ESSAYS ON AMERICAN OPTIONS PRICING UNDER LEVY MODELS WITH STOCHASTIC VOLATILITY AND JUMPS

ESSAYS ON AMERICAN OPTIONS PRICING UNDER LEVY MODELS WITH STOCHASTIC VOLATILITY AND JUMPS

A Dissertation Presented

by

YE CHEN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2019

Management

ESSAYS ON AMERICAN OPTIONS PRICING UNDER LEVY MODELS WITH STOCHASTIC VOLATILITY AND JUMPS

A Dissertation Presented

by

YE CHEN

Approved as to form and content by:

Sanjay Nawalkha, Chair

Nikunj Kapadia, Member

Matthew Linn, Member

Michael Sullivan, Outside Member, Department of Mathematics and Statistics, UMass Amherst

George R. Milne, Ph.D. Program Director
Isenberg School of Management

DEDICATION

To my beloved family, for their love, support, and encouragement.

ACKNOWLEDGMENTS

I am deeply indebted to my supervisor, Dr. Sanjay Nawalkha, for his guidance and support through each stage of the process. I am extremely grateful to my dissertation committee members, Dr. Nikunj Kapadia, Dr. Matthew Linn, and Dr. Michael Sullivan, for their valuable insights and perspectives on my research. I also would like to thank the faculty members, staff, and PhD students of the Department of Finance for their mentoring and support.

Most importantly, I would like to extend my deepest gratitude to my family, for their unwavering support and encouragement during these years.

ABSTRACT

ESSAYS ON AMERICAN OPTIONS PRICING UNDER LEVY MODELS WITH STOCHASTIC VOLATILITY AND JUMPS

SEPTEMBER 2019

YE CHEN, B.S., WUYI UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Sanjay Nawalkha

In "A Multi-dimensional Transform for Pricing American Options Under Stochastic Volatility Models", we present a new transform-based approach for pricing American options under low-dimensional stochastic volatility models which can be used to construct multi-dimensional path-independent lattices for all low-dimensional stochastic volatility models given in the literature, including SV, SV2, SVJ, SV2J, and SVJ2 models. We demonstrate that the prices of European options obtained using the path-independent lattices converge rapidly to their true prices obtained using quasi-analytical solutions. Our transform-based approach is computationally more efficient than all other methods given in the literature for a large class of low-dimensional stochastic volatility models.

In "A Multi-dimensional Transform for Pricing American Options Under Levy Models", We extend the multi-dimensional transform to Levy models with stochastic volatility and jumps in the underlying stock price process. Efficient path-independent tree can be constructed for both European and American options. Our path-independent

lattice method can be applied to almost all Levy models in the literature, suach as Merton (1976), Bates (1996, 2000, 2006), Pan (2002), the NIG model, the VG model and the CGMY model. The numerical results show that our method is extemly accurate and fast.

In "Empirical performance of Levy models for American Options", we investigate in-sample fitting and out-of-sample pricing performance on American call options under Levy models. The drawback of the BS model has been well documented in the literatures, such as negative skewness with excess kurtosis, fat tail, and non-normality. Therefore, many models have been proposed to resolve known issues associated the BS model. For example, to resolve volatility smile, local volatility, stochastic volatility, and diffusion with jumps have been considered in the literatures; to resolve non-normality, non-Markov processes have been considered, e.g., Poisson process, variance gamma process, and other type of Levy processes. One would ask: what is the gain from each of the generalized models? Or, which model is the best for option pricing? We address these problems by examining which model results in the lowest pricing error for American style contracts. For in-sample analysis, the rank (from best to worst) is Pan, CGMYsv, VGsv, Heston, CGMY, VG and BS. And for out-of-sample pricing performance, the rank (from best to worst) is CGMYsv, VGsv, Pan, Heston, BS, VG, and CGMY. Adding stochastic volatility and jump into a model improves American options pricing performance, but pure jump models are worse than the BS model in American options pricing. Our empirical results show that pure jump model are over-fitting, but not improve American options pricing when they are applied to out-of-sample data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

INTRODUCTION

The breakthroughs in European option pricing have outpaced those in American option pricing with the discovery of stochastic volatility model of Heston (1993) followed by different types of stochastic volatility jump (SVJ) models. This paper shows how to price American options under the most advanced types of SVJ option pricing models by generalizing the transform-based approach of Nelson and Ramaswamy (1990) to a multidimensional framework. The generalized framework allows the construction of simple path-independent lattices for pricing American options under the five classes of option pricing models given below:

1. One volatility factor-based stochastic volatility (SV) models (e.g., Chesney and Scott (1989), Heston (1993), Hull and White (1987), Stein and Stein (1991), Wiggins (1987), Aıt-Sahalia and Kimmel (2007), and Medvedev and Scaillet (2010));

2. Two volatility factors-based stochastic volatility (SV2) models (e.g., Bates (2000), and Jiang and Oomen (2007), and Medvedev and Scaillet (2010));

3. Stochastic volatility jump models with one volatility factor and one jump factor (SVJ models) (e.g., Andersen, Benzoni, and Lund (2002), Bates (1996, 2003), Bakshi, Cao, and Chen (1997, 2000), Pan (2002));

4. Levy models (e.g., the Variance Gamma (VG) model of Madan and Seneta (1990) and the CGMY model of Carr and Geman (2002); Carr, Geman, Madan, and Yor (2003a)).

5. Levy models extended with stochastic volatility given in this paper.

To our knowledge, this is the first paper to provide path-independent lattices for any of the above five classes of SVJ models. We demonstrate that the prices of European options obtained using the path-independent lattices converge rapidly to their true prices obtained using quasi-analytical solutions. The transform-based approach presented in this paper is more efficient and accurate than other methods proposed in the literature for the pricing of American options under the five classes of SVJ models.

The hugely popular book, The Volatility Surface (Gatheral (2011)), reveals immense practitioner interest in the SVJ models for pricing equity options. Much econometric research has already investigated the empirical properties of the above models. Aıt-Sahalia and Kimmel (2007) show how to estimate the one volatility factor-based stochastic volatility (SV) models using the maximum likelihood method. The studies of Bates (2000, 2006), Chacko and Viceira (2003), Chernov and Ghysels (1999), Eraker (2004), Eraker, Johannes, and Polson (2003), Jiang and Oomen (2007), and Pan (2002) find jumps to be an important feature in addition to stochastic volatility for explaining the features of equity index returns and option prices. The existence of a rich econometric literature on estimation of low-dimensional stochastic volatility models makes the results of our paper useful for future empirical investigations into American option pricing under these models.

To get more insight on the challenges in pricing American equity options, consider the following classification of the option pricing models, based upon the type of stochastic processes followed by the state variables:

1. low-dimensional models with non-stochastic volatility,

2. high-dimensional models with non-stochastic volatility,

3. low-dimensional models with stochastic volatility,

4. high-dimensional models with stochastic volatility.

Examples of the first type of models include the BS model of Black and Scholes (1973) with constant volatility, the CEV (constant elasticity of variance) model of Cox and Ross (1976) with state variable-dependent volatility, and the jump-diffusion models of Merton (1976) and Kou (2002) with constant volatility and constant jump distribution parameters. At least three methods, given as the lattice method, the finite difference method, and the analytical approximation method, have been successful at pricing American options under these models. The lattice method was proposed by Cox, Ross, and Rubinstein (1979) for the BS model. Nelson and Ramaswamy (1990) extend this method to build efficient recombining trees for the CEV model of Cox and Ross (1976) with state variable-dependent volatility. Amin (1993) extends this method to build efficient multinomial recombining trees for the Merton's jump diffusion model. The finite difference method was initially proposed by Brennan and Schwartz (1978) for the BS model, and has been extended by Zhang (1997) and Carr, Hirsa, and Stanley (2003b) for the Merton's jump diffusion model. Finally, analytical approximations for pricing American options have been derived by Barone-Adesi and Whaley (1987), Kim (1990) and Carr, Jarrow, and Myneni (1992) for the BS model, and have been extended by Gukhal (2001) and Chiarella and Ziogas (2005) for the Merton's jump diffusion model, and by Kou and Wang (2004) for the double exponential jump diffusion model of Kou (2002).

The absence of analytical approximations and practical limitations on computational time prevent the three methods mentioned above for pricing American options under the second type of models with many state variables. The significant breakthrough for pricing American options under the high-dimensional models with non-stochastic volatility is provided by Longstaff and Schwartz (2001) (LS), who building on the initial work by Carriere (1996) and Tsitsiklis and Roy (1999), develop an innovative least squares Monte Carlo (LSM) approach that uses least square regressions to estimate the conditional expected payoff to the option holder from continuation.

LS demonstrate the effectiveness of this approach for high-dimensional models using the examples of pricing an American swaption using a twenty-factor string model, and an American option on the maximum of five risky asset prices. The solutions, which would normally take hours or days to compute using the traditional methods, are obtained within minutes using the LSM approach.

Though the LSM approach is powerful for the second type of models, it is not generally recommended for the first type of models. [1] For example, the traditional binomial method, the Nelson and Ramaswamy (1990) extension of the binomial method, and the Amin (1993) multinomial extension of the binomial method are significantly more efficient than the LSM approach for pricing American options under the BS model, the Cox and Ross CEV models, and the Merton jump diffusion model, respectively. The traditional methods can generate the price of an American option within a fraction of time it takes for the LSM approach to do so, by running least squares regressions using simulated data with tens of thousands of paths.

Clearly, the traditional methods (i.e., lattice methods, finite difference methods, and analytical approximation methods) are well suited for the first type of models, and the LSM approach is advantageous for the second type of models. However, it is not immediately obvious if the LSM approach is useful for the third type of models characterized by low-dimensional stochastic processes with stochastic volatility. The main difficulty here is that the ordinary least squares regressions can no longer be used for estimating the conditional expected payoff to the option holder from continuation, due to the heteroscedastic error terms when the underlying processes follow stochastic volatility, as in the models of Heston (1993), Hull and White (1987), Bates (2000,

[1] The only exception is when the underlying state variables follow non-Markovian processes, making the lattice methods difficult to use for pricing American options. However, as shown by Amin and Morton (1994), for low-dimensional, non-Markovian HJM models (Heath, Jarrow and Morton (1991, 1992)), non-recombining trees are quite accurate even with as few as 8 to 10 steps. This finding may not hold over other classes of non-Markovian models, however.

2006), Jiang and Oomen (2007), Bakshi et al. (1997, 2000), Pan (2002), Duffie, Pan, and Singleton (2000), and the stochastic volatility-based extensions of the Levy jump models of Madan and Seneta (1990) and Carr and Geman (2002); Carr et al. (2003a). Though, in principle, some type of non-linear regression such as GMM, etc., could be used, this would further slow down the computational speed of the LSM procedure, as doing GMM would require time-consuming non-linear optimizations to be performed at every time step of the simulation.

As an alternative to the LSM approach, this paper derives a multi-dimensional transform that allows efficient path-independent lattice construction for virtually all low-dimensional stochastic volatility models (with and without jumps) given in the literature, using which American options can be priced accurately within seconds. Hence, even if some non-linear regression methods can be used to make the LSM procedure applicable to the third type of models, doing this would be relatively inefficient since the lattice methods derived in this paper can achieve the same objective with significantly higher computational efficiency.

Before outlining our approach to American option pricing for the above models, we would like to highlight a related approach that uses GARCH models for pricing equity options in the presence of stochastic volatility. This approach, developed by Duan (1995) for European option pricing using simulation methods, has been extended to American option pricing by Ritchken and Trevor (1999) (RT) using a lattice approximation method. Duan (1996, 1997) also shows that virtually all of the bivariate diffusion stochastic volatility models (e.g., Heston (1993), Hull and White (1987), and others) can be obtained as special cases of a family of generalized GARCH models. This allows the lattice algorithm of RT to be applicable to all of the bivariate diffusion stochastic volatility models, as well.

Since the variance process under the GARCH method is a path-dependent process, which is implying an explosive number of variances at any given node of the asset

price tree, RT use the approximation methods of Hull and White (1993) and Ritchken, Sankarasubramanian, and Vijh (1993) (which keep track of only the minimum and the maximum values of variances at each node) and linear interpolation to generate K variances at that node. With suitable parameterizations, RT find that a value of $K = 20$ is generally sufficient for accurate pricing of options using the GARCH model. Due to the volatility interpolations and other approximations involved in the modeling of the path-dependent variance process, the RT lattice algorithm is relatively slower and less accurate than a typical path-independent binomial or trinomial model. The RT lattice algorithm also requires a forward dynamic program to determine the values of the minimum and maximum variance at each node.

In contrast to the RT algorithm, this paper derives a multi-dimensional transform in order to model both the asset price process and the volatility process as path-independent trees. The efficiency and accuracy of our approach is of similar order as a two-factor model of Cox et al. (1979), which represents a significant improvement over the RT algorithm for the special case of the bivariate diffusion stochastic volatility models. Our approach can also generalize parsimoniously to other low-dimensional stochastic volatility models, such as those with two volatility state variables, or jump factors in the asset returns. In contrast, the GARCH option pricing models cannot allow more than one volatility state variable, and it is virtually impossible to extend the RT lattice approximation method to the GARCH-jump models (e.g., Duan, Ritchken, and Sun (2006)), which allow jumps in the asset return and/or the volatility process.

Though a few other researchers have provided different approaches for pricing American options under low-dimensional stochastic volatility models, none are as general in terms of applicability to a diverse set of models with stochastic volatility and jumps, or numerically as efficient as our approach (e.g., see Clarke and Parrott (1999), Finucane and Tomas (1996), Guan and Xiaoqiang (2000), Hilliard and Schwartz (1996), Leisen (2000), Tzavalis and Wang (2003), Chiarella and Ziogas

(2005), Chiarella, Kang, Meyer, and Ziogas (2009), Ballestra and Sgarra (2010), Li (2009), Zhylyevskyy (2009), Ballestra and Pacelli (2013), Salmi, Toivanen, and Sydow (2014), Rad and Parand (2014), and others). For example, Leisen (2000) shows how to construct recombining lattices for the class of bivariate diffusion stochastic volatility models. Leisen (2000) approach requires both forward induction and backward induction with a complex tree structure of the order $O(N^4)$ for an N-step tree. The requirement of forward induction puts huge burden on the computer memory and the high order (i.e., $O(N^4)$) for the tree structure slows down the approach of Leisen (2000) significantly. Though Hilliard and Schwartz (1996) generate a path independent lattice for the special case of Hull and White (1987) model using a two-dimensional transform (along the lines of Nelson and Ramaswamy (1990)), their approach cannot be generalized to any other stochastic volatility models.

Chiarella and Ziogas (2005) extend McKean (1965) incomplete Fourier transform method to solve the two-factor partial differential equation for the price and early exercise surface using numerical methods. However, Chiarella and Ziogas (2005) approach is limited only to the case of Heston model. Clarke and Parrott (1999) consider a multigrid implicit finite difference scheme for pricing American options under bivariate diffusion stochastic volatility models. This approach requires analytic strike-price related transformation of asset prices and adaptive time-stepping for efficient computation of option prices. It cannot be generalized easily to models with jumps in asset prices and/or volatility. Guan and Xiaoqiang (2000) use an interpolation-based approach to construct recombining trees for pricing American options under stochastic volatility models. However, their interpolation-based approximations lead to large pricing errors for long-dated options.

The analytical methods for American option pricing by Li (2009) and Zhylyevskyy (2009) apply only to diffusion-based stochastic volatility models. Ballestra and Sgarra (2010) present a finite element method, which efficiently combines an implicit/explicit

time stepping and an operator splitting technique, for pricing stochastic volatility jump models given in Bates (1996). However, given their method uses implicit/explicit time stepping and the approach presented in this paper uses path independent lattice method, which is equivalent to using a purely explicit finite difference scheme, our method is computationally more efficient.

Ballestra and Pacelli (2013) develops a radial basis function (RBF) approach to price American and European options, which is based on operator splitting technique and on Gaussian radial functions. But their approach is limited only to the case of the BS and Heston model. It is shown that the RBF approach performs fairly better than the finite difference schemes, but it is difficult to extend to the SVJ models and Levy models. Recently, there are few papers applying complicated numerical method for pricing American options under stochastic volatility models with/without jumps, such as Salmi et al. (2014) (STS) and Rad and Parand (2014) (RP). STS employs adaptive multigrid method, LU decomposition, and operator splitting method to pricing European and American options under the Bates (1996) model. RP proposes a local weak form meshless methods to price options under the Bates (1996) model. However, they calculate American option prices by extrapolation of the price of Bermudan options, which will reduce the accuracy for pricing American options. Both STS and RP methods are not easy to apply to more general Levy models, such as the VG model and the CGMY model. The transform-based path independent lattice method presented in this paper is much more simple and intuitive. It can be applied to almost all low dimensional Levy models with stochastic volatility and jump.

This dissertation is organized as follows. Chapter 1 proposes a multi-dimensional transform that is conditionally independent of the volatility state variable(s). A zero conditional correlation between the transform process and the volatility processes allows path-independent lattice construction without the occurrence of negative tran-

sition probabilities for virtually all stochastic volatility models given in the literature with one and two volatility factors (i.e., all SV and SV2 models). Section 1.1 of Chapter 1 introduces the discrete processes, notation used for the discrete processes, the assumptions required for the convergence of the discrete processes to the corresponding continuous processes, and the main convergence theorem. Section 1.2 of Chapter 1 constructs path-independent trinomial trees for the transform process and the volatility processes, and obtains two schemes for the joint path-independent lattice construction for these processes for valuing American options. In Section 1.3 of Chapter 1, we present option prices obtained by applying the multi-dimensional transform approach under SV and SVJ models. The transform-based path-independent lattices method is extended to Levy models with stochastic volatility in Chapter 2. It is shown that path-independent lattices method proposed in this disseration can be applied to almost all Levy models with or without stochastic volatility and jumps. Chapter 3 investigates the empirical performance of Levy models for American options. The conclusion is followed.

CHAPTER 1

A MULTI-DIMENSIONAL TRANSFORM FOR PRICING AMERICAN OPTIONS UNDER STOCHASTIC VOLATILITY MODELS

Assume that the stock price S_t, follows a two-factor stochastic process under the risk-neutral measure, given as

$$\frac{dS_t}{S_t} = rdt + \psi_1(V_{1,t}, t)dZ_{1,t} + \psi_2(V_{2,t}, t)dZ_{2,t}, \tag{1.1}$$

where, each of the state variables $V_{1,t}$ and $V_{2,t}$ follow a one-factor stochastic process, given as

$$dV_{i,t} = \phi_i(V_{i,t}, t)dt + \varphi_i(V_{i,t}, t)dW_{i,t}, \quad \text{for } i = 1, 2, \tag{1.2}$$

with

$$dZ_{i,t}dW_{i,t} = \rho_i dt, \ |\rho_i| < 1, \quad \text{for } i = 1, 2. \tag{1.3}$$

and zero correlations among all other pairs of Wiener processes, $Z_{1,t}$, $Z_{2,t}$, $W_{1,t}$, $W_{2,t}$. [1]

By appropriately choosing the functions $\psi_i, \phi_i, \varphi_i$ defined on $\mathbb{R} \times [0, \infty)$, equations (1.1) through (1.3) nest all continuous-time stochastic volatility (SV) models in the equity option pricing literature (e.g., the SV models of Chesney and Scott (1989), Heston (1993), Hull and White (1987), Johnson and Shanno (1987), Stein and Stein

[1] With perfect correlation, a path-independent lattice cannot be built even for simple two-factor binomial models with constant volatilities.

(1991), and Wiggins (1987); and the SV2 model of Alizadeh, Brandt, and Diebold (2002), Bates (2000), Chacko and Viceira (2003), and Jiang and Oomen (2007)).

The above model uses four factors, given as two diffusion factors and two volatility factors, of which only two pairs are correlated as shown in equation (1.3). The following proposition outlines the main contribution of this paper, given as the multi-dimensional transform which reformulates the Markovian system given by equations (1.1) through (1.3) with two pairs of conditionally correlated state variables into an equivalent Markovian system with *conditionally independent* state variables. This transform allows construction of path independent lattices for all SV and SV2 models given in the literature.

Proposition 1.0.1. For all stochastic processes nested in equations (1.1) through (1.3) and let $\mathbf{V}_t = [V_{1,t}, V_{2,t}]$, the multi-dimensional transform, given as,

$$X(S_t, \mathbf{V}_t, t) = \ln(S_t) - \sum_{i=1}^{2} \rho_i \int^{V_{i,t}} \frac{\psi_i(\xi, t)}{\varphi_i(\xi, t)} d\xi \tag{1.4}$$

is conditionally independent of each of the volatility processes, i.e.,

$$dX(S_t, \mathbf{V}_t, t) dV_{i,t} = 0, \quad \text{for } i = 1, 2. \tag{1.5}$$

Proof. Using Ito's lemma, the stochastic process for X is given as

$$dX(S_t, \mathbf{V}_t, t) = \mu(\mathbf{V}_t, t)dt + \sum_{i=1}^{2} \psi_i(V_{i,t}, t)(dZ_{i,t} - \rho_i dW_{i,t}), \tag{1.6}$$

where

$$\begin{aligned}\mu(\mathbf{V}_t,t) &= r-\frac{1}{2}\sum_{i=1}^{2}\psi_i^2(V_{i,t},t)-\sum_{i=1}^{2}\rho_i\frac{\psi_i(V_{i,t},t)}{\varphi_i(V_{i,t},t)}\phi_i(V_{i,t},t)\\ &\quad -\frac{1}{2}\sum_{i=1}^{2}\rho_i\left(\varphi_i(V_{i,t},t)\frac{\partial\psi_i(V_{i,t},t)}{\partial v_i}-\psi_i(V_{i,t},t)\frac{\partial\varphi_i(V_{i,t},t)}{\partial v_i}\right) \qquad (1.7)\\ &\quad -\sum_{i=1}^{2}\rho_i\int^{V_{i,t}}\left(\frac{1}{\varphi_i(\xi,t)}\frac{\partial\psi_i(\xi,t)}{\partial t}-\frac{\psi_i(\xi,t)}{\varphi_i^2(\xi,t)}\frac{\partial\varphi_i(\xi,t)}{\partial t}\right)d\xi.\end{aligned}$$

Equation (1.5) follows by multiplying equations (1.6) and (1.2), and using the definition of the correlation terms in equation (1.3), for $i = 1, 2$. Hence, the multi-dimensional transform $X(S_t, \mathbf{V}_t, t)$ is conditionally independent of the volatility processes $V_{i,t}$ for $i = 1, 2$. The definition of $X(S_t, \mathbf{V}_t, t)$ in equation (1.4) contains two terms. The first term is the standard log transformation of the stock price, which is needed for the construction of a path-independent tree for the stock price process even in absence of the volatility factors, as in the case of a path-independent binomial tree construction for the standard BS model. The second term eliminates the conditional correlation between $X(S_t, \mathbf{V}_t, t)$ and the two state variables $V_{i,t}$.

□

In order to price American options under stochastic volatility models, section 1.2 shows how to build path-independent trees for each of the two state variables $V_{i,t}$, and a path-independent tree for the X_t process (where, $X_t := X(S_t, \mathbf{V}_t, t)$). We show that the conditional independence of the transform $X(S_t, \mathbf{V}_t, t)$ from the volatility processes $V_{i,t}$, obtained in proposition 1.0.1, allows the marginal transitional probabilities to remain non-negative when constructing path-independent trees (in section 1.2). The next section gives the main assumptions, notation used for the discrete stochastic processes, and the main convergence theorem for convergence of the discrete processes to the corresponding continuous processes.

1.1 Discrete Approximations and Convergence Theorem

The main convergence theorem in this section requires four assumptions. The first two assumptions concern the existence and uniqueness conditions for the stochastic processes $X_t := X(S_t, \mathbf{V}_t, t)$ and $V_{i,t}$. The last two assumptions concern the continuity and moment convergence conditions for the discrete processes to converge to the corresponding continuous processes. Before stating the assumptions, we first provide the notation describing the discrete processes corresponding to the $V_{i,t}$ and X_t processes in the following two subsections.

1.1.1 The Discrete Processes for $V_{i,t}$-process

Let the time interval $[0, T]$ be partitioned into n equal, left-closed right-open subintervals, $[t_k, t_{k+1})$, where $0 = t_0 < t_1 < \ldots < t_n = T$, and let $h = T/n$. For each $h > 0$, the scalar valued functions $V^u_{i,h}(v_i, hk), V^m_{i,h}(v_i, hk)$, and $V^d_{i,h}(v_i, hk)$ are defined on $\mathbb{R} \times [0, \infty)$ and satisfy

$$-\infty < V^d_{i,h}(v_i, hk) \leq V^m_{i,h}(v_i, hk) \leq V^u_{i,h}(v_i, hk) < \infty,$$

and

$$\begin{aligned} V^u_{i,h}(v_i, hk) &= v_i + \Delta V^u_{i,h}(v_i, hk), \\ V^m_{i,h}(v_i, hk) &= v_i + \Delta V^m_{i,h}(v_i, hk), \\ V^d_{i,h}(v_i, hk) &= v_i + \Delta V^d_{i,h}(v_i, hk), \end{aligned} \tag{1.8}$$

for $i = 1, 2$ and $k = 0, 1, 2, \cdots, n$, where, $\Delta V^u_{i,h}(v_i, hk)$, $\Delta V^m_{i,h}(v_i, hk)$, and $\Delta V^d_{i,h}(v_i, hk)$ are three scalar valued functions defined on $\mathbb{R} \times [0, \infty)$, representing the size of up, middle, and down movements. For each function, the first argument v_i is volatility (spatial variable), and the second argument $hk = t_k$ is the time variable. Three transition probabilities corresponding to these movements are defined on $\mathbb{R} \times [0, \infty)$ as

$q^u_{i,h}(v_i, hk)$, $q^m_{i,h}(v_i, hk)$, and $q^d_{i,h}(v_i, hk) = 1 - q^u_{i,h}(v_i, hk) - q^m_{i,h}(v_i, hk)$, These three probabilities satisfy $0 \leq q^j_{i,h}(v_i, hk) \leq 1$ for all $v_i \in \mathbb{R}$, and $k = 0, 1, 2, \cdots, n$, and $j = u, m, d$.

For each $h > 0$ and $i = 1, 2$, the discrete approximation of the $V_{i,t}$ process is denoted as ${}_hV_{i,t}$ on the time interval $[0, T]$. The ${}_hV_{i,t}$-process is constant between nodes, i.e., ${}_hV_{i,0} = v_{i,0}$ for $t = 0$, and ${}_hV_{i,t} = {}_hV_{i,hk}$ for $hk \leq t < h(k+1)$. At any given time node hk, it either moves up, stays the same, or moves down, with the probabilities of up, middle, and down movements defined as, respectively,

$$\begin{aligned}
P\left[{}_hV_{i,h(k+1)} = V^u_{i,h}({}_hV_{i,hk}, hk) \,|hk, {}_hV_{i,hk}\right] &= q^u_{i,h}({}_hV_{i,hk}, hk), \\
P\left[{}_hV_{i,h(k+1)} = V^m_{i,h}({}_hV_{i,hk}, hk) \,|hk, {}_hV_{i,hk}\right] &= q^m_{i,h}({}_hV_{i,hk}, hk), \\
P\left[{}_hV_{i,h(k+1)} = V^d_{i,h}({}_hV_{i,hk}, hk) \,|hk, {}_hV_{i,hk}\right] &= q^d_{i,h}({}_hV_{i,hk}, hk).
\end{aligned}$$

For any $h > 0$, the local drift function $\phi_{i,h}(v_i, t)$ and the local second moment $\varphi^2_{i,h}(v_i, t)$ are defined as

$$\begin{aligned}
\phi_{i,h}(v_i, t) &= \frac{q^u_{i,h}(v_i, t^*)\Delta V^u_{i,h}(v_i, t^*)}{h} + \frac{q^m_{i,h}(v_i, t^*)\Delta V^m_{i,h}(v_i, t^*)}{h} \\
&\quad + \frac{q^d_{i,h}(v_i, t^*)\Delta V^d_{i,h}(v_i, t^*)}{h}, \qquad (1.9) \\
\varphi^2_{i,h}(v_i, t) &= \frac{q^u_{i,h}(v_i, t^*)[\Delta V^u_{i,h}(v_i, t^*)]^2}{h} + \frac{q^m_{i,h}(v_i, t^*)[\Delta V^m_{i,h}(v_i, t^*)]^2}{h} \\
&\quad + \frac{q^d_{i,h}(v_i, t^*)[\Delta V^d_{i,h}(v_i, t^*)]^2}{h}, \qquad (1.10)
\end{aligned}$$

for all $v_i \in \mathbb{R}$, where $t^* = h[\text{INT}(t/h)]$, and $\text{INT}(t/h)$ is the integer part of t/h.

1.1.2 The Discrete Processes for X_t-process

For each $h > 0$ and $\mathbf{v} = [v_1, v_2]$, three scalar valued functions $X^u_h(x, \mathbf{v}, hk)$, $X^m_h(x, \mathbf{v}, hk)$, and $X^d_h(x, \mathbf{v}, hk)$ satisfying $0 < X^d_h(x, \mathbf{v}, hk) \leq X^m_h(x, \mathbf{v}, hk) \leq X^u_h(x, \mathbf{v}, hk) < \infty$, are defined on $\mathbb{R} \times \mathbb{R}^2 \times [0, \infty)$ as

$$
\begin{aligned}
X_h^u(x, \mathbf{v}, hk) &= x + \Delta X_h^u(\mathbf{v}, hk), \\
X_h^m(x, \mathbf{v}, hk) &= x + \Delta X_h^m(\mathbf{v}, hk), \\
X_h^d(x, \mathbf{v}, hk) &= x + \Delta X_h^d(\mathbf{v}, hk),
\end{aligned}
$$

where $k = 0, 1, \ldots, n$. Here $\Delta X_h^u(\mathbf{v}, hk)$, $\Delta X_h^m(\mathbf{v}, hk)$, and $\Delta X_h^d(\mathbf{v}, hk)$ are three scalar valued functions defined on $\mathbb{R}^2 \times [0, \infty)$ representing the size of up, middle, and down movements. These functions are assumed *independent* of x, because the drift and the variance of X_t depend only upon the state variables $\mathbf{V}_t$ and time t (see equations (1.6) and (1.7)). The three transition probabilities corresponding to these movements to be also specified *independently* of x (for the same reason noted above) as three scalar valued functions $p_h^u(\mathbf{v}, hk)$, $p_h^m(\mathbf{v}, hk)$, and $p_h^d(\mathbf{v}, hk) = 1 - p_h^u(\mathbf{v}, hk) - p_h^m(\mathbf{v}, hk)$, defined on $\mathbb{R}^2 \times [0, \infty)$. These three probabilities satisfy $0 \leq p_h^j(\mathbf{v}, hk) \leq 1$ for all $\mathbf{v} \in \mathbb{R}^2$, $k = 0, 1, 2, \ldots, n$, and $j = u, m, d$.

Given $h > 0$, the discrete approximation of the X_t process is denoted as ${}_hX_t$ on the time interval $[0, T]$. The ${}_hX_t$ process is constant between nodes, i.e., ${}_hX_0 = x_0$ for $t = 0$, and ${}_hX_t = {}_hX_{hk}$ for $hk \leq t < h(k+1)$. At any given time node hk, it either moves up, stays same, or moves down, with the probabilities of up, middle, and down movements defined as, respectively,

$$
\begin{aligned}
P\left[{}_hX_{h(k+1)} = X_h^u({}_hX_{hk}, {}_h\mathbf{V}_{hk}, hk) \,|hk, {}_hX_{hk}, {}_h\mathbf{V}_{hk}\right] &= p_h^u({}_h\mathbf{V}_{hk}, hk), \\
P\left[{}_hX_{h(k+1)} = X_h^m({}_hX_{hk}, {}_h\mathbf{V}_{hk}, hk) \,|hk, {}_hX_{hk}, {}_h\mathbf{V}_{hk}\right] &= p_h^m({}_h\mathbf{V}_{hk}, hk), \\
P\left[{}_hX_{h(k+1)} = X_h^d({}_hX_{hk}, {}_h\mathbf{V}_{hk}, hk) \,|hk, {}_hX_{hk}, {}_h\mathbf{V}_{hk}\right] &= p_h^d({}_h\mathbf{V}_{hk}, hk),
\end{aligned}
$$

where, ${}_h\mathbf{V}_{hk} = ({}_hV_{1,hk}, {}_hV_{2,hk})$.

For any $h > 0$, the local drift function $\mu_h(\mathbf{v}, t)$ and the local second moment $g_h^2(\mathbf{v}, t)$ are defined as

$$\begin{aligned}
\mu_h(\mathbf{v},t) &= \frac{p_h^u(\mathbf{v},t^*)\Delta X_h^u(\mathbf{v},t^*) + p_h^m(\mathbf{v},t^*)\Delta X_h^m(\mathbf{v},t^*) + p_h^d(\mathbf{v},t^*)\Delta X_h^d(\mathbf{v},t^*)}{h}, \\
g_h^2(\mathbf{v},t) &= \frac{p_h^u(\mathbf{v},t^*)[\Delta X_h^u(\mathbf{v},t^*)]^2 + p_h^m(\mathbf{v},t^*)[\Delta X_h^m(\mathbf{v},t^*)]^2 + p_h^d(\mathbf{v},t^*)[\Delta X_h^d(\mathbf{v},t^*)]^2}{h},
\end{aligned}$$

for all $\mathbf{v} \in \mathbb{R}^2$, where, $t^* = h[\mathrm{INT}(t/h)]$.

1.1.3 The Assumptions

We now state four assumptions that are sufficient for the discrete multi-dimensional process $\{{}_hX_t, {}_h\mathbf{V}_t\}$ to converge weakly to the corresponding continuous multi-dimensional process $\{X_t, \mathbf{V}_t\}$ as $h \to 0$.

A1. The SDEs for the $\mathbf{V}_t$ processes in equation (1.2) have unique and strong solutions on $[0,T]$.

A2. The SDE for X_t in equation (1.6) has a distributionally unique and weak solution on $[0,T]$.

A3. For all $\delta_i > 0, i = 1,2$, and for $j = u,d,m$,

$$\begin{aligned}
\lim_{h\to 0}\left(\sup_{\substack{|v_i|<\delta_i \\ 0\le t\le T}}\left|V_{i,h}^j(v_i,t) - v_i\right|\right) &= \lim_{h\to 0}\left(\sup_{\substack{|v_i|<\delta_i \\ 0\le t\le T}}\left|\Delta V_{i,h}^j(v_i,t)\right|\right) = 0, \\
\lim_{h\to 0}\left(\sup_{\substack{|v_i|<\delta_i \\ 0\le t\le T}}\left|X_h^j(\mathbf{v},t) - x\right|\right) &= \lim_{h\to 0}\left(\sup_{\substack{|v_i|<\delta_i \\ 0\le t\le T}}\left|\Delta X_h^j(\mathbf{v},t)\right|\right) = 0.
\end{aligned}$$

A4. For all $\delta_i > 0, i = 1,2$,

$$\begin{aligned}
&\lim_{h\to 0}\left(\sup_{\substack{|v_i|<\delta_i \\ 0\le t\le T}}\left|\phi_{i,h}(v_i,t) - \phi_i(v_i,t)\right|\right) = 0, \quad \lim_{h\to 0}\left(\sup_{\substack{|v_i|<\delta_i \\ 0\le t\le T}}\left|\varphi_{i,h}^2(v_i,t) - \varphi_i^2(v_i,t)\right|\right) = 0, \\
&\lim_{h\to 0}\left(\sup_{\substack{|v_i|<\delta_i \\ 0\le t\le T}}\left|\mu_h(\mathbf{v},t) - \mu(\mathbf{v},t)\right|\right) = 0, \qquad \lim_{h\to 0}\left(\sup_{\substack{|v_i|<\delta_i \\ 0\le t\le T}}\left|g_h^2(\mathbf{v},t) - g^2(\mathbf{v},t)\right|\right) = 0.
\end{aligned}$$

Here, $g^2(\mathbf{v}, t)$ is the local second moment of the X_t process, which is obtained using equation (1.6), as follows,

$$g^2(\mathbf{v}, t) = \sum_{i=1}^{2} \psi_i^2(v_i, t)(1 - \rho_i^2). \tag{1.11}$$

1.1.4 The Convergence Theorem

We now state the convergence theorem that establishes the convergence of the discrete processes to the corresponding continuous processes.

Theorem 1.1.1. Under assumptions A1 through A4, the discrete multi-dimensional process $\{{}_hX_t, {}_h\mathbf{V}_t\}$ converge weakly to the continuous multi-dimensional process $\{X_t, \mathbf{V}_t\}$ as $h \to 0$.

Theorem 1.1.1 implies that we can approximate these continuous models with discrete processes, which we proceed to do. [2]

1.2 Pricing American Options under Stochastic Volatility Models using Path-Independent Trinomial Trees

In this section, we will apply the multi-dimensional transform (see Proposition 1.0.1) to various stochastic volatility models nested in equation (1.1) through (1.3). Table 1 lists the five types of option pricing models in the literature, which we are considering to use path-independent trinomial tree method to price American options with the multi-dimensional transform. The first three rows of Table 1 specify the functions $\phi_i(v), \varphi_i(v)$, and $\psi_i(v)$ for models with both one volatility factor and two volatility factors, i.e., the SV models of Hull and White (1987), Chesney and Scott (1989), Stein and Stein (1991), and Heston (1993); the SV2 models given by Alizadeh et al.

[2] Sami Zreik (2012) shows how the continuous models we consider in this paper (see Appendix A for a detailed list) satisfy Assumptions A1 through A4.

(2002), Bates (2000), Chacko and Viceira (2003), and Jiang and Oomen (2007). Row 4 gives the transformed process X_t. Row 5 and 6 are drift and diffusion parts for X_t-process in equation (1.6). The last row is the inverse transform. The SDEs of these five models are given in Appendix A.

The transformed process X_t is independent with the volatility process V_t, which allows us to construct path-independent trees to obtain option price for both American options and European options. We discuss below the details of creating trinomial trees for V_t and X_t processes, respectively. Then, we use these technique to compute both European and American option prices.

1.2.1 Construction of Path-Independent Trinomial Trees for ${}_h\mathbf{V}_t$ Processes

When the volatility process V_t in the first four models in Table 1 follows a log-normal distribution (Hull-White model), mean-reverting process (Chesney-Scott model and Stein-Stein model), or square root process (Heston model), a recombining tree is constructed. When the two volatility processes in the SV2 model are uncorrelated, a recombining tree is constructed separately for each of the volatility process.

Let $\bar{\sigma}_i$, a constant parameter, represent the value used for spacing the adjacent nodes on the grid for the trinomial tree for the ${}_hV_{it}$ process, such that $\bar{\sigma}_i\sqrt{h}$ is the vertical distance between any two adjacent nodes on the entire grid. At each node $v_i \in \mathbb{R}$, and $t \in [0, \infty)$, the three size of movements in equation (1.8) are defined as

$$\begin{aligned} \Delta V_h^u(v_i, t) &= J_i(v_i, t)\bar{\sigma}_i\sqrt{h}, \\ \Delta V_h^m(v_i, t) &= 0, \\ \Delta V_h^d(v_i, t) &= -J_i(v_i, t)\bar{\sigma}_i\sqrt{h}, \end{aligned} \tag{1.12}$$

where $J_i(v_i, t)$ is the first positive integer such that

$$\left(\frac{\varphi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\right)^2 + \frac{|\phi_i(v_i,t)|}{J_i(v_i,t)\bar{\sigma}_i}\sqrt{h} \leq 1.$$

The main role of the positive integer $J_i(v_i,t)$ is to ensure that the span of the nodes (distance between the up and the down node) is sufficient to generate a positive middle probability. The three transition probabilities are computed as

$$\begin{aligned} q^u_{i,h}(v_i,t) &= \frac{1}{2}\left(\frac{\varphi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\right)^2 + \frac{1}{2}\frac{\phi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\sqrt{h}, \\ q^m_{i,h}(v_i,t) &= 1 - \left(\frac{\varphi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\right)^2, \\ q^d_{i,h}(v_i,t) &= \frac{1}{2}\left(\frac{\varphi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\right)^2 - \frac{1}{2}\frac{\phi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\sqrt{h}. \end{aligned} \tag{1.13}$$

Generally, the up and down probabilities $q^u_{i,h}(v_i,t)$ and $q^d_{i,h}(v_i,t)$ will remain *non-negative* with small enough values of h. However, since under stochastic volatility models $\varphi_i(v_i,t)$ can become arbitrarily small (or even go to zero with a non-zero probability as in the Heston model), there is no guarantee that the up and down probabilities will always remain non-negative for any small h. To ensure their non-negativity, the three probabilities are redefined if the following condition is satisfied,

$$\left(\frac{\varphi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\right)^2 < \frac{|\phi_i(v_i,t)|}{J_i(v_i,t)\bar{\sigma}_i}\sqrt{h}. \tag{1.14}$$

By introducing the indicator function $I_{v_i}, i = 1, 2,$

$$I_{v_i} = \begin{cases} 1, & \text{if } \phi_i(v_i,t) > 0, \\ 0, & \text{otherwise}, \end{cases}$$

the transition probabilities in equation (1.13) can be replaced as

$$
\begin{aligned}
q^u_{i,h}(v_i,t) &= \frac{1}{2}\left(\frac{\varphi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\right)^2 + \frac{1}{2}\frac{\phi_i(v_i,t)I_{v_i}}{J_i(v_i,t)\bar{\sigma}_i}\sqrt{h}, \\
q^m_{i,h}(v_i,t) &= 1 - \left(\frac{\varphi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\right)^2 - \frac{1}{2}\frac{\phi_i(v_i,t)\,(2I_{v_i}-1)}{J_i(v_i,t)\bar{\sigma}_i}\sqrt{h}, \\
q^d_{i,h}(v_i,t) &= \frac{1}{2}\left(\frac{\varphi_i(v_i,t)}{J_i(v_i,t)\bar{\sigma}_i}\right)^2 - \frac{1}{2}\frac{\phi_i(v_i,t)(1-I_{v_i})}{J_i(v_i,t)\bar{\sigma}_i}\sqrt{h}.
\end{aligned}
\tag{1.15}
$$

It is straightforward to check that the three probabilities in equation (1.15) are *always* valid, i.e., non-negative probabilities. Note that, unlike in equation (1.13), the term $\phi_i(v_i,t)$ is contained in the middle probability, and is either a part of the up probability or the down probability, but not both due to the indicator function I_{v_i}. This allows the up and down probabilities to remain positive when $\varphi_i(v_i,t)$ becomes significantly smaller than $\phi_i(v_i,t)$ such that the inequality (1.14) applies.

The lattice construction uses the probabilities in equation (1.13), and only switches to the probabilities in equation (1.15), when the condition given in equation (1.14) applies. The probabilities in equation (1.15) always ensure convergence of the first two moments (regardless of whether or not the condition in equation (1.14) applies), however, the rate of convergence of the second moment is faster if the probabilities in equation (1.15) are used only when condition in equation (1.14) applies, and otherwise the probabilities in equation (1.13) are used. Finally, note that in models such as the Heston model in which $\varphi_i(v_i,t) = 0$, the down probability is zero (since $\phi_i(v_i,t)$ is greater than zero in the Heston model, when $\varphi_i(v_i,t) = 0$, the indicator function $I_{v_i} = 1$).

After constructing the trinomial tree for volatility process, the next task is to construct a trinomial tree for the transformed process X_t, which is detailed in the next subsection.

1.2.2 Construction of Path-Independent Trinomial Tree for ${}_hX_t$ Process

Construction of a path-independent trinomial tree for the ${}_hX_t$ process requires that the size of movements $\Delta X^u_h(\mathbf{v},t), \Delta X^m_h(\mathbf{v},t), \Delta X^d_h(\mathbf{v},t)$ (that define the up, middle,

and down movements, respectively, in equation (1.16)) and the corresponding transition probabilities $p_h^u(\mathbf{v},t),\ p_h^m(\mathbf{v},t),\ p_h^d(\mathbf{v},t)$ (defined in equation (1.17)) are specified in a manner that satisfy assumptions A3 and A4. The specifications of functions $\Delta X_h^u(\mathbf{v},t), \Delta X_h^m(\mathbf{v},t), \Delta X_h^d(\mathbf{v},t)$ have to also ensure that future nodes recombine, making the tree path-independent.

Let $\bar{\sigma}$, a constant parameter, represent the value used for spacing the adjacent nodes on the grid for the trinomial tree for the ${}_hX_t$ process, such that $\bar{\sigma}\sqrt{h}$ is the vertical distance between any two adjacent nodes on the entire grid. For any $\mathbf{v} \in \mathbb{R}^2$, and $t \in [0,\infty)$, the size of movements are defined as follows

$$\begin{aligned} \Delta X_h^u(\mathbf{v},t) &= J(\mathbf{v},t)\bar{\sigma}\sqrt{h}, \\ \Delta X_h^m(\mathbf{v},t) &= 0, \\ \Delta X_h^d(\mathbf{v},t) &= -J(\mathbf{v},t)\bar{\sigma}\sqrt{h}, \end{aligned} \tag{1.16}$$

where $J(\mathbf{v},t)$ is the first positive integer such that

$$\left(\frac{g(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\right)^2 + \frac{|\mu(\mathbf{v},t)|}{J(\mathbf{v},t)\bar{\sigma}}\sqrt{h} \leq 1.$$

The three functions that define the transition probabilities corresponding to the up, middle, and down states are computed as

$$\begin{aligned} p_h^u(\mathbf{v},t) &= \frac{1}{2}\left(\frac{g(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\right)^2 + \frac{1}{2}\frac{\mu(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\sqrt{h}, \\ p_h^m(\mathbf{v},t) &= 1 - \left(\frac{g(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\right)^2, \\ p_h^d(\mathbf{v},t) &= \frac{1}{2}\left(\frac{g(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\right)^2 - \frac{1}{2}\frac{\mu(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\sqrt{h}. \end{aligned} \tag{1.17}$$

Note that selection of $J(\mathbf{v},t)$ ensures that $p_h^u(\mathbf{v},t) \leq \frac{1}{2}$, $p_h^d(\mathbf{v},t) \leq \frac{1}{2}$, and $p_h^m(\mathbf{v},t) \leq 1$. To ensure their non-negativity, the three probabilities are redefined if the following condition is satisfied,

$$\left(\frac{g(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\right)^2 < \frac{|\mu(\mathbf{v},t)|}{J(\mathbf{v},t)\bar{\sigma}}\sqrt{h}. \tag{1.18}$$

Then transition probabilities in equation (1.17) is replaced by

$$\begin{aligned} p_h^u(\mathbf{v},t)) &= \frac{1}{2}\left(\frac{g(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\right)^2 + \frac{1}{2}\frac{\mu(\mathbf{v},t)I_A}{J(\mathbf{v},t)\bar{\sigma}}\sqrt{h}, \\ p_h^m(\mathbf{v},t)) &= 1 - \left(\frac{g(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\right)^2 - \frac{1}{2}\frac{\mu(\mathbf{v},t)(2I_A-1)}{J(\mathbf{v},t)\bar{\sigma}}\sqrt{h}, \\ p_h^d(\mathbf{v},t)) &= \frac{1}{2}\left(\frac{g(\mathbf{v},t)}{J(\mathbf{v},t)\bar{\sigma}}\right)^2 - \frac{1}{2}\frac{\mu(\mathbf{v},t)(1-I_A)}{J(\mathbf{v},t)\bar{\sigma}}\sqrt{h}, \end{aligned} \tag{1.19}$$

where I_A is the indicator function

$$I_A = \begin{cases} 1, & \text{if } \mu(\mathbf{v},t) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

It is straightforward to check that the three probabilities in (1.19) are *always* valid, i.e., non-negative probabilities.

After building path-independent trinomial tree for volatility process and transformed process for stock price, separately, these independent trees will be combined into one multi-dimensional tree, and options' price can be obtained by the usual backward induction. The next section shows some numerical results of both American and European option prices, using the transform-based approach presented in this paper.

1.3 Tree Simulations for Pricing American Options under Stochastic Volatility Models

The transform method is applied to SV and SV2 models in this section. Table 2 documents the performance of Heston model. It shows the European and American

put option prices computed by using the transform-based lattice method and the Heston model closed form solution for the case of European option. The computations have been performed for the options with strikes $X = \$90, \$100$, and $\$110$ maturing in $1, 3$ and 6 months for the starting volatility value of $\sqrt{V(0)} = 0.1, 0.2, 0.3$ and starting firm value of $S(0) = \$100$. The other parameters are given as follows, $r = 0.03, \sigma = 0.38, \alpha = 5.3, m = 0.0242, \rho = -0.57$. The number of time steps in a tree is $N = 100$ and 300. The transform-based method presented in this paper is computational efficient both for European and American options. The European put prices obtained by tree method converge to the closed-form solution within seconds, and the average percent error is 1.36×10^{-3}.

Table 3 presents European put option prices under the Hull-White model. The parameters are adopted from Leisen (2000) and given as follows, $r = 0$, $\sigma = 1.0$, $\alpha = 0$, $m = 0$, $v(0) = 0.1$, maturity = 0.5 year, and starting firm value $S(0) = 100$. The put options are evaluated at strike price from 95 to 105 with dollar increment and $\rho = -0.5, 0$, and 0.5, separately. The transform-based tree method converges to the closed-form solution rapidly (within seconds for time step $N = 100$) and much faster than the Leisen (2000) method and the Monte Carlo method.

Table 4 documents the performance of SV2 model. It shows the European and American put option prices and closed form solution for the case of European option. The computations have been performed for the options with strikes $X = \$90, \$100$, and $\$110$ maturing in $1, 3$ and 6 months for the starting volatility value of $\sqrt{v(0)} = 0.1$, and 0.3 and starting firm value of $S(0) = \$100$. The other parameters are adopted from Christoffersen, Heston, and Jacobs (2009) and listed as follows, $r = 0.03, \sigma_1 = 1.072, \alpha_1 = 0.2563, m_1 = 0.004, \rho_1 = -0.8084, \sigma_2 = 0.2875, \alpha_2 = 1.8817, m_2 = 0.0233, \rho_2 = -0.6997$. The number of time steps in a tree is $N = 50$ and 150. The transform-based tree method has high accuracy and computational efficiency.

CHAPTER 2

A MULTI-DIMENSIONAL TRANSFORM FOR PRICING AMERICAN OPTIONS UNDER LEVY MODELS

Under stochastic volatility models (mentioned in the previous sections, and/or others), the volatility is modeled as a diffusion process. Therefore, it does not have enough variation in the short run to generate high short-term kurtosis and therefore prices of short-term options might be biased. To deal with this problem, many papers extend stochastic volatility models with jumps in asset returns (for example, see Bates (1996, 2000), Bakshi et al. (1997), Pan (2002), Andersen et al. (2002), and Chernov, Gallant, Ghysels, and Tauchen (1999)). It has been suggested that jumps could account for the skewness and high kurtosis in option prices. There are not many papers that use or develop numerical procedures for pricing American options under stochastic volatility and jumps in asset returns. For example, Bates (1996) uses a bivariate explicit finite-difference method influenced by Omberg (1988) to compute American option prices. He subsequently adjusts the obtained prices for observed biases in European prices using the control variate technique. Another approach using Monte Carlo simulation that works with multiple factors was developed by Longstaff and Schwartz (2001). The key to their approach is to use a cross-sectional least squares regression to estimate the conditional expected payoff to the option holder from continuation. This way they are able to obtain a complete specification of the optimal exercise strategy along each sample path. With this specification, American options can then be valued by simulation. But, those methods are computationally inefficient. In the second part of this paper we further extend our transform-based approach to Levy models with stochastic volatility and allowing jumps in asset returns,

for example, the Variance Gamma (VG) model of Madan and Seneta (1990) and the CGMY model of Carr and Geman (2002); Carr et al. (2003a). The numerical results show that our transform-based recombining lattices method is much more efficient than the previous methods mentioned.

2.1 Approximation of Levy Models with Compound Poisson Process and Brownian Motion

Under the risk-neutral world, the SDE of Levy models for log-price process is given as

$$d\ln S_t = \left(r - \frac{1}{2}\psi_1^2(V_{1,t},t) - \frac{1}{2}\psi_2^2(V_{2,t},t)\right)dt + \psi_1(V_{1,t},t)dZ_{1,t} + \psi_2(V_{2,t},t)dZ_{2,t} + dL_{1,t} + dL_{2,t}. \quad (2.1)$$

Here, each of the state variables $V_{1,t}$ and $V_{2,t}$ follows the one-factor stochastic processes given in equation (1.2), and all correlations between state variables are zero except those given in equation (1.3); furthermore, $L_{1,t}$ and $L_{2,t}$ are two Levy components, which are defined by the corresponding triplets

$$[0, \quad \boldsymbol{\kappa}_i(V_{i,t})\nu_i, \quad \boldsymbol{\gamma}_i(V_{i,t})], \qquad i = 1,2, \quad (2.2)$$

where, ν_i is the intensity measure associated with the traditional Levy processes given in the literature such as the VG, CGMY, inverse Gaussian, and damped power law models, etc. For a review of Levy characteristic triplets see Cont and Tankov (2003) (Page 80). Modifying the traditional Levy intensities as $\boldsymbol{\kappa}_i(V_{i,t})\nu_i$, where $\boldsymbol{\kappa}(V_{i,t})$ is an increasing function of $V_{i,t}$, allows the intensity to increase with an increase in the

corresponding volatility state variable. Under the risk-neutral condition, $\boldsymbol{\gamma}_i(V_{i,t})$ must satisfies (see Proposition 3.18 in Cont and Tankov (2003) p.98)

$$\begin{aligned}\boldsymbol{\gamma}_i(V_{i,t}) &= -\int_{\mathbb{R}} \left(e^y - 1 - y1_{|y|\leq 1}\right) \boldsymbol{\kappa}_i(V_{i,t})\nu_i(dy) \\ &= -\boldsymbol{\kappa}(V_{i,t}) \int_{\mathbb{R}} \left(e^y - 1 - y1_{|y|\leq 1}\right) \nu_i(dy). \end{aligned} \tag{2.3}$$

As shown in Section 6.3–6.4 of Cont and Tankov (2003), a Levy process with infinite jump activity can be approximated using Brownian motion and Poisson process by modeling the finite number of jumps of size greater or equal to ε using a compound Poisson process and either ignoring the uncertainty related to the infinite number of jumps of size less than ε (the finite variation case) or modeling the infinite number of jumps of size less than ε using a Brownian motion (the infinite variation case). In this paper, to allow discrete approximation of Levy process consistent with tree methods, we approximate the Levy processes $L_{i,t}$ with $L_{i,t}^{\varepsilon_i}$, where $L_{i,t}^{\varepsilon_i}$ is a new Levy process with compound Poisson and Brownian components.

The Levy triplets associated with $L_{i,t}^{\varepsilon_i}$ depend upon the type of Levy processes assumed. For convenient, here we define three new variables as follows,

$$\delta_i(\varepsilon_i) = \int_{|y|\geq\varepsilon_i} (e^y - 1)\, \nu_i(dy), \tag{2.4}$$

$$\lambda_i(\varepsilon_i) = \int_{|y|\geq\varepsilon_i} \nu_i(dy), \tag{2.5}$$

$$\begin{aligned}\gamma_i(V_{i,t}, \varepsilon_i) &= -\int_{|y|\geq\varepsilon_i} \left(e^y - 1 - y1_{|y|\leq 1}\right) \boldsymbol{\kappa}_i(V_{i,t})\nu_i(dy), \\ &= -\boldsymbol{\kappa}_i(V_{i,t}) \int_{|y|\geq\varepsilon_i} \left(e^y - 1 - y1_{|y|\leq 1}\right) \nu_i(dy). \end{aligned} \tag{2.6}$$

Then we define the triplets associated with $L_{i,t}^{\varepsilon_i}$ as well as the stochastic differential of the Levy process in the following three cases, which correspond to the three types of Levy processes.

Case 1. Finite jump activity Levy processes.

In this case $\varepsilon_i = 0$, and the approximated Levy process is the same as the original Levy process, i.e., $L^0_{i,t} = L_{i,t}$. And the Levy triplet is defined as

$$\left[0, \quad \boldsymbol{\kappa}_i(V_{i,t})\nu_i, \quad \gamma_i(V_{i,t}, 0)\right]. \tag{2.7}$$

The stochastic differential of the Levy process can be represented as a compound Poisson process due to finite jump activity, and is given as follows

$$dL^0_{i,t} = dL_{i,t} = -\boldsymbol{\kappa}_i(V_{i,t})\delta_i(0)dt + J_{i,t}dN^0_{i,t}, \tag{2.8}$$

where, $dN^0_{i,t}$ is a Poisson process with jump intensity $\boldsymbol{\kappa}_i(V_{i,t})\lambda_i(0)$, and the probability density of the size of jump $J_{i,t}$ is

$$p(J_{i,t}) = \frac{\boldsymbol{\kappa}_i(V_{i,t})\nu_i(J_{i,t})}{\boldsymbol{\kappa}_i(V_{i,t})\lambda_i(0)} = \frac{\nu_i(J_{i,t})}{\lambda_i(0)}. \tag{2.9}$$

Case 2. Infinite jump activity Levy processes with finite variation.

In this case, the approximated Levy process, $L^{\varepsilon_i}_{i,t}$, is defined by the following triplet

$$\left[0, \quad \boldsymbol{\kappa}_i(V_{i,t})\nu_i 1_{|y|\geq\varepsilon_i}, \quad \gamma_i(V_{i,t}, \varepsilon_i)\right]. \tag{2.10}$$

The stochastic differential of the Levy process is

$$dL^{\varepsilon_i}_{i,t} = -\boldsymbol{\kappa}_i(V_{i,t})\delta_i(\varepsilon_i)dt + J_{i,t}dN^{\varepsilon_i}_{i,t}, \tag{2.11}$$

where, $dN_{i,t}^{\varepsilon_i}$ is a Poisson process with jump intensity $\boldsymbol{\kappa}_i(V_{i,t})\lambda_i(\varepsilon_i)$. The probability density of the size of the big jumps $J_{i,t}$ (conditional on $|J_{i,t}| \geq \varepsilon_i$) is given as

$$p(J_{i,t}) = \frac{\boldsymbol{\kappa}_i(V_{i,t})\nu_i(J_{i,t})}{\boldsymbol{\kappa}_i(V_{i,t})\lambda_i(\varepsilon_i)} = \frac{\nu_i(J_{i,t})}{\lambda_i(\varepsilon_i)}. \tag{2.12}$$

When Levy process is finite variation, the approximation $L_{i,t}^{\varepsilon_i}$ ignores the infinite number of small jumps, and models the finite number of big jumps of size $\geq \varepsilon_i$.

Case 3. Infinite jump activity Levy processes with infinite variation.
Under this case, the approximated Levy process, $L_{i,t}^{\varepsilon_i}$, is defined by the following triplet

$$\left[\boldsymbol{\kappa}_i(V_{i,t})u_i^2(\varepsilon_i), \quad \boldsymbol{\kappa}_i(V_{i,t})\nu_i 1_{|y|\geq\varepsilon_i}, \quad \gamma_i(V_{i,t},\varepsilon_i) - \boldsymbol{\kappa}(V_{i,t})u_i^2(\varepsilon_i)/2\right], \tag{2.13}$$

where,

$$u_i^2(\varepsilon_i) = \int_{-\varepsilon_i}^{\varepsilon_i} y^2 \nu_i(dy). \tag{2.14}$$

And the stochastic differential of the Levy process is

$$dL_{i,t}^{\varepsilon_i} = -\boldsymbol{\kappa}_i(V_{i,t})\left(\delta_i(\varepsilon_i) + \frac{1}{2}u_i^2(\varepsilon_i)\right)dt + u_i(\varepsilon_i)\sqrt{\boldsymbol{\kappa}_i(V_{i,t})}dU_{i,t} + J_{i,t}dN_{i,t}^{\varepsilon_i}, \tag{2.15}$$

where, $dU_{i,t}, (i = 1, 2)$ are mutually independent Wiener processes that are independent of all other processes. The Poisson process $dN_{i,t}^{\varepsilon_i}$ and the size of the jump $J_{i,t}, (|J_{i,t}| \geq \varepsilon_i)$ are defined exactly as in Case 2. However, unlike in Case 2, the small jumps of size less than ε_i cannot be ignored in Case 3, and are modeled by using the Brownian motion given by the second term in equation (2.15).

Note that by setting $u_i(\varepsilon_i) = 0$ with $\varepsilon_i > 0$ in equations (2.13) and (2.15), we obtain equations (2.10) and (2.11), respectively, which are consistent with Case 2.

Similarly, by setting $\varepsilon_i = 0$ (which implies $u_i(\varepsilon_i) = 0$) in equations (2.13) and (2.15), we obtain equations (2.7) and (2.8), respectively, which are consistent with Case 1. Hence, we will focus on Case 3 only from now on. By appropriate substitutions for ε_i and $u_i(\varepsilon_i)$, the results apply to both Case 1 and Case 2. A multi-dimensional transform is presented in the next sub-section for the approximated Levy models, which could obtain option price efficiently, using recombining lattices approach.

2.2 A Multi-dimensional Transform for Approximated Levy Models with Stochastic Volatility

Replacing the Levy processes in equation (2.1) by using their approximations in equation (2.15), the log-stock price process will be governed by the SDE

$$d\ln S_t = \left(r - \frac{1}{2}\psi_1^2(V_{1,t},t) - \frac{1}{2}\psi_2^2(V_{2,t},t)\right)dt + \psi_1(V_{1,t},t)dZ_{1,t} + \psi_2(V_{2,t},t)dZ_{2,t} + dL_{1,t}^{\varepsilon_1} + dL_{2,t}^{\varepsilon_2}. \quad (2.16)$$

The rest of this paper assumes that the above specification of stock price process as the exogenously given true process for the purpose of pricing options. We have demonstrated numerically that for small enough values of ε_i, the option prices converge to the corresponding quasi-analytical solutions with $\varepsilon_i = 0$. This is also proved in Cont and Voltchkova (2005) (see Proposition 12.10 in Cont and Tankov (2003)). However, our convergence theorem below assume specific values of ε_i close to zero are exogenously given and represent the true stock price process.

The following proposition gives a multi-dimensional transform for Levy models which makes correlated state variables into conditionally independent state variables. The transform allows us to construct path independent lattices for Levy models which include up to two volatility factors and/or jump factors.

Proposition 2.2.1. For all Levy processes nested in equations (2.16), (1.2), and (1.3), and let $\mathbf{V}_t = [V_{1,t}, V_{2,t}]$, the multi-dimensional transform given as:

$$F(S_t, \mathbf{V}_t, t) = \ln(S_t) - \sum_{i=1}^{2} \rho_i \int^{V_{i,t}} \frac{\psi_i(\xi, t)}{\varphi_i(\xi, t)} d\xi \tag{2.17}$$

is conditionally independent of each of the volatility processes, i.e.,

$$dF(S_t, \mathbf{V}_t, t) dV_{i,t} = 0, \quad \text{for } i = 1, 2. \tag{2.18}$$

Proof. Using Ito's lemma, the stochastic process for F is given as

$$\begin{aligned} dF(S_t, \mathbf{V}_t, t) = \mu_{\text{F}}(\mathbf{V}_t, t)dt + \sum_{i=1}^{2} \psi_i(V_{i,t}, t)(dZ_{i,t} - \rho_i dW_{i,t}) \\ + \sum_{i=1}^{2} u_i(\varepsilon_i)\sqrt{\boldsymbol{\kappa}_i(V_{i,t})} dU_{i,t} + \sum_{i=1}^{2} J_{i,t} dN_{i,t}^{\varepsilon_i}, \end{aligned} \tag{2.19}$$

where,

$$\mu_{\text{F}}(\mathbf{V}_t, t) = \mu(\mathbf{V}_t, t) - \sum_{i=1}^{2} \boldsymbol{\kappa}_i(V_{i,t}) \left(\delta_i(\varepsilon_i) + \frac{1}{2} u_i^2(\varepsilon_i) \right), \tag{2.20}$$

and $\mu(\mathbf{V}_t, t)$ is defined in equation (1.7).

Equation (2.18) follows by multiplying equations (2.19) and (1.2), and using the definition of the correlation terms in equation (1.3), for $i = 1, 2$, and noting that Brownian motions $dU_{i,t}$, jump size variables $J_{i,t}$, and Poisson variables $dN_{i,t}^{\varepsilon_i}$ are all mutually independent (for $i = 1, 2$) and independent of all other processes. Hence, the multi-dimensional transform $F(S_t, \mathbf{V}_t, t)$ is conditionally independent of the volatility processes, $V_{i,t}$, for $i = 1, 2$. The definition of $F(S_t, \mathbf{V}_t, t)$ in equation (2.17) contains two terms. The first term is the standard log transformation of the stock price, which is needed for the construction of a path-independent tree for the stock price process even in absence of the volatility factors, as in the case of a path-independent

binomial tree construction for the standard BS model. The second term eliminates the conditional correlation between $F(S_t, \mathbf{V}_t, t)$ and the two state variables $V_{i,t}$. $\square$

From equation (2.19), the local second moment of F_t-process can be obtained, which is given as

$$g_{\mathrm{F}}^2(\mathbf{V}_t, t) = \sum_{i=1}^{2} \left[\psi_i^2(V_{i,t}, t)(1 - \rho_i^2) + \frac{1}{2}\boldsymbol{\kappa}_i(V_{i,t}) u_i^2(\varepsilon_i) \right]. \tag{2.21}$$

When the multi-dimensional transform (2.17) is applied to Levy models in recombining lattices method (which we will present later in this paper), will the discrete stochastic process converge to the corresponding continuous stochastic process? The convergence theorem in next session guarantees that the discrete approximation of the stochastic process will converge to its corresponding continuous one.

2.3 Convergence Theorem for Approximated Levy Models with Stochastic Volatility

We approximate the Levy jump processes using compound Poisson processes and Brownian motions, as shown in equation (2.15) and (2.16). Taking stochastic integral on both sides of equation (2.19), the transform F_t can be expressed as

$$F(S_t, \mathbf{V}_t, t) = F_t = F_C(t) + F_J(t), \tag{2.22}$$

where,

$$\begin{aligned} F_C(t) &= F(S_0, \mathbf{V}_0, 0) \\ &+ \int_0^t \left[\mu_{\mathrm{F}}(\mathbf{V}_s, s)ds + \sum_{i=1}^{2} \psi_i(V_{i,s}, s)(dZ_{i,s} - \rho_i dW_{i,s}) + \sum_{i=1}^{2} u_i(\varepsilon_i)\sqrt{\boldsymbol{\kappa}_i(V_{i,s})}dU_{i,s} \right], \end{aligned} \tag{2.23}$$

and

$$F_J(t) = \int_0^t \left(\sum_{i=1}^{2} J_{i,t} dN_{i,s}^{\varepsilon_i} \right). \tag{2.24}$$

Consider ${}_hF_t$, ${}_hF_{C,t}$, and ${}_hF_{J,t}$ as the discrete approximations of F_t, $F_C(t)$, and $F_J(t)$, respectively, also consider $\mu_{F,h}$ and $g_{F,h}$ as the discrete approximations of $\mu_{\mathrm{F}}(\mathbf{V}_t, t)$ and $g_{\mathrm{F}}(\mathbf{V}_t, t)$ (defined in equation (2.20) and (2.21)), respectively, then follow the same procedure as in 1.1.1, we have the following convergence theorem for Levy jump processes.

Theorem 2.3.1. [1] Under assumptions A1 through A4, the discrete multi-dimensional process $\{{}_hF_t, {}_h\mathbf{V}_t\}$ converge weakly to the continuous multi-dimensional process $\{F_t, \mathbf{V}_t\}$ as $h \to 0$.

After the multi-dimensional transform and convergence theorem have been established, we will apply them to Levy models with stochastic volatility to evaluate both American and European option prices.

2.4 General Quasi-Analytical solution for Levy Models

Before we apply the multi-dimensional transform to construct path-independent trees for Levy models, a general quasi-analytical (closed-form) solution is derived in this sub-section. We assume the following specifications to generate different affine Levy models and derive the corresponding quasi-analytical solutions. Let

$$\begin{aligned} \psi_i(V_{i,t}, t) &= \sqrt{V_{i,t}}, & \phi_i(V_{i,t}, t) &= \alpha_i(m_i - V_{i,t}), \\ \varphi_i(V_{i,t}, t) &= \sigma_i\sqrt{V_{i,t}}, & \boldsymbol{\kappa}_i(V_{i,t}) &= a_i + b_i V_{i,t}. \end{aligned} \tag{2.25}$$

[1] The proof is not given here. Readers can refer to Amin (1993), which says that the convergence is enough if the probability distribution and first two moments are matching.

Combining equation (2.25) with equations (2.15) and (2.16), we can obtain the stochastic process for the log-price $x_t = \ln(S_t)$ as follows

$$dx_t = \left[r - \sum_{i=1}^{2} \left(\frac{1}{2} V_{i,t} - (a_i + b_i V_{i,t}) \left(\delta_i(\varepsilon_i) + \frac{1}{2} u_i^2(\varepsilon_i) \right) \right) \right] dt + \sum_{i=1}^{2} \left(\sqrt{V_{i,t}} dZ_{i,t} + u_i(\varepsilon_i) \sqrt{a_i + b_i V_{i,t}} dU_{i,t} + J_i dN_{i,t}^{\varepsilon_i} \right). \quad (2.26)$$

Suppose the characteristic function of x_t in equation (2.26) has the format of

$$\begin{aligned} f(x_t, V_{1,t}, V_{2,t}, \varepsilon_{1,t}, \varepsilon_{2,t}, \tau, \varsigma) &= E_t \left[\exp(\imath \varsigma x_t) \right], \\ &= \exp\left[C(\tau, \varsigma) + D_1(\tau, \varsigma) V_{1,t} + D_2(\tau, \varsigma) V_{2,t} + \imath \varsigma x_t \right], \end{aligned} \quad (2.27)$$

where, $\imath^2 = -1$, $\tau = T - t$ is the time to maturity of an option, and the initial conditions are $C(0, \varsigma) = 0$, $D_1(0, \varsigma) = 0$, and $D_2(0, \varsigma) = 0$.

Following Heston (1993), Appendix in Bakshi et al. (1997), and Secion 15.4 in Cont and Tankov (2003), we compute the characteristic function (2.27) in Appendix B. After the characteristic function is known, the price of European call option with maturity τ and exercise price K can be calculate as (see equation (8) in Bakshi et al. (1997), and others)

$$C_t = S_t \Pi_1 - K e^{-r\tau} \Pi_2, \quad (2.28)$$

where,

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left[\frac{e^{-\imath \varsigma \ln(K)} f(x_t, V_{1,t}, V_{2,t}, \varepsilon_{1,t}, \varepsilon_{2,t}, \tau, \varsigma - \imath)}{\imath \varsigma S_t e^{r\tau}} \right] d\varsigma, \quad (2.29)$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left[\frac{e^{-\imath \varsigma \ln(K)} f(x_t, V_{1,t}, V_{2,t}, \varepsilon_{1,t}, \varepsilon_{2,t}, \tau, \varsigma)}{\imath \varsigma} \right] d\varsigma. \quad (2.30)$$

Using put-call parity, we can obtain the price for European put options. Note that even though we use the compound Poisson framework (a small number ε_i is included) to derive closed-from solution for European call options in equation (2.28), for the quasi-analytical solution for Levy models, we set $\varepsilon_i = 0$ (2.28). In the path-independent lattice method (see section 2.5), to avoid explosion in the numerical computation, we set ε_i to be a small positive number, say 10^{-4}.

2.5 Construction of Path-Independent Trees for Levy Models

The multi-dimensional transform and convergence theorem make it possible to construct path-independent trees for Levy models to price both American and European options. In the following two subsections, we demonstrate the construction of both trees. The tree construction for Levy models with one volatility state factor and one Levy jump process can be extended to Levy models with two volatility state factor and/or two Levy jump processes in a straightforward manner. For simplicity, from now on we focus our discussion on one-factor Levy models only, i.e., we will drop the index notation i in the corresponding equations and variables, and denote the volatility process as V_t, whose process has the following format

$$dV_t = \phi(V_t, t)dt + \varphi(V_t, t)dW_t. \tag{2.31}$$

2.5.1 Construction of Path-Independent Tree for Volatility Process

The construction of a path-independent tree for the volatility process is almost identical to the one outlined in section 1.2.1 for stochastic volatility models, except that truncation from the above and below of the tree is applied.

Under the general specifications of equation (2.25), the one-factor volatility process (2.31) can be written as

$$dV_t = \mu_{\mathrm{V}}(V_t, t)dt + g_{\mathrm{V}}(V_t, t)dW_t. \tag{2.32}$$

Here,

$$\mu_{\mathrm{V}}(V_t, t) = \alpha(m - V_t), \tag{2.33}$$

$$g_{\mathrm{V}}(V_t, t) = \sigma\sqrt{V_t}. \tag{2.34}$$

The discretization for the volatility process (2.32) will follow the same procedure as in section 1.1.1. In the trinomial tree, if at time step k ($t = hk$) and at any spacial node the volatility is V_{hk}, then the value $V_{h(k+1)}$ of volatility at next time step $k+1$ will follow the trinomial model

$$V_{h(k+1)} = \begin{cases} V_{hk}^u = V_{hk} + J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}\sqrt{h}, & \text{with probability } p_{\mathrm{V}}^u, \\ V_{hk}^m = V_{hk}, & \text{with probability } p_{\mathrm{V}}^m, \\ V_{hk}^d = V_{hk} - J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}\sqrt{h}, & \text{with probability } p_{\mathrm{V}}^d. \end{cases} \tag{2.35}$$

Here, $J_{\mathrm{V}}(V_t, t)$ is the smallest positive integer such that

$$\left(\frac{g_{\mathrm{V}}(V_t, t)}{J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}}\right)^2 + \frac{|\mu_{\mathrm{V}}(V_t, t)|}{J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h} \leq 1, \tag{2.36}$$

where, $\bar{\sigma}_{\mathrm{V}} = \sigma\sqrt{V_0}$, and V_0 is the value of volatility at time $t = 0$. The probabilities of up, middle, and down movements are defined as

$$\begin{aligned} p_{\mathrm{V}}^u &= \frac{1}{2}\left(\frac{g_{\mathrm{V}}(V_t, t)}{J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}}\right)^2 + \frac{1}{2}\frac{\mu_{\mathrm{V}}(V_t, t)}{J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h}, \\ p_{\mathrm{V}}^m &= 1 - \left(\frac{g_{\mathrm{V}}(V_t, t)}{J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}}\right)^2, \\ p_{\mathrm{V}}^d &= \frac{1}{2}\left(\frac{g_{\mathrm{V}}(V_t, t)}{J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}}\right)^2 - \frac{1}{2}\frac{\mu_{\mathrm{V}}(V_t, t)}{J_{\mathrm{V}}(V_t, t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h}. \end{aligned} \tag{2.37}$$

Here, the selection of $J_{\mathrm{V}}(V_t,t)$ (see equation (2.36)) ensures that $0 \leq p_{\mathrm{V}}^m \leq 1$, $p_{\mathrm{V}}^u \leq 1$, and $p_{\mathrm{V}}^d \leq 1$. Generally, p_{V}^u and p_{V}^d will also remain non-negative with small value of h. However, under stochastic volatility, since $g_{\mathrm{V}}(V_t,t)$ can become arbitrarily small (or even be zero with a non-zero probability as in Heston model) and/or $|\mu_{\mathrm{V}}(V_t,t)|$ can become large, there is no guarantee that p_{V}^u and p_{V}^d will always remain non-negative for very small h. To ensure non-negativity of p_{V}^u and p_{V}^d, we define the indicator function I_{V} as

$$I_{\mathrm{V}} = \begin{cases} 1, & \text{if } \mu_{\mathrm{V}}(V_t,t) > 0, \\ 0, & \text{otherwise }, \end{cases}$$

and replace the probabilities in equations (2.37) with

$$\begin{aligned} p_{\mathrm{V}}^u &= \frac{1}{2}\left(\frac{g_{\mathrm{V}}(V_t,t)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\right)^2 + \frac{1}{2}\frac{\mu_{\mathrm{V}}(V_t,t)I_{\mathrm{V}}}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h}, \\ p_{\mathrm{V}}^m &= 1 - \left(\frac{g_{\mathrm{V}}(V_t,t)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\right)^2 - \frac{1}{2}\frac{\mu_{\mathrm{V}}(V_t,t)(2I_{\mathrm{V}}-1)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h}, \\ p_{\mathrm{V}}^d &= \frac{1}{2}\left(\frac{g_{\mathrm{V}}(V_t,t)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\right)^2 - \frac{1}{2}\frac{\mu_{\mathrm{V}}(V_t,t)(1-I_{\mathrm{V}})}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h}, \end{aligned} \tag{2.38}$$

if the following condition is satisfied

$$\left(\frac{g_{\mathrm{V}}(V_t,t)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\right)^2 < \frac{|\mu_{\mathrm{V}}(V_t,t)|}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h}. \tag{2.39}$$

To accelerate the speed of convergence, the tree is truncated both from below and from above, which saves considerable computational time. At the bottom, it is required that volatility is greater than zero. The tree is truncated from the above when the total number of nodes at any time step is more than a preset number V_n. Usually, V_n is a constant between 51 and 101. If V_n is too small, then it will increase efficiency but lose accuracy; if V_n is too large, the computational cost is too high.

When the tree is truncated from below, we set the value of down node to zero, and the probability of down movement is added to the probability of middle movement. That is, when the tree is truncated from below, eqations (2.35) and (2.37) are replaced with

$$V_{h(k+1)} = \begin{cases} V_{hk}^u = V_{hk} + J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}\sqrt{h}, & \text{with probability } p_{\mathrm{V}}^u, \\ V_{hk}^m = V_{hk}, & \text{with probability } p_{\mathrm{V}}^m, \\ V_{hk}^d = 0, & \text{with probability } p_{\mathrm{V}}^d, \end{cases} \tag{2.40}$$

and

$$\begin{aligned} p_{\mathrm{V}}^u &= \frac{1}{2}\left(\frac{g_{\mathrm{V}}(V_t,t)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\right)^2 + \frac{1}{2}\frac{\mu_{\mathrm{V}}(V_t,t)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h}, \\ p_{\mathrm{V}}^m &= 1 - p_{\mathrm{V}}^u, \\ p_{\mathrm{V}}^d &= 0. \end{aligned} \tag{2.41}$$

When the tree is truncated from above, we set the value of up node to zero, and the probability of up movement is added to the probability of middle movement. That is, when the tree is truncated from above, eqations (2.35) and (2.37) are replaced with

$$V_{h(k+1)} = \begin{cases} V_{hk}^u = 0, & \text{with probability } p_{\mathrm{V}}^u, \\ V_{hk}^m = V_{hk}, & \text{with probability } p_{\mathrm{V}}^m, \\ V_{hk}^d = V_{hk} - J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}\sqrt{h}, & \text{with probability } p_{\mathrm{V}}^d. \end{cases} \tag{2.42}$$

and

$$\begin{aligned} p_{\mathrm{V}}^u &= 0, \\ p_{\mathrm{V}}^m &= 1 - p_{\mathrm{V}}^d, \\ p_{\mathrm{V}}^d &= \frac{1}{2}\left(\frac{g_{\mathrm{V}}(V_t,t)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\right)^2 - \frac{1}{2}\frac{\mu_{\mathrm{V}}(V_t,t)}{J_{\mathrm{V}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\sqrt{h}. \end{aligned} \tag{2.43}$$

The tree constructed for the volatility process is path-independent, which will be combined into one 2-dimensional tree with the path-independent tree for transformed process.

2.5.2 Construction of Path-Independent Tree for the Transformed Process

Now we discuss the lattice construction for the transformed process. Under the one-factor version of general specification (2.25), the transformed process follows the dynamic

$$dF_t = \mu_{\text{F}}(V_t, t)dt + g_{\text{F}}(V_t, t)dB_t + J_t dN_t^{\varepsilon}, \tag{2.44}$$

where,

$$\mu_{\text{F}}(V_t, t) = r - \frac{\rho}{\sigma}\alpha m + V_t\left(\frac{\rho\alpha}{\sigma} - \frac{1}{2}\right) - (a + bV_t)\left(\delta(\varepsilon) + \frac{1}{2}u_{\varepsilon}^2\right), \tag{2.45}$$

$$g_{\text{F}}(V_t, t) = \sqrt{V_t(1-\rho^2) + (a+bV_t)\frac{u^2(\varepsilon)}{2}}, \tag{2.46}$$

$$dB_t = \sqrt{V_t}(dZ_t - \rho dW_t) + u(\varepsilon)\sqrt{a + bV_t}dU_t. \tag{2.47}$$

The corresponding discrete-time approximation of equation (2.44) is

$$\Delta F_t = \mu_{\text{F}}(V_t, t)\Delta t + g_{\text{F}}(V_t, t)\Delta B_t + J_t \Delta N_t^{\varepsilon}, \tag{2.48}$$

where, the term ΔB_t represents the discrete-time approximation of the continuous change in the Brownian motion dB_t, and ΔN_t^{ε} represent the discrete-time approximations of the continuous changes in the Poisson process dN_t^{ε}. The value of

ΔN_t^{ε} is assumed to be 1 with probability $(a+bV_t)\lambda(\varepsilon)\Delta t$, and 0 with probability $1-(a+bV_t)\lambda(\varepsilon)\Delta t$. This implies that equation (2.48) can be approximated as follows

$$\Delta F_t = \begin{cases} \mu_{\mathrm{F}}(V_t,t)\Delta t + g_{\mathrm{F}}(V_t,t)\Delta B_t, & \text{with probability } 1-(a+bV_t)\lambda(\varepsilon)\Delta t, \\ J_t, & \text{with probability } (a+bV_t)\lambda(\varepsilon)\Delta t. \end{cases} \tag{2.49}$$

In order to approximate the F_t-process in equation (2.44), a multinomial jump diffusion tree is constructed. To superimpose jump part onto diffusion part, at any time step $t = hk$ at any node in the jump diffusion tree, there are F_n branches emanating in the next time step $t = h(k+1)$, as shown in Figure 1. Here, F_n must be an odd positive integer number.

At time step $t = h(k+1)$, the value at each node emanating out from node F_{hk} is computed as

$$\begin{aligned} F^l_{h(k+1)} &= F_{hk} + \left(\frac{F_n+1}{2} - l\right)\bar{\sigma}_{\mathrm{F}}\sqrt{h}, \\ \bar{\sigma}_{\mathrm{F}} &= \sqrt{V_0(1-\rho^2) + (a+bV_0)\frac{u^2(\varepsilon)}{2}}. \end{aligned} \tag{2.50}$$

Here, $l = 1, 2, 3, \cdots, F_n$ is the position (level) of the node at time $t = h(k+1)$ (see Figure 1). The diffusion part in equation (2.44) is approximated by a trinomial model. And the value of the three diffusion nodes (among the F_n nodes at time $t = h(k+1)$) are defined as

$$\begin{cases} F^u_{hk} = F_{hk} + J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}\sqrt{h}, & \text{with probability } p^u_{\mathrm{F}}, \\ F^m_{hk} = F_{hk}, & \text{with probability } p^m_{\mathrm{F}}, \\ F^d_{hk} = F_{hk} - J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}\sqrt{h}, & \text{with probability } p^d_{\mathrm{F}}. \end{cases} \tag{2.51}$$

Here, $J_{\mathrm{F}}(V_t,t)$ is the smallest positive integer such that

$$\left(\frac{g_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\right)^2+\frac{|\mu_{\mathrm{F}}(V_t,t)|}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\sqrt{h}\leq 1, \tag{2.52}$$

The probabilities of up, middle, and down movements conditional on no-jump occurrence sum up to 1, and are defined as

$$\begin{aligned}
p_{\mathrm{F}}^{*u}&=\frac{1}{2}\left(\frac{g_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\right)^2+\frac{1}{2}\frac{\mu_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\sqrt{h},\\
p_{\mathrm{F}}^{*m}&=1-\left(\frac{g_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\right)^2,\\
p_{\mathrm{F}}^{*d}&=\frac{1}{2}\left(\frac{g_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\right)^2-\frac{1}{2}\frac{\mu_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\sqrt{h}.
\end{aligned} \tag{2.53}$$

Here, the selection of $J_{\mathrm{F}}(V_t,t)$ (equation (2.52)) ensures that $0\leq p_{\mathrm{F}}^{*m}\leq 1$, $p_{\mathrm{F}}^{*u}\leq 1$, and $p_{\mathrm{F}}^{*d}\leq 1$. Generally, p_{F}^{*u} and p_{F}^{*d} will also remain non-negative with small enough time step value of h. However, under stochastic volatility, since $g_{\mathrm{F}}(V_t,t)$ can become arbitrarily small (or even go to zero with a non-zero probability as in Heston model) and/or $|\mu_{\mathrm{F}}(V_t,t)|$ can become large, there is no guarantee that p_{F}^{*u} and p_{F}^{*d} will always remain non-negative for any h. To ensure non-negativity of p_{F}^{*u} and p_{F}^{*d}, we introduce a indicator function I_{F} as

$$I_{\mathrm{F}}=\begin{cases}1, & \text{if } \mu_{\mathrm{F}}(V_t,t)>0,\\ 0, & \text{otherwise ,}\end{cases}$$

and replace the probabilities in equations (2.53) with

$$\begin{aligned}
p_{\mathrm{F}}^{*u}&=\frac{1}{2}\left(\frac{g_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\right)^2+\frac{1}{2}\frac{\mu_{\mathrm{F}}(V_t,t)I_{\mathrm{F}}}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\sqrt{h},\\
p_{\mathrm{F}}^{*m}&=1-\left(\frac{g_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\right)^2-\frac{1}{2}\frac{\mu_{\mathrm{F}}(V_t,t)(2I_{\mathrm{F}}-1)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\sqrt{h},\\
p_{\mathrm{F}}^{*d}&=\frac{1}{2}\left(\frac{g_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{V}}}\right)^2-\frac{1}{2}\frac{\mu_{\mathrm{F}}(V_t,t)(1-I_{\mathrm{F}})}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\sqrt{h}.
\end{aligned} \tag{2.54}$$

if the following condition is satisfied

$$\left(\frac{g_{\mathrm{F}}(V_t,t)}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\right)^2 < \frac{|\mu_{\mathrm{F}}(V_t,t)|}{J_{\mathrm{F}}(V_t,t)\bar{\sigma}_{\mathrm{F}}}\sqrt{h}. \tag{2.55}$$

Since the Poisson jumps occur with probability $(a+bV_t)\lambda(\varepsilon)\Delta t$, the unconditional probabilities of the three diffusion nodes are give as

$$\begin{aligned} p_{\mathrm{F}}^{u} &= p_{\mathrm{F}}^{*u}\Big(1-(a+bV_t)\lambda(\varepsilon)\Delta t\Big), \\ p_{\mathrm{F}}^{m} &= p_{\mathrm{F}}^{*m}\Big(1-(a+bV_t)\lambda(\varepsilon)\Delta t\Big), \\ p_{\mathrm{F}}^{d} &= p_{\mathrm{F}}^{*d}\Big(1-(a+bV_t)\lambda(\varepsilon)\Delta t\Big). \end{aligned} \tag{2.56}$$

And the procedure to calculate the probability for each jump node at time $t = h(k+1)$ is as follows. Define the value of y at each node at time step $t = h(k+1)$ (see Figure 1) for the Levy intensity $\nu(dy)$ as

$$y^l = \left(\frac{F_n+1}{2} - l\right)\bar{\sigma}_{\mathrm{F}}\sqrt{h}, \quad l = 1,2,3,\cdots,F_n, \tag{2.57}$$

and suppose $\jmath$ is the position of one jump node at time $t = h(k+1)$, $\jmath^+$ is the jump node just above node $\jmath$, and $\jmath^-$ is the jump node just below node $\jmath$. Here, $1 \le \jmath \le F_n$, but, to exclude the three diffusion nodes, they are not equal to $\frac{F_n+1}{2} - J_{\mathrm{F}}(V_t,t)$, $\frac{F_n+1}{2}$, and $\frac{F_n+1}{2} + J_{\mathrm{F}}(V_t,t)$, respectively, then, in general, we can calculate the probability of jump node $\jmath$ as

$$\begin{aligned} \text{Prob(jump node } \jmath) &= \frac{\int_{y_1}^{y_2}\nu(dy)}{\lambda(\varepsilon)}\,(a+bV_t)\,\lambda(\varepsilon)\Delta t, \\ &= \left(\int_{y_1}^{y_2}\nu(dy)\right)(a+bV_t)\,\Delta t. \end{aligned} \tag{2.58}$$

Here, y_1 and y_2 are defined as (following Amin (1993))

- If node $\jmath$ is the most top jump node, $y_1 = \frac{1}{2}\left(y^{\jmath} + y^{\jmath^-}\right)$, and $y_2 = +\infty$.
- If node $\jmath$ is the first jump node above the middle diffusion node, $y_1 = +\varepsilon$, and $y_2 = \frac{1}{2}\left(y^{\jmath} + y^{\jmath^-}\right)$.
- If node $\jmath$ is the first jump node below the middle diffusion node, $y_1 = \frac{1}{2}\left(y^{\jmath} + y^{\jmath^+}\right)$, and $y_2 = -\varepsilon$.
- If node $\jmath$ is the most bottom node, $y_1 = -\infty$, and $y_2 = \frac{1}{2}\left(y^{\jmath} + y^{\jmath^-}\right)$.
- Otherwise, $y_1 = \frac{1}{2}\left(y^{\jmath} + y^{\jmath^-}\right)$, and $y_2 = \frac{1}{2}\left(y^{\jmath} + y^{\jmath^+}\right)$.

To increase the efficiency, the tree for F_t-process is truncated both from top and from bottom. At any time step, the total number of nodes in the F_t-tree is F_n. If the up-diffusion node is truncated, the probability of up-diffusion node is added to the probability of middle-diffusion node. If the down-diffusion node is truncated, the probability of down-diffusion node is added to the probability of middle-diffusion node. If jump nodes are truncated from the top, then the probability of all truncated jump nodes are added to the first top jump node. If jump nodes are truncated from the bottom, then the probability of all truncated jump nodes are added to the last bottom jump node.

Once V_t-tree and F_t-tree are obtained, the two trees must be combined to generate the assumed correlation structure between the two processes into a two-dimensional tree to calculate option price. At any time step $t = hk$, for each node V_{hk} in V_t-tree, each of the three branches $\left(V_{hk}^u, V_{hk}^m, \text{ and } V_{hk}^d\right)$ emanating out from V_{hk} will combine with the F_n branches emanating out from F_{hk} in F_t-tree to generate the nodes at next time step $t = h(k+1)$ in the two-dimensional tree (see Figure 2). The stock price at each node in the two-dimensional tree is computed by the inverse transform of equation (2.17). Under the general specification (2.25), it is $S_t = \exp\left(F_t + \frac{\rho}{\sigma}V_t\right)$.

The probability for each node in the two-dimensional tree (joint probability) is the product of corresponding marginal probabilities of the V_t-tree and the F_t-tree, as

shown in Figure 2, because the V_t-tree and the F_t-tree are uncorrelated. The nodes which are truncated in V_t-tree and F_t-tree are not considered in the two-dimensional tree. So, at any time step, at most, there are $V_n \times F_n$ nodes. In our numerical tests (see next section), we use $V_n = 51, F_n = 201$, and the number of time step $T_n = 181$. Keeping V_n, F_n, and T_n at an optimal range will increase the efficiency.

2.6 Numerical results of Levy models

In this section, we will apply the quasi-analytical solution (in Section 2.4) and transform-based tree method (in Section 2.5) to some particular interesting Levy models to price both European and American options. The five Levy models considered in this paper are listed in Appendix C. The computations have been performed for put options with strike price \$90, \$100, and \$110, maturing in 1, 3, and 6 months for the starting volatility value of $\sqrt{V(0)} = 0.1, 0.2, 0.3$ and starting stock price of $S(0) = \$100$.

The results for the Pan model, the VGsv model, and the CGMYsv model are presented in Table 5 through Table 7. In each table, it shows European and American put option prices obtained using transform-based tree method and quasi-analytical (closed-form) solution of each model for the case of European put options. The model parameters are set as follows, respectively.

- Pan model: $r = 0.03$, $\sigma = 0.2087$, $\alpha = 10$, $m = 0.0032$, $\rho = -0.99$, $a = 0$, $b = 21.51$, $\mu_J = -0.0356$, and $\sigma_J = 0.0286$.

- VGsv model: $r = 0.03$, $\sigma = 0.3433$, $\alpha = 10$, $m = 0.0059$, $\rho = -0.99$, $a = 0.1$, $b = 10$, $C = 13.2626$, $G = 12.6496$, and $M = 34.9055$.

- CGMYsv model: $r = 0.03$, $\sigma = 0.5457$, $\alpha = 10$, $m = 0.0149$, $\rho = -0.99$, $a = 0.1$, $b = 5.3518$, $C_p = 500$, $G = 16.9655$, $M = 612.4486$, $Y_p = -77.1569$, $Y_n = -36.1443$ and $\zeta = 1.2771$.

For tree method, we divide the maturity of each options into 181 time steps. After the truncation, the maximum number of nodes in volatility process tree at each time step is 51, and the maximum number of nodes in transformed process tree at each time step is 201.

The striking result here is that European put option prices obtained using the path-independent lattice converge rapidly to the closed-form solution. For each put option, it takes less than 15 seconds for the tree method converging to analytical solution for each model.

CHAPTER 3

EMPIRICAL PERFORMANCE OF LEVY MODELS WITH STOCHASTIC VOLATILITY FOR AMERICAN OPTIONS

The multi-dimensional transform developed in the previous parts can be used to construct path-independent lattices for all low-dimensional stochastic volatility models with or without jump given in the literature. The drawback of the BS model has been well documented in the literatures, such as negative skewness with excess kurtosis, fat tail, and non-normality. Therefore, many models have been proposed to resolve known issues associated the BS model. For example, to resolve volatility smile, local volatility, stochastic volatility, and diffusion with jumps have been considered in the literatures; to resolve non-normality, non-Markov processes have been considered, e.g., Poisson process, variance gamma process, and other type of Levy processes. One would ask: what is the gain from each of the generalized models? Or, which model is the best for option pricing? Bakshi et al. (1997) examine BS, SV, SVJ and other models using S&P 500 Index options (European style contracts) from three perspectives: in-sample consistency of implied volatility, out-of-sample pricing error, and hedging error. They find that incorporating SV and jumps is important for pricing and internal consistency, but for hedging, modeling SV alone yields the best performance.

In this part, we will investigate that which model results in the lowest pricing error for American style contracts. The following seven models are considered here: the BS model, the Heston model, the Pan model, the VG model, the CGMY model, the VGsv model, and the CGMYsv model. The data used in the empirical study is

described in Section 3.1. The procedure and methodology we use are introduced in Section 3.2. And the empirical results are presented in Section 3.3.

3.1 Data

SPY pays quarterly cash dividend, which is discrete. When considering discrete dividend payment using tree methods, the underling price will be reduced the same amount as dividend on ex-dividend date, which will make the tree structure much more complicated: the tree will not recombine after dividend payment and the number of nodes will increase dramatically. For simplification, we consider options with only one discrete dividend payment during the remaining time to expire. The procedure described in this paper could be extended straightforward to options with multiple discrete dividend payments during the option life time. The four ex-dividend dates for SPY in 2014 are March 21, June 20, September 19, and December 19. To avoid January effect, and because the interest rate is relatively higher in March and April, our sample period is from March 24 to April 25. And the options' expiration date will be no later than September 19.

It is well known that the price of an American call and European call without dividends should not diverge; for put options, the price of American and European options can diverge even for underlying assets that do not pay dividends. Based on this consideration and to make our procedure more efficient, we use SPY call options in our empirical experiments, because we can use quasi-analytical solution for European options to calibrate model parameters for in-sample fitting, then use the calibrated model parameters to investigate out-of-sample pricing performance.

Therefore, we use SPY daily call options for our empirical study, which is one of the most actively traded American style contracts. The sample period is from March 24 to April 25 in 2014. The daily bid-ask quotes and SPY end of day price are obtained from OptionMetrics. The discrete dividend payment on SPY is also

obtained from OptionMetrics. Three-month treasury-bill rate is used as the risk-free rate, and obtained from CRSP.

Some exclusion filters are applied to construct SPY options sample data set. We generally follow Bakshi et al. (1997) to filter the data. Options with less than six days to expiration may include liquidity related bias, and so are excluded from the sample. Options which have trading volume less than 500 are removed from the sample because those options are not traded actively. To mitigate the impact of price discreteness on options valuation, bid prices lower than \$0.75 are not included. Finally, quotes not satisfying the arbitrage restriction

$$C_t \geq \max\big(0, \quad S_t - X, \quad S_t - Xe^{-r\tau} - \text{PV(dividend)}\big) \tag{3.1}$$

are taken out of the sample. Here call price C_t is computed as mid-point of bid and ask prices, and PV(dividend) is the cumulative present value at current time of dividend payment during the life of one call options. If there is no dividend payment during the life of the option, PV(dividend) is zero.

3.2 Procedure and Methodology

In the empirical part, our goal is to find out which model results the lowest out-of-sample pricing error. To do that, for each day in the sample period, we split the SPY call options into two groups: one group (in-sample data) is the options with maturity longer than six days and expiration date before June 20 in 2014 (which will be used as in-sample fitting), the other group (out-of-sample data) is options with expiration date on or after June 20 but before September 19 in 2014 (which will be used as out-of sample pricing).

3.2.1 In-sample fitting

For in-sample fitting, there is no dividend payment during the life of each call option, so SPY American call option price will not diverge from its equivalent European call option price (if there is European style contract on SPY). Therefore, we will use quasi-analytical solution of each model for European style call option when we calibrate model parameters using in-sample data. To estimate parameters for each model on a daily basis, we use the following objective function (which is widely used in the literature, e.g., Bakshi et al. (1997) and others)

$$\min\left(\mathrm{SSE}_t\right) = \min\left(\sum_{n=1}^{N}\left|C_n^{\mathrm{market}} - C_n^{\mathrm{model}}\right|^2\right), \tag{3.2}$$

which is to minimize the sum of squared dollar pricing error (SSE). Here, C_n^{market} is the observed call price for the n^{th} contract on day t in the sample, C_n^{model} is the call price for the n^{th} contract on day t computed using quasi-analytical solution for European style contract for each model, and N is the total number of contracts on day t.

3.2.2 Methodology

On each day t, we need to find parameters for each model so that the SSE_t in the objective function in equation (3.2) is minimized. The optimization problem in the objective function is not convex and the parameter space has dimensions of 4 to 13. The standard optimization techniques based on gradient search or direction search will be difficult to calibrate. Levenber-Marquardt method, gradient descent algorithm, Gauss-Newton algorithm, and Nelder-Mead algorithm are examples of standard optimization techniques. If applied to our objective function, they may fail to convergence, because they need good starting point; or they might be trapped in a local minimum. Here, we are looking for a global minimum, so we adopt a stochastic search method known as differential evolution (DE).

DE, introduced in Storn and Price (1997), is a type of genetic algorithm, which has a remarkable performance finding global optimization because it uses a random parameter population instead a single point in parameter space as in the standard optimization techniques. We will closely follow Vollrath and Wendland (2009) to briefly describe the DE algorithm here.

We put all parameters (needed to calibrate) of a model into a vector $p = (\beta_1, \beta_2, \cdots, \beta_n)$, and refer this vector as an individual. A sample of individuals is called a population: $p_1,\ p_2,\ \cdots, p_{NP}$, where $NP \geq 4$ is the population size. A modified population is called a generation. DE algorithm has the following four stages.

(1) *Initialization:* Give the starting population $p_i = (\beta_1, \beta_2, \cdots, \beta_{n-1}, \beta_n), i = 1, 2, \cdots, NP$ initial values. These values could be generated randomly, though good starting points will make DE converge faster. And NP is usually set to be $10 \times n$, n is the number of parameter to be calibrated. In the initialization, it is shown that, at one time, DE evolves a family of potential parameter set instead of one single point in the parameter space.

(2) *Mutation:* Each individual in a population is mutated as follows. Randomly select an individual p_i in the population, then select other three individuals randomly in the same population: $p_a,\ p_b$, and p_c, where $p_i,\ p_a,\ p_b$, and p_c are mutually different, a new individual is set as $p_i^{'} = \left(\beta_1^{'}, \beta_2^{'}, \cdots, \beta_{n-1}^{'}, \beta_n^{'}\right) = p_a + F\left(p_b - p_c\right)$. Here, $F \in (0, 2)$ is the mutation factor. After the mutation procedure is repeated NP times, a mutated population is created $p_1^{'},\ p_2^{'},\ \cdots,\ p_{NP}^{'}$.

(3) *Crossover:* After the mutation step, the crossover population is formed by combining the original population and its mutated population. For each $i, (1 \leq i \leq NP)$, a new individual is created $p_i^{''} = \left(\beta_1^{''}, \beta_2^{''}, \cdots, \beta_{n-1}^{''}, \beta_n^{''}\right)$. Here, $\beta_j^{''}$ is give by

$$\beta_j^{''} = \begin{cases} \beta_j^{'}, & \text{if } \eta \leq CR \text{ or } j = \iota, \\ \beta_j, & \text{if } \eta > CR \text{ and } j \neq \iota, \end{cases}$$

where, η is a number selected uniformly from the range $[0, 1]$, and ι is a randomly selected index number from 1, 2, $\cdots$, n. And $CR \in (0, 1)$ is the crossover ratio. The use of ι ensures that at least one mutated parameter is accepted. Then $p_1^{''}$, $p_2^{''}$, $\cdots$, $p_{NP}^{''}$ form the crossover population.

(4) *Selection:* At the final step, a new generation is created. For each $i = 1, 2, \cdots, NP$, the value of the objective function is computed for $p_i^{''}$ and p_i, respectively. The individual (p_i or $p_i^{''}$) who has smaller value is selected into the next generation.

Steps (2)-(4) are repeated until the DE algorithm converges, or the maximum number of generation is reached, or the maximum time is run out.

The DE algorithm has been shown to be the fastest evolutionary algorithm, and it can be parallelized to use on a cluster (super-computer). The values of F and CR are very important in determining how the algorithm behaves and need to be selected with care. In most cases, F and CR can be set to be constants, but it has been found that selecting F from the interval $[0.5, 1.0]$ randomly and CR from the interval $[0.2, 0.9]$ randomly for each generation or for each difference vector improves convergence behavior significantly.

3.2.3 Out-of-sample pricing

After using the DE algorithm to fit option pricing model with market data on a daily basis, we will obtain parameters for each model on each day in the sample period. Then, we can use the calibrated model parameters to examine each model's out-of-sample pricing performance.

On each day t for each model, we use the parameters calibrated on the same day to price each American style option in the out-of-sample data. Then we compute the absolute pricing error for each contract

$$\text{abserr}_{t,n} = \left| C_{t,n}^{\text{market}} - C_{t,n}^{\text{model}} \right|. \tag{3.3}$$

Here. $C_{t,n}^{\text{market}}$ is the observed call price for n^{th} contract in the out-of-sample data on day t, and $C_{t,n}^{\text{model}}$ is the call price for n^{th} contract in the out-of-sample data on day t computed by using path-independent tree method based on the multidimensional transform in equation (2.17). At the end, for each model, we compute the average absolute pricing error as follow

$$\text{avgerr} = \frac{\sum_{t} \sum_{n=1}^{N_t} \text{abserr}_{t,n}}{\sum_{t} N_t}, \tag{3.4}$$

where, N_t is the total number of contracts in the out-of-sample data on day t in the sample period. Therefore, the out-of-sample pricing performance for each model is evaluated based on the average absolute pricing error.

To compute $C_{t,n}^{\text{model}}$ for n^{th} contract on day t for each model, we must make adjustment for the one-time discrete dividend payment on June 20 in 2014. We follow the idea in Haug, Haug, and Lewis (2003). For each contract in the out-of-sample data set, the maturity is divided into two parts: one part is from the current day t to June 20, the other part is from June 20 to its expiration date. In the second part of maturity, there is no dividend payment, which implies that American all price will not diverge from European call price. Thus we can use the quasi-analytical solution for European style contract to evaluate the call option in the second part of maturity, and obtain the call price on the ex-dividend date. In the first part of the maturity, one two-dimensional path-independent tree is constructed. When the tree is grow-

ing up to ex-dividend date (June 20), it is stopped. Then the underlying price at each node in the tree on June 20 is reduced the same amount as dividend. And the quasi-analytical solution is applied to evaluate the call option in the second part of maturity for each node in the tree on June 20 with the starting underlying price is the stock price less the dividend mentioned before. We refer the price computed for the second part of maturity as C_{divi}. Then the payoff for each node at the end of tree (on ex-dividend date) is calculated as $\max(C_{\text{divi}}, \text{stock price} - \text{strike price})$. After we obtain the payoffs at the end of tree, we can compute $C_{t,n}^{\text{model}}$ as usual. The empirical results are present in the next subsection.

3.3 Empirical Results

Our sample period is from March 24 to April 25 in 2014, five weeks in total. After some exclusion filters described in Section 3.1, we have 876 contracts in total in our in-sample data set, and 466 contracts in total in the out-of-sample data set. Table 8 describes some properties of our in- and out-of-sample data sets. We report the total number of contracts, maturities, and strikes for each date in the sample period for in- and out-of-sample, respectively. We also report the average SPY call price of each date for the two sample data sets. The total number of contracts for each date in the in-sample date set is between 16 and 52, which is larger than the number of parameters to be estimated for each model (the CGMYsv has the most number of parameters (13) to be estimated). Also, for in-sample data, the total number of maturities (3 to 7 distinct maturity) and strikes (11 to 24 distinct strike price) on each given day indicate that our in-sample data has a good cross sectional spread among the volatility surface.

On a given day, we use all SPY call options in the in-sample data as inputs to calibrate that day's model parameters for each model, separately. Table 9 reports the daily average of calibrated parameters for each of the seven models: BS, Heston, Pan,

VG, VGsv, CGMY, and CGMYsv models. The spot volatility $\sqrt{V_0}$ is quite different across model types: pure jump Levy models (VG and CGMY) have least average calibrated spot volatility, 1.32% and 1.95% respectively. The spot volatility increases when stochastic volatility and jumps are incorporated into the models. They are 12.03%, 12.11%, and 12.67% under VGsv, Pan, and CGMYsv model, respectively. When jumps are not allowed but volatility is stochastic, the spot volatility is increase a little bit to 13.08% under Heston model. And it is the most under the BS model, which is 35.86%.

The parameters which governing the volatility process are α, m, ρ and σ. The speed-of-volatility adjustment α is the most for Heston model ($\alpha = 17.55$), and it reduces a lot when jumps are allowed, 9.88, 9.74, and 1.56 under Pan, VGsv, and CGMYsv, respectively. The long-term mean volatility m is 2.01%, 1.43%, 1.4%, and 1.82% for Heston, Pan, VGsv, and CGMYsv, respectively. The correlation between stock price process and volatility process ρ controls the negative skewness and the variance coefficient of volatility process σ controls the excess kurtosis. It is shown in Table 9 that, when jumps are allowed to occur in stock price, these two parameters are smaller in terms of magnitude.

When doing in-sample fitting to calibrate parameter for each model on each day, we minimized the sum of squared dollar pricing error (SSE), see equation (3.2). The in-sample fitting error is reported in Table 10. Panel A lists the minimized SSE for each model on each day, and Panel B reports the statistics of SSE for each model. Looking at the daily average SSE (mean) in Panel B, one can conclude that the BS model has the highest in-sample fitting error, which is 1.9247; Adding SV in the mode improves in-sample fitting, which is 0.1528 under the Heston model; Allowing jumps in stock price process (SVJ mdoels) can improve the in-sample fitting further. The in-sample fitting error is 0.0852, 0.0897, and 0.0905 for Pan, VGsv, and CGMYsv, respectively. But, for pure jump models (VG and CGMY), the in-sample fitting is

much worse compared to SVJ model, though improved a lot compared to the BS model. Note that the lowest SSE is 0.0012, which is occurred on April 21 under the Pan model; and the highest SSE is 3.6202, which is occurred on April 11 under the BS model.

As it is said in Bakshi et al. (1997) that every candidate model is misspecified, then one straightforward question is which models will result least options pricing error. After calibrating parameters for each model on each day, now we turn to examine each model's out-of-sample pricing performance for American call options. The procedure to perform out-of-sample analysis is described in Section 3.2.3. Table 11 reports the out-of-sample pricing results. The out-of-sample pricing error rank CGMYsv first, then follows VGsv and Pan models. These are SVJ models and the results confirm that incorporating stochastic volatility and jump into the model improves options pricing performance, which agrees arguments in the literature (see Bakshi et al. (1997) and others). The next better model is the Heston model, a stochastic volatility model, which reduces the absolute pricing error by 5.66% compared to the BS model. It is surprising that the pure jump models (VG and CGMY) are worse than the BS model in terms of options pricing. It is because that VG and CGMY give much larger pricing error for OTM options though the pricing error for ITM and ATM is much smaller, compared to the BS model. But VG and CGMY have better in-sample fitting, which implies that VG and CGMY are over-fitting and penalized because the extra parameters cannot improve the models. The standard deviation is relatively large for each model. That is because every model prices ITM and ATM options much better than pricing OTM options. So we have very small pricing error for ITM and ATM options, but very large pricing error for OTM options. This is also can be seen from the range of min and max of each model.

CONCLUSION

We present a transform-based approach for building simple path-independent lattices for pricing American options under low-dimensional stochastic volatility jump models. Our approach generalizes the work of Nelson and Ramaswamy (1990) to a multi-dimensional setting allowing the construction of efficient and accurate path-independent trees for virtually all low-dimensional stochastic volatility models with or without jump given in the literature, including one volatility factor-based stochastic models (SV), two volatility factors-based stochastic volatility models (SV2), stochastic volatility jump models (SVJ), Levy models such as the Variance Gamma model and the CGMY model, and Levy models extended with stochastic volatility proposed in this paper (the VGsv and CGMYsv models). The lattice based approximations of the prices of European options converge rapidly to their true prices obtained with quasi-analytical solutions using our approach.

The path-independent transform-based approach presented in this paper is more efficient and accurate than all other methods proposed in the literature, including the LSM approach of Longstaff and Schwartz (2001), the path-dependent GARCH method of Ritchken and Trevor (1999), the incomplete Fourier transform method of Chiarella and Ziogas (2005), the multigrid implicit finite difference scheme of Clarke and Parrott (1999), the finite element method of Ballestra and Sgarra (2010), the RBF approach of Ballestra and Pacelli (2013), and the operator splitting method of Salmi et al. (2014), and others for pricing American options under low-dimensional stochastic volatility jump models.

We also investigate empirical performance of Levy models with stochastic volatility and jumps for American options. By conducting in-sample fitting and out-of-sample

pricing performance analysis, using SPY call options data, our results show that, for in-sample analysis, the rank (from best to worst) is Pan, CGMYsv, VGsv, Heston, CGMY, VG and BS; and for out-of-sample pricing performance, the rank (from best to worst) is CGMYsv, VGsv, Pan, Heston, BS, VG, and CGMY. Adding stochastic volatility and jump into a model improves American options pricing performance, but pure jump models are worse than the BS model in American options pricing. Our empirical results show that pure jump model are over-fitting, but not improve American options pricing when they are applied to out-of-sample data.

APPENDIX A

THE FIVE STOCHASTIC VOLATILITY MODELS

For the five models listed in Table 1, we present the dynamics equations (SDEs) for the volatility $\mathbf{V}_t$, the asset price S_t, the transformed process X_t, as well as the log-price $Y_t = \ln(S_t)$. Since the first 4 models are one-factor volatility models, to simplify notation we let $V_t = V_{1,t}, dZ_t = dZ_{1,t}, dW_t = dW_{1,t}$ and $dZ_t dW_t = \rho_1 dt = \rho dt$ for these models.

Hull-White model:

$$\begin{aligned}
dV_t &= \alpha(m - V_t)dt + \sigma V_t dW_t, \\
\frac{dS_t}{S_t} &= rdt + \sqrt{V_t}dZ_t, \\
dY_t &= \left(r - \frac{V_t}{2}\right)dt + V_t dZ_t, \\
dX_t &= \left(r - \frac{\alpha\rho(m - V_t)}{\sigma\sqrt{V_t}} - \frac{V_t}{2} + \frac{\rho\sigma\sqrt{V_t}}{4}\right)dt + \sqrt{V_t(1-\rho^2)}dZ_t^*;
\end{aligned}$$

Chesney-Scott model:

$$\begin{aligned}
dV_t &= \alpha(m - V_t)dt + \sigma dW_t, \\
\frac{dS_t}{S_t} &= rdt + e^{V_t}dZ_t, \\
dY_t &= \left(r - \frac{e^{2V_t}}{2}\right)dt + e^{V_t}dZ_t, \\
dX_t &= \left(r - \frac{\alpha\rho e^{V_t}(m - V_t)}{\sigma} - \frac{e^{2V_t} + \rho\sigma e^{V_t}}{2}\right)dt + e^{V_t}\sqrt{1-\rho^2}\, dZ_t^*;
\end{aligned}$$

Stein-Stein model:

$$\begin{aligned}
dV_t &= \alpha(m - V_t)dt + \sigma dW_t, \\
\frac{dS_t}{S_t} &= rdt + V_t dZ_t, \\
dY_t &= \left(r - \frac{V_t^2}{2}\right)dt + V_t dZ_t, \\
dX_t &= \left(r - \frac{\rho\sigma}{2} - \frac{\alpha\rho V_t\,(m - V_t)}{\sigma} - \frac{V_t^2}{2}\right)dt + V_t\sqrt{1-\rho^2}\; dZ_t^*;
\end{aligned}$$

Heston model :

$$\begin{aligned}
dV_t &= \alpha(m - V_t)dt + \sigma\sqrt{V_t}dW_t, \\
\frac{dS_t}{S_t} &= rdt + \sqrt{V_t}dZ_t, \\
dY_t &= \left(r - \frac{V_t}{2}\right)dt + \sqrt{V_t}dZ_t, \\
dX_t &= \left(r - \frac{\rho\alpha(m - V_t)}{\sigma} - \frac{\sigma(\rho+2)}{4}\right)dt + \sqrt{V_t\,(1-\rho^2)}\; dZ_t^*;
\end{aligned}$$

SV2 models:

$$\begin{aligned}
dV_{i,t} &= \alpha_i(m_i - V_{i,t})dt + \sigma_i\sqrt{V_{i,t}}dW_{i,t}, \\
\frac{dS_t}{S_t} &= rdt + \sqrt{V_{1,t}}dZ_{1,t} + \sqrt{V_{2,t}}dZ_{2,t}, \\
dY_t &= \left(r - \left(\frac{V_{1,t}}{2} + \frac{V_{2,t}}{2}\right)\right)dt + \sqrt{V_{1,t}}dZ_{1,t} + \sqrt{V_{2,t}}dZ_{2,t}, \\
dX_t &= \left(r - \sum_{i=1}^{2}\frac{\rho_i\alpha_i(m_i - V_{i,t})}{\sigma_i} - \frac{\sigma_i(\rho_i+2)}{4}\right)\; dt \\
&\quad + \sqrt{V_{1,t}\,(1-\rho_1^2)}\; dZ_{1,t}^* + \sqrt{V_{2,t}\,(1-\rho_2^2)}\; dZ_{2,t}^*.
\end{aligned}$$

APPENDIX B

CHARACTERISTIC FUNCTION FOR LEVY MODELS

Following Heston (1993), (Bakshi et al., 1997, Appendix), and Cont and Tankov (2003) (Secion 15.4), we compute the characteristic function (2.27) for Levy models. Using Ito's lemma and the law of iterated expectations on the characteristic function (2.27), also combining equations (1.2) and (2.26), we get the differential of the characteristic function as follows

$$\begin{aligned} df = &-\frac{\partial f}{\partial \tau} + \frac{\partial f}{\partial x_t}\left[r - \sum_{i=1}^{2}\left(\frac{1}{2}V_{i,t} - (a_i + b_i V_{i,t})\left(\delta_i(\varepsilon_i) + \frac{1}{2}u_i^2(\varepsilon_i)\right)\right)\right] \\ &+ \sum_{i=1}^{2}\frac{\partial f}{\partial V_{i,t}}\left(\alpha_i(m_i - V_{i,t})\right) + \frac{1}{2}\frac{\partial^2 f}{\partial x_t^2}\sum_{i=1}^{2}\left[V_{i,t} + (a_i + b_i V_{i,t})u_i^2(\varepsilon_i)\right] \\ &+ \frac{1}{2}\sum_{i=1}^{2}\frac{\partial^2 f}{\partial V_{i,t}^2}\sigma_i^2 V_{i,t} + \sum_{i=1}^{2}\frac{\partial^2 f}{\partial x_t \partial V_{i,t}}\rho_i \sigma_i V_{i,t} + \sum_{i=1}^{2}(a_i + b_i V_{i,t})\omega_i(\varepsilon_i, \varsigma)f, \qquad \text{(B.1)} \\ = &\ 0. \qquad \text{(B.2)} \end{aligned}$$

Here,

$$\omega_i(\varepsilon_i, \varsigma) = \int_{|y| \geq \varepsilon_i} \left(e^{\imath \varsigma y} - 1\right) \nu_i(dy). \qquad \text{(B.3)}$$

Taking partial derivatives in the characteristic function (2.27), we get

$$
\begin{aligned}
\frac{\partial f}{\partial x_t} &= \imath\varsigma f, & \frac{\partial^2 f}{\partial V_{i,t}^2} &= D_i^2(\tau,\varsigma) f, \\
\frac{\partial^2 f}{\partial x_t^2} &= -\varsigma^2 f, & \frac{\partial^2 f}{\partial x_t \partial V_{i,t}} &= \imath\varsigma D_i(\tau,\varsigma) f, \\
\frac{\partial f}{\partial V_{i,t}} &= D_i(\tau,\varsigma) f, & \frac{\partial f}{\partial \tau} &= \left(\frac{\partial C(\tau,\varsigma)}{\partial \tau} + \sum_{i=1}^{2} \frac{\partial D_i(\tau,\varsigma)}{\partial \tau} V_{i,t} \right) f.
\end{aligned}
\tag{B.4}
$$

Substituting the partial derivatives in equation (B.4) into equation (B.1), we have

$$
\frac{\partial C(\tau,\varsigma)}{\partial \tau} = H + \sum_{i=1}^{2} \alpha_i m_i D_i(\tau,\varsigma), \tag{B.5}
$$

$$
\frac{\partial D_i(\tau,\varsigma)}{\partial \tau} = P_i + Q_i D_i(\tau,\varsigma) + R_i D_i^2(\tau,\varsigma). \tag{B.6}
$$

Here,

$$
H = r\imath\varsigma - a \sum_{i=1}^{2} \left[\frac{1}{2} \imath\varsigma u_i^2(\varepsilon_i) - \frac{1}{2} \varsigma^2 u_i^2(\varepsilon_i) - \imath\varsigma \delta_i(\varepsilon_i) + \omega_i(\varepsilon_i,\varsigma) \right], \tag{B.7}
$$

$$
P_i = -\imath\varsigma \left[\frac{1}{2} \left(1 + b_i u_i^2(\varepsilon_i) \right) + b_i \delta_i(\varepsilon_i) \right] - \frac{1}{2} \varsigma^2 (1 + b_i u_i^2(\varepsilon_i)) + b_i \omega_i(\varepsilon_i,\varsigma), \tag{B.8}
$$

$$
Q_i = \imath\varsigma \rho_i \sigma_i - \alpha_i, \tag{B.9}
$$

$$
R_i = \frac{1}{2} \sigma_i^2. \tag{B.10}
$$

Since $D_i(0,\varsigma) = 0$, it can be shown that

$$
D_i(\tau,\varsigma) = -\frac{2P_i}{Q_i - L_i \coth(\frac{\tau L_i}{2})}, \tag{B.11}
$$

where, $L_i = \sqrt{Q_i^2 - 4P_i R_i}$. Similarly, given $C(0,\varsigma) = 0$, we can get

$$
C(\tau,\varsigma) = \tau \left(H - \sum_{i=1}^{2} \frac{\alpha_i m_i Q_i}{2R_i} \right) - \sum_{i=1}^{2} \frac{\alpha_i m_i}{R_i} \ln \left[\cosh(\frac{\tau L_i}{2}) - \frac{Q_i}{L_i} \sinh(\frac{\tau L_i}{2}) \right]. \tag{B.12}
$$

To compute $C(\tau, \varsigma)$ and $D_i(\tau, \varsigma)$, we need to numerically evaluate $\delta_i(\varepsilon_i)$, $u_i^2(\varepsilon_i)$, and $\omega_i(\varepsilon_i, \varsigma)$. For computation efficiency, we eliminate the need for numerically computing $\delta_i(\varepsilon_i)$ by noting that from equation (2.4) and (B.3)

$$\delta_i(\varepsilon_i) = \omega_i(\varepsilon_i, \varsigma = -\imath). \tag{B.13}$$

Furthermore, $\omega_i(\varepsilon_i, \varsigma)$ can be expressed as follows

$$\omega_i(\varepsilon_i, \varsigma) = \omega_i(0, \varsigma) - \bar{\omega}_i(\varepsilon_i, \varsigma), \tag{B.14}$$

where,

$$\bar{\omega}_i(\varepsilon_i, \varsigma) = \int_{-\varepsilon_i}^{\varepsilon_i} (e^{\imath \varsigma y} - 1)\nu_i(dy), \tag{B.15}$$

and $\omega_i(0, \varsigma)$ is the unit time log characteristic function associated with Levy density ν_i, which, in most cases, has analytical (closed-form) solution. For example, Carr et al. (2003a) provides closed-form solutions of the log characteristic function for various Levy densities such as the VG model and the CGMY model, etc.

Hence, the values of $\bar{\omega}_i(\varepsilon_i, \varsigma)$ and $u_i^2(\varepsilon_i)$ are the only numerical approximations required to solve the characteristic function (2.27).

APPENDIX C

THE FIVE LEVY MODELS

The five Levy models considered in Section 2.6 are Pan(2002), the VG model, the VGsv model, the CGMY model, and the CGMYsv model. Below we present the Levy density and the unit time log characteristic function for each of the five Levy models.

- Pan model: the Levy intensity in Pan model (Pan, 2002, Section 2.1) is

$$\nu(dy) = bf(y),$$

where, $f(y)$ is the normal density of jumps with mean μ_J and volatility σ_J. The other specification are same as in the general Levy specification (2.25) expect that $\boldsymbol{\kappa}(V_t) = bV_t$. Since there is no singular point near zero in the integration of $\lambda(\varepsilon)$ (see equation (2.5)), $\varepsilon = 0$ is allowed under the Pan model. The unit time log characteristic function in equation (B.14) is

$$\omega(0,\varsigma) = b\exp\left(\imath\varsigma\mu_J - \frac{1}{2}\sigma_J^2\varsigma^2\right) - b.$$

- VG model: the Levy density for the VG model (Madan and Seneta (1990) and (Carr et al., 2003a, Section 2.2)) is

$$\nu(dy) = \begin{cases} \frac{C\exp(Gy)}{-y}, & y < 0, \\ \frac{C\exp(-My)}{y}, & y > 0. \end{cases}$$

And the unit time log characteristic function in equation (B.14) is

$$\omega(0,\varsigma) = C\log\left(\frac{GM}{GM+(M-G)\imath\varsigma+\varsigma^2}\right).$$

Since the volatility in the VG model is assumed to be constant, in the general Levy specification (2.25), we set $\alpha = 0, \sigma = 0, a = 1$, and $b = 0$.

- VGsv model: the VG model with stochastic volatility. The VG model is extended by allowing volatility to be stochastic, following one-factor version of equation (1.2). Therefore, the VGsv model will have the same Levy intensity and $\omega(0,\varsigma)$ as the VG model has. Also, the VGsv model will have the general Levy specification (2.25), which means that α, σ, a, and b are not specified to be zeros.

- CGMY model: The Levy density for the CGMY model (Carr et al., 2003a, Section 2.3) is defined as

$$\nu(dy) = \begin{cases} \frac{C\exp(Gy)}{(-y)^{1+Y}}, & y < 0, \\ \frac{C\exp(-My)}{y^{1+Y}}, & y > 0. \end{cases}$$

Here, C, G, M are positive real numbers, and $Y < 2$. In Carr et al. (2003a), C and Y are allowed to take different values for $y > 0$ and $y < 0$. Letting C_p and Y_p be the values for positive y, and letting C_n and Y_n be the values for negative y, the Levy density could be rewritten as

$$\nu(dy) = \begin{cases} \frac{C_n\exp(Gy)}{(-y)^{1+Y_n}}, & y < 0, \\ \frac{C_p\exp(-My)}{y^{1+Y_p}}, & y > 0. \end{cases}$$

And the unit time log characteristic function in equation (B.14) is

$$\omega(0,\varsigma) = C_p \left[\Gamma(-Y_p)\left((M-\imath\varsigma)^{Y_p} - M^{Y_p}\right) + \zeta\Gamma(-Y_n)\left((G+\imath\varsigma)^{Y_n} - G^{Y_n}\right)\right],$$

where, $\zeta = \frac{C_n}{C_p}$. Since the volatility in the CGMY model is constant, in the general Levy specification (2.25), we set $\alpha = 0, \sigma = 0, a = 1$, and $b = 0$.

- CGMYsv model: the CGMY model with stochastic volatility. The CGMY model is extended by allowing volatility to be stochastic, following one-factor version of equation (1.2). Therefore, the CGMYsv model will have the same Levy intensity and $\omega(0,\varsigma)$ as the CGMY model has. Also, the CGMYsv will have the general Levy specification (2.25), which means that α, σ, a, and b are not specified to be zeros.

Table 1. Five Stochastic Volatility Models

The first three rows specify the functions $\phi_i(v), \varphi_i(v)$, and $\psi_i(v)$ for the five models. Row 4 give the $X-$transform. Row 5 and 6 are drift and diffusion parts for X_t-process in equation (1.6). The last row is the inverse transform.

	Hull-White (1987)	Chesney-Scott (1989)	Stein-Stein (1991)	Heston (1993)	SV2 $i = 1, 2$
$\phi(v)$	$\alpha(m - v(t))$	$\alpha(m - v(t))$	$\alpha(m - v(t))$	$\alpha(m - v(t))$	$\phi_i(v_i) = \alpha_i(m_i - v_i(t))$
$\varphi(v)$	$\sigma v(t)$	σ	σ	$\sigma\sqrt{v(t)}$	$\varphi_i(v_i) = \sigma_i\sqrt{v_i(t)}$
$\psi(v)$	$\sqrt{v(t)}$	$\exp(v(t))$	$v(t)$	$\sqrt{v(t)}$	$\psi_i(v_i) = \sqrt{v_i(t)}$
X_t	$\ln(S_t) - \frac{2\rho}{\sigma}\sqrt{v(t)}$	$\ln(S_t) - \frac{\rho}{\sigma}\exp(v(t))$	$\ln(S_t) - \frac{\rho}{2\sigma}v^2(t)$	$\ln(S_t) - \frac{\rho}{\sigma}v(t)$	$\ln(S_t) - \sum_{i=1}^{2}\frac{\rho_i}{\sigma_i}v_i(t)$
$\mu_X(t)$	$r - \frac{\rho\alpha(m - v(t))}{\sigma\sqrt{v(t)}}$ $-\frac{1}{2}v(t) + \frac{\rho\sigma}{4}\sqrt{v(t)}$	$r - \frac{\rho}{\sigma}\exp(v(t))\alpha(m - v(t))$ $-\frac{1}{2}\left(\exp(2v(t)) + \rho\sigma\exp(v(t))\right)$	$r - \frac{1}{2}\rho\sigma - \frac{1}{2}v^2(t)$ $-\frac{\rho}{\sigma}\alpha(v - v(t))v(t)$	$r - \frac{\rho}{\sigma}\alpha m$ $+v(t)\left(\frac{\rho\alpha}{\sigma} - \frac{1}{2}\right)$	$r - \sum_{i=1}^{2}\frac{\rho_i\alpha_i m_i}{\sigma_i}$ $+\sum_{i=1}^{2}v_i(t)\left(\frac{\rho_i\alpha_i}{\sigma_i} - \frac{1}{2}\right)$
$\sigma_X(t)$	$\sqrt{v(t)(1-\rho^2)}$	$\sqrt{\exp(2v(t))(1-\rho^2)}$	$\sqrt{v^2(t)(1-\rho^2)}$	$\sqrt{v(t)(1-\rho^2)}$	$\sqrt{\sum_{i=1}^{2}v_i(t)(1-\rho_i^2)}$
S_t	$\exp(X_t + \frac{2\rho}{\sigma}\sqrt{v(t)})$	$\exp(X_t + \frac{\rho}{\sigma}e^{v(t)}))$	$\exp(X_t + \frac{\rho}{2\sigma}v^2(t))$	$\exp(X_t + \frac{\rho}{\sigma}v(t))$	$\exp(X_t + \sum_{i=1}^{2}\frac{\rho_i}{\sigma_i}v_i(t))$

Table 2. Heston model, European and American put prices.

Other parameters are set as: underling price = \$100, $r = 0.03$, $\sigma = 0.38$, $\alpha = 5.3$, $m = 0.0242$, and $\rho = -0.57$. Column 6 is the Heston put prices, computed using closed-form solution for European options. In the tree method, N is the number of time steps.

$\sqrt{V(0)}$	Maturity	Strike	European Tree		Heston	American Tree	
			$N = 100$	$N = 300$	(closed)	$N = 100$	$N = 300$
0.1	0.0833	90	0.0083	0.0086	0.0086	0.0087	0.0087
		100	1.1518	1.1523	1.1524	1.1669	1.1671
		110	9.1849	9.1975	9.1994	9.4681	9.4357
	0.25	90	0.2201	0.2222	0.2232	0.2254	0.2254
		100	2.1202	2.1263	2.1267	2.1746	2.1745
		110	9.2840	9.2953	9.2954	9.6351	9.5874
	0.5	90	0.7022	0.7149	0.7141	0.7319	0.7308
		100	3.0987	3.1125	3.1122	3.2268	3.2274
		110	9.7257	9.7257	9.7257	10.1524	10.0822
0.2	0.0833	90	0.1031	0.1034	0.1036	0.1039	0.1040
		100	2.0698	2.0700	2.0700	2.2865	2.2861
		110	9.7605	9.7588	9.7577	9.9994	9.9998
	0.25	90	0.6196	0.6229	0.6232	0.6298	0.6300
		100	3.1827	3.1843	3.185	3.2503	3.2511
		110	9.8524	9.8732	9.873	10.2471	10.2485
	0.5	90	1.2076	1.2145	1.2179	1.2481	1.2489
		100	4.0260	4.0318	4.0344	4.1913	4.1933
		110	9.9098	9.9116	9.911	10.5276	10.529
0.3	0.0833	90	0.3945	0.3963	0.3972	0.3984	0.3984
		100	3.0567	3.0582	3.0589	3.0756	3.0758
		110	10.1382	10.1369	10.1362	10.3401	10.3409
	0.25	90	1.3036	1.3123	1.3168	1.3306	1.3310
		100	4.4596	4.4651	4.4678	4.5414	4.5427
		110	10.6921	10.6893	10.6877	10.9511	10.9533
	0.5	90	1.9832	2.0132	2.0117	2.0655	2.0670
		100	5.2543	5.2753	5.2746	5.4699	5.4726
		110	11.0172	11.0174	11.0169	11.5545	11.5582

Table 3. Hull-White model, European put prices

Other parameters are adopted from Leisen (2000): $r = 0.0$, $\sigma = 1.0$, $\alpha = 0.0$, $m = 0.0$, and $V(0) = 0.1$. Column 5 is the put prices computed by using Leisen (2000) method. In the tree method, N is the number of time steps.

ρ	Strike	European Tree		Leisen method	Monte Carlo
		$N = 40$	$N = 100$		
0.5	95	0.7904	0.7931	0.7930	0.7880
	96	1.0632	1.0682	1.0680	1.0610
	97	1.3948	1.3930	1.4030	1.3930
	98	1.7877	1.7974	1.8000	1.7870
	99	2.2439	2.2544	2.2590	2.2430
	100	2.7613	2.7614	2.7750	2.7610
	101	3.3382	3.3589	3.3590	3.3370
	102	3.9698	3.9909	3.9930	3.9690
	103	4.6528	4.6635	4.6780	4.6520
	104	5.3813	5.4029	5.4080	5.3810
	105	6.1512	6.1723	6.1790	6.1510
0	95	0.9252	0.9227	0.9220	0.9220
	96	1.1835	1.1808	1.1810	1.1800
	97	1.4936	1.4912	1.4910	1.4900
	98	1.8575	1.8558	1.8580	1.8550
	99	2.2802	2.2815	2.2830	2.2790
	100	2.7595	2.7616	2.7620	2.7630
	101	3.3078	3.3089	3.3100	3.3070
	102	3.9112	3.9093	3.9100	3.9080
	103	4.5699	4.5656	4.5660	4.5650
	104	5.2778	5.2749	5.2730	5.2730
	105	6.0309	6.0277	6.0270	6.0280
-0.5	95	1.0280	1.0289	1.0220	1.0290
	96	1.2714	1.2726	1.2650	1.2720
	97	1.5607	1.5512	1.5530	1.5620
	98	1.8995	1.9008	1.8910	1.9010
	99	2.2930	2.2941	2.2820	2.2940
	100	2.7434	2.7339	2.7250	2.7450
	101	3.2537	3.2342	3.2370	3.2550
	102	3.8235	3.8035	3.8040	3.8250
	103	4.4528	4.4304	4.4290	4.4540
	104	5.1386	5.1185	5.1120	5.1400
	105	5.8779	5.8513	5.8490	5.8790

Table 4. SV2 model, European and American put prices.(continued on next page)

Other parameters are adopted from Christoffersen et al. (2009): $r = 0.03$, $\sigma_1 = 1.072$, $\alpha_1 = 0.2563$, $m_1 = 0.004$, $\rho_1 = -0.8084$, $\sigma_2 = 0.2875$, $\alpha_2 = 1.8817$, $m_2 = 0.0233$, and $\rho_2 = -0.6997$. In the tree method, N is the number of time steps.

$\sqrt{V_1(0)}$	$\sqrt{V_2(0)}$	Maturity	Strike	European Tree		Closed	American Tree	
				$N = 50$	$N = 150$	Form	$N = 50$	$N = 150$
0.1	0.1	0.0833	90	0.0858	0.0879	0.0875	0.0877	0.0912
			100	1.4246	1.4249	1.4248	1.4389	1.4392
			110	8.9607	8.9698	8.9703	9.1051	9.1943
		0.25	90	0.4710	0.4764	0.4788	0.4832	0.4826
			100	2.2875	2.2897	2.2913	2.3396	2.3401
			110	9.2724	9.2753	9.2767	9.5344	9.5315
		0.5	90	0.9750	0.9913	1.0000	1.0210	1.0191
			100	3.0953	3.1020	3.1112	3.2214	3.2211
			110	9.7267	9.7266	9.7266	10.0096	10.0105
	0.3	0.0833	90	0.5609	0.5649	0.5669	0.5682	0.5683
			100	3.3828	3.3874	3.3898	3.4043	3.4047
			110	10.2955	10.2963	10.2965	10.3803	10.3812
		0.25	90	1.9292	1.9505	1.9608	1.9733	1.9742
			100	5.3075	5.3396	5.3405	5.3953	5.3978
			110	11.3447	11.3528	11.3564	11.5391	11.5443
		0.5	90	3.1210	3.1945	3.1989	3.2426	3.2472
			100	6.6380	6.6704	6.7072	6.8443	6.8534
			110	12.1942	12.2256	12.2273	12.5890	12.6036

$\sqrt{V_1(0)}$	$\sqrt{V_2(0)}$	Maturity	Strike	European Tree		Closed Form	American Tree	
				$N = 50$	$N = 150$		$N = 50$	$N = 150$
0.3	0.1	0.0833	90	0.7953	0.7973	0.7983	0.7997	0.8000
			100	3.3766	3.3768	3.3769	3.3896	3.3899
			110	10.0468	10.0432	10.0429	10.1526	10.1539
		0.25	90	2.4661	2.4705	2.4737	2.4851	2.4855
			100	5.2339	5.2346	5.2355	5.2820	5.2831
			110	10.6953	10.6887	10.6849	10.9113	10.9129
		0.5	90	3.7903	3.8095	3.8053	3.8430	3.8441
			100	6.5255	6.5328	6.5305	6.6435	6.6463
			110	11.3837	11.3744	11.3710	11.7701	11.7711
	0.3	0.0833	90	1.3394	1.3442	1.3466	1.3489	1.3490
			100	4.5789	4.5832	4.5852	4.5985	4.5990
			110	11.0074	11.0071	11.0070	11.0617	11.0628
		0.25	90	3.5924	3.6098	3.6192	3.6357	3.6369
			100	7.2201	7.2365	7.2451	7.2950	7.2978
			110	12.8930	12.8995	12.9029	13.0357	13.0410
		0.5	90	5.3733	5.4328	5.4331	5.4854	5.4905
			100	9.0509	9.1086	9.1067	9.2333	9.2424
			110	14.3213	14.3529	14.3526	14.6334	14.6477

Table 5. Pan model, European and American put prices.

Other parameters are set as: underline price \$100, $r = 0.03$, $\sigma = 0.2087$, $\alpha = 10$, $m = 0.0032$, $\rho = -0.99$, $b = 21.51$, $\mu_J = -0.0356$, and $\sigma_J = 0.0286$. Column 5 is the Pan model put prices, computed using quasi-analytical (closed-form) solution for European options. In the tree method, the maturity is partitioned into 181 time steps. At each time step, the nodes in volatility process tree is ≤ 51, and the nodes in transformed stock price process tree is ≤ 201.

$\sqrt{V(0)}$	Maturity	Strike	European Tree	Pan model (closed-form)	American Tree
0.1	0.0833	90	0.0670	0.0669	0.0838
		100	1.2383	1.2383	1.2686
		110	9.7254	9.7254	10.0000
	0.25	90	0.2534	0.2535	0.2874
		100	1.8237	1.8238	1.8418
		110	9.1790	9.1790	10.0000
	0.5	90	0.6261	0.6215	0.6513
		100	2.9680	2.9675	2.9818
		110	8.3877	8.3874	10.0000
0.2	0.0833	90	0.2936	0.2930	0.2956
		100	2.5366	2.5367	2.5424
		110	9.8062	9.8056	10.0000
	0.25	90	0.6644	0.6648	0.6698
		100	3.1958	3.1956	3.2696
		110	9.6058	9.6060	10.0893
	0.5	90	0.8061	0.8067	0.9530
		100	3.2936	3.2946	3.6605
		110	9.1938	9.1937	10.3611
0.3	0.0833	90	0.4640	0.4589	0.4672
		100	3.8419	3.8409	3.8708
		110	10.4754	10.4754	10.4954
	0.25	90	1.5558	1.5562	1.8935
		100	4.7864	4.7862	4.8307
		110	10.8557	10.8560	10.8647
	0.5	90	1.7049	1.7046	1.8755
		100	4.8405	4.8403	4.9901
		110	10.5707	10.5698	11.1387

Table 6. VGsv model, European and American put prices.

Other parameters are set as: underling price \$100, $r = 0.03$, $\sigma = 0.3433$, $\alpha = 10$, $m = 0.0059$, $\rho = -0.99$, $a = 0.1$, $b = 10$, $C = 13.2626$, $G = 12.6496$, and $M = 34.9055$. Column 5 is the VGsv model put prices, computed using quasi-analytical (closed-form) solution for European options. In the tree method, the maturity is partitioned into 181 time steps. At each time step, the nodes in volatility process tree is ≤ 51, and the nodes in transformed stock price process tree is ≤ 201.

$\sqrt{V(0)}$	Maturity	Strike	European Tree	VGsv model (closed-form)	American Tree
0.1	0.0833	90	0.0542	0.0543	0.0831
		100	1.1894	1.1888	1.2893
		110	8.5116	8.5117	10.0000
	0.25	90	0.2333	0.2341	0.2605
		100	1.8460	1.8480	1.8632
		110	9.1781	9.1781	10.0351
	0.5	90	0.6641	0.6659	0.6947
		100	3.1281	3.1238	3.2475
		110	9.7253	9.7255	10.1432
0.2	0.0833	90	0.2158	0.2153	0.2205
		100	2.1593	2.1593	2.1954
		110	9.0420	9.0425	10.0000
	0.25	90	0.5981	0.5971	0.6316
		100	2.8767	2.8768	2.9712
		110	9.3594	9.3592	10.0000
	0.5	90	0.8470	0.8500	0.9207
		100	3.2204	3.2210	3.3141
		110	9.7345	9.7345	10.2813
0.3	0.0833	90	0.5799	0.5800	0.6280
		100	3.2032	3.2042	3.2203
		110	9.9880	9.9887	10.0657
	0.25	90	1.2140	1.2154	1.2905
		100	4.0985	4.0974	4.1775
		110	10.0740	10.0730	10.0971
	0.5	90	1.4738	1.4728	1.8460
		100	4.3304	4.3336	4.8964
		110	10.1797	10.1802	10.3354

Table 7. CGMYsv model, European and American put prices.

Other parameters are set as: underling price \$100, $r = 0.03$, $\sigma = 0.5457$, $\alpha = 10$, $m = 0.0149$, $\rho = -0.99$, $a = 0.1$, $b = 5.3518$, $C_p = 500$, $G = 16.9655$, $M = 612.4486$, $Y_p = -77.1569$, $Y_n = -36.1443$ and $\zeta = 1.2771$. Column 5 is the CGMYsv model put prices, computed using quasi-analytical (closed-form) solution for European options. In the tree method, the maturity is partitioned into 181 time steps. At each time step, the nodes in volatility process tree is ≤ 51, and the nodes in transformed stock price process tree is ≤ 201.

$\sqrt{V(0)}$	Maturity	Strike	European Tree	CGMYsv model (closed-form)	American Tree
0.1	0.0833	90	0.0627	0.0630	0.0664
		100	1.1027	1.1026	1.1893
		110	9.7253	9.7249	10.0000
	0.25	90	0.2664	0.2667	0.2849
		100	1.9194	1.9198	1.9826
		110	9.1783	9.1783	10.0000
	0.5	90	0.6871	0.6875	0.6916
		100	2.9713	2.9712	2.9965
		110	8.5768	8.5768	10.0000
0.2	0.0833	90	0.1972	0.1975	0.2178
		100	1.9403	1.9406	1.9892
		110	9.7253	9.7253	10.0000
	0.25	90	0.6870	0.6870	0.7054
		100	2.7326	2.7320	2.7754
		110	9.2411	9.2414	10.0000
	0.5	90	1.1401	1.1401	1.1923
		100	3.3483	3.3487	3.3677
		110	9.9734	9.9737	10.0000
0.3	0.0833	90	0.4961	0.4960	0.5286
		100	2.8613	2.8612	2.8975
		110	9.8743	9.8742	10.0169
	0.25	90	1.2150	1.2152	1.2606
		100	3.7731	3.7722	3.8503
		110	9.7813	9.7813	10.2358
	0.5	90	1.6627	1.6636	1.7410
		100	4.2455	4.2456	4.2629
		110	9.7036	9.7046	10.5285

Table 8. Sample properties of SPY call options.(continued on next page).

This table reports some properties of our SPY call options sample. The first column is the trade dates in our sample period. For each date, the total number of contracts, maturities, and strikes are reported, respectively. We also report the average SPY call price of each date (last column).

date	in-sample data			
	contract	maturity	strike	avg call
20140324	49	6	19	3.0327
20140325	35	6	18	3.1209
20140326	35	6	17	2.6297
20140327	45	5	20	3.2033
20140328	38	5	17	2.9180
20140331	37	7	17	2.6227
20140401	33	6	15	2.9353
20140402	39	5	19	3.1969
20140403	36	6	17	2.9183
20140404	34	6	16	2.4197
20140407	40	5	17	2.5435
20140408	33	5	18	3.0406
20140409	41	4	19	3.2727
20140410	37	6	16	2.2170
20140411	50	5	20	2.8832
20140414	39	4	21	2.7908
20140415	52	5	24	3.3881
20140416	41	5	20	3.0477
20140417	37	5	18	3.3551
20140421	21	3	17	3.2010
20140422	29	5	16	2.7859
20140423	16	3	11	2.2650
20140424	27	5	15	2.7278
20140425	32	6	15	2.2967

date	out-of-sample data			
	contract	maturity	strike	avg call
20140324	14	2	13	3.8932
20140325	15	3	13	4.7083
20140326	13	3	12	3.5927
20140327	16	3	14	4.0544
20140328	17	3	12	3.8582
20140331	7	1	7	3.7186
20140401	11	3	9	3.6577
20140402	21	3	12	3.0752
20140403	14	2	13	3.9779
20140404	17	3	12	2.8465
20140407	20	2	14	3.8235
20140408	20	3	15	3.5540
20140409	16	3	11	3.4138
20140410	23	3	13	3.0667
20140411	21	3	16	3.6471
20140414	20	3	15	4.0453
20140415	30	3	15	4.2818
20140416	16	3	10	4.1822
20140417	19	2	17	4.1887
20140421	19	3	13	3.1468
20140422	29	3	14	3.8469
20140423	21	4	12	3.7310
20140424	35	4	18	5.2444
20140425	32	4	15	4.0105

Table 9. Calibrated model parameters using in-sample data.

Daily average model parameters calibrated using in-sample data are listed in the table. On each given date, we calibrate model parameters for each model, then calculate the average in the sample period. $\sqrt{V_0}$ is the spot volatility, α is the speed-of-volatility adjustment, m is the long-term mean volatility, ρ is correlation between stock price process and volatility process (control the negative skewness), σ is the variance coefficient of volatility process (control the excess kurtosis), a and b control the Levy intensity for each model, μ_J and σ_J are jump parameters of Pan model, C, G, and M are parameters of VG model, C_p, G, M, Y_p, Y_n and ζ are parameters of CGMY model.

	BS	Heston	Pan	VG	VGsv	CGMY	CGMYsv
$\sqrt{V_0}$	0.358583	0.130786	0.121128	0.013228	0.120258	0.019467	0.126739
α		17.550536	9.877792		9.741594		10.564926
m		0.020991	0.014349		0.014015		0.018232
ρ		-0.941273	-0.890562		-0.889387		-0.919393
σ		0.842803	0.517278		0.514004		0.612909
a					0.103802		3.323084
b			5.925754		8.726458		8.324288
μ_J			-31.458971				
σ_J			2.722529				
C				10.950624	0.532185		
G				24.584957	1.916942	42.364002	19.531235
M				62.099591	429.215395	412.430575	540.855547
C_p						74.256335	149.527553
Y_p						-0.025796	-51.801174
Y_n						-0.113848	-48.107276
ζ						2.216887	4.393478

Table 10. In-sample fitting error.

This table reports in-sample fitting error for each model on each day in Panel A. The statistics of the in-sample fitting error for each model is in Panel B. The in-sample fitting error on day t is defined as to minimize the sum of squared dollar pricing error: $\text{SSE}_t = \sum_{n=1}^{N} \left| C_n^{\text{market}} - C_n^{\text{model}} \right|^2$, see equation (3.2).

Panel A: Daily in-sampel fitting error							
date	BS	Hston	Pan	VG	VGsv	CGMY	CGMYsv
20140324	3.4282	0.1984	0.1235	1.1221	0.1180	0.5646	0.1394
20140325	2.2730	0.1084	0.0354	0.4739	0.0384	0.2809	0.0365
20140326	2.0994	0.1428	0.0958	1.6595	0.1017	0.7048	0.1000
20140327	2.7629	0.2185	0.1240	1.5647	0.1250	0.7168	0.1279
20140328	2.2643	0.1552	0.0790	1.1742	0.0826	0.5226	0.0853
20140331	1.4869	0.1968	0.1274	0.7166	0.1310	0.3306	0.1312
20140401	1.9353	0.1156	0.0388	0.2100	0.0393	0.2085	0.0482
20140402	2.3335	0.1336	0.0407	0.2754	0.0430	0.2146	0.0436
20140403	1.8527	0.1627	0.0821	0.5284	0.0849	0.2783	0.0871
20140404	1.4663	0.1000	0.0343	0.3408	0.0334	0.1573	0.0350
20140407	1.8539	0.1564	0.0897	0.9940	0.0846	0.3747	0.0931
20140408	1.8150	0.1051	0.0517	0.5142	0.0483	0.2072	0.0563
20140409	2.2571	0.1365	0.0465	0.2318	0.0466	0.1262	0.0476
20140410	2.3007	0.2399	0.2530	2.5812	0.2869	1.2616	0.1975
20140411	3.6202	0.2336	0.2588	3.5204	0.3212	1.7455	0.2861
20140414	1.9584	0.1204	0.0254	0.9676	0.0281	0.2942	0.0410
20140415	3.3988	0.4753	0.3437	1.4597	0.3530	0.7572	0.3689
20140416	1.7021	0.1425	0.0472	0.4135	0.0518	0.1846	0.0577
20140417	1.5380	0.1667	0.0477	0.1491	0.0446	0.1046	0.0589
20140421	0.8587	0.0601	0.0012	0.2967	0.0032	0.0635	0.0054
20140422	1.1012	0.0864	0.0062	0.4462	0.0062	0.1130	0.0071
20140423	0.4084	0.0345	0.0036	0.3753	0.0076	0.0727	0.0080
20140424	0.5848	0.0766	0.0210	0.5238	0.0212	0.0589	0.0211
20140425	0.8919	0.1025	0.0685	1.0177	0.0725	0.2401	0.0699
Panel B: Statistics of daily in-sample fitting error							
mean	1.9247	0.1528	0.0852	0.8982	0.0905	0.3993	0.0897
std	0.8418	0.0866	0.0866	0.8100	0.0963	0.4055	0.0876
min	0.4084	0.0345	0.0012	0.1491	0.0032	0.0589	0.0054
max	3.6202	0.4753	0.3437	3.5204	0.3530	1.7455	0.3689

Table 11. Out-of-sample pricing error for American call options.

This table reports the average absolute dollar pricing error for each model in the out-of-sample analysis. On a given date, for each model, we us the parameters calibrated on the same date as inputs to compute the model based price for each American style contract in the out-of-sample data set. Then we compute the average absolute dollar pricing error for each model using equation (3.4).

model	average absolute pricing error			
	mean	std	min	max
BS	0.363549	0.242326	0.002532	1.196025
Heston	0.342985	0.333444	0.001404	2.635035
Pan	0.337033	0.379298	0.000141	2.732665
VG	0.508083	0.372884	0.000087	1.830751
VGsv	0.334102	0.502581	0.001032	1.965002
CGMY	0.852510	0.845815	0.004286	2.837980
CGMYsv	0.319356	0.292688	0.000319	1.353551

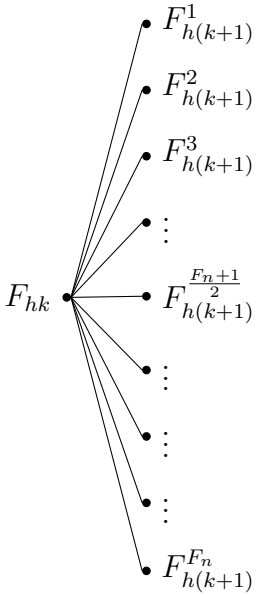

Figure 1. One node at time $t = hk$ in the tree for transformed process F_t, without considering truncation. There are F_n nodes at the next time step $t = h(k+1)$, three of them are diffusion nodes, the others are jump nodes. The nodes are labeled as level $l = 1, 2, 3, \cdots, F_n$ from top to bottom. Here, F_n must be an odd number.

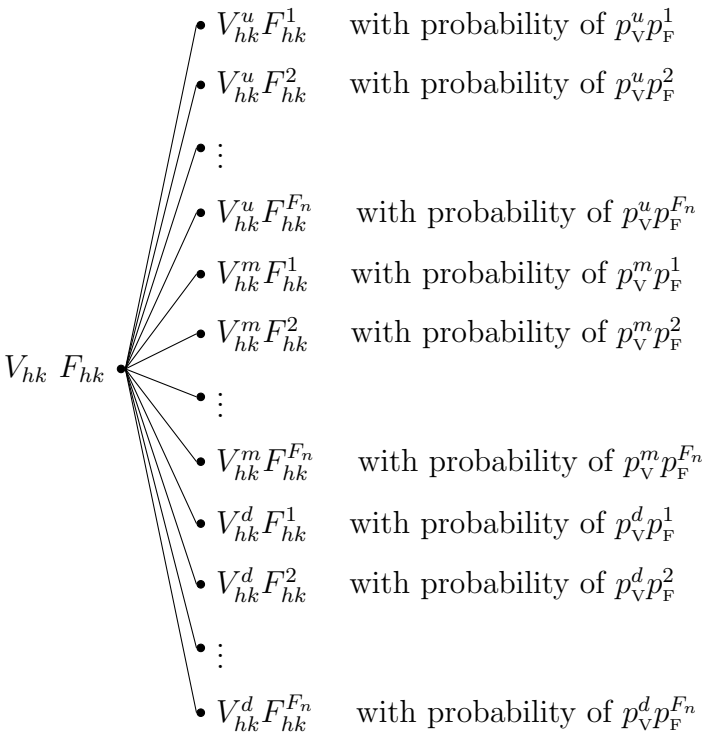

Figure 2. One node at time $t = hk$ in the two-dimensional tree. Here, F^l_{hk} is the one node in F_t-tree at time $t = hk$ labeled as l, and the corresponding probability for that node is $p^l_{\rm F}$, where $l = 1, 2, 3, \cdots, F_n$. V^u_{hk}, V^m_{hk}, and V^d_{hk} are up, middle, and down node emanating out from the node V_{hk} in V_t tree.

BIBLIOGRAPHY

Aıt-Sahalia, Yacine, and Robert Kimmel, 2007, Maximum likelihood estimation of stochastic volatility models, *Journal of Financial Economics* 83, 413–452.

Alizadeh, Sassan, Michael W. Brandt, and Francis X. Diebold, 2002, Range-Based Estimation of Stochastic Volatility Models, *The Journal of Finance* 57, 1047–1091.

Amin, Kaushik I., 1993, Jump Diffusion Option Valuation in Discrete Time, *The Journal of Finance* 48, 1833–1863.

Amin, Kaushik I., and Andrew J. Morton, 1994, Implied volatility functions in arbitrage-free term structure models, *Journal of Financial Economics* 35, 141–180.

Andersen, Torben G., Luca Benzoni, and Jesper Lund, 2002, An Empirical Investigation of Continuous-Time Equity Return Models, *The Journal of Finance* 57, 1239–1284.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical Performance of Alternative Option Pricing Models, *The Journal of Finance* 52, 2003–2049.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 2000, Do Call Prices and the Underlying Stock Always Move in the Same Direction?, *Review of Financial Studies* 13, 549–584.

Ballestra, Luca Vincenzo, and Graziella Pacelli, 2013, Pricing European and American options with two stochastic factors: A highly efficient radial basis function approach, *Journal of Economic Dynamics and Control* 37, 1142–1167.

Ballestra, Luca Vincenzo, and Carlo Sgarra, 2010, The evaluation of American options in a stochastic volatility model with jumps: An efficient finite element approach, *Computers & Mathematics with Applications* 60, 1571–1590.

Barone-Adesi, Giovanni, and Robert E. Whaley, 1987, Efficient Analytic Approximation of American Option Values, *The Journal of Finance* 42, 301–320.

Bates, David S., 1996, Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options, *Review of Financial Studies* 9, 69–107.

Bates, David S., 2000, Post-'87 crash fears in the S&P 500 futures option market, *Journal of Econometrics* 94, 181–238.

Bates, David S., 2003, Empirical option pricing: A retrospection, *Journal of Econometrics* 116, 387–404.

Bates, David S., 2006, Maximum Likelihood Estimation of Latent Affine Processes, *Review of Financial Studies* 19, 909–965.

Black, Fischer, and Myron Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637–654.

Brennan, Michael J., and Eduardo S. Schwartz, 1978, Finite Difference Methods and Jump Processes Arising in the Pricing of Contingent Claims: A Synthesis, *Journal of Financial and Quantitative Analysis* 13, 461–474.

Carr, Peter, and Helyette Geman, 2002, The fine structure of asset returns: An empirical investigation, *The Journal of Business* 75, 305–332.

Carr, Peter, Hélyette Geman, Dilip B. Madan, and Marc Yor, 2003a, Stochastic Volatility for Lévy Processes, *Mathematical Finance* 13, 345–382.

Carr, Peter, Ali Hirsa, and Morgan Stanley, 2003b, Why be Backward ? Forward equations for American options, *Risk* 103–107.

Carr, Peter, Robert Jarrow, and Ravi Myneni, 1992, Alternative Characterizations of American Put Options, *Mathematical Finance* 2, 87–106.

Carriere, Jacques F., 1996, Valuation of the early-exercise price for options using simulations and nonparametric regression, *Insurance: Mathematics and Economics* 19, 19–30.

Chacko, George, and Luis M. Viceira, 2003, Spectral GMM estimation of continuous-time processes, *Journal of Econometrics* 116, 259–292.

Chernov, Mikhail, A. Ronald Gallant, Eric Ghysels, and George Tauchen, 1999, A New Class of Stochastic Volatility Models with Jumps: Theory and Estimation, SSRN Scholarly Paper ID 189628, Social Science Research Network, Rochester, NY.

Chernov, Mikhail, and Eric Ghysels, 1999, Estimation of Stochastic Volatility Models for the Purpose of Option Pricing, SSRN Scholarly Paper ID 158229, Social Science Research Network, Rochester, NY.

Chesney, Marc, and Louis Scott, 1989, Pricing European Currency Options: A Comparison of the Modified Black-Scholes Model and a Random Variance Model, *Journal of Financial and Quantitative Analysis* 24, 267–284.

Chiarella, Carl, Boda Kang, Gunter H. Meyer, and Andrew Ziogas, 2009, The evaluation of american option prices under stochastic volatility and jump-diffusion dynamics using the method of lines, *International Journal of Theoretical and Applied Finance* 12, 393–425.

Chiarella, Carl, and Andrew Ziogas, 2005, Pricing American Options on Jump-Diffusion Processes Using Fourier Hermite Series Expansions, SSRN Scholarly Paper ID 893090, Social Science Research Network, Rochester, NY.

Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2009, The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well, *Management Science* .

Clarke, Nigel, and Kevin Parrott, 1999, Multigrid for American option pricing with stochastic volatility, *Applied Mathematical Finance* 6, 177–195.

Cont, Rama, and Peter Tankov, 2003, *Financial Modelling with Jump Processes*, first edition (Chapman and Hall/CRC, Boca Raton, Fla).

Cont, Rama, and Ekaterina Voltchkova, 2005, A Finite Difference Scheme for Option Pricing in Jump Diffusion and Exponential Lévy Models, *SIAM Journal on Numerical Analysis* 43, 1596–1626.

Cox, John C., and Stephen A. Ross, 1976, The valuation of options for alternative stochastic processes, *Journal of Financial Economics* 3, 145–166.

Cox, John C., Stephen A. Ross, and Mark Rubinstein, 1979, Option pricing: A simplified approach, *Journal of Financial Economics* 7, 229–263.

Duan, Jin-Chuan, 1995, The Garch Option Pricing Model, *Mathematical Finance* 5, 13–32.

Duan, Jin-Chuan, 1996, *A Unified Theory of Option Pricing under Stochastic Volatility from GARCH to Diffusion* (Working Paper).

Duan, Jin-Chuan, 1997, Augmented GARCH (p,q) process and its diffusion limit, *Journal of Econometrics* 79, 97–127.

Duan, Jin-Chuan, Peter Ritchken, and Zhiqiang Sun, 2006, Approximating Garch-Jump Models, Jump-Diffusion Processes, and Option Pricing, *Mathematical Finance* 16, 21–52.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-diffusions, *Econometrica* 68, 1343–1376.

Eraker, Bjørn, 2004, Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, *The Journal of Finance* 59, 1367–1403.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The Impact of Jumps in Volatility and Returns, *The Journal of Finance* 58, 1269–1300.

Finucane, Thomas J., and Michael J. Tomas, 1996, American stochastic volatility call option pricing: A lattice based approach, *Review of Derivatives Research* 1, 183–201.

Gatheral, Jim, 2011, *The Volatility Surface: A Practitioner's Guide* (John Wiley & Sons).

Guan, Lim Kian, and Guo Xiaoqiang, 2000, Pricing American options with stochastic volatility: Evidence from S&P 500 futures options, *Journal of Futures Markets* 20, 625–659.

Gukhal, Chandrasekhar Reddy, 2001, Analytical Valuation of American Options on Jump-Diffusion Processes, *Mathematical Finance* 11, 97–115.

Haug, E. G., J. Haug, and A. Lewis, 2003, Back to Basics: A New Approach to the Discrete Dividend Problem, *Wilmott Magazine* September, 37–47.

Heston, Steven L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 327–343.

Hilliard, Jimmy E, and Adam Schwartz, 1996, Binominal Option Pricing Under Stochastic Volatility and Correlated State Variables, *The Journal of Derivatives* 4, 23–39.

Hull, John, and Alan White, 1987, The Pricing of Options on Assets with Stochastic Volatilities, *The Journal of Finance* 42, 281–300.

Hull, John C, and Alan D White, 1993, Efficient Procedures for Valuing European and American Path-Dependent Options, *The Journal of Derivatives* 1, 21–31.

Jiang, George J., and Roel C. A. Oomen, 2007, Estimating Latent Variables and Jump Diffusion Models Using High-Frequency Data, *Journal of Financial Econometrics* 5, 1–30.

Johnson, Herb, and David Shanno, 1987, Option Pricing when the Variance Is Changing, *Journal of Financial and Quantitative Analysis* 22, 143–151.

Kim, In Joon, 1990, The Analytic Valuation of American Options, *Review of Financial Studies* 3, 547–572.

Kou, S. G., 2002, A Jump-Diffusion Model for Option Pricing, *Management Science* 48, 1086–1101.

Kou, S. G., and Hui Wang, 2004, Option Pricing under a Double Exponential Jump Diffusion Model, *Management Science* 50, 1178–1192.

Leisen, Dietmar P.J, 2000, Stock Evolution Under Stochastic Volatility: A Discrete Approach, *The Journal of Derivatives* 8, 9–27.

Li, Minqiang, 2009, Analytical approximations for the critical stock prices of American options: A performance comparison, *Review of Derivatives Research* 13, 75–99.

Longstaff, Francis A., and Eduardo S. Schwartz, 2001, Valuing American Options by Simulation: A Simple Least-Squares Approach, *Review of Financial Studies* 14, 113–147.

Madan, Dilip B., and Eugene Seneta, 1990, The variance gamma (v.g.) model for share market returns, *The Journal of Business* 63, 511–524.

McKean, Henry P., Jr., 1965, *Appendix: A Free Boundary Problem for the Heat Equation Arising from a Problem in Mathematical Economics* (Academic Press, New York).

Medvedev, Alexey, and Olivier Scaillet, 2010, Pricing American options under stochastic volatility and stochastic interest rates, *Journal of Financial Economics* 98, 145–159.

Merton, Rc, 1976, Option Pricing When Underlying Stock Returns Are Discontinuous, *Journal of Financial Economics* 3, 125–144, WOS:A1976CF62300004.

Nelson, Daniel B., and Krishna Ramaswamy, 1990, Simple Binomial Processes as Diffusion Approximations in Financial Models, *Review of Financial Studies* 3, 393–430.

Omberg, Edward, 1988, Efficient Discrete Time Jump Process Models in Option Pricing, *Journal of Financial and Quantitative Analysis* 23, 161–174.

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

Rad, Jamal Amani, and Kourosh Parand, 2014, Numerical pricing of American options under two stochastic factor models with jumps using a meshless local Petrov-Galerkin method, *arXiv:1412.6064 [cs, q-fin]* .

Ritchken, Peter, L. Sankarasubramanian, and Anand M. Vijh, 1993, The Valuation of Path Dependent Contracts on the Average, *Management Science* 39, 1202–1213.

Ritchken, Peter, and Rob Trevor, 1999, Pricing Options under Generalized GARCH and Stochastic Volatility Processes, *The Journal of Finance* 54, 377–402.

Salmi, S., J. Toivanen, and L. Von Sydow, 2014, An IMEX-Scheme for Pricing Options under Stochastic Volatility Models with Jumps, *SIAM Journal on Scientific Computing* 36, B817–B834.

Stein, Elias, and Jeremy Stein, 1991, Stock Price Distributions with Stochastic Volatility: An Analytic Approach, *Review of Financial Studies* 4, 727–752.

Storn, Rainer, and Kenneth Price, 1997, Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces, *Journal of Global Optimization* 11, 341–359.

Tsitsiklis, J. N., and B. Van Roy, 1999, Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives, *IEEE Transactions on Automatic Control* 44, 1840–1851.

Tzavalis, Elias, and Shijun Wang, 2003, Pricing American Options Under Stochastic Volatility: A New Method Using Chebyshev Polynomials to Approximate the Early Exercise Boundary, SSRN Scholarly Paper ID 381220, Social Science Research Network, Rochester, NY.

Vollrath, Ian, and Juergen Wendland, 2009, Calibration of Interest Rate and Option Models Using Differential Evolution, SSRN Scholarly Paper ID 1367502, Social Science Research Network, Rochester, NY.

Wiggins, James B., 1987, Option values under stochastic volatility: Theory and empirical estimates, *Journal of Financial Economics* 19, 351–372.

Zhang, Xiao Lan, 1997, Numerical Analysis of American Option Pricing in a Jump-Diffusion Model, *Mathematics of Operations Research* 22, 668–690.

Zhylyevskyy, Oleksandr, 2009, A fast Fourier transform technique for pricing American options under stochastic volatility, *Review of Derivatives Research* 13, 1–24.