q(z)$ and $f^P(z)$ are the densities associated with the absolutely continuous measures Q and P, respectively. In the discrete case we will denote the symmetric relative entropy by $DS(Q,P)$ where

$$DS(Q,P) = \left\{ \begin{array}{cc} \sum\limits_{k=1}^{N} \log\left(\frac{q_k}{p_k}\right)(q_k - p_k) & \text{if } Q \sim P, \\ \infty & \text{otherwise} \end{array} \right\}, \tag{6.9}$$

see Kapur and Kesavan (1992), p. 159, and Reesor (2001) pp. 21-22.

In this chapter we assume that the absolutely continuous measure P is fixed, and has a fixed associated density $f^P(z)$. However, the absolutely continuous measure $Q(\theta)$ is assumed to depend on the parameter θ, thus having an associated density $f^Q(z;\theta)$. For some $b > 0$, we will consider a left-endpoint uniform discretization of $z \in [-b, b]$ with grid spacing Δz. According to this discretization, in equation (6.9) above, set

$$\begin{aligned} p_k &= f^P(z_k)\,\Delta z, \text{ and} \\ q_k &= f^Q(z_k;\theta)\,\Delta z, \\ k &= 1, 2, ..., N. \end{aligned}$$

Let $\Delta f(z_k;\theta) = \left(f^Q(z_k;\theta) - f^P(z_k)\right)\Delta z$. Then equation (6.9) becomes

$$DS(Q(\theta), P) = \left\{ \begin{array}{cc} \sum_{k=1}^{N} \log\left(\frac{f^Q(z_k;\theta)}{f^P(z_k)}\right)\Delta f(z_k;\theta) & \text{if } Q \sim P, \\ \infty & \text{otherwise} \end{array} \right\}. \quad (6.10)$$

The quantity $DS(Q(\theta), P)$ given in equation (6.10) above forms a Riemann approximation to the integral in equation (6.8) for the continuous symmetric relative entropy $HS(Q(\theta), P)$, where $Q(\theta)$ depends on θ.

6.2.3 Computing Symmetric Relative Entropy by FFT

Remark 6.7 (L^1 Integrability) *Let finite $T > 0$ be the latest option price maturity date, and let t be any maturity date on $(0, T]$. Assume that the initial stock price S_0 is constant, and define*

$$Z_t = \log\left(\frac{S_t}{S_0}\right).$$

For the SVJ and SVSJ models the characteristic function

$$\phi_{Z_t}(u) = e^{C(u)+D(u)\sigma_0^2} \quad (6.11)$$

is known in closed form with the affine coefficients $C(u)$ *and* $D(u)$ *given in equations* (A.22) *and* (A.23) *respectively in Appendix A. Moreover, for the SVJ and SVSJ models, since* $t > 0$*, by Theorem 3.17 from subsection 3.3.3, the hypothesis that* $\rho < 0$ *implies*

$$\phi_{Z_t}(u) \in L^1. \tag{6.12}$$

For the SVJ and SVSJ models, we can price an option under the fixed parameters of the P-measure in Ω_P, where Ω_P is defined in equation (6.3) of subsection 6.2.1 above. Moreover by Remark 6.7 above, for any $t \in (0, T]$, $\phi_{Z_t^P}(u) \in L^1$. Thus, by the Fourier inversion theorem, Theorem 3.6, there exists a density $f^P(z_t)$ such that

$$f^P(z_t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux} \phi_{Z_t^P}(u)\, du, \text{ and } f^P(z_t) \text{ is continuous.} \tag{6.13}$$

Similarly, for $\theta \in \Omega_Q$, where the risk-neutral parameter space Ω_Q is defined in equation (6.4) of subsection 6.2.1 above, $\phi_{Z_t^Q}(u;\theta) \in L^1$, and there exists a density $f^Q(z_t;\theta)$ such that $f^Q(z_t;\theta)$ is continuous for all $\theta \in \Omega_Q$, and

$$f^Q(z_t;\theta) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux} \phi_{Z_t^Q}(u;\theta)\, du. \tag{6.14}$$

Remark 6.8 (Absolute Continuity) *Fix* $t \in (0, T]$*, and let* $F^P(z_t)$ *and* $F^Q(z_t;\theta)$ *be the distribution functions of* Z_t^P *and* Z_t^Q *respectively. Since* $\phi_{Z_t^P}(u)$ *and* $\phi_{Z_t^Q}(u;\theta)$ *are both* L^1*, equations* (6.13) *and* (6.14) *respectively imply that each of* $F^P(z_t)$ *and* $F^Q(z_t;\theta)$ *has a continuous and bounded derivative for each* $[a, b] \subset \mathbb{R}$*. Hence, by the mean value theorem, each of* $F^P(z_t)$ *and* $F^Q(z_t;\theta)$ *is Lipschitz on each* $[a, b] \subset \mathbb{R}$*. This implies that each of* $F^P(z_t)$ *and* $F^Q(z_t;\theta)$ *is absolutely continuous on each* $[a, b] \subset \mathbb{R}$*, see Royden and Fitzpatrick (2010), p. 122. Thus, the measures induced by* $F^P(z_t)$ *and* $F^Q(z_t;\theta)$ *respectively, are absolutely continuous measures for all Borel sets* B*, see Billingsley (1995), p. 413.*

For some fixed $t \in (0, T]$, based on Remark 6.8 above, we define P_t and $Q_t(\theta)$ to be the absolutely continuous measures associated with the continuous densities $f^P(z_t)$ such that $f^Q(z_t;\theta)$ defined respectively by equations (6.13) and (6.14). Thus, from equation (6.8) in subsection 6.2.2 we obtain a continuous symmetric relative entropy of the form

$$HS(Q_t(\theta), P_t) = \left\{ \begin{array}{cc} \int_{-\infty}^{+\infty} \log\left(\frac{f^Q(z_t;\theta)}{f^P(z_t)}\right) \Delta f(z_t;\theta) & \text{if } Q_t \sim P_t, \\ \infty & \text{otherwise} \end{array} \right\}, \quad (6.15)$$

where we define $\Delta f(z_t;\theta) = \left(f^Q(z_t;\theta) - f^P(z_t)\right) dz_t$. In the summary below we show how to approximate equation (6.15) above with two fast Fourier transforms. First, we give some necessary details of FFT.

Let $t \in (0, T]$ be fixed. The Fourier inversions in equations (6.13) and (6.14) above that we wish to approximate by FFT are both of the form

$$f(z_t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuz_t} \phi_{Z_t}(u)\, du. \quad (6.16)$$

Recall from subsection 3.2.1 that by Corollary 3.7, since a density is purely real, equation (6.16) above simplifies to

$$f(z_t) = \frac{1}{\pi} \operatorname{Re} \int_0^{\infty} e^{-iuz_t} \phi_{Z_t}(u)\, du. \quad (6.17)$$

Similar to subsection 3.2.3, the approximation of equation (6.17) under the DFT begins with the uniform discretization of the domain of integration,

$$u_j = (j-1)\Delta u, \;\; j = 1, 2, ..., N, \quad (6.18)$$

intending a left-endpoint Riemann sum for $f(z_{t,k})$ at each $z_{t,k} \in [-b, b]$ under the uniform discretization of the transform domain given by

$$z_{t,k} = -b + (k-1)\Delta z_t, \;\; k = 1, 2, ..., N, \text{ where } b = \frac{N}{2}\Delta z_t. \quad (6.19)$$

For the relative entropy methods of this chapter, we recommend the choices $N = 2^{12}$ with $\Delta u = \frac{1}{4}$, specified in Carr and Madan (1999), see p. 69. As in subsection 3.2.3, this leads under Nyquist optimal sampling to

$$\Delta z_t = \frac{\pi}{512}, \text{ and } b = 4\pi. \tag{6.20}$$

For some $t \in (0, T]$, and for each $z_{t,k} \in [-b, b]$, the Riemann approximation with quadrature weights w_j for the real part of the integral in equation (6.17) above is given by

$$\begin{aligned} f(z_{t,k}) &\approx \text{Re} \int_0^A e^{-iuz_{t,k}} \phi_{Z_t}(u)\, du \\ &\approx \text{Re} \sum_{j=1}^N e^{-iu_j z_{t,k}} \phi_{Z_t}(u_j)\, w_j \Delta u, \;\; k = 1, 2, ..., N, \end{aligned} \tag{6.21}$$

where $A = N\Delta u$. We use the trapezoidal rule for the quadrature weights w_j, see subsection 3.4.1. Definition 3.12 in subsection 3.2.3 gives the standard DFT. From Carr and Madan (1999), p. 68, the Riemann approximation in equation (6.21) above can be written as a non-centred and shifted standard DFT. By this analysis, we obtain for each $z_{t,k}$, $k = 1, 2, ..., N$,

$$\begin{aligned} f(z_{t,k}) &\approx \text{Re} \sum_{j=1}^N e^{-iu_j z_{t,k}} \phi_{Z_t}(u_j)\, w_j \Delta u \\ &= \text{Re} \sum_{j=1}^N e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \phi_{Z_t}(u_j)\, w_j \Delta u. \end{aligned} \tag{6.22}$$

Then, since $N = 2^{12}$ is a power of 2, the DFT in equation (6.22) above can be evaluated using FFT, see subsection 3.2.4. However, for the purposes of this chapter, we only need a subset of the N points computed by FFT. For a given $t \in (0, T]$, we will keep only the points corresponding to

$$z_{t,k} \in \left(-\pi\sqrt{t}, \pi\sqrt{t}\right), \; k = 1, 2, ..., N_t^*, \tag{6.23}$$

where N_t^* counts the number of FFT points such that $z_{t,k} \in \left(-\pi\sqrt{t}, \pi\sqrt{t}\right)$.

Summary of FFT Based Evaluation

Fix $t \in (0, T]$. Then let P_t and $Q_t(\theta)$ be absolutely continuous measures with associated densities

$$f^P(z_t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuz_t} \phi_{Z_t^P}(u)\, du, \text{ and} \tag{6.24}$$

$$f^Q(z_t;\theta) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuz_t} \phi_{Z_t^Q}(u;\theta)\, du, \tag{6.25}$$

respectively, where $Q_t(\theta)$ is a risk-neutral measure under the condition $\theta \in \Omega_Q$, see equation (6.4) in subsection 6.2.1, and where both $\phi_{Z_t^P}(u)$ and $\phi_{Z_t^Q}(u;\theta)$ are L^1. Define $\Delta f(z_t;\theta) = \left(f^Q(z_t;\theta) - f^P(z_t)\right) dz_t$. We approximate the continuous symmetric relative entropy from equation (6.15) above, given by

$$HS(Q_t(\theta), P_t) = \left\{ \begin{array}{ll} \int_{-\infty}^{+\infty} \log\left(\frac{f^Q(z_t;\theta)}{f^P(z_t)}\right) \Delta f(z_t;\theta) & \text{if } Q_t \sim P_t, \\ \infty & \text{otherwise} \end{array} \right\}, \tag{6.26}$$

on two separate levels, as follows.

First of all, by equation (6.21) above, the densities $f^P(z_t)$ and $f^Q(z_t;\theta)$, defined respectively by equations (6.24) and (6.25), have FFT approximations given by

$$f^P(z_{t,k}) \approx \text{Re} \sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \phi_{Z_t^P}(u_j)\, w_j \Delta u, \tag{6.27}$$

$$f^Q(z_{t,k};\theta) \approx \text{Re} \sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \phi_{Z_t^Q}(u_j;\theta)\, w_j \Delta u, \tag{6.28}$$

both for $k = 1, 2, ..., N$,

where in each case $z_{t,k}$ lies on a uniform discretization of the interval $[-4\pi, 4\pi]$. In accordance with equation (6.23) above, given $t \in (0, T]$, for each of the FFT approximations above, we only retain the points corresponding to the condition $z_{t,k} \in \left(-\pi\sqrt{t}, \pi\sqrt{t}\right)$ and this entails that $k = 1, 2, ..., N_t^*$.

Secondly, for convenience, for $k = 1, 2, ..., N_t^*$, define

$$\Delta g\left(z_{t,k};\theta\right) = \log\left(\frac{f^Q\left(z_{t,k};\theta\right)}{f^P\left(z_{t,k}\right)}\right)\left(f^Q\left(z_{t,k};\theta\right) - f^P\left(z_{t,k}\right)\right)\Delta z_t, \tag{6.29}$$

where Δz_t is the uniform grid spacing under each FFT approximation. Then we obtain

$$DS^*\left(Q_t\left(\theta\right), P_t\right) = \left\{ \begin{array}{cc} \sum\limits_{k=1}^{N_t^*} \Delta g\left(z_{t,k};\theta\right) & \text{if } Q_t \sim P_t, \\ \infty & \text{otherwise} \end{array} \right\}. \tag{6.30}$$

Notice that the function of θ defined by $DS^*\left(Q_t\left(\theta\right), P_t\right)$ in equations (6.29) and (6.30) above forms a Riemann approximation to the continuous symmetric relative entropy $HS\left(Q_t\left(\theta\right), P_t\right)$ in equation (6.26). This function $DS^*\left(Q_t\left(\theta\right), P_t\right)$ defined above is the function that we will use to approximate the symmetric relative entropy between two absolutely continuous measures P_t and $Q_t\left(\theta\right)$, for some $t \in (0, T]$, and for $\theta \in \Omega_Q$, for the remainder of this chapter. The next subsection treats the average symmetric relative entropy in terms of the above approximating function $DS^*\left(Q_t\left(\theta\right), P_t\right)$.

6.2.4 The Average Symmetric Relative Entropy

Recall that Ω_P and Ω_Q were respectively defined in equations (6.3) and (6.4) of subsection 6.2.1 above. The data for the Parametric Minimum Entropy Martingale Measure (PMEMM) has been filtered into the two sets of parameter estimates, $\theta_P \in \Omega_P$ and $\theta_Q^* \in \Omega_Q$. The estimates for $\theta_Q^* \in \Omega_Q$ were obtained from spot call options prices (January 19th, 2011) on Apple stock, see Table 5.3 in subsection 5.3.2, and will be used as the prior for each respective model. The role of the prior will be discussed along with the two stage optimization scheme in Section 6.3 below. The estimates for $\theta_P \in \Omega_P$ were obtained from historical data on Apple stock log-returns (1991-2011) by the Approximate Maximum Likelihood method (AML), see Table 4.2 in

subsection 4.4.2. Recall that our approximation to the symmetric relative entropy is the function of $\theta \in \Omega_Q$ given by $DS^*\left(Q_t\left(\theta\right), P_t\right)$ in equations (6.29) and (6.30) from subsection 6.2.3 above, for fixed $\theta_P \in \Omega_P$. Let M be the number of annually based maturity dates used to compute the average symmetric relative entropy, and consider the right endpoint uniform discretization of the maturity range $t \in (0, T]$ given by

$$t\left(m\right) = m\Delta t,\ 1, 2, ..., M, \tag{6.31}$$
$$\text{where } \Delta t = \frac{T}{M}.$$

The function of $\theta \in \Omega_Q$ that we effectively minimize is the average symmetric relative entropy defined by

$$AS\left(\theta\right) = \frac{1}{M}\sum_{m=1}^{M} DS^*\left(Q_{t(m)}\left(\theta\right), P_{t(m)}\right). \tag{6.32}$$

Notice that by the definition of $DS^*\left(Q_{t(m)}\left(\theta\right), P_{t(m)}\right)$ in equation (6.30) from subsection 6.2.3 above , if $AS\left(\theta\right) < \infty$, then $Q_{t(m)}\left(\theta\right) \sim P_{t(m)}$, for all $m = 1, 2..., M$. Based on this, we expect that the stronger condition, $Q\left(\theta\right) \sim P$, for $t \in (0, T]$, will hold approximately, provided that $AS\left(\theta\right) < \infty$.

6.3 Minimizing the Average Relative Entropy

Recall from equation (6.32) in subsection 6.2.4 immediately above that we wish to minimize the average symmetric relative entropy given by

$$AS\left(\theta\right) = \frac{1}{M}\sum_{m=1}^{M} DS^*\left(Q_{t(m)}\left(\theta\right), P_{t(m)}\right), \tag{6.33}$$

where by equation (6.30) of subsection 6.2.3, $DS^*\left(Q_{t(m)}\left(\theta\right), P_{t(m)}\right)$ is our proposed FFT approximation to the parametric symmetric relative entropy at maturity $t\left(m\right) = m\Delta t,\ m = 1, 2, ..., M$. But, as we mentioned at the

outset, relative entropy is not necessarily convex in the model parameters, see Cont and Tankov (2004b), p. 11. Hence, our function $AS(\theta)$, for the average symmetric relative entropy in equation (6.33) above, may or may not be convex in the model parameters $\theta \in \Omega_Q$. This is similar to the situation in Cont and Tankov (2004b), see p. 13, where it is believed that the sum of squared call options pricing errors, $\|C(\theta) - C\|^2$, is not necessarily convex. Moreover, it is expected that due to this lack of convexity in the objective function, nonlinear least squares minimization may exhibit flat regions near the solution, see Cont and Tankov (2004b), p. 13. Furthermore, there may be local minima, see Cont and Tankov (2004b), p. 14. Since lack of convexity makes the minimization of the sum of squared call options pricing errors, $\|C(\theta) - C\|^2$, an unstable problem in the manner described above, we expect that minimization of the average symmetric relative entropy, $AS(\theta)$, is also an unstable problem, given that $AS(\theta)$ is not necessarily convex.

In subsection 6.3.1 below we propose a two stage optimization routine to stabilize the minimization of $AS(\theta)$. This method is adapted from Engl, Hanke, and Neubauer (1996), see p. 250. Then in subsection 6.3.2 we show that our stabilization method is equivalent to the minimization of $AS(\theta)$ combined with a weighted squared L^2 norm penalty function. This will explain why we choose to view θ_Q^* as the prior. This equivalence is also discussed in Engl, Hanke, and Neubauer (1996), see p. 250.

6.3.1 The Two Stage Optimization Routine

Following the same procedure as in Cont and Tankov (2004b), pp. 25-26, stage one of our stabilization routine for the minimization of the average symmetric relative entropy $AS(\theta)$ consists of estimating the model error δ by the unstabilized minimum value

$$\delta = \min_{\theta \in \Omega_Q} AS(\theta). \tag{6.34}$$

If $AS(\theta)$ suffers from a flat region near the solution, then δ will be an accurate estimate. However, if $AS(\theta)$ suffers from local minima, then the estimate δ will be high. For the results in Section 6.5 below we also keep track of the unstabilized solution for the parametric minimum entropy martingale, PMEMM(0), whose parameter vector we will denote by

$$\theta_Q^{(0)} = \underset{\theta \in \Omega_Q}{\arg\min} AS(\theta). \tag{6.35}$$

Recall that we refer to $\theta_Q^* \in \Omega_Q$ as the prior, and for each model this is given by the parameters implied by call options on Apple stock from January 19th, 2011, see Table 5.3 in subsection 5.3.2. In stage two we obtain

$$\begin{aligned} \widetilde{\theta}_Q = \underset{\theta \in \Omega_Q}{\arg\min} \left\| \theta - \theta_Q^* \right\|^2 \\ \text{subject to } AS(\theta) \leq \delta, \end{aligned} \tag{6.36}$$

as the stabilized minimizer of $AS(\theta)$, see Engl, Hanke, and Neubauer (1996), Eq. (10.19), p. 250. We refer to $\widetilde{\theta}_Q$ as the parametric minimum entropy martingale (PMEMM). It represents the closest parameterized measure to the prior θ_Q^* such that the average symmetric relative entropy $AS(\theta)$ is bounded above by its unstabilized minimum value δ. The above two stage optimization routine is what we used to obtain the PMEMM results in Section 6.5 below.

6.3.2 Equivalence to Tikhonov Regularization

For some value of $\gamma > 0$, determined as a function of the model error δ, a Tikhonov regularization for the minimization of the average symmetric relative entropy $AS(\theta)$, with θ_Q^* acting as the prior, is given by

$$\widetilde{\theta}_Q = \underset{\theta \in \Omega_Q}{\arg\min} \left[AS(\theta) + \gamma \left\| \theta - \theta_Q^* \right\|^2 \right], \tag{6.37}$$

see Engl, Hanke, and Neubauer (1996), p. 243. The penalty parameter θ_Q^* is referred to as the prior in this type of regularization scheme, see Cont and Tankov (2004b), p. 17. The penalty coefficient $\gamma > 0$ in equation (6.37) may be determined from the two stage optimization routine of subsection 6.3.1 above as follows, see Engl, Hanke, and Neubauer (1996), p. 250. Assume that the model error δ, appearing as the upper bound of the constraint in equation (6.36) of subsection 6.3.1 above, has been determined in some suitable fashion. The parameter δ is often called the noise level. The Lagrangian corresponding to equation (6.36) in the two stage method of subsection 6.3.1 above is given by

$$\widetilde{\theta}_Q(\lambda) = \underset{\theta \in \Omega_Q}{\arg\min} \left[\left\| \theta - \theta_Q^* \right\|^2 + \lambda \left(AS(\theta) - \delta \right) \right], \tag{6.38}$$

where $\lambda > 0$ is the Lagrange multiplier. Let λ^* be the optimal Lagrange multiplier in equation (6.38) above, and let

$$\gamma = \frac{1}{\lambda^*}. \tag{6.39}$$

Then, after re-arranging terms, and ignoring the constant δ, equation (6.38) above becomes the Tikhonov regularization scheme in equation (6.37). The idea behind a regularization scheme such as equation (6.37) above is that for an optimal value of the penalty coefficient γ the penalized minimization problem becomes sufficiently convex, see Cont and Tankov (2004b), p. 20. We have illustrated above that we achieve the same thing by using our two stage optimization routine from subsection 6.3.1 which determines an optimal Lagrange multiplier $\lambda^* = \frac{1}{\gamma}$.

6.4 Relative Entropy in the Finance Literature

In this section we briefly summarize some of the foundational papers from the literature on relative entropy minimization in finance.

6.4.1 The Minimum Relative Entropy Distribution

The Minimum Relative Entropy Distribution (MRED), see Reesor (2001), pp. 22-23, is known from the field of Information Science, see Cover and Thomas (1991), p. 277, and has been widely adopted in the finance literature as a nonparametric method to obtain a risk-neutral measure, see for example Avellaneda (1998), pp. 451-453, and Stutzer (1996), pp. 1639-41. Below we present the MRED minimization problem in the discrete case. But, it also generalizes to the continuous case, see Avellaneda (1998), p. 451. The discrete relative entropy of Q given P is simply the discrete case of equation (6.5) from subsection 6.2.2 above, given by

$$D\left(Q|P\right) = \left\{ \begin{array}{cc} \sum\limits_{k=1}^{N} \log\left(\frac{q_k}{p_k}\right) q_k & \text{if } Q \ll P, \\ \infty & \text{otherwise} \end{array} \right\}, \tag{6.40}$$

see Cover and Thomas (1991), p. 277, and Reesor (2001), pp. 20-21. The Minimum Relative Entropy Distribution (MRED) is the solution, with $D\left(Q|P\right)$ given by equation (6.40) above, to the problem

$$\min_{Q} D\left(Q|P\right) \tag{6.41}$$

$$\text{subject to } E^{Q}\left[G_j\left(X\right)\right] = C_j,\ j = 1, 2, ..., M, \tag{6.42}$$

$$q_i \geq 0,\ i = 1, 2, ..., N, \text{ and } \sum_{i=1}^{N} q_i = 1,$$

see Reesor (2001), p. 22-23. The quantity X in equation (6.42) above is a state-variable for the economy, and the C_j, $j = 1, 2, ..., M$, are discounted cash-flows of traded assets, see Avellaneda (1998), p. 451. The typical assumption is that the functions $G_j\left(X\right)$ in equation (6.42) above are the discounted payoffs of European call options, $e^{-rT}\left[S_T - K_j\right]^{+}$, and the C_j are the corresponding call options prices, see Cont and Tankov (2004b), p. 21. The problem (6.41) above is convex in Q, and has a closed form

solution for the MRED in terms of the Lagrangian, see Cover and Thomas (1991), p. 277, Reesor (2001), p. 23, and Avellaneda (1998), p. 451. The main drawback of the Minimum Relative Entropy Distribution is that for a given P-measure, P, the equality constraints in (6.42) above may not be met within reasonable bounds. In such cases, the solution does not exist, see Cont and Tankov (2004b), p. 21.

6.4.2 The Method of Cont and Tankov

Cont and Tankov (2004b), see p. 18, considers the least squares options price calibration problem penalized by relative entropy. A similar problem was considered in Avellaneda et al (2001), see p. 99. In Cont and Tankov (2004b), the log-price Y_t is modeled as a Lévy process, and further specialized into a jump-diffusion with finite-activity jumps. This allows the log-price to be described solely by the volatility σ, and a nonparametric Lévy measure ν, on a finite grid. After a suitable discretization of the problem, see Cont and Tankov (2004a) pp. 443-44, pricing is handled by a version of the Carr and Madan (1999) formula, see Cont and Tankov (2004a) p. 445, and the relative entropy of Q given P may be expressed as

$$H(\nu) = \frac{T}{2\sigma^2}\left[\sum_{j=1}^{N}\left(e^{y_j}-1\right)\left(\nu_j-\nu_j^P\right)\right]^2 + T\sum_{j=1}^{N}\left[\nu_j\log\left(\frac{\nu_j}{\nu_j^P}\right)+\nu_j^P-\nu_j\right], \tag{6.43}$$

for maturity T, with $\nu_j^Q = \nu_j$, see Cont and Tankov (2004b), p. 41. Notice that σ is not a parameter on the left in $H(\nu)$ as defined in equation (6.43). This is because two Lévy processes are equivalent if and only if they share the same volatility σ, and have equivalent Lévy measures, see Cont and Tankov (2004b), p. 6. Hence, $\sigma = \sigma^P$ is fixed in the model, see Cont and

Tankov (2004b), p. 10. Thus, the problem becomes

$$\min_{\nu} \left[\|C(\nu) - C\|^2 + \gamma H(\nu) \right], \tag{6.44}$$

for some penalty coefficient $\gamma > 0$, see Cont and Tankov (2004b), p. 18. Notice that in the solution to problem (6.44) above, $H(\nu) < \infty$ implies that $Q \ll P$. It is further proven in Cont and Tankov (2006), see p. 10, that if the log-price C.F. under P is L^1, then $Q \sim P$.

The penalty coefficient $\gamma > 0$ is estimated as follows, see Cont and Tankov (2004b), pp. 25-26. First estimate the model error by

$$\delta^2 = \min_{\nu} \|C(\nu) - C\|^2 . \tag{6.45}$$

Then, define the function of $\gamma > 0$ given by

$$\nu(\gamma) = \arg\min_{\nu} \left[\|C(\nu) - C\|^2 + \gamma H(\nu) \right]. \tag{6.46}$$

By the Morozov discrepancy principle, see Engl, Hanke, and Neubauer (1996), p. 84, for some $a > 1$,

$$\gamma = \sup_{\gamma > 0} \{ \|C(\nu(\gamma)) - C\| \leq a\delta \}. \tag{6.47}$$

Cont and Tankov (2004b), see p. 16, seeks to stabilize the least squares calibration problem. In Section 6.3 above, we stabilized the minimization of the average symmetric relative entropy, as defined in equation (6.32) from subsection 6.2.4, in effect with a squared L^2 norm penalty. While we estimate the model error similar to equation (6.45) above, we are able to obtain $\gamma > 0$ by a simple Lagrange multiplier technique.

Placing the Method of Cont and Tankov in the Literature

Citing from Cont and Tankov (2004b), pp. 20-21, there is a method for incomplete markets from Kallsen (2002) that produces the minimal entropy

consistent martingale measure (MECMM). The MECMM minimizes the relative entropy of Q given P, but is also consistent with a basket of market call options prices on the evaluation day. While Kallsen (2002) provides theory in support of the MECMM, and related measures, the paper offers no method of numerical implementation. Strictly speaking, the Method of Cont and Tankov is not equivalent to the Kallsen (2002) MECMM, see Cont and Tankov (2004b), p. 21. This is because the Method of Cont and Tankov restricts the log-price to the class of Lévy processes, or more specifically jump-diffusions. However, given this restriction, the Method of Cont and Tankov provides a computable approximation to the MECMM, see Cont and Tankov (2004b), p. 21.

6.4.3 The Minimal Entropy Martingale Measure

Let P be the objective measure for the discounted stock price $e^{-rt}S_t$, let $Q_M = \left\{Q \ll P \mid \textit{such that } Q \textit{ is a martingale measure for } e^{-rt}S_t\right\}$, and also let $Q_{ME} = \{Q \in Q_M \mid Q \sim P\}$. Recall from subsection 6.2.2 above that in equation (6.5) we defined relative entropy to be the function of Q given by

$$I\left(Q|P\right) = \left\{ \begin{array}{cc} E^Q\left[\log\left(\frac{dQ}{dP}\right)\right] = E^P\left[\frac{dQ}{dP}\log\left(\frac{dQ}{dP}\right)\right] & \text{if } Q \ll P, \\ \infty & \text{otherwise} \end{array} \right\}. \quad (6.48)$$

As given in Frittelli (2000), p. 41, the Minimal Entropy Martingale Measure (MEMM) is defined by the solution Q_0 to the problem

$$\inf_{Q \in Q_M} I\left(Q|P\right), \quad (6.49)$$

see Frittelli (2000), p. 41. Note that the MEMM is not necessarily structure preserving. The model under the MEMM may differ in structure from the model under P. This happens in the Barndorff-Nielsen and Shephard stochastic volatility model, see Rheinländer and Steiger (2006), p. 1320.

The sufficient conditions for the existence of the MEMM given in Frittelli (2000), see p. 42, are as follows. Each discounted stock price process is bounded, and there exists $Q \in Q_M$ such that $I(Q|P) < \infty$. It is also shown that if there exists $Q \in Q_{ME}$ such that $I(Q|P) < \infty$, then the MEMM, Q_0, is equivalent to P, and this is true even if the processes are unbounded, see Frittelli (2000), p. 43. Note that uniqueness of the solution Q_0 for the MEMM follows directly from the convexity of relative entropy in problem (6.49) above. Moreover, an options price computed under the MEMM may be viewed as the limit of the exponential utility indifference price, as the rate of risk aversion approaches zero, see Cont and Tankov (2004a), p. 344. In the case of an exponential Lévy process for the discounted stock price under P, the conditions for the existence of the MEMM simplify considerably, and the existence merely of a solution to a system containing an equality and an upper bound determines the existence of the MEMM, see Fujiwara and Miyahara (2003), p. 515. Moreover, the MEMM for an exponential Lévy process is another exponential Lévy process, and when this MEMM exists it is the Esscher transform of a modified log-price with a specified transform parameter, see Fujiwara and Miyahara (2003), pp. 515-16.

Pricing Call Options under the Heston MEMM and PMEMM

To the best of our knowledge, the existing literature on the MEMM has not treated either SVJ model or the SVSJ model. However, it was shown in Hobson (2004), see pp. 554-55, that the MEMM for the Heston model exists. In terms of the stock price S_t, the change of measure corresponding to the Heston MEMM is given by

$$M_t^S = \frac{dQ_t^S}{dP_t^S},\ t \in [0,T], \tag{6.50}$$

see Hobson (2004), p. 554. The expression for the Heston MEMM change of measure, M_t^S in equation (6.50) above, even if $t \in [0,T]$ is fixed, is quite

complicated, see Hobson (2004), p. 554. For example, to the best of our knowledge, its Fourier transform is not known in closed form. Hence, numerical evaluation of M_t^S, in equation (6.50) above, is beyond the scope of this thesis. However, we will show how to price a call option under the Heston MEMM, assuming that M_t^S in equation (6.50) above can be evaluated. First, fix some $t \in (0, T]$. Then the Heston MEMM European call price is given by

$$\begin{aligned} C_t^{MEMM}(K) &= e^{-rt} E^Q \left[(S_t - K)^+ \right] \\ &= e^{-rt} \int_{\Omega_S} M_t^S (s_t - K)^+ \, dP_t^S. \end{aligned} \tag{6.51}$$

A more direct method of pricing call options under the Heston MEMM is provided in He and Zhu (2016). With the Heston MEMM process described by a suitable SDE, He and Zhu (2016), see p. 237, propose a boundary value problem (BVP) with the Heston MEMM call price as its solution. Since, the numerical solution of this BVP appears tractable, we propose it as our method of computing Heston MEMM call prices for future research.

By contrast, the pricing of a European call option, $C_t^{PMEMM}(K)$, under the Heston PMEMM is simple. Recall from subsection 6.2.1 above that the Heston PMEMM is structure preserving, and that with $\mu_0 = r$, only the other model parameters change. Thus, the exponentially affine risk-neutral log-price characteristic function is known in closed form, see equation (6.11) from subsection 6.2.3, with PMEMM parameters $\widetilde{\theta}_Q \in \Omega_Q$. Hence, we can price a call option with maturity $t \in (0, T]$ using the log-moneyness version of the Carr and Madan formula from subsection 5.2.2.

Comparing the PMEMM to the MEMM

Similar to Cont and Tankov (2004b), pp. 20-21, in comparing their method to the MECMM of Kallsen (2002), we say that the PMEMM is not equivalent

to the MEMM. For one, the PMEMM minimizes symmetric relative entropy, whereas the MEMM minimizes relative entropy. But more importantly, the PMEMM is structure preserving with changing parameters, whereas the MEMM allows model structure to change while preserving the parameters. To the extent that the changing parameters of the PMEMM may actually correspond to the changing structural form of the MEMM, in terms of the graph of the risk-neutral density, or the values of the computed call prices, there is room to consider the PMEMM as a computable approximation to an alternative version of the MEMM that minimizes symmetric relative entropy in place of minimizing relative entropy. However, in the absence of a means for direct comparison, we may only conclude that the PMEMM is an easily computable alternative solution to a particular minimum relative entropy problem that may or may not differ from the MEMM.

6.5 PMEMM Results for Apple Stock

All results in this section were computed on an Intel Xeon 2xE5-2643v3 dual processor at 3.4 GHz on 12 cores. The two stage routine for the minimization of the average symmetric relative entropy $AS(\theta)$, as described in subsection 6.3.1 above was implemented for all five models, using the function *fmincon* from Matlab R2016a. Based on the January 19th, 2011 Apple stock call options data, we chose $T = 2$ years, with $M = 24$ months in equation (6.32) for $AS(\theta)$ from subsection 6.2.4.

6.5.1 Main Line Results and Forecasts

The SVSJ Merton model, the SVJ Merton model, and the SVJ Variance Gamma model, each for Apple stock (1991-2011), form the main line cases for this section. Results for these three cases are given below.

PMEMM: Main Line Results									
$\boldsymbol{\Theta}$	μ_0	σ_0^2	κ	η	ω	ρ	λ	α	β
SVSJ Merton									
θ_P	0.360	0.167	3.849	0.167	0.643	-0.278	91.38	0.062	0.009
θ_Q^*	r	0.051	1.805	0.103	0.441	-0.706	64.25	0.082	-0.003
$\widetilde{\theta}_Q$	r	0.138	1.827	0.155	0.264	-0.606	64.25	0.104	0.007
$\theta_Q^{(0)}$	r	0.569	812.0	0.090	4.658	-0.154	136.4	0.104	-0.046
Total CPU Time				67.1 sec	**Risk-Free Rate** r				0.293%
Model Error δ				0.572	**Lagrange Multiplier** λ^*				1.332
SVJ Merton									
θ_P	0.361	0.175	3.995	0.175	0.726	-0.268	10.22	0.070	0.007
θ_Q^*	r	0.058	1.767	0.136	0.567	-0.662	1.637	0.095	-0.008
$\widetilde{\theta}_Q$	r	0.198	1.809	0.208	0.346	-0.309	1.647	0.149	-0.022
$\theta_Q^{(0)}$	r	11.02	8799	0.160	37.25	0.450	2.789	0.145	-0.091
Total CPU Time				74.7 sec	**Risk-Free Rate** r				0.293%
Model Error δ				0.578	**Lagrange Multiplier** λ^*				6.449
SVJ Variance Gamma									
θ_P	0.367	0.161	3.421	0.161	0.691	-0.293	13.537	0.068	0.008
θ_Q^*	r	0.057	1.736	0.135	0.559	-0.668	1.337	0.111	-0.014
$\widetilde{\theta}_Q$	r	0.196	1.777	0.214	0.343	-0.349	1.349	0.165	-0.018
$\theta_Q^{(0)}$	r	1.436	1333	0.162	6.145	-0.056	2.520	0.156	-0.093
Total CPU Time				107 sec	**Risk-Free Rate** r				0.293%
Model Error δ				0.599	**Lagrange Multiplier** λ^*				5.927

θ_P	The Historical P-Measure Estimated by AML
θ_Q^*	The Prior Q-Measure Estimated by Nonlinear Least Squares
$\widetilde{\theta}_Q$	The PMEMM Estimated by the Two Stage Method
$\theta_Q^{(0)}$	The Unstabilized PMEMM from Stage One to Obtain Model Error δ

Table 6.2: PMEMM Parameter Estimates: Apple Stock: Main Cases.

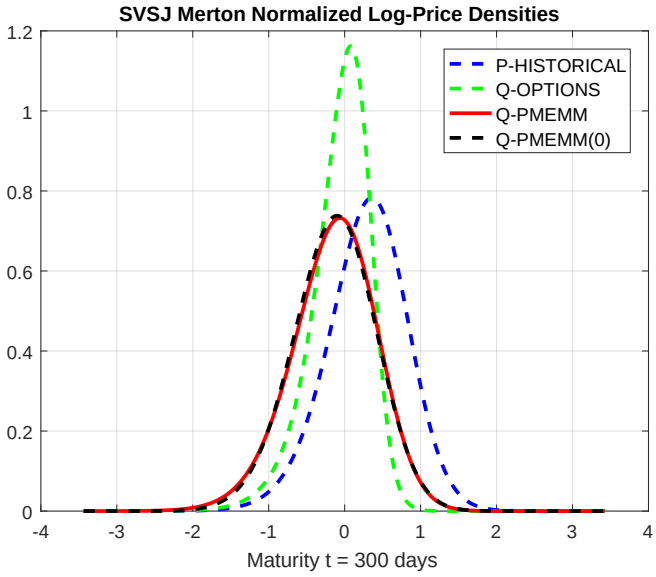

Figure 6.1: PMEMM: Normalized Log-Price Densities: SVSJ Model.

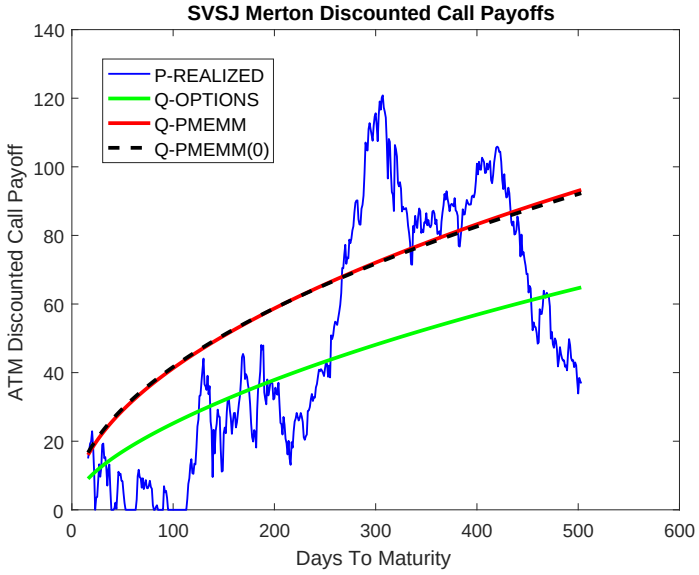

Figure 6.2: SVSJ PMEMM: ATM Discounted Call Payoff Forecast.

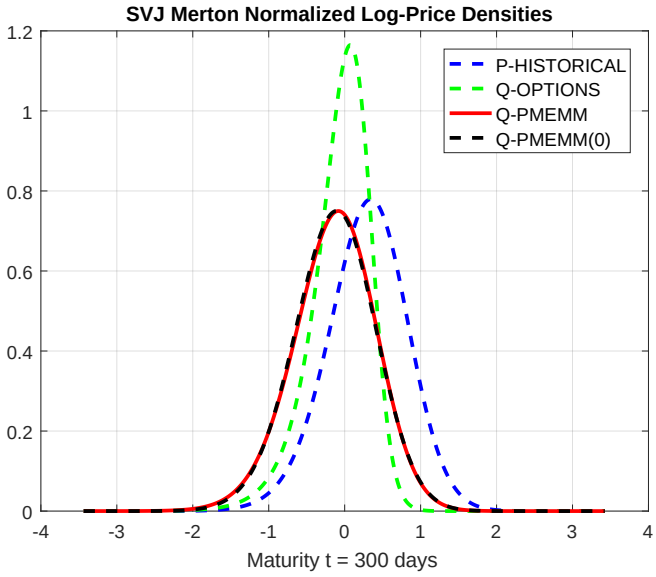

Figure 6.3: PMEMM: Normalized Log-Price Densities: SVMJ Model.

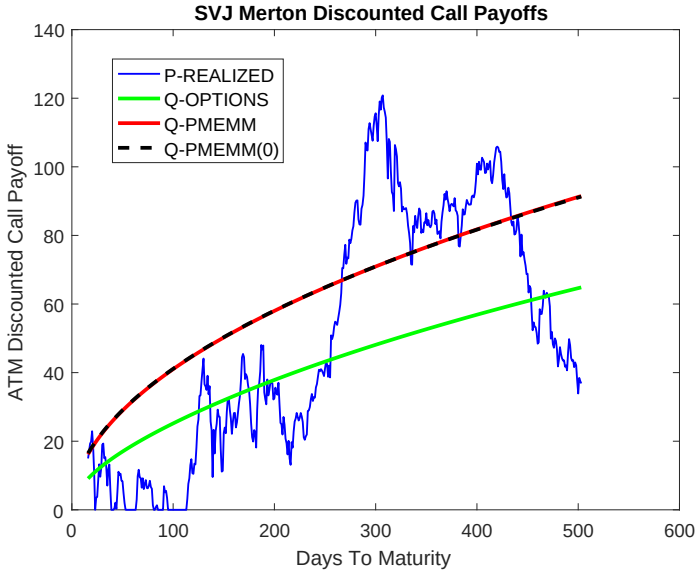

Figure 6.4: SVMJ PMEMM: ATM Discounted Call Payoff Forecast.

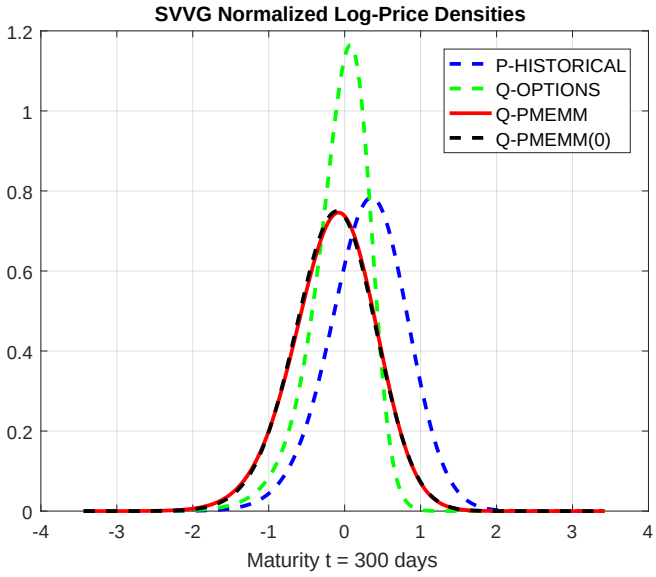

Figure 6.5: PMEMM: Normalized Log-Price Densities: SVVG Model.

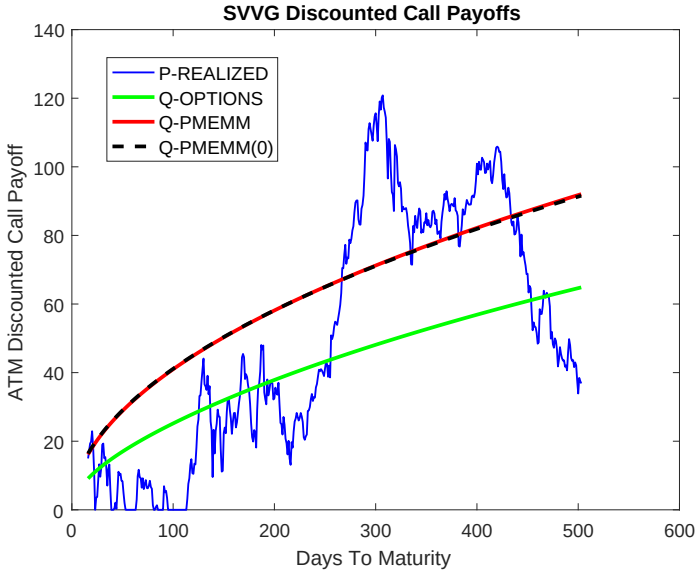

Figure 6.6: SVVG PMEMM: ATM Discounted Call Payoff Forecast.

Analysis of Main Line Results

We will focus here on the results for the SVSJ Merton model from Table 6.2, Figure 6.1, and Figure 6.2. It is clear from Figure 6.3 and Figure 6.4 for the SVJ Merton model, as well as Figure 6.5 and Figure 6.6 for the SVVG model, that these cases are similar. Table 6.2 shows that the unstabilized estimate, $\theta_Q^{(0)}$, of the PMEMM for the SVSJ Merton model has a very high value of κ, with an unrealistically high value of ω, suggesting a sharp struggle to contain the volatility of the volatility. This is very different from the stabilized estimate, $\widetilde{\theta}_Q$, of the PMEMM for the SVSJ Merton model in Table 6.2 which has been influenced by the prior θ_Q^*. The stabilization process has lead to an estimate of the PMEMM, $\widetilde{\theta}_Q$, for the SVSJ model in Table 6.2 with a modest value of κ, and a realistic value of ω. But, despite the difference between these two estimates, $\widetilde{\theta}_Q$ and $\theta_Q^{(0)}$, of the PMEMM for the SVSJ Merton model, Figure 6.1 indicates that the graphs of the two corresponding normalized log-price densities are essentially the same. The maturity of 300 days was chosen in Figure 6.1 since it lies off the monthly maturity grid, and is also relatively late within the two year time horizon.

Main Line Forecasts

In Figure 6.2, Figure 6.4, and Figure 6.6, respectively, we have plotted a daily sequence of discounted call option payoffs on Apple stock at-the-money for the SVSJ Merton, SVJ Merton, and SVVG models. The evaluation day is January 19th, 2011, and the maturities run from 15 days to 504 days past this date. The true values under the P-measure are based on realized historical prices of Apple stock. For the three Q-measure based forecasts considered in each respective figure, the Carr and Madan (1999) formula was used, under the parameters in Table 6.2, to obtain the forecasts of the at-the-money discounted call option payoffs. Focusing again on the SVSJ

Merton model in Figure 6.2 above, observe that for the stabilized and unstabilized PMEMM estimates, $\widetilde{\theta}_Q$ and $\theta_Q^{(0)}$, from Table 6.2, the discounted call option payoff forecasts are essentially the same, suggesting that only the graph of the log-price density matters, and not the parameters. Also, this shows that in the main line case, the prior θ_Q^* does not influence the PMEMM options price forecasts. The prior only influences our choice of the optimal parametric solution $\widetilde{\theta}_Q$. But, we prefer the stabilized estimate of the PMEMM, $\widetilde{\theta}_Q$, since it is consistent with our prior beliefs about parametric structure. For the practical benefits of the PMEMM forecast, as compared to the option implied Q-measure forecast, we refer to Figure 6.2 where the case is illustrated for the SVSJ Merton model. Observe that in the last 250 days, the forecast of the at-the-money call options priced under the PMEMM is much closer to the graph of the realized discounted at-the-money call payoffs based on the historical prices of Apple stock, than the forecast of the at-the-money call options priced under the option implied Q-measure is. As pointed out in Chernov and Ghysels (2000), see p. 447, the option implied Q-measure is only sufficient for shorter maturities. Longer maturities require the excess kurtosis of the return based P-measure in order to be well priced. Observe in Figure 6.1 for the SVSJ Merton model that the density for the risk-neutral PMEMM has the same shape as the graph of the density for the historical P-measure, suggesting that $Q \sim P$ holds at the stated level of the maturity. Moreover, observe in particular that the PMEMM density has the same wide hips as the historical return based density. This indicates excess kurtosis, and it is why the PMEMM outperforms for later maturities. However, the option implied Q-measure is clearly better for shorter maturities. Lastly, since each model produces a similar PMEMM forecast, we believe it is the PMEMM itself, and not the choice of model, that leads to long term forecasting success.

6.5.2 Additional Results and Forecasts

The additional cases are the SVJ Meixner model, and the Heston model. Their behaviour is essentially the same as the main line cases. However, the SVJ Meixner model appears to suffer from a high estimate of the model error δ, and the P-measure for the Heston model is quite different from the prior Q-measure. The analysis will investigate these differences, and provide a modified approach in the Heston case.

$\boldsymbol{\Theta}$	μ_0	σ_0^2	κ	η	ω	ρ	λ	α	β
SVJ Meixner									
θ_P	0.368	0.152	3.256	0.152	0.715	-0.299	5.227	0.166	0.277
θ_Q^*	r	0.056	1.720	0.134	0.554	-0.670	0.396	0.293	-0.380
$\widetilde{\theta}_Q$	r	0.190	1.741	0.231	0.389	-0.500	0.448	0.414	-0.362
$\theta_Q^{(0)}$	r	1.097	1588	0.222	21.62	-0.889	225.2	0.011	-0.601
Total CPU Time				56.4 sec	**Risk-Free Rate** r				0.293%
Model Error δ				0.601	**Lagrange Multiplier** λ^*				0.5392
Heston									
θ_P	0.254	0.226	21.98	0.226	2.428	-0.125	-		
θ_Q^*	r	0.071	1.973	0.148	0.604	-0.556	-		
$\widetilde{\theta}_Q$	r	0.229	1.997	0.208	0.436	-0.259	-		
$\theta_Q^{(0)}$	r	0.300	28.05	0.210	2.873	-0.161	-		
Total CPU Time				50.4 sec	**Risk-Free Rate** r				0.293%
Model Error δ				0.286	**Lagrange Multiplier** λ^*				0.9562

θ_P	The Historical P-Measure Estimated by AML
θ_Q^*	The Prior Q-Measure Estimated by Nonlinear Least Squares
$\widetilde{\theta}_Q$	The PMEMM Estimated by the Two Stage Method
$\theta_Q^{(0)}$	The Unstabilized PMEMM from Stage One to Obtain Model Error δ

Table 6.3: PMEMM Parameter Estimates: Apple Stock: Extra Cases.

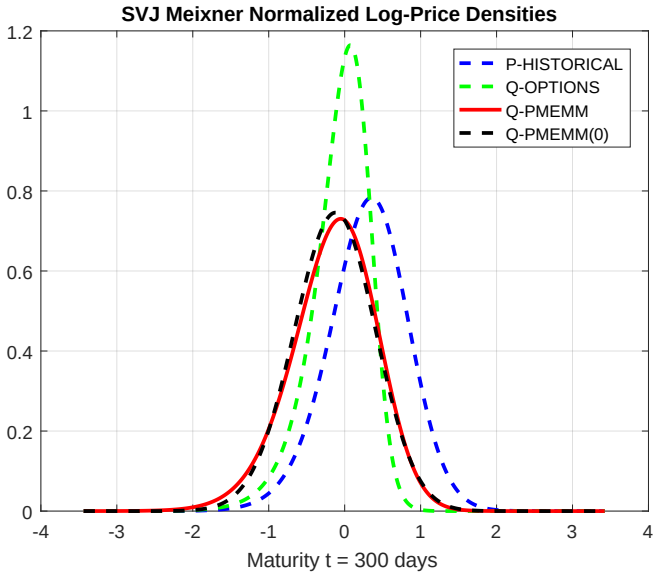

Figure 6.7: PMEMM: Normalized Log-Price Densities: SVMX Model.

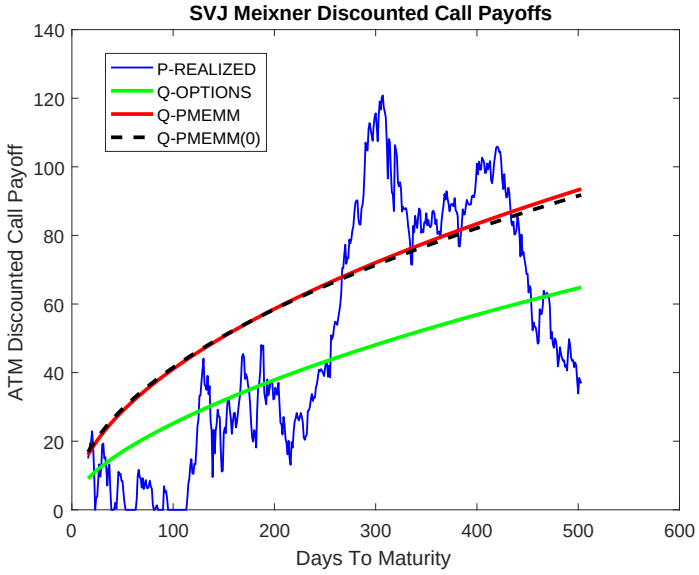

Figure 6.8: SVMX PMEMM: ATM Discounted Call Payoff Forecast.

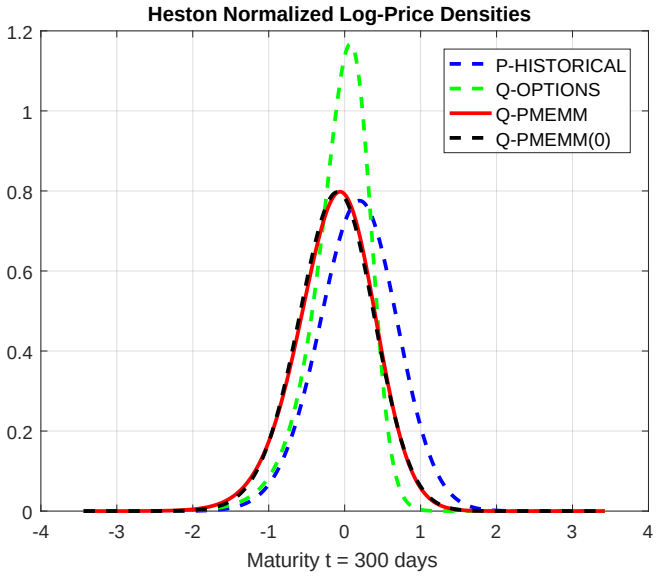

Figure 6.9: PMEMM: Normalized Log-Price Densities: Heston Model.

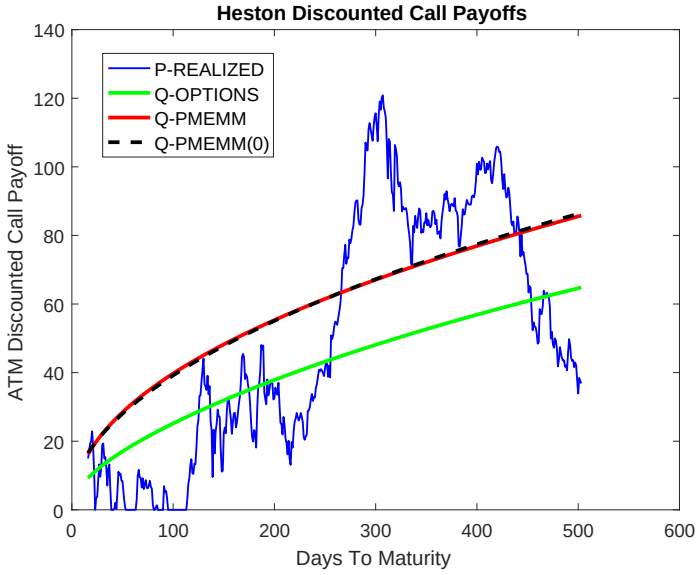

Figure 6.10: Heston PMEMM: ATM Discounted Call Payoff Forecast.

Analysis of Additional Results and Forecasts

As illustrated in Figure 6.7 and Figure 6.8 above for the SVJ Meixner model, as well as Figure 6.9 and Figure 6.10 above for the Heston model, the general analysis provided in subsection 6.5.1 above will suffice for the additional results and forecasts. But, notice in Figure 6.7 for the SVJ Meixner model that the graphs of the PMEMM density and the unstabilized PMEMM(0) density do not perfectly coincide. Accordingly, in Figure 6.8 the corresponding forecasts of the discounted call options payoffs fail to coincide exactly. But, also notice that the model error δ for the SVJ Meixner PMEMM is the highest, and the optimal Lagrange multiplier λ^* is the lowest, for any of the models considered in this chapter, see Table 6.2 and Table 6.3 above. We believe that this δ is a local minimum of problem (6.34), stage one of subsection 6.3.1. Thus, it is too high. This causes slack in the constraint of problem (6.36), stage two of subsection 6.3.1. Thus, the optimal Lagrange multiplier λ^* is too low, implying that the weight $\gamma = \frac{1}{\lambda^*}$ on the prior is too high. This unbalances the PMEMM slightly. Looking at Figure 6.7 for the SVJ Meixner model, we see that as the PMEMM approaches the prior, it also moves closer to P. Thus, in Figure 6.8 for the SVJ Meixner forecast, the PMEMM discounted payoff forecast lies above the unstabilized PMEMM(0) discounted payoff forecast.

For the Heston case, observe in Table 6.3 that the parameter vector, θ_P, for the P-measure is quite far away from the prior, θ_Q^*. The values of κ and ω under the P-measure are too large. This occurs because the Heston model has no jumps, and it is a poor fit to the Apple stock log-return data which has many jumps. In order to obtain a solution for the PMEMM under this particular Heston model, we need to induce extra weight on the prior. We modify the two stage optimization from subsection 6.3.1 as follows. Stage

one is the same with

$$\delta = \min_{\theta \in \Omega_Q} AS\left(\theta\right), \tag{6.52}$$

as in equation (6.34). But, in stage two we scale up δ in the constraint of equation (6.36) to $a\delta$, for some factor $a > 1$. This is similar to the Morozov discrepancy principle in equation (6.47) of subsection 6.4.2 above, covering the method of Cont and Tankov. For the Heston case in Table 6.3 above, we used $a = 1.01$. Stage two of the two stage optimization becomes

$$\widetilde{\theta}_Q = \arg\min_{\theta \in \Omega_Q} \left\|\theta - \theta_Q^*\right\|^2 \tag{6.53}$$
$$\text{subject to } AS\left(\theta\right) \leq a\delta, \text{ for some } a > 1.$$

Equation (6.53) above leads to a consistent solution for the Heston PMEMM in Figure 6.9 above. But, notice in Figure 6.10 above that the distance between the Heston PMEMM discounted call payoff forecast and the forecast based on the prior is slightly less than the same for the other models.

6.5.3 On the Choice of M in the PMEMM

In this subsection we table the PMEMM results for Apple stock, again with time horizon $T = 2$, but with various uniform maturity date sets corresponding to $M = 3$, 6, 12, 24, 48, and 96. We refer to this simply as the PMEMM(M). We give results for all five models, finding that as M reaches 96, only two of the models remain self-consistent in both the parameters and the tabled call prices, while one of the models remains self-consistent beyond $M = 24$ only in the tabled call prices, and the other two models diverge only when $M = 96$. Because only some of the models appear to have a robust PMEMM(M) for large M, for practical purposes we recommend monthly maturities, with $M = 24$ when $T = 2$, to cover all five models.

SVJ Merton PMEMM(M)

	σ_0^2	κ	η	ω	ρ	λ	α	β
$M = 3$	0.173	1.791	0.184	0.409	-0.581	1.650	0.164	0.083
$M = 6$	0.179	1.790	0.194	0.413	-0.574	1.649	0.165	0.065
$M = 12$	0.185	1.792	0.200	0.395	-0.555	1.648	0.160	0.053
$M = 24$	0.198	1.809	0.208	0.346	-0.309	1.647	0.149	-0.022
$M = 48$	0.200	1.809	0.212	0.350	-0.325	1.646	0.144	-0.013
$M = 96$	0.200	1.808	0.213	0.346	-0.348	1.646	0.144	-0.006

Table 6.4: PMEMM(M) Parameters: Various M: SVMJ Model.

	$M = 3$	$M = 6$	$M = 12$	$M = 24$	$M = 48$	$M = 96$
Time	16.7	30.7	42.5	74.7	367	242

Table 6.5: Time in Seconds to Compute PMEMM(M): SVMJ Model.

	3 mo.	**6 mo.**	**12 mo.**	**18 mo.**	**24 mo.**	S_0	338.84
$M = 3$	32.047	45.493	64.341	78.677	90.643	r	0.2934%
$M = 6$	32.174	45.668	64.617	79.042	91.083	-	
$M = 12$	32.311	45.840	64.847	79.314	91.386	-	
$M = 24$	32.508	46.045	65.061	79.517	91.567	-	
$M = 48$	32.507	46.043	65.079	79.558	91.630	-	
$M = 96$	32.506	46.051	65.104	79.598	91.682	-	

Table 6.6: ATM Call Prices at PMEMM(M): SVMJ Model.

As M reaches 96, the SVJ Merton PMEMM(M) is the most self-consistent, in both parameters and call prices, of any of the four models with jumps. Reasons for choosing $M = 24$ in this case are that the PMEMM(M) and the call prices have sufficiently converged, and $M = 24$ is faster.

Heston PMEMM(M)

	σ_0^2	κ	η	ω	ρ
$M = 3$	0.225	1.995	0.211	0.436	-0.288
$M = 6$	0.226	1.993	0.213	0.452	-0.309
$M = 12$	0.227	1.994	0.212	0.450	-0.296
$M = 24$	0.229	1.997	0.208	0.436	-0.259
$M = 48$	0.225	1.993	0.215	0.459	-0.315
$M = 96$	0.229	1.997	0.209	0.437	-0.260

Table 6.7: PMEMM(M) Parameters: Various M: Heston Model.

	$M = 3$	$M = 6$	$M = 12$	$M = 24$	$M = 48$	$M = 96$
Time	8.1	11.4	25.6	50.4	44.3	67.2

Table 6.8: Time in Seconds to Compute PMEMM(M): Heston Model.

	3 mo.	**6 mo.**	**12 mo.**	**18 mo.**	**24 mo.**	S_0	338.84
$M = 3$	31.557	44.130	61.563	74.792	85.863	r	0.2934%
$M = 6$	31.645	44.247	61.723	74.990	86.096	-	
$M = 12$	31.709	44.310	61.761	75.000	86.082	-	
$M = 24$	31.802	44.378	61.733	74.875	85.871	-	
$M = 48$	31.636	44.269	61.822	75.161	86.328	-	
$M = 96$	31.785	44.369	61.750	74.915	85.931	-	

Table 6.9: ATM Call Prices at PMEMM(M): Heston Model.

The PMEMM(M) for the Heston model is fully self-consistent as M reaches 96, as are the at-the-money call prices. But, this is partly because of the extra coercion that was necessary in this case, in order to conform to the prior. See equation (6.53) in subsection 6.5.2 above.

SVJ Meixner PMEMM(M)

	σ_0^2	κ	η	ω	ρ	λ	α	β
$M = 3$	0.155	1.755	0.148	0.214	-0.066	0.793	0.426	-1.099
$M = 6$	0.187	1.745	0.217	0.350	-0.472	0.459	0.439	-0.356
$M = 12$	0.189	1.743	0.223	0.365	-0.478	0.453	0.425	-0.359
$M = 24$	0.190	1.741	0.231	0.389	-0.500	0.448	0.414	-0.362
$M = 48$	0.192	4.436	0.208	0.749	-0.263	1.292	0.245	-0.400
$M = 96$	0.181	4.948	0.198	0.826	-0.282	2.187	0.212	-0.232

Table 6.10: PMEMM(M) Parameters: Various M: SVMX Model.

	$M = 3$	$M = 6$	$M = 12$	$M = 24$	$M = 48$	$M = 96$
Time	105	19.5	36.3	56.4	511	1071

Table 6.11: Time in Seconds to Compute PMEMM(M): SVMX Model.

	3 mo.	**6 mo.**	**12 mo.**	**18 mo.**	**24 mo.**	S_0	338.84
$M = 3$	32.503	46.221	65.331	79.753	91.735	r	0.2934%
$M = 6$	32.430	46.198	65.729	80.641	93.072	-	
$M = 12$	32.431	46.221	65.825	80.810	93.306	-	
$M = 24$	32.398	46.221	65.933	81.027	93.619	-	
$M = 48$	32.439	46.090	65.423	80.108	92.315	-	
$M = 96$	32.405	46.080	65.436	80.127	92.334	-	

Table 6.12: ATM Call Prices at PMEMM(M): SVMX Model.

Beyond $M = 24$, the SVJ Meixner PMEMM(M) is self-consistent in the call prices only. The PMEMM(M) parameters themselves diverge from the pattern at both $M = 48$ and $M = 96$. However, the choice $M = 24$ for the SVJ Meixner PMEMM(M) appears to be reasonable.

SVVG PMEMM(M)

	σ_0^2	κ	η	ω	ρ	λ	α	β
$M = 3$	0.161	1.754	0.197	0.440	-0.593	1.353	0.206	0.037
$M = 6$	0.188	1.769	0.204	0.312	-0.505	1.352	0.183	0.011
$M = 12$	0.186	1.761	0.209	0.370	-0.543	1.350	0.181	0.026
$M = 24$	0.196	1.777	0.214	0.343	-0.349	1.349	0.165	-0.019
$M = 48$	0.196	1.775	0.215	0.335	-0.386	1.349	0.165	-0.008
$M = 96$	0.242	34.993	0.241	1.315	-0.104	0.033	1.967	-5.315

Table 6.13: PMEMM(M) Parameters: Various M: SVVG Model.

	$M = 3$	$M = 6$	$M = 12$	$M = 24$	$M = 48$	$M = 96$
Time	14.3	21.9	58.2	107	157	231

Table 6.14: Time in Seconds to Compute PMEMM(M): SVVG Model.

	3 mo.	**6 mo.**	**12 mo.**	**18 mo.**	**24 mo.**	S_0	338.84
$M = 3$	31.534	45.126	64.471	79.285	91.657	r	0.2934%
$M = 6$	32.424	46.053	65.238	79.829	91.987	-	
$M = 12$	32.327	45.961	65.216	79.895	92.140	-	
$M = 24$	32.441	46.054	65.274	79.917	92.125	-	
$M = 48$	32.441	46.068	65.317	79.984	92.211	-	
$M = 96$	35.287	50.694	72.722	89.602	103.70	-	

Table 6.15: ATM Call Prices at PMEMM(M): SVVG Model.

For $M = 96$, the PMEMM(M) for the SVVG model diverges from its well established pattern, and in this case the at-the-money call prices are also affected at $M = 96$. However, the choices $M = 24$ and $M = 48$ both appear to be reasonable for the SVVG model PMEMM(M).

SVSJ Merton PMEMM(M)

	σ_0^2	κ	η	ω	ρ	λ	α	β
$M = 3$	0.117	1.817	0.136	0.345	-0.676	64.252	0.105	0.055
$M = 6$	0.109	1.811	0.138	0.395	-0.690	64.252	0.116	0.041
$M = 12$	0.130	1.830	0.143	0.233	-0.597	64.252	0.114	-0.003
$M = 24$	0.138	1.827	0.155	0.264	-0.606	64.252	0.104	0.007
$M = 48$	0.143	1.827	0.162	0.272	-0.600	64.252	0.098	0.009
$M = 96$	0.011	0.854	0.508	0.450	-0.382	69.371	1E-4	-0.107

Table 6.16: PMEMM(M) Parameters: Various M: SVSJ Model.

	$M = 3$	$M = 6$	$M = 12$	$M = 24$	$M = 48$	$M = 96$
Time	13.2	17.4	57.8	67.1	76.3	408

Table 6.17: Time in Seconds to Compute PMEMM(M): SVSJ Model.

	3 mo.	**6 mo.**	**12 mo.**	**18 mo.**	**24 mo.**	S_0	338.84
$M = 3$	32.098	45.783	65.209	80.095	92.538	r	0.2934%
$M = 6$	31.677	45.448	65.185	80.390	93.112	-	
$M = 12$	32.599	46.448	65.966	80.821	93.194	-	
$M = 24$	32.598	46.436	66.008	80.934	93.375	-	
$M = 48$	32.635	46.467	66.050	80.992	93.447	-	
$M = 96$	21.676	39.927	72.151	99.542	122.95	-	

Table 6.18: ATM Call Prices at PMEMM(M): SVSJ Model.

Again, for $M = 96$, the PMEMM(M) for the SVSJ Merton model diverges from its pattern. Also in this case, the at-the-money call prices are affected at $M = 96$. But, the choices $M = 24$ and $M = 48$ both remain reasonable for the SVSJ Merton model PMEMM(M).

In summary, we have found that the Heston and SVJ Merton models each have a robust PMEMM(M) as M reaches 96. For the SVVG, and SVSJ Merton models, we have found that for $T = 2$, M should not exceed 48. However, for the SVJ Meixner model it appears that we must choose $M = 24$. For reasons of both stability and computational speed, we have chosen $M = 24$ months, when $T = 2$, for all five models.

6.6 Conclusion

1. For each of the five models, the PMEMM provides a good forecast of longer term at-the-money discounted call option payoffs. This is consistent with the main result from Chernov and Ghysels (2000), that the kurtosis found in historical returns is a necessary ingredient for the pricing of late maturity options. But, the option implied Q-measure provides a better forecast of the same, for shorter maturities. Note that no particular model stands out, suggesting that it is the PMEMM itself that leads to superior long term forecasting success.

2. In future research, we propose to obtain bootstrap standard errors for the parameters of the PMEMM as follows. First, using the asymptotic normal distribution from Remark 5.1 in subsection 5.3.2, simulate 100 copies of $\theta_Q^*(m) \in \Omega_Q$, $m = 1, 2, ..., 100$. Next, by using the same P-measure parameters, $\theta_P \in \Omega_P$, for each run, execute the two stage optimization routine from subsection 6.3.1 for the PMEMM, obtaining $\widetilde{\theta}_Q(m) \in \Omega_Q$, $m = 1, 2, ..., 100.$, corresponding to each $\theta_Q^*(m) \in \Omega_Q$, $m = 1, 2, ..., 100$. Then compute bias, RMSE, and standard errors.

3. The PMEMM is an easily computable approximation to a version of the minimum relative entropy Q-measure, proposed as a solution for incomplete market models. In future research, we propose to make a

direct numerical comparison between Heston call options priced under the MEMM and the PMEMM, respectively. The Heston call options under the MEMM will be priced in a manner similar to the numerical boundary value problem solved in He and Zhu (2016). The Heston call options under the PMEMM will be priced by the Carr and Madan (1999) formula. This comparison will reveal any possible similarity between the MEMM and the PMEMM.

4. For a fixed time horizon T, as the number M of uniformly placed maturity dates used to compute the PMEMM gets large, the Heston and SVJ Merton models appear to be the most self-consistent of the five models considered for the PMEMM, in the sense that the PMEMM parameters and at-the-money call prices stay on the same course as M gets large for these two models in particular.

Part III

Monte Carlo Methods

Preface to Part III

This part covers aspects of the thesis that could only be treated by Monte Carlo. In Chapter 7 we introduce a new conditional Monte Carlo pricing technique, based on a conditional version of the Carr and Madan (1999) formula. We call our method conditional FFT. Instead of averaging the conditional call prices, we average the Fourier transforms of the respective conditional damped call prices, and prove that this latter average converges almost surely to the Fourier transform of the damped call price. Thus, conditional FFT requires only one fast Fourier transform. Also in Chapter 7, by conditioning on the jump intensity, we are able to calibrate the non-affine Leveraged Jump Intensity (LJI) model from Chapter 2 to the selection of spot options prices on Apple stock, as introduced in Chapter 5 above. These results show that the model for the leverage effect provided by the LJI model leads to a much lower RMSE than the Barndorff-Nielsen and Shephard (2001) stochastic volatility model. Then in Chapter 8 we provide a new acceptance-rejection technique for simulating Meixner increments, called the method of least squares rejection. We use a normal inverse-Gaussian rejection density. This technique was used for simulating the SVJ Meixner model in Chapter 4 above. We also show that our technique simulates European options prices under the Meixner model more accurately than the method illustrated in Madan and Yor (2008).

Chapter 7

Conditional FFT and the LJI Meixner Model

7.1 Introduction

The main goal of this chapter is to calibrate by least squares the new LJI model from subsection 2.7.2, under the Meixner jumps of subsection 2.3.3, to the closing book of call options prices on Apple stock, for January 19th, 2011. However, since the LJI model does not have a closed form log-price characteristic function, we require an alternative pricing mechanism in order to calibrate. In the early literature on stochastic volatility, conditional Monte Carlo was proposed as a European call option pricing technique, see Hull and White (1987), pp. 289-90, where the call price is conditioned on the average variance of the stock price over the life of the option. It was believed for some time that this application of conditional Monte Carlo required the correlation ρ between the stock price and the variance of the stock price to be zero. Willard (1997) shows that if instead the entire path of the Brownian motion driving the stock price variance is conditioned

upon, then any $\rho \in [-1, 1]$ is possible for call option pricing under stochastic volatility by conditional Monte Carlo. It is also shown in Willard (1997), see Exhibit 4, p. 56, that conditional quasi-Monte Carlo (QMC) outperforms conditional Monte Carlo for the Heston model. However, as pointed out in Imai and Tan (2006), see pp. 149-50, the Willard (1997) analysis relies on the Black-Scholes formula for conditional pricing. Thus, the Willard (1997) analysis is strictly limited to diffusions.

In this chapter we propose a new conditional Monte Carlo call option pricing method that is based on conditional characteristic functions and the inverse Fourier transform. Specifically, for conditional call option pricing, our method employs a conditional version of the Carr and Madan (1999) formula, to be derived in subsection 7.4.1 below. The method is broad enough to cover models with jumps and stochastic jump intensity. Thus, it is suitable for the LJI model. In the development below we assume that the LJI model has Meixner jumps. But, LJI model results for VG and Merton jumps are also given in subsection 7.5.3, at the end of this chapter.

In Section 7.2 we begin with methods of pathwise simulation for the CIR process by both Monte Carlo and QMC, since we ultimately condition on either the volatility or the jump intensity. Then in Section 7.3 we derive a conditional log-price characteristic function for each of the time changed stochastic exponential Meixner model (TCMX), the LJI Meixner model, and the SVSJ Merton model. In Section 7.4 we formally derive the conditional Carr and Madan formula, and the new call option pricing method that we refer to as conditional FFT. Also, we provide benchmark pricing tests against the unconditional Carr and Madan (1999) formula for the TCMX and SVSJ models. Then in Section 7.5 we undertake benchmark calibration of the TCMX and SVSJ models. For comparison to the LJI Meixner model, we also calibrate the Barndorff-Nielsen and Shephard stochastic volatility

model (BN-S) from subsection 2.7.1 by exact FFT. Lastly, we calibrate the LJI Meixner model, and the LJI model under both VG and Merton jumps, each by conditional FFT. We conclude in Section 7.6.

7.2 Pathwise Simulation of the CIR Process

In this section we illustrate the simulation of CIR process paths. Below in subsection 7.2.1 and subsection 7.2.2 we treat the simulation of Brownian motion paths by Monte Carlo, and quasi-Monte Carlo (QMC), respectively. Then in subsection 7.2.3 we present a drift-implicit Milstein scheme for simulating the CIR process from Brownian motion paths. This drift-implicit Milstein scheme remains positive for $4\kappa\eta > \omega^2$, see Kahl, Günther, and Rossberg (2008), p. 289, and has strong convergence for $2\kappa\eta > \omega^2$, see Kloeden and Neuenkirch (2013), p. 71.

7.2.1 Simulating Brownian Paths by Monte Carlo

For Monte Carlo, we simulate antithetic standard Brownian motion paths, as recommended for European call option pricing by conditional Monte Carlo in Hull and White (1987), p. 289, and in Willard (1997), p. 51.

Let W_t, $t \in [0, T]$ be a standard Brownian motion. For a suitable number of time steps J, we discretize time by

$$\Delta t = \frac{T}{J}, \text{ and} \tag{7.1}$$

$$t_j = j\Delta t,\ j = 0, 1, 2, ..., J-1. \tag{7.2}$$

Let M be the number of standard Brownian motion paths to be simulated, and assume that M is even. Then simulate

$$U_{m,j+1} \sim Unif\left[0, 1\right),\ i.i.d. \tag{7.3}$$

$$\text{for } m = 1, 2, ..., \frac{M}{2}, \text{ and } j = 0, 1, 2, ..., J-1.$$

In traditional Monte Carlo, each $U_{m,j+1}$ is treated as a random number. In quasi-Monte Carlo, we relax this assumption. One consequence is that traditional Monte Carlo estimates have a sample standard error.

Given the $U_{m,j+1}$ from equation (7.3), we define antithetic standard Brownian motion paths as follows. For $j = 0, 1, 2, ..., J-1$, let

$$Z_{m,j+1} = \Phi^{-1}\left(U_{m,j+1}\right), \text{ for } m = 1, 2, ..., \frac{M}{2}, \text{ and} \tag{7.4}$$

$$Z_{m,j+1} = -\Phi^{-1}\left(U_{m-\frac{M}{2},j+1}\right), \text{ for } m = \frac{M}{2} + 1, ..., M. \tag{7.5}$$

Then, for $m = 1, 2, ..., M$, starting from $W_{m,0} = 0$, iterate

$$W_{m,j+1} = W_{m,j} + \sqrt{\Delta t} Z_{m,j+1},\ j = 0, 1, 2, ..., J-1. \tag{7.6}$$

Furthermore, for $m = 1, 2, ..., M$, letting

$$\Delta W_{m,t_{j+1}} = W_{m,t_{j+1}} - W_{m,t_j},\ j = 0, 1, 2, ..., J-1, \tag{7.7}$$

we obtain the increments $\Delta W_{m,j+1} \sim N(0, \Delta t)$ of the M standard Brownian motion paths, each with J equal time steps of length $\Delta t = \frac{T}{J}$. For the Brownian motion paths sampled by Monte Carlo, we simulate a total of $M_{mc} = 8192$ paths, twice the number that we generate for QMC, see subsection 7.2.2 below. In the benchmark pricing of subsection 7.4.4, this shows that QMC prices more efficiently than Monte Carlo, even with only half the number of simulations. The numbers of time steps J_T per maturity are the same for Monte Carlo as for QMC, see subsection 7.2.2 below.

7.2.2 Simulating Brownian Paths by Quasi-Monte Carlo

A QMC sequence is a non-random sequence of the form $\left\{\mathbf{u}_m | \mathbf{u}_m \in [0,1)^J\right\}$ for some positive integer dimension J, offering an alternative to a Monte Carlo. The main idea behind a QMC sequence is to distribute points within $[0,1)^J$ as evenly as possible, see Glasserman (2004), p. 281. A set

of QMC points is defined by $P_M = \left\{ \mathbf{u}_m | \mathbf{u}_m \in [0,1)^J \right\}_{m=1}^{M}$ for some positive integer M which is typically less than the number of simulations required for ordinary Monte Carlo. Sobol' sequences are a particular implementation of QMC, see Glasserman (2004), §5.2.3. In this thesis we use the high dimensional Sobol' sequences with maximum dimension 1,111 proposed in Joe and Kuo (2003), see p. 52. Specifically, we use the Matlab R2016a Sobol' sequence generator based on the Joe and Kuo (2003) paper.

For Brownian motion paths generated by QMC with Sobol' sequences, we use $M_{qmc} = 4096$ paths. This is half the number of paths used for Monte Carlo, see subsection 7.2.1 above. For QMC we do not use antithetic paths, as they are of no consequence in this case, see Willard (1997), p. 59, note 14. In each of our applications, the dimension J of each Sobol' sequence is the number J_T of time steps to maturity T along the generated path of the Brownian motion driving the CIR process, where the CIR process models the latent factor that we condition on. Recall from subsection 5.3.1 that for the Apple call options data set of this thesis, based on the January 19th, 2011 close, the exact maturity dates are given by 22, 41, 61, 123, 192, 253, and 503 days, respectively. Given a 21 day month, these maturities are approximately 1, 2, 3, 6, 9, 12, and 24 months. Thus, we use the exact number of days, divided by 252, for each maturity T in years, and we determine the number of time steps J_T to maturity by the approximate number of months to maturity. Thus, with 46 time steps per month, we require Sobol' sequence dimensions of 46, 92, 138, 276, 414, 552, and 1104, respectively.

The QMC approximation is typically described as an integral over the unit hypercube of dimension J. The effective dimension J' is the integer number of important variables in the problem, in the sense that these important variables determine most of the solution, see Lemieux (2008), p. 216.

High dimensional problems are more likely to succeed when J' is much less than J. Typically J' is unknown and can only be estimated, see Lemieux (2008), pp. 225-28. However, Glasserman (2004), pp. 86-88, shows how to illustrate the effective dimension obtained from the principal components construction of a standard Brownian motion path, see Table 7.1 following the analysis below. As recommended in Acworth, Broadie, and Glasserman (1998), see pp. 11-12, we use the principal components construction for the quasi-Monte Carlo approach to simulating Brownian motion paths.

We now proceed in a manner similar to Monte Carlo in subsection 7.2.1 above. We let W_t, $t \in [0, T]$ be a standard Brownian motion, and for a suitable number of time steps J, we discretize time by

$$\Delta t = \frac{T}{J}, \text{ and} \tag{7.8}$$

$$t_j = j\Delta t,\ j = 0, 1, 2, ..., J. \tag{7.9}$$

Let M be the number of standard Brownian motion paths to be generated, and to improve the uniformity of the Sobol' sequence, assume that M is a power of 2, see Imai and Tan (2006), p. 143. Then, for $m = 1, 2, ..., M$, with $W_m(0) = 0$, each vector

$$W_m = [W_m(t_1), W_m(t_2), ..., W_m(t_J)]', \tag{7.10}$$

is a discrete standard Brownian motion path. The construction of the paths proceeds as follows. Generate

$$u_{m,j+1} \in [0, 1), \tag{7.11}$$

$$\text{for } m = 1, 2, ..., M, \text{ and } j = 0, 1, 2, ..., J - 1,$$

from a J dimensional Sobol' sequence $\left\{\mathbf{u}_m | \mathbf{u}_m \in [0, 1)^J\right\}$. Then set

$$Z_{m,j+1} = \Phi^{-1}(u_{m,j+1}), \tag{7.12}$$

$$\text{for } m = 1, 2, ..., M, \text{ and } j = 0, 1, 2, ..., J - 1.$$

We now define the collection of vectors

$$Z_m = [Z_{m,1}, Z_{m,2}, ...Z_{m,J}]', \quad m = 1, 2, ..., M. \tag{7.13}$$

Thus, each $J \times 1$ vector W_m, see equation (7.10) above, is given by

$$W_m = AZ_m, \tag{7.14}$$

where the $J \times J$ matrix A satisfies $AA' = \Sigma$, such that Σ is the covariance matrix of the discrete standard Brownian motion path W_m. This is discussed in Glasserman (2004), see p. 87. Moreover, the autocovariance function for standard Brownian motion is given by

$$c(s,t) = \min\{s,t\}, \text{ for all } s,t \geq 0, \tag{7.15}$$

see Grimmett and Stirzaker (2001), p. 516. Therefore, assuming equal time steps Δt, equation (7.15) implies that for each $m = 1, 2, ..., M$,

$$\Sigma = AA' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & \cdots & J \end{bmatrix} \Delta t, \tag{7.16}$$

is the covariance matrix of W_m. Thus, since Σ is symmetric, it is orthogonally diagonalizable such that

$$A = PD^{\frac{1}{2}}, \tag{7.17}$$

where D is a diagonal matrix whose nonzero entries are the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_J$ of the covariance matrix Σ, and the matrix P has columns given by the corresponding normalized eigenvectors $v_1, v_2, ..., v_J$. The main idea of the principal components construction is to rank the eigenvalues from largest to smallest. Then since the *jth* column of A is given by

$$A_j = \sqrt{\lambda_j} v_j, \; j = 1, 2, ..., J, \tag{7.18}$$

see Glasserman (2004), p. 87, it follows from equation (7.12) above that the standard Brownian paths given by $W_m = AZ_m$, $m = 1, 2, ..., M$, see equation (7.14) above, are determined primarily by the first few dimensions of the Sobol' sequence. This is desirable since the first few dimensions of a QMC sequence are known to have better equidistributional properties, see Acworth, Broadie, and Glasserman (1998), p. 7.

We now illustrate the effective dimension obtained from the principal components construction, as in Glasserman (2004), p. 88. Let $J_T(\alpha)$ be the smallest index k such that

$$\frac{\lambda_1 + \lambda_2 + ... + \lambda_k}{\lambda_1 + \lambda_2 + ... + \lambda_J} > \alpha \in (0, 1). \tag{7.19}$$

Table 7.1 below considers $J_T(\alpha)$ for the dimensions used in this chapter.

T **in Days**	**Dimension** J_T	$J_T(.95)$	$J_T(.99)$	**P.C. CPU Time**
22	46	4	18	0.011 sec
41	92	5	20	0.021 sec
61	138	5	20	0.032 sec
123	276	5	21	0.069 sec
192	414	5	21	0.115 sec
253	552	5	21	0.175 sec
503	1104	5	21	0.453 sec

Table 7.1: P.C. Analysis: An Illustration of the Effective Dimension.

It is suggested in Table 7.1 above that the effective dimension of the Brownian motion paths is made small by the principal components construction, particularly for the 253 day and 503 day maturities. However, the computational complexity of the principal components strategy is $O\left(MJ_T^2\right)$, see Lemieux (2008), p. 223. Thus, the CPU time to process $M = 4096$ paths with dimension $J_T = 1,104$ is disproportionately large.

Along each path $m = 1, 2, ..., M$, the Brownian motion increments

$$\Delta W_{m,t_{j+1}},\ j = 0, 1, 2, ..., J-1, \tag{7.20}$$

are obtained by taking differences over each path vector

$$W_m = \left[W_m\left(t_1\right), W_m\left(t_2\right), ..., W_m\left(t_J\right)\right]', \text{ with}$$
$$W_m\left(0\right) = 0.$$

Lastly, there is a field of study called randomized quasi-Monte Carlo (RQMC), see Lemieux (2008), §6.2, which allows a standard error to be computed. However, in this thesis we only require the exact QMC estimates for option price calibration in Section 7.5 below. Thus, we do not pursue RQMC.

7.2.3 Simulating CIR Paths from Brownian Motion Paths

In this subsection we present the drift-implicit Milstein approximation to the CIR process from Kahl, Günther, and Rossberg (2008), see p. 289, as the method that we use to simulate CIR paths in this chapter. The main features of this method are that it remains positive whenever $4\kappa\eta > \omega^2$, and also exhibits strong convergence whenever $2\kappa\eta > \omega^2$. The drift-implicit Milstein scheme for the CIR process is given by

$$\overline{X}_{t_{j+1}} = \overline{X}_{t_j} + \kappa\left(\eta - \overline{X}_{t_{j+1}}\right)\Delta t + \omega\sqrt{\overline{X}_{t_j}}\Delta W_{t_{j+1}} + \frac{\omega^2}{4}\left(\Delta W_{t_{j+1}}^2 - \Delta t\right).$$

Lemma 7.1 *The drift-implicit Milstein scheme for the CIR process remains positive for all $4\kappa\eta > \omega^2$ and may be written in explicit form as*

$$\overline{X}_{t_{j+1}} = \frac{\left(\sqrt{\overline{X}_{t_j}} + \frac{\omega}{2}\Delta W_{t_{j+1}}\right)^2 + \left(\kappa\eta - \frac{\omega^2}{4}\right)\Delta t}{1 + \kappa\Delta t}, \tag{7.21}$$

as in Kloeden and Neuenkirch (2013), p. 70.

Proof. Working backwards from equation (7.21) above we obtain

$$\begin{aligned}(1+\kappa\Delta t)\,\overline{X}_{t_{j+1}} &= \left(\sqrt{\overline{X}_{t_j}} + \frac{\omega}{2}\Delta W_{t_{j+1}}\right)^2 + \left(\kappa\eta - \frac{\omega^2}{4}\right)\Delta t \\ &= \overline{X}_{t_j} + \omega\sqrt{\overline{X}_{t_j}}\Delta W_{t_{j+1}} + \frac{\omega^2}{4}\Delta W_{t_{j+1}}^2 + \kappa\eta\Delta t - \frac{\omega^2}{4}\Delta t \\ &= \overline{X}_{t_j} + \kappa\eta\Delta t + \omega\sqrt{\overline{X}_{t_j}}\Delta W_{t_{j+1}} + \frac{\omega^2}{4}\left(\Delta W_{t_{j+1}}^2 - \Delta t\right),\end{aligned}$$

and this implies that

$$\overline{X}_{t_{j+1}} = \overline{X}_{t_j} + \kappa\left(\eta - \overline{X}_{t_{j+1}}\right)\Delta t + \omega\sqrt{\overline{X}_{t_j}}\Delta W_{t_{j+1}} + \frac{\omega^2}{4}\left(\Delta W_{t_{j+1}}^2 - \Delta t\right),$$

as required. That $\overline{X}_{t_j}$ in the drift-implicit scheme remains positive for all $4\kappa\eta > \omega^2$ follows from the explicit form in equation (7.21) above. ■

For the case $\omega^2 > 4\kappa\eta$, Kloeden and Neuenkirch (2013), see p. 73, uses the truncated scheme given by

$$\overline{X}_{t_{j+1}} = \frac{\left(\sqrt{\overline{X}_{t_j}^+} + \frac{\omega}{2}\Delta W_{t_{j+1}}\right)^2 + \left(\kappa\eta - \frac{\omega^2}{4}\right)\Delta t}{1+\kappa\Delta t}. \tag{7.22}$$

But, in this thesis for the case $\omega^2 > 4\kappa\eta$ we use the scheme given by

$$\overline{X}_{t_{j+1}} = \left[\frac{\left(\sqrt{\overline{X}_{t_j}} + \frac{\omega}{2}\Delta W_{t_{j+1}}\right)^2 + \left(\kappa\eta - \frac{\omega^2}{4}\right)\Delta t}{1+\kappa\Delta t}\right]^+, \tag{7.23}$$

in the drift-implicit Milstein scheme for the CIR process.

Proposition 7.2 *Let $2\kappa\eta > \omega^2$, $x_0 > 0$, and $T > 0$. Then the drift-implicit Milstein scheme in equation* (7.21) *above has strong convergence such that*

$$\lim_{n\to\infty} E\left[\max_{j=0,1,\ldots,n}\left|X_{t_j} - \overline{X}_{t_j}^n\right|^2\right] = 0. \tag{7.24}$$

Proof. See Kloeden and Neuenkirch (2013), pp. 71-73. ■

In the recent literature there is another well known scheme for simulating the CIR process that has a lot in common with the drift-implicit Milstein

scheme of equation (7.21) above. However, in this thesis we ultimately choose the drift-implicit Milstein scheme of equation (7.21) above for pricing and calibration. The alternative scheme is referred to as the drift-implicit square-root Euler method. Letting $Y_t = \sqrt{X_t}$ in the CIR process, after applying Itô's lemma we obtain the drift-implicit form

$$\overline{Y}_{t_{j+1}} = \overline{Y}_{t_j} + \left(\frac{4\kappa\eta - \omega^2}{8\overline{Y}_{t_{j+1}}} - \frac{\kappa}{2}\overline{Y}_{t_{j+1}} \right) \Delta t + \frac{\omega}{2}\Delta W_{t_{j+1}}, \tag{7.25}$$

which has solution given by

$$\overline{Y}_{t_{j+1}} = \frac{\overline{Y}_{t_j} + \frac{\omega}{2}\Delta W_{t_{j+1}}}{2\left(1 + \frac{\kappa}{2}\Delta t\right)} + \sqrt{\frac{\left(\overline{Y}_{t_j} + \frac{\omega}{2}\Delta W_{t_{j+1}}\right)^2}{4\left(1 + \frac{\kappa}{2}\Delta t\right)^2} + \frac{\left(4\kappa\eta - \omega^2\right)\Delta t}{8\left(1 + \frac{\kappa}{2}\Delta t\right)}}, \tag{7.26}$$

see Kloeden and Neuenkirch (2013), p. 69. The idea of the drift-implicit square-root Euler method for the CIR process is to set $\overline{X}_{t_j} = \overline{Y}^2_{t_j}$, for each $j = 0, 1, 2, ...$, and to thereby obtain a positive solution for all possible CIR parameters. While this may seem attractive, it does not appear to lead to superior strong form convergence rates relative to the drift-implicit Milstein method for all scenarios tested, see Kloeden and Neuenkirch (2013), p. 74. These results are also cited in Table 7.2 below. With $\overline{X}_{t_j} = \overline{Y}^2_{t_j}$, for each $j = 0, 1, 2, ...$, based on equation (7.26), the drift-implicit square-root Euler scheme for the CIR process may be written as

$$\begin{aligned} \overline{X}_{t_{j+1}} = & \frac{\left(\sqrt{\overline{X}_{t_j}} + \frac{\omega}{2}\Delta W_{t_{j+1}}\right)^2 + \left(\kappa\eta - \frac{\omega^2}{4}\right)\Delta t}{1 + \kappa\Delta t} \\ & - \frac{1}{1 + \kappa\Delta t}\left(\frac{4\kappa\eta - \omega^2}{8\sqrt{\overline{X}_{t_{j+1}}}} - \frac{\kappa}{2}\sqrt{\overline{X}_{t_{j+1}}} \right)^2 (\Delta t)^2 , \end{aligned} \tag{7.27}$$

see Kloeden and Neuenkirch (2013), p. 70. Comparing equation (7.27) to equation (7.21) above, we see that for the CIR process, the drift-implicit Milstein scheme agrees with the drift-implicit square-root Euler scheme up to a second order term.

Kloeden and Neuenkirch (2013), see pp. 73-75, employs Monte Carlo simulation to estimate the strong rate of convergence for the drift-implicit Milstein scheme in equation (7.21) above, and the drift-implicit square-root Euler scheme in equation (7.27) above. In Table 7.2 below we present these results. For each scenario, the higher of two similar estimates from Kloeden and Neuenkirch (2013), p. 74, is given. As Table 7.2 indicates, if $2\kappa\eta > \omega^2$, then a strong convergence rate near the Milstein ideal of 1.0 is evident, surprisingly for both schemes. But, if $\omega^2 > 4\kappa\eta$, then the rate appears to be less than the Euler ideal of 0.5, for both schemes.

		Est. Strong Convergence Rate	
Scenario	**Ratio** $\frac{\omega^2}{\kappa\eta}$	**Implicit Milstein**	**Square Root Euler**
$2\kappa\eta > \omega^2$	0.9944	0.9447	0.9281
$\omega^2 > 4\kappa\eta$	5.556	0.3096	0.2734

Table 7.2: Convergence Rates: Implicit Milstein vs. Square Root Euler.

In subsection 7.4.4 below we compare the drift-implicit Milstein and drift-implicit square-root Euler schemes on the basis of call option pricing.

7.3 Main Conditional Characteristic Functions

In this section we derive a risk-neutral conditional characteristic function (CCF) for each of the stochastic exponential time changed Meixner (TCMX) model, the LJI Meixner model, and SVSJ Merton (SVSJ) model, for the purposes European call option pricing, by the method of conditional FFT. In each of these models, the latent factor is modeled by the CIR process. The main idea from Willard (1997), for the Heston model, is to condition on the entire path of the Brownian motion driving the latent CIR process for $t \in [0, T]$. By doing this, we can reduce the problem to conditioning

on terminal functions taken at time T along this path. Thus, we consider terminal time $T = \tau$, with initial time $t_0 = 0$. Letting S_0 be the initial stock price, we define

$$Z_T = \log\left(\frac{S_T}{S_0}\right). \tag{7.28}$$

Specific to each model, we condition Z_T on H_T, where H_T is either one or two dimensional. Thus, each CCF takes the form

$$\phi_{Z_T}(u|H_T) = E\left[e^{iuZ_T}|H_T\right]. \tag{7.29}$$

In subsection 7.3.1 we derive the TCMX model CCF. The LJI Meixner model CCF is derived in subsection 7.3.2. For the LJI Meixner model, we have no prior knowledge of the joint distribution of $H_T = (W_T, \Lambda_T)$, where $\Lambda_T = \int_0^T \lambda_t dt$ and $W_T = \int_0^T dW_t$, with W_t being the Brownian motion driving the jump intensity λ_t. Thus, we must use paths to simulate $H_T = (W_T, \Lambda_T)$ in this case. Lastly, in subsection 7.3.3 we derive the CCF for the SVSJ Merton model.

7.3.1 The Time Changed Meixner Model

Recall from equation (2.70) in subsection 2.5.3 and equations (2.39) and (2.40) from subsection 2.3.3 that the time changed Meixner stochastic exponential model for the risk-neutral log-price is given by

$$dY_t = (r - \lambda_t \psi_X(-i))\, dt + dX(\Lambda_t), \tag{7.30}$$

$$d\lambda_t = \kappa(\eta - \lambda_t)\, dt + \omega\sigma_t dW_t, \tag{7.31}$$

$$d\Lambda_t = \lambda_t dt,\ \lambda_0 = 1, \text{ with} \tag{7.32}$$

$$\psi_{MX}(u) = 2\lambda \log\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cosh\left(\frac{1}{2}(\alpha u - i\beta)\right)}\right), \text{ and} \tag{7.33}$$

$$\psi_{MX}(-i) = 2\lambda \log\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cos\left(\frac{\alpha+\beta}{2}\right)}\right). \tag{7.34}$$

For convenience, based on equations (7.33) and (7.34) above, we define

$$\widehat{\psi}_{MX}(u) = \psi_{MX}(u) - iu\psi_{MX}(-i). \tag{7.35}$$

The following result is a consequence of Theorem 2.5, the Lévy-Khinchin formula, from subsection 2.2.1, see Carr and Wu (2004), pp. 117-19.

Lemma 7.3 *Let $Y_t = X(\Lambda_t)$ be a time changed Lévy process such that the time change process Λ_t is independent of the Lévy process X. Assume that the Lévy process X has a characteristic exponent given by $\psi_X(u)$. Then*

$$\phi_{Y_t}(u|\Lambda_t) = E\left[e^{iuY_t}|\Lambda_t\right] = e^{\Lambda_t\psi_X(u)}. \tag{7.36}$$

Proof. Let $s \geq 0$. Then

$$\begin{aligned}\phi_{Y_t}(u|\Lambda_t = s) &= E\left[e^{iuY_t}|\Lambda_t = s\right] \\ &= E\left[e^{iuX_s}|\Lambda_t = s\right] = e^{\Lambda_t\psi_X(u)},\end{aligned}$$

by Theorem 2.5, the Lévy-Khinchin formula. ■

Lemma 7.4 *Let $H_T = \Lambda_T = \int_0^T \lambda_t dt$. Then, the CCF for the TCMX model is given by*

$$\phi_{Z_T}(u|H_T) = e^{iurT + \Lambda_T\widehat{\psi}_{MX}(u)}. \tag{7.37}$$

Proof. By integrating equation (7.30) w.r.t. $t \in [0,T]$ we obtain

$$Y_T = Y_0 + rT - \psi_{MX}(-i)\int_0^T \lambda_t dt + X(\Lambda_T). \tag{7.38}$$

Therefore, given $H_T = \Lambda_T = \int_0^T \lambda_t dt$,

$$Z_T = Y_T - Y_0 = rT + X(\Lambda_T) - \Lambda_T\psi_{MX}(-i). \tag{7.39}$$

Hence, by Lemma 7.3, and equation (7.35) above,

$$\begin{aligned}\phi_{Z_T}(u|H_T) &= E\left[e^{iuZ_T}|\Lambda_T\right] \\ &= e^{iurT + \Lambda_T\widehat{\psi}_{MX}(u)},\end{aligned}$$

as required. ■

7.3.2 The LJI Meixner Model

Recall from subsection 2.7.2 that for some pure jump Lévy process X, the Q-dynamics of the LJI model are given by

$$dY_t = \left(r - \lambda_t \psi_X(-i) - \frac{1}{2}\gamma^2\right) dt + dX(\Lambda_t) - \gamma dW_t, \tag{7.40}$$

$$d\lambda_t = \kappa(\eta - \lambda_t)\, dt + \omega\sqrt{\lambda_t} dW_t, \tag{7.41}$$

$$d\Lambda_t = \lambda_t dt \ , \ \lambda_0 = 1, \text{ with } \gamma > 0. \tag{7.42}$$

As in subsection 7.3.1 above for the TCMX model, we let the Lévy process X for the LJI model be Meixner so that $\psi_{MX}(u)$ and $\psi_{MX}(-i)$ are the same as in equations (7.33) and (7.34) respectively,

$$\text{and } \widehat{\psi}_{MX}(u) = \psi_{MX}(u) - iu\psi_{MX}(-i). \tag{7.43}$$

Observe that the TCMX model of subsection 7.3.1 above is a special case of the LJI Meixner model with the parameter γ in equation (7.40) set to zero.

Lemma 7.5 *For the LJI model given above, let* $W_T = \int_0^T dW_t$, *and let* $\Lambda_T = \int_0^T \lambda_t dt$. *Then, where* $H_T = (W_T, \Lambda_T)$, *the CCF is given by*

$$\phi_{Z_T}(u|H_T) = e^{iu\left(rT - \frac{1}{2}\gamma^2 T - \gamma W_T\right) + \Lambda_T \widehat{\psi}_{MX}(u)}. \tag{7.44}$$

Proof. By integrating equation (7.40) w.r.t. $t \in [0,T]$ we obtain

$$Y_T = Y_0 + rT - \psi_{MX}(-i)\int_0^T \lambda_t dt - \frac{1}{2}\gamma^2 T + X(\Lambda_T) - \gamma\int_0^T dW_t. \tag{7.45}$$

Therefore, given $W_T = \int_0^T dW_t$, and $\Lambda_T = \int_0^T \lambda_t dt$

$$Z_T = Y_T - Y_0 = rT - \frac{1}{2}\gamma^2 T - \gamma W_T + X(\Lambda_T) - \Lambda_T \psi_{MX}(-i). \tag{7.46}$$

Hence, by Lemma 7.3 from subsection 7.3.1, and equation (7.43) above,

$$\begin{aligned}\phi_{Z_T}(u|H_T) &= E\left[e^{iuZ_T}|(W_T,\Lambda_T)\right] \\ &= e^{iu\left(rT-\frac{1}{2}\gamma^2T-\gamma W_T\right)+\Lambda_T\widehat{\psi}_{MX}(u)},\end{aligned}$$

as required. ■

7.3.3 The SVSJ Merton Model

Recall from equation (2.74) in subsection 2.6.2 that the Q-dynamics of the SVSJ Merton model are given by

$$dY_t = \left(r-\frac{1}{2}\sigma_t^2-\sigma_t^2\psi_X(-i)\right)dt+\sigma_t dW_t^{(S)}+dX(V_t), \tag{7.47}$$

$$d\sigma_t^2 = \kappa\left(\eta-\sigma_t^2\right)dt+\omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right]=\rho dt, \tag{7.48}$$

$$dV_t = \sigma_t^2 dt. \tag{7.49}$$

Note that in equations (7.47) and (7.48) above, for an independent pair of Brownian motions $W_t^{(1)}$ and $W_t^{(2)}$, we may write

$$W_t^{(S)} = \rho W_t^{(1)}+\sqrt{1-\rho^2}W_t^{(2)}, \tag{7.50}$$

$$\text{and } W_t^{(V)} = W_t^{(1)}. \tag{7.51}$$

Also, in the SVSJ model the jumps are Merton. Thus, by equations (2.23) and (2.24) from subsection 2.3.1 we have

$$\psi_{MJ}(u) = \lambda\left(\exp\left(i\beta u-\frac{1}{2}\alpha^2u^2\right)-1\right), \tag{7.52}$$

$$\psi_{MJ}(-i) = \lambda\left(\exp\left(\beta+\frac{1}{2}\alpha^2\right)-1\right), \tag{7.53}$$

$$\text{and we define } \widehat{\psi}_{MJ}(u) = \psi_{MJ}(u)-iu\psi_{MJ}(-i). \tag{7.54}$$

Lemma 7.6 *Let σ_t be a given deterministic function of time, and let B_t be a standard Brownian motion. Define the process*

$$A_t = \int_0^t \sigma_s dB_s. \tag{7.55}$$

Then $A_t \sim N\left(0, \int_0^t \sigma_s^2 ds\right)$.

Proof. See Björk (2009), p. 57. ■

Lemma 7.7 *In the SVSJ model, condition on $\mathcal{F}_T^{(1)} = \sigma\left\{W_t^{(1)}, t \in [0,T]\right\}$, such that the random variables $G_T = \int_0^T \sigma_t dW_t^{(1)}$, and $V_T = \int_0^T \sigma_t^2 dt$ are given. Then, where $H_T = (G_T, V_T)$, the CCF reduces to*

$$\phi_{Z_T}(u|H_T) = e^{iu\left(rT - \frac{1}{2}V_T + \rho G_T\right) - \frac{1}{2}V_T\left(1-\rho^2\right)u^2 + V_T\widehat{\psi}_{MJ}(u)}. \tag{7.56}$$

Proof. By integrating equation (7.47) w.r.t. $t \in [0,T]$, and taking note that $V_T = \int_0^T \sigma_t^2 dt$, after re-arranging we have

$$Y_T = Y_0 + rT - \frac{1}{2}V_T + \int_0^T \sigma_t dW_t^{(S)} + X(V_T) - V_T\psi_{MJ}(-i), \tag{7.57}$$

where by equation (7.50)

$$\int_0^T \sigma_t dW_t^{(S)} = \rho\int_0^T \sigma_t dW_t^{(1)} + \sqrt{1-\rho^2}\int_0^T \sigma_t dW_t^{(2)}. \tag{7.58}$$

Thus, since $G_T = \int_0^T \sigma_t dW_t^{(1)}$, and $V_T = \int_0^T \sigma_t^2 dt$, we have

$$\begin{aligned} Z_T &= Y_T - Y_0 \\ &= rT - \frac{1}{2}V_T + G_T + A_T + X(V_T) - V_T\psi_{MJ}(-i), \end{aligned} \tag{7.59}$$

$$\text{where } A_T = \sqrt{1-\rho^2}\int_0^T \sigma_t dW_t^{(2)}. \tag{7.60}$$

However, since $W_t^{(V)} = W_t^{(1)}$, by the variance equation (7.48) of the SVSJ model, for $t \in [0,T]$ the process σ_t is deterministic given $\mathcal{F}_T^{(1)} = \sigma\left\{W_t^{(1)}, t \in [0,T]\right\}$. Hence, by Lemma 7.6,

$$A_T = \sqrt{1-\rho^2}\int_0^T \sigma_t dW_t^{(2)} \sim N\left(0, \left(1-\rho^2\right)V_T\right). \tag{7.61}$$

Moreover, since $W_t^{(1)}$ and $W_t^{(2)}$ are independent, given V_T the quantity A_T is independent of the jumps, and the rest of the model. Thus, by equation (7.61), and by Lemma 7.3 from subsection 7.3.1 with equation (7.54),

$$\begin{aligned}\phi_{Z_T}(u|H_T) &= E\left[e^{iuZ_T}|\,(G_T,V_T)\right] \\ &= \phi_{Z_T}(u|H_T) = e^{iu\left(rT-\frac{1}{2}V_T+\rho G_T\right)-\frac{1}{2}V_T\left(1-\rho^2\right)u^2+V_T\widehat{\psi}_{MJ}(u)},\end{aligned}$$

as required. ■

7.3.4 Main CCF Realization Summary

In this subsection we summarize with the definitions and evaluation methods for the simulated values of the three main CCFs of this section.

TCMX Model CCF

Based on equation (7.37) from subsection 7.3.1 above we will denote the time changed Meixner model CCF realization by

$$\phi_{Z_T}\left(u|\Lambda_T = \xi_T^{(m)}\right) = e^{iurT+\xi_T^{(m)}\widehat{\psi}_{MX}(u)}. \tag{7.62}$$

To evaluate the TCMX model CCF realization, simply simulate the CIR intensity $\overline{\lambda}_{t_j}^{(m)}$ by the drift-implicit Milstein scheme from subsection 7.2.3 above, and then set

$$\xi_T^{(m)} = \sum\nolimits_{j=0}^{J-1}\overline{\lambda}_{t_j}^{(m)}\Delta t. \tag{7.63}$$

LJI Meixner Model CCF

Based on equation (7.44) from subsection 7.3.2 above we will denote the LJI Meixner model CCF realization by

$$\phi_{Z_T}\left(u|W_T = w_T^{(m)}, \Lambda_T = \xi_T^{(m)}\right) = e^{iu\left(rT - \frac{1}{2}\gamma^2 T - \gamma w_T^{(m)}\right) + \xi_T^{(m)}\widehat{\psi}_{MX}(u)}. \quad (7.64)$$

To evaluate the LJI Meixner model CCF realization, simply simulate the CIR intensity $\overline{\lambda}_{t_j}^{(m)}$ by the drift-implicit Milstein scheme from subsection 7.2.3 above, and track the underlying standard Brownian motion increments $\Delta W_{m,t_{j+1}}$. Then given these values set

$$w_T^{(m)} = \sum_{j=0}^{J-1} \Delta W_{m,t_{j+1}}, \quad (7.65)$$

$$\text{and } \xi_T^{(m)} = \sum_{j=0}^{J-1} \overline{\lambda}_{t_j}^{(m)} \Delta t. \quad (7.66)$$

SVSJ Merton Model CCF

Based on equation (7.56) from subsection 7.3.3 above we will denote the SVSJ Merton model CCF realization by

$$\phi_{Z_T}\left(u|G_T = g_T^{(m)}, V_T = v_T^{(m)}\right)$$
$$= e^{iu\left(rT - \frac{1}{2}v_T^{(m)} + \rho g_T^{(m)}\right) - \frac{1}{2}v_T^{(m)}\left(1-\rho^2\right)u^2 + v_T^{(m)}\widehat{\psi}_{MJ}(u)}. \quad (7.67)$$

To evaluate the SVSJ Merton model CCF realization, simply simulate the CIR variance $\overline{\sigma}_{m,t_j}^2$ by the drift-implicit Milstein scheme from subsection 7.2.3 above, and track the underlying standard Brownian motion increments $\Delta W_{m,t_{j+1}}$. Then given these values set

$$g_T^{(m)} = \sum_{j=0}^{J-1} \overline{\sigma}_{m,t_j} \Delta W_{m,t_{j+1}}, \quad (7.68)$$

$$\text{and } v_T^{(m)} = \sum_{j=0}^{J-1} \overline{\sigma}_{m,t_j}^2 \Delta t. \quad (7.69)$$

7.4 Call Option Pricing with Conditional FFT

In this section we present the conditional Monte Carlo call option pricing method that we refer to as conditional FFT. This is the method we rely on for calibration of the LJI Meixner model in subsection 7.5.3 below. The method is based on the conditional version of the Carr and Madan (1999) formula introduced in subsection 7.4.1, and can be used with the conditional log-price CFs summarized in subsection 7.3.4 above. In subsection 7.4.2 we formally define the conditional FFT estimator. In subsection 7.4.3 we derive asymptotic results for conditional FFT. In subsection 7.4.4 we benchmark conditional FFT prices for the TCMX and the SVSJ Merton models to the corresponding results from the unconditional Carr and Madan (1999) formula, and provide a pricing analysis.

7.4.1 The Conditional Carr and Madan Formula

Let T be the maturity date of a call option, and define $Z_T = \log\left(\frac{S_T}{S_0}\right)$. Then for some H_T let $f_T^Q(z|h)$ be the conditional risk-neutral density of $Z_T|(H_T = h)$. When H_T is independent of Z_T the analysis of this subsection leads back to the log-moneyness version of the Carr and Madan formula from subsection 5.2.2 above. Choose some damping parameter $a > 0$, and define the conditional damped call price to be

$$\begin{aligned} C_T^a(\chi|H_T = h) &= e^{-rT} S_0 e^{a\chi} E^Q\left[\left(e^{Z_T} - e^{\chi}\right)^+ |H_T = h\right] \\ &= e^{-rT} S_0 \int_{-\infty}^{+\infty} e^{a\chi}(e^z - e^{\chi})^+ f_T^Q(z|h)\, dz. \end{aligned} \tag{7.70}$$

Lemma 7.8 *Let the damping parameter be $a > 0$. Then*

$$\begin{aligned} &E^Q\left[S_T^{a+1}|H_T\right] < \infty \\ &\text{implies } C_T^a(\chi|H_T = h) \in L^1, \text{ for all } h \in H_T. \end{aligned} \tag{7.71}$$

Proof.

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \left|C_T^a\left(\chi|H_T=h\right)\right| d\chi = \int_{-\infty}^{+\infty} C_T^a\left(\chi|H_T=h\right) d\chi \\
&= e^{-rT} S_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{a\chi}\left(e^z - e^{\chi}\right)^+ f_T^Q\left(z|h\right) dz d\chi \\
&= e^{-rT} S_0 \int_{-\infty}^{+\infty} f_T^Q\left(z|h\right) \int_{-\infty}^{z} \left(e^{z+a\chi} - e^{(a+1)\chi}\right) d\chi dz \\
&= e^{-rT} S_0 \int_{-\infty}^{+\infty} f_T^Q\left(z|h\right) \frac{e^{(a+1)z}}{a\left(a+1\right)} dz \\
&= \frac{e^{-rT} E^Q\left[S_T^{a+1}|H_T=h\right]}{S_0^a a\left(a+1\right)} < \infty,
\end{aligned}$$

by the main hypothesis, as required. ■

We will denote the Fourier transform of the conditional damped call price by $\widehat{C}_T^a\left(u|H_T=h\right)$. By the hypothesis that $E^Q\left[S_T^{a+1}|H_T\right] < \infty$, Lemma 7.8 above implies that $\widehat{C}_T^a\left(u|H_T=h\right)$ exists. Therefore, by Definition 3.1 from subsection 3.2.1

$$\widehat{C}_T^a\left(u|H_T=h\right) = \int_{-\infty}^{+\infty} e^{iu\chi} C_T^a\left(\chi|H_T=h\right) d\chi. \tag{7.72}$$

Parallel to the original analysis from Carr and Madan (1999), see p. 64, $\widehat{C}_T^a\left(u|H_T=h\right)$ may be expressed as

$$\begin{aligned}
\widehat{C}_T^a\left(u|H_T=h\right) &= \int_{-\infty}^{+\infty} e^{iu\chi} C_T^a\left(\chi|H_T=h\right) d\chi \\
&= e^{-rT} S_0 \int_{-\infty}^{+\infty} e^{iu\chi} \int_{-\infty}^{+\infty} e^{a\chi}\left(e^z - e^{\chi}\right)^+ f_T^Q\left(z|h\right) dz d\chi \\
&= e^{-rT} S_0 \int_{-\infty}^{+\infty} f_T^Q\left(z|h\right) \int_{-\infty}^{z} e^{iu\chi}\left(e^{z+a\chi} - e^{(a+1)\chi}\right) d\chi dz \\
&= e^{-rT} S_0 \int_{-\infty}^{+\infty} f_T^Q\left(z|h\right) \left[\frac{e^{(a+1+iu)z}}{a+iu} - \frac{e^{(a+1+iu)z}}{a+1+iu}\right] dz \\
&= \frac{e^{-rT} S_0 \phi_{Z_T}\left(u-(a+1)\,i|H_T=h\right)}{a^2+a-u^2+i\left(2a+1\right)u}.
\end{aligned} \tag{7.73}$$

For simplicity, in this chapter we will refer to the function

$$\begin{aligned}\phi_{Z_T}\left(u-(a+1)\,i|H_T=h\right) &= E\left[e^{i(u-(a+1)i)Z_T}|H_T=h\right] \\ &= E\left[e^{iuZ_T+(a+1)Z_T}|H_T=h\right], \end{aligned} \tag{7.74}$$

appearing in equation (7.73) above, as the conditional damped log-price CF. We sample from this function in conditional Monte Carlo.

Lemma 7.9 *Let the damping parameter be $a>0$. If $E^Q\left[S_T^{a+1}|H_T\right]<\infty$, then there exists a finite constant $A>0$ such that*

$$\left|\phi_{Z_T}\left(u-(a+1)\,i|H_T=h\right)\right|<A, \text{ for all } h\in H_T. \tag{7.75}$$

Proof. Let $h\in H_T$. Then by equation (7.74)

$$\begin{aligned}\left|\phi_{Z_T}\left(u-(a+1)\,i|H_T=h\right)\right| &\leq E\left[\left|e^{iuZ_T+(a+1)Z_T}\right||H_T=h\right] \\ &= E\left[e^{(a+1)Z_T}|H_T=h\right] \\ &= \frac{1}{S_0}E^Q\left[S_T^{a+1}|H_T=h\right]<\infty,\end{aligned}$$

by the hypothesis that $E^Q\left[S_T^{a+1}|H_T\right]<\infty$. Hence, since $h\in H_T$ was arbitrary, the result (7.75) follows, as required. ■

The following result is also generalized from the original analysis of Carr and Madan (1999), see pp. 64-65. This result when combined with the L^1 result for the conditional damped call price $C_T^a\left(\chi|H_T=h\right)$ in Lemma 7.8 above allows for Fourier inversion via Theorem 3.6 to take place.

Lemma 7.10 *Let the damping parameter be $a>0$. Then*

$$\begin{aligned}&E^Q\left[S_T^{a+1}|H_T\right]<\infty \\ &\text{implies } \widehat{C}_T^a\left(u|H_T=h\right)\in L^1, \text{ for all } h\in H_T.\end{aligned} \tag{7.76}$$

Proof. By equation (7.73)

$$\begin{aligned}\left|\widehat{C}_T^a\left(u|H_T=h\right)\right| &= \left|\frac{e^{-rT}S_0\phi_{Z_T}\left(u-(a+1)\,i|H_T=h\right)}{a^2+a-u^2+i\left(2a+1\right)u}\right| \\ &= e^{-rT}S_0\frac{\left|E^Q\left[e^{iuZ_T}\left(\frac{S_T}{S_0}\right)^{a+1}|H_T=h\right]\right|}{\left|a^2+a-u^2+i\left(2a+1\right)u\right|} \\ &\leq \frac{e^{-rT}}{S_0^a}\frac{E^Q\left[S_T^{a+1}|H_T=h\right]}{\sqrt{\left(a^2+a-u^2\right)^2+i\left(2a+1\right)^2u^2}} \leq \frac{A}{u^2}, \qquad (7.77)\end{aligned}$$

for some finite $A>0$, by Lemma 7.9 above, since $E^Q\left[S_T^{a+1}|H_T\right]<\infty$. Moreover, by Lemma 7.8, $C_T^a\left(\chi|H_T=h\right)\in L^1$, since $E^Q\left[S_T^{a+1}|H_T\right]<\infty$. Hence, by Lemma 3.3 from subsection 3.2.1, $\widehat{C}_T^a\left(u|H_T=h\right)$ is a continuous in u, and by Proposition 3.5 item *5.*, $\left|\widehat{C}_T^a\left(u|H_T=h\right)\right|$ is an even function of u, since $C_T^a\left(\chi|H_T=h\right)$ is purely real. Hence, the inequality (7.77) above implies that $\widehat{C}_T^a\left(u|H_T=h\right)\in L^1$, by the Lebesgue dominated convergence theorem, as required. ■

The main result of this subsection is that if $E^Q\left[S_T^{a+1}|H_T\right]<\infty$, then by Lemma 7.8, $C_T^a\left(\chi|H_T=h\right)$ is L^1, and by Lemma 7.10, $\widehat{C}_T^a\left(u|H_T=h\right)$ is L^1, both for all $h\in H_T$. Hence, by the Fourier inversion theorem, the conditional damped call price is given by

$$C_T^a\left(\chi|H_T=h\right)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-iu\chi}\widehat{C}_T^a\left(u|H_T=h\right)du, \qquad (7.78)$$

for all $h\in H_T$.

Moreover, the undamped conditional call price is simply given by

$$\begin{aligned}C_T\left(\chi|H_T=h\right) &= \frac{e^{-a\chi}}{2\pi}\int_{-\infty}^{+\infty}e^{-iu\chi}\widehat{C}_T^a\left(u|H_T=h\right)du \qquad (7.79)\\ &= \frac{e^{-a\chi}}{\pi}\,\mathrm{Re}\int_0^{\infty}e^{-iu\chi}\widehat{C}_T^a\left(u|H_T=h\right)du, \qquad (7.80)\end{aligned}$$

for all $h\in H_T$,

where equation (7.80) follows from Corollary 3.7 in subsection 3.2.1, since the conditional damped call price is purely real. Hence, for some choice of quadrature weights, the log-moneyness version of the conditional Carr and Madan formula is given by

$$C_T(\chi_k|h) \approx \frac{e^{-a\chi_k}}{\pi} \operatorname{Re} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \widehat{C}_T^a(u_j|h)\, w_j \Delta u, \quad (7.81)$$

for all $h \in H_T$.

In this chapter the quadrature weights w_j in equation (7.81) above are assumed to follow the trapezoidal rule given by $w_1 = w_N = \frac{1}{2}$, with $w_j = 1$, $j = 2, 3, ..., N-1$. Moreover, if H_T is independent of $Z_T = \log\left(\frac{S_T}{S_0}\right)$, then the analysis of the above subsection leads back to the log-moneyness version of the Carr and Madan formula from subsection 5.2.2 above.

7.4.2 The Method of Conditional FFT

Recall from subsection 7.4.1 on the conditional Carr and Madan formula that if $E^Q\left[S_T^{a+1}|H_T\right] < \infty$, then by Lemma 7.8 the conditional damped call price $C_T^a(\chi|H_T = h)$ is L^1, and by Lemma 7.10 the Fourier transform of the conditional damped call price $\widehat{C}_T^a(u|H_T = h)$ is L^1, both for any $h \in H_T$. Hence, by the Fourier inversion theorem, the conditional damped call price is given by

$$C_T^a(\chi|H_T = h) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iu\chi} \widehat{C}_T^a(u|H_T = h)\, du, \quad (7.82)$$

for all $h \in H_T$.

Moreover, the undamped conditional call price is simply given by

$$C_T(\chi|H_T = h) = e^{-a\chi} C_T^a(\chi|H_T = h), \quad (7.83)$$

for all $h \in H_T$.

Let $h^{(1)}, h^{(2)}, ..., h^{(M)}$ be independent realizations of H_T. Then define the damped Monte Carlo averages

$$\overline{C}_T^a(\chi) = \frac{1}{M}\sum_{m=1}^{M} C_T^a\left(\chi|H_T = h^{(m)}\right), \tag{7.84}$$

$$\text{and } \Psi_T^a(u) = \frac{1}{M}\sum_{m=1}^{M} \widehat{C}_T^a\left(u|H_T = h^{(m)}\right). \tag{7.85}$$

The standard conditional Monte Carlo estimator for the damped call price is given by $\overline{C}_T^a(\chi)$ in equation (7.84) above. However, notice that it takes a total of M inverse Fourier transforms to compute each $C_T^a\left(\chi|H_T = h^{(m)}\right)$, $m = 1, 2, ..., M$, in the right hand side of equation (7.84).

Lemma 7.11 (Conditional FFT) *Assume that* $E^Q\left[S_T^{a+1}|H_T\right] < \infty$. *Then the standard conditional Monte Carlo estimator for the damped call price* $\overline{C}_T^a(\chi)$ *in equation* (7.84) *may be written as*

$$\overline{C}_T^a(\chi) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iu\chi}\Psi_T^a(u)\,du, \tag{7.86}$$

where $\Psi_T^a(u)$ *is defined in equation* (7.85). *Moreover,* $\overline{C}_T^a(\chi)$ *in equation* (7.84) *is the inverse Fourier transform of* $\Psi_T^a(u)$.

Proof. Since $E^Q\left[S_T^{a+1}|H_T\right] < \infty$, we have by equation (7.82) above that

$$C_T^a\left(\chi|H_T = h^{(m)}\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iu\chi}\widehat{C}_T^a\left(u|H_T = h^{(m)}\right)du,$$
$$m = 1, 2, ..., M.$$

Thus, by linearity we have that

$$\begin{aligned}
\overline{C}_T^a(\chi) &= \frac{1}{M}\sum_{m=1}^{M} C_T^a\left(\chi|H_T = h^{(m)}\right) \\
&= \frac{1}{M}\sum_{m=1}^{M}\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iu\chi}\widehat{C}_T^a\left(u|H_T = h^{(m)}\right)du \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iu\chi}\frac{1}{M}\sum_{m=1}^{M}\widehat{C}_T^a\left(u|H_T = h^{(m)}\right)du \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iu\chi}\Psi_T^a(u)\,du.
\end{aligned}$$

Moreover, by Lemma 7.8 from subsection 7.4.1, each $C_T^a\left(\chi|H_T=h^{(m)}\right)$ is L^1, $m=1,2,...,M$, since $E^Q\left[S_T^{a+1}|H_T\right]<\infty$. Hence, $\overline{C}_T^a(\chi)$ is L^1. Also, by Lemma 7.10 from subsection 7.4.1, each $\widehat{C}_T^a\left(u|H_T=h^{(m)}\right)$ is L^1, for each $m=1,2,...,M$, since $E^Q\left[S_T^{a+1}|H_T\right]<\infty$. Hence, $\Psi_T^a(u)$ is L^1. Thus, since we have shown above that

$$\overline{C}_T^a(\chi)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-iu\chi}\Psi_T^a(u)\,du,$$

$\overline{C}_T^a(\chi)$ is the inverse Fourier transform of $\Psi_T^a(u)$, as required. ■

We will denote the undamped versions of these two estimators by

$$\overline{C}_T^{(0)}(\chi)=\frac{e^{-a\chi}}{M}\sum\nolimits_{m=1}^{M}C_T^a\left(\chi|H_T=h^{(m)}\right),\text{ and} \tag{7.87}$$

$$\overline{C}_T(\chi)=\frac{e^{-a\chi}}{2\pi}\int_{-\infty}^{+\infty}e^{-iu\chi}\Psi_T^a(u)\,du, \tag{7.88}$$

where $\overline{C}_T^{(0)}(\chi)$ in equation (7.87) is the standard undamped estimator, and $\overline{C}_T(\chi)$ in equation (7.88) is conditional FFT. By equation (7.85) above, the Fourier transform $\Psi_T^a(u)$ in conditional FFT is the Monte Carlo average of the Fourier transforms of the conditional damped call price $\widehat{C}_T^a\left(u|H_T=h^{(m)}\right)$, where by equation (7.73) from subsection 7.4.1, for $m=1,2,...,M$,

$$\widehat{C}_T^a\left(u|H_T=h^{(m)}\right)=\frac{e^{-rT}S_0\phi_{Z_T}\left(u-(a+1)\,i|H_T=h^{(m)}\right)}{a^2+a-u^2+i\,(2a+1)\,u}. \tag{7.89}$$

The generalized conditional CF in the right hand side of equation (7.89)

$$\phi_{Z_T}\left(u-(a+1)\,i|H_T=h^{(m)}\right)=E\left[e^{iuZ_T+(a+1)Z_T}|H_T=h^{(m)}\right], \tag{7.90}$$

will be called the conditional damped log-price CF. In this sense, conditional FFT samples directly from the conditional damped log-price CF, and is faster since it only requires one inverse Fourier transform. But, because the standard estimator samples directly from the conditional damped call prices, it leads directly to standard errors for the estimated call prices. Thus, we use

conditional FFT for speed, and the standard estimator only when standard errors are needed. Similar to equation (7.81) from subsection 7.4.1, the FFT formula for conditional FFT is given by

$$C_T(\chi_k) \approx \frac{e^{-a\chi_k}}{\pi} \operatorname{Re} \sum\nolimits_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} e^{ibu_j} \Psi_T^a(u)\, w_j \Delta u. \tag{7.91}$$

Strike Price	**Cond. FFT**	**Cond. MC**	**Difference**	**MCSE**
85	18.2545395	18.2545395	-4.62E-14	9.25E-02
100	8.5557387	8.5557387	-2.31E-14	5.47E-02
115	2.8546183	2.8546183	7.10E-15	2.09E-02
CPU Time	0.573 seconds	1.324 seconds	**-**	

Table 7.3: Conditional FFT vs. Standard Conditional Monte Carlo.

Table 7.3 above was computed on an Intel 2xE5-2643v3 CPU at 3.4 GHz. The pricing model is the SVSJ Merton, with $T = 0.5$, and $M = 8192$. Otherwise the setting is as in Table 7.9 below. Identical sets of pseudo-random numbers are used for each method. Notice that there is almost no numerical difference between conditional FFT and standard conditional Monte Carlo, but that conditional FFT is more than twice as fast.

7.4.3 Convergence for Conditional FFT

Proposition 7.12 (Convergence) *Assume that* $E^Q\left[S_T^{a+1}|H_T\right] < \infty$ *and that* $E^Q\left[S_T^{a+1}\right] < \infty$. *Let* $\widehat{C}_T^a(u)$ *be the Fourier transform of the damped call price, and* $C_T(\chi)$ *be the unconditional call price. Then*

1. $\Psi_T^a(u) = \frac{1}{M}\sum\nolimits_{m=1}^{M} \widehat{C}_T^a\left(u|H_T = h^{(m)}\right) \underset{a.s.}{\to} \widehat{C}_T^a(u)$.
2. $\overline{C}_T(\chi) = \frac{e^{-a\chi}}{2\pi} \int_{-\infty}^{+\infty} e^{-iu\chi} \Psi_T^a(u)\, du \to C_T(\chi)$.

Proof. To prove *1.*, by equation (7.89) in subsection 7.4.2 above, consider

$$\begin{aligned}
&E\left[\widehat{C}_T^a\left(u|H_T=h^{(m)}\right)\right] \\
&= E\left[\frac{e^{-rT}S_0\phi_{Z_T}\left(u-(a+1)\,i|H_T=h^{(m)}\right)}{a^2+a-u^2+i\,(2a+1)\,u}\right] \\
&= \frac{e^{-rT}S_0E\left[E\left[e^{i(u-(a+1)i)Z_T}|H_T=h\right]\right]}{a^2+a-u^2+i\,(2a+1)\,u} \\
&= \frac{e^{-rT}S_0E\left[e^{i(u-(a+1)i)Z_T}\right]}{a^2+a-u^2+i\,(2a+1)\,u} \\
&= \frac{e^{-rT}S_0\phi_{Z_T}\left(u-(a+1)\,i\right)}{a^2+a-u^2+i\,(2a+1)\,u} \\
&= \widehat{C}_T^a\left(u\right),
\end{aligned}$$

by equation (7.73) of subsection 7.4.1.with H_T independent of Z_T. But, if H_T is independent of Z_T, then the hypothesis $E^Q\left[S_T^{a+1}\right]<\infty$ implies that $C_T^a\left(\chi\right)\in L^1$ by Lemma 7.8 of subsection 7.4.1. Hence, since $\widehat{C}_T^a\left(u\right)$ is the Fourier transform of $C_T^a\left(\chi\right)$,

$$\begin{aligned}
\left|\widehat{C}_T^a\left(u\right)\right| &= \left|\int_{-\infty}^{+\infty}e^{iu\chi}C_T^a\left(\chi\right)d\chi\right| \\
&\leq \int_{-\infty}^{+\infty}\left|C_T^a\left(\chi\right)\right|d\chi<\infty.
\end{aligned}$$

Therefore, by the strong law of large numbers,

$$\Psi_T^a\left(u\right)=\frac{1}{M}\sum\nolimits_{m=1}^{M}\widehat{C}_T^a\left(u|H_T=h^{(m)}\right)\underset{a.s.}{\rightarrow}\widehat{C}_T^a\left(u\right). \tag{7.92}$$

To prove *2.*, again by equation (7.89) in subsection 7.4.2 above,

$$\begin{aligned}
\left|\widehat{C}_T^a\left(u|H_T=h^{(m)}\right)\right| &= \left|\frac{e^{-rT}S_0\phi_{Z_T}\left(u-(a+1)\,i|H_T=h^{(m)}\right)}{a^2+a-u^2+i\,(2a+1)\,u}\right| \\
&= e^{-rT}S_0\frac{\left|E^Q\left[e^{iuZ_T}\left(\frac{S_T}{S_0}\right)^{a+1}|H_T=h^{(m)}\right]\right|}{\left|a^2+a-u^2+i\,(2a+1)\,u\right|} \\
&\leq \frac{e^{-rT}}{S_0^a}\frac{E^Q\left[S_T^{a+1}|H_T=h^{(m)}\right]}{\sqrt{\left(a^2+a-u^2\right)^2+i\,(2a+1)^2\,u^2}}\leq\frac{A}{u^2},
\end{aligned}$$

$m = 1, 2, ..., M$, for the same finite $A > 0$, by Lemma 7.9 of subsection 7.4.1, since $E^Q\left[S_T^{a+1}|H_T\right] < \infty$. Hence,

$$\begin{aligned}\left|\Psi_T^a(u)\right| &= \left|\frac{1}{M}\sum\nolimits_{m=1}^{M}\widehat{C}_T^a\left(u|H_T = h^{(m)}\right)\right| \\ &\leq \frac{1}{M}\sum\nolimits_{m=1}^{M}\left|\widehat{C}_T^a\left(u|H_T = h^{(m)}\right)\right| \leq \frac{A}{u^2}.\end{aligned}$$

Therefore, by the Lebesgue dominated convergence theorem,

$$\begin{aligned}\lim_{M\to\infty}\overline{C}_T(\chi) &= \frac{e^{-a\chi}}{2\pi}\lim_{M\to\infty}\int_{-\infty}^{+\infty}e^{-iu\chi}\Psi_T^a(u)\,du \\ &= \frac{e^{-a\chi}}{2\pi}\int_{-\infty}^{+\infty}e^{-iu\chi}\lim_{M\to\infty}\Psi_T^a(u)\,du \\ &= \frac{e^{-a\chi}}{2\pi}\int_{-\infty}^{+\infty}e^{-iu\chi}\widehat{C}_T^a(u)\,du \\ &= C_T(\chi),\end{aligned}$$

by the Fourier inversion theorem, and undamping, as required. ■

7.4.4 Choosing the Milstein Scheme and QMC Accuracy

All computations in this subsection were done on an Intel 2xE5-2643v3 processor at 3.4GHz. First, in Tables 7.4 and 7.5 below, we show that there is essentially no pricing difference, based on SVSJ model call pricing by conditional FFT, between the drift-implicit Milstein scheme of equation (7.21) and the drift-implicit square-root Euler scheme of equation (7.26), each from subsection 7.2.3 above. Thus, for the remainder of the thesis, we use the simpler drift-implicit Milstein scheme. Secondly, in Tables 7.6 to 7.11 below we test the accuracy of conditional FFT under the drift-implicit Milstein scheme in the TCMX and SVSJ models, which both have a closed form log-price CF, see Carr, Geman, Madan, and Yor (2003), p. 359, and Appendix A of this thesis, respectively. Thus, we compare call options prices obtained from the Carr and Madan (1999) formula to those

computed by conditional FFT. The three scenarios considered are $2\kappa\eta > \omega^2$, $4\kappa\eta > \omega^2 > 2\kappa\eta$, and $\omega^2 > 4\kappa\eta$, each taken at maturities $T = 0.25$, 1.00, and 2.00. Pricing errors for Monte Carlo are compared to pricing errors for QMC. But, first we compare the Milstein and square-root Euler schemes.

Table 7.4: SVSJ: Milstein vs. Sqrt Euler: QMC $M_{qmc} = 4096$: $\mathbf{T = 1.00}$								
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
$2\kappa\eta > \omega^2$	0.05	1.80	0.10	0.45	-0.70	65	0.08	-0.04
$S_0 = 100$	$r = 0.003$		**QMC Pricing Errors**					
Strike	**Exact Price**		**Implicit Milstein**			**Implicit Sqrt Euler**		
85	21.6754622		6.24E-03			7.90E-03		
100	12.8859381		5.85E-03			8.62E-03		
115	6.8111756		2.65E-03			6.10E-03		
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
$4\kappa\eta > \omega^2 > 2\kappa\eta$	0.05	1.60	0.08	0.65	-0.70	65	0.08	-0.04
$S_0 = 100$	$r = 0.003$		**QMC Pricing Errors**					
Strike	**Exact Price**		**Implicit Milstein**			**Implicit Sqrt Euler**		
85	20.7163333		1.90E-02			4.05E-02		
100	11.1910863		1.19E-02			5.17E-02		
115	4.8596344		-3.71E-04			4.58E-02		
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
$\omega^2 > 4\kappa\eta$	0.05	1.50	0.06	0.85	-0.70	65	0.08	-0.04
$S_0 = 100$	$r = 0.003$		**QMC Pricing Errors**					
Strike	**Exact Price**		**Implicit Milstein**			**Implicit Sqrt Euler**		
85	19.7571064		-1.45E-01			-2.86E-01		
100	9.4622219		-2.39E-01			-3.14E-01		
115	3.1454599		-2.18E-01			-6.73E-02		

Table 7.4: SVSJ: Implicit Milstein vs. Square Root Euler: QMC.

Table 7.5: SVSJ: Milstein vs. Sqrt Euler: MC $M_{mc} = 8192$: T = 1.00								
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
$2\kappa\eta > \omega^2$	0.05	1.80	0.10	0.45	-0.70	65	0.08	-0.04
$S_0 = 100$, $r = .003$	**Implicit Milstein**				**Implicit Sqrt Euler**			
Strike	**Pricing Error**				**Pricing Error**			
85	1.7E-02 (1.2E-01)				1.8E-02 (1.2E-01)			
100	1.0E-03 (8.5E-02)				3.9E-03 (8.5E-02)			
115	-7.3E-03 (4.9E-02)				-3.7E-03 (4.9E-02)			
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
$4\kappa\eta > \omega^2 > 2\kappa\eta$	0.05	1.60	0.08	0.65	-0.70	65	0.08	-0.04
$S_0 = 100$, $r = .003$	**Implicit Milstein**				**Implicit Sqrt Euler**			
Strike	**Pricing Error**				**Pricing Error**			
85	9.3E-02 (9.5E-02)				1.2E-01 (9.4E-02)			
100	4.7E-02 (5.5E-02)				8.9E-02 (5.4E-02)			
115	1.9E-02 (2.3E-02)				6.7E-02 (2.3E-02)			
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
$\omega^2 > 4\kappa\eta$	0.05	1.50	0.06	0.85	-0.70	65	0.08	-0.04
$S_0 = 100$, $r = .003$	**Implicit Milstein**				**Implicit Sqrt Euler**			
Strike	**Pricing Error**				**Pricing Error**			
85	-1.2E-01 (7.3E-02)				-3.0E-01 (7.4E-02)			
100	-2.1E-01 (3.5E-02)				-3.2E-01 (3.6E-02)			
115	-2.0E-01 (1.6E-02)				-5.4E-02 (1.9E-02)			

Table 7.5: SVSJ: Implicit Milstein vs. Square Root Euler: Monte Carlo.

Table 7.4 and Table 7.5 above show no significant difference between the drift-implicit Milstein and drift-implicit square-root Euler schemes. Thus, we choose the simpler drift-implicit Milstein scheme. Next, we compare QMC to Monte Carlo for our choice of the drift-implicit Milstein scheme.

TCMX: Implicit Milstein Pricing Accuracy I: $2\kappa\eta > \omega^2$								
Parameters	κ	η	ω	λ	α	β	λ_0	
	2.00	3.70	3.20	1.50	0.20	-0.30	1.00	
T = 0.25	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	30.08736		4.68E-04		3.83E-04		5.31E-04	
85	15.52358		2.35E-03		2.51E-03		4.11E-03	
100	3.955470		5.02E-03		3.70E-03		1.34E-02	
115	0.6210140		2.82E-03		3.13E-03		5.64E-03	
130	0.1169158		9.44E-04		8.58E-04		1.63E-03	
CPU Time			0.258 sec		0.447 sec			
T = 1.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	31.42651		2.89E-03		8.29E-03		1.19E-02	
85	19.37717		4.99E-03		9.42E-03		2.39E-02	
100	10.62375		6.41E-03		8.11E-03		3.20E-02	
115	5.481658		6.02E-03		1.05E-02		2.98E-02	
130	2.839680		5.01E-03		1.26E-02		2.28E-02	
CPU Time			0.422 sec		0.569 sec			
T = 2.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	34.30663		5.48E-03		7.91E-03		2.45E-02	
85	24.24516		7.41E-03		7.16E-03		3.59E-02	
100	16.67959		8.67E-03		6.77E-03		4.25E-02	
115	11.35497		8.80E-03		7.87E-03		4.33E-02	
130	7.754281		8.48E-03		9.94E-03		4.01E-02	
CPU Time			0.732 sec		0.722 sec			

Table 7.6: TCMX Conditional FFT Pricing Case I : $2\kappa\eta > \omega^2$.

TCMX: Implicit Milstein Pricing Accuracy II: $4\kappa\eta > \omega^2 > 2\kappa\eta$								
Parameters	κ	η	ω	λ	α	β	λ_0	
	1.70	3.30	3.90	1.50	0.20	-0.30	1.00	
T = 0.25	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	30.08733		4.56E-04		6.16E-04		6.84E-04	
85	15.48288		1.93E-03		2.39E-03		4.87E-03	
100	3.630266		2.62E-03		1.90E-03		1.65E-02	
115	0.5644410		2.21E-03		2.78E-03		6.65E-03	
130	0.1140778		8.59E-04		1.30E-03		2.03E-03	
CPU Time			0.258 sec		0.445 sec			
T = 1.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	31.23742		2.33E-03		5.31E-03		1.42E-02	
85	18.73204		4.88E-03		6.56E-03		2.90E-02	
100	9.566588		7.36E-03		8.42E-03		4.03E-02	
115	4.633184		6.08E-03		7.57E-03		3.64E-02	
130	2.363877		4.18E-03		7.76E-03		2.72E-02	
CPU Time			0.422 sec		0.573 sec			
T = 2.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	33.66971		6.14E-03		-4.21E-03		3.07E-02	
85	23.06475		8.80E-03		-5.23E-03		4.62E-02	
100	15.17291		1.03E-02		-5.70E-03		5.56E-02	
115	9.889547		1.05E-02		-6.23E-03		5.60E-02	
130	6.551812		9.89E-03		-6.28E-03		5.08E-02	
CPU Time			0.734 sec		0.723 sec			

Table 7.7: TCMX Conditional FFT Pricing Case II : $4\kappa\eta > \omega^2 > 2\kappa\eta$.

TCMX: Implicit Milstein Pricing Accuracy III: $\omega^2 > 4\kappa\eta$								
Parameters	κ	η	ω	λ	α	β	λ_0	
	1.40	1.80	4.50	1.50	0.20	-0.30	1.00	
T = 0.25	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	30.08219		2.18E-04		3.45E-04		8.07E-04	
85	15.39271		-1.34E-03		-1.01E-03		5.24E-03	
100	2.989395		-3.07E-02		-3.48E-02		1.91E-02	
115	0.4413152		-2.17E-03		-1.75E-03		7.11E-03	
130	0.0949362		6.18E-05		4.86E-04		2.31E-03	
CPU Time			0.253 sec		0.443 sec			
T = 1.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	30.80159		-2.43E-02		-1.98E-02		1.36E-02	
85	17.30355		-9.14E-02		-9.55E-02		2.94E-02	
100	6.887874		-2.17E-01		-2.37E-01		4.54E-02	
115	2.740314		-1.24E-01		-1.31E-01		3.73E-02	
130	1.341687		-5.83E-02		-5.53E-02		2.62E-02	
CPU Time			0.421 sec		0.570 sec			
T = 2.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	32.07722		-1.04E-01		-1.05E-01		3.02E-02	
85	19.83681		-2.56E-01		-2.37E-01		4.96E-02	
100	10.70267		-4.08E-01		-3.68E-01		6.40E-02	
115	5.795585		-3.36E-01		-3.08E-01		6.14E-02	
130	3.460687		-2.20E-01		-2.10E-01		5.19E-02	
CPU Time			0.736 sec		0.732 sec			

Table 7.8: TCMX Conditional FFT Pricing Case III : $\omega^2 > 4\kappa\eta$.

SVSJ: Implicit Milstein Pricing Accuracy I: $2\kappa\eta > \omega^2$								
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
	0.05	1.80	0.10	0.45	-0.70	65	0.08	-0.04
T = 0.25	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	30.24372		2.24E-03		-2.06E-02		8.65E-02	
85	16.40626		6.10E-03		-2.47E-03		7.09E-02	
100	5.758612		3.19E-03		1.13E-02		3.74E-02	
115	0.8713932		-2.87E-03		-9.47E-04		7.10E-03	
130	0.0688712		-3.71E-04		-4.83E-04		6.54E-04	
CPU Time			0.374 sec		0.712 sec			
T = 1.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	32.93455		3.02E-03		-7.09E-02		1.54E-01	
85	21.67546		1.57E-03		-1.25E-02		1.21E-01	
100	12.88594		3.91E-04		1.98E-02		8.21E-02	
115	6.811176		-1.37E-03		2.19E-02		4.73E-02	
130	3.190422		-2.21E-03		1.03E-02		2.28E-02	
CPU Time			0.543 sec		0.819 sec			
T = 2.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	36.67352		6.24E-03		-7.78E-02		2.18E-01	
85	27.00539		6.53E-03		-5.86E-02		1.80E-01	
100	19.23431		6.00E-03		-4.21E-02		1.40E-01	
115	13.26743		4.53E-03		-2.93E-02		1.03E-01	
130	8.884514		2.84E-03		-1.93E-02		7.24E-02	
CPU Time			0.885 sec		0.995 sec			

Table 7.9: SVSJ Conditional FFT Pricing Case I : $2\kappa\eta > \omega^2$.

SVSJ: Implicit Milstein Pricing Accuracy II: $4\kappa\eta > \omega^2 > 2\kappa\eta$								
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
	0.05	1.60	0.08	0.65	-0.70	65	0.08	-0.04
T = 0.25	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	30.28552		2.50E-03		1.26E-02		8.11E-02	
85	16.39234		5.96E-03		4.95E-03		6.33E-02	
100	5.344542		7.78E-04		-3.98E-03		2.81E-02	
115	0.5738374		-3.37E-03		-6.62E-04		4.28E-03	
130	0.0471860		-2.22E-04		-2.03E-04		5.97E-04	
CPU Time			0.374 sec		0.722 sec			
T = 1.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	32.61858		5.03E-03		-2.78E-02		1.29E-01	
85	20.71633		2.29E-03		-9.22E-03		9.35E-02	
100	11.19109		-8.11E-04		-3.22E-03		5.42E-02	
115	4.859634		-3.81E-03		-1.01E-02		2.33E-02	
130	1.768320		-3.33E-03		-9.34E-03		1.03E-02	
CPU Time			0.543 sec		0.826 sec			
T = 2.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	35.57758		2.71E-02		-8.62E-02		1.74E-01	
85	25.05308		2.42E-02		-8.25E-02		1.33E-01	
100	16.50783		1.81E-02		-6.11E-02		9.14E-02	
115	10.11918		9.83E-03		-3.28E-02		5.51E-02	
130	5.799368		2.91E-03		-9.45E-03		2.95E-02	
CPU Time			0.883 sec		0.994 sec			

Table 7.10: SVSJ Conditional FFT Pricing Case II : $4\kappa\eta > \omega^2 > 2\kappa\eta$.

SVSJ: Implicit Milstein Pricing Accuracy III: $\omega^2 > 4\kappa\eta$								
Parameters	σ_0^2	κ	η	ω	ρ	λ	α	β
	0.05	1.50	0.06	0.85	-0.70	65	0.08	-0.04
T = 0.25	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	30.3291595		1.03E-02		3.43E-03		7.49E-02	
85	16.3722971		9.83E-03		-3.40E-03		5.56E-02	
100	4.8984780		-1.89E-02		-3.16E-02		2.07E-02	
115	0.4273847		-6.37E-03		-4.29E-03		4.27E-03	
130	0.0435029		-3.78E-04		2.33E-04		7.53E-04	
CPU Time			0.378 sec		0.721 sec			
T = 1.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	32.2852980		-2.41E-02		-8.02E-02		1.06E-01	
85	19.7571064		-1.08E-01		-1.50E-01		7.19E-02	
100	9.4622219		-2.26E-01		-2.59E-01		3.48E-02	
115	3.1454599		-2.16E-01		-2.44E-01		1.61E-02	
130	1.0031177		-7.33E-02		-9.03E-02		1.02E-02	
CPU Time			0.546 sec		0.831 sec			
T = 2.00	$S_0 = 100$		$M_{qmc} = 4096$		$M_{mc} = 8192$		$r = 0.003$	
Strike	**Exact Price**		**QMC Error**		**MC Error**		**MCSE**	
70	34.4842415		-8.72E-02		-1.09E-01		1.37E-01	
85	23.0651351		-2.39E-01		-2.62E-01		9.78E-02	
100	13.6767609		-3.99E-01		-4.22E-01		5.78E-02	
115	6.9985285		-4.63E-01		-4.81E-01		2.77E-02	
130	3.2461703		-3.46E-01		-3.57E-01		1.59E-02	
CPU Time			0.882 sec		0.996 sec			

Table 7.11: SVSJ Conditional FFT Pricing Case III : $\omega^2 > 4\kappa\eta$.

Pricing Error Analysis: Drift-Implicit Milstein Case

In Table 7.6 above we have $2\kappa\eta > \omega^2$ in the TCMX model, and the pricing errors are on order of 10^{-3} for both QMC with 4096 simulations, and Monte Carlo with 8192 simulations. In Table 7.7 above for the TCMX model, where $4\kappa\eta > \omega^2 > 2\kappa\eta$, the pricing errors again are on order of 10^{-3} for both QMC with 4096 simulations, and Monte Carlo with 8192 simulations. Note that in Table 7.7 above the TCMX model parameters are set close to the least squares estimates for the LJI Meixner model from Table 7.15 in subsection 7.5.3 below, where the TCMX model is a special case of the LJI model. We regard this level of pricing accuracy to be suitable for calibration. However, notice in Table 7.8 above for TCMX model that when $\omega^2 > 4\kappa\eta$ there is a serious degradation of the pricing errors for both QMC and Monte Carlo, especially at the two year maturity. Moreover, recall from subsection 7.2.3 above in equation (7.23) that $\omega^2 > 4\kappa\eta$ is the case in which the drift-implicit Milstein scheme must be truncated. But, the calibration results in Table 7.15 of subsection 7.5.3 below show that the case $\omega^2 > 4\kappa\eta$ typically lies far away from the solution.

For the SVSJ Merton model, the QMC pricing results with 4096 simulations, for $2\kappa\eta > \omega^2$ in Table 7.9 and for $4\kappa\eta > \omega^2 > 2\kappa\eta$ in Table 7.10, are both also on order of 10^{-3}, similar to the TCMX model. However, notice in Table 7.9 and Table 7.10 above that the corresponding Monte Carlo pricing results, for the SVSJ Merton model with 8192 simulations, are typically only of order of 10^{-2}. That is, QMC is clearly the better choice for the SVSJ Merton model which has a more complicated conditional characteristic function as shown in subsection 7.3.3. Lastly, for the SVSJ Merton model in Table 7.11 where $\omega^2 > 4\kappa\eta$, the pricing errors are of poor quality, since truncation is required in the drift-implicit Milstein scheme for the CIR process. Overall, QMC with half the number of simulations is best.

7.5 Model Calibration and Model Comparison

In this section we present selected least squares calibration results for the Apple call options data set of this thesis, based on the January 19th, 2011 close. All computations were done on an Intel 2xE5-2643v3 CPU at 3.4 GHz. In subsection 7.5.1 we cover calibration of the TCMX and the SVSJ Merton models, each by both exact FFT and conditional FFT. For comparison to the LJI model leverage effect, in subsection 7.5.2 we calibrate, by exact FFT, both the stationary gamma and stationary inverse-Gaussian versions of the BN-S model, see Barndorff-Nielsen and Shephard (2001), pp. 177-81. Lastly, in subsection 7.5.3 we calibrate the LJI model under Meixner, VG, and Merton jumps by conditional FFT.

7.5.1 Calibration to Exact Benchmarks

Table 7.12 below presents the TCMX model calibration results for the Apple call options data set, using both exact and conditional FFT.

Parameter	κ	η	ω	λ	α	β
Time Changed Meixner (Exact FFT)						
Estimate	0.527	1.654	0.014	1.567	0.284	-1.321
Std. Err.	0.109	0.244	0.011	0.678	0.088	0.127
t-statistic	4.846	6.770	1.318	2.311	3.224	-10.369
Iterations	231	**Time**	18.5 seconds		**RMSE**	0.6249
Time Changed Meixner (Conditional FFT)						
Estimate	0.528	1.653	0.005	1.564	0.284	-1.321
Std. Err.	0.408	0.751	0.160	0.262	0.021	0.181
t-statistic	1.296	2.201	0.033	5.962	13.630	-7.299
Iterations	87	**Time**	22.1 minutes		**RMSE**	0.6243

Table 7.12: TCMX Calibration Results: Exact vs. Conditional FFT.

Recall from equation (7.32) in subsection 7.3.1 above that we assume $\lambda_0 = 1$ in the TCMX model. Moreover, the marginal log-price CF for the stochastic exponential time changed Lévy model for any Lévy process characteristic exponent is given in Carr, Geman, Madan, and Yor (2003), see p. 352 and p. 359. For the TCMX model we apply the Meixner log-price characteristic exponent from equation (7.33) in subsection 7.3.1 above. Notice in Table 7.12 above that the parameter estimates and the root mean squared errors are virtually identical for both conditional and exact FFT. However, the t-statistic for the volatility ω of the jump intensity λ_t is small in the exact FFT case. This causes the standard errors to differ between the two methods, since the intensity process is insignificant. Moreover, this occurs because the TCMX model has no leverage parameter.

Table 7.13 below gives the SVSJ Merton model calibration results for the Apple call options data set, using both exact and conditional FFT. For details of the SVSJ Merton marginal log-price CF see Appendix A.

Parameter	κ	η	ω	ρ	λ	α	β
SVSJ Merton (Exact FFT)							
Estimate	1.805	0.103	0.441	-0.706	64.251	0.082	-0.003
Std. Err.	0.100	0.009	0.039	0.043	22.644	0.009	0.006
t-statistic	17.976	11.528	11.355	-16.596	2.838	9.171	-0.538
Iterations	120	**Time**	12.5 seconds		**RMSE**	0.1960	**-**
SVSJ Merton (Conditional FFT)							
Estimate	1.805	0.104	0.447	-0.699	66.664	0.081	-0.004
Std. Err.	0.089	0.001	0.020	0.021	15.975	0.011	0.005
t-statistic	20.350	96.842	21.995	-32.657	4.173	7.567	-0.831
Iterations	94	**Time**	30.5 minutes		**RMSE**	0.1947	**-**

Table 7.13: SVSJ Calibration Results: Exact vs. Conditional FFT.

In Table 7.13 above we have made the assumption that $\sigma_0^2 = 0.051$ is known for estimation purposes, similar to the assumption that $\lambda_0 = 1$ in the TCMX model above. The estimate of $\sigma_0^2 = 0.051$ is from Table 5.3 in subsection 5.3.2 for exact FFT calibration of the SVSJ model to the Apple options data. Notice in Table 7.13 for the SVSJ model that again the parameter estimates and the root mean squared errors are virtually identical for both conditional and exact FFT. However, in this case the standard errors are also relatively close between the two methods. This is because all of the t-statistics are significant for the SVSJ model, which has a well known leverage parameter ρ. Considering both the TCMX and SVSJ models above, it appears that provided the parameter estimates are significant, calibration by conditional FFT using the drift-implicit Milstein scheme of subsection 7.2.3 above will produce both reliable parameter estimates, and also reasonable asymptotic standard errors.

7.5.2 The BN-S Model with Stationary Gamma Variance

In subsection 2.7.1 we gave technical details of the stationary $\Gamma(\nu, \delta)$ version of the Barndorff-Nielsen and Shephard (2001) stochastic volatility model, to be known simply as the BNS-SG model, given by

$$dY_t = \left(r - \kappa \psi_U(-i\rho; \nu, \delta) - \frac{1}{2}\sigma_t^2 \right) dt + \sigma_t dW_t + \rho dU(\nu, \delta; \kappa t), \quad (7.93)$$

$$d\sigma_t^2 = -\kappa \sigma_t^2 dt + dU(\nu, \delta; \kappa t), \text{ with } \rho < 0. \quad (7.94)$$

The function $\psi_U(u; \nu, \delta)$ is the characteristic exponent of a positive Lévy process $U(\nu, \delta; \kappa t)$ such that the non-Gaussian Ornstein-Uhlenbeck equation (7.94) for the variance has a stationary $\Gamma(\nu, \delta)$ distribution. The other well known version of the BN-S model is identical to the model above, but with Lévy process $U_{\kappa t} = U(\theta, \xi; \kappa t)$, and characteristic exponent $\psi_U(u; \theta, \xi)$, such that the non-Gaussian Ornstein-Uhlenbeck equation (7.94) for the variance

has a stationary inverse-Gaussian distribution, $IG(\theta, \xi)$, with characteristic function given by

$$\phi_{IG}(u) = \exp\left(-\theta\left(\sqrt{-2iu + \xi^2} - \xi\right)\right). \tag{7.95}$$

We refer to this alternative model as the BNS-SIG model. To make a fair comparison to the LJI model, we need to show that the BNS-SG model outperforms the BNS-SIG model for the Apple stock call options data set of this thesis. The log-price CFs for both the BNS-SG and BNS-SIG models are given in Schoutens (2003), see pp. 87-88.

Parameter	σ_0^2	κ	ν	δ	ρ
BN-S Stationary Gamma (BNS-SG)					
Estimate	0.053	1.790	1.975	0.032	-1.987
Std. Err.	0.001	0.102	0.123	0.002	0.092
t-statistic	60.084	17.483	16.045	16.509	-21.511
Iterations	61	**Time**	4.3 sec	**RMSE**	0.2485
Parameter	σ_0^2	κ	θ	ξ	ρ
BN-S Stationary Inverse-Gaussian (BNS-SIG)					
Estimate	0.066	0.271	1.003	0.532	-0.182
Std. Err.	0.001	0.067	0.269	0.105	0.035
t-statistic	89.782	4.012	3.729	5.077	-5.191
Iterations	53	**Time**	4.3 sec	**RMSE**	0.5475

Table 7.14: BNS-SG and BNS-SIG Calibration Results: Exact FFT.

We mention here that in removing the final maturity of 503 days from the Apple stock call options data set, the RMSE for the BNS-SIG model falls to 0.3051, while the RMSE for the BNS-SG model stays roughly the same at 0.2394. This suggests that the BNS-SG model is better than the BNS-SIG model for late maturities. In any case, we have established that

the BNS-SG is the better of the two BN-S models for the entire Apple stock call options data set, and the BNS-SG results in Table 7.14 above are what we will compare to the LJI model results. Notice in Table 7.14 above that the t-statistics for the BNS-SG model are all highly significant, especially for the leverage parameter $\rho < 0$. However, recall from subsection 2.7.1 that we predict that the BNS-SG model will underperform the LJI model because of the following two structural weaknesses that together impede the BNS-SG leverage effect.

1. In equation (7.93) above for the BNS-SG log-price, the Lévy process $U_{\kappa t} = U(\nu, \delta; \kappa t)$ takes on positive values only.

2. In the BNS-SG log-price of equation (7.93) above, W_t is uncorrelated with $U_{\kappa t}$ driving the variance in equation (7.94), so that $\int_0^t \sigma_s dW_s$ is a symmetric diffusion for all values of $U_{\kappa t} > 0$.

As we argued back in subsection 2.7.1, these two effects together mean that leverage in the BNS-SG model is less effective specifically when the Lévy process $U_{\kappa t} = U(\nu, \delta; \kappa t)$ is small.

It may be remarked at this point that for the Apple stock call options data set, the RMSE for the BNS-SG model in Table 7.14 above is 0.2485, and this is noticeably larger than the RMSE of 0.1960 for the SVSJ Merton model in Table 7.13 above under exact FFT.

7.5.3 Calibrating the LJI Model

Because the LJI model does not have a closed form log-price characteristic function, all calibration results in this subsection were necessarily obtained by the method of conditional FFT. We begin by focusing on the LJI Meixner model calibration results for the Apple call options data set. It turns out that similar to subsection 5.3.2 above, where we calibrated all three

SVJ models to the Apple call options data set, the behaviour of the LJI Meixner model is not significantly different from the LJI Merton and LJI Variance Gamma models, under the Apple call options data set of this thesis. Results for all three LJI models together will be given towards the end of this subsection. Table 7.15 below gives the LJI Meixner model calibration results for the Apple call options data set under conditional FFT.

Parameter	γ	κ	η	ω	λ	α	β
LJI Meixner: Conditional FFT							
Estimate	0.208	1.718	3.361	3.913	1.564	0.199	-0.284
Std. Err.	0.008	0.092	0.316	0.441	0.371	0.024	0.089
t-statistic	25.163	18.623	10.648	8.880	4.213	8.291	-3.178
Iterations	152	**Time**	45.4 minutes		**RMSE**	0.1981	**-**

Table 7.15: LJI Meixner Calibration Results: Conditional FFT.

Notice in Table 7.15 above for the LJI Meixner model calibration that all of the t-statistics are highly significant, in particular the t-statistic for the volatility ω of the jump intensity λ_t is highly significant. Recall that in the LJI model the parameter $\gamma > 0$ doubles as the diffusion coefficient in the log-price as well as the leverage parameter. The t-statistic of γ is also highly significant. Also, the RMSE of 0.1981 for the LJI Meixner model in Table 7.15 above is much lower than the RMSE of 0.6243 for the TCMX model from Table 7.12. Thus, the time changed stochastic exponential model of Carr, Geman, Madan, and Yor (2003) can perform well, when a diffusion and a suitable leverage effect are both added to the model.

Recall from subsection 2.7.2 that we predict that the LJI model will outperform the Barndorff-Nielsen and Shephard (2001) model, on the basis of RMSE, because of the following two structural strengths of the LJI model that together make for a superior leverage effect. Furthermore, recall that

from subsection 2.7.2 the LJI model is defined by

$$dY_t = \left(r - \lambda_t \psi_X(-i) - \frac{1}{2}\gamma^2\right) dt + dX(\Lambda_t) - \gamma dW_t, \tag{7.96}$$

$$d\lambda_t = \kappa(\eta - \lambda_t)\, dt + \omega\sqrt{\lambda_t} dW_t, \tag{7.97}$$

$$d\Lambda_t = \lambda_t dt \ , \ \lambda_0 = 1, \text{ with } \gamma > 0. \tag{7.98}$$

1. The Brownian motion leverage term, W_t, in the log-price of the LJI model, see equation (7.96) above, takes positive and negative values.

2. The time changed jumps, $X(\Lambda_t)$, in the log-price of the LJI model, see equation (7.96) above, have inherent skewness governed by the parameter β, where $\beta < 0$ is anticipated.

As we argued back in subsection 2.7.2, since $\gamma > 0$, $-\gamma dW_t$ in equation (7.96) above is always perfectly negatively correlated with dW_t driving the LJI model jump intensity in equation (7.97) above. This is the main basis of the LJI leverage effect. However, noting that $\Delta X(\Lambda_t) \approx X(\lambda_t \Delta t)$, see equation (2.88) of subsection 2.7.2, it can be shown that when the intensity λ_t in the LJI model gets large, and the log-prices falls, the negative skewness of the time changed jump increment dissipates, see Table 2.1 in subsection 2.3.4, so that the contribution from the jumps is more neutral. Accordingly, when λ_t in the LJI model gets small, and the log-price rises, the negative skewness of the time changed jump increment grows, thus the contribution from the jumps is more negatively skewed, and this damps the rise in the log-price. We believe that this leads to a superior leverage effect.

It may be observed at this point that for the Apple stock call options data set, the RMSE for the LJI Meixner model in Table 7.15 above is 0.1981, and this is substantially lower than the RMSE of 0.2485 for the BNS-SG model in Table 7.14 above. The LJI Meixner RMSE of 0.1981, obtained by conditional FFT, is also competitive with the RMSE of 0.1960 for the

SVSJ Merton model in Table 7.13 above, under exact FFT. In Table 7.16 below, we present the LJI Meixner, LJI Variance Gamma, and LJI Merton calibration results, based on the Apple stock call options data set.

Parameter	γ	κ	η	ω	λ	α	β
LJI Meixner: Conditional FFT							
Estimate	0.208	1.718	3.361	3.913	1.564	0.199	-0.284
Std. Err.	0.008	0.092	0.316	0.441	0.371	0.024	0.089
t-statistic	25.163	18.623	10.648	8.880	4.213	8.291	-3.178
Iterations	152	**Time**	45.4 minutes		**RMSE**	0.1981	**-**
LJI Variance Gamma: Conditional FFT							
Estimate	0.209	1.720	3.410	3.991	5.178	0.077	-0.008
Std. Err.	0.008	0.083	0.315	0.372	1.152	0.007	0.003
t-statistic	24.721	20.793	10.834	10.722	4.495	10.492	-2.538
Iterations	66	**Time**	20.6 minutes		**RMSE**	0.1983	**-**
LJI Merton: Conditional FFT							
Estimate	0.210	1.729	3.502	4.157	6.705	0.067	-0.006
Std. Err.	0.010	0.100	0.383	0.467	1.628	0.007	0.003
t-statistic	21.430	17.353	9.135	8.893	4.119	9.883	-1.901
Iterations	102	**Time**	31.5 minutes		**RMSE**	0.1987	**-**

Table 7.16: LJI Meixner, LJI VG, and LJI Merton: Calibration Results.

Table 7.16 above shows that each of the three LJI models outperforms the BNS-SG model of Table 7.14 above on the basis of RMSE. While we believe that is mainly due to a superior leverage effect in the LJI model, some of this difference may be accounted for by the superiority of the continuous CIR process as a model for the jump intensity, as compared to a pure jump non-Gaussian OU process as a model for the variance in the BNS-SG model. Notice in Table 7.16 above that while the Feller condition, $2\kappa\eta > \omega^2$, fails for

all three LJI models, the positivity condition, $4\kappa\eta > \omega^2$, of the drift-implicit Milstein scheme for the CIR process does not fail for any of the three LJI models. Moreover, notice in Table 7.16 above that when leveraged, jump intensity is a wide process. Lastly, in Table 7.16 above, notice that the significance of the skewness parameter $\beta < 0$ is the highest for the LJI Meixner model, but that the LJI VG model converges the most rapidly.

7.6 Conclusion

1. The Leveraged Jump Intensity (LJI) model appears to be a better model for the leverage effect, as compared to the BN-S stochastic volatility model of Barndorff-Nielsen and Shephard (2001). However, some of this difference may be due to the superiority of the CIR process as a latent factor in the LJI model, as compared to the choice of a pure jump non-Gaussian Ornstein-Uhlenbeck process as the latent factor in the BN-S model.

2. In future research we will consider an LJI model with a CIR process for the jump intensity, but also with simultaneous Poisson jumps in the log-price and the jump intensity, where the normal and exponential jump sizes, respectively, are correlated, similar to the "double-jump" model of Duffie, Pan, and Singleton (2000), see pp. 1360-62.

3. The method of conditional FFT, using the log-moneyness version of the conditional Carr and Madan (1999) formula for conditional pricing, and the drift-implicit Milstein scheme of Kahl, Günther, and Rossberg (2008) to simulate the CIR process as a model for the latent factor, is both reasonably fast and sufficiently accurate enough to calibrate a single daily closing book of call options prices under the non-affine LJI model.

4. In future research we will consider an LJI model with a Vasiček (1977) mean-reverting Ornstein-Uhlenbeck process for the jump intensity. In this case, a closed form log-price characteristic function can be found by using the analysis of Carr, Geman, Madan, and Yor (2003), see §6.1. Conditional Monte Carlo will become the benchmark price. The Vasiček LJI model has a time change with some amount of mass on the negative real line, and this causes a small but identifiable bias in pricing. We intend to approximate this bias with a random variable, in order to debias pricing in the Vasiček LJI model.

Chapter 8

The Simulation of Meixner Process Increments

8.1 Introduction

The Meixner process of Schoutens and Teugels (1998), see subsection 2.3.3 above, has been used in this thesis to represent infinite-activity jumps with infinite variation. As the Meixner process is relatively new, the literature on simulating Meixner increments is limited. In this chapter we propose a new method for simulating Meixner increments based on least squares estimation. This new method was used in Chapter 4 above to simulate Meixner daily increments for testing the Bates (2006) AML method.

In the literature there are two main methods of Meixner simulation, one from Grigoletto and Provasi (2009), and one from Madan and Yor (2008). Similar to our method, in Grigoletto and Provasi (2009), acceptance-rejection is used. However, their choice of rejection density is based on an inverse hyperbolic sine transform of the normal distribution, see Johnson, Kotz, and Balakrishnan (2004), p. 34. Moreover, the parameters of their rejection

density are fitted by the method of moments. Note that for the distribution of the Meixner increment, the scale parameter α does not scale time, see Schoutens (2003), p. 62. Thus, the time increment must be scaled into the other parameters, specifically λ. We found that for daily increments, the method of moments as proposed in Grigoletto and Provasi (2009), did not have a stable solution. The method proposed in Madan and Yor (2008) constructs the Meixner process as a time changed Brownian motion. The time change process is simulated by a one-sided stable process with stability index of one half, but with a subset of the small jumps removed by a suitable truncation function. However, this method is difficult to execute successfully. In our trial implementation of the Madan and Yor (2008) method, the sample mean and variance matched the true Meixner mean and variance relatively well, but the skewness and kurtosis did not match at all, and it was not clear how to correct this. Consequently, we constructed a Meixner simulation method of our own, one that is both robust to daily increments, and more transparent to the end user. Our method of simulating Meixner process increments uses acceptance-rejection with the marginal density for the Normal inverse-Gaussian (NIG) process of Barndorff-Nielsen (1997) as the rejection density. We use least squares to estimate the parameters of the rejection density from a grid of points mapped out by the target density under the proposed parameters. All of the computations for this chapter were done on an Intel Xeon 2xE5-2643v3 dual processor at 3.4 GHz with 12 cores. In Section 8.2 we give the technical details of least squares rejection in the context of Meixner-NIG rejection. In Section 8.3 we test the sample moments of Meixner-NIG rejection for large increments by simulation, and compare with the Meixner European options pricing results from Madan and Yor (2008). In Section 8.4 we test the sample moments of Meixner-NIG rejection for daily increments by simulation. Section 8.5 concludes.

8.2 The Method of Least Squares Rejection

For a suitably well chosen rejection density, the method of least squares rejection provides a parametric fit of the rejection density that lies quite close to the target density in the sense that the expected number of simulations from the rejection density in order to simulate from the target density can be surprisingly low. In this section we describe the technique in the context of Meixner-NIG rejection. In subsection 8.2.1 we describe how to simulate NIG process increments. In subsection 8.2.2 we describe the least squares estimation of the rejection density. Then in subsection 8.2.3 we present the Meixner-NIG rejection algorithm.

8.2.1 Simulating NIG Process Increments

The Normal inverse-Gaussian process may be expressed as a time changed Brownian motion, where the time change is an inverse Gaussian (IG) process with a specified combination of NIG process parameters, see Barndorff-Nielsen (1997), §3.1, p. 3. Consequently, we first examine how to simulate from the inverse-Gaussian distribution. The parameterization of the IG distribution associated with the NIG process has a density given by

$$g_1(x;\delta,\gamma) = \frac{\delta}{\sqrt{2\pi}} \exp(\delta\gamma)\, x^{-\frac{3}{2}} \exp\left(-\frac{1}{2}\left(\frac{\delta^2}{x} + \gamma^2 x\right)\right). \tag{8.1}$$

Equation (8.1) is the density of the first passage time to the level δ of a Brownian motion with drift γ, see Barndorff-Nielsen (1997), §3.1, p. 3. However, the IG distribution is typically simulated assuming an alternative parameterization given by

$$g_2(x;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(\frac{-\lambda}{2\mu^2 x}(x-\mu)^2\right), \tag{8.2}$$

see Devroye (1986), §4.3, p. 148. Equation (8.2) is also the parameterization for the inverse-Gaussian distribution used by Matlab R2016a.

Lemma 8.1 *The inverse-Gaussian densities* $g_1(x;\delta,\gamma)$ *and* $g_2(x;\mu,\lambda)$*, from equations* (8.1) *and* (8.2) *respectively, are equivalent when*

$$\mu = \frac{\delta}{\gamma} \text{ and } \lambda = \delta^2. \tag{8.3}$$

Proof.

$$\begin{aligned}
g_2(x;\mu,\lambda) &= \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(\frac{-\lambda}{2\mu^2 x}(x-\mu)^2\right) \\
&= \frac{\sqrt{\lambda}}{\sqrt{2\pi}} x^{-\frac{3}{2}} \exp\left(-\frac{1}{2}\frac{\lambda\left(x^2-2\mu x+\mu^2\right)}{\mu^2 x}\right) \\
&= \frac{\sqrt{\lambda}}{\sqrt{2\pi}} x^{-\frac{3}{2}} \exp\left(-\frac{1}{2}\left(\frac{\lambda}{\mu^2}x - 2\frac{\lambda}{\mu} + \lambda\frac{1}{x}\right)\right) \\
&= \frac{\sqrt{\lambda}}{\sqrt{2\pi}} x^{-\frac{3}{2}} \exp\left(\frac{\lambda}{\mu}\right) \exp\left(-\frac{1}{2}\left(\frac{\lambda}{x} + \frac{\lambda}{\mu^2}x\right)\right) \\
&= \frac{\delta}{\sqrt{2\pi}} x^{-\frac{3}{2}} \exp(\delta\gamma) \exp\left(-\frac{1}{2}\left(\frac{\delta^2}{x} + \gamma^2 x\right)\right) \\
&= g_1(x;\delta,\gamma),
\end{aligned}$$

as required. ■

The following algorithm for simulating from the inverse-Gaussian distribution is taken from Devroye (1986), §4.3, p. 149. It is attributed to Michael, Schucany, and Haas (1976), and assumes the (μ,λ) parameterization as in equation (8.2) above.

Algorithm 8.2 *To simulate* $X \sim IG(\mu,\lambda)$*,*

Generate $Z \sim N(0,1)$*.*

Set $Y = Z^2 \sim \chi^2(1)$*.*

Set $X_1 = \mu + \frac{\mu^2 Y}{2\lambda} - \frac{\mu}{2\lambda}\sqrt{4\mu\lambda Y + \mu^2 Y^2}$*.*

Generate $U \sim Unif(0,1)$*.*

If $U \leq \frac{\mu}{\mu+X_1}$ *then return* $X = X_1$*.*

Else return $X = \frac{\mu^2}{X_1}$*.*

Let X_t be a Normal inverse-Gaussian (NIG) process. For parameter identification purposes, $X_t \sim NIG(a, b, m; t)$ has a characteristic function given by

$$\phi_{X_t}(u) = \exp\left(mt\left(\sqrt{a^2 - b^2} - \sqrt{a^2 - (b + iu)^2}\right)\right), \tag{8.4}$$

see Barndorff-Nielsen (1997), p. 4. The process X_t may be expressed as a time changed Brownian motion with drift,

$$X_t = bIG_t + W_{IG_t}, \tag{8.5}$$

where the marginal density of the time change process IG_t is inverse-Gaussian with parameters

$$\delta = mt \text{ and } \gamma = \sqrt{a^2 - b^2}, \tag{8.6}$$

under the parameterization $g_1(x; \delta, \gamma)$ from equation (8.1) above, see Barndorff-Nielsen (1997), §3.1, p. 3. However, by Lemma 8.1, the time change process IG_t in equation (8.5) above is equivalently inverse-Gaussian with parameters

$$\mu = \frac{mt}{\sqrt{a^2 - b^2}} \text{ and } \lambda = m^2 t^2, \tag{8.7}$$

under the parameterization $g_2(x; \mu, \lambda)$ from equation (8.2) above. Hence, we may use Algorithm 8.2 above, with the parameters specified in equation (8.7), to simulate from the marginal distribution of the requisite inverse-Gaussian process in equation (8.5), specified by equation (8.6). This result leads directly to the following algorithm for simulating the increment of a Normal inverse-Gaussian process.

Algorithm 8.3 *Let X be the increment of an $NIG(a, b, m; t)$ process, and let $\tau = \Delta t$. To simulate X,*

Set $\mu = \frac{m\tau}{\sqrt{a^2 - b^2}}$ and $\lambda = m^2\tau^2$.

By Algorithm 8.2, generate $Y \sim IG(\mu, \lambda)$.

Generate $Z \sim N(0, 1)$.

Return $X = bY + \sqrt{Y}Z$.

8.2.2 The Least Squares Envelope

We propose to simulate the drift-free Meixner process increment by the acceptance-rejection technique with the marginal density of the drift-free NIG process as our rejection density. Thus, let X be the increment of a $Meixner\,(\alpha, \beta, \lambda; t)$ process, and let $\tau = \Delta t$. Recall from equation (2.46) in subsection 2.3.3 above that the density of X is given by

$$f\,(x; \theta_{Meix}, \tau) = \frac{\left(2\cos\left(\frac{\beta}{2}\right)\right)^{2\lambda\tau} e^{\frac{\beta x}{\alpha}}}{2\pi\Gamma\,(2\lambda\tau)} \left|\Gamma\left[\lambda\tau + \frac{ix}{\alpha}\right]\right|^2, \tag{8.8}$$
$$\alpha > 0,\ |\beta| < \pi,\ \lambda > 0.$$

where $\theta_{Meix} = (\alpha, \beta, \lambda)$ is the parameter vector. Note that evaluation of the complex-valued gamma function in equation (8.8), for the Meixner density, has an above average computational expense. This adds to both the setup cost and the execution time for our method. From Schoutens (2003), p. 63, the density of the Meixner increment has semi-heavy tails with asymptotic behaviour given by

$$f\,(x; \theta_{Meix}, \tau) \sim \left\{ \begin{array}{ll} C_-\,|x|^{\rho_-} \exp\,(-\eta_-\,|x|) & \text{as } x \to -\infty \\ C_+\,|x|^{\rho_+} \exp\,(-\eta_+\,|x|) & \text{as } x \to +\infty \end{array} \right\}, \tag{8.9}$$

for some positive constants C_-,C_+, where $\rho_\pm = 2\lambda - 1$, and $\eta_\pm = \frac{(\pi \pm \beta)}{\alpha}$. The density of an $NIG\,(a, b, m; t)$ process increment, where $\tau = \Delta t$, is given by the function

$$g\,(x; \theta_{NIG}, \tau) = \frac{am\tau}{\pi} \exp\left(m\tau\sqrt{a^2 - b^2} + bx\right) \frac{K_1\left(a\sqrt{m^2\tau^2 + x^2}\right)}{\sqrt{m^2\tau^2 + x^2}}, \tag{8.10}$$
$$a > 0,\ |b| < a,\ m > 0,$$

where $\theta_{NIG} = (a, b, m)$ is the parameter vector, see Schoutens (2003), pp. 59-60. The function $K_1\,(\cdot)$ in equation (8.10) is the modified Bessel function

of the second kind of order 1. Numerical evaluation of this particular special function is relatively fast. Also, as may be obtained from Schoutens (2003), p. 60, the density of the NIG increment has semi-heavy tails with asymptotic behaviour given by

$$g\left(x;\theta_{NIG},\tau\right)\sim\left\{\begin{array}{ll} D_{-}\left|x\right|^{-\frac{3}{2}}\exp\left(-\gamma_{-}\left|x\right|\right) & \text{as } x\rightarrow-\infty \\ D_{+}\left|x\right|^{-\frac{3}{2}}\exp\left(-\gamma_{+}\left|x\right|\right) & \text{as } x\rightarrow+\infty \end{array}\right\}, \tag{8.11}$$

for some positive constants D_{-},D_{+}, where $\gamma_{\pm}=a\mp b$. For simplicity let $H_{-}=\frac{C_{-}}{D_{-}}$ and $H_{+}=\frac{C_{+}}{D_{+}}$. Then the asymptotic behaviour of the rejection ratio is given by

$$\frac{f\left(x;\theta_{Meix},\tau\right)}{g\left(x;\theta_{NIG},\tau\right)}\sim\left\{\begin{array}{ll} H_{-}\left|x\right|^{2\lambda+\frac{1}{2}}\exp\left(-\delta_{-}\left|x\right|\right) & \text{as } x\rightarrow-\infty \\ H_{+}\left|x\right|^{2\lambda+\frac{1}{2}}\exp\left(-\delta_{+}\left|x\right|\right) & \text{as } x\rightarrow+\infty \end{array}\right\}, \tag{8.12}$$

where

$$\begin{aligned} \delta_{-} &= \frac{\left(\pi-\beta\right)}{\alpha}-\left(a+b\right), \text{ and} \\ \delta_{+} &= \frac{\left(\pi+\beta\right)}{\alpha}-\left(a-b\right). \end{aligned} \tag{8.13}$$

Thus, evidently the rejection ratio vanishes as $x\rightarrow\pm\infty$ when

$$\begin{aligned} a+b &< \frac{\pi-\beta}{\alpha}, \text{ and} \\ a-b &< \frac{\pi+\beta}{\alpha}. \end{aligned} \tag{8.14}$$

Under these conditions we can guarantee that

$$c=\sup_{x\in\mathbb{R}}\frac{f\left(x;\theta_{Meix},\tau\right)}{g\left(x;\theta_{NIG},\tau\right)}<\infty. \tag{8.15}$$

The main idea behind least squares rejection is the following. Let X be the increment that we wish to simulate. In the case at hand $X\sim f\left(x;\theta_{Meix},\tau\right)$, where $\theta_{Meix}=\left(\alpha,\beta,\lambda\right)$ is known. We propose $g\left(x;\theta_{NIG},\tau\right)$ as the rejection

density, where $\theta_{NIG} = (a, b, m)$ is unknown. For some positive constant $b' > 0$, and some positive integer N, let

$$x_k \in \left[-b', b'\right], \; k = 1, 2, ..., N, \tag{8.16}$$

be a uniform discretization over the real numbers. We recommend that b' be fifteen standard deviations of the target density, and that N be chosen such that $\Delta x \sim 10^{-4}$. Then, for fixed inputs $\theta_{Meix} = (\alpha, \beta, \lambda)$ and $\tau = \Delta t$, evaluate

$$f_k = f\left(x_k; \theta_{Meix}, \tau\right), \; k = 1, 2, ..., N. \tag{8.17}$$

The optimal parameters of the NIG rejection density that guarantee a bounded rejection ratio for all $x \in \mathbb{R}$, and satisfy the NIG parameter constraints, are given by the least squares problem

$$\widehat{\theta}_{NIG} = \underset{\theta_{NIG} \in \Omega}{\arg\min} \sum_{k=1}^{N} \left[f_k - g\left(x_k; \theta_{NIG}, \tau\right)\right]^2 \tag{8.18}$$

$$\begin{aligned} &\text{subject to} \\ &a + b < \frac{\pi - \beta}{\alpha}, \\ &a - b < \frac{\pi + \beta}{\alpha}, \\ &|b| < a. \end{aligned}$$

For fixed $\theta_{Meix} = (\alpha, \beta, \lambda)$ and $\tau = \Delta t$, the method of moments may be used to obtain starting values for $\theta_{NIG} = (a, b, m)$ in equation (8.18).

8.2.3 Meixner-NIG Rejection

There are two items to be computed before the execution of Algorithm 8.4 below for Meixner-NIG rejection. The time to compute these two items comprises the setup cost. First, given $\theta_{Meix} = (\alpha, \beta, \lambda)$ and $\tau = \Delta t$, we need to compute $\widehat{\theta}_{NIG} = \left(\widehat{a}, \widehat{b}, \widehat{m}\right)$ by equations (8.17) and (8.18) from

subsection 8.2.2 above. Then, the rate of rejection c is given by

$$c = \sup_{x \in \mathbb{R}} \frac{f\left(x; \theta_{Meix}, \tau\right)}{g\left(x; \widehat{\theta}_{NIG}, \tau\right)}, \tag{8.19}$$

see McLeish (2005), pp. 99-100. It is suggested in Grigoletto and Provasi (2009) that equation (8.19) above may be evaluated either graphically or numerically. We use a graph to determine the local maxima of the rejection ratio, and then numerical optimization over a floating point representation of the real numbers given suitable starting values for each local maxima. Specifically, we use the Matlab R2016a function *fmincon*. Illustrations will be given in the two examples that follow.

Algorithm 8.4 (Meixner-NIG Rejection) *To simulate from the marginal density $f\left(x; \theta_{Meix}, \tau\right)$ of the Meixner process increment, where $\tau = \Delta t$,*

1. *By Algorithm 8.3 generate $X \sim g\left(x; \widehat{\theta}_{NIG}, \tau\right)$.*
2. *Generate $V \sim Unif\left(0, 1\right)$ independent of X.*
3. *Set $Y = V \times cg\left(X; \widehat{\theta}_{NIG}, \tau\right)$.*
4. *If $Y \leq f\left(X; \theta_{Meix}, \tau\right)$ then return X.*
5. *Otherwise go to step 1.*

8.3 Meixner European Option Price Simulation

8.3.1 Simulated Meixner Moments for Large Increments

For the Meixner parameter values $\theta_{Meix} = (\alpha, \beta, \lambda) = (.25, -.5, 1)$, as in Madan and Yor (2008), p. 44, with $\tau = 0.50$, the NIG parameters estimated by equation (8.18) are given by $\widehat{\theta}_{NIG} = \left(\widehat{a}, \widehat{b}, \widehat{m}\right) = (8.521, -2.01, 0.274)$. The rejection rate c computed by maximizing the rejection ratio on the right hand side of equation (8.19) in subsection 8.2.3 above is $c = 1.02437$. In Figure 8.1 below we illustrate that there are two main local maxima at very

nearly the same level. The local maximum value on the left is 1.02106, and the maximum value on the right is $c = 1.02437$. Moreover, we have observed that if $\beta < 0$, then the local maximum on the right is typically also the global maximum.

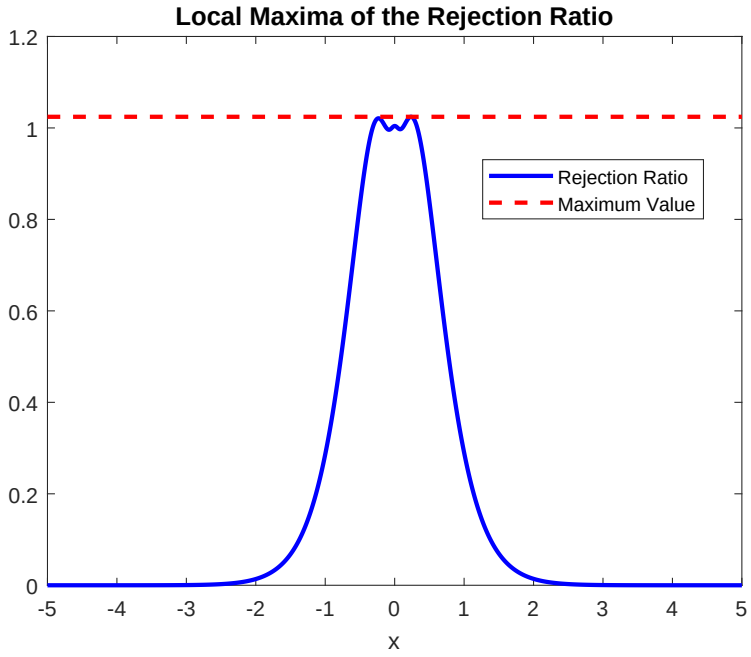

Figure 8.1: Meixner-NIG Rejection Ratio: $\tau = 0.50$, c=1.02437.

The rejection rate $c > 1$ on the left hand side of equation (8.19) above may be interpreted as the expected number of iterations executed in the rejection loop. Thus, a low value of c implies that the rejection method is efficient, see McLeish (2005), p. 100. In Grigoletto and Provasi (2009), §3, a rejection algorithm for simulating the Meixner distribution is proposed, using the method of moments to identify the rejection density. Values of the rejection rate c between 1.0772 and 1.5527 are reported, and c is larger when the skewness and kurtosis of the underlying Meixner distribution are more extreme, see Grigoletto and Provasi (2009), Table 1, p. 64. Because it fits

the shape of the target density directly, the method of least squares rejection is robust to extremities in the underlying skewness and kurtosis. Relative to Grigoletto and Provasi (2009), Table 1, p. 64, the true values of the skewness and kurtosis for the Meixner increment with $(\alpha, \beta, \lambda) = (.25, -.5, 1)$, and $\tau = 0.50$, given in Table 8.1 below are moderately extreme. However, least squares rejection still obtains an efficiency rate of $c = 1.02437$.

	N	**Mean**	**Variance**	**Skewness**	**Kurtosis**
True Value		**-0.03192**	**0.01664**	**-0.49481**	**5.24480**
$Bias$	**5,000**	-0.00012	0.00001	-0.00779	-0.00018
	10,000	-0.00001	0.00003	-0.00187	-0.02317
	20,000	-0.00006	0.00002	-0.00077	-0.01494
$RMSE$	**5,000**	0.00177	0.00043	0.09280	0.42993
	10,000	0.00133	0.00033	0.06238	0.31527
	20,000	0.00083	0.00024	0.04497	0.21278
$\sqrt{MSE + Bias^2}$	**5,000**	0.00177	0.00043	0.09247	0.42993
	10,000	0.00133	0.00033	0.06236	0.31442
	20,000	0.00083	0.00024	0.04496	0.21225
Total Simulation Time			6.13 minutes		

Table 8.1: Simulations of 100 Sample Meixner Moments: $\tau = 0.50$.

Using a batch size of 100, Table 8.1 above assesses the bias and variability for the first four sample moments of a Meixner random sample simulated by the Meixner-NIG rejection technique at three sample sizes. For samples of size 5,000, both the biases and the RMSEs reported in Table 8.1 are modestly reasonable for each sample moment. But, $\sqrt{MSE + Bias^2}$ is twice as low for samples of size 20,000 relative to samples of size 5,000, suggesting that a Monte Carlo sample size of 20,000 will be effective.

8.3.2 Comparison to the Madan and Yor Results

In Table 8.2 below we compare the simulated Meixner options prices from Madan and Yor (2008), Table 1, p. 45, to our Meixner options prices as obtained by the Meixner-NIG least squares rejection method.

Strike	τ	**Madan & Yor**	**Meixner FFT**	**Our Prices**	**MCSE**
80	**0.25**	0.1174	0.1266	0.1244	0.0086
90	**0.25**	0.6022	0.6179	0.6320	0.0195
100	**0.25**	3.3690	3.4127	3.4195	0.0400
110	**0.25**	0.5934	0.6543	0.6609	0.0216
120	**0.25**	0.1293	0.1587	0.1548	0.0103
80	**0.5**	0.3150	0.3284	0.3148	0.0135
90	**0.5**	1.2362	1.2868	1.2919	0.0291
100	**0.5**	5.1135	5.2081	5.2888	0.0575
110	**0.5**	1.5889	1.6870	1.6665	0.0362
120	**0.5**	0.4656	0.5282	0.5177	0.0211
80	**0.75**	0.5593	0.5714	0.5778	0.0192
90	**0.75**	1.8617	1.9000	1.8990	0.0364
100	**0.75**	6.5256	6.6170	6.6352	0.0708
110	**0.75**	2.7066	2.7717	2.7794	0.0502
120	**0.75**	1.0410	1.0710	1.0591	0.0330
80	**1**	0.8117	0.8323	0.8174	0.0236
90	**1**	2.3644	2.4458	2.4545	0.0424
100	**1**	7.7168	7.8220	7.8262	0.0841
110	**1**	3.7458	3.8095	3.7933	0.0630
120	**1**	1.6819	1.7152	1.7103	0.0452
Average CPU Time: Our Prices				3.76 seconds	

Table 8.2: Simulated Meixner European Options Prices.

Table 8.2 above is based on Madan and Yor (2008), Table 1, p. 45. The options are European style, the risk-free rate is 3%, the dividend yield is 1%, and the spot price is 100. For strikes below the spot price, the option is a put, and for strikes at or above the spot price, the option is a call. The model is a risk-neutral pure jump Meixner model. The FFT prices were taken from Madan and Yor (2008), Table 1, p. 45. For our prices we used a sample size of 20,000. This lead to adequate Monte Carlo standard errors, and all of our prices lie within a 95% confidence interval based on the FFT price. We used the Meixner-NIG least squares rejection technique described in Section 8.2 above to construct sample prices. Notice that our prices are more accurate than Madan and Yor (2008), Table 1, p. 45. The average time in Table 8.2 is net of the setup cost, computed on an Intel 2xE5-2643v3 at 3.4GHz with 12 cores in parallel.

8.4 Simulating Daily Meixner Increments

We have shown in Section 8.3 above that the method of Meixner-NIG least squares rejection works well for simulating large Meixner increments. In this section we show that the method is also quite robust for simulating small Meixner increments. This is the purpose that the method was intended to serve. We simulate the sample moments of a random sample of daily Meixner increments. For $\theta_{Meix} = (\alpha, \beta, \lambda) = (.166, .277, 5.23)$, from the Apple stock (1991-2011) SVJ model estimates in Table 4.2, with $\tau = \frac{1}{252}$, the NIG daily increment parameters estimated by equation (8.18) are $\widehat{\theta}_{NIG} = \left(\widehat{a}, \widehat{b}, \widehat{m}\right) = (9.203, 1.666, 0.870)$. The rejection rate c computed from equation (8.19) in subsection 8.2.3 above is $c = 1.023251$. In Figure 8.2 below, again there are two local maxima at very nearly the same level. This time, $\beta > 0$, the global maximum value is on the left at $c = 1.02351$, and the local maximum on the right occurs at 1.02325.

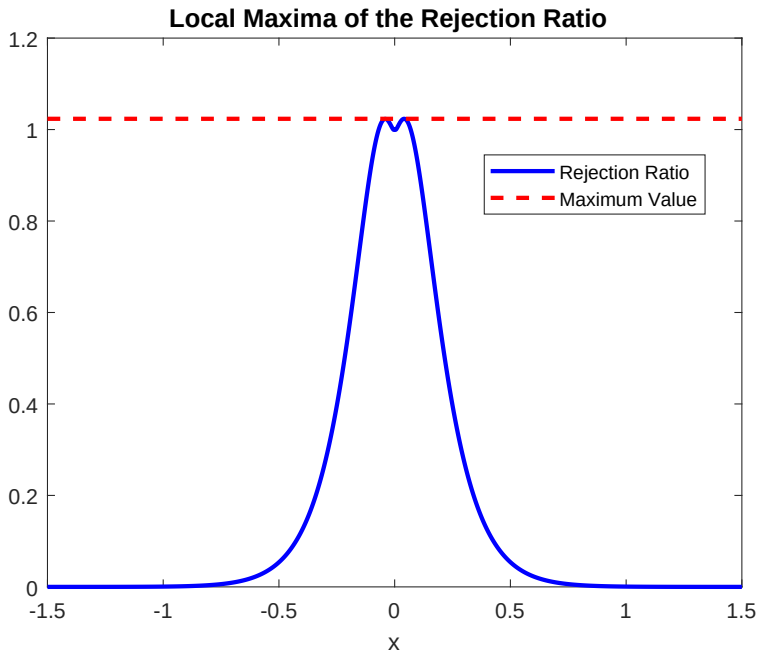

Figure 8.2: Meixner-NIG Rejection Ratio: $\tau = 1/252$, c = 1.02351.

Daily jump increments in stock prices exhibit leptokurtosis. However, least squares rejection is capable of closely fitting the NIG rejection density to the Meixner daily increment density. The rejection rate of $c = 1.02351$ is low and indicates very high efficiency for the proposed rejection scheme. The extreme kurtosis is evident from the true values of the moments in Table 8.3 and Table 8.4 below, for Meixner and VG daily increments, respectively. The Meixner parameters are $(\alpha, \beta, \lambda) = (.166, .277, 5.23)$ in Table 8.3, and for VG the parameters are $(\alpha, \beta, \lambda) = (.068, .008, 13.5)$ in Table 8.4, both from the SVJ model estimates for Apple stock (1991-2011) from Table 4.2. Below, Table 8.3 assesses the bias and variability for sample moments of Meixner daily increments simulated by the Meixner-NIG rejection technique. Variance Gamma (VG) in Table 8.4 is included for comparison.

	N	Mean	Variance	Skewness	Kurtosis
True Value		**0.000480**	**0.000292**	**1.35530**	**53.0200**
$Bias$	**5,000**	-0.000016	-0.000002	-0.05484	-3.51590
	10,000	0.000005	0.000000	-0.00996	-2.03670
	20,000	0.000003	0.000000	-0.00305	-0.40429
$RMSE$	**5,000**	0.000267	0.000026	1.18290	16.2890
	10,000	0.000172	0.000018	0.92164	13.5420
	20,000	0.000123	0.000014	0.63687	10.2630
$\sqrt{MSE + Bias^2}$	**5,000**	0.000267	0.000026	1.18160	15.9050
	10,000	0.000172	0.000018	0.92158	13.3880
	20,000	0.000123	0.000014	0.63686	10.2550
Total Simulation Time			5.87 minutes		

Table 8.3: Simulations of 100 Sample Meixner Moments: $\tau = 1/252$.

	N	Mean	Variance	Skewness	Kurtosis
True Value		**0.000429**	**0.000251**	**1.50750**	**60.5190**
$Bias$	**5,000**	-0.000020	0.000002	0.04644	-2.23940
	10,000	-0.000011	0.000000	-0.02535	-1.49680
	20,000	0.000002	0.000000	0.00318	-0.43669
$RMSE$	**5,000**	0.000227	0.000026	1.15930	15.1520
	10,000	0.000171	0.000018	1.05930	13.1980
	20,000	0.000117	0.000013	0.80543	11.6690
$\sqrt{MSE + Bias^2}$	**5,000**	0.000226	0.000026	1.15840	14.9850
	10,000	0.000170	0.000018	1.05900	13.1130
	20,000	0.000117	0.000013	0.80543	11.6610
Total Simulation Time			1.04 seconds		

Table 8.4: Simulations of 100 Sample V.G. Moments: $\tau = 1/252$.

The parameters underlying both Table 8.3 and Table 8.4, respectively above, are estimated from the daily log-returns of Apple stock (1991-2011). Thus, VG simulation in Table 8.4 makes a useful comparison to our Meixner simulation. Recall from equation (2.27) in subsection 2.3.2, that the VG process is simply a time changed Brownian motion, where the time change is a gamma process. This means that the increment of the VG process can be simulated from one gamma r.v. and one independent normal r.v., see Appendix E for further details. Thus, very little simulation error is expected to occur in the simulation of VG increments. But, notice that Table 8.3 above for simulated Meixner increments, and Table 8.4 above for simulated VG increments, have highly similar corresponding entries. This suggests that the method of Meixner-NIG least squares rejection performs as well as a simple time changed Brownian motion simulator.

Nonetheless, there are potential problems suggested by the results in Table 8.3 and Table 8.4. The biases for all moments in both tables are adequate at each sample size. However, the variability for each of the sample skewness and the sample kurtosis is a bit high in both tables for samples of size 5,000. The fact that this occurs in both Table 8.3 and Table 8.4 suggests that the problem may be inherent with infinite-activity jumps for small increments such as daily.

In Chapter 4 we chose to estimate the SVJ and SVSJ models from samples of twenty years of daily log-returns. This corresponds roughly to samples of size 5,000. We made this choice for consistency with other literature, for example Eraker, Johannes, and Polson (2003), and Li, Wells, and Yu (2008). The levels of bias in both Table 8.3 and Table 8.4 above, for samples of size 5,000, suggest that our choice was reasonable. But, a larger sample size could lead to a slight improvement in the variability of the higher moments, according to Table 8.3 and Table 8.4 above.

8.5 Conclusion

1. The method of Meixner-NIG least squares rejection has been shown to be a highly efficient method of simulating Meixner increments both large and small. The rejection density is based on a well-known time changed Brownian motion, the Normal inverse-Gaussian process of Barndorff-Nielsen (1997). Thus, it is not surprising that we found our method of simulation to perform as well as the direct simulation of another well-known time changed Brownian motion, the Huang and Wu (2004) version of the Variance Gamma process. The only real drawback of our method is the lack of computational speed in the complex-valued gamma function contained in the Meixner density.

2. Under the present setup with both the rejection and target densities known in closed form, our other finding is that for both large and small Meixner increments, the simulated higher moments have excess variability with standard sample sizes of 5,000. We also found this to be true for the Huang and Wu (2004) type Variance Gamma daily increments. Thus, when simulating the daily increments of infinite-activity jumps, it could be of benefit to consider a larger sample size such as 10,000. Based on the example from Bates (2006) with 11,076 daily log-returns, see subsection 4.4.5 above, we believe that both the SVJ and SVSJ models can handle 10,000 data points. However, in the interests of maintaining speed in AML estimation, we also believe that 5,000 data points, as used in this thesis, will suffice.

Part IV

The Conclusion of the Thesis

This thesis is about the estimation of models for the log stock price with a continuous component, stochastic volatility, jumps, and stochastic jump intensity. Particular emphasis has been placed on Fourier based estimation techniques, and consequently the fast Fourier transform (FFT) has been employed in all estimation methods. We have treated models from the affine Markov class, and we have also introduced a new non-affine model. Each of these two streams leads to a main contribution, each with supporting contributions. After an outline of these contributions, we will close with a list of five future research topics.

Contributions to Affine Markov Model Estimation

Our main contribution to the estimation of the affine Markov models is:

- the Parametric Minimum Entropy Martingale Measure (PMEMM) of Chapter 6.

Many supporting contributions were made in the development of the PMEMM, and these came largely in our extensions to the theory of AML from Bates (2006) for the Heston, SVJ, and SVSJ models:

- proof in Section 4.3 that L^1 integrability holds for each of the three Fourier transforms of AML, and implementation of AML by FFT,
- proof in Appendix B and Appendix C that our joint affine coefficients are continuous on the principal branch,
- proof in Appendix D that we may differentiate the primary Fourier transform of AML twice under the integral with respect to $v \in (u, v)$,
- AML estimation results showing that certain banking and technology stocks, and evidently also the S&P 500 index, may be characterized by infinite-activity jumps, from Section 4.6

Other affine Markov model contributions include:

- the least squares rejection algorithm for simulating Meixner increments with an NIG rejection density from Chapter 8,
- proof in Section 3.3 that the marginal log-price characteristic functions are each L^1 for the Heston, SVJ, and SVSJ models, as required for the PMEMM in Chapter 6,
- the trapezoidal rule as an easy solution to the problem of negative call prices deep out-of-the-money in the Carr and Madan (1999) formula, from Chapter 5.

Contributions to Non-Affine Model Estimation

Our main contribution to non-affine log stock price models is:

- the Leveraged Jump Intensity (LJI) model of Chapter 7.

Supporting contributions for the LJI model are:

- a generalization of the Carr and Madan (1999) formula to a particular conditional Monte Carlo and quasi-Monte Carlo pricing technique that we refer to as conditional FFT, from Section 7.4,
- least squares calibration evidence in Section 7.5 that the LJI model provides a superior model for the leverage effect as compared to the Barndorff-Nielsen and Shephard (2001) stochastic volatility model.

Future Research Topics

1. The future of AML is with high-frequency data. We have already further investigated the AML estimation of S&P 500 log-returns at

the 78-minute data frequency with annually based parameters. To increase the data frequency to five minutes, we plan to switch to daily based parameters, as are commonly used in the MCMC literature, see for example Eraker, Johannes, and Polson (2003). Recall that for daily data on an annual basis, $\Delta t = \frac{1}{252}$. But, with five-minute data on a daily basis, $\Delta t = \frac{1}{78} > \frac{1}{252}$. Hence, it can be shown in this case that oscillation in the integrands of the three Fourier transforms of the AML method will decay at a faster rate, thereby improving accuracy. We will consider three month samples of approximately 5,000 data points at five-minute frequency. Thus, it will be worth updating the estimate of the P-measure each day, along with a daily re-estimate of the option implied Q-measure. Each newly estimated P-measure and Q-measure will form the basis a fresh daily PMEMM estimate.

2. So far we have shown that the PMEMM provides a good forecast of late maturity at-the-money discounted call option payoffs, and this is consistent with Chernov and Ghysels (2000). However, we may be able to show more than this. In future research, by numerically solving the boundary value problem for the Heston MEMM call price, as proposed in He and Zhu (2016), given $\theta_P \in \Omega_P$ for the Heston model, we can price Heston MEMM call options. These can be compared to the corresponding Heston PMEMM call prices, given $\theta_P \in \Omega_P$. If these sets of call prices are similar, it will be suggested that the PMEMM is a more easily computable approximation to the MEMM.

3. The method of conditional FFT is a highly flexible pricing technique that only requires the conditional log-price characteristic function given the latent factor to be known in closed form. Thus, in future research we will extend the LJI model to include simultaneous Poisson jumps in the jump intensity and the log-price, with correlated exponential and

normal jump sizes, respectively, similar to the "double-jump" model of Duffie, Pan, and Singleton (2000). The stochastic jump intensity will be simulated as a CIR process augmented with Poisson jumps, and exponential jump sizes. Then, given the jump intensity, and given its jumps, the conditional log-price CF will be easily known in closed form. This should lower the implied volatilities of out-of-the-money call options, similar to Duffie, Pan, and Singleton (2000).

4. We developed the LJI model expressly so that the CIR process could be used to model the jump intensity. But, this leads to a non-affine model. Consider instead, the same LJI model with a Brownian motion and time changed jumps in the log-price, but with the jump intensity modeled by a Vasiček (1977) process. By using the analysis of Carr, Geman, Madan, and Yor (2003), see §6.1, a closed form log-price CF for the Vasiček LJI model can be found. However, this model has a time change with some mass on the negative real line, resulting in biased prices. In future research, we intend to quantify this bias using the method of conditional FFT for benchmark prices, and develop a random variable to model the bias.

5. Bates (2012) considers AML for estimating infinite-activity Lévy process models for the log-price, time changed by the CIR process. This poses a challenge in that a time changed infinite-activity Lévy process is a hidden Markov model, see Bates (2012), p. 230, and is not necessarily Markov. In this thesis we developed the gamma approximation of AML in subsection 4.2.3, based on Bates (2006), see p. 917, under the assumption that the log-price model is a Markov process. It remains an open problem to re-work this analysis, so as to allow for hidden Markov models.

Appendix A

Characteristic Functions: SVJ and SVSJ Models

Recall from subsection 2.6.1 that under the P-measure, the SVJ model with compensated drift is given by

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \psi_X\left(-i\right)\right)dt + \sigma_t dW_t^{(S)} + dX_t, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right] = \rho dt. \end{aligned} \tag{A.1}$$

Hence, by Remark 2.33 from subsection 2.6.1, the SVJ model under the Q-measure is simply $(A.1)$ above with μ_0 replaced by the risk-free rate r. Moreover, it was shown in subsection 2.6.2 that the SVSJ model under the P-measure with compensated drift is given by

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2 - \psi_X\left(-i\right)\sigma_t^2\right)dt + \sigma_t dW_t^{(S)} + dX\left(V_t\right), \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right] = \rho dt, \\ dV_t &= \sigma_t^2 dt. \end{aligned} \tag{A.2}$$

Similarly, by Remark 2.37 from subsection 2.6.2, the SVSJ model under the Q-measure is given by $(A.2)$ with μ_0 replaced by the risk-free rate r.

By Bates (2006), pp. 953-955, the SVJ and SVSJ models are both affine, with an exponential-affine C.F. for the terminal joint process $Z_T = \left(Y_T, \sigma_T^2\right)$ given an initial point $Z_{t_0} = \left(Y_{t_0}, \sigma_{t_0}^2\right)$, written as

$$\phi_{Z_T|Z_{t_0}}(u,v) = \exp\left[iuY_{t_0} + \underline{C}(u,v;\tau) + \underline{D}(u,v;\tau)\,\sigma_{t_0}^2\right], \tag{A.3}$$

in conditional form where Y_t is the log-price, σ_t^2 is the latent factor, and $\tau = T - t_0 > 0$ is the gap time. The functional forms

$$\underline{C} = \underline{C}(u,v;\tau), \text{ and } \underline{D} = \underline{D}(u,v;\tau),$$

are referred to as the affine coefficients. In this thesis we further denote the affine coefficients of the corresponding conditional log-price characteristic function by

$$C = C(u;\tau) = \underline{C}(u,0;\tau), \text{ and } D = D(u;\tau) = \underline{D}(u,0;\tau).$$

Bates (2006), pp. 953-954, nests the SVJ and SVSJ models into one model, and derives the affine coefficients of the joint characteristic function as a solution to the related Kolmogorov backward equation with a suitable boundary condition, similar to Heston (1993), pp. 340-342. Recall that is was shown in subsection 2.6.2 that a time changed compound Poisson Merton process is equivalent to a Poisson stochastic integral where the integrator is a non-homogeneous Poisson process with affine stochastic jump intensity $\lambda_t = \lambda\sigma_t^2$, consistent with Bates (2006), p. 953.

This appendix presents joint affine coefficients $\underline{C}(u,v;\tau)$ and $\underline{D}(u,v;\tau)$, for the SVJ and SVSJ models assuming $\tau > 0$, that are shown under the condition $\rho < 0$ from Appendix B, and a key result from Appendix C.2, to be principal branch continuous for all $(u,v) \in \mathbb{R} \times (-\delta, \delta)$, for some $\delta > 0$. This is based on Lord and Kahl (2010), §2-3, where the problem is studied for the log-price C.F. of the Heston model, with the initial point known. Theorem A.1 of Appendix A.3 proves that our affine coefficients are equivalent to those given in Bates (2006), p. 954.

A.1 Specifications in Common with Bates (2006)

The main auxiliary variables and joint coefficient format presented in this introductory part of Appendix A, are the same across both our specification of the affine coefficients in Appendix A.2, and the Bates (2006) specification presented in Appendix A.3. However, note that Bates (2006) only pursued Merton (1976) jumps. We will provide for the other two jump types from Section 2.3 as well.

We define the compensated characteristic exponent of the jumps by

$$\widehat{\psi}_X(u) = \psi_X(u) - iu\psi_X(-i), \tag{A.4}$$

and for convenience we also define the same for Brownian motion,

$$\widehat{\psi}_W(u) = -\frac{1}{2}\left(u^2 + iu\right). \tag{A.5}$$

Recall from subsection 2.3.1 that for Merton jumps, as in Bates (2006)

$$\begin{aligned}\widehat{\psi}_{MJ}(u) = \lambda&\left(\exp\left(i\beta u - \frac{1}{2}\alpha^2u^2\right) - 1\right) \\ &- iu\lambda\left(\exp\left(\beta + \frac{1}{2}\alpha^2\right) - 1\right),\end{aligned} \tag{A.6}$$

while from subsection 2.3.2 for Huang and Wu type Variance Gamma jumps

$$\begin{aligned}\widehat{\psi}_{VG}(u) = -\lambda&\log\left(1 - i\beta u + \frac{1}{2}\alpha^2u^2\right) \\ &+ iu\lambda\log\left(1 - \beta - \frac{1}{2}\alpha^2\right),\end{aligned} \tag{A.7}$$

and from subsection 2.3.3 for Meixner jumps

$$\begin{aligned}\widehat{\psi}_{MX}(u) = 2\lambda&\log\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cosh\left(\frac{1}{2}\left(\alpha u - i\beta\right)\right)}\right) \\ &- iu2\lambda\log\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cos\left(\frac{\alpha+\beta}{2}\right)}\right).\end{aligned} \tag{A.8}$$

The primary auxiliary variable that distinguishes between the SVJ case and the SVSJ case is

$$\psi(u) = \left\{ \begin{array}{c} \psi^{J0}(u) \\ \psi^{J1}(u) \end{array} \right\} = \left\{ \begin{array}{cc} \widehat{\psi}_W(u) & \text{for SVJ} \\ \widehat{\psi}_W(u) + \widehat{\psi}_{MJ}(u) & \text{for SVSJ} \end{array} \right\}. \tag{A.9}$$

The style of specification for $\psi(u)$ in $(A.9)$ is taken from Huang and Wu (2004). We now define the two primary auxiliary variables,

$$\kappa_G(u) = \kappa - iu\rho\omega, \text{ and} \tag{A.10}$$

$$\gamma(u) = \left\{ \begin{array}{c} \gamma^{J0}(u) \\ \gamma^{J1}(u) \end{array} \right\} = \sqrt{\kappa_G(u)^2 - 2\omega^2\psi(u)}. \tag{A.11}$$

As above, for functions and variables that depend on $\gamma(u)$, or more precisely $\psi(u)$, a superscript $J0$ will indicate the SVJ case, similarly $J1$ will indicate the SVSJ case, and either a superscript ψ or no superscript will indicate both. A common format for both our specification and that of Bates (2006), p. 954, is given by

$$\begin{aligned} \underline{C} &= \underline{C}(u, v; \tau) \\ &= C(u; \tau) - \frac{2\kappa\eta}{\omega^2} \log\left(1 - K^{\psi}(u; \tau)\, iv\right), \end{aligned} \tag{A.12}$$

$$\begin{aligned} \underline{D} &= \underline{D}(u, v; \tau) \\ &= D(u; \tau) + \frac{R^{\psi}(u; \tau)\, iv}{1 - K^{\psi}(u; \tau)\, iv}, \end{aligned} \tag{A.13}$$

$$C(u; \tau) = \left\{ \begin{array}{c} C^{J0}(u; \tau) \\ C^{J1}(u; \tau) \end{array} \right\} = \left\{ \begin{array}{c} C^{SV}(u; \tau) \\ C^{J1}(u; \tau) \end{array} \right\} + \mathbf{1}_{SVJ} \widehat{\psi}_X(u)\, \tau, \tag{A.14}$$

$$\text{and} \quad D(u; \tau) = D^{\psi}(u; \tau). \tag{A.15}$$

Note that in equation $(A.14)$, the term $C^{SV}(u)$ is simply the C coefficient for the Heston model.

The differences between our specification and that of Bates (2006) are found in the definitions of the marginal coefficients, $C(u)$, and $D(u)$, and auxiliary variables $K(u)$, and $R(u)$.

A.2 Affine Coefficients used in this Thesis

If the sole aim of this thesis were to estimate the SVJ and SVSJ models from daily log-returns by AML with $\tau = \frac{1}{252}$, then we would not need the alternative set of affine coefficients presented here. However, the relative entropy method from Chapter 6, and the calibration in Chapter 5, both invert the log-price characteristic function at potentially large maturity dates τ. To accommodate this, continuity on the principal branch is desired, see Lord and Kahl (2010), p. 676, concerning the Heston model. Thus, we offer a unified set of affine coefficients to cover every aspect of this thesis.

By generalizing the arguments from Lord and Kahl (2010), §2-3, we prove in Appendix B and Appendix C that our affine coefficients are principal branch continuous. Other known cases of this phenomenon include Bakshi, Cao, and Chen (1997), p. 2046, and in particular Duffie, Pan, and Singleton (2000), p. 1360. Note that the affine coefficients derived in Heston (1993), pp. 340-342, are not principal branch continuous despite being algebraically equivalent to a version that is, see Lord and Kahl (2010), p. 674-675.

Recall that the general specifications from Appendix A.1 are

$$\underline{C}(u,v;\tau) = C(u;\tau) - \frac{2\kappa\eta}{\omega^2}\log\left(1 - K^{\psi}(u;\tau)\, iv\right), \tag{A.16}$$

$$\underline{D}(u,v;\tau) = D(u;\tau) + \frac{R^{\psi}(u;\tau)\, iv}{1 - K^{\psi}(u;\tau)\, iv}, \tag{A.17}$$

$$C(u;\tau) = \left\{ \begin{array}{c} C^{SV}(u;\tau) \\ C^{J1}(u;\tau) \end{array} \right\} + \mathbf{1}_{SVJ}\widehat{\psi}_X(u)\,\tau, \tag{A.18}$$

$$\text{and } D(u;\tau) = D^{\psi}(u;\tau). \tag{A.19}$$

We now give $C(u)$, $D(u)$, $K(u)$, and $R(u)$ defining the proposed principal branch version of the affine coefficients for the SVJ and SVSJ models.

We begin by defining two additional auxiliary variables

$$A(u) = A^{\psi}(u) = \frac{\kappa_G(u) - \gamma(u)}{\kappa_G(u) + \gamma(u)}, \text{ and} \tag{A.20}$$

$$H(u;\tau) = H^{\psi}(u;\tau) = (\kappa_G(u) + \gamma(u))\left(1 - A(u)\, e^{-\gamma(u)\tau}\right). \tag{A.21}$$

The SVJ and SVSJ joint affine coefficients are now fulfilled simultaneously with marginal coefficients and auxiliary variables defined by

$$\begin{aligned} C(u;\tau) = {} & iu\mu_0\tau + \frac{\kappa\eta}{\omega^2}(\kappa_G(u) - \gamma(u))\tau \\ & - \frac{2\kappa\eta}{\omega^2}\log\left(\frac{A(u)\, e^{-\gamma(u)\tau} - 1}{A(u) - 1}\right) + \mathbf{1}_{SVJ}\widehat{\psi}_X(u)\,\tau, \end{aligned} \tag{A.22}$$

$$D(u;\tau) = \frac{\kappa_G(u) - \gamma(u)}{\omega^2}\left(\frac{1 - e^{-\gamma(u)\tau}}{1 - A(u)\, e^{-\gamma(u)\tau}}\right), \tag{A.23}$$

$$K(u;\tau) = \frac{\omega^2\left(1 - e^{-\gamma(u)\tau}\right)}{H(u;\tau)}, \text{ and} \tag{A.24}$$

$$R(u;\tau) = 4e^{-\gamma(u)\tau}\left(\frac{\gamma(u)}{H(u;\tau)}\right)^2. \tag{A.25}$$

A.3 Affine Coefficients from Bates (2006)

We assume the same general specifications for the joint coefficients as applied above in equations $(A.16)$ to $(A.19)$ from Appendix A.2. However, in this case we define the additional auxiliary variable given by

$$B(u;\tau) = B^{\psi}(u;\tau) = \frac{e^{\gamma(u)\tau} + 1}{e^{\gamma(u)\tau} - 1}. \tag{A.26}$$

It may be deduced from Bates (2006), p. 954, that the marginal coefficients and auxiliary variables $C(u)$, $D(u)$, $K(u)$, and $R(u)$ fulfilling the

definition of the joint affine coefficients for both the SVJ and SVSJ models simultaneously are given by

$$C(u;\tau) = iu\mu_0\tau + \frac{\kappa\eta}{\omega^2}(\kappa_G(u) + \gamma(u))\tau \tag{A.27}$$
$$- \frac{2\kappa\eta}{\omega^2}\log\left(1 + \frac{(\kappa_G(u) + \gamma(u))\left(e^{\gamma(u)\tau} - 1\right)}{2\gamma(u)}\right) + \mathbf{1}_{SVJ}\widehat{\psi}_X(u)\tau,$$

$$D(u;\tau) = \frac{2\psi(u)}{\gamma(u)B(u;\tau) + \kappa_G(u)}, \tag{A.28}$$

$$K(u;\tau) = \frac{\omega^2}{\gamma(u)B(u;\tau) + \kappa_G(u)}, \text{ and} \tag{A.29}$$

$$R(u;\tau) = \frac{B(u;\tau)^2 - 1}{\left(B(u;\tau) + \frac{\kappa_G(u)}{\gamma(u)}\right)^2}. \tag{A.30}$$

As we prove in Appendix B, it is the change that we make in this thesis to $C(u)$ that guarantees principal branch continuity. Appendix B also shows that the changes that we make to $D(u)$, $K(u)$, and $R(u)$ improve stability, since on the principal branch $\operatorname{Re}\gamma(u) > 0$ for all $u \in \mathbb{R}$, whereby $e^{-\gamma(u)\tau}$ decays as $u \to \pm\infty$. The property from Bates (2006), p. 954, that $\underline{C}(u,v;\tau)$ and $\underline{D}(u,v;\tau)$ are both easily twice differentiable with respect v, has also been preserved.

Theorem A.1 (Equivalence of Affine Coefficients) *Equations* (*A*.22) *through* (*A*.25) *of Appendix A.2 are equivalent to equations* (*A*.27) *through* (*A*.30) *of Appendix A.3.*

Proof. From equation (*A*.27),

$$\left(C_{Bates} - iu\mu_0\tau - \mathbf{1}_{SVJ}\widehat{\psi}_X\tau\right)\frac{\omega^2}{\kappa\eta}$$
$$= (\kappa_G + \gamma)\tau - 2\log\left(1 + \frac{(\kappa_G + \gamma)(e^{\gamma\tau} - 1)}{2\gamma}\right)$$
$$= (\kappa_G - \gamma)\tau - 2\log e^{-\gamma\tau} - 2\log\left(1 + \frac{(\kappa_G + \gamma)(e^{\gamma\tau} - 1)}{2\gamma}\right)$$

$$
\begin{aligned}
&= (\kappa_G - \gamma)\,\tau - 2\log\left(\frac{2\gamma e^{-\gamma\tau} + (\kappa_G + \gamma)\,(1 - e^{-\gamma\tau})}{2\gamma}\right) \\
&= (\kappa_G - \gamma)\,\tau - 2\log\left(\frac{\kappa_G + \gamma}{2\gamma} - \frac{\kappa_G - \gamma}{2\gamma}e^{-\gamma\tau}\right) \\
&= (\kappa_G - \gamma)\,\tau - 2\log\left(\left(\frac{\kappa_G + \gamma}{\kappa_G + \gamma} - \frac{\kappa_G - \gamma}{\kappa_G + \gamma}e^{-\gamma\tau}\right)\frac{\kappa_G + \gamma}{2\gamma}\right) \\
&= (\kappa_G - \gamma)\,\tau - 2\log\left(\frac{Ae^{-\gamma\tau} - 1}{A - 1}\right), \\
&\text{since } A - 1 = \frac{-2\gamma}{\kappa_G + \gamma}, \text{ with } A \text{ from } (A.20)\,, \\
&= \left(C - iu\mu_0\tau - \mathbf{1}_{SVJ}\widehat{\psi}_X\tau\right)\frac{\omega^2}{\kappa\eta}, \text{ from equation } (A.22)\,.
\end{aligned}
$$

To see that the D coefficients are equivalent, first note that by squaring both sides of equation $(A.11)$ we obtain

$$
2\psi = \frac{\kappa_G^2 - \gamma^2}{\omega^2} = \frac{(\kappa_G + \gamma)\,(\kappa_G - \gamma)}{\omega^2}. \tag{A.31}
$$

Thus, from equation $(A.28)$,

$$
\begin{aligned}
&D_{Bates} = \frac{2\psi}{\gamma B + \kappa_G} \\
&= \frac{(\kappa_G + \gamma)\,(\kappa_G - \gamma)}{\omega^2}\left(\frac{1 - e^{-\gamma\tau}}{\gamma\,(1 + e^{-\gamma\tau}) + \kappa_G\,(1 - e^{-\gamma\tau})}\right), \\
&\text{where } B = \frac{e^{\gamma\tau} + 1}{e^{\gamma\tau} - 1} \text{from equation } (A.26)\,, \\
&= \frac{(\kappa_G + \gamma)\,(\kappa_G - \gamma)}{\omega^2}\left(\frac{1 - e^{-\gamma\tau}}{(\kappa_G + \gamma) - (\kappa_G - \gamma)\,e^{-\gamma\tau}}\right), \\
&= \frac{\kappa_G - \gamma}{\omega^2}\left(\frac{1 - e^{-\gamma\tau}}{1 - Ae^{-\gamma\tau}}\right), \\
&\text{since } A = \frac{\kappa_G - \gamma}{\kappa_G + \gamma} \text{ from } (A.20)\,, \\
&= D, \text{ from equation } (A.23)\,.
\end{aligned}
$$

The auxiliary variable K is similar to D. From equation $(A.30)$,

$$
\begin{aligned}
K_{Bates} &= \frac{\omega^2}{\gamma B + \kappa_G} \\
&= \frac{\omega^2 \left(1 - e^{-\gamma\tau}\right)}{\gamma\left(1 + e^{-\gamma\tau}\right) + \kappa_G\left(1 - e^{-\gamma\tau}\right)}, \\
&\text{where } B = \frac{e^{\gamma\tau} + 1}{e^{\gamma\tau} - 1} \text{from equation } (A.26), \\
&= \frac{\omega^2 \left(1 - e^{-\gamma\tau}\right)}{\left(\kappa_G + \gamma\right) - \left(\kappa_G - \gamma\right) e^{-\gamma\tau}}, \\
&= \frac{\omega^2 \left(1 - e^{-\gamma\tau}\right)}{\left(\kappa_G + \gamma\right)\left(1 - Ae^{-\gamma\tau}\right)}, \\
&\text{since } A = \frac{\kappa_G - \gamma}{\kappa_G + \gamma} \text{ from } (A.20), \\
&= \frac{\omega^2 \left(1 - e^{-\gamma\tau}\right)}{H} = K, \text{ from equation } (A.24), \\
&\text{since } H = \left(\kappa_G + \gamma\right)\left(1 - Ae^{-\gamma\tau}\right) \text{ from } (A.21).
\end{aligned}
$$

For R, first note that in the numerator of equation $(A.30)$

$$
\begin{aligned}
B^2 - 1 &= \left(\frac{e^{\gamma\tau} + 1}{e^{\gamma\tau} - 1}\right)^2 - 1 \\
&= \left(\frac{1 + e^{-\gamma\tau}}{1 - e^{-\gamma\tau}}\right)^2 - \left(\frac{1 - e^{-\gamma\tau}}{1 - e^{-\gamma\tau}}\right)^2 \\
&= \frac{\left(1 + e^{-\gamma\tau}\right)^2 - \left(1 - e^{-\gamma\tau}\right)^2}{\left(1 - e^{-\gamma\tau}\right)^2} \\
&= \frac{\left(1 + e^{-\gamma\tau} + 1 - e^{-\gamma\tau}\right)\left(1 + e^{-\gamma\tau} - \left(1 - e^{-\gamma\tau}\right)\right)}{\left(1 - e^{-\gamma\tau}\right)^2}, \\
&\text{with the above by difference of squares in the numerator,} \\
&= \frac{2\left(2e^{-\gamma\tau}\right)}{\left(1 - e^{-\gamma\tau}\right)^2} = \frac{4e^{-\gamma\tau}}{\left(1 - e^{-\gamma\tau}\right)^2}. \qquad \text{(A.32)}
\end{aligned}
$$

Using $(A.32)$, we obtain from $(A.30)$

$$
R_{Bates} = \frac{B^2 - 1}{\left(B + \frac{\kappa_G}{\gamma}\right)^2}
$$

$$
\begin{aligned}
&= \frac{4e^{-\gamma\tau}}{\left(B + \frac{\kappa_G}{\gamma}\right)^2 \left(1 - e^{-\gamma\tau}\right)^2} \\
&= \frac{4e^{-\gamma\tau}\gamma^2}{\left(\gamma\left(1 + e^{-\gamma\tau}\right) + \kappa_G\left(1 - e^{-\gamma\tau}\right)\right)^2}, \\
&\text{since } \; B = \frac{e^{\gamma\tau} + 1}{e^{\gamma\tau} - 1} = \frac{1 + e^{-\gamma\tau}}{1 - e^{-\gamma\tau}}, \\
&= \frac{4e^{-\gamma\tau}\gamma^2}{\left(\gamma\left(1 + e^{-\gamma\tau}\right) + \kappa_G\left(1 - e^{-\gamma\tau}\right)\right)^2}, \\
&= \frac{4e^{-\gamma\tau}\gamma^2}{\left(\left(\kappa_G + \gamma\right) + \left(\kappa_G - \gamma\right)e^{-\gamma\tau}\right)^2}, \\
&= \frac{4e^{-\gamma\tau}\gamma^2}{\left(\left(\kappa_G + \gamma\right)\left(1 - Ae^{-\gamma\tau}\right)\right)^2}, \\
&\text{since } A = \frac{\kappa_G - \gamma}{\kappa_G + \gamma} \text{ from } \; (A.20), \\
&= \frac{4e^{-\gamma\tau}\gamma^2}{H^2}, \text{ since } H = \left(\kappa_G + \gamma\right)\left(1 - Ae^{-\gamma\tau}\right) \text{ from } \; (A.21), \\
&= 4e^{-\gamma\tau}\left(\frac{\gamma}{H}\right)^2 = R, \text{ from equation } \; (A.25).
\end{aligned}
$$

Thus, the affine coefficients and auxiliary variables used in this thesis are equivalent to Bates (2006). ■

Remark A.2 *For the log-price coefficients $C(u)$ and $D(u)$, the parameter dependency under the SVJ and SVSJ models is summarized as follows.*

- *SVJ:* $C(u) = C(u; \mu_0, \kappa, \eta, \omega, \rho, \lambda, \alpha, \beta)$ *and* $D(u) = D(u; \kappa, \omega, \rho)$
- *SVSJ:* $C(u) = C(u; \mu_0, \kappa, \eta, \omega, \rho, \lambda, \alpha, \beta)$ *and* $D(u) = D(u; \kappa, \omega, \rho, \lambda, \alpha, \beta)$

These results follow from the definitions in Appendix A.1, and either Appendix A.2 or this Appendix A.3. Note in particular that $D(u)$ does not depend on either the drift parameter μ_0 or the long run variance parameter η in either model, and that in the SVSJ model $D(u)$ depends on the jump parameters λ,α, and β, whereas in the SVJ model $D(u)$ has no dependence on the jump parameters.

Appendix B

Continuity: SVJ and SVSJ Models

Recall from Appendix A that for the SVJ and SVSJ models the joint affine coefficients, $\underline{C}(u, v; \tau)$ and $\underline{D}(u, v; \tau)$ in equations $(A.16)$ and $(A.17)$, are determined by the marginal log-price coefficients, $C(u)$ and $D(u)$, as well as the auxiliary variables, $K(u)$ and $R(u)$, of the joint C.F. In this appendix we show that the versions of the main variables $C(u)$, $D(u)$, $K(u)$ and $R(u)$ in equations $(A.22)$ to $(A.25)$ respectively from Appendix A.2 are, assuming $\tau > 0$, continuous for all $u \in \mathbb{R}$. In particular, Appendix B.2 shows that $C(u)$ is continuous on the principal branch of the logarithm in equation $(A.22)$, given the sole condition that $\rho < 0$, as introduced in Appendix B.1. The task that we have in this appendix is similar to the treatment of the Heston model in Lord and Kahl (2010), §2-3. However, there is at least one important difference. We consider the parameter u to be purely real, whereas Lord and Kahl (2010) considers $u = a + ib$ to be complex. Consequently, though our argument follows Lord and Kahl (2010), it is simpler. Moreover, since the jumps are independent of the

Heston process in the SVJ model, the SVJ extension of Lord and Kahl (2010), §2-3, is trivial. However, the extension of this analysis to the SVSJ model with Merton jumps, to the best of our knowledge, has not been seen before in the previous literature.

The result from Appendix B.2 that for $\tau > 0$, and given $\rho < 0$, the log-price coefficient $C(u)$ is continuous on the principal branch for both the SVJ and SVSJ models, is used in Chapter 3 to prove that the marginal log-price C.F. is L^1. This particular L^1 result is used primarily in Chapter 6 on relative entropy. The continuity of $C(u)$, given $\rho < 0$, is used again in Chapter 4 on the Bates (2006) AML method to prove that the joint conditional C.F. is L^1 in the u-variable.

We will continue to use the convention from Appendix A that a superscript $J0$ will indicate the SVJ case, a superscript $J1$ will indicate the SVSJ case, and either a superscript ψ or no superscript will indicate both. A superscript SV will continue to denote the Heston model.

B.1 Continuity of Auxiliary Variables

In this appendix we treat the continuity of all auxiliary variables. From Appendix A, this list includes $\widehat{\psi}_X(u)$, $\psi(u)$, $\kappa_G(u)$, $\gamma(u)$, and $A(u)$ from equations $(A.4)$, $(A.9)$, $(A.10)$, $(A.11)$, and $(A.20)$ respectively, as well as $H(u)$, $K(u)$, and $R(u)$ respectively from equations $(A.21)$, $(A.24)$, and $(A.25)$. The first group is used primarily to prove continuity of the log-price coefficients in Appendix B.2. The second group is used to prove continuity for the joint coefficients in Appendix C.3, further requiring a result from Appendix C.2 proved by asymptotics. The auxiliary variables $\gamma(u)$ and $A(u)$ have additional properties to be treated in this appendix. The main result is Lemma B.11 proving that for both the SVJ and SVSJ models, $\rho < 0$ implies $|A(u)| < 1$, for all $u \in \mathbb{R}$. First, a few obvious points.

Lemma B.1 *If $\psi_X(-i) < \infty$ then for any pure jump Lévy process X_t*

$$\widehat{\psi}_X(u) = \psi_X(u) - iu\psi_X(-i),$$

is continuous for all $u \in \mathbb{R}$.

Proof. By hypothesis $\psi_X(-i)$ is a finite constant. Then by Theorem 2.5

$$e^{t\widehat{\psi}_X(u)} = \phi_{X_t}(u)\, e^{-iut\psi_X(-i)}.$$

But, $\phi_{X_t}(u)$ is a characteristic function, so it is continuous for all $u \in \mathbb{R}$ by Proposition 3.11 *3*. Moreover, the complex exponential is continuous. ■

Corollary B.2 *The auxiliary variable*

$$\psi(u) = \left\{ \begin{array}{c} \psi^{J0}(u) \\ \psi^{J1}(u) \end{array} \right\} = \left\{ \begin{array}{c} \widehat{\psi}_W(u) \quad \textit{for SVJ} \\ \widehat{\psi}_W(u) + \widehat{\psi}_{MJ}(u) \quad \textit{for SVSJ} \end{array} \right\}, \tag{B.1}$$
$$\textit{where } \widehat{\psi}_W(u) = -\frac{1}{2}\left(u^2 + iu\right),$$

is clearly continuous for all $u \in \mathbb{R}$ in both the SVJ and SVSJ cases.

Furthermore, it is clear from equation $(A.11)$ that for all $u \in \mathbb{R}$,

$$\kappa_G(u) = \kappa - iu\rho\omega$$

is continuous, with strictly positive real part when u is purely real.

The complications begin here. From equation $(A.11)$, the auxiliary variable $\gamma(u)$ is defined by

$$\gamma(u) = \left\{ \begin{array}{c} \gamma^{J0}(u) \\ \gamma^{J1}(u) \end{array} \right\} = \left\{ \begin{array}{c} \sqrt{\kappa_G(u)^2 - 2\omega^2\psi^{J0}(u)} \quad \text{for SVJ} \\ \sqrt{\kappa_G(u)^2 - 2\omega^2\psi^{J1}(u)} \quad \text{for SVSJ} \end{array} \right\}.$$

Note further that $\gamma^{J0}(u) = \gamma^{SV}(u)$. Define

$$z(u) = \gamma(u)^2 = \kappa_G(u)^2 - 2\omega^2\psi(u). \tag{B.2}$$

Furthermore, define the real and imaginary parts of z by

$$x(u) = \operatorname{Re} z(u), \tag{B.3}$$

$$\text{and} \quad y(u) = \operatorname{Im} z(u). \tag{B.4}$$

In polar co-ordinates $\gamma(u)$ is defined by

$$\gamma(u) = \pm\sqrt{|z(u)|}\exp\left(i\frac{\theta}{2}\right), \tag{B.5}$$

$$\text{where } \tan(\theta) = \frac{y(u)}{x(u)}. \tag{B.6}$$

To uniquely satisfy equation $(B.6)$, we will always choose the unique value $\theta_0 \in (-\pi, \pi]$ denoting the principal value of θ. That is, we will use the principal branch. However, to obtain a unique representation of $\gamma(u)$ in equation $(B.5)$, we must choose between the positive and negative roots arbitrarily. In this thesis we assume the positive root in equation $(B.5)$. This is the same assumption made by Lord and Kahl (2010), p. 675.

For future reference, and for the part of the next lemma corresponding to the SVSJ model, we will need the real and imaginary parts of $\widehat{\psi}_{MJ}(u)$. From equation $(A.6)$ in Appendix A we obtain

$$\operatorname{Re}\widehat{\psi}_{MJ}(u) = \lambda\left(e^{-\frac{1}{2}\alpha^2u^2}\cos(\beta u) - 1\right), \text{ and} \tag{B.7}$$

$$\operatorname{Im}\widehat{\psi}_{MJ}(u) = \lambda\left(e^{-\frac{1}{2}\alpha^2u^2}\sin(\beta u) - u\left(e^{\beta+\frac{1}{2}\alpha^2} - 1\right)\right). \tag{B.8}$$

Lemma B.3 *Let u be purely real, and for the SVJ and SVSJ models let*

$$z(u) = \gamma(u)^2 = \kappa_G(u)^2 - 2\omega^2\psi(u).$$

Then for all $u \in \mathbb{R}$

$$x(u) = \operatorname{Re} z(u) > 0. \tag{B.9}$$

Proof. Since u is purely real, for the SVJ model we have

$$\begin{aligned} z^{J0}(u) &= \gamma^{J0}(u)^2 = \kappa_G(u)^2 - 2\omega^2\psi^{J0}(u) \\ &= (\kappa - iu\rho\omega)^2 + \omega^2\left(u^2 + iu\right) \\ &= \kappa^2 + \left(1-\rho^2\right)u^2\omega^2 + i\left(\omega^2 - 2\kappa\omega\rho\right)u \\ &= x^{J0}(u) + iy^{J0}(u). \end{aligned} \tag{B.10}$$

Thus, for all $u \in \mathbb{R}$, $x^{J0}(u) > 0$, given $\kappa > 0$, $\omega > 0$, and $\rho \in [-1, 1]$.

Again, since u is purely real, for the SVSJ model we have

$$\begin{aligned} z^{J1}(u) &= \gamma^{J1}(u)^2 = \kappa_G(u)^2 - 2\omega^2\psi^{J1}(u) \\ &= (\kappa - iu\rho\omega)^2 + \omega^2\left(u^2 + iu\right) - 2\omega^2\widehat{\psi}_{MJ}(u) \\ &= \kappa^2 + \left(1-\rho^2\right)u^2\omega^2 - 2\omega^2\operatorname{Re}\widehat{\psi}_{MJ}(u) \\ &+ i\left[\left(\omega^2 - 2\kappa\omega\rho\right)u - 2\omega^2\operatorname{Im}\widehat{\psi}_{MJ}(u)\right] \\ &= x^{J1}(u) + iy^{J1}(u). \end{aligned} \tag{B.11}$$

Hence, by equation $(B.7)$

$$\begin{aligned} x^{J1}(u) &= \kappa^2 + \left(1-\rho^2\right)u^2\omega^2 - 2\omega^2\operatorname{Re}\widehat{\psi}_{MJ}(u) \\ &= x^{J0}(u) + 2\omega^2\lambda\left(1 - e^{-\frac{1}{2}\alpha^2u^2}\cos(\beta u)\right) \end{aligned}$$

But, $0 \le 1 - e^{-\frac{1}{2}\alpha^2u^2}\cos(\beta u) \le 1$, for all $u \in \mathbb{R}$. Hence, since $\lambda > 0$, the previous part implies $x^{J1}(u) > 0$. ■

Proposition B.4 *Let $u \in \mathbb{R}$, and let $\gamma(u)$ be evaluated on the principal branch. Then, for both the SVJ and SVSJ models, for all $u \in \mathbb{R}$*

1. *$\gamma(u)$ is continuous*
2. *$\operatorname{Re}\gamma(u) > 0$, and*
3. *$\operatorname{Im}\gamma(u)$ has the same sign as $y(u) = \operatorname{Im} z(u) = \operatorname{Im}\gamma(u)^2$.*

Proof. Assuming the positive root, equations $(B.5)$ and $(B.6)$ state that

$$\gamma(u) = \sqrt{|z(u)|}\exp\left(i\frac{\theta(u)}{2}\right), \text{ where } \tan(\theta(u)) = \frac{y(u)}{x(u)}. \tag{B.12}$$

Since $\kappa_G(u)$ is continuous, and by Corollary B.2, $\psi(u)$ is continuous for both the SVJ and SVSJ models, we have that

$$z(u) = \gamma(u)^2 = \kappa_G(u)^2 - 2\omega^2\psi(u),$$

is also continuous. Thus, $x(u) = \operatorname{Re} z(u)$, and $y(u) = \operatorname{Im} z(u)$ are both continuous. But, by Lemma B.3, $x(u) > 0$, when $u \in \mathbb{R}$. Therefore, assuming the principal branch, $\theta(u)$ has the unique solution

$$\theta_0(u) = \arctan\left(\frac{y(u)}{x(u)}\right) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \tag{B.13}$$

with the sign of $\theta_0(u)$ matching the sign of $y(u)$ for each $u \in \mathbb{R}$. Then, since equation $(B.13)$ indicates that $\theta_0(u)$ is continuous, we have that

$$\gamma(u) = \sqrt{|z(u)|}\exp\left(i\frac{\theta_0(u)}{2}\right) \tag{B.14}$$

is continuous on the principal branch. This proves *1.* Moreover,

$$\operatorname{Re}\gamma(u) = \sqrt{|z(u)|}\cos\left(\frac{\theta_0(u)}{2}\right) > 0, \text{ since } \theta_0(u) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

when $u \in \mathbb{R}$, and this proves *2.* Furthermore, on the principal branch,

$$\operatorname{Im}\gamma(u) = \sqrt{|z(u)|}\sin\left(\frac{\theta_0(u)}{2}\right), \text{ with } \theta_0(u) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right). \tag{B.15}$$

However, when $u \in \mathbb{R}$, $x(u) > 0$ by Lemma B.3. Hence, $\theta_0(u)$ has the same sign as $y(u)$ by equation $(B.13)$ where

$$y(u) = \operatorname{Im} z(u) = \operatorname{Im}\gamma(u)^2.$$

But, then $\operatorname{Im}\gamma(u)$ has the same sign as $\theta_0(u)$ by equation $(B.15)$. Thus, the result *3.* follows. ■

The next few results concerning the auxiliary variable,

$$A(u) = \frac{\kappa_G(u) - \gamma(u)}{\kappa_G(u) + \gamma(u)}, \tag{B.16}$$

form the backbone of the Lord and Kahl (2010) case for principal branch continuity of the affine coefficient $C(u)$. As mentioned, these results follow more easily when u is purely real. Moreover, we extend these results to the SVSJ case. For evaluation of $\gamma(u)$, the principal branch will henceforth be assumed for convenience and simplicity.

Lemma B.5 *Let u be purely real. Then $A(u)$ in equation $(B.16)$ is continuous, and for all $u \in \mathbb{R}$, $A(u) \neq 1$.*

Proof. When $u \in \mathbb{R}$, $\operatorname{Re}\kappa_G(u) = \kappa > 0$, and, by Proposition B.4 *2.*, $\operatorname{Re}\gamma(u) > 0$. Hence, the denominator of $A(u) \neq 0$. Then, since $\kappa_G(u)$ is continuous, and by Proposition B.4 *1.*, $\gamma(u)$ is continuous, clearly $A(u)$ is continuous. Moreover, since $\kappa_G(u)$ is never zero if $u \in \mathbb{R}$, $A(u) = 1$ if and only if $\gamma(u) = 0$. But, when $u \in \mathbb{R}$, $\operatorname{Re}\gamma(u) > 0$. Therefore, $A(u) \neq 1$. ■

Lemma B.6 *For the auxiliary variable $A(u)$ in equation $(B.16)$,*

$$|A(u)| < 1 \quad \textit{if and only if}$$
$$\operatorname{Re}\kappa_G(u)\operatorname{Re}\gamma(u) + \operatorname{Im}\kappa_G(u)\operatorname{Im}\gamma(u) > 0. \tag{B.17}$$

Proof. By definition of $A(u)$ in equation $(B.16)$, we have $|A(u)|^2 < 1$

$$\begin{aligned} &\text{iff} \quad |\kappa_G(u) + \gamma(u)|^2 > |\kappa_G(u) - \gamma(u)|^2 \qquad (B.18)\\ &\text{iff} \quad 4\left[\operatorname{Re}\kappa_G(u)\operatorname{Re}\gamma(u) + \operatorname{Im}\kappa_G(u)\operatorname{Im}\gamma(u)\right] > 0, \end{aligned}$$

since the squared terms in the expansion of $(B.18)$ cancel. Thus, the result in $(B.17)$ follows. ■

We now introduce the reasonable assumption that $\rho < 0$. When u is purely real, this leads to the following key results.

Lemma B.7 *Let $\rho < 0$, and assume that u is purely real. If $\operatorname{Im}\gamma(u)$ has the same sign as u, for all $u \in \mathbb{R}$, then*

$$|A(u)| < 1, \text{ for all } u \in \mathbb{R}.$$

Proof. When $u \in \mathbb{R}$, $\operatorname{Re}\kappa_G(u) = \kappa > 0$, and by Proposition B.4 *2.*, $\operatorname{Re}\gamma(u) > 0$. Furthermore, when $u \in \mathbb{R}$ and $\rho < 0$, $\operatorname{Im}\kappa_G(u) = -u\rho\omega$ has the same sign as u. Therefore, given $\rho < 0$, the hypothesis that $\operatorname{Im}\gamma(u)$ has the same sign as u, for all $u \in \mathbb{R}$, implies that

$$\operatorname{Re}\kappa_G(u)\operatorname{Re}\gamma(u) + \operatorname{Im}\kappa_G(u)\operatorname{Im}\gamma(u) > 0.$$

Hence, by Lemma B.6, $|A(u)| < 1$, for all $u \in \mathbb{R}$. ■

Lemmas B.8 and B.10 below parallel Lemma 3.4, p. 678 from Lord and Kahl (2010) which considers the Heston model. The main differences are that we have assumed u is purely real throughout our analysis, and we treat both the SVJ and SVSJ models in addition to the Heston model.

Lemma B.8 *Let $\rho < 0$, and assume that u is purely real. Then for both the Heston and SVJ models,*

$$|A(u)| < 1, \text{ for all } u \in \mathbb{R}.$$

Proof. We can view the Heston model as a special case of the SVJ model with no jumps. By equation $(B.1)$ from Corollary B.2 we have

$$\psi^{J0}(u) = -\frac{1}{2}\left(u^2 + iu\right).$$

Thus, similar to the proof of Lemma B.3 in the SVJ case, if $u \in \mathbb{R}$, then

$$\begin{aligned} z^{J0}(u) = \gamma^{J0}(u)^2 &= \kappa_G(u)^2 - 2\omega^2\psi^{J0}(u) \\ &= (\kappa - iu\rho\omega)^2 + \omega^2\left(u^2 + iu\right) \\ &= \kappa^2 + \left(1-\rho^2\right)u^2\omega^2 + i\left(\omega^2 - 2\kappa\omega\rho\right)u \\ &= x^{J0}(u) + iy^{J0}(u). \end{aligned} \tag{B.19}$$

That is, from equation $(B.19)$, if $u \in \mathbb{R}$, then

$$y^{J0}(u) = \operatorname{Im} z^{J0}(u) = \left(\omega^2 - 2\kappa\omega\rho\right) u.$$

Thus, by the hypothesis that $\rho < 0$, it is clear that $y^{J0}(u) = \operatorname{Im} z^{J0}(u)$ has the same sign as u. Moreover, Proposition B.4 *3.* further implies that $\operatorname{Im} \gamma^{J0}(u)$ also has the same sign as u. Therefore, again by the assumption that $\rho < 0$, Lemma B.7 implies

$$|A(u)| < 1,$$

for all $u \in \mathbb{R}$, as required. ■

For the SVSJ case, recall from equations $(B.7)$ and $(B.8)$ that

$$\operatorname{Re} \widehat{\psi}_{MJ}(u) = \lambda \left(e^{-\frac{1}{2}\alpha^2 u^2} \cos(\beta u) - 1 \right), \text{ and} \tag{B.20}$$

$$\operatorname{Im} \widehat{\psi}_{MJ}(u) = \lambda \left(e^{-\frac{1}{2}\alpha^2 u^2} \sin(\beta u) - u \left(e^{\beta + \frac{1}{2}\alpha^2} - 1 \right) \right). \tag{B.21}$$

Lemma B.9 *The function $g(\alpha, \beta)$ defined by*

$$g(\alpha, \beta) = e^{\beta + \frac{1}{2}\alpha^2} - 1 - \beta \tag{B.22}$$

is strictly positive for all $\alpha > 0$ and $\beta \in \mathbb{R}$.

Proof. Consider $h : \mathbb{R} \to \mathbb{R}$, given by $h(\xi) = e^{\xi} - 1 - \xi$. It can be shown by basic calculus that $\xi = 0$ is the unique global minimizer of $h(\xi)$. Furthermore, $h(0) = 0$. So, if we let $\xi = \beta + \frac{1}{2}\alpha^2$, then

$$h\left(\beta + \frac{1}{2}\alpha^2\right) = e^{\beta + \frac{1}{2}\alpha^2} - 1 - \beta - \frac{1}{2}\alpha^2 \geq 0.$$

But since $\alpha > 0$, this implies that

$$g(\alpha, \beta) = e^{\beta + \frac{1}{2}\alpha^2} - 1 - \beta \geq \frac{1}{2}\alpha^2 > 0,$$

for all $\alpha > 0$ and $\beta \in \mathbb{R}$, as required. ■

To the best of our knowledge, the following result for the SVSJ model has never been presented before. We continue with the reasonable assumption that $\rho < 0$, as in the SVJ case.

Lemma B.10 *Let $\rho < 0$, and assume that u is purely real. Then for the SVSJ model,*

$$|A(u)| < 1, \text{ for all } u \in \mathbb{R}.$$

Proof. By equation $(B.1)$ from Corollary B.2 we have

$$\psi^{J1}(u) = -\frac{1}{2}\left(u^2 + iu\right) + \widehat{\psi}_{MJ}(u),$$

where the real and imaginary parts of $\widehat{\psi}_{MJ}(u)$ are given in equations $(B.20)$ and $(B.21)$ above. Recall the SVSJ part of the proof of Lemma B.3. Similar to that proof, if $u \in \mathbb{R}$, then

$$\begin{aligned} z^{J1}(u) = \gamma^{J1}(u)^2 &= \kappa_G(u)^2 - 2\omega^2\psi^{J1}(u) \\ &= (\kappa - iu\rho\omega)^2 + \omega^2\left(u^2 + iu\right) - 2\omega^2\widehat{\psi}_{MJ}(u) \\ &= \kappa^2 + \left(1-\rho^2\right)u^2\omega^2 - 2\omega^2 \operatorname{Re}\widehat{\psi}_{MJ}(u) \\ &+ i\left[\left(\omega^2 - 2\kappa\omega\rho\right)u - 2\omega^2 \operatorname{Im}\widehat{\psi}_{MJ}(u)\right] \\ &= x^{J1}(u) + iy^{J1}(u). \end{aligned} \tag{B.23}$$

Hence, from equations $(B.23)$ and $(B.21)$, if $u \in \mathbb{R}$, then

$$\begin{aligned} y^{J1}(u) &= \operatorname{Im} z^{J1}(u) = \left(\omega^2 - 2\kappa\omega\rho\right)u - 2\omega^2 \operatorname{Im}\widehat{\psi}_{MJ}(u) \\ &= \left(\omega^2 - 2\kappa\omega\rho\right)u - 2\omega^2\lambda\left(e^{-\frac{1}{2}\alpha^2u^2}\sin(\beta u) - u\left(e^{\beta+\frac{1}{2}\alpha^2} - 1\right)\right) \\ &= \left[\omega^2 - 2\kappa\omega\rho + 2\lambda\omega^2\left(e^{\beta+\frac{1}{2}\alpha^2} - 1 - \beta e^{-\frac{1}{2}\alpha^2u^2}\frac{\sin(\beta u)}{\beta u}\right)\right]u. \end{aligned} \tag{B.24}$$

It follows directly from equation $(B.24)$ that if

$$\omega^2 - 2\kappa\omega\rho + 2\lambda\omega^2\left(e^{\beta+\frac{1}{2}\alpha^2} - 1 - \beta e^{-\frac{1}{2}\alpha^2u^2}\frac{\sin(\beta u)}{\beta u}\right) > 0 \tag{B.25}$$

then $y^{J1}(u) = \operatorname{Im} z^{J1}(u)$ will have the same sign as u. But, we assume that $\rho < 0$. Moreover κ, ω, and λ are all strictly positive in the SVSJ model parameter space. Hence, the condition

$$e^{\beta+\frac{1}{2}\alpha^2} - 1 - \beta e^{-\frac{1}{2}\alpha^2u^2}\frac{\sin(\beta u)}{\beta u} > 0, \tag{B.26}$$

will suffice for $(B.25)$. However, the secondary condition $(B.26)$ simplifies even further. By a property of the function $\frac{\sin(a)}{a}$, known as the "sinc" function, see for example Briggs and Henson (1995), pp. 96-97, $\frac{\sin(\beta u)}{\beta u} \leq 1$ for all $u \in \mathbb{R}$, and $\beta \in \mathbb{R}$. Moreover, it is clear that $e^{-\frac{1}{2}\alpha^2 u^2} \leq 1$ for all $u \in \mathbb{R}$, and $\alpha > 0$. Hence,

$$e^{\beta + \frac{1}{2}\alpha^2} - 1 - \beta > 0 \tag{B.27}$$

suffices for $(B.26)$, and therefore also $(B.25)$. However, the inequality $(B.27)$ follows directly by Lemma B.9. Thus, we have shown that when $\rho < 0$, the inequality $(B.25)$ holds, and as stated above this implies that $y^{J1}(u) = \operatorname{Im} z^{J1}(u)$ will have the same sign as u. Hence, similar to the proof of Lemma B.8, Proposition B.4 *3.* now implies that $\operatorname{Im} \gamma^{J1}(u)$ will also have the same sign as u. Therefore, by Lemma B.7, we obtain the result that $|A(u)| < 1$, for all $u \in \mathbb{R}$, as required. ■

For convenience, the above analysis is summarized as follows.

Lemma B.11 *Let $\rho < 0$, and assume that u is purely real. Then for both the SVJ and SVSJ models,*

$$|A(u)| < 1, \text{ for all } u \in \mathbb{R}.$$

Proof. For the SVJ model see Lemma B.8 above, and for the SVSJ model see Lemma B.10 immediately above. ■

Remark B.12 *Under the assumption that $u = a + ib$ is a complex number, Lord and Kahl (2006), Lemma 3, proves that for the Heston model*

$$\text{if } \rho \leq \frac{\kappa}{\omega}, \text{ then } |A(u)| \leq 1, \text{ for all } u \in \mathbb{C}.$$

Note that if $u \in \mathbb{C}$, then $\gamma(u) = 0$ is possible, implying that $A(u) = 1$. Hence, the inequality $|A(u)| \leq 1$ is not necessarily strict when u is a complex number. However, if $\kappa \geq \omega$, which is often true in practice, then the above result from Lord and Kahl (2006) holds for all $\rho \in [-1, 1]$.

There remains one more small but critically important result concerning the auxiliary variable $A(u)$.

Lemma B.13 *Let u be purely real, and assume that $\rho < 0$. Then for $\tau > 0$, in both the SVJ and SVSJ models, $A(u)\, e^{-\gamma(u)\tau} \neq 1$, for all $u \in \mathbb{R}$.*

Proof. Suppose that for some $u \in \mathbb{R}$, $A(u)\, e^{-\gamma(u)\tau} = 1$. This implies

$$\left| A(u)\, e^{-\gamma(u)\tau} \right| = 1. \tag{B.28}$$

But, by Lemma B.11 $\rho < 0$ implies $|A(u)| < 1$ for all $u \in \mathbb{R}$. Moreover, $\left|e^{-\gamma(u)\tau}\right| = e^{-\operatorname{Re}\gamma(u)\tau} < 1$, by Proposition B.4 *2.*, where $\operatorname{Re}\gamma(u) > 0$, and we assume $\tau > 0$. This contradicts equation $(B.28)$. Therefore,

$$A(u)\, e^{-\gamma(u)\tau} \neq 1, \text{ for all } u \in \mathbb{R},$$

as required. ■

We close this appendix with the continuity of the auxiliary variables related to the joint characteristic function. Recall from equations $(A.21)$, $(A.24)$, and $(A.25)$, respectively in Appendix A, that for $\tau > 0$,

$$H(u) = (\kappa_G(u) + \gamma(u)) \left(1 - A(u)\, e^{-\gamma(u)\tau}\right), \tag{B.29}$$

$$K(u) = \frac{\omega^2 \left(1 - e^{-\gamma(u)\tau}\right)}{H(u)}, \text{ and} \tag{B.30}$$

$$R(u) = 4e^{-\gamma(u)\tau} \left(\frac{\gamma(u)}{H(u)}\right)^2. \tag{B.31}$$

Proposition B.14 *Let u be purely real, and assume that $\rho < 0$. Then, for all $u \in \mathbb{R}$, and for $\tau > 0$,*

1. *$H(u)$ is continuous and never zero,*
2. *$K(u)$ is continuous, and*
3. *$R(u)$ is continuous.*

Proof. By Lemma B.13, if $\rho < 0$ and $\tau > 0$, then $1 - A(u)\, e^{-\gamma(u)\tau}$ is never zero, for any $u \in \mathbb{R}$. Moreover, when $u \in \mathbb{R}$, Proposition B.4 *2.* implies that $\operatorname{Re}\gamma(u) > 0$. Also, if $u \in \mathbb{R}$ then $\operatorname{Re}\kappa_G(u) = \kappa > 0$. This shows that $H(u) \neq 0$, for all $u \in \mathbb{R}$. Furthermore, $A(u)$ is continuous by Lemma B.5, $\kappa_G(u)$ is clearly continuous, and $\gamma(u)$ is continuous by Proposition B.4 *1.* Thus, $H(u)$ is continuous. Hence, we have proven *1.* Then, since $H(u)$ is continuous and never zero, and $\gamma(u)$ is continuous, clearly both $K(u)$ and $R(u)$ are continuous. This proves both *2.* and *3.* ■

B.2 Continuity of Log-Price Affine Coefficients

Using the results from Appendix B.1 above, we are now in a position to prove the main result of this appendix, Proposition B.16. Recall from equations $(A.22)$ and $(A.23)$ in Appendix A that the log-price affine coefficients $C(u)$ and $D(u)$ used in this thesis are given by

$$C(u) = iu\mu_0\tau + \frac{\kappa\eta}{\omega^2}\left(\kappa_G(u) - \gamma(u)\right)\tau \tag{B.32}$$
$$- \frac{2\kappa\eta}{\omega^2}\log\left(\frac{A(u)\, e^{-\gamma(u)\tau} - 1}{A(u) - 1}\right) + \mathbf{1}_{SVJ}\widehat{\psi}_X(u)\,\tau, \text{ and}$$
$$D(u) = \frac{\kappa_G(u) - \gamma(u)}{\omega^2}\left(\frac{1 - e^{-\gamma(u)\tau}}{1 - A(u)\, e^{-\gamma(u)\tau}}\right), \tag{B.33}$$

where the auxiliary variables $\gamma(u)$ and $A(u)$ each differ between the SVJ and SVSJ model specifications. The main result, Proposition B.16, proves that for $\tau > 0$ in both the SVJ and SVSJ models, the primary condition $\rho < 0$ from Appendix B.1 above, implies that $C(u)$ in equation $(B.32)$ is continuous on the principal branch, for all $u \in \mathbb{R}$, and that $D(u)$ in equation $(B.33)$ is continuous, for all $u \in \mathbb{R}$. We continue to assume that the auxiliary variable $\gamma(u)$ corresponds to the positive root, and that for simplicity $\gamma(u)$ is evaluated on the principal branch.

The following result is based on the proof from Lord and Kahl (2010), Theorem 3.5, p. 678. Since we assume that u is purely real, our version is simpler, thus having fewer exceptional cases.

Theorem B.15 *Let u be purely real, and assume that $\rho < 0$. Further assume that $\tau > 0$. Then, in equation $(B.32)$ for the log-price affine coefficient $C(u)$ of both the SVJ and SVSJ models, the argument of the logarithm,*

$$\frac{A(u)\, e^{-\gamma(u)\tau} - 1}{A(u) - 1}, \tag{B.34}$$

never crosses the negative real axis, for any $u \in \mathbb{R}$.

Proof. Suppose that $(B.34)$ does cross the negative real axis. This implies that for some $u \in \mathbb{R}$, and some $\xi \geq 0$,

$$\frac{A(u)\, e^{-\gamma(u)\tau} - 1}{A(u) - 1} = -\xi. \tag{B.35}$$

Note that by Lemma B.11, the hypothesis $\rho < 0$ implies that $|A(u)| < 1$, holds for all $u \in \mathbb{R}$. Further note that since $u \in \mathbb{R}$, Lemma B.5 implies $A(u) \neq 1$. Moreover, since $\rho < 0$, and we assume $\tau > 0$, Lemma B.13 implies that $A(u)\, e^{-\gamma(u)\tau} \neq 1$. Thus, ξ is strictly positive. Consider equation $(B.35)$. Observe that it may be re-written as

$$A(u)\left(e^{-\gamma(u)\tau} + \xi\right) = 1 + \xi. \tag{B.36}$$

Taking the norm of the left-hand side in equation $(B.36)$ yields

$$\begin{aligned}
&\left|A(u)\left(e^{-\gamma(u)\tau} + \xi\right)\right| \leq |A(u)|\left(\left|e^{-\gamma(u)\tau}\right| + \xi\right) \\
&< \left(\left|e^{-\gamma(u)\tau}\right| + \xi\right), \text{ since } |A(u)| < 1 \\
&= e^{-\operatorname{Re}\gamma(u)\tau} + \xi < 1 + \xi, \text{ since } \tau > 0, \text{ and} \\
&\text{since } \operatorname{Re}\gamma(u) > 0 \text{ by Proposition B.4 } \textit{2}.
\end{aligned} \tag{B.37}$$

But the inequality in $(B.37)$ is strict. This contradicts the equality stated in $(B.36)$. Therefore, $(B.34)$ never crosses the negative real axis. ■

Proposition B.16 *Let u be purely real, and assume that $\rho < 0$. Then, for all $u \in \mathbb{R}$, when $\tau > 0$, in both the SVJ and SVSJ models,*

1. *$C(u)$ in equation $(B.32)$ is continuous on the principal branch of the logarithm, and*

2. *$D(u)$ in equation $(B.33)$ is continuous.*

Proof. Since $\kappa_G(u)$ is clearly continuous, and and $\gamma(u)$ is continuous by Proposition B.4 *1.*, the term

$$iu\mu_0\tau + \frac{\kappa\eta}{\omega^2}\left(\kappa_G(u) - \gamma(u)\right)\tau$$

in equation $(B.32)$ for $C(u)$ is clearly continuous. Moreover, $\widehat{\psi}_X(u)$ is continuous by Lemma B.1. Then since $A(u)$ is continuous by Lemma B.5, and we assume $\tau > 0$, the term

$$-\frac{2\kappa\eta}{\omega^2}\log\left(\frac{A(u)\,e^{-\gamma(u)\tau} - 1}{A(u) - 1}\right)$$

in equation $(B.32)$ for $C(u)$ is continuous on the principal branch by the critical result, Theorem B.15. This proves *1.* Furthermore, since $\rho < 0$, and we assume $\tau > 0$, Lemma B.13, implies that

$$1 - A(u)\,e^{-\gamma(u)\tau}$$

is never zero in the denominator of equation $(B.33)$ $D(u)$. Given the other results cited above, this is sufficient for *2.* ■

B.3 Summary

For affine characteristic functions, principal branch continuity is preferable, since with it the tracking of the branch of the complex logarithm in the $C(u)$ coefficient is unnecessary. We may simply use the principal branch.

We have proposed a systematic way to ensure this type of continuity, based on Lord and Kahl (2010). Once the arrangement of the affine coefficients from Appendix A.2 has been reached, one only needs to establish the key property that

$$\begin{aligned} &|A(u)| < 1, \text{ for all } u \in \mathbb{R}, \text{ where} \\ &A(u) = \frac{\kappa_G(u) - \gamma(u)}{\kappa_G(u) + \gamma(u)}. \end{aligned} \tag{B.38}$$

For both the SVJ and SVSJ models, $\rho < 0$ is sufficient for $(B.38)$ by Lemma B.11, summarizing Appendix B.1. Consequently, when $\tau > 0$, the primary hypothesis that $\rho < 0$ is sufficient for principal branch continuity of the log-price coefficients by Theorem B.15 and Proposition B.16, in Appendix B.2. Also, regarding the case $\rho \geq 0$ for the Heston model, see the Lord and Kahl (2006) result in Remark B.12 of Appendix B.1. For continuity of the joint affine coefficients, see Appendix C.3. Joint continuity requires a result from Appendix C.2 that uses an asymptotic argument.

The Riccati equations from Heston (1993), p. 341, have two solutions. One is principal branch continuous, while the other is not. Heston chose the latter solution. This went unnoticed for many years, as those who solved the problem, for example Duffie, Pan, and Singleton (2000), chose not to explain the situation, and in many scenarios the complex discontinuities make very little difference numerically.

Albrecher, Mayer, Schoutens, and Tistaert (2007), Proposition 2, shows that there is a threshold value of τ, depending on the model parameters, below which the Heston characteristic function with complex discontinuities will produce accurate results, but above which inaccuracies will only grow. They also provide an example with realistic parameters where the threshold value of τ is less than one. Hence, the problem is relevant when τ is a late maturity date, in particular in Chapter 6 of this thesis.

Appendix C

The Asymptotics of the Affine Coefficients

For the SVJ and SVSJ models, as established in Appendix A, the joint affine coefficients $\underline{C}(u, v; \tau)$ and $\underline{D}(u, v; \tau)$, from equations $(A.16)$ and $(A.17)$, are determined by the log-price affine coefficients $C(u)$ and $D(u)$, along with the joint auxiliary variables $K(u)$ and $R(u)$. However, the jumps in the SVJ model are an independent Lévy process which can be handled separately without resort to asymptotics. Moreover, in Appendix C.1, we show that the auxiliary variable $\gamma^{J1}(u)$ for the SVSJ model is asymptotically equivalent to $\gamma^{J0}(u)$ for the SVJ model, which is identical to $\gamma^{SV}(u)$. This means that if we agree to handle the jumps in the SVJ model separately, then essentially all of the asymptotics for both the SVJ and SVSJ models are equivalent to the asymptotics of the Heston model. The asymptotics of the log-price coefficients $C(u)$ and $D(u)$ for the Heston model are given a rudimentary treatment in the appendix of Kahl and Jäckel (2005), similar to Appendix C.1 below. However, in Appendix C.2, we further provide asymptotics for the joint auxiliary variables $K(u)$ and $R(u)$. Also, in both

Appendix C.1 and Appendix C.2, we provide asymptotic forms in the limit as $u \to -\infty$ in addition to the limit as $u \to +\infty$. Both limits are needed in nearly all of our proven results. In combination with the continuity results from Appendix B.2, the asymptotic results of Appendix C.1 are used in Chapter 3 to prove that the marginal log-price characteristic functions for the SVJ and SVSJ models are L^1. Also, Appendix C.2 provides a specific result, Lemma C.11, concerning the joint auxiliary variable $K(u)$. In addition to the asymptotics, this result is used in Chapter 4 on the Bates (2006) AML method to prove that the joint conditional characteristic function and its v-derivatives are L^1 in the u-variable, for the SVJ and SVSJ models. Furthermore, in Appendix C.3, Lemma C.11 is used to prove that for the joint affine coefficients $\underline{C}(u, v; \tau)$ and $\underline{D}(u, v; \tau)$ of the SVJ and SVSJ models, assuming $\tau > 0$, the primary hypothesis that $\rho < 0$ is sufficient for principal branch continuity, for all $u \in \mathbb{R}$, and for all $|v| < \delta$, for some $\delta > 0$. Lastly, the results of this appendix are particularly useful in Appendix D where we prove that we can differentiate twice under the Fourier integral in the Bates (2006) AML volatility filter.

C.1 Asymptotics of Log-Price Affine Coefficients

Recall from equation $(A.10)$ in Appendix A that

$$\kappa_G(u) = \kappa - iu\rho\omega. \tag{C.1}$$

From equation $(C.1)$, it is clear that

$$\lim_{u \to \pm\infty} \frac{\kappa_G(u)}{u} = -i\rho\omega, \text{ and} \tag{C.2}$$

$$\kappa_G(u) \sim -i\rho\omega u, \text{ as } u \to \pm\infty. \tag{C.3}$$

Note that $(C.3)$ is consistent with the proof of Lemma B.7 in that when $\rho < 0$, $\operatorname{Im} \kappa_G(u)$ has the same sign as u, for all $u \in \mathbb{R}$.

Further recall from equations $(A.5)$, $(A.9)$, and $(A.11)$ in Appendix A that for the SVJ model

$$\psi^{J0}(u) = \widehat{\psi}_W(u) = -\frac{1}{2}\left(u^2 + iu\right), \text{ and} \tag{C.4}$$

$$\gamma^{J0}(u) = \sqrt{\kappa_G(u)^2 - 2\omega^2\widehat{\psi}_W(u)}. \tag{C.5}$$

Moreover,

$$\gamma^{SV}(u) = \gamma^{J0}(u), \tag{C.6}$$

since the Heston model can be viewed as a special case of the SVJ model with no jumps.

Lemma C.1 *For the SVJ model*

$$\lim_{u\to\pm\infty}\frac{\gamma^{J0}(u)}{u} = \pm\omega\sqrt{1-\rho^2}, \quad \textit{and} \tag{C.7}$$

$$\gamma^{J0}(u) \sim \pm\omega\sqrt{1-\rho^2}u, \textit{ as } u \to \pm\infty. \tag{C.8}$$

Proof. By equations $(C.1)$, $(C.4)$, and $(C.5)$

$$\begin{aligned}
\gamma^{J0}(u) &= \sqrt{\kappa_G(u)^2 - 2\omega^2\widehat{\psi}_W(u)} \\
&= \sqrt{(\kappa - iu\rho\omega)^2 + \omega^2(u^2 + iu)} \\
&= \sqrt{\kappa^2 + (1-\rho^2)u^2\omega^2 + i(\omega^2 - 2\kappa\omega\rho)u}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\lim_{u\to+\infty}\frac{\gamma^{J0}(u)}{u} \\
&= \lim_{u\to+\infty}\sqrt{\frac{\kappa^2 + (1-\rho^2)u^2\omega^2 + i(\omega^2 - 2\kappa\omega\rho)u}{u^2}} \\
&= \omega\sqrt{1-\rho^2}, \text{ and} \\
&\lim_{u\to-\infty}\frac{\gamma^{J0}(u)}{u} = \lim_{u\to+\infty}\frac{\gamma^{J0}(-u)}{-u} = -\lim_{u\to+\infty}\frac{\gamma^{J0}(-u)}{u} \\
&= -\lim_{u\to+\infty}\sqrt{\frac{\kappa^2 + (1-\rho^2)(-u)^2\omega^2 + i(\omega^2 - 2\kappa\omega\rho)(-u)}{u^2}} \\
&= -\omega\sqrt{1-\rho^2}.
\end{aligned}$$

Hence,

$$\gamma^{J0}(u) \sim \pm\omega\sqrt{1-\rho^2}u, \text{ as } u \to \pm\infty, \tag{C.9}$$

as required. ■

Note that the asymptotic form of $\gamma^{J0}(u)$ in $(C.9)$ is consistent with Proposition B.4 *2.* whereby $\operatorname{Re}\gamma(u) > 0$, for all $u \in \mathbb{R}$.

For the next result, recall from equation $(A.6)$ in Appendix A that

$$\begin{aligned}\widehat{\psi}_{MJ}(u) = \lambda\left(\exp\left(i\beta u - \frac{1}{2}\alpha^2u^2\right) - 1\right) \\ - iu\lambda\left(\exp\left(\beta + \frac{1}{2}\alpha^2\right) - 1\right).\end{aligned} \tag{C.10}$$

Lemma C.2

$$\lim_{u\to+\infty}\frac{\widehat{\psi}_{MJ}(\pm u)}{u^2} = 0.$$

Proof.

$$\begin{aligned}&\lim_{u\to+\infty}\frac{\widehat{\psi}_{MJ}(u)}{u^2} \\ &= \lim_{u\to+\infty}\frac{\lambda\left(e^{-\frac{1}{2}\alpha^2u^2}e^{i\beta u} - 1\right)}{u^2} - \lim_{u\to+\infty}\frac{iu\lambda\left(e^{\beta+\frac{1}{2}\alpha^2} - 1\right)}{u^2} \\ &= \lim_{u\to+\infty}\frac{\lambda e^{-\frac{1}{2}\alpha^2u^2}e^{i\beta u}}{u^2} \\ &= \lim_{u\to+\infty}\frac{\lambda e^{-\frac{1}{2}\alpha^2u^2}\left(\cos(\beta u) + i\sin(\beta u)\right)}{u^2} = 0, \text{ and similarly,} \\ &\lim_{u\to+\infty}\frac{\widehat{\psi}_{MJ}(-u)}{u^2} \\ &= \lim_{u\to+\infty}\frac{\lambda e^{-\frac{1}{2}\alpha^2u^2}e^{-i\beta u}}{u^2} \\ &= \lim_{u\to+\infty}\frac{\lambda e^{-\frac{1}{2}\alpha^2u^2}\left(\cos(\beta u) - i\sin(\beta u)\right)}{u^2} = 0,\end{aligned}$$

as required. ■

Further recall from equations $(A.9)$ and $(A.11)$ in Appendix A that for the SVSJ model

$$\psi^{J1}(u) = \widehat{\psi}_W(u) + \widehat{\psi}_{MJ}(u), \text{ so that} \tag{C.11}$$

$$\gamma^{J1}(u) = \sqrt{\kappa_G(u)^2 - 2\omega^2\widehat{\psi}_W(u) - 2\omega^2\widehat{\psi}_{MJ}(u)}, \tag{C.12}$$

where $\widehat{\psi}_W(u) = -\frac{1}{2}\left(u^2 + iu\right)$, as in $(C.4)$.

Lemma C.3 *For the SVJ and SVSJ models $\gamma^{J1}(u)$ and $\gamma^{J0}(u)$ are asymptotically equivalent as $u \to \pm\infty$, and we write*

$$\lim_{u\to\pm\infty} \frac{\gamma(u)}{u} = \pm\omega\sqrt{1-\rho^2}, \quad \textit{with} \tag{C.13}$$

$$\gamma(u) \sim \pm\omega\sqrt{1-\rho^2}u, \textit{ as } u \to \pm\infty, \tag{C.14}$$

for both the SVJ and SVSJ models.

Proof. Combining both limits, and using equation $(C.12)$, we obtain

$$\begin{aligned}
&\lim_{u\to\pm\infty} \frac{\gamma^{J1}(u)}{u} = \lim_{u\to+\infty} \frac{\gamma^{J1}(\pm u)}{\pm u} \\
&= \pm \lim_{u\to+\infty} \sqrt{\frac{\kappa_G(\pm u)^2 - 2\omega^2\widehat{\psi}_W(\pm u) - 2\omega^2\widehat{\psi}_{MJ}(\pm u)}{u^2}} \\
&= \pm \lim_{u\to+\infty} \sqrt{\frac{\kappa_G(\pm u)^2 - 2\omega^2\widehat{\psi}_W(\pm u)}{u^2}}, \text{ by Lemma C.2,} \\
&= \lim_{u\to\pm\infty} \frac{\gamma^{J0}(u)}{u} = \pm\omega\sqrt{1-\rho^2}, \text{ from the proof of Lemma C.1.}
\end{aligned}$$

Therefore, $\gamma^{J1}(u)$ and $\gamma^{J0}(u)$ are asymptotically equivalent as $u \to \pm\infty$, and we may write

$$\gamma(u) \sim \pm\omega\sqrt{1-\rho^2}u, \text{ as } u \to \pm\infty,$$

in both cases, as required. ■

Recall from equation $(A.20)$ in Appendix A that for both the SVJ and SVSJ models, we have the auxiliary variable

$$A(u) = \frac{\kappa_G(u) - \gamma(u)}{\kappa_G(u) + \gamma(u)}. \tag{C.15}$$

Proposition C.4 *Using $A(u)$ from equation (C.15), define*

$$A_{+\infty} = \lim_{u \to +\infty} A(u), \text{ and} \tag{C.16}$$

$$A_{-\infty} = \lim_{u \to -\infty} A(u) = \lim_{u \to +\infty} A(-u). \tag{C.17}$$

Then we have

1. $A_{+\infty} = -1 + 2\rho^2 - 2i\rho\sqrt{1-\rho^2}$ *is constant,*
2. $A_{-\infty} = \overline{A_{+\infty}} = -1 + 2\rho^2 + 2i\rho\sqrt{1-\rho^2}$ *is also constant, and*
3. $|A_{+\infty}| = |A_{-\infty}| = 1$.

Proof. *1.* Using equation (C.2) and the positive limit in equation (C.13), equation (C.15) yields

$$\begin{aligned}
A_{+\infty} = \lim_{u \to +\infty} A(u) &= \frac{-i\rho\omega - \omega\sqrt{1-\rho^2}}{-i\rho\omega + \omega\sqrt{1-\rho^2}} \\
&= \frac{\left(-i\rho\omega - \omega\sqrt{1-\rho^2}\right)\left(i\rho\omega + \omega\sqrt{1-\rho^2}\right)}{\left|-i\rho\omega + \omega\sqrt{1-\rho^2}\right|^2} \\
&= \frac{\rho^2\omega^2 - 2i\omega^2\rho\sqrt{1-\rho^2} - \omega^2\left(1-\rho^2\right)}{\rho^2\omega^2 + \omega^2\left(1-\rho^2\right)} \\
&= \frac{2\rho^2\omega^2 - 2i\omega^2\rho\sqrt{1-\rho^2} - \omega^2}{\omega^2} \\
&= -1 + 2\rho^2 - 2i\rho\sqrt{1-\rho^2} \text{ is constant.}
\end{aligned}$$

2. Similarly, using the negative limit in equation (C.13)

$$\begin{aligned}
A_{-\infty} = \lim_{u \to -\infty} A(u) &= \frac{-i\rho\omega + \omega\sqrt{1-\rho^2}}{-i\rho\omega - \omega\sqrt{1-\rho^2}} \\
&= \frac{i\rho\omega - \omega\sqrt{1-\rho^2}}{i\rho\omega + \omega\sqrt{1-\rho^2}} = \overline{\left(\frac{-i\rho\omega - \omega\sqrt{1-\rho^2}}{-i\rho\omega + \omega\sqrt{1-\rho^2}}\right)} \\
&= \overline{A_{+\infty}} = -1 + 2\rho^2 + 2i\rho\sqrt{1-\rho^2}, \text{ from } \textit{1.}, \text{ is also constant.}
\end{aligned}$$

3. From *2.* we obtain

$$\begin{aligned}
&|A_{+\infty}| = \left|\overline{A_{+\infty}}\right| = |A_{-\infty}| \\
&= \sqrt{\left(1-2\rho^2\right)^2 + 4\rho^2\left(1-\rho^2\right)} \\
&= \sqrt{1-4\rho^2+4\rho^4+4\rho^2-4\rho^4} = 1,
\end{aligned}$$

as required. ■

Note that the result in Proposition C.4 *1.* is consistent with the appendix of Kahl and Jäckel (2005) where

$$\lim_{u\to+\infty}\frac{\kappa_G(u)+\gamma(u)}{\kappa_G(u)-\gamma(u)} = \lim_{u\to+\infty}\frac{1}{A(u)}$$

is computed instead. Also, note that the result $A_{-\infty} = \overline{A_{+\infty}}$ from Proposition C.4 *2.* is consistent with Lord and Kahl (2010), p. 689, where

$$A(-\overline{u}) = \overline{A(u)}, \text{ for all } u \in \mathbb{C}$$

is given.

We are now in a position to present the two main results on the asymptotic forms of the log-price affine coefficients $C(u)$ and $D(u)$ for the SVJ and SVSJ models. These are Lemmas C.6 and C.7 below. First, pursuant to our plan to treat the jumps in the SVJ model separately, we modify equations $(A.18)$ and $(A.22)$ from Appendix A.2 to obtain

$$\widetilde{C}(u) = C(u) - \mathbf{1}_{SVJ}\widehat{\psi}_X(u)\tau = \left\{\begin{array}{c} C^{SV}(u;\tau) \\ C^{J1}(u;\tau) \end{array}\right\}, \text{ that is} \tag{C.18}$$

$$\begin{aligned}
\widetilde{C}(u) &= iu\mu_0\tau + \frac{\kappa\eta}{\omega^2}\left(\kappa_G(u)-\gamma(u)\right)\tau \\
&\quad - \frac{2\kappa\eta}{\omega^2}\log\left(\frac{A(u)\,e^{-\gamma(u)\tau}-1}{A(u)-1}\right).
\end{aligned} \tag{C.19}$$

Lemma C.5 *Let $\tau > 0$. Then for the SVJ and SVSJ models with $\rho < 0$,*

$$\lim_{u\to\pm\infty}\frac{\log\left(\frac{A(u)e^{-\gamma(u)\tau}-1}{A(u)-1}\right)}{u} = 0. \tag{C.20}$$

Proof. By Lemma C.3, $\gamma(u) \sim \pm\omega\sqrt{1-\rho^2}u$, as $u \to \pm\infty$, for both the SVJ and SVSJ models, and by Proposition C.4 *1.* and *2.*

$$\lim_{u\to\pm\infty} A(u) = A_{\pm\infty} = -1 + 2\rho^2 \mp 2i\rho\sqrt{1-\rho^2}.$$

Therefore, since $\tau > 0$,

$$\lim_{u\to\pm\infty} \frac{A(u)\, e^{-\gamma(u)\tau} - 1}{A(u) - 1} = \frac{1}{1 - A_{\pm\infty}}, \text{ is constant.}$$

Moreover,

$$\begin{aligned}
&\frac{1}{1-A_{\pm\infty}} = \frac{1}{2(1-\rho^2) \pm 2i\rho\sqrt{1-\rho^2}} \\
&= \frac{2\left(1-\rho^2\right) \mp 2i\rho\sqrt{1-\rho^2}}{\left|2(1-\rho^2) \pm 2i\rho\sqrt{1-\rho^2}\right|^2} = \frac{2\left(1-\rho^2\right) \mp 2i\rho\sqrt{1-\rho^2}}{4(1-\rho^2)}, \\
&\text{so that, } \operatorname{Re}\left(\frac{1}{1-A_{\pm\infty}}\right) = \frac{1}{2} > 0.
\end{aligned}$$

Since $\tau > 0$, and $\rho < 0$, it now follows from Theorem B.15 that

$$\lim_{u\to\pm\infty} \frac{\log\left(\frac{A(u)e^{-\gamma(u)\tau}-1}{A(u)-1}\right)}{u} = \frac{\log\left(\frac{1}{1-A_{\pm\infty}}\right)}{\lim\limits_{u\to\pm\infty} u} = 0,$$

as required. ■

Lemma C.6 *Let $\tau > 0$. Then when $\rho < 0$ in both the SVJ model without jumps, and the SVSJ model, we have*

$$\lim_{u\to\pm\infty} \frac{\widetilde{C}(u)}{u} = \mp\frac{\kappa\eta\tau}{\omega}\sqrt{1-\rho^2} + i\tau\left(\mu_0 - \frac{\kappa\eta\rho}{\omega}\right), \quad \textit{with} \tag{C.21}$$

$$\widetilde{C}(u) \sim \mp\frac{\kappa\eta\tau}{\omega}\sqrt{1-\rho^2}u + i\tau\left(\mu_0 - \frac{\kappa\eta\rho}{\omega}\right)u, \textit{ as } u \to \pm\infty. \tag{C.22}$$

Proof. Equation $(C.19)$ defines

$$\begin{aligned}
\widetilde{C}(u) = iu\mu_0\tau + \frac{\kappa\eta}{\omega^2}\left(\kappa_G(u) - \gamma(u)\right)\tau \\
- \frac{2\kappa\eta}{\omega^2}\log\left(\frac{A(u)\, e^{-\gamma(u)\tau} - 1}{A(u) - 1}\right).
\end{aligned} \tag{C.23}$$

However, by Lemma C.5, with $\tau > 0$ and $\rho < 0$,

$$\lim_{u \to \pm\infty} \frac{\log\left(\frac{A(u)e^{-\gamma(u)\tau}-1}{A(u)-1}\right)}{u} = 0.$$

Moreover, by equation $(C.2)$ and Lemma C.3 combined,

$$\lim_{u \to \pm\infty} \frac{\kappa_G(u) - \gamma(u)}{u} = -i\rho\omega \mp \omega\sqrt{1-\rho^2}.$$

$$\text{Lastly, } \lim_{u \to \pm\infty} \frac{iu\mu_0\tau}{u} = i\mu_0\tau.$$

Substituting into equation $(C.23)$ yields the required results. ■

Note that the limit as $u \to +\infty$ in equation $(C.21)$ is consistent with Kahl and Jäckel (2005), p. 101, equation (65). Observe that from their equation (34) on p. 98, $d_\infty = \omega\sqrt{1-\rho^2}$, and from their equation (15) on p. 95, $\alpha = \frac{\kappa\eta}{\omega^2}$. Also, note that they treat the $C(u)$ coefficient without a drift. The asymptotic form in equation $(C.22)$ is consistent with Lemma 3.15 from subsection 3.3.2 of this thesis, whereby $\operatorname{Re} C(u) \leq 0$, for all $u \in \mathbb{R}$.

The other main result is the asymptotic form of the log-price affine coefficient $D(u)$. Note that $D(u)$ is unaffected by the jumps in the SVJ model, and that by Lemma C.3 the SVSJ case it is asymptotically equivalent to the Heston and SVJ cases which are the same. Recall from equation $(A.23)$ in Appendix A that

$$D(u) = \frac{\kappa_G(u) - \gamma(u)}{\omega^2}\left(\frac{1-e^{-\gamma(u)\tau}}{1-A(u)\,e^{-\gamma(u)\tau}}\right). \tag{C.24}$$

Lemma C.7 *For both the SVJ and SVSJ models, with $\tau > 0$, we have*

$$\lim_{u \to \pm\infty} \frac{D(u)}{u} = \frac{\mp\sqrt{1-\rho^2} - i\rho}{\omega}, \quad \textit{with} \tag{C.25}$$

$$D(u) \sim \frac{\mp\sqrt{1-\rho^2}u}{\omega} - i\frac{\rho u}{\omega}, \textit{ as } u \to \pm\infty. \tag{C.26}$$

Proof. Since Lemma C.3 states that $\gamma(u) \sim \pm\omega\sqrt{1-\rho^2}u$, as $u \to \pm\infty$, and by Proposition C.4 *1.* and *2.*, $A_{\pm\infty} = \lim_{u\to\pm\infty} A(u)$ is constant,

$$\text{and since } \tau > 0, \lim_{u\to\pm\infty} \frac{1-e^{-\gamma(u)\tau}}{1-A(u)\,e^{-\gamma(u)\tau}} = 1.$$

Therefore, by the definition of $D(u)$ in equation $(C.24)$,

$$\lim_{u\to\pm\infty} \frac{D(u)}{u} = \frac{\lim_{u\to\pm\infty} \frac{\kappa_G(u)}{u} - \lim_{u\to\pm\infty} \frac{\gamma(u)}{u}}{\omega^2},$$

so that by equation $(C.2)$ and Lemma C.3 combined, after re-arranging terms,

$$\lim_{u\to\pm\infty} \frac{D(u)}{u} = \frac{\mp\sqrt{1-\rho^2} - i\rho}{\omega},$$

as required. ■

Note that the limit as $u \to +\infty$ in equation $(C.25)$ is consistent with Kahl and Jäckel (2005), p. 101, equation (61). Moreover, the asymptotic form in equation $(C.26)$ is consistent with Lemma 3.15 from subsection 3.3.2 of this thesis, whereby $\operatorname{Re} D(u) \leq 0$, for all $u \in \mathbb{R}$.

C.2 Asymptotics of Joint Auxiliary Variables

Recall from equations $(A.24)$ and $(A.25)$ in Appendix A.2, where the joint affine coefficients $\underline{C}(u, v; \tau)$ and $\underline{D}(u, v; \tau)$ used in this thesis are specified, that the joint auxiliary variables $K(u)$ and $R(u)$ are given by

$$K(u) = \frac{\omega^2\left(1-e^{-\gamma(u)\tau}\right)}{H(u)}, \text{ and} \tag{C.27}$$

$$R(u) = 4e^{-\gamma(u)\tau}\left(\frac{\gamma(u)}{H(u)}\right)^2. \tag{C.28}$$

Since both $K(u)$ and $R(u)$ depend on the auxiliary variable $H(u)$, we begin with the asymptotics of $H(u)$. Recall from equation $(A.21)$ in Appendix A.2 that

$$H(u) = [\kappa_G(u) + \gamma(u)]\left[1 - A(u)\,e^{-\gamma(u)\tau}\right]. \tag{C.29}$$

Lemma C.8 *For both the SVJ and SVSJ models, with $\tau > 0$, we have*

$$\lim_{u\to\pm\infty} \frac{H(u)}{u} = \pm\omega\sqrt{1-\rho^2} - i\rho\omega, \quad \textit{with} \tag{C.30}$$

$$H(u) \sim \pm\omega\sqrt{1-\rho^2}u - i\rho\omega u, \textit{ as } u \to \pm\infty. \tag{C.31}$$

Proof. By Lemma C.3, $\gamma(u) \sim \pm\omega\sqrt{1-\rho^2}u$, as $u \to \pm\infty$, and also by Proposition C.4 *1.* and *2.*, $A_{\pm\infty} = \lim_{u\to\pm\infty} A(u)$ is constant, thus

$$\lim_{u\to\pm\infty} 1 - A(u)\, e^{-\gamma(u)\tau} = 1.$$

Hence, by equation $(C.14)$ in Lemma C.3, and by equation $(C.2)$, after re-arranging terms,

$$\lim_{u\to\pm\infty} \frac{H(u)}{u} = \pm\omega\sqrt{1-\rho^2} - i\rho\omega,$$

as required. ■

The next two results, for $K(u)$ and $R(u)$ respectively, both show that the limits as $u \to \pm\infty$ are zero. In particular, $R(u) \to 0$ exponentially as $u \to \pm\infty$.

Lemma C.9 *For both the SVJ and SVSJ models, with $\tau > 0$,*

$$K(u) \sim \frac{\pm\omega\sqrt{1-\rho^2} + i\rho\omega}{u}, \textit{ as } u \to \pm\infty. \tag{C.32}$$

$$\textit{Hence, } \lim_{u\to\pm\infty} K(u) = 0. \tag{C.33}$$

Proof. From equation $(C.27)$

$$K(u) = \frac{\omega^2\left(1 - e^{-\gamma(u)\tau}\right)}{H(u)}.$$

But by Lemma C.3,

$$\gamma(u) \sim \pm\omega\sqrt{1-\rho^2}u, \text{ as } u \to \pm\infty,$$

and by Lemma C.8,

$$H(u) \sim \pm\omega\sqrt{1-\rho^2}u - i\rho\omega u, \text{ as } u \to \pm\infty.$$

Therefore,

$$\begin{aligned} K(u) &\sim \frac{\omega^2}{\pm\omega\sqrt{1-\rho^2}u - i\rho\omega u} \\ &= \frac{\omega^2\left(\pm\omega\sqrt{1-\rho^2} + i\rho\omega\right)u}{\left|\pm\omega\sqrt{1-\rho^2}u - i\rho\omega u\right|^2} \\ &= \frac{\pm\omega\sqrt{1-\rho^2} + i\rho\omega}{u}, \text{ as } u \to \pm\infty, \end{aligned}$$

as required. ■

Lemma C.10 *For both the SVJ and SVSJ models, with $\tau > 0$,*

$$R(u) \sim 4e^{\mp\omega\sqrt{1-\rho^2}u\tau}\left[\left(1-\rho^2\right) \pm i\rho\sqrt{1-\rho^2}\right]^2, \text{ as } u \to \pm\infty. \tag{C.34}$$

$$\text{Hence, } R(u) \to 0 \text{ exponentially as } u \to \pm\infty. \tag{C.35}$$

Proof. From equation (*C*.28)

$$R(u) = 4e^{-\gamma(u)\tau}\left(\frac{\gamma(u)}{H(u)}\right)^2.$$

But by Lemma C.3, $\gamma(u) \sim \pm\omega\sqrt{1-\rho^2}u$, as $u \to \pm\infty$, and by Lemma C.8, $H(u) \sim \pm\omega\sqrt{1-\rho^2}u - i\rho\omega u$, as $u \to \pm\infty$. Therefore,

$$\begin{aligned} \frac{\gamma(u)}{H(u)} &\sim \frac{\pm\omega\sqrt{1-\rho^2}u}{\left(\pm\omega\sqrt{1-\rho^2} - i\rho\omega\right)u} \\ &= \frac{\left(\pm\omega\sqrt{1-\rho^2}\right)\left(\pm\omega\sqrt{1-\rho^2} + i\rho\omega\right)}{\left|\pm\omega\sqrt{1-\rho^2} - i\rho\omega\right|^2} \\ &= \frac{\omega^2\left(1-\rho^2\right) \pm i\omega^2\rho\sqrt{1-\rho^2}}{\omega^2} \\ &= \left(1-\rho^2\right) \pm i\rho\sqrt{1-\rho^2}, \text{ as } u \to \pm\infty. \end{aligned} \tag{C.36}$$

Moreover, again since $\gamma(u) \sim \pm\omega\sqrt{1-\rho^2}u$, as $u \to \pm\infty$, the result $(C.36)$ gives

$$R(u) \sim 4e^{\mp\omega\sqrt{1-\rho^2}u\tau}\left[\left(1-\rho^2\right) \pm i\rho\sqrt{1-\rho^2}\right]^2, \text{ as } u \to \pm\infty,$$

when $\tau > 0$, as required. ■

This next result, Lemma C.11, arises frequently in this thesis. It is used in Chapter 4 to show that the joint transforms related to the Bates (2006) AML volatility filter are L^1 in the u-variable. Furthermore, it is used in the Appendix D proof that we can differentiate under the Fourier integral under the Bates (2006) AML volatility filtration method. Lastly, Lemma C.11, will be used in Appendix C.3 below to show that, when $\tau > 0$, the joint affine coefficients $\underline{C}(u, v; \tau)$ and $\underline{D}(u, v; \tau)$ for the SVJ and SVSJ models are, provided $|A(u)| < 1$, for all $u \in \mathbb{R}$, as in Appendix B, principal branch continuous for all $u \in \mathbb{R}$, and for all $|v| < \delta$, for some $\delta > 0$.

Lemma C.11 *Assume that $\rho < 0$. Then, for both the SVJ and SVSJ models, with $\tau > 0$, we have that*

$$\begin{aligned} &\operatorname{Re}(1 - K(u)\, iv) > 0, \textit{ for all } u \in \mathbb{R}, \\ &\textit{and for all } |v| < \delta, \textit{ for some } \delta > 0 \textit{ sufficiently small.} \end{aligned} \tag{C.37}$$

Proof. First note that

$$\operatorname{Re}(1 - K(u)\, iv) = 1 + \operatorname{Im} K(u)\, v. \tag{C.38}$$

We may assume that $\tau > 0$. Then, since by Proposition B.14 2., given $\rho < 0$, $K(u)$ is continuous, and by Lemma C.9 $\lim_{u \to \pm\infty} K(u) = 0$, we have that $\operatorname{Im} K(u)$ is bounded above and below, for all $u \in \mathbb{R}$. Therefore, by equation $(C.38)$, there exists $\delta > 0$ sufficiently small, such that $\operatorname{Re}(1 - K(u)\, iv) > 0$, for all $u \in \mathbb{R}$, and for all $|v| < \delta$, as required. ■

C.3 Continuity of Joint Affine Coefficients

We now fulfill the promise from Appendix A to show that the joint affine coefficients, $\underline{C}(u, v; \tau)$ and $\underline{D}(u, v; \tau)$, used in this thesis, for the SVJ and SVSJ models combined, are principal branch continuous for all $u \in \mathbb{R}$, and for all $|v| < \delta$, for some $\delta > 0$, when $\tau > 0$ and $\rho < 0$, as considered in Appendix B, and given the results in Appendix C.2. In doing so, we will highlight some of the main results from Appendix A, Appendix B, and Appendix C.

Recall from equations $(A.16)$ and $(A.17)$ in Appendix A.2 that the joint affine coefficients used in this thesis are given by

$$\underline{C}(u, v) = C(u) - \frac{2\kappa\eta}{\omega^2} \log(1 - K(u)\, iv), \text{ and} \tag{C.39}$$

$$\underline{D}(u, v) = D(u) + \frac{R(u)\, iv}{1 - K(u)\, iv}. \tag{C.40}$$

Proposition C.12 *Let u be purely real, and assume both that $\tau > 0$ and $\rho < 0$, in the SVJ and SVSJ models. Then, for all $u \in \mathbb{R}$, and for all $|v| < \delta$, for some $\delta > 0$ sufficiently small,*

1. *$\underline{C}(u, v)$ in equation $(C.39)$ is continuous on the principal branch, and*
2. *$\underline{D}(u, v)$ in equation $(C.40)$ is continuous.*

Proof. Since, u is purely real, $\tau > 0$, and $\rho < 0$, Proposition B.16 *1.* gives that the log-price coefficient $C(u)$ is continuous on the principal branch for all $u \in \mathbb{R}$. Moreover, by Lemma C.11, given $\tau > 0$ and $\rho < 0$, $\mathrm{Re}(1 - K(u)\, iv) > 0$, for all $u \in \mathbb{R}$, and for all $|v| < \delta$, for some $\delta > 0$ sufficiently small. Hence,

$$\log(1 - K(u)\, iv)$$

is continuous on the principal branch. This proves *1.*

To prove *2.*, consider that by Proposition B.16 *2.*, given $\tau > 0$ and $\rho < 0$, the log-price coefficient $D(u)$ is continuous for all $u \in \mathbb{R}$. Furthermore, by Proposition B.14 3., $\rho < 0$ with $\tau > 0$ implies that $R(u)$ is continuous for all $u \in \mathbb{R}$. Moreover, by Lemma C.11, $\tau > 0$ with $\rho < 0$ implies that $\mathrm{Re}(1 - K(u)\, iv) > 0$, for all $u \in \mathbb{R}$, and for all $|v| < \delta$, for some $\delta > 0$ sufficiently small. Hence, the numerator of

$$\frac{R(u)\, iv}{1 - K(u)\, iv}$$

is continuous, and its denominator is never zero. Therefore, $\underline{D}(u, v)$ is continuous for all $u \in \mathbb{R}$, and for all $|v| < \delta$, for some $\delta > 0$ sufficiently small, as required. ■

Appendix D

Differentiating Under the Fourier Integral: AML

This Appendix is an extension of Section 4.3 on L^1 transforms for the AML method. As in equations (4.61) and (4.62) from subsection 4.3.1, define

$$\begin{aligned}\underline{F}\left(u,v|y_n^H\right) &= e^{\underline{C}(u,v;\Delta t)}G_{n|n}\left[\underline{D}\left(u,v;\Delta t\right)\right] \\ &= e^{\underline{C}(u,v;\Delta t)}\left[1-b_n\underline{D}\left(u,v;\Delta t\right)\right]^{-a_n}. \end{aligned} \tag{D.1}$$

In this Appendix we prove that for the SVJ and SVSJ models,

$$\frac{\partial}{\partial v}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u,v|y_n^H\right)du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_v\left(u,v|y_n^H\right)du, \tag{D.2}$$

and

$$\frac{\partial^2}{\partial v^2}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}\left(u,v|y_n^H\right)du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}}\underline{F}_{vv}\left(u,v|y_n^H\right)du, \tag{D.3}$$

as we require for items *2.* and *3.* in Proposition 4.14 located near the end of subsection 4.3.1. To prove $(D.2)$ and $(D.3)$ above, we apply the theory of uniform convergence for improper integrals with a parameter, in particular Abel's test, and one other result from Zorich (2004), cited below. In this Appendix we take the hypothesis $\tau = \Delta t > 0$ to be given.

D.1 Results from Classical Analysis

Definition D.1 (Uniform Convergence, Zorich (2004), pp. 415-16) *Consider the improper integral with a parameter given by*

$$G(y) = \int_a^\infty g(x,y)\,dx, \tag{D.4}$$

over the interval $[a,\infty) \subset \mathbb{R}$, *with* $y \in [c,d]$. *We say that the improper integral* $(D.4)$ *converges uniformly for all* $y \in [c,d]$ *if for all* $\varepsilon > 0$ *there exists a neighbourhood* $U_{[a,\infty)}(\infty)$ *of infinity with* $U_{[a,\infty)}(\infty) \subset [a,\infty)$ *such that*

$$\left| \int_b^\infty g(x,y)\,dx \right| < \varepsilon \tag{D.5}$$

for all $b \in U_{[a,\infty)}(\infty)$, *and for all* $y \in [c,d]$.

Proposition D.2 (Zorich (2004), p. 426) *Let the functions* $g(x,y)$ *and* $g_y(x,y) = \frac{\partial}{\partial y} g(x,y)$ *be continuous for every for every* $x \in [a,\infty)$, *and for each* $y \in [c,d]$. *If*

1. *the integral* $\int_a^\infty g_y(x,y)\,dx$ *converges uniformly, for each* $y \in [c,d]$, *and*
2. *the integral* $\int_a^\infty g(x,y)\,dx$ *converges for some* $y_0 \in [c,d]$,

then
the integral $\int_a^\infty g(x,y)\,dx$ *converges uniformly for all* $y \in [c,d]$, *and*

$$\frac{\partial}{\partial y} \int_a^\infty g(x,y)\,dx = \int_a^\infty g_y(x,y)\,dx. \tag{D.6}$$

Proof. See Zorich (2004), pp. 426-27. ■

The final classical result that we will put to use in this appendix is called Abel's test. It is cited together with the more well known result, Dirichlet's test, in Zorich (2004), Proposition 3, p. 421.

Proposition D.3 (Abel's Test, Zorich (2004), pp. 421-22) *Assume that the functions $f(x,y)$ and $g(x,y)$ are integrable with respect to x on every closed interval $[a,b] \subset [a,\infty)$ at each $y \in [c,d]$. Then a sufficient condition for the uniform convergence of the improper integral*

$$\int_a^{\infty} f(x,y)\, g(x,y)\, dx, \tag{D.7}$$

for each $y \in [c,d]$, is that the following pair of conditions holds.

1. *The integral $\int_a^{\infty} g(x,y)\, dx$ converges uniformly, for each $y \in [c,d]$.*
2. *For each $y \in [c,d]$, the function $f(x,y)$ is monotone in x on the interval $[a,\infty)$, and there exists $M \in \mathbb{R}$ such that $|f(x,y)| < M$ for every $x \in [a,\infty)$, and for each $y \in [c,d]$.*

Proof. See Zorich (2004), p. 422. ■

Remark D.4 (Zorich (2004), p. 423) *The uniform convergence of improper integrals as presented here applies to integrands that are real-valued, vector-valued, complex-valued, and in general over any vector space for the integrand that is complete. However, any treatment involving Abel's test must provide that the function assumed to be monotone is real-valued.*

Pursuant to Remark D.4 above, when the function $f(x,y)$ in equation $(D.7)$ of Proposition D.3 (Abel's test) is complex-valued we will take its real and imaginary parts, test each part for monotonicity and boundedness, then apply Abel's test to the resulting pair of improper integrals.

D.2 The First Derivative

From equation $(D.1)$ we can write $\underline{F} = \underline{F}\left(u, v | y_n^H\right)$ as

$$\underline{F}\left(u, v | y_n^H\right) = e^{\underline{C}(u,v;\Delta t) - a_n \log(1 - b_n \underline{D}(u,v;\Delta t))}. \tag{D.8}$$

Thus, due to the exponential form, we obtain

$$\underline{F}_v = \underline{f}_v \underline{F}, \text{ where } \underline{f} = \log \underline{F}. \tag{D.9}$$

The first goal is to show that we can obtain the result that

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v \left(u, v | y_n^H\right) du \tag{D.10}$$

$$= \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{f}_v \left(u, v | y_n^H\right) \underline{F} \left(u, v | y_n^H\right) du, \tag{D.11}$$

converges uniformly, for every $v \in (-\delta, \delta)$, for some $\delta > 0$, by Abel's test, which is Proposition D.3 above. But, because $\underline{f}_v$ is complex-valued, by Remark D.4 above, we must take its real and imaginary parts. This can be done using the asymptotic form of $\underline{f}_v$, as illustrated in Lemma D.5 below. The asymptotic forms of $\operatorname{Re} \underline{f}_v$ and $\operatorname{Im} \underline{f}_v$, as $u \to \pm\infty$, turn out to be sufficient to obtain the result $(D.11)$ above, and this is demonstrated in Lemma D.6 and its proof below.

Lemma D.5 *Let $\delta > 0$, and for the SVJ and SVSJ models, assume that $\rho < 0$. For a sufficiently large real number $a > 0$, define the respective neighbourhoods $U_{[a,+\infty)}(+\infty) \subset [a, +\infty)$, and $U_{(-\infty,-a]}(-\infty) \subset (-\infty, -a]$ of $\pm\infty$. Then, for each $v \in (-\delta, \delta)$,*

1. *$\operatorname{Re} \underline{f}_v \sim \widetilde{\alpha}_v^{(\pm)}$ as $u \to \pm\infty$, such that $\widetilde{\alpha}_v^{(+)}$ is monotone and bounded for all $u \in U_{[a,+\infty)}(+\infty)$, and $\widetilde{\alpha}_v^{(-)}$ is monotone and bounded for all $u \in U_{(-\infty,-a]}(-\infty)$.*

2. *$\operatorname{Im} \underline{f}_v \sim \widetilde{\beta}_v^{(\pm)}$ as $u \to \pm\infty$, such that $\widetilde{\beta}_v^{(+)}$ is monotone and bounded for all $u \in U_{[a,+\infty)}(+\infty)$, and $\widetilde{\beta}_v^{(-)}$ is monotone and bounded for all $u \in U_{(-\infty,-a]}(-\infty)$.*

Proof. Recall from equations (4.56) and (4.58) in subsection 4.3.1 that

$$\underline{C}_v = \frac{\partial}{\partial v} \underline{C}(u, v) = \frac{\frac{2\kappa\eta}{\omega^2} K(u)\, i}{1 - K(u)\, iv}, \text{ and} \tag{D.12}$$

$$\underline{D}_v = \frac{\partial}{\partial v}\underline{D}\,(u,v) = \frac{R\,(u)\,i}{(1-K\,(u)\,iv)^2}. \tag{D.13}$$

Further recall from subsection 4.3.1 that by equation (4.66)

$$\underline{f}_v = \underline{C}_v + \frac{a_n b_n \underline{D}_v}{1 - b_n \underline{D}}. \tag{D.14}$$

However, by Lemma 3.15 from subsection 3.3.2, $\mathrm{Re}\,\underline{D}\,(u,v;\tau) \leq 0$, for all $(u,v) \in \mathbb{R} \times (-\delta,\delta)$, for some $\delta > 0$. Thus, since $b_n > 0$, $\mathrm{Re}\,[1 - b_n\underline{D}] > 0$, for all $(u,v) \in \mathbb{R} \times (-\delta,\delta)$. Also, by Lemma C.11 from Appendix C.2, $\mathrm{Re}\,(1 - K\,(u)\,iv) > 0$, for all $(u,v) \in \mathbb{R} \times (-\delta,\delta)$, for $\delta > 0$ sufficiently small. Furthermore, by Lemma C.10 from Appendix C.2, $R\,(u) \to 0$ exponentially as $u \to \pm\infty$. Thus, altogether we have

$$\underline{f}_v \sim \underline{C}_v = \frac{\frac{2\kappa\eta}{\omega^2} K\,(u)\,i}{1 - K\,(u)\,iv}, \text{ as } u \to \pm\infty. \tag{D.15}$$

But, recall from Lemma C.9 in Appendix C.2 that

$$K\,(u) \sim \frac{\pm\omega\sqrt{1-\rho^2} + i\rho\omega}{u}, \text{ as } u \to \pm\infty. \tag{D.16}$$

Upon substituting $(D.16)$ into $(D.15)$ and simplifying, we obtain

$$\underline{f}_v \sim \frac{\frac{2\kappa\eta}{\omega^2}\left[-\left(\rho\omega u + \omega^2 v\right) \pm \omega\sqrt{1-\rho^2}ui\right]}{u^2 + 2\rho\omega uv + \omega^2 v^2}, \text{ as } u \to \pm\infty. \tag{D.17}$$

By the result $(D.17)$ we have

$$\widetilde{\alpha}_v^{(\pm)} = \frac{\frac{-2\kappa\eta}{\omega^2}\left(\rho\omega u + \omega^2 v\right)}{u^2 + 2\rho\omega uv + \omega^2 v^2} \sim \frac{-2\kappa\eta\rho}{u\omega}, \text{ as } u \to \pm\infty, \text{ and} \tag{D.18}$$

$$\widetilde{\beta}_v^{(\pm)} = \frac{\frac{\pm 2\kappa\eta}{\omega^2}\left(\omega\sqrt{1-\rho^2}u\right)}{u^2 + 2\rho\omega uv + \omega^2 v^2} \sim \frac{\pm 2\kappa\eta\sqrt{1-\rho^2}}{u\omega}, \text{ as } u \to \pm\infty. \tag{D.19}$$

Let $v \in (-\delta,\delta)$. Then, by the result $(D.18)$, $\widetilde{\alpha}_v^{(+)}$ is monotone and bounded for all $u \in U_{[a,+\infty)}\,(+\infty)$. Moreover, $\widetilde{\alpha}_v^{(-)}$ is also monotone and bounded for all $u \in U_{(-\infty,-a]}\,(-\infty)$. Likewise by the result $(D.19)$, $\widetilde{\beta}_v^{(+)}$ is monotone and bounded for all $u \in U_{[a,+\infty)}\,(+\infty)$, and $\widetilde{\beta}_v^{(-)}$ is monotone and bounded for all $u \in U_{(-\infty,-a]}\,(-\infty)$. This completes the proof. ■

Lemma D.6 *Let $\delta > 0$, and assume that the hypotheses and definitions of Lemma D.5 above hold. If in addition, for every $v \in (-\delta, \delta)$,*

1. *$\underline{F}\left(u, v|y_n^H\right)$ is uniformly continuous in u, for all $u \in \mathbb{R}$,*
2. *$\underline{F}\left(u, v|y_n^H\right)$ is L^1 in the u-variable, and*
3. *$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du$ converges uniformly,*

then
$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u, v|y_n^H\right) du \text{ converges uniformly,} \tag{D.20}$$
for every $v \in (-\delta, \delta)$.

Proof. We will treat one asymptotic form in detail. The other cases are similar. Let $a > 0$ be sufficiently large. Recall from Lemma D.5 above that $\operatorname{Re} \underline{f}_v \sim \widetilde{\alpha}_v^{(+)}$, as $u \to +\infty$, where $\widetilde{\alpha}_v^{(+)}$ is monotone and bounded for all u in some neighbourhood $U_{[a,+\infty)}(+\infty)$ of $+\infty$. Thus, by item *3.* above, and Proposition D.3 (Abel's test),
$$\int_a^{\infty} e^{-iuy_{n+1}} \widetilde{\alpha}_v^{(+)} \underline{F} du \text{ converges uniformly,} \tag{D.21}$$
for all $v \in (-\delta, \delta)$. Hence, by Definition D.1, for every $\varepsilon > 0$,
$$\left| \int_b^{\infty} e^{-iuy_{n+1}} \widetilde{\alpha}_v^{(+)} \underline{F} du \right| < \frac{\varepsilon}{2}, \tag{D.22}$$
for all $b \in U_{[a,+\infty)}(+\infty)$, for each $v \in (-\delta, \delta)$. Also, $\underline{f}_v$ is bounded for all $u \in \mathbb{R}$, and for all $v \in (-\delta, \delta)$, by Proposition 4.9 in subsection 4.3.1; hence, so is $\operatorname{Re} \underline{f}_v$. Moreover, by Lemma D.5, $\widetilde{\alpha}_v^{(+)}$ is bounded for all $u \in U_{[a,+\infty)}(+\infty)$, and for all $v \in (-\delta, \delta)$. Thus, since for all $v \in (-\delta, \delta)$, $\underline{F}\left(u, v|y_n^H\right)$ is both uniformly continuous in u, for all $u \in \mathbb{R}$, and L^1 in u, we have that for every $\varepsilon > 0$,
$$\int_b^{\infty} |\underline{F}| \left| \operatorname{Re} \underline{f}_v - \widetilde{\alpha}_v^{(+)} \right| du < \frac{\varepsilon}{2}, \tag{D.23}$$

for all $b \in U_{[a,+\infty)}(+\infty)$, for each $v \in (-\delta, \delta)$. The critical result follows from the pair of results $(D.22)$ and $(D.23)$ above. For every $\varepsilon > 0$,

$$\begin{aligned}
&\left| \int_b^{\infty} e^{-iuy_{n+1}} \operatorname{Re} \underline{f}_v \underline{F} du \right| \\
&= \left| \int_b^{\infty} e^{-iuy_{n+1}} \widetilde{\alpha}_v^{(+)} \underline{F} du + \int_b^{\infty} e^{-iuy_{n+1}} \left(\operatorname{Re} \underline{f}_v - \widetilde{\alpha}_v^{(+)} \right) \underline{F} du \right| \\
&\leq \left| \int_b^{\infty} e^{-iuy_{n+1}} \widetilde{\alpha}_v^{(+)} \underline{F} du \right| + \left| \int_b^{\infty} e^{-iuy_{n+1}} \underline{F} \left(\operatorname{Re} \underline{f}_v - \widetilde{\alpha}_v^{(+)} \right) du \right| \\
&\leq \left| \int_b^{\infty} e^{-iuy_{n+1}} \widetilde{\alpha}_v^{(+)} \underline{F} du \right| + \int_b^{\infty} e^{-iuy_{n+1}} |\underline{F}| \left| \operatorname{Re} \underline{f}_v - \widetilde{\alpha}_v^{(+)} \right| du \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,
\end{aligned}$$

for all $b \in U_{[a,+\infty)}(+\infty)$, for each $v \in (-\delta, \delta)$. Hence, by Definition D.1 for the uniform convergence of an improper integral with a parameter,

$$\int_a^{\infty} e^{-iuy_{n+1}} \operatorname{Re} \underline{f}_v \underline{F} du \text{ converges uniformly,}$$

for each $v \in (-\delta, \delta)$. Similarly, based on the other results from Lemma D.5 above, each of the following tail integrals

$$\begin{aligned}
&\int_a^{\infty} e^{-iuy_{n+1}} \operatorname{Im} \underline{f}_v \underline{F} du, \\
&\int_{-\infty}^{-a} e^{-iuy_{n+1}} \operatorname{Re} \underline{f}_v \underline{F} du, \text{ and} \\
&\int_{-\infty}^{-a} e^{-iuy_{n+1}} \operatorname{Im} \underline{f}_v \underline{F} du,
\end{aligned}$$

converges uniformly, for every $v \in (-\delta, \delta)$. Thus, since by Proposition 4.10 from subsection 4.3.1, $\underline{F}_v = \underline{f}_v \underline{F}$ is continuous on the principal branch, for all $(u, v) \in \mathbb{R} \times (-\delta, \delta)$, we have

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v \left(u, v | y_n^H \right) du \text{ converges uniformly,} \tag{D.24}$$

for every $v \in (-\delta, \delta)$, as required. ■

Theorem D.7 (Differentiation I) *Assume that $\rho < 0$ in the SVJ and SVSJ models, and note that this is sufficient for the asymptotic forms in Lemma D.5 above. From equation $(D.1)$ define*

$$\begin{aligned} \underline{F}\left(u,v|y_n^H\right) &= e^{\underline{C}(u,v)} G_{n|n}\left[\underline{D}\left(u,v\right)\right] \\ &= e^{\underline{C}(u,v;\Delta t)}\left[1 - b_n \underline{D}\left(u,v;\Delta t\right)\right]^{-a_n}, \qquad \text{(D.25)} \\ \text{with } \underline{F}_v\left(u,v|y_n^H\right) &= \frac{\partial}{\partial v}\underline{F}\left(u,v|y_n^H\right). \end{aligned}$$

Let $\delta > 0$. The main hypotheses are that for every $v \in (-\delta,\delta)$, $\underline{F}\left(u,v|y_n^H\right)$ is uniformly continuous in u, for all $u \in \mathbb{R}$, $\underline{F}\left(u,v|y_n^H\right)$ is L^1 in u, and the improper integral

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u,v|y_n^H\right) du \text{ converges uniformly.} \qquad \text{(D.26)}$$

If Lemma D.5 and the main hypotheses hold, then for every $v \in (-\delta,\delta)$,

1. $\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u,v|y_n^H\right) du$ *converges uniformly,*
2. $\frac{\partial}{\partial v}\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u,v|y_n^H\right) du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u,v|y_n^H\right) du$, *and*
3. $\underline{F}_v\left(u,v|y_n^H\right)$ *is uniformly continuous in u, for all $u \in \mathbb{R}$.*

Proof. Given that Lemma D.5 holds, the main hypotheses of the current theorem satisfy Lemma D.6, and this proves item *1.* Furthermore, item *1.* and the hypothesis $(D.26)$ imply item *2.*, by Proposition D.2 from Appendix D.1. Moreover, with item *2.* having been established, equation (4.80) from subsection 4.3.1 indicates that $\underline{F}_v\left(u,v|y_n^H\right)$ is the Fourier transform in the u-variable of $\phi_{n+1|n+1}^{(1)}(v)\, p\left(y_{n+1}|y_n^H\right)$, for every $v \in (-\delta,\delta)$. But, by Proposition 4.14 *1.* from subsection 4.3.1, for every $v \in (-\delta,\delta)$, $\phi_{n+1|n+1}^{(1)}(v)\, p\left(y_{n+1}|y_n^H\right)$ is L^1 in y_{n+1}. Therefore, we have by Lemma 3.3 from subsection 3.2.1 that $\underline{F}_v\left(u,v|y_n^H\right)$ is uniformly continuous in u, for all $u \in \mathbb{R}$, for every $v \in (-\delta,\delta)$. This completes the proof. ■

D.3 The Second Derivative

Recall from equation $(D.9)$ in Appendix D.2 above that

$$\underline{F}_v = \underline{f}_v \underline{F}, \text{ where } \underline{f} = \log \underline{F}. \tag{D.27}$$

By the product rule, the second v-derivative of $\underline{F}\left(u, v|y_n^H\right)$ is given by

$$\underline{F}_{vv} = \underline{f}_{vv}\underline{F} + \underline{f}_v \underline{F}_v. \tag{D.28}$$

The main goal is to show that we can obtain the result that

$$\begin{aligned} &\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_{vv}\left(u, v|y_n^H\right) du \\ &= \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \left(\underline{f}_{vv}\underline{F} + \underline{f}_v \underline{F}_v\right) du, \end{aligned} \tag{D.29}$$

converges uniformly, for every $v \in (-\delta, \delta)$, for some $\delta > 0$. Using the form of the integral in $(D.29)$, this can be done in two steps. The step described in Lemma D.8 below is trivial given the results from Appendix D.2. The other step requires the asymptotic form of $\underline{f}_{vv}$.

Lemma D.8 *Let $\delta > 0$, and assume that the hypotheses and definitions of Lemma D.5 from Appendix D.2 hold. If in addition, for every $v \in (-\delta, \delta)$,*

1. *$\underline{F}_v\left(u, v|y_n^H\right)$ is uniformly continuous in u, for all $u \in \mathbb{R}$,*
2. *$\underline{F}_v\left(u, v|y_n^H\right)$ is L^1 in the u-variable, and*
3. *$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u, v|y_n^H\right) du$ converges uniformly,*

then $$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{f}_v \underline{F}_v du \text{ converges uniformly,} \tag{D.30}$$

for every $v \in (-\delta, \delta)$.

Proof. With $\underline{F}_v$ in place of $\underline{F}$, the above Lemma is identical to Lemma D.6 from Appendix D.2 above. The proof is otherwise the same. ■

The next goal is to arrive at an asymptotic result for $\underline{f}_{vv}$, similar to Lemma D.5 in Appendix D.2 above for $\underline{f}_{v}$. Lemma D.11 below is the result for $\underline{f}_{vv}$. In order to simplify the proof of Lemma D.11, we introduce two items. These are Remark D.9, and Lemma D.10 below. Lemma D.10 isolates the main technical needs for the proof of Lemma D.11, making the proof of the latter similar to the proof of Lemma D.5 in Appendix D.2 above.

Remark D.9 *Let $z = re^{i\theta}$ be a complex number with $\theta \in (-\pi, \pi]$. If we maintain the principal branch, then $\sqrt{z} = \sqrt{r}e^{i\frac{\theta}{2}}$, where in this thesis we assume the positive root. In particular, on the principal branch*

$$\left(\sqrt{z}\right)^2 = re^{i\theta} = z. \tag{D.31}$$

Furthermore, letting $\sqrt{z} = a + ib$, we have that

$$\begin{aligned} z = \left(\sqrt{z}\right)^2 &= (a+ib)^2 \\ &= a^2 - b^2 + i2ab. \end{aligned} \tag{D.32}$$

That is, when $\sqrt{z} = a + ib$,

$$\mathrm{Re}\,(z) = a^2 - b^2, \text{ and } \mathrm{Im}\,(z) = 2ab. \tag{D.33}$$

Lemma D.10 *Let $\delta > 0$, and for all $(u, v) \in \mathbb{R} \times (-\delta, \delta)$ define*

$$g^{(\pm)}(u,v) = \frac{\left(\omega\sqrt{1-\rho^2} \pm i\rho\omega\right)^2}{\left(u + \rho\omega v \mp iv\omega\sqrt{1-\rho^2}\right)^2}. \tag{D.34}$$

Then

$$\mathrm{Re}\, g^{(\pm)}(u,v) = \frac{\omega^2\left(\left(1-2\rho^2\right)u^2 - 2\rho\omega uv - \omega^2v^2\right)}{\left(u^2 + 2\rho\omega uv + \omega^2v^2\right)^2}, \tag{D.35}$$

$$\text{and } \mathrm{Im}\, g^{(\pm)}(u,v) = \frac{\pm 2\omega^2\sqrt{1-\rho^2}\left(\rho u^2 + \omega uv\right)}{\left(u^2 + 2\rho\omega uv + \omega^2v^2\right)^2}. \tag{D.36}$$

Proof. First note from equation $(D.34)$ that $g^{(+)}(u,v)$ and $g^{(-)}(u,v)$ are complex conjugates. This implies that

$$\operatorname{Re} g^{(+)} = \operatorname{Re} g^{(-)}, \text{ and } \operatorname{Im} g^{(+)} = -\operatorname{Im} g^{(-)}. \tag{D.37}$$

Thus, it suffices to consider only $\operatorname{Re} g^{(+)}(u,v)$, $\operatorname{Im} g^{(+)}(u,v)$, and in general $g^{(+)}(u,v)$. By equation $(D.34)$ we have

$$g^{(+)}(u,v) = \frac{\left(\omega\sqrt{1-\rho^2} + i\rho\omega\right)^2}{\left(u + \rho\omega v - iv\omega\sqrt{1-\rho^2}\right)^2}. \tag{D.38}$$

Recall that for $z \in \mathbb{C}$, $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$. By treating the denominator of equation $(D.38)$ in this way, we obtain the result

$$g^{(+)}(u,v) = \frac{\left(\omega\sqrt{1-\rho^2} + i\rho\omega\right)^2 \left(u + \rho\omega v + iv\omega\sqrt{1-\rho^2}\right)^2}{\left(u^2 + 2\rho\omega uv + \omega^2 v^2\right)^2}. \tag{D.39}$$

Pursuant to Remark D.9 above, we will take the square root of $g^{(+)}(u,v)$ in equation $(D.39)$, and simplify the numerator. This leads to

$$\sqrt{g^{(+)}(u,v)} = \frac{\omega\sqrt{1-\rho^2}u + i\omega(\rho u + \omega v)}{\left(u^2 + 2\rho\omega uv + \omega^2 v^2\right)}. \tag{D.40}$$

Hence, by the two results in equation $(D.33)$ of Remark D.9,

$$\operatorname{Re} g^{(+)}(u,v) = \frac{\left(\omega\sqrt{1-\rho^2}u\right)^2 - (\omega(\rho u + \omega v))^2}{\left(u^2 + 2\rho\omega uv + \omega^2 v^2\right)^2}, \tag{D.41}$$

$$\text{and } \operatorname{Im} g^{(+)}(u,v) = \frac{2\left(\omega\sqrt{1-\rho^2}u\right)(\omega(\rho u + \omega v))}{\left(u^2 + 2\rho\omega uv + \omega^2 v^2\right)^2}. \tag{D.42}$$

Equations $(D.41)$ and $(D.42)$ simplify to

$$\operatorname{Re} g^{(+)}(u,v) = \frac{\omega^2\left(\left(1-2\rho^2\right)u^2 - 2\rho\omega uv - \omega^2 v^2\right)}{\left(u^2 + 2\rho\omega uv + \omega^2 v^2\right)^2}, \tag{D.43}$$

$$\text{and } \operatorname{Im} g^{(+)}(u,v) = \frac{2\omega^2\sqrt{1-\rho^2}\left(\rho u^2 + \omega uv\right)}{\left(u^2 + 2\rho\omega uv + \omega^2 v^2\right)^2}. \tag{D.44}$$

The two results in equation $(D.37)$ complete the proof. ■

Lemma D.11 *Let $\delta > 0$, and for the SVJ and SVSJ models, assume that $\rho < 0$. For a sufficiently large real number $a > 0$, define the respective neighbourhoods $U_{[a,+\infty)}(+\infty) \subset [a, +\infty)$, and $U_{(-\infty,-a]}(-\infty) \subset (-\infty, -a]$ of $\pm\infty$. Then, for each $v \in (-\delta, \delta)$,*

1. $\operatorname{Re} \underline{f}_{vv} \sim \widetilde{\alpha}_{vv}^{(\pm)}$ *as $u \to \pm\infty$, such that $\widetilde{\alpha}_{vv}^{(+)}$ is monotone and bounded for all $u \in U_{[a,+\infty)}(+\infty)$, and $\widetilde{\alpha}_{vv}^{(-)}$ is monotone and bounded for all $u \in U_{(-\infty,-a]}(-\infty)$.*

2. $\operatorname{Im} \underline{f}_{vv} \sim \widetilde{\beta}_{vv}^{(\pm)}$ *as $u \to \pm\infty$, such that $\widetilde{\beta}_{vv}^{(+)}$ is monotone and bounded for all $u \in U_{[a,+\infty)}(+\infty)$, and $\widetilde{\beta}_{vv}^{(-)}$ is monotone and bounded for all $u \in U_{(-\infty,-a]}(-\infty)$.*

Proof. From equations (4.57), (4.58), and (4.59) in subsection 4.3.1,

$$\underline{C}_{vv} = \frac{\partial^2}{\partial v^2}\underline{C}(u,v) = \frac{-\frac{2\kappa\eta}{\omega^2}K(u)^2}{(1-K(u)\,iv)^2}, \tag{D.45}$$

$$\underline{D}_{v} = \frac{\partial}{\partial v}\underline{D}(u,v) = \frac{R(u)\,i}{(1-K(u)\,iv)^2}, \text{ and} \tag{D.46}$$

$$\underline{D}_{vv} = \frac{\partial^2}{\partial v^2}\underline{D}(u,v) = \frac{-2K(u)\,R(u)}{(1-K(u)\,iv)^3}. \tag{D.47}$$

Moreover, by equation (4.67) in subsection 4.3.1,

$$\underline{f}_{vv} = \underline{C}_{vv} + \frac{a_n b_n \underline{D}_v + a_n b_n^2\left(\underline{D}_v^2 - \underline{D}\underline{D}_{vv}\right)}{(1-b_n\underline{D})^2}. \tag{D.48}$$

However, by Lemma 3.15 from subsection 3.3.2, $\operatorname{Re}\underline{D}(u,v;\tau) \leq 0$, for all $(u,v) \in \mathbb{R}\times(-\delta,\delta)$, for some $\delta > 0$. Thus, since $b_n > 0$, $\operatorname{Re}[1-b_n\underline{D}] > 0$, for all $(u,v) \in \mathbb{R}\times(-\delta,\delta)$. Also, by Lemma C.11 from Appendix C.2, $\operatorname{Re}(1-K(u)\,iv) > 0$, for all $(u,v) \in \mathbb{R}\times(-\delta,\delta)$, for $\delta > 0$ sufficiently small. Furthermore, by Lemma C.9 from Appendix C.2, $K(u) \to 0$ as $u \to \pm\infty$, and by Lemma C.10 from Appendix C.2, $R(u) \to 0$ exponentially, as $u \to \pm\infty$. Thus, altogether we have

$$\underline{f}_{vv} \sim \underline{C}_{vv} = \frac{-\frac{2\kappa\eta}{\omega^2}K(u)^2}{(1-K(u)\,iv)^2}, \text{ as } u \to \pm\infty. \tag{D.49}$$

But, recall also from Lemma C.9 in Appendix C.2 that

$$K(u) \sim \frac{\pm\omega\sqrt{1-\rho^2}+i\rho\omega}{u}, \text{ as } u \to \pm\infty. \tag{D.50}$$

By substituting $(D.50)$ into $(D.49)$ and simplifying, as $u \to \pm\infty$ we obtain

$$\underline{f}_{vv} \sim \frac{\frac{-2\kappa\eta}{\omega^2}\left(\omega\sqrt{1-\rho^2} \pm i\rho\omega\right)^2}{\left(u+\rho\omega v \mp iv\omega\sqrt{1-\rho^2}\right)^2} \tag{D.51}$$

$$= \frac{-2\kappa\eta}{\omega^2} g^{(\pm)}(u,v), \tag{D.52}$$

where $g^{(\pm)}(u,v)$ is defined by equation $(D.34)$ from Lemma D.10 above. Thus, it follows from equation $(D.52)$ and Lemma D.10 above that

$$\widetilde{\alpha}_{vv}^{(\pm)} = \frac{-2\kappa\eta\left(\left(1-2\rho^2\right)u^2 - 2\rho\omega uv - \omega^2v^2\right)}{\left(u^2+2\rho\omega uv+\omega^2v^2\right)^2}$$

$$\sim \frac{-2\kappa\eta\left(1-2\rho^2\right)}{u^2}, \text{ as } u \to \pm\infty, \tag{D.53}$$

$$\text{and } \widetilde{\beta}_{vv}^{(\pm)} = \frac{\mp 4\kappa\eta\sqrt{1-\rho^2}\left(\rho u^2+\omega uv\right)}{\left(u^2+2\rho\omega uv+\omega^2v^2\right)^2}$$

$$\sim \frac{\mp 4\kappa\eta\rho\sqrt{1-\rho^2}}{u^2}, \text{ as } u \to \pm\infty. \tag{D.54}$$

Let $v \in (-\delta,\delta)$. Then, by the result $(D.53)$, $\widetilde{\alpha}_{vv}^{(+)}$ is monotone and bounded for all $u \in U_{[a,+\infty)}(+\infty)$. Moreover, $\widetilde{\alpha}_{vv}^{(-)}$ is also monotone and bounded for all $u \in U_{(-\infty,-a]}(-\infty)$. Likewise by the result $(D.54)$, $\widetilde{\beta}_{vv}^{(+)}$ is monotone and bounded for all $u \in U_{[a,+\infty)}(+\infty)$, and $\widetilde{\beta}_{vv}^{(-)}$ is monotone and bounded for all $u \in U_{(-\infty,-a]}(-\infty)$. This completes the proof. ■

Lemma D.12 below follows easily from Lemma D.11 above, given that the proof is similar to the proof of Lemma D.6 from Appendix D.2. By combining the results from Lemma D.8 above, and Lemma D.12 below, we will be able to obtain the result that for all $v \in (-\delta,\delta)$, for some $\delta > 0$,

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_{vv}\left(u,v|y_n^H\right) du \text{ converges uniformly.} \tag{D.55}$$

Lemma D.12 *Let $\delta > 0$, and assume that the hypotheses and definitions of Lemma D.11 above hold. If in addition, for every $v \in (-\delta, \delta)$,*

1. $\underline{F}\left(u, v|y_n^H\right)$ *is uniformly continuous in u, for all $u \in \mathbb{R}$,*
2. $\underline{F}\left(u, v|y_n^H\right)$ *is L^1 in the u-variable, and*
3. $\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du$ *converges uniformly,*

then $$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{f}_{vv} \underline{F} du \text{ converges uniformly,} \tag{D.56}$$

for every $v \in (-\delta, \delta)$.

Proof. With the asymptotic results from Lemma D.11 for $\underline{f}_{vv}$ in place of similar results from Lemma D.5 for $\underline{f}_v$, the above Lemma is identical to Lemma D.6 from Appendix D.2. The proof is otherwise the same. ■

Theorem D.13 (Differentiation II) *Assume that $\rho < 0$ in the SVJ and SVSJ models, noting that this is sufficient for the asymptotic forms in both Lemma D.5 from Appendix D.2, and Lemma D.11 above. From equation $(D.1)$ at the start of this Appendix define*

$$\begin{aligned} \underline{F}\left(u, v|y_n^H\right) &= e^{\underline{C}(u,v)} G_{n|n}\left[\underline{D}\left(u, v\right)\right] \\ &= e^{\underline{C}(u,v;\Delta t)}\left[1 - b_n \underline{D}\left(u, v; \Delta t\right)\right]^{-a_n}, \qquad \text{(D.57)} \\ \text{with } \underline{F}_v\left(u, v|y_n^H\right) &= \frac{\partial}{\partial v} \underline{F}\left(u, v|y_n^H\right), \\ \text{and } \underline{F}_{vv}\left(u, v|y_n^H\right) &= \frac{\partial^2}{\partial v^2} \underline{F}\left(u, v|y_n^H\right). \end{aligned}$$

Let $\delta > 0$. The main hypotheses are that for every $v \in (-\delta, \delta)$, $\underline{F}\left(u, v|y_n^H\right)$ is uniformly continuous in u, for all $u \in \mathbb{R}$, both $\underline{F}\left(u, v|y_n^H\right)$ and $\underline{F}_v\left(u, v|y_n^H\right)$ are L^1 in the u-variable, and the improper integral

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du \text{ converges uniformly.} \tag{D.58}$$

If Lemma D.5, Lemma D.11, and the main hypotheses as stated above each hold, then for every $v \in (-\delta, \delta)$,

1. $\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_{vv}\left(u, v|y_n^H\right) du$ *converges uniformly,*
2. $\frac{\partial^2}{\partial v^2} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_{vv}\left(u, v|y_n^H\right) du.$

Proof. Given $\rho < 0$, the main hypotheses as stated above are sufficient for Theorem D.7 from Appendix D.2. This implies that for all $v \in (-\delta, \delta)$,

a) $\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u, v|y_n^H\right) du$ converges uniformly,

b) $\underline{F}_v\left(u, v|y_n^H\right)$ is uniformly continuous in u, for all $u \in \mathbb{R}$, and

c) $\frac{\partial}{\partial v} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}\left(u, v|y_n^H\right) du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v\left(u, v|y_n^H\right) du.$

But, given that Lemma D.5 holds, we have that a), b), and the main hypothesis that $\underline{F}_v$ is L^1 in u, satisfy Lemma D.8, and this implies that for all $v \in (-\delta, \delta)$,

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{f}_v \underline{F}_v du \text{ converges uniformly.} \tag{D.59}$$

Moreover, given that Lemma D.11 holds, the main hypotheses of the current theorem satisfy Lemma D.12, and this implies that for all $v \in (-\delta, \delta)$,

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{f}_{vv} \underline{F} du \text{ converges uniformly.} \tag{D.60}$$

Hence, by equation $(D.29)$, the pair of results $(D.59)$ and $(D.60)$ imply that for all $v \in (-\delta, \delta)$,

$$\int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_{vv}\left(u, v|y_n^H\right) du \text{ converges uniformly,} \tag{D.61}$$

and this proves item *1.* To complete the proof, we need to show that for all $v \in (-\delta, \delta)$,

$$\frac{\partial^2}{\partial v^2} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F} du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_{vv} du. \tag{D.62}$$

By differentiating both sides of c) above w.r.t. v, outside of the integrals, we obtain that for all $v \in (-\delta, \delta)$,

$$\frac{\partial^2}{\partial v^2} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F} du = \frac{\partial}{\partial v} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v du. \tag{D.63}$$

Hence, by equation $(D.63)$, it is sufficient to show that for all $v \in (-\delta, \delta)$,

$$\frac{\partial}{\partial v} \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_v du = \int_{-\infty}^{+\infty} e^{-iuy_{n+1}} \underline{F}_{vv} du. \tag{D.64}$$

Moreover, by Proposition D.2 from Appendix D.1, a) above and the result $(D.61)$ are sufficient for $(D.64)$. Since this implies item *2.*, this completes the proof. ■

Appendix E

Path Simulation

In this Appendix we present the path simulation methods that were used to simulate the Heston, SVJ, and SVSJ daily log-returns in Chapter 4 on the Bates (2006) AML method. Appendix E.1 covers the simulation of Heston model daily log-returns by first simulating the variance process at a faster than daily rate, as suggested in Bates (2006), pp. 957-58. Appendix E.2 treats the SVJ model by adding independent jump increments to the Heston model log-returns, where the jumps are either Merton, Variance Gamma, or Meixner. Then Appendix E.3 covers the SVSJ Merton model by adding conditionally independent time changed Merton jump increments to the Heston model log-returns, given the variance.

E.1 Simulating Heston Model Log-Returns

Let $W_t^{(1)}$ and $W_t^{(2)}$ be independent standard Brownian motions. Then the P-dynamics of the Heston model log-price may be written as

$$dY_t = \left(\mu_0 - \frac{1}{2}\sigma_t^2\right) dt + \sigma_t \left[\rho dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)}\right] \tag{E.1}$$

$$d\sigma_t^2 = \kappa\left(\eta - \sigma_t^2\right) dt + \omega\sigma_t dW_t^{(1)}. \tag{E.2}$$

Equations $(E.1)$ and $(E.2)$ above may be derived from equation (2.59) in subsection 2.5.2. Next we consider the following definitions. Given the path of σ_t^2 over $[t, t+\Delta t]$, and given Y_t, let

$$\Delta Y_{t+\Delta t} = Y_{t+\Delta t} - Y_t, \tag{E.3}$$

$$\Delta \sigma_{t+\Delta t}^2 = \sigma_{t+\Delta t}^2 - \sigma_t^2, \tag{E.4}$$

$$\Sigma_t \Delta t = \int_t^{t+\Delta t} \sigma_s^2 ds, \text{ and} \tag{E.5}$$

$$B_t = \frac{1}{\omega}\left[\Delta \sigma_{t+\Delta t}^2 - \kappa\eta\Delta t + \kappa\Sigma_t \Delta t\right]. \tag{E.6}$$

By the main result of Broadie and Kaya (2006), see §3.3, pp. 221-22, we have

$$\Delta Y_{t+\Delta t} = \left(\mu_0 - \frac{1}{2}\Sigma_t\right)\Delta t + \rho B_t + \sqrt{(1-\rho^2)\,\Sigma_t \Delta t} Z_{t+\Delta t}, \tag{E.7}$$

$$Z_{t+\Delta t} \sim N(0,1).$$

To obtain equation $(E.7)$ we only need to simulate $\Delta \sigma_{t+\Delta t}^2 = \sigma_{t+\Delta t}^2 - \sigma_t^2$, and

$$\Sigma_t = \frac{1}{\Delta t}\int_t^{t+\Delta t} \sigma_s^2 ds. \tag{E.8}$$

But, Σ_t in equation $(E.8)$ can be approximated by the average intraday variance process, see Bates (2006), p. 958. Also, simulating the intraday variance will yield values for $\sigma_{t+\Delta t}^2$ and σ_t^2. In this thesis we used $N_s = 78$ intraday subsamples per day. This gives $N = 252 N_s$ subsamples, assuming a 252 day year, with subsampling frequency $\Delta u = \frac{1}{N}$. Also assume that there are T daily log-returns to simulate. Note from subsection 4.2.2 that for AML we assume σ_0^2 is generated from a $\Gamma\left(\frac{2\kappa\eta}{\omega^2}, \frac{\omega^2}{2\kappa}\right)$ distribution, so that when the Feller condition $2\kappa\eta > \omega^2$ also holds, the variance process is strictly stationary and strongly ergodic. In the following algorithm for simulating the CIR process by its scaled noncentral chi-squared transition law we assume for simplicity that $4\kappa\eta > \omega^2$. The case $4\kappa\eta \leq \omega^2$ is provided for in Glasserman (2004), see p. 124, if needed.

Algorithm E.1 (Glasserman (2004), p. 124) *Assume* $d = \frac{4\kappa\eta}{\omega^2} > 1$, *and that* σ_0^2 *is input. To simulate the CIR process* σ_s^2 *on a discrete time grid* $u_k = k\Delta u$, *for* $k = 1, 2, ...NT$,

Begin For $k = 0, 1, ..., NT - 1$

Set $c = \frac{\omega^2}{4\kappa}\left(1 - e^{-\kappa\Delta u}\right)$, *and set* $b = \sigma_{u_k}^2 \frac{e^{-\kappa\Delta u}}{c}$.

Generate $Z \sim N(0, 1)$.

Generate $X \sim \chi_{d-1}^2$.

Set $\sigma_{u_{k+1}}^2 = c\left[\left(Z + \sqrt{b}\right)^2 + X\right]$.

End For

$$\text{Then, } \Sigma_t \approx \frac{1}{N_s}\sum\nolimits_{j=0}^{N_s-1} \sigma_{t+j\Delta u}^2. \tag{E.9}$$

E.2 Simulating SVJ Model Log-Returns

From equation $(A.1)$ in Appendix A the SVJ model may be written as

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2\right)dt + \sigma_t dW_t^{(S)} + dX_t - \psi_X(-i)\,dt, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)}, \\ \rho dt &= E\left[dW_t^{(S)} dW_t^{(V)}\right]. \end{aligned} \tag{E.10}$$

Equation $(E.10)$ is evidently the Heston process from Appendix E.1 above, plus an independent compensated Lévy process,

$$\widehat{X}_t = X_t - \psi_X(-i)\,t. \tag{E.11}$$

Thus, the SVJ log-return is simply a Heston log-return plus $\Delta\widehat{X}_{t+\Delta t}$ where

$$\begin{aligned} \Delta\widehat{X}_{t+\Delta t} &= \widehat{X}_{t+\Delta t} - \widehat{X}_t = \widehat{X}(t + \Delta t - t) \\ &= \widehat{X}(\Delta t) = X(\Delta t) - \psi_X(-i)\,\Delta t. \end{aligned}$$

Merton Jumps

Recall from equation (2.24) in subsection 2.3.1 that for Merton jumps

$$\psi_X(-i) = \lambda\left(\exp\left(\beta + \frac{1}{2}\alpha^2\right) - 1\right). \tag{E.12}$$

Algorithm E.2 (Merton Jumps) *For the Merton jump process X the following algorithm simulates one increment $\Delta X_{t+\Delta t} = X(\Delta t)$.*

Generate $N_{t+\Delta t} \sim Poiss(\lambda \Delta t)$.

Generate $Z_{t+\Delta t} \sim N(0,1)$.

Set $\Delta X_{t+\Delta t} = \beta N_{t+\Delta t} + \alpha\sqrt{N_{t+\Delta t}} Z_{t+\Delta t}$.

See Glasserman (2004), §3.5, pp. 138-139.

Variance Gamma Jumps

Variance Gamma jumps may be expressed as a time changed Brownian motion. Recall from equation (2.32) in subsection 2.3.2 that for Variance Gamma jumps of the Huang and Wu (2004) type

$$\psi_X(-i) = -\lambda \log\left(1 - \beta - \frac{1}{2}\alpha^2\right), \tag{E.13}$$

and from equation (2.29) in subsection 2.3.2 the Huang and Wu (2004) version of the process has a time change given by

$$G_t \sim \Gamma(\lambda t, 1). \tag{E.14}$$

Algorithm E.3 (Variance Gamma Jumps) *For the Variance Gamma jump process X of Huang and Wu (2004) type the following algorithm simulates one increment $\Delta X_{t+\Delta t} = X(\Delta t)$.*

Generate $G_{t+\Delta t} \sim \Gamma(\lambda \Delta t, 1)$.

Generate $Z_{t+\Delta t} \sim N(0,1)$.

Set $\Delta X_{t+\Delta t} = \beta G_{t+\Delta t} + \alpha\sqrt{G_{t+\Delta t}} Z_{t+\Delta t}$.

Adapted from Glasserman (2004), §3.5, pp. 143-144.

Meixner Jumps

Recall from equation (2.40) in subsection 2.3.3 that for Meixner jumps

$$\psi_X(-i) = 2\lambda \log\left(\frac{\cos\left(\frac{\beta}{2}\right)}{\cos\left(\frac{\alpha+\beta}{2}\right)}\right). \tag{E.15}$$

For the simulation of a single Meixner increment, see Algorithm 8.4 from subsection 8.2.3, with $\tau = \Delta t$.

E.3 Simulating SVSJ Model Log-Returns

From equation $(A.2)$ in Appendix A the SVSJ model may be written as

$$\begin{aligned} dY_t &= \left(\mu_0 - \frac{1}{2}\sigma_t^2\right)dt + \sigma_t dW_t^{(S)} + dX(V_t) - \sigma_t^2\psi_X(-i)\,dt, \\ d\sigma_t^2 &= \kappa\left(\eta - \sigma_t^2\right)dt + \omega\sigma_t dW_t^{(V)},\ E\left[dW_t^{(S)}dW_t^{(V)}\right] = \rho dt, \\ dV_t &= \sigma_t^2 dt. \end{aligned} \tag{E.16}$$

Similar to equation $(E.10)$ for the SVJ model in Appendix E.2 above, equation $(E.16)$ for the SVSJ model is simply a Heston process plus jumps, and with a few modifications we may simulate on this basis. The difference is that the compensated jump process,

$$\widehat{X}(V_t) = \left[X(V_t) - \psi_X(-i)\,V_t\right], \tag{E.17}$$

is time changed by the integrated variance process from the Heston model. However, the SVSJ jumps are conditionally independent given the variance process. Moreover, the variance process is dictated by the Heston process. Thus, we can simply simulate a Heston process, and pass the simulated variance to the time changed jumps.

Recall from equation $(E.8)$ in Appendix E.1 that

$$\Sigma_t = \frac{1}{\Delta t}\int_t^{t+\Delta t} \sigma_s^2 ds. \tag{E.18}$$

Hence, in continuous time the integrated variance $V_{t+\Delta t}$ has a decomposition given by

$$\begin{aligned} V_{t+\Delta t} &= \int_0^t \sigma_s^2 ds + \int_t^{t+\Delta t} \sigma_s^2 ds \\ &= V_t + \Sigma_t \Delta t. \end{aligned} \tag{E.19}$$

Therefore, the continuous time increment to the compensated time changed jumps in the SVSJ model is given by

$$\begin{aligned} \Delta \widehat{X}\left(V_{t+\Delta t}\right) &= \widehat{X}\left(V_{t+\Delta t}\right) - \widehat{X}\left(V_t\right) \overset{law}{=} \widehat{X}\left(V_{t+\Delta t} - V_t\right) \\ &\overset{law}{=} \widehat{X}\left(\Sigma_t \Delta t\right), \text{ by equation } (E.19) \\ &= X\left(\Sigma_t \Delta t\right) - \psi_X\left(-i\right) \Sigma_t \Delta t. \end{aligned} \tag{E.20}$$

In particular, in continuous time, given Σ_t, the increment to the time changed jump process,

$$\Delta X\left(V_{t+\Delta t}\right) \overset{law}{=} X\left(\Sigma_t \Delta t\right), \tag{E.21}$$

has the distribution of a Merton jump with time increment $\Sigma_t \Delta t$. Thus, we can simulate from $\Delta X\left(V_{t+\Delta t}\right)$ by generating $N_{t+\Delta t} \sim Poiss\left(\lambda \Sigma_t \Delta t\right)$ in place of $N_{t+\Delta t} \sim Poiss\left(\lambda \Delta t\right)$ in Algorithm E.2 for Merton jumps. However, note that this is subject to the accuracy of the approximation

$$\Sigma_t \approx \frac{1}{N_s} \sum\nolimits_{j=0}^{N_s - 1} \sigma_{t+j\Delta u}^2,$$

made in equation $(E.9)$ from Appendix E.1 above on the simulation of log-returns for the Heston model.

Bibliography

[1] Acworth, P. A., M. Broadie, and P. Glasserman: 1998, 'A comparison of some Monte Carlo and quasi Monte Carlo techniques for option pricing'. In: *Monte Carlo and Quasi-Monte Carlo Methods 1996*. pp. 1–18.

[2] Aït-Sahalia, Y. and J. Jacod: 2012, 'Analyzing the spectrum of asset returns: jump and volatility components in high frequency data'. *Journal of Economic Literature* **50**(4), 1007–1050.

[3] Aït-Sahalia, Y. and R. Kimmel: 2007, 'Maximum likelihood estimation of stochastic volatility models'. *Journal of Financial Economics* **83**(2), 413–452.

[4] Albrecher, H., P. Mayer, W. Schoutens, and J. Tistaert: 2007, 'The little Heston trap'. *Wilmott Magazine* **1**, 73–82.

[5] Andersen, L. and V. Piterbarg: 2007, 'Moment explosions in stochastic volatility models'. *Finance and Stochastics* **11**(1), 29–50.

[6] Andersen, T. G., L. Benzoni, and J. Lund: 2002, 'An empirical investigation of continuous-time equity return models'. *The Journal of Finance* **57**(3), 1239–1284.

[7] Applebaum, D.: 2009, *Lévy Processes and Stochastic Calculus*. Cambridge University Press, 2nd edition.

[8] Avellaneda, M.: 1998, 'Minimum-relative-entropy calibration of asset-pricing models'. *International Journal of Theoretical and Applied Finance* **1**(4), 447–472.

[9] Avellaneda, M., R. Buff, C. Friedman, N. Grandechamp, L. Kruk, and J. Newman: 2001, 'Weighted Monte Carlo: a new technique for calibrating asset-pricing models'. *International Journal of Theoretical and Applied Finance* **4**(1), 91–119.

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

[11] Barndorff-Nielsen, O.: 1997, 'Normal inverse Gaussian distributions and stochastic volatility modelling'. *Scandinavian Journal of Social Statistics* **24**(1), 1–13.

[12] Barndorff-Nielsen, O.: 1998, 'Processes of normal inverse Gaussian type'. *Finance and Stochastics* **2**(1), 41–68.

[13] Barndorff-Nielsen, O. and N. Shephard: 2001, 'Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics'. *Journal of the Royal Statistical Society: Series B* **63**(2), 167–241.

[14] Barndorff-Nielsen, O. and A. Shiryaev: 2010, *Change of Time and Change of Measure*, Vol. 13 of *Advanced Series on Statistics & Applied Probability*. World Scientific, New Jersey.

[15] Bartlett, M.: 1938, 'The characteristic function of a conditional statistic'. *Journal of the London Mathematical Society* **1**(1), 62–67.

[16] Bates, D.: 1996, 'Jumps and stochastic volatility: exchange rate processes implicit in Deutsche mark options'. *Review of Financial Studies* **9**(1), 69–107.

[17] Bates, D.: 2000, 'Post-'87 crash fears in the S&P 500 futures option market'. *Journal of Econometrics* **94**(1-2), 181–238.

[18] Bates, D.: 2006, 'Maximum likelihood estimation of latent affine processes'. *Review of Financial Studies* **19**(3), 909–965.

[19] Bates, D.: 2012, 'U.S. stock market crash risk, 1926-2010'. *Journal of Financial Economics* **105**(2), 229–259.

[20] Bernard, C., Z. Cui, and D. McLeish: 2014, 'On the martingale property in stochastic volatility models based on time-homogeneous diffusions'. *Mathematical Finance*. DOI: 10.1111/mafi.12084.

[21] Billingsley, P.: 1995, *Probability and Measure*. John Wiley & Sons, New York, 3rd edition.

[22] Björk, T.: 2009, *Arbitrage Theory in Continuous Time*. Oxford University Press, 3rd edition.

[23] Black, F.: 1976, 'Studies of stock price volatility changes'. In: *Proceedings of the American Statistical Association 1976, Business and Economics Section (American Statistical Association, Washington, D.C.)*. pp. 177–181.

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

[25] Briggs, W. L. and V. E. Henson: 1995, *The DFT*. SIAM.

[26] Broadie, M. and O. Kaya: 2006, 'Exact simulation of stochastic volatility and other affine jump diffusion processes'. *Operations Research* **54**(2), 217–231.

[27] Brown, J. and R. Churchill: 2004, *Complex Variables and Applications*. McGraw-Hill, Boston, 7th edition.

[28] Burden, R. and J. Faires: 1997, *Numerical Analysis*. Brooks Cole, Pacific Grove, CA, 6th edition.

[29] Carr, P., H. Geman, D. Madan, and M. Yor: 2002, 'The fine structure of asset returns: an empirical investigation'. *Journal of Business* **75**(2), 305–333.

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

[31] Carr, P. and D. Madan: 1999, 'Option valuation using the fast Fourier transform'. *Journal of Computational Finance* **2**(4), 61–73.

[32] Carr, P. and D. Madan: 2009, 'Saddlepoint methods for option pricing'. *Journal of Computational Finance* **13**(1), 49–61.

[33] Carr, P. and L. Wu: 2003, 'What type of process underlies options? A simple robust test'. *The Journal of Finance* **58**(6), 2581–2610.

[34] Carr, P. and L. Wu: 2004, 'Time-changed Lévy processes and option pricing'. *Journal of Financial Economics* **71**(1), 113–141.

[35] Chernov, M. and E. Ghysels: 2000, 'A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation'. *Journal of Financial Economics* **56**(3), 407–458.

[36] Christoffersen, P., S. Heston, and K. Jacobs: 2009, 'The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well'. *Management Science* **55**(12), 1914–1932.

[37] Cont, R. and P. Tankov: 2004a, *Financial Modelling with Jump Processes*. Chapman & Hall, Boca Raton, Florida.

[38] Cont, R. and P. Tankov: 2004b, 'Nonparametric calibration of jump-diffusion processes'. *Journal of Computational Finance* **7**(3), 1–49.

[39] Cont, R. and P. Tankov: 2006, 'Retrieving Lévy processes from option prices: Regularization of an ill-posed inverse problem'. *SIAM Journal on Control and Optimization* **45**(1), 1–25.

[40] Cooley, J. W. and J. W. Tukey: 1965, 'An algorithm for the machine calculation of complex Fourier series'. *Mathematics of Computation* **19**(90), 297–301.

[41] Cover, T. M. and J. A. Thomas: 1991, *Elements of Information Theory*. John Wiley & Sons, New York.

[42] Cox, J., J. Ingersoll Jr, and S. Ross: 1985, 'A theory of the term structure of interest rates'. *Econometrica* **53**(2), 385–408.

[43] Davis, P. J. and P. Rabinowitz: 1984, *Methods of Numerical Integration*. Academic Press, 2nd edition.

[44] Devroye, L.: 1986, *Non-Uniform Random Variate Generation*. Springer-Verlag, New York.

[45] Duffie, D., D. Filipović, and W. Schachermayer: 2003, 'Affine processes and applications in finance'. *The Annals of Applied Probability* **13**(3), 984–1053.

[46] Duffie, D., J. Pan, and K. Singleton: 2000, 'Transform analysis and asset pricing for affine jump-diffusions'. *Econometrica* **68**(6), 1343–1376.

[47] Engl, H., M. Hanke, and A. Neubauer: 1996, *Regularization of Inverse Problems*, Vol. 375. Kluwer, London.

[48] Eraker, B., M. Johannes, and N. Polson: 2003, 'The impact of jumps in equity index volatility and returns'. *Journal of Finance* **58**(3), 1269–1300.

[49] Fang, F. and C. Oosterlee: 2008, 'A novel pricing method for European options based on Fourier-cosine series expansions'. *SIAM Journal on Scientific Computing* **31**(2), 826–848.

[50] Feller, W.: 1951, 'Two singular diffusion problems'. *Annals of Mathematics* **54**(1), 173–182.

[51] Frittelli, M.: 2000, 'The minimal entropy martingale measure and the valuation problem in incomplete markets'. *Mathematical Finance* **10**(1), 39–52.

[52] Fujiwara, T. and Y. Miyahara: 2003, 'The minimal entropy martingale measures for geometric Lévy processes'. *Finance and Stochastics* **7**(4), 509–531.

[53] Glasserman, P.: 2004, *Monte Carlo Methods in Financial Engineering*. Springer, New York.

[54] Gradshteyn, I. and I. Ryzhik: 1965, *Table of Integrals, Series, and Products*. Adademic Press, New York.

[55] Grigelionis, B.: 1999, 'Processes of Meixner type'. *Lithuanian Mathematical Journal* **39**(1), 33–41.

[56] Grigoletto, M. and C. Provasi: 2009, 'Simulation and estimation of the Meixner distribution'. *Communications in Statistics-Simulation and Computation* **38**(1), 58–77.

[57] Grimmett, G. and D. Stirzaker: 2001, *Probability and Random Processes*. Oxford University Press, 3rd edition.

[58] He, X.-J. and S.-P. Zhu: 2016, 'Pricing European options with stochastic volatility under the minimal entropy martingale measure'. *European Journal of Applied Mathematics* **27**(2), 233–247.

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

[60] Hobson, D.: 2004, 'Stochastic volatility models, correlation, and the q-optimal measure'. *Mathematical Finance* **14**(4), 537–556.

[61] Honoré, P.: 1998, 'Pitfalls in estimating jump-diffusion models'. Working Paper 18, Centre for Analytical Finance, University of Aarhus School of Business.

[62] Huang, J.-Z. and L. Wu: 2004, 'Specification analysis of option pricing models based on time-changed Lévy processes'. *Journal of Finance* **59**(3), 1405–1439.

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

[64] Hull, J. C.: 2009, *Options, Futures, and other Derivatives*. Prentice Hall, New Jersey, 7th edition.

[65] Imai, J. and K. S. Tan: 2006, 'A general dimension reduction technique for derivative pricing'. *Journal of Computational Finance* **10**(2), 129–155.

[66] Jacod, J. and A. Shiryaev: 2003, *Limit Theorems for Stochastic Processes*. Springer-Verlag, Heidelberg, 2nd edition.

[67] Jeanblanc, M., M. Yor, and M. Chesney: 2009, *Mathematical Methods for Financial Markets*. Springer.

[68] Joe, S. and F. Y. Kuo: 2003, 'Remark on algorithm 659: Implementing Sobol's quasirandom sequence generator'. *ACM Transactions on Mathematical Software (TOMS)* **29**(1), 49–57.

[69] Johnson, N., S. Kotz, and N. Balakrishnan: 2004, *Continuous Univariate Distributions Volume 1*. John Wiley & Sons, New York, 2nd edition.

[70] Kahl, C., M. Günther, and T. Rossberg: 2008, 'Structure preserving stochastic integration schemes in interest rate derivative modeling'. *Applied Numerical Mathematics* **58**(3), 284–295.

[71] Kahl, C. and P. Jäckel: 2005, 'Not-so-complex logarithms in the Heston model'. *Wilmott Magazine* **19**, 94–103.

[72] Kallsen, J.: 2002, 'Utility-based derivative pricing in incomplete markets'. In: H. Geman, D. Madan, S. Pliska, and T. Vorst (eds.): *Mathematical Finance - Bachelier Congress 2000*. pp. 313–338.

[73] Kapur, J. N. and H. K. Kesavan: 1992, *Entropy Optimization Principles with Applications*. Academic Press.

[74] Karlin, S. and H. M. Taylor: 1981, *A Second Course in Stochastic Processes*. Academic Press, New York.

[75] Kawata, T.: 1972, *Fourier Analysis in Probability Theory*, Vol. 15. Academic Press, New York.

[76] Kloeden, P. and A. Neuenkirch: 2013, 'Convergence of numerical methods for SDEs in finance'. In: P. Gerstner, Thomas ans Kloeden (ed.): *Recent Developments in Computational Finance: Foundations, Algorithms, and Applications*, Vol. 14.

[77] Körner, T. W.: 1989, *Fourier Analysis*. Cambridge University Press.

[78] Kullback, S. and R. A. Leibler: 1951, 'On information and sufficiency'. *The Annals of Mathematical Statistics* **22**(1), 79–86.

[79] Lehmann, E. L. and G. Casella: 1998, *Theory of Point Estimation*. Springer, New York, 2nd edition.

[80] Lemieux, C.: 2008, *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer.

[81] Li, H., M. Wells, and C. L. Yu: 2008, 'A Bayesian analysis of return dynamics with Lévy jumps'. *Review of Financial Studies* **21**(5), 2345–2378.

[82] Li, H., M. Wells, and C. L. Yu: 2011, 'MCMC estimation of Lévy jump models using stock and option prices'. *Mathematical Finance* **21**(3), 383–422.

[83] Lord, R. and C. Kahl: 2006, 'Why the rotation count algorithm works'. Technical report, Tinbergen Institute Discussion Paper. SSRN: 921335.

[84] Lord, R. and C. Kahl: 2010, 'Complex logarithms in Heston-like models'. *Mathematical Finance* **20**(4), 671–694.

[85] Lysy, M.: 2012, 'The Method of Batch Inference for Multivariate Diffusions'. Ph.D. thesis, Harvard University.

[86] Madan, D., P. Carr, and E. Chang: 1998, 'The variance gamma process and option pricing'. *European Finance Review* **2**(1), 79–105.

[87] Madan, D. and M. Yor: 2008, 'Representing the CGMY and Meixner Lévy processes as time changed Brownian motions'. *Journal of Computational Finance* **12**(1), 27–47.

[88] Mandelbrot, B.: 1963, 'The variation of certain speculative prices'. *Journal of Business* **36**(4), 394–419.

[89] McLeish, D.: 1982, 'A robust alternative to the normal distribution'. *Canadian Journal of Statistics* **10**(2), 89–102.

[90] McLeish, D.: 2005, *Monte Carlo Simulation and Finance*. Wiley Finance, Hoboken, NJ.

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

[92] Michael, J., W. Schucany, and R. Haas: 1976, 'Generating random variates using transformations with multiple roots'. *The American Statistician* **30**(2), 88–90.

[93] Nicolato, E. and E. Venardos: 2003, 'Option pricing in stochastic volatility models of the Ornstein-Uhlenbeck type'. *Mathematical Finance* **13**(4), 445–466.

[94] Nocedal, J. and S. Wright: 2006, *Numerical Optimization*. Springer, New York, 2nd edition.

[95] Pan, J.: 2002, 'The jump-risk premia implicit in options: Evidence from an integrated time-series study'. *Journal of Financial Economics* **63**(1), 3–50.

[96] Reesor, R.: 2001, 'Relative Entropy, Distortion, the Bootstrap, and Risk'. Ph.D. thesis, University of Waterloo.

[97] Rheinländer, T. and G. Steiger: 2006, 'The minimal entropy martingale measure for general Barndorff-Nielsen/Shephard models'. *The Annals of Applied Probability* **16**(3), 1319–1351.

[98] Royden, H. and P. Fitzpatrick: 2010, *Real Analysis*. Prentice Hall, Upper Saddle River, 4th edition.

[99] Rudin, W.: 1987, *Real and Complex Analysis*. McGraw-Hill, Boston, 3rd edition.

[100] Sato, K.-I.: 1999, *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

[101] Schoutens, W.: 2003, *Lévy Processes in Finance: Pricing Financial Derivatives*. Wiley, Chichester, West Sussex.

[102] Schoutens, W. and J. Teugels: 1998, 'Lévy processes, polynomials and martingales'. *Communications in Statistics: Stochastic Models* **14**(1-2), 335–349.

[103] Seber, G. and C. Wild: 2003, *Nonlinear Regression*. Wiley, New York.

[104] Stuart, A. and K. Ord: 1994, *Kendall's Advanced Theory of Statistics Vol. 1: Distribution Theory*. Wiley, Chichester, West Sussex, 6th edition.

[105] Stutzer, M.: 1996, 'A simple nonparametric approach to derivative security valuation'. *The Journal of Finance* **51**(5), 1633–1652.

[106] Vasiček, O.: 1977, 'An equilibrium characterization of the term structure'. *Journal of Financial Economics* **5**(2), 177–188.

[107] Walker, J. S.: 1996, *Fast Fourier Transforms*. CRC Press, 2nd edition.

[108] Whaley, R. E.: 2000, 'The investor fear gauge'. *The Journal of Portfolio Management* **26**(3), 12–17.

[109] Willard, G. A.: 1997, 'Calculating prices and sensitivities for path-independent derivatives securities in multifactor models'. *Journal of Derivatives* **5**(1), 45–61.

[110] Zorich, V. A.: 2004, *Mathematical Analysis II*. Springer-Verlag, Berlin. First English edition translated from Russian by Roger Cooke.