Exchange Options under Stochastic Volatility and Jump-Diffusion (SVJD) Dynamics

by

Len Patrick Dominic M. Garces
M.Appl.Math. (Mathematical Finance), Ateneo de Manila University, Philippines

A thesis submitted in fulfillment of the degree of

Doctor of Philosophy

Centre for Industrial and Applied Mathematics
UniSA STEM



University of
South Australia

Principal Supervisor: Dr Gerald Cheang
Associate Supervisors: Associate Professor John van der Hoek
Associate Professor Anatoli Torokhti

October 7, 2021

Contents

List of Tables vii

List of Figures ix

Thesis Summary xi

Declaration xiii

Acknowledgments xv

List of Publications and Presentations xvii

1 Introduction 1

1.1 The Option Pricing Problem and the Limitations of the Classical Model: A Background 1

1.2 Beyond Black-Scholes-Merton: A Literature Review 4

1.2.1 American Options 4

1.2.2 Jump-Diffusion and Lévy-Driven Models 5

1.2.3 Stochastic Volatility Models 8

1.2.4 Stochastic Volatility and Jump-Diffusion Models 11

1.2.5 Exchange Options and Other Multi-Asset Options 13

1.2.6 Numerical Methods and Deep Learning Approaches 21

1.3 Statement of the Problem 24

1.4 Structure of the Thesis 26

2 Probabilistic Representations of Exchange Option Prices 29

2.1 Introduction 29

2.2 A Stochastic Volatility Jump-Diffusion Model 30

2.3 A Change of Measure Mechanism 34

2.4 Exchange Option Pricing IPDE 42

2.5 A Representation of the European Exchange Option Price 47

2.5.1 A Change-of-Numéraire Approach 48

2.5.2 A Fourier Transform Approach 53

2.6 A Representation of the American Exchange Option Price 65

2.6.1 Properties of American Exchange Options under SVJD Dynamics 65

2.6.2 The Early Exercise Representation under SVJD Dynamics 74
2.7 Chapter Summary and Conclusions 79

3 An Integral Transform Approach to Exchange Option Pricing 81
3.1 Introduction 81
3.2 The Proportional SVJD Model 83
3.3 Deriving the IPDE for Option Prices 87
3.4 Fourier and Laplace Transform of the Pricing IPDE 90
3.5 Solving the First-Order PDE 98
3.6 Inverting the Laplace Transform 98
3.7 The European Exchange Option Price 99
3.8 The American Exchange Option Price 105
3.9 Exchange Option Prices under Log-Normal Jumps 115
3.10 Chapter Summary and Conclusion 118

4 A Put-Call Transformation of the Exchange Option Problem 121
4.1 Introduction 121
4.2 The Proportional SVJD Model and a Change of Measure 123
4.3 The Exchange Option Pricing IPDE 131
4.4 An Early Exercise Representation 137
4.5 Limit of the Early Exercise Boundary at Maturity 145
4.6 The Transition Density Function 151
4.7 The European Exchange Option Price 159
4.8 The Early Exercise Premium 162
4.9 Chapter Summary and Conclusion 165

5 A Method of Lines (MOL) Approach 167
5.1 Introduction 167
5.2 Method of Lines Implementation 170
5.2.1 MOL Approximations of the Pricing IPDEs 171
5.2.2 Solution via the Riccati Transform 175
5.2.3 Adjustments for the American Exchange Option Price 178
5.2.4 Venttsel Boundary Conditions at $v = v_M$ 181
5.3 Finite Difference with PSOR 183
5.3.1 Spatial Discretization 183
5.3.2 The Semi-Discrete IPDE 186
5.3.3 Completing the Finite Difference Approximation 191
5.3.4 Overview of the PSOR Algorithm for General LCPs 193
5.3.5 The PSOR for LCP (5.31) 194
5.4 Numerical Results and Discussion 195
5.4.1 MOL Approximation of Exchange Option Prices 198
5.4.2 Comparison with Option Prices Approximated with the LSMC and PSOR Algorithms 204
5.4.3 Effect of Venttsel Boundary Conditions at $v = v_M$ 207
5.5 Impact of Jumps and Stochastic Volatility 207
5.5.1 Impact on the Early Exercise Boundary 209
5.5.2 Impact of Model Parameters on Exchange Option Prices 210

5.5.3 Impact of Jumps and Stochastic Volatility on Exchange Option Prices 214
5.6 Chapter Summary and Conclusion 216

6 The Affine Two-Asset SVJD Model and Volatility Indices 219
6.1 Introduction 219
6.2 A Two-Asset Affine SVJD Model with Volatility Jumps and Common Jumps 220
6.3 Affine Dependence of the Volatility Index 226
6.4 Chapter Summary and Conclusion 230

7 Conclusions and Directions for Future Research 233
7.1 Thesis Summary and Conclusions 233
7.2 Directions for Future Research 236

A Preliminary Concepts in Stochastic Calculus 239
A.1 Itô's Lemma for Lévy Processes 239
A.2 Martingale Property of Exponential Processes 241
A.3 Girsanov Theorem for Lévy Processes 243
A.4 The Minimum Entropy Martingale Measure 247
A.5 Transition Probabilities for the Proportional SVJD Model 250

B Proofs of Selected Results in Chapter 3 253
B.1 Proof of Proposition 3.7 253
B.2 Proof of Proposition 3.8 261
B.3 Proof of Proposition 3.9 266
B.4 Proof of Lemma 3.10 268
B.5 Proof of Lemma 3.11 269
B.6 Proof of Proposition 3.14 271
B.7 Proof of Proposition 3.19 277

C Proofs of Selected Results in Chapter 4 283
C.1 Proof of Proposition 4.15 283
C.2 Proof of Lemma 4.16 284
C.3 Proof of Proposition 4.17 286
C.4 Proof of Proposition 4.18 288

D The Trapezoidal Rule and Least Squares Monte Carlo 293
D.1 Trapezoidal Rule for the Sweep Equations 293
D.1.1 The Riccati Transform and the Sweep Equations 293
D.1.2 The Trapezoidal Rule 294
D.2 Least-Squares Monte Carlo Simulation 296
D.2.1 Simulating the Underlying State Variables 296
D.2.2 The Least-Squares Monte Carlo Algorithm 298

E Methods for Parameter Estimation and Calibration of the SVJD Model 301
E.1 Parameter Estimation 301
E.1.1 Maximum Likelihood Estimation and Other Likelihood-Based Approaches . . 302
E.1.2 Generalized Method of Moments (GMM) Estimation 305
E.1.3 Indirect Inference or Efficient Method of Moments (EMM) Estimation 307
E.1.4 Markov Chain Monte Carlo (MCMC) Approach 308

E.1.5 Jump Activity Analysis 311
E.2 Calibration 312
E.2.1 Main Ideas 313
E.2.2 Calibrating Multi-Asset Models 315

Bibliography **317**

List of Tables

4.1 Comparison of notation and coefficients of equation (4.57), Chiarella et al. (2010, Equation 29) and Cheang et al. (2013, Equation 46). 157

5.1 Default SVJD model parameter values and spatial domain discretization used for numerical experiments 197
5.2 Comaparison of discounted American exchange option prices and early exercise boundary estimates using MOL, PSOR, and LSMC 206

6.1 Poisson counting processes for individual and common jumps for a two-asset affine SVJD model 222

List of Figures

2.1 The early exercise boundary and the continuation and stopping regions for the American exchange option 72

4.1 The early exercise boundary and the continuation and stopping regions for the American exchange option in terms of the asset yield ratio $\tilde{S}$. 136
4.2 Illustration of rebalancing costs 143
4.3 Behavior of the early exercise boundary with respect to q_1 149
4.4 Behavior of the early exercise boundary with respect to $\tilde{\lambda}_1$ 150

5.1 Typical number of variance line iterations per integral term iteration at the nth time step for the MOL for European and American exchange options. 198
5.2 Method of lines approximation of the discounted European exchange option price and delta at $\tau = T$ 199
5.3 Method of lines approximation of the early exercise boundary surface 200
5.4 The early exercise boundary $s_1 = B(\tau, v)s_2$ on the s_1s_2-plane 201
5.5 Method of lines approximation of the discounted American exchange option price and delta at $\tau = T$ 202
5.6 Method of lines approximation of the discounted early exercise premium and the difference between the deltas of the American and European exchange options at time $\tau = T$ 203
5.7 Comparison of early exercise boundaries, discounted option prices, and discounted deltas for various choices of boundary conditions 208
5.8 Effect of changes in key jump and stochastic volatility parameters on the early exercise boundary when $v = 0.56$. 209
5.9 Price differences in European and American exchange option prices for various values of the correlation ρ_w between the asset price processes at $t = 0$ and $v = 0.56$. 211
5.10 Price differences in European and American exchange option prices for various values of the jump intensity parameters $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ at $t = 0$ and $v = 0.56$. 212
5.11 Price differences in European and American exchange option prices for various values of the jump size means α_{j_1} and α_{j_2} at $t = 0$ and $v = 0.56$. 212
5.12 Price differences in European and American exchange option prices for various values of the volatility coefficients σ_1 and σ_2 of the asset price processes at $t = 0$ and $v = 0.56$. 213

5.13 Price differences in European and American exchange option prices for various values of the volatility of volatility ω at $t = 0$ and $v = 0.56$. 213
5.14 Comparison of discounted European exchange option prices at $t = 0$ and $v = 0.56$ across different asset price model specifications 215

Thesis Summary

This thesis addresses the problem of pricing European and American exchange options when the underlying asset price processes are modelled as stochastic volatility and jump-diffusion (SVJD) processes. While it is known that European exchange option prices can be computed exactly given the joint characteristic function for the underlying market variables (see for example, Caldana and Fusai, 2013; Cufaro-Petroni and Sabino, 2020), there has been little attention paid to American exchange options due to complications in characterizing the early exercise boundary, more so when using an SVJD model for asset prices. Furthermore, a numerical solution of the corresponding integro-partial differential equation (IPDE) is by no means trivial due to its multi-dimensional nature and the early exercise feature. Thus, this thesis characterizes exchange option prices when the underlying asset price model is a two-asset version of the Bates (1996) SVJD model. Furthermore, it extends existing work that have examined exchange options in the jump-diffusion framework (Cheang and Chiarella, 2011; Cufaro-Petroni and Sabino, 2020) and the stochastic volatility framework (Antonelli and Scarlatti, 2010; Alòs and Rheinlander, 2017; Kim and Park, 2017; Parischa and Goel, 2021) separately.

We construct equivalent martingale measures with respect to our asset price model and derive the IPDEs and boundary conditions for the exchange option prices. We then derive probabilistic representations of the European exchange option price using two approaches: the change-of-numéraire technique (Geman et al., 1995) and Fourier inversion formulas (Caldana and Fusai, 2013). We then derive an early exercise representation for the American exchange option price. From this representation, we find that the presence of jumps significantly impacts the early exercise option since American option holders stand to lose on their investment when asset prices suddenly jump in an undesirable direction, rendering the exercise suboptimal. This loss is referred to as the rebalancing costs due to jumps (Chiarella and Ziogas, 2009).

We also consider a proportional SVJD model for asset prices. The proportional SVJD model

assumes that the two price processes share a single stochastic volatility process and, by construction, belongs to the class of affine jump-diffusions (Duffie et al., 2000). This model specification allows us to directly solve the exchange option pricing IPDE and the corresponding Kolmogorov backward equation for the joint transition density function of the market variables using Fourier and Laplace transforms, thereby extending work on single-asset American options (Chiarella and Ziogas, 2009; Cheang et al., 2013).

To further simplify the American exchange option pricing problem, we use the put-call transformation technique introduced by Bjerskund and Stensland (1993). This technique involves taking one of the assets as the numéraire, reducing the problem to an American call option on the ratio of the two assets. While this technique has been applied to the European case (Margrabe, 1978; Cheang et al., 2006; Antonelli and Scarlatti, 2010; Alòs and Rheinlander, 2017; Parischa and Goel, 2021), its implementation to the American case under SVJD dynamics has yet to be seen. With this simplification, we also analyze the behavior of the early exercise boundary near maturity and discuss the substantial effect of jumps, consistent with the idea of rebalancing costs due to jumps.

Finally, we propose a method of lines (MOL) based approach to solve the option pricing IPDE. The MOL is considered efficient since it can accommodate the free-boundary case with minimal adjustments and is able to generate the option price, delta, and gamma in a single implementation. The MOL also performs faster and more accurately than comparable methods for American option pricing, namely the least squares Monte Carlo (LSMC) method and a projected successive overrelaxation (PSOR) solution of the linear complementary problem. With exchange option prices approximated using the MOL, we find that option prices and the early exercise boundary are significantly affected by the jump and stochastic volatility model parameters.

Declaration

I declare that:

1. this thesis presents work carried out by myself and does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university;
2. to the best of my knowledge it does not contain any materials previously published or written by another person except where due reference is made in the text; and
3. all substantive contributions by others to the work presented, including jointly authored publications, are clearly acknowledged.

Len Patrick Dominic M. Garces

October 7, 2021

Acknowledgments

I thank Dr. Gerald Cheang, my Principal Supervisor, for his guidance and support throughout, which started even before my candidature, and for sharing his knowledge on my research topic. I am also grateful to A/Prof. John van der Hoek and A/Prof. Anatoli Torokhti, my Associate Supervisors, for the invaluable advice they have offered during my studies.

I also express my gratitude to Tricia Huffa, Frances Shantl, and the rest of the ITMS and UniSA STEM research administration team for their assistance in navigating the vast administrative dimension of research at UniSA. I also thank Prof. John Boland, Prof. Regina Burachik, and Dr. Diane Velasquez for their guidance in their capacity as REPL. I am also grateful to Prof. Robert J. Elliott for the conversations and his support, especially with regard to my future career as a researcher. I also thank Dr. Monica Behrend for her invaluable guidance throughout my research proposal preparation and thesis writing.

I also thank the anonymous reviewers of our journal submissions and several participants of the Quantitative Methods in Finance 2019 Conference at Sydney, the online Bernoulli-IMS One World Symposium 2020, and the online 10th Advanced Mathematical Methods for Finance (AMaMeF) Conference for their comments and suggestions that helped improve my work. I am also grateful to my thesis examiners, Prof. Wolfgang Runggaldier and A/Prof. Jonathan Ziveyi for their constructive and thought-provoking comments and suggestions which helped improve the presentation of this manuscript.

The research presented in this thesis has been generously supported by an Australian Government Research Training Program Scholarship (RPTi). I also thank UniSA STEM for financial assistance towards my attendance at the Quantitative Methods in Finance 2019 Conference and Ateneo de Manila University Loyola Schools for financial assistance in my move from the Philippines to Adelaide.

I am also thankful to Dr. Belinda Chiera, Dr. Nick Fewster-Young, A/Prof. Anatoli Torokhti,

Dr. Kevin White, Dr. Rhoda Abadia, and Dr. Neil Bretaña for their confidence in my teaching abilities and providing me the opportunity to teach during my PhD candidature. Special thanks go to Dr. Belinda Chiera for getting me involved in her research projects as a Research Assistant, a role from which I certainly learned a lot.

I am deeply grateful to Engel John Dela Vega for his friendship and support, both academic or personal, before and throughout my candidature. I also thank my friends in the Philippines for checking up on me every now and then despite their busy schedules.

Most importantly, I am profoundly thankful to my parents, Len and Ellen, my siblings, SK and Josef, and to my extended family back in Metro Manila, Philippines and in California, USA for their unending and generous love, support, and encouragement.

Publications and Presentations Based on the Results of the Thesis

The work presented in this thesis serves as the basis of the following manuscripts and conference presentations.

Publications and Preprints

1. G. H. L. Cheang & L. P. D. M. Garces. Representation of exchange option prices under stochastic volatility and jump-diffusion dynamics. *Quantitative Finance*, 20(2): 291-310. 2020. DOI: https://doi.org/10.1080/14697688.2019.1655785.

   Note: This paper is based on the results contained in Chapter 2 of this thesis.

2. L. P. D. M. Garces & G. H. L. Cheang. Put-call transformation of the exchange option problem under stochastic volatility and jump-diffusion dynamics. Preprint, https://arxiv.org/pdf/2002.10194.pdf. 2020.

   Note: This paper is based on the results contained in Chapter 4 of this thesis.

3. L. P. D. M. Garces & G. H. L. Cheang. A numerical approach to pricing exchange options under stochastic volatility and jump-diffusion dynamics. *Quantitative Finance*. 2021. DOI: https://doi.org/10.1080/14697688.2021.1926534.

   Note: This paper summarizes the first half of Chapter 4 (up until the analysis of the near-maturity behavior of the early exercise boundary) and presents the numerical results obtained in Chapter 5 of this thesis.

Conference Presentations

1. Representation of exchange option prices under stochastic volatility and jump-diffusion dynamics. *Quantitative Methods in Finance 2019 Conference*, Sydney NSW, Australia, 19 December

2019.

2. Put-call transformation of the exchange option problem under stochastic volatility and jump-diffusion dynamics. *Bernoulli-IMS One World Symposium 2020*, Online, 24 to 28 August 2020.

3. A numerical approach to pricing exchange options under stochastic volatility and jump-diffusion dynamics. *10th General Advanced Mathematical Methods for Finance (AMaMeF) Conference*, Online, 22 to 25 June 2021.

All presentations were prepared and delivered by Len Patrick Dominic M. Garces based on the journal manuscript or preprint with the same title enumerated in the list of publications.

CHAPTER 1

Introduction

1.1 The Option Pricing Problem and the Limitations of the Classical Model: A Background

An option is a financial instrument that gives the holder the right, but not the obligation, to buy or sell an underlying asset on or before a particular date, referred to as the maturity of the option. A call option is the option to buy the asset, whereas a put option is the option to sell the underlying. An option is a particular class of financial derivatives, financial instruments whose value derives from the value of one or more underlying assets (Hull, 2012). The issue of valuing or ascribing a "fair price" to options and other financial derivatives is a central question in financial mathematics.

Options may roughly be classified (although not exhaustively) by the manner in which it is exercised relative to its maturity. European options may only be exercised at the maturity of the option, whereas American options may be exercised any time before maturity (Black and Scholes, 1973; Merton, 1973). An intermediate between these two types is the Bermudan option, an option that can be exercised on pre-identified dates prior to, and including, the maturity date of the option. It is by virtue of these early exercise opportunities that the American and Bermudan options are not of the same value as the identical European option (Merton, 1973).

Options may also be defined on more than one underlying asset or with other exercise mechanics. A basic example of a multi-asset option is the exchange option, the option to exchange one risky asset for another (Margrabe, 1978). Other examples include the spread option, min/max options, and correlation options (Broadie and Detemple, 1997). Options with "non-standard" mechanics are known as exotic options, of which examples are forward start options, gap options, Asian options, and barrier options. Davis (2010a) and Hull (2012) provide additional examples of exotic options.

While the origins of option pricing is attributed to Bachelier (1900), the first mathematically

satisfactory analysis of the problem came from Black and Scholes (1973) and Merton (1973), who used continuous-time arbitrage and hedging arguments and stochastic analysis to derive a closed form equation for the price of European options. The key assumption made by Black and Scholes (1973) and Merton (1973) is that the underlying asset price process is described by a geometric Brownian motion

$$\mathrm{d}S(t) = \mu S(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}W(t),$$

where μ is the instantaneous expected return on the asset, σ^2 is the instantaneous variance of the return, and W is a standard Brownian motion. This specification encodes the assumption that asset returns are normally distributed and asset volatility is constant. In their derivation of option pricing formulas, Black and Scholes (1973) mainly employed hedging arguments while Merton (1973) used methods in stochastic analysis.

The classical Black and Scholes (1973) and Merton (1973) result was further refined by Cox and Ross (1976) who observe that the option valuation is independent of investor preferences. From this observation, Cox et al. (1979), in their discrete-time binomial model for option pricing, were able to express the option price as an expectation of its discounted payoff in a risk-neutral world. This notion was then solidified in continuous-time by Harrison and Kreps (1979) and Harrison and Pliska (1981), who introduced the notions of equivalent martingale measures and complete markets and showed that the current option price is equal to the expected value, with respect to an equivalent martingale measure, of the discounted option payoff. A detailed historical account of the development and solutions of the option pricing problem and the Black and Scholes (1973) theory are given by Benth (2010), Davis (2010b), and Jarrow (2010).

Since then, contemporary literature in financial mathematics has grown with characterizations of option prices under asset price models that are clear departures or extensions from the classical Black and Scholes (1973) framework. It has been shown that the geometric Brownian motion is insufficient to capture typical empirical features of asset price data (also referred to as *stylized features*) in the market (see for example Cont, 2001; Schoutens, 2003; Kou, 2002, 2008, and references therein). Most notably, these empirical studies of daily asset price returns show that: (1) the distribution of asset returns exhibit tail-ends which are heavier than the tails postulated in a Gaussian distribution (implied by a geometric Brownian motion assumption on asset prices), (2) volatility is not constant and exhibits strong autocorrelation over periods of time, and (3) volatility is negatively correlated with asset returns. Other stylized facts include disparities in asset return behavior and volatilities over different time scales, asymmetry in upward and downward price movements, and absence of linear serial correlations among asset returns.

As such, one major strand in financial literature is the development of alternative asset price models that are able to capture these empirical characteristics more accurately. Examples of such

alternative asset price models include stochastic volatility models (Hull and White, 1987; Stein and Stein, 1991; Heston, 1993), jump-diffusion models (Merton, 1976; Kou, 2002), and combinations of these two models (Bates, 1996; Scott, 1997; Bakshi et al., 1997; Duffie et al., 2000). There has also been, and continues to be, great interest in developing models based on Lévy processes (Schoutens, 2003; Cont and Tankov, 2004a; Eberlein, 2009), a rich class of stochastic processes with independent and stationary increments which are able to represent skewness and excess kurtosis that are typical in asset price returns. However, markets in which asset prices exhibit these features are generally incomplete because they contain risks which cannot be fully eliminated by just trading in the basic traded assets. Therefore, there exist multiple equivalent martingale measures, each of which correspond to a pricing rule that is consistent with the no-arbitrage condition (Harrison and Pliska, 1981). Furthermore, option pricing via perfect hedging is no longer feasible, so alternative methodologies must be considered to characterize option prices and the corresponding risk management framework.

In addition to formulating alternative models for asset price dynamics, many studies have also moved from pricing plain vanilla instruments (such as the European stock option) and delved into analyzing more sophisticated instruments such as exotic options, options with early exercise features (for example, American-style or Bermudan options), and options written on assets or commodities other than stocks. Analyses of multi-asset options (for example, exchange options, spread options, min/max options) under myriad underlying asset price dynamics are also present in financial and mathematical literature.

With increasingly complicated asset price models and option payoffs, often with consideration for multi-dimensional cases, closed form expressions for option prices become more elusive. As such, researchers have resorted to a variety of analytical approximations, simulations, and numerical methods to characterize option prices under more sophisticated asset price dynamics. With the increased availability of powerful computational capabilities in computers, researchers have also applied techniques in deep learning to address option pricing, hedging, model calibration, and other problems in mathematical finance.

The rest of this introductory chapter is organized as follows. In Section 1.2, we review salient research related to the topic of this thesis. In particular, we describe in this section existing work on American options, alternative asset price models (stochastic volatility, jump-diffusion, stochastic volatility and jump-diffusion), analyses of multi-asset options, and the development of numerical methods to address these option pricing problems. We then formulate our main research problem and list our main contributions in Section 1.3. The outline of the thesis and the key results in each chapter are briefly described in Section 1.4.

1.2 Beyond Black-Scholes-Merton: A Literature Review

Before we present the main research problem of this thesis, we review existing results related to American option pricing, option pricing under alternative asset price models, exchange options and other multi-asset options, and numerical methods for option pricing.

1.2.1 American Options

American options differ from their European counterpart in that American options may be exercised on or before the specified maturity date. The possibility of early exercise therefore provides an additional mathematical challenge to the option pricing problem. The pricing of American options is also significant since most exchange-traded options are American. In this section, we briefly look into the significant contributions made towards addressing the American option pricing problem following the outline of Myneni (1992), Salopek (1997), and Barone-Adesi (2005).

The American option pricing problem has historically been tackled as an optimal stopping problem (Bensoussan, 1984; Karatzas, 1988). Within this framework, the price of the American option is the largest of the set of all discounted expected rewards at each possible exercise date (represented as stopping times with respect to a given filtration) from the date of valuation until the option's maturity. At each point in time, the decision to exercise is driven by an unknown early exercise boundary that defines the stopping and continuation regions. That is, if the underlying asset price is within the stopping region, the option holder should exercise the option, whereas if it is in the continuation region, the option holder should hold off from exercising the option. Properties of the put option price with respect to the early exercise boundary are discussed in Myneni (1992) and have been expanded to outside the Black-Scholes framework by Pham (1997) and Touzi (1999).

In the analysis of McKean (1965), the pricing problem was transformed from an optimal stopping problem to a free boundary problem for partial differential equations. In this formulation, additional boundary conditions known in the literature as the "smooth-pasting" or "smooth-fit" condition (Myneni, 1992; Pham, 1997; Touzi, 1999), must be specified to accommodate the unknown early exercise boundary. The connection between the optimal stopping problem and the free boundary problem for the American put option was elucidated by Van Moerbeke (1976).

Closely related to the free boundary problem formulation of the American option price is its formulation as a solution of a variational inequality, as put forward by Bensoussan and Lions (1982) and Jalliet et al. (1990). In this particular formulation, the early exercise or free boundary does not appear explicitly, and so the variational inequality formulation lends itself nicely to a number of computational methods.

Following the major developments in American option pricing, it became well-known that the

price of the American option can be decomposed into the price of the corresponding European option and an early exercise premium (Kim, 1990; Jacka, 1991; Carr et al., 1992; Jamshidian, 1992; Myneni, 1992). This is known as the early exercise representation of the American option. This representation is particularly important since it suggests that one may use existing methods to price European options under any given underlying asset price dynamics and characterize the corresponding early exercise premium. The second step, however, is far from trivial and largely depends on the assumed underlying asset price model.

Parallel to all the work being done on the analysis of American option prices, significant contributions were also made toward finding numerical solutions and analytic approximations of option prices and the early exercise boundary. Among numerous studies, notable ones include the binomial tree approach by Cox et al. (1979), the compound options approach by Geske and Johnson (1984), the quadratic approximation by Barone-Adesi and Whaley (1987), a numerical solution by Allegretto et al. (1995), and the Monte Carlo simulation approaches by Longstaff and Schwartz (2001) and Ibáñez and Zapatero (2004). Pre-1997 work on American put options are surveyed more comprehensively by Salopek (1997) and the book by Chiarella et al. (2015) contains an extensive treatment of American option prices in the advent of alternative asset pricing formula. Computational methods for American options on underlying assets driven by a geometric Brownian motion are surveyed in Broadie and Detemple (1996).

1.2.2 Jump-Diffusion and Lévy-Driven Models

One of the primary motivations of incorporating jumps in asset pricing models is empirical evidence that asset prices tend to exhibit sudden jumps, especially when observed in finer time resolutions (for example, intra-day or weekly) (Cont and Tankov, 2004a). Furthermore, the distribution of asset returns tend to possess tails that are heavier than postulated with a Gaussian distribution, which is used in the classical analysis (Cont, 2001; Cont and Tankov, 2004a; Kou, 2008). Furthermore, jumps in stock prices may also be rationalized as the way market participants react to innovations or news in the market; that is, sudden movements in asset prices represent market reactions to good or bad news, which may be assumed to arrive according to a Poisson process (Merton, 1976; Kou, 2002).

The basic jump-diffusion model for the asset price $S(t)$ has the form

$$\frac{\mathrm{d}S(t)}{S(t-)} = (\mu - \lambda\kappa)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t) + (Y-1)\,\mathrm{d}N(t),$$

where $S(t-) = \lim_{u\to t^-} S(u)$ is the pre-jump price of the stock (equal to $S(t)$ if no jump occurs at time t), μ and σ^2 are the instantaneous expected return and variance of the return (conditional on no jumps), W is the standard Wiener process, N is a Poisson process (independent of the

Wiener component) with intensity λ, Y is the random jump-size variable, and $\kappa \equiv \mathbb{E}_{\mathbb{P}}[Y-1]$ is the mean percentage change in the stock price when the jump event occurs. The diffusion part $\sigma\,\mathrm{d}W(t)$ represents the normal market noise, whereas the jump part $\mathrm{d}N_t$ represents "unusual" market shocks (for example, availability of new market information). The solution $S(t)$ of the stochastic differential equation above is given by

$$S(t) = S(0)e^{X(t)}, \qquad \text{where} \qquad X(t) = (\mu - \lambda\kappa)t - \frac{1}{2}\sigma^2 t + \sigma W(t) + \sum_{n=1}^{N(t)} Y_n.$$

The two most popular jump-diffusion model specifications are those of Merton (1976) and Kou (2002). Merton (1976) assumes that $\log Y$ has a normal distribution while Kou (2002) assumes that $\log Y$ has an asymmetric double exponential distribution.

If the market consists only of the risky underlying asset and a riskless bond, then it is not possible to construct a replicating portfolio corresponding to an option on the underlying asset (Naik and Lee, 1990). In this regard, the market is said to be incomplete and consequently there exist many equivalent martingale measures. Merton (1976) works around this by essentially assuming that the jump risk is not priced (Cheang and Chiarella, 2012). In contrast, Cheang and Chiarella (2012) considered a hedging argument involving two options on the underlying asset with different maturity dates in order to properly account for jump risk. A crucial result in Merton (1976) and Cheang and Chiarella (2012) is that if $\log Y$ is normally distributed, then the price of a European call option can be expressed as an infinite sum of Black-Scholes-type option prices weighted by Poisson probabilities emanating from the number of jumps occurring up until the option's maturity.

The double-exponential model of Kou (2002) retains the attractive properties of the Merton (1976) jump-diffusion model, such as the ability to replicate the leptokurtic and heavy-tailed features of asset returns. The double-exponential model is also motivated by the varying degree (over- or under-reaction) by which investors react to sudden market shocks or news. As opposed to the normal jump-diffusion model of Merton (1976), the double-exponential model is analytically tractable for path-dependent options, such as perpetual American options and barrier options.

Numerous other studies have continued to use jump-diffusion models in pricing options. Naik (1993) considers the option pricing and hedging problem where the volatility of asset prices is subject to discontinuities, in contrast to Merton's approach where jumps occur in the stock price return. Pham (1997) and Gukhal (2001) discuss a model for American option prices under jump-diffusion dynamics. They find that the early exercise decision is significantly impacted by the presence of jumps. Ekström and Tysk (2007a,b) discuss properties of option prices, including convexity and monotonicity in parameters, when assets have jump-diffusion dynamics. Chiarella and Ziogas (2004, 2009) extend the Fourier transform approach in pricing American options of McKean (1965) to the case of jump-diffusion dynamics. Cheang and Teh (2014) extend Merton's

jump-diffusion model to the case where the discount factor is also a modelled as a jump-diffusion process. Wang and Chen (2009) formulate a regime-switching version of the Merton (1976) jump-diffusion model. The calibration of jump-diffusion processes has been tackled by Bjursell and Gentle (2012) and Xu and Jia (2017) (see also Chapter 6). Cont et al. (2007) and Mina et al. (2015) consider the hedging problem in an incomplete market driven by jump-diffusion processes.

The presence of jumps implies that option prices are now solutions of an *integro*-partial differential equation (IPDE) that involves integrals of the yet-unknown option price function. As analytical solutions are rarely available for IPDEs, numerical methods are typically employed to solve the option price. Some numerical methods for solving IPDEs shall be discussed in Section 1.2.6.

The jump-diffusion models of Merton (1976) and Kou (2002), by and large, assume that the frequency of jumps is finite. This assumption allows us to write the cumulative jump movement over the period $[0, t]$ as a compound Poisson process, as shown in the exponent $X(t)$ of $S(t) = S(0)e^{X(t)}$. The inclusion of jumps in the asset price dynamics can be generalized in at least two ways.

The first extension is to assume that X is Lévy process, a stochastic process with independent and stationary increments. This leads to an *exponential Lévy model* for S (see for example Cont and Tankov, 2004b). With exponential Lévy models, the exponent X can be an infinite activity Lévy process, a class of Lévy processes with infinitely many jumps in every interval. Unlike the finite activity Lévy process assumed in the Merton (1976) or Kou (2002) jump-diffusion, there is no need to specify a Brownian component in the infinite activity case since it is able to capture frequent small movements and rare large movements (Carr et al., 2002). Well-known examples of infinite activity exponential Lévy models are the variance-gamma model (Madan et al., 1998), the normal inverse Gaussian model (Barndorff-Nielsen, 1998), the CGMY model (Carr et al., 2002), and the class of hyperbolic models (Eberlein et al., 1998). Lévy processes and their applications to finance have been extensively surveyed by Schoutens (2003), Cont and Tankov (2004a), and Eberlein (2009). The key advantage of exponential Lévy models is the availability of the characteristic function of the Lévy process X via the Lévy-Khintchine formula (see for example Cont and Tankov, 2004a, Theorem 3.1), thereby allowing the efficient pricing of European options using Fourier inversion techniques (see for example Carr and Madan, 1999; Lee, 2004).

Another way to extend the Merton (1976) and Kou (2002) jump-diffusion models is to assume that S is the solution of a stochastic differential equation with jumps

$$\mathrm{d}S(t) = \mu(t, S(t-))\,\mathrm{d}t + \sigma(t, S(t-))\,\mathrm{d}W(t) + \int_E \gamma(t, S(t-), z)p(\mathrm{d}t, \mathrm{d}z),$$

where $\mu : [0,T] \times \mathbb{R}^+ \to \mathbb{R}$, $\sigma : [0,T] \times \mathbb{R}^+ \to \mathbb{R}$, and $\gamma : [0,T] \times \mathbb{R}^+ \times E \to \mathbb{R}$ are predictable stochastic processes and $p(\mathrm{d}z, \mathrm{d}t)$ is a σ-finite random counting measure on the mark space $(E, \mathcal{E})$

with intensity $\lambda_t(\mathrm{d}z)$ (i.e. the compensator of $p(\mathrm{d}z, \mathrm{d}t)$ is $\lambda_t(\mathrm{d}z)\,\mathrm{d}t$) (Brémaud, 1981; Runggaldier, 2003). The intensity $\lambda_t(\mathrm{d}z)$ can be assumed to have the form $\lambda(t)m(t, \mathrm{d}y)$, where λ is a nonnegative predictable process representing the intensity of a Poisson counting process $N(t, E)$ and $m(t, \mathrm{d}y)$ is a probability measure on E. It is also typically assumed that $\lambda_t(E) < \infty$ a.s. for all finite t (Björk et al., 1997b; Runggaldier, 2003; Platen and Bruti-Liberati, 2010). The random measure $p(\mathrm{d}z, \mathrm{d}t)$ may also be taken as the random measure corresponding to a square-integrable Lévy process (see for example Øksendal and Sulem, 2019). SDEs with jumps such as the one specified above have been used as underlying asset price models for option pricing (see for example Duffie et al., 2000; Pan, 2002; Runggaldier, 2003) or as models for bond markets (Björk et al., 1997b,a). This specification is also particularly useful as it easily lends itself to an extension to the stochastic volatility case.

In light of stylized facts discovered around asset returns, it is noted that jump-diffusion models are unable to capture volatility clustering effects (Cont, 2001; Kou, 2008). This implies that jump-diffusion models perform optimally in pricing short-term derivatives where volatility clustering is less evident (Kou, 2002, 2008). The same can be said for exponential Lévy models, such as those stated above. A commonly proposed fix to improving the long-term fit of jump-diffusion models is the incorporation of stochastic volatility, which has been done by Bates (1996), Bakshi et al. (1997), Scott (1997), and Duffie et al. (2000), among numerous others. For models based on pure-jump Lévy processes, i.e. the exponential Lévy model for asset prices, stochastic volatility can be achieved via a stochastic time change (Geman et al., 2001) using the integral of a Cox et al. (1985) process (Carr et al., 2003).

1.2.3 Stochastic Volatility Models

The view that asset price volatility is stochastic, in contrast to the constant volatility assumption of the classical Black and Scholes (1973) analysis, is motivated by empirical evidence of non-constant implied volatilities and the presence of high autocorrelation in the time series of squared asset returns (Cont, 2001; Kou, 2008). The first phenomenon is frequently referred to as the implied volatility smile or skew, a phenomenon in which the implied volatility is significantly dependent on the strike price and maturity of the option. Implied volatility smiles are also related to the observation that asset price returns and volatility are negatively correlated, exhibiting a leverage effect. High autocorrelation in squared asset returns is connected to volatility clustering, the situation where periods of high volatility are more likely to be followed by periods of similarly high volatility. These empirical findings thus prompted a more specialized treatment of asset price volatility.

One major area of development in this regard is the use of stochastic volatility models, models where volatility itself is assumed to follow a certain stochastic process. For a stochastic process to be suitable to model volatility, it has to have positive values. Furthermore, mean-reversion is

another attractive quality of a stochastic volatility model. The choice of stochastic process to model volatility is largely motivated by these qualitative properties, in addition to analytical tractability. Similar to markets where asset prices jump, a market in which volatility is stochastic is incomplete since there are more sources of risk than there are assets (the risky and riskless assets).

One of the earlier research incorporating stochastic volatility to the Black-Scholes model was that of Hull and White (1987). They assumed that the stock price $S(t)$ and the instantaneous variance $\sigma^2(t)$ are modelled using independent geometric Brownian motions, i.e. their stochastic variance process is not mean-reverting. Their model also assumes that volatility is uncorrelated with aggregate consumption (thereby nullifying the effects of investors' risk preferences). The primary drawback of this model is that no analytic form for the option price can be obtained. To this end, Hull and White (1987) produced a power series approximation for the option price.

While the Hull and White (1987) model ensures that volatility is positive, it does not provide mean-reversion properties. To this end, mean-reverting models for stochastic volatility were introduced by Scott (1987), Stein and Stein (1991) and Ball and Roma (1994). Scott (1987) and Stein and Stein (1991) used an Ornstein-Ulhenbeck process to model volatility, while Ball and Roma (1994) used the Cox et al. (1985) square root process. Both models retain the Hull and White (1987) assumption that the stock price and volatility processes are uncorrelated and obtain option pricing formulas in terms of characteristic functions.

Using a geometric Brownian motion for the asset price dynamics and the square root pricess for the volatility process, Heston (1993) considers the case where the asset price and volatility are correlated. The Heston (1993) model is specified as

$$\begin{aligned} \mathrm{d}S(t) &= \mu S(t)\,\mathrm{d}t + \sqrt{v(t)}S(t)\,\mathrm{d}W_1(t), \\ \mathrm{d}v(t) &= \kappa(\theta - v(t))\,\mathrm{d}t + \sigma\sqrt{v(t)}\,\mathrm{d}W_2(t), \end{aligned}$$

where $\mu, \kappa, \theta, \sigma$ are constants and W_1 and W_2 are correlated Wiener processes with correlation coefficient ρ. The parameter σ is known as the volatility of volatility parameter. Taking $\sigma \to 0$ allows one to recover a time-dependent volatility version of the Black and Scholes (1973) result. To properly price volatility risk, Heston (1993) assumes that the market price of volatility risk is proportional to the current volatility level. Within this framework, the price of the European call option has the form

$$C(S, v, t) = SP_1 - Ke^{-r(T-t)}P_2,$$

where P_1 and P_2 are risk-neutral probabilities that the option expires in-the-money. These probabilities are obtained via an inverse Fourier transform of their associated characteristic functions. The Heston (1993) model is particularly useful in that these probabilities are analytically available.

An alternative to the Heston (1993) model is the 3/2 model (Heston, 1997; Platen, 1997; Lewis,

2000) where the volatility process is given by

$$\mathrm{d}v(t) = \kappa v(t)(\theta - v(t))\,\mathrm{d}t + \sigma(v(t))^{3/2}\,\mathrm{d}W_2(t).$$

Since the rate of mean reversion is now $\kappa v(t)$ instead of just κ in the Heston (1993) model, the 3/2 model allows for a faster reduction in volatility following unusually large volatility spikes (Christoffersen et al., 2010). Grasselli (2017) proposed the 4/2 model, an analytically tractable unified approach that combines the Heston (1993) "1/2" model and the 3/2 model. Cheng et al. (2019) considers a multivariate, mean-reverting extension of the 4/2 model.

In order for stochastic volatility models to be tenable for option pricing under the arbitrage framework, the existence of an equivalent martingale measure must be established. Sin (1998) provides conditions for the existence of equivalent martingale measures in a general stochastic volatility framework. Care must also be taken in stochastic volatility models to ensure that the volatility process does not venture to zero or explode to infinity during the life of the option. Properties of stochastic volatility models in relation to this consideration have been analyzed by Feller (1951), Lewis (2000), Andersen and Piterbarg (2007), and Bertini and Passalacqua (2008).

Since their introduction, stochastic volatility models have appeared prominently in financial and econometric literature. Shephard (2005b) presents a compilation of works in relation to building stochastic volatility models, parameter estimation and inference, and applications to option pricing. Shephard (2005a) provides a comprehensive treatment of the origins and development of stochastic volatility in financial literature and practice. Fouque et al. (2011) provide a textbook treatment to stochastic volatility models with applications to pricing options and other derivatives. Gatheral (2006) discusses stochastic volatility models and provides an in-depth view of modelling volatility surfaces and pricing exotic options with stochastic volatility. Antonelli et al. (2009) present a general stochastic volatility framework for pricing European call options via a power series approach on the correlation parameter between the asset price process and the volatility process. Christoffersen et al. (2009) considered a "double" Heston stochastic volatility specification where the asset price has two separate volatility factors each driven by a square-root process. They argue that a multi-factor stochastic volatility specification improves the fit of volatility smiles and the volatility term structure. Christoffersen et al. (2010) consider further alternative specifications when modelling the volatility dynamics of the S& P500 index. Integral transform methods (i.e. Fourier and Laplace transforms) have been employed by Chiarella et al. (2010), for which an extensive numerical implementation was provided by Adolfsson et al. (2013). Zhylyevskyy (2010) used Fast Fourier Transform (FFT) methods to price American options under stochastic volatility dynamics.

While stochastic volatility models have been able to capture volatility clustering, smiles, and skews, they are unable to characterize discontinuities in stock price returns (i.e. jumps) and short-

term behavior in implied volatility (Cont and Tankov, 2004a). As shall be seen in the next section, a movement to improve the short-term performance of stochastic volatility models is one that is focused on incorporating jumps in asset prices.

Local volatility models represent an alternative approach to accounting for the empirical properties of implied volatility. Developed by Derman and Kani (1994) and Dupire (1994) in the discrete- and continuous-time settings, respectively, the basic local volatility model assumes that the asset price volatility is a deterministic function of time and the asset price,

$$\mathrm{d}S(t) = \mu(t)S(t)\,\mathrm{d}t + \sigma(t, S(t))\,\mathrm{d}W(t).$$

The local volatility function $\sigma(t, s)$ is not known, but may be assumed to be a parametric function whose parameters are estimated from the data. Local volatility models are constructed to perfectly replicate the observed prices of call and put options. Specifically, the local volatility model is the unique diffusion process whose risk-neutral distribution is derived completely from the market prices of European options. However, since European option prices are not available at a continuum of strike prices and maturity, the estimation of local volatility models are heavily dependent on numerical methods. Local volatilities are also inadequate in capturing dynamic properties of spot asset prices and volatilities and are unable to accurately price exotic options. To address these shortcomings, researchers have considered local stochastic volatility models where the best features of local volatility and stochastic volatility models are considered. For a more comprehensive overview of local volatility and local stochastic volatility models, see Musiela and Rutkowski (2005, Chapter 7). See also Tian et al. (2015) and Saporito et al. (2019), and the references therein, for a more detailed discussion of local stochastic volatility models.

Rough volatility has also recently emerged as an area of interest in volatility modelling, motivated by the observation from high-frequency data that the time series of realized variance behaves very similar to a fractional Brownian motion. Recent evidence has also emerged that the usual local, stochastic, and local-stochastic volatility models where volatility is taken to be a diffusion process is unable to properly fit the volatility skew of at-the-money (ATM) options (also known as the power law the volatility skew). Thus, so-called rough volatility models have been proposed to address these shortcomings of traditional volatility models and improve the fit on volatility time series data and volatility skews. For an overview of rough volatility models, its motivation, and its applications, see for example Bayer et al. (2016), Gatheral et al. (2018), Fukasawa (2020).

1.2.4 Stochastic Volatility and Jump-Diffusion Models

Asset price models with both stochastic volatility and jumps serve to encompass more of the observed empirical properties of asset returns and to improve the pricing and hedging of short-

term and long-term derivatives (Bates, 1996; Bakshi et al., 1997; Scott, 1997). In this section, we look into some of the prominent investigations done on the use of stochastic volatility jump diffusion models, their application in option pricing, and their empirical performance.

Having analyzed options on the Deutsche Mark, Bates (1996) first proposed a stochastic volatility jump-diffusion model that combines the features of the jump-diffusion model of Merton (1976) and the square-root stochastic volatility model of Heston (1993). The Bates (1996) model is given by the pair of stochastic differential equations

$$\begin{aligned} \mathrm{d}S(t) &= (\mu - \lambda\kappa)S(t)\,\mathrm{d}t + \sqrt{v(t)}\,\mathrm{d}W_1(t) + (Y-1)\,\mathrm{d}N(t), \\ \mathrm{d}v(t) &= \xi(\eta - v(t))\,\mathrm{d}t + \sigma\sqrt{v(t)}\,\mathrm{d}W_2(t), \end{aligned}$$

where we use the same notation used by Merton (1976) and Heston (1993). Under this model, the price of European options may be derived in terms of the characteristic function of the log-price, which is available in closed form.

Scott (1997) considers the Merton (1976) jump-diffusion model with stochastic volatility and stochastic interest rates in pricing European stock options. In this model, the instantaneous interest rate is taken to be a linear combination of two independent Cox et al. (1985) processes, one of which drives the volatility of the stock price. Option prices were then derived using Fourier inversion techniques based on the characteristic function of the stock price. While it was shown that the inclusion of jumps and stochastic volatility have an empirically significant impact on option prices, Scott (1997) makes no remark on how close his option price estimates are to observed option prices.

Bakshi et al. (1997) provide an empirical evaluation of competing stock price models, including SVJD models, and their performance in pricing options. In their analysis, they found that SVJD models are misspecified, but are the least so compared to the Black-Scholes model, the stochastic volatility model, and the stochastic interest rate model. They also note that the SVJD model yields the least out-of-sample error compared to its competitors, the jumps contributes to a better fit in short-term options, and the stochastic interest rate improves fit with respect to long-term options. They also showed that SVJD models performed satisfactorily in terms of minimum-variance hedging and delta-hedging due to its control for stochastic volatility. Similar to Bates (1996) and Scott (1997), Bakshi et al. (1997) obtain their option pricing formulas via Fourier inversion.

It is no coincidence that European option pricing with the stochastic volatility jump-diffusion models specified by Bates (1996), Bakshi et al. (1997), and Scott (1997) is achievable using the characteristic function of the underlying state variables (asset price and variance). Duffie et al. (2000) found that the SVJD model specifications of Bates (1996), Bakshi et al. (1997), and Scott (1997) belong to the class of affine jump-diffusion (AJD) processes. Affine jump-diffusion models are jump-diffusion models (including stochastic volatility models) whose parameters (i.e. drift

coefficient, instantaneous covariances, and jump intensities) are assumed to be affine functions of some underlying state vector. Duffie et al. (2000) then show that the characteristic function of the AJD process is known analytically up to the solution of two complex-valued ordinary differential equations that depend on the model parameters. They then formulate a series of transforms based on the characteristic function that allows for an efficient pricing of options via the Lévy inversion formula (see for example Williams, 1991, Section 16.6).

Eraker et al. (2003) proposed including jumps not only in the stock price returns, but also in the return of the volatility process in response to contemporaneous findings that models with diffusive volatility process, such as those used by Bates (1996) and Scott (1997), are statistically misspecified. This is motivated by empirical findings by Bates (2000) and Pan (2002) who analyze option prices under a two-factor SVJD model for stock prices. By including jumps in the volatility process, Eraker et al. (2003) consequently allow volatility to increase rapidly, possibly to represent phenomena such as market stress or crashes. To estimate model parameters, jumps, and volatilities, a Markov Chain Monte Carlo (MCMC) method was employed. While an empirical analysis of the effect jumps in volatility on European call option prices was conducted, Eraker et al. (2003) do not derive option pricing formulas based on their proposed model. Nonetheless, it was noted that jumps in volatility improve the model's fit to prices by providing a persistent driving factor for sharp increases in volatility and that jumps in volatility lead to a significant increase in implied volatilities from option prices. Further evidence supporting a stochastic volatility with asset return jumps and volatility jumps for accurate equity option pricing is provided by Bakshi et al. (2021).

Despite econometric issues that hound these models (see Bates, 2000; Andersen et al., 2002; Pan, 2002; Chernov et al., 2003; Eraker et al., 2003), the Bates (1996) and Scott (1997) stochastic volatility jump-diffusion models have been extensively applied in option pricing literature, especially for the pricing of European-type contracts. The pricing of American options when asset prices are driven by SVJD processes have received less attention in the literature. To this end, Cheang et al. (2013) derive probabilistic representations of American option prices when the underlying asset is modelled using the Bates (1996) SVJD model. Chiarella et al. (2009) provide a numerical solution to the American option pricing problem under SVJD dynamics using the method of lines.

1.2.5 Exchange Options and Other Multi-Asset Options

Classical and alternative asset price models have also been applied in the pricing of multi-asset options (options written on two or more underlying assets). In this section, we provide an overview of extant literature on the multi-asset option pricing, with a particular focus on exchange and spread options.

An exchange option is an agreement to exchange one risky asset for another Margrabe (1978). As with other European-style options, the European exchange option may only be exercised at the

maturity date of the option, which we denote by T. If the contract stipulates that the holder of the option receives one share of asset 1 with price $S_1(T)$ in exchange for one share of asset 2 with price $S_2(T)$, then the terminal payoff of the European exchange option is given by

$$V(T) = \big(S_1(T) - S_2(T)\big)^+ = \max\{S_1(T) - S_2(T), 0\}. \tag{1.1}$$

From here, it can be seen that if the price of asset 2 is fixed, then the exchange option reduces to a European call option on asset 1 with strike price $S_2(T)$; similarly, if the price of asset 1 is fixed, then the exchange option is a European put option on asset 2.

Margrabe (1978) first analyzed the European exchange option under the classical Black-Scholes assumptions, i.e. the asset price processes are modelled as a pair of correlated geometric Brownian motions with constant volatilities. However, instead of solving the resulting pricing PDE directly, Margrabe (1978) transforms the exchange option problem into a European call option problem on the ratio of the prices of the two assets, a procedure that is now known as a change-of-numéraire technique (Geman et al., 1995). This technique takes advantage of the homogeneity of the payoff function with respect to the underlying asset prices and reduces the exchange option pricing problem to the problem of pricing a call option on $S_1(T)/S_2(T)$ with strike price 1. Fischer (1978) considered a closely related problem of determining European call option prices when the exercise price is also a diffusion process. McDonald and Siegel (1985), in their analysis of firm valuation, were able to extend the Margrabe formula to the case in which the underlying assets pay a continuous dividend yield. Bernard and Cui (2010) extend the Margrabe formula to consider stochastic interest rates, possibly correlated to the asset price processes. As these studies have been conducted in the Black-Scholes framework, the results are subject to the same criticisms challenging the classical results.

The valuation of the European exchange option has since then been conducted under alternative asset price models, such as jump-diffusion models, stochastic volatility models, and exponential Lévy models:

- Jamshidian (2010) considers the pricing of European exchange options and constructing hedging portfolios when asset prices are driven by pure-diffusion processes with deterministic volatility and when they are modelled using exponential Poisson processes. The analysis of Jamshidian (2010) emphasizes a solution via self-financing trading strategies instead of a PDE approach.

- Antonelli and Scarlatti (2010), Alòs and Rheinlander (2017), Kim and Park (2017), and Parischa and Goel (2021) provide prices for the European exchange option under stochastic volatility dynamics. Except Kim and Park (2017), these aforementioned authors followed the

change-of-numéraire technique followed by Margrabe (1978) to reduce the dimensionality of the exchange option pricing problem.

Antonelli and Scarlatti (2010) extend the correlation expansion approach proposed by Antonelli et al. (2009) to the case where the underlying assets prices are modelled with fairly general stochastic volatility dynamics, covering both the Hull and White (1987) and the Heston (1993) model specifications. The correlation expansion approach leads to a first-order approximation of the solution of the pricing PDE as a Taylor polynomial in the correlations between the Brownian motions present in the asset price and volatility dynamics. The authors report varied levels of performance quality depending on the maturity of the options and the assumed values of the correlation coefficients, although propose a slight modification to their first-order approximation to improve option price estimates.

Alòs and Rheinlander (2017) propose a decomposition of the exchange option price in terms of the Black and Scholes (1973) price of a call option on the ratio of the two asset prices and several integrals with respect to square-integrable martingales in the equivalent martingale measure implied by taking the second asset as the numéraire. The authors then derive explicit expressions for the decomposition formula using Malliavin calculus and discuss mean-variance hedging strategies within this framework. Alòs and Rheinlander (2017) show that their results hold for both the Heston (1993) and the 3/2 stochastic volatility models.

Kim and Park (2017) assume that asset prices follow the fast mean-reverting stochastic volatility model of Fouque et al. (2011) and obtain an approximation of the exchange option price using a multiscale asymptotic expansion. They then provide a series of numerical experiments illustrating the effect of stochastic volatility relative to prices calculated using the Margrabe (1978) formula.

Parischa and Goel (2021), using two correlated Heston (1993) stochastic volatility models for asset prices, provide a solution using the characteristic function of the log of the ratio of the two asset prices. With additional restrictions on the correlation coefficients leading to an exponential-affine characteristic function, the authors obtained a pricing formula based on the Lévy inversion formula (see also Duffie et al., 2000).

- Quittard-Pinon and Randrianarivony (2010), Cheang and Chiarella (2011), Li et al. (2018), Ma et al. (2020), and Cufaro-Petroni and Sabino (2020) analyzed European exchange option prices when asset prices are driven by jump-diffusion models.

  Quittard-Pinon and Randrianarivony (2010) and Li et al. (2018) lay out the details of the change-of-numéraire technique for the case where one asset has a jump-diffusion price dynamics while the other is a pure diffusion process. Quittard-Pinon and Randrianarivony (2010) also consider the case where jumps follow a double exponential distribution as in Kou (2002).

With an emphasis on the application of the change-of-numéraire technique, both analyses provide a probabilistic representation of exchange option prices, rather than going through the integro-partial differential equation approach. Both analyses provide a comparison of exchange option prices computed using the Margrabe (1978) formula and those which incorporate jumps and find a significant deviation from prices under a misspecified diffusion model.

Cheang and Chiarella (2011) extended the Merton (1976) jump-diffusion model to the case of two assets and assumed the existence of common jumps between the two asset prices emulating macroeconomic or industry-wide shocks. The model specified by Cheang and Chiarella (2011) also allows for correlated jump sizes. The authors provided additional details on the construction of a risk-neutral measure and the derivation of the integro-partial differential equation. The authors eventually provide probabilistic representations of the European exchange option price.[1]

Cufaro-Petroni and Sabino (2020) generalized the results of Cheang and Chiarella (2011) by considering more general asset price dynamics (going beyond the Merton (1976) specification) and more modes of possible correlation between jumps (i.e. independent jumps, correlation via common jumps, and correlation via self-decomposability). Cufaro-Petroni and Sabino (2020) also considered the more general delay conditions on how macroeconomic shocks impact asset prices. Cufaro-Petroni and Sabino (2020) provided various numerical illustrations applying their pricing formulas to exchange options in energy markets.

Ma et al. (2020) consider the case where the underlying assets are driven by Hawkes jump-diffusion processes, emulating the phenomenon of clustered jumps. Recognizing that a solution via IPDEs is challenging given their model specification, Ma et al. (2020) instead proposed a solution using the characteristic function of the ratio of the two asset prices. While the resulting model is affine, its characteristic function has to be solved numerically. The authors also discussed the calculation of hedge ratios for the European exchange option price and illustrated how one can measure systematic risk given clustered jumps. Ma et al. (2020) also discuss that the exchange option pricing framework can be used to value other similarly structured financial contracts in firm valuation and portfolio management.

- Boyarchenko and Levendorskiĭ (2002, Chapter 9), Eberlein and Papapantoleon (2005), Fajardo and Mordecki (2006) tackled the exchange option pricing under the assumption that the underlying prices are driven by exponential Lévy processes. These authors take advantage of the change-of-numéraire technique, also referred to as the dual market method by

[1] Refinements to the representation of European exchange options derived by Cheang and Chiarella (2011) are provided by Caldana et al. (2015).

Fajardo and Mordecki (2006), to simplify the exchange option pricing problem. Eberlein and Papapantoleon (2005), in particular, focused on obtaining symmetry relations between the exchange option and a plain vanilla European option on the ratio of the two asset prices, while Fajardo and Mordecki (2006) derived pricing formulas based on the inversion of the characteristic function of the exponential Lévy process. We note however that the analysis of Fajardo and Mordecki (2006) is fairly general, allowing for other option payoff functions as long as they are homogenenous in the underlying asset prices.

Analysis of exchange options with non-European payoffs have also gained traction in financial literature. Bjerskund and Stensland (1993) analyzed the American exchange option where asset prices follow correlated geometric Brownian motions and proposed a procedure called the put-call transformation to reduce the two-dimensional optimal stopping problem into a one-dimensional problem involving a finite-lived American call or put option. This then allows the use of analytical approximations (see for example Barone-Adesi and Whaley, 1987) to obtain exchange option prices. Carr (1995) extended the Geske and Johnson (1984) method for the American put option to price American exchange options where the underlying dividend-paying assets follow a geometric Brownian motion. Gerber and Shiu (1996) consider the pricing of a perpetual exchange option using the same model specification as Margrabe (1978). Following their analysis of the European exchange option under jump-diffusion dynamics, Cheang and Chiarella (2011) obtain a linked system of integro partial-differential equations characterizing the price of the American exchange option and the associated early exercise boundary. Cheang and Lian (2015) develop a model to price perpetual exchange options under jump-diffusion dynamics, while Peng and Peng (2016) analyze the price of Bermudan-style exchange options under a jump-diffusion model. Fajardo and Mordecki (2006) also analyzed perpetual exchange options under their exponential Lévy asset price models. Typical of American-style and early exercise derivatives, however, is the unavailability of a closed-form option pricing formula, and so one must resort to approximations or numerical solutions. Bounds for the value of a finite-expiration American exchange option, where asset prices are driven by correlated geometric Brownian motions, are constructed in Morozov and Khizhnyak (2016).

Barrier-type exchange options as extensions of the Margrabe-type European exchange option have also been analyzed. Haug and Haug (2002) and Haug (2007) provided formulas for down-and-in and up-and-in exchange options in terms of Black and Scholes (1973) call and put option formulas following Margrabe's asset price model specification. In these types of exchange options, the option is "knocked in" if the ratio of the asset prices S_1/S_2 hits a barrier H before the option matures. Lindset and Persson (2006) focused the knock-out exchange option where the option is "knocked out" at the first instance $S_1(t)$ hits $S_2(t)$ from above. Using a model-free arbitrage argument, Lindset and Persson (2006) show that the value at time t of a knock-out exchange option

is simply the difference $S_1(t) - S_2(t)$. As these basic barrier exchange options knock in or out the option at the instant the barrier is reached, holders of such options may be left vulnerable to short-term market fluctuations that trigger the barrier event. To this end, Chen and Suchanecki (2011) consider Parisian exchange options, a type of barrier option that knocks the option in or out when one option is worth more than the other for a certain period of time. Chen and Suchanecki (2011) provide Parisian exchange option pricing formula under the Margrabe (1978) asset price model specification by considering the excursion of the asset price ratio process and a change-of-numéraire.

The exchange option is a special case of the spread option, an option whose payoff is dependent on the difference of the prices of two risky assets (known as the spread) relative to a given strike price. For example, the payoff function of a European call spread option on two assets with prices $S_1(t)$ and $S_2(t)$, strike price K, and maturity T is given by (Carmona and Durrleman, 2003a)

$$V\text{spr}(T) = \big(S_1(T) - S_2(T) - K\big)^+ = \max\big\{S_1(T) - S_2(T) - K, 0\big\}. \tag{1.2}$$

In particular, the exchange option is a spread option whose strike price is 0. However, as Carmona and Durrleman (2003a) point out, closed-form expressions (similar to that of Margrabe (1978)) are not available for European spread options, even in the pure-diffusion or log-normal setting. As such, a substantial body of research has been conducted with respect to the approximate pricing of spread options and obtaining semi-closed form pricing formulas.

A prominent area of research in the pricing of European spread options is the construction of analytical pricing approximations for spread option prices when the underlying asset prices are modelled as diffusion processes. One of the most popular analytical approximations under the log-normal asset price model is that of Kirk (1995). Accurate approximations via lower bounds for both spread option prices and sensitivities (i.e. the greeks), which are important for hedging, have been derived by Carmona and Durrleman (2003b). They also consider approximation formulas for when asset prices are driven by jump-diffusion processes. Using a quadratic approximation of the exercise boundary, Deng et al. (2008) derive closed-form approximations for spread option prices and sensitivities as sums of one-dimensional integrals. Li et al. (2010) proposed an extension of the Kirk (1995) approximation and a second-order approximation for spread options on more than two assets. Venkatramanan and Alexander (2011) developed approximation formulas by expressing the spread option price as a sum of two compound exchange options, i.e. exchange options on two exchange options. Bjerskund and Stensland (2011) develop a closed-form lower bound approximation for European call spread option prices that was shown to perform better than the Kirk (1995) approximation. van der Hoek and Korolkiewicz (2012) consider an analytical approximation of spread options via a perturbation expansion on the partial differential equation

satisfied by the spread option price, provided that the asset prices are driven by geometric Brownian motions with nonzero constant correlation.

Another strategy employed in deriving spread option pricing formulas is the use of Fourier transforms. Following the Fourier inversion formula derived by Carr and Madan (1999) for the modified European call option, Dempster and Hong (2002) provide two-dimensional Fourier integral formulas and propose fast Fourier transform (FFT) algorithms to price spread options. Hurd and Zhou (2010), following the approach of Lewis (2001) for single-asset options, considered an expression of the Fourier transform of the payoff function in terms of the complex gamma function. Caldana and Fusai (2013) used the Fourier transform method of Dempster and Hong (2002) to provide analytical approximations for spread option prices under general asset price dynamics, provided that the joint characteristic function is available, as an extension of the valuation formula developed by Bjerskund and Stensland (2011). In contrast to the FFT approach of Hurd and Zhou (2010), exact European exchange option prices can be obtained directly from the approximation of Caldana and Fusai (2013) by setting the spread option strike price to zero. Informed by the FFT approach of Hurd and Zhou (2010), Cane and Olivares (2014) use a two-dimensional Bates (1996) model to price European spread options. Alfeus and Schlögl (2019) showed that the Hurd and Zhou (2010) method is a particular application of the two-dimensional Parseval's Identity. A striking feature of the Fourier transform approach is that closed form expressions or approximations for European spread option prices can be obtained with knowledge of the joint characteristic function of the log-prices of the underlying assets, which is available in most contexts.

Parallel to the developments in European spread option pricing, investigations have also been made into American spread option pricing. Gerber and Shiu (1996) provide a martingale treatment to the pricing of perpetual American options on two assets under the assumption that prices are driven by a two-dimensional Wiener process. Broadie and Detemple (1997) provide a framework for pricing American-style multi-asset options and analyzing the associated early exercise boundaries assuming that asset prices are modelled using correlated pure-diffusion processes. In addition to spread and exchange options, Broadie and Detemple (1997) also considered other options such as the option on the maximum, the arithmetic mean, or the products or powers of two assets and higher-dimensional options. Ekström and Tysk (2007b) provide general convexity and monotonicity properties of multi-asset American options written on assets that exhibit jumps in prices, extending the single-asset case considered by Pham (1997). Chiarella and Ziveyi (2011) use a method of lines procedure to provide a numerical solution to the pricing of American spread options under stochastic volatility dynamics. In addition to their analysis on European spread options, Venkatramanan and Alexander (2011) also derive early exercise representations for American spread options within the geometric Brownian motion framework. Morozov and Khizhnyak (2013) construct an upper bound for the value of perpetual options on the sum or difference of two assets driven by geometric

Brownian motions. In particular, they analyze the perpetual exchange call option and the infinite basket call option. Following the methods of McKean (1965) and Jamshidian (1992), an analysis of American options written on two underlying assets in the pure-diffusion setting was tackled via partial differential equations and Fourier integral transforms by Chiarella and Ziveyi (2014).

As analyzed by Broadie and Detemple (1997), the shape of the early exercise boundary for multi-asset American options vastly differs depending on the option being considered. Significant attention has therefore been spent characterizing the geometry of the early exercise boundary for multi-asset American options for a variety of payoff functions and asset price dynamics. As mentioned before, Broadie and Detemple (1997) have studied this problem when the underlying assets are modelled as correlated geometric Brownian motions. Villeneuve (1999) extended the analysis to include more general American rainbow options, providing theoretical results for n-asset American options under more general diffusion dynamics and providing results that complement those of Broadie and Detemple (1997) for specific payoff functions. The main result established by Villeneuve (1999) is the non-emptiness of the exercise region for the n-asset American option. Rapuch (2005) extend the results of Villeneuve (1999) to the case where assets follow a local volatility specification. Rapuch (2005) also showed that American option prices are convex in the volatility parameter. For multi-asset American options written on assets whose prices are modelled as general diffusion processes, Nyström (2008) examine the regularity properties of the early exercise boundary near option maturity. Mishura and Shevchenko (2009) provide specialized results for the finite-maturity exchange option written on assets whose prices are modelled as geometric Brownian motions and derive an integral equation for the early exercise boundary.

Exchange options may also be seen as special case of power exchange options which have a payoff function

$$V^{\text{peo}}(T) = (\beta_1 S_1^{\alpha_1}(T) - \beta_2 S_2^{\alpha_2}(T))^+,$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are real constants. Power exchange options simultaneously generalize exchange options and single-asset power options and are useful in pricing contracts related to executive stock options and firm valuation. Blenman and Clark (2005) were first to consider the pricing of power exchange option under a pure diffusion asset pricing model. Wang (2016) extended this to the case where the assets are modelled as jump-diffusion processes with correlated jumps (similar to Cheang and Chiarella (2011)). Wang et al. (2017) further extends this, accounting for counterparty risk. Parischa and Goel (2019) use a Hawkes jump-diffusion model specification in their analysis of power exchange options.

1.2.6 Numerical Methods and Deep Learning Approaches

As noted prior, the primary complication with the use of stochastic volatility and SVJD dynamics is the unavailability of analytical option price formulas. Alongside the developments made in the Fourier transform approach to solving the option pricing problem, researchers have also analyzed a variety of numerical methods to solve the IPDEs or numerically evaluate integrals associated to the option price. A numerical approach to solving option price PDEs in the Black-Scholes-Merton framework is not a new concept and have thus been treated comprehensively in a number of textbooks on quantitative finance (see Brandimarte, 2006; Duffy, 2006; Cerny, 2009; Seydel, 2017; Radi and El Hami, 2018, for example). In this section, we spend some time here reviewing how numerical methods have been applied particularly to pricing American-style options given the problem's intimate relation to free-boundary problems.

Finite difference methods (FDM) are based on approximating partial derivatives using difference quotients derived from a Taylor's approximation of the function (Brandimarte, 2006). Explicit and implicit finite difference methods result to algebraic equations which will be solved to find the value of the unknown function at the grid points. The FDM is particularly useful when the domain of the equation is rectangular, but complications arise when dealing with non-uniform grids or non-rectangular grids (Seydel, 2017). Furthermore, finite difference schemes in higher dimensions result to systems of equations for which solutions are numerically expensive to calculate (Brandimarte, 2006). In higher dimensions, the Alternating Directions Implicit (ADI) method, a variant of the FDM that employs an operator splitting approach, has been used to price European options under jump-diffusion dynamics (Andersen and Andreasen, 2000). Similar applications for pricing American options using the FDM are presented in Ikonen and Toivanen (2004, 2007, 2009). Applications of the FDM in pricing European and American options under jump-diffusion dynamics are found in Carr and Hirsa (2003), Briani et al. (2004), d'Halluin et al. (2004b), d'Halluin et al. (2004a), and Almendral and Oosterlee (2005). Finite differences have also been applied by Toivanen (2010a) and Salmi et al. (2013) to price American options whose underlying asset follows the Bates (1996) SVJD dynamics. These studies feature a number of adjustments in handling the integral term, such as quadrature schemes, arising from the jump in addition to usual PDE discretizations. Finite difference methods for options with early exercise have been surveyed by Toivanen (2010b).

An alternative to FDM is the finite element method (FEM). While a variety of numerical algorithms have been placed under the umbrella term *finite element method*, the basic idea is that the domain of the PDE/IPDE is divided into sub-domains, and the solution is approximated via a weighted sum of basis functions that satisfy certain error-minimization criteria (Gander and Kwok, 2018). FEM is particularly useful in dealing with high-dimensional problems and the method can be designed to handle non-uniform and/or non-rectangular grids (Seydel, 2017). A finite element

method was employed in Achdou and Tchou (2002) to solve the Black-Scholes equation under a mean-reverting Ornstein-Ulhenbeck equation for the stochastic volatility. Matache et al. (2005) provided a wavelet-based finite element method to solving pricing IPDEs under discontinuous Lévy processes (which include jump-diffusion processes). Ballestra and Sgarra (2010) apply a finite element approach to pricing American options under SVJD dynamics, taking advantage of a Richardson extrapolation of Bermudan options to deal with early-exercise considerations. Wang and Chen (1993) formulated a numerical method for solving IPDEs arising from jump-diffusion asset price dynamics by considering a finite difference approximation of the spatial derivatives and a discontinuous Galerkin FE discretization with respect to time. A comprehensive exposition of finite element methods for derivative pricing is given in Hilber et al. (2013).

A third approach to solving the pricing IPDEs is the method of lines (MOL). The method of lines consists in transforming a partial differential equation into a system of ordinary differential equations using algebraic approximations in all but one of the independent variables (Schiesser and Griffiths, 2009; Chiarella et al., 2015; Meyer, 2015). In a PDE with a spatial and temporal dimension, the time-discrete MOL or Rothe's method arises when the spatial variable is left continuous (Meyer, 2015).[2] The approximations may be based on difference quotients, finite elements, or series expansions, but one that employs finite differences are the most basic (Meyer, 2015). In pricing American options, the method of lines is particularly advantageous in that the option price, the early exercise boundary, the option delta, and the option gamma are all determined as part of the numerical algorithm (Chiarella et al., 2009). This method has been applied in pricing American options in the Black-Scholes framework by Meyer and van der Hoek (1997) and under jump-diffusion dynamics by Meyer (1998). When the method of lines is applied to the option pricing PDE, the result is a second-order ordinary differential equation. Meyer and van der Hoek (1997) and Meyer (1998) apply a Riccati transformation to further simplify the second-order ODE as a system of first order ODEs for which numerical solutions are readily available. Adolfsson et al. (2013) adapt this particular implementation of the MOL to pricing American options under Heston (1993) SV dynamics and Chiarella et al. (2009) do the same for American options under the Bates (1996) SJVD model. Kang and Meyer (2014) also used the MOL to price American options under stochastic volatility and stochastic interest rates. Ballestra and Sgarra (2010) point out, however, that the numerical implementation in Chiarella et al. (2009) is not numerically efficient due to their use of a fixed-point iteration procedure. Chiarella and Ziveyi (2014) propose a method of lines solution to pricing American spread options where asset prices share one stochastic volatility process. Under the Black-Scholes-Merton framework, Kovalov et al. (2007) employ a method of lines based on finite elements and a penalty term in pricing American-style spread options, geometric mean puts,

[2] This is in contrast to the situation where the spatial variable is instead discretized, resulting to ordinary differential equations in t. This is the typical MOL procedure in engineering problems, whereas the time-discrete MOL is what is used in financial problems (Meyer, 2015).

and max/min options.

Advances have also been made towards the numerical implementation of Fourier inversion formulas in option pricing. Carr and Madan (1999), Lewis (2001), Lee (2004), and Kwok et al. (2012) have provided a Fourier transform analysis and a fast Fourier transform technique for single-asset options. Dempster and Hong (2002), Hurd and Zhou (2010), Caldana and Fusai (2013), and Cane and Olivares (2014) have proposed extensions of Fourier transform techniques to price multi-asset European-style options (for example, European spread and exchange options).

Recently, the BENCHOP (The BENCHmarking Project in Option Pricing) was launched in order to provide researchers a host of option pricing problems that can be used to compare the performance of numerical methods and to serve as a springboard for the development of new computational methods for option pricing.[3] This project looks at six benchmark models corresponding to the assumed underlying asset price dynamics, namely: (1) the Black and Scholes (1973) and Merton (1973) geometric Brownian motion for one asset; (2) the Black-Scholes-Merton model with discrete dividends; (3) the Black-Scholes-Merton model with local volatility; (4) the Heston (1993) stochastic volatility model for one underlying asset; (5) the Merton (1976) jump-diffusion model for one underlying asset; and (6) the correlated Black-Scholes-Merton model for two underlying assets (i.e. the Margrabe (1978) model). Under each model, the price of European call options (or the European spread option in the case of the sixth model) was computed. In the first and second benchmark problems, the prices of American calls, American puts, and barrier options were also considered. Each problem was solved via Monte Carlo methods, Fourier methods, finite difference methods, and radial basis function methods. Von Sydow et al. (2015) compare the performance of these numerical methods and discuss the individual merits and disadvantages of each computational method.

The aforementioned numerical methods for solving partial differential equations and integro-partial differential equations are susceptible to the curse of dimensionality. That is, these methods become increasingly more complicated to implement the higher the dimensionality of the option pricing problem (i.e. when the option is written on more assets or when there are additional stochastic processes relevant to the underlying asset price dynamics). To this end, solutions for high-dimensional partial differential equations based on deep learning methodologies (see for example Lagaris et al., 1998; Sirignano and Spiliopoulos, 2018) have been proposed in financial literature, given the ability of feedforward neural networks to approximate any arbitrary function and its derivatives (Hornik et al., 1990). Deep learning methods for solving high-dimensional partial differential equations arising in financial applications have been proposed by E et al. (2017), Han et al. (2018), Beck et al. (2019), and Liu et al. (2019b), among others. Deep learning approaches

[3] Details about the project can be found in http://www.it.uu.se/research/scientific_computing/project/compfin/benchop and are discussed in the paper of Von Sydow et al. (2015).

have also been used to analyze financial time series and other financial market data (Cavalcante et al., 2016; Bao et al., 2017) and model calibration Itkin (2019); Liu et al. (2019a); Horvath et al. (2021).

1.3 Statement of the Problem

Based on the literature review in the previous section, an analysis of American exchange options under such asset price dynamics has not yet been done. In addition, the effect of stochastic volatility and jumps have only been analyzed separately for the European exchange option. Therefore, this thesis contributes to existing literature by providing an in-depth probabilistic analysis of European and American exchange option prices and provides numerical methods for approximating exchange option prices when asset price dynamics exhibit both stochastic volatility and jumps. In particular, this thesis presents the following contributions towards the resolution of this problem:

1. Assuming that the two underlying asset prices are modelled as a pair of correlated Bates (1996) stochastic volatility jump-diffusion processes, we derive a probabilistic version of the early exercise representation for American exchange option prices, i.e. the American exchange option price as the sum of its European counterpart and an early exercise premium, both of which are written as expectations under a risk-neutral measure. From these representations, we obtain key insights about the effect of jumps on option prices.

2. Under a modified stochastic volatility and jump-diffusion model specification for underlying assets, we present a solution of the corresponding exchange option pricing integro-partial differential equation using Fourier and Laplace integral transforms and verify an integral version of the early exercise representation for American exchange options.

3. With the same model specification as in (2), we apply the *put-call transformation* (initially suggested by Bjerskund and Stensland (1993) as a strategy for pricing American exchange options under GBM assumptions) to reduce the dimensionality of the original exchange option pricing problem. While this technique has been applied to the European case (Alòs and Rheinlander, 2017; Antonelli and Scarlatti, 2010; Cheang et al., 2006; Margrabe, 1978; Parischa and Goel, 2021), its implementation to the American case under SVJD dynamics has yet to be seen. Under this lower-dimensional version of the problem, we solve the Kolmogorov backward equation for the joint transition density function of the underlying state variables and present the corresponding integral representations of the European and American exchange option prices.

4. For the lower-dimensional exchange option pricing IPDE resulting from the put-call transformation, we formulate a numerical method based on the method of lines (MOL) to numerically

approximate European and American exchange option prices. We also provide a numerical solution of the linear complementarity problem (LCP) corresponding to American exchange option prices using the projected successive overrelaxation (PSOR) algorithm. The PSOR-approximated prices serve as the baseline to which the MOL method is assessed (see for example Chiarella et al., 2009; Ikonen and Toivanen, 2007). These numerical solutions are by no means trivial due to the presence of jumps, the multidimensional nature of the IPDE, and the presence of an unknown free boundary.

Before proceeding, we state some general assumptions and limitations of this research. We focus mainly on the pricing of exchange options assuming that model parameters have been estimated and that the model has been appropriately calibrated. In Appendix E, we provide an overview of parameter estimation and model calibration techniques that can be applied to our model following the discussion of the affine structure of our SVJD model in Chapter 6. In addition, we maintain some of the classical Black and Scholes (1973) assumptions on the market, namely (1) the risk-free interest rate is a known constant, (2) there are no transaction costs, (3) fractional sales or purchases of shares is permitted, and (4) short-selling is possible. Additional assumptions shall be stated in each chapter as required.

Furthermore, we emphasize that our approach to pricing options is that of arbitrage-free or risk-neutral pricing, rather than through utility indifference pricing and hedging (Becherer, 2004; Biagini et al., 2010; Delbaen et al., 2002; Henderson and Hobson, 2009; Hodges and Neuberger, 1989; Musiela and Zariphopoulou, 2004), super-replicating strategies (Bellamy and Jeanblanc, 1999; Eberlein and Jacod, 1997; El Karoui and Quenez, 1995; Kramkov, 1996), or pricing rules implied by risk measures (Artzner et al., 1999; Klöppel and Schweizer, 2007; Xu, 2006). The financial market implied by SVJD dynamics is incomplete, so the equivalent martingale measure (EMM) is not unique. Thus, we assume that the pricing measure has already been selected from the collection of martingale measures equivalent to the historical or objective market measure and we price the exchange options as expectations under the chosen pricing measure. In Chapter 2, we characterize a class of EMMs (via a Radon-Nikodým derivative process) corresponding to the choice of the money market account as the numéraire, while in Chapter 4, we consider EMMs that arise from choosing the second asset yield process as the numéraire asset (also via an appropriate Radon-Nikodým derivative process). The pricing of exchange options then take place under the probability measure $\mathbb{Q}$ (or $\hat{\mathbb{Q}}$ following the notation in Chapter 4) that is assumed to belong to the class of EMMs characterized by our specific Radon-Nikodým derivative process. As such, in this thesis, we do not discuss ways to identify a specific equivalent martingale measure satisfying certain distance conditions[4] as the pricing measure, although in Chapter 2, we provide initial steps

[4] Delbaen and Schachermayer (1996) and Schweizer (1996), for instance, consider the minimum martingale measure which minimizes the $L^2(\mathbb{P})$-norm of the density of $\mathbb{Q}$. Goll and Rüschendorf (2001) characterize minimal

towards the identification of the minimum entropy martingale measure discussed in Frittelli (2000); Fujiwara and Miyahara (2003); Miyahara (1996, 2012). Rather, we focus on the mathematical and computational methods for determining option prices under any EMM as characterized by our specific Radon-Nikodým derivative process.

While the thesis focuses on exchange options typically written on stocks, one may consider exchange options written on other assets such as indices and foreign currencies, using the results contained in this thesis as a starting point. For foreign currencies, in particular, the dividend yields are replaced by risk-free interest rates in the domestic and foreign money markets. Siegel (1995) explains how exchange options can be used to estimate the "implicit beta" between an underlying stock and a given market index. The exchange option framework may be adapted to investigate real options (Kensinger, 1988; Carr, 1995; Armada et al., 2007), outperformance options (Cheang and Chiarella, 2011)[5], energy market options (surveyed in Benth and Zdanowicz, 2015)[6], the option to enter/exit an emerging market (Miller, 2012), and problems in corporate finance where companies bid for another firm using their own shares (Chen and Suchanecki, 2011), among others. Ma et al. (2020) provide additional examples of financial contracts which can be priced under the exchange option framework. As such, the exchange option pricing framework can be extended to any problem in which an investor is interested in analyzing the relative performance of two assets.

1.4 Structure of the Thesis

Below, we summarize the main structure and significant findings of each chapter in the thesis.

Chapter 2 We specify a two-asset stochastic volatility and jump-diffusion model based on the model specifications of Bates (1996), Heston (1993), and Merton (1976). A Radon-Nikodým derivative process is also introduced to facilitate the shift from the objective or historical market measure to other equivalent probability measures. We state additional conditions on the form of the Radon-Nikodým to characterize equivalent martingale measures under which option pricing is performed. Under the equivalent martingale measure, we derive the integro-partial differential equation that characterizes the exchange option prices. We also

distance martingale measures for general semimartingales via the f-divergence distance with respect to the market measure $\mathbb{P}$. The f-divergence, defined as $f(\mathbb{Q}|\mathbb{P}) = \int_\Omega f(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}})\,\mathrm{d}\mathbb{P}$ (where $f : (0,\infty) \to \mathbb{R}$ is convex), if it exists, generalizes the relative entropy distance (Frittelli, 2000; Miyahara, 1996), among others, but does not include the minimum martingale measure. Under exponential utility functions, it has been shown (Delbaen et al., 2002; Goll and Rüschendorf, 2001) that the problem of finding the minimal distance martingale measure is the dual problem of the utility maximization problem.

[5] Cheang and Chiarella (2011) assumed that only one asset price process had jumps while the other was modelled as a pure-diffusion process. Quittard-Pinon and Randrianarivony (2010) discuss in greater detail the European exchange option pricing problem under a similar model specification.

[6] SVJD model specifications similar to ours may be appropriate for most commodity classes, such as oil, agricultural commodities, or precious metals. However for electricity markets, prices jump depending on the interaction of demand and supply, so jump arrivals may be driven by hitting times instead of the usual Poisson counting process (Carmona and Coulon, 2012). For details on the particularities of modelling commodity spot prices, see for example Geman (2005) or Dempster and Tang (2016).

derive representations of the European exchange option price using the change-of-numéraire technique proposed by Geman et al. (1995) and the Fourier inversion formula defined by Caldana and Fusai (2013), and show with detailed calculations that these two representations are comparable. Lastly, we show that the American exchange option can be decomposed into the price of the European exchange option and an early exercise representation.[7]

Chapter 3 We assume that asset prices are modelled using a *proportional stochastic volatility jump-diffusion model*, an alternative SVJD model specification in which the two asset price processes share the same stochastic volatility process. Under this model specification, we again derive the integro-partial differential equation for the exchange option prices. This model induces a structure of the integro-partial differential equation which can be solved using Fourier and Integral transforms as in Cheang et al. (2013). Using the integral transform method, we derive integral representations of the European and American exchange option prices.

Chapter 4 With the same asset price dynamics assumed in Chapter 3, we consider the pricing of European and American exchange options under an equivalent martingale measure obtained by setting the second asset yield process as the numéraire. Such a choice for the numéraire reduces the exchange option pricing problem, a two-dimensional problem, to pricing a call option written on the ratio of the yield processes of the two assets, a one-dimensional problem. The joint transition density function of the asset yield ratio process and the instantaneous variance process is then determined from the corresponding Kolmogorov backward equation via integral transforms. We then determine integral representations for the European exchange option price and the early exercise premium and state a linked system of integral equations that characterizes the American exchange option price and the associated early exercise boundary. Properties of the early exercise boundary near maturity are also discussed.[8]

Chapter 5 We propose a numerical solution, based on the method of lines (MOL), of the IPDE arising from the change-of-numéraire considered in Chapter 4. We discuss its specific implementation using computing software and possible alternative boundary conditions at the far limits of the computational domain. Our analytical and numerical investigation shows that the near-maturity behavior of the early exercise boundary of the American exchange option is significantly influenced by the dividend yields and the presence of jumps in the underlying asset prices. Furthermore, with the numerical results generated by the MOL, we are able to show that key jump and stochastic volatility parameters significantly affect the

[7] This is adapted from the abstract of the paper titled "Representation of Exchange Option Prices under Stochastic Volatility and Jump-Diffusion Dynamics", which is based on this chapter and is published in *Quantitative Finance*.

[8] This abstract is adapted from the arXiv preprint titled "A Put-Call Transformation of the Exchange Option Problem under Stochastic Volatility and Jump Diffusion Dynamics" which is based on this chapter.

early exercise boundary and exchange option prices. Our numerical analysis also verifies that the MOL performs more efficiently, compared to other finite difference methods or simulation approaches for American options, since the MOL integrates the computation of option prices, greeks, and the early exercise boundary and does so with the least error. The MATLAB codes used in this chapter can be found here: https://drive.google.com/drive/folders/1KrLDaiXiXPfaJ1JulRyDVKFgwJae82_v?usp=sharing.

Chapter 6 One notable complication one faces in parameter estimation is that the instantaneous variance process is a latent or unobservable variable. While there exist a number of approaches that can handle unobserved processes, we consider an alternative method that takes advantage of the fact that the volatility index defined on an asset (for example, the VIX for the S&P 500) can be written as an affine function of the instantaneous variance process if the asset price process is affine. We show how this property holds under an extension of the two-asset stochastic volatility jump-diffusion specification we considered in Chapter 2. This chapter presents initial work towards the parameter estimation of the SVJD model using the methods for fully-observed processes discussed in Appendix E.

CHAPTER 2

Probabilistic Representations of Exchange Option Prices

2.1 Introduction

In this chapter, we derive representations of the European and American exchange option prices when the underlying asset prices are characterized by a stochastic volatility jump-diffusion model. After establishing the two-dimensional Bates (1996) model of the financial market and the necessary change-of-measure mechanisms, we derive the integro-partial differential equation (and the associated boundary conditions) that characterizes exchange option prices under SVJD dynamics. We then employ probabilistic arguments to obtain expressions for the European and American exchange option prices arising from the IPDE.

Our discussion of the European exchange option employs two methods: a probabilistic approach and a Fourier transform approach. First, we demonstrate the use of the Geman et al. (1995) change-of-numéraire technique to obtain an alternative probabilistic representation of European exchange option prices that resembles the original Margrabe (1978) formula. In this analysis, we also establish some conditions on the volatility processes to ensure that asset and option prices are well-defined after changes in probability measures. Second, with the imposition of additional assumptions on the correlation structure of the market model, we derive the the joint characteristic function of the log-prices of the stocks (following Cont and Tankov (2004a) and Cane and Olivares (2014)) and use the result of Caldana and Fusai (2013) to obtain the price of the European exchange option. Our analysis shows that the Caldana and Fusai (2013) formulation is compatible with the representation of option prices obtain via the change-of-numéraire procedure, linking the probabilities of option exercise (under the secondary probability measures) to Fourier inversion formulas.

To the best of our knowledge, an analysis of American exchange options under both stochastic volatility and jump-diffusion dynamics is yet to be formulated (see Section 1.2.5). As such, this

chapter aims to extend the probabilistic analysis of American exchange options by Cheang and Chiarella (2011) to the case of stochastic volatility jump-diffusion dynamics. Under the SVJD model, we were able to show that the American exchange option price can also be decomposed into the sum of the price of the European exchange option and an early exercise premium, the latter of which can be decomposed further into diffusive and jump components (a feature that was also shown by Cheang et al. (2013) for the single-asset American option). In this regard, we also derive the coupled system of integral equations that determine both the American exchange option price and the unknown early exercise boundary.

The rest of the chapter is organized as follows. Section 2.2 presents the stochastic volatility jump-diffusion model based on Bates (1996) and some pertinent results surrounding the stochastic volatility process; Section 2.3 presents the Radon-Nikodým derivative that will be used to shift to an equivalent martingale measure as well as to the secondary probability measures required in the change-of-numéraire technique; Section 2.4 contains the derivation of the exchange option pricing integro-partial differential equation; Section 2.5 discusses a representation of the European exchange option price based on the change-of-numéraire technique (Section 2.5.1) and the Fourier transform method of Caldana and Fusai (2013) (Section 2.5.2); Section 2.6 shows the derivation of the decomposition of the American exchange option price; Section 2.7 provides a summary and some concluding remarks.

2.2 A Stochastic Volatility Jump-Diffusion Model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. The probability measure $\mathbb{P}$ is interpreted to be the market probability measure. Let S_1 and S_2 denote the prices of two stocks at time t, modelled as adapted *cadlág* stochastic processes (adapted stochastic processes that are right-continuous with left limits). Denote by $S_i(t-) = \lim_{u \to t^-} S_i(u)$ the price of stock i $(i = 1, 2)$ immediately before time t. If a jump in stock price occurs at time t, then $S_i(t-)$ represents the pre-jump price of stock i. Assume that stock i pays a (constant) continuously compounded dividend yield q_i. We assume that the evolution of the price of stock i is given by a stochastic volatility jump-diffusion (SVJD) model specified as

$$\mathrm{d}S_i(t) = \mu_i S_i(t-)\,\mathrm{d}t + \sqrt{v_i(t)} S_i(t-)\,\mathrm{d}W_i(t) + S_i(t-)\int_{\mathbb{R}} (e^y - 1)\left(p_i(\mathrm{d}y, \mathrm{d}t) - \lambda_i m_{\mathbb{P}}^{(i)}(\mathrm{d}y)\,\mathrm{d}t\right), \tag{2.1}$$

$$\mathrm{d}v_i(t) = \xi_i(\eta_i - v_i(t))\,\mathrm{d}t + \sigma_i\sqrt{v_i(t)}\,\mathrm{d}Z_i(t), \tag{2.2}$$

where μ_i is the instantaneous return on asset i per unit time, $v_i(t)$ is the instantaneous variance per unit time, ξ_i is the rate of mean reversion of v_v, η_i is the long-run mean for v_i, σ_i is the

instantaneous volatility of v_i, and $\{W_i(t)\}$ and $\{Z_i(t)\}$ are standard Wiener processes under $\mathbb{P}$. It is assumed that μ_i, λ_i, ξ_i, η_i, and σ_i are positive constants. We assume that the initial values $S_i(0)$ and $v_i(0)$ are known positive constants.

Furthermore, for $i = 1, 2$, the counting measure $p_i(\mathrm{d}y, \mathrm{d}t)$, defined on the mark space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with $\mathcal{B}(\mathbb{R})$ the Borel σ-algebra on $\mathbb{R}$, is associated to a marked Poisson process $(T_{i,n}, Y_{i,n})$, where the marks $Y_{i,1}, Y_{i,2}, \ldots$ are i.i.d. real-valued random variables with a non-atomic $\mathbb{P}$-density $m_{\mathbb{P}}^{(i)}(\mathrm{d}y)$ and $\{N_i(t)\}$, given by

$$N_i(t) = \sum_{n=1}^{\infty} \mathbf{1}(T_n \leq t)\mathbf{1}(Y_n \in \mathbb{R}),$$

is a Poisson process with (constant) intensity λ_i under $\mathbb{P}$. It is thus said that $p_i(\mathrm{d}y, \mathrm{d}t)$ has $\mathbb{P}$-local characteristics $(\lambda_i, m_{\mathbb{P}}^{(i)}(\mathrm{d}y))$ (Brémaud, 1981; Runggaldier, 2003). The random variable $Y_{i,n}$ is interpreted to be the size of the jump in the log-return of stock i occurring at time $T_{i,n}$. The measure $\upsilon_i(\mathrm{d}y) = \lambda_i m_{\mathbb{P}}^{(i)}(\mathrm{d}y)$ is also referred to as the Lévy measure corresponding to $p_i(\mathrm{d}y, \mathrm{d}t)$, which we assume satisfies the condition

$$\int_{\mathbb{R}_0} \min\{1, y^2\}\upsilon_i(\mathrm{d}y) = \int_{\mathbb{R}_0} \min\{1, y^2\}\lambda_i m_{\mathbb{P}}^{(i)}(\mathrm{d}y) < \infty,$$

where $\mathbb{R}_0 = \mathbb{R}\backslash\{0\}$.

We assume that the marks $Y_{i,n}$ and the Poisson process $\{N_{i,t}\}$ are independent of each other and independent of the Wiener processes defined above. We further assume that the marked Poisson processes of the two stocks are independent of each other. We equip the probability space with the natural filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ generated by W_1, W_2, Z_1, Z_2, $p_1(\mathrm{d}y, \mathrm{d}t)$, and $p_2(\mathrm{d}y, \mathrm{d}t)$ and we further assume that $\{\mathcal{F}_t\}_{t\in[0,T]}$ satisfies the *usual conditions* (see for example, Karatzas and Shreve, 1988, Definition 2.25), i.e.

$$\mathcal{F}_t = \sigma\left(W_1(s), W_2(s), Z_1(s), Z_2(s), p_1(A, (0, s]), p_2(A, (0, s]), B : 0 \leq s \leq t, A \in \mathcal{B}(\mathbb{R}), B \in \mathcal{N}\right)$$

where $\mathcal{N}$ is the collection of all $\mathbb{P}$-null sets in $\mathcal{F}$.

This model is a two-asset extension of the stock price dynamics in Cheang et al. (2013), which combines the jump-diffusion model of Merton (1976) and the square-root volatility process of Heston (1993) (as what was done by Bates (1996)). Likewise, it is a stochastic volatility version of the model introduced in Cheang and Chiarella (2011) and Cheang and Lian (2015) for exchange options.

We assume that the correlation structure of the model is summarized as follows:

	$W_1(t)$	$W_2(t)$	$Z_1(t)$	$Z_2(t)$
$W_1(t)$	1	ρ_w	ρ_{wz_1}	0
$W_2(t)$	ρ_w	1	0	ρ_{wz_2}
$Z_1(t)$	ρ_{wz_1}	0	1	ρ_z
$Z_2(t)$	0	ρ_{wz_2}	ρ_z	1

Let $\mathbf{\Sigma}$ denote the correlation matrix of the Wiener processes as given above,

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \rho_w & \rho_{wz_1} & 0 \\ \rho_w & 1 & 0 & \rho_{wz_2} \\ \rho_{wz_1} & 0 & 1 & \rho_z \\ 0 & \rho_{wz_2} & \rho_z & 1 \end{bmatrix}. \tag{2.3}$$

To avoid trivial cases, we assume that neither ρ_w nor ρ_z are equal to ± 1. In the succeeding analysis, the correlation coefficients ρ_w and ρ_z may be set to zero as simplifying assumptions.

Remark 2.1. *We assume that* $\mathbf{\Sigma}$ *permits a Cholesky decomposition* $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^\top$*, where* $\mathbf{L}$ *is a lower triangular matrix. If*

$$\mathbf{W} = (W_1, W_2, Z_1, Z_2)^\top$$

is a random vector containing the Wiener processes defined above, we can then write $\mathbf{W}(t) = \mathbf{L}\mathbf{B}(t)$ *where* $\mathbf{B}$ *is a standard 4-dimensional* $\mathbb{P}$*-Brownian motion (see for example, Cohen and Elliott, 2015, Definition 5.5.1.). The associated increments can also be written as* $\mathrm{d}\mathbf{W}_t = \mathbf{L}\,\mathrm{d}\mathbf{B}_t$*. By assuming that the correlation coefficients are not equal to 1 in magnitude, the diagonal entries of the lower triangular matrix* $\mathbf{L}$ *are nonzero, and so* $\mathbf{L}^{-1}$ *exists.* △

In our analysis, we assume that all model parameters have been estimated. See the discussion in Chapter 6 and Appendix E for an overview of parameter estimation methods that can be implemented to our model.

We make the following assumptions on the correlation coefficients and the coefficients of the volatility processes.

Assumption 2.2. The coefficients ξ_i, η_i, and σ_i are positive and satisfy $2\xi_i\eta_i \geq \sigma_i^2$, $i = 1, 2$ and the correlation coefficients ρ_{wz_1} and ρ_{wz_2}

$$-1 < \rho_{wz_i} < \min\left\{\frac{\xi_i}{\sigma_i}, 1\right\}, \qquad i = 1, 2.$$

The condition $2\xi_i\eta_i \geq \sigma_i^2$, also known as the *Feller condition*, is required to ensure that the volatility processes do not hit zero or explode in finite time under the objective market measure $\mathbb{P}$ (Andersen and Piterbarg, 2007; Bertini and Passalacqua, 2008; Cheang et al., 2011). This also

implies that v_1 and v_2 have continuous trajectories for $t \in [0, T]$. The condition on ρ_{wz_i} ensures that the volatility processes do not explode in finite time under the equivalent martingale measure and other equivalent probability measures considered in the change-of-numéraire procedure. This will be justified in the succeeding sections.

Denote by κ_i the expected jump-size increment under $\mathbb{P}$, which is given by

$$\kappa_i = \int_{\mathbb{R}} (e^y - 1) m_{\mathbb{P}}^{(i)}(\mathrm{d}y) = \mathbb{E}_{\mathbb{P}}[e^{Y_i} - 1].$$

Then equation (2.1) may be rewritten as

$$\mathrm{d}S_i(t) = (\mu_i - \lambda_i \kappa_i) S_i(t-)\,\mathrm{d}t + \sqrt{v_i(t)} S_i(t-)\,\mathrm{d}W_i(t) + S_i(t) \int_{\mathbb{R}} (e^y - 1) p_i(\mathrm{d}y, \mathrm{d}t). \tag{2.4}$$

We assume that the moment generating function of jump sizes, denoted by $M_{\mathbb{P},Y_i}(u) = \mathbb{E}_{\mathbb{P}}[e^{uY_i}]$, exists but the distribution is not specified.

By Itô's Lemma for jump-diffusion processes (Theorem A.1), the log-price $X_i = \ln S_i$ satisfies the stochastic differential equation

$$\mathrm{d}X_i(t) = \left(\mu_i - \lambda_i \kappa_i - \frac{1}{2} v_i(t)\right) \mathrm{d}t + \sqrt{v_i(t)}\,\mathrm{d}W_i(t) + \int_{\mathbb{R}} y p_i(\mathrm{d}y, \mathrm{d}t), \tag{2.5}$$

for $i = 1, 2$. This implies that equation (2.1) admits a solution of the form

$$S_i(t) = S_i(0) \exp\left\{ (\mu_i - \lambda_i \kappa_i) t - \frac{1}{2} \int_0^t v_i(s)\,\mathrm{d}s + \int_0^t \sqrt{v_i(s)}\,\mathrm{d}W_i(s) + \sum_{n=1}^{N_i(t)} Y_{i,n} \right\}, \tag{2.6}$$

for $i = 1, 2$. Let $\tilde{S}_i(t) = e^{-(r-q)t} S_i(t)$. Then $\{\tilde{S}_i(t)\}$ is the discounted yield process for stock i and its $\mathbb{P}$-dynamics is given by the SDE

$$\begin{aligned} \mathrm{d}\tilde{S}_i(t) = {} & (\mu_i + q_i - r - \lambda_i \kappa_i) \tilde{S}_i(t-)\,\mathrm{d}t + \sqrt{v_i(t)} \tilde{S}_i(t-)\,\mathrm{d}W_i(t) \\ & + \tilde{S}_i(t-) \int_{\mathbb{R}} (e^y - 1) p_i(\mathrm{d}y, \mathrm{d}t). \end{aligned}$$

We investigate later the dynamics of the discounted yield process under the equivalent risk-neutral measure in line with the option pricing problem.

With respect to the expressions obtained for the log-price and stock price dynamics, the non-explosion of the volatility process ensures that the integrals $\int_0^t v_i(s)\,\mathrm{d}s$ and $\int_0^t \sqrt{v_i(s)}\,\mathrm{d}W_i(s)$ are

properly defined.[1] Furthermore, it also guarantees that the process $\{E_i(t)\}$, where

$$E_i(t) = \exp\left\{-\frac{1}{2}\int_0^t v_i(s)\,\mathrm{d}s + \int_0^t \sqrt{v_i(s)}\,\mathrm{d}W_i(s) - \lambda_i\kappa_i t + \sum_{n=1}^{N_i(t)} Y_{i,n}\right\}, \qquad i = 1, 2, \tag{2.7}$$

is a $\mathbb{P}$-martingale. To establish this martingale property, we first note that

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left[\exp\left\{-\lambda_i\kappa_i t + \sum_{n=1}^{N_i(t)} Y_{i,n}\right\}\right] &= e^{-\lambda_i\kappa_i t}\mathbb{E}_{\mathbb{P}}\left[\exp\left\{\sum_{n=1}^{N_i(t)} Y_{i,n}\right\}\right] \\ &= e^{-\lambda_i\kappa_i t}\exp\left\{\lambda_i t\left(\mathbb{E}_{\mathbb{P}}(e^{Y_i}) - 1\right)\right\} = e^{-\lambda_i\kappa_i t}e^{\lambda_i\kappa_i t} = 1.\end{aligned}$$

Next we examine if the process

$$E_i^c(t) = \exp\left\{-\frac{1}{2}\int_0^t v_i(s)\,\mathrm{d}s + \int_0^t \sqrt{v_i(s)}\,\mathrm{d}W_i(s)\right\}$$

is a $\mathbb{P}$-martingale. From Kuo (2006, Theorem 8.7.3.) and Wong and Heyde (2004, Theorem 1), $E_i^c(t)$ is a martingale if and only if $\mathbb{E}_{\mathbb{P}}[E_i^c(t)] = 1$ for all $t \in [0, T]$. Kuo (2006) notes that this condition is generally difficult to verify, but the stronger Novikov's condition, $\mathbb{E}_{\mathbb{P}}[\exp\{\frac{1}{2}\int_0^T v_i(t)\,\mathrm{d}t\}] < \infty$, may be used instead.[2] Following Andersen and Piterbarg (2007, Proposition 2.1), the process $\{v_i(t)\}$ satisfying the conditions in Assumption 2.2 does not hit 0 and does not explode in in finite time (see also Lewis, 2000, Chapter 9). Hence, $0 < v_i(t) < \infty$ almost surely for all $t \in [0, T]$, and so $\int_0^T v_i(t)\,\mathrm{d}t < \infty$ almost surely. Thus, Novikov's condition is satisfied, allowing us to conclude that $\{E_i^c(t)\}$ is a martingale and that $\mathbb{E}_{\mathbb{P}}[E_i^c(t)] = 1$. See also Appendix A.2.

2.3 A Change of Measure Mechanism

In order to price of the exchange option, we must find a suitable Radon-Nikodým derivative that can be used to define probability measures $\mathbb{Q}$ on $(\Omega, \mathcal{F})$, equivalent to $\mathbb{P}$, under which the discounted stock yield processes $\{S_1(t)e^{-(r-q_1)t}\}_{t\in[0,T]}$ and $\{S_2(t)e^{-(r-q_2)t}\}_{t\in[0,T]}$ are martingales. Such a probability measure $\mathbb{Q}$ shall be referred to as an *equivalent martingale measure (EMM)*.

As in the previous section, we let $\mathbf{W}(t) = (W_1(t), W_2(t), Z_1(t), Z_2(t))^\top$ be a vector of $\mathbb{P}$-Wiener processes with correlation matrix $\mathbf{\Sigma}$. Let $Q_1(t) = \sum_{n=1}^{N_1(t)} Y_{1,n}$ and $Q_2(t) = \sum_{n=1}^{N_2(t)} Y_{2,n}$ be compound Poisson processes generated by the Poisson random measures $p_1(\mathrm{d}y, \mathrm{d}t)$ and $p_2(\mathrm{d}y, \mathrm{d}t)$.

We state the following Radon-Nikodým derivative and Girsanov Theorem to facilitate a change of measure from $\mathbb{P}$ to $\mathbb{Q}$, inducing a drift in the components of $\mathbf{W}_t$ under $\mathbb{Q}$, a change in intensity

[1] See Shreve (2004) and Kuo (2006) for a discussion of properties that must be satisfied by integrands for stochastic integrals to be well-defined.

[2] As Novikov's condition is stronger as pointed out by Kuo (2006), weaker alternative conditions for the martingale property are discussed in Wong and Heyde (2004).

of the Poisson processes, and a change in density of the jump size variables.

Proposition 2.3. *Let* $\mathbf{B}$ *be a standard 4-dimensional Brownian motion which can be transformed to* $\mathbf{W}$ *via* $\mathbf{W} = \mathbf{LB}$*, where* $\mathbf{L}$ *is a lower triangular matrix such that* $\boldsymbol{\Sigma} = \mathbf{LL}^\top$ *(see Remark 2.1). Suppose* $\boldsymbol{\theta} : [0,T] \times \Omega \to \mathbb{R}^4$ *is a predictable process and let* ν_1*,* ν_2*,* γ_1*, and* γ_2 *be constants such that the process* $L(t)$ *be given by the equation*

$$\begin{aligned} L(t) = \exp\left\{ -\int_0^t \left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s)\right)^\top \mathrm{d}\mathbf{W}(s) - \frac{1}{2}\int_0^t \boldsymbol{\theta}^\top(s)\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s)\,\mathrm{d}s \right\} \\ \times \exp\left\{ \sum_{n=1}^{N_1(t)} (\gamma_1 Y_{1,n} + \nu_1) - \lambda_1 t\left(e^{\nu_1}\mathbb{E}_{\mathbb{P}}[e^{\gamma_1 Y_1}] - 1\right)\right\} \\ \times \exp\left\{ \sum_{n=1}^{N_{2,t}} (\gamma_2 Y_{2,n} + \nu_2) - \lambda_2 t\left(e^{\nu_2}\mathbb{E}_{\mathbb{P}}[e^{\gamma_2 Y_2}] - 1\right)\right\} \end{aligned} \tag{2.8}$$

is well-defined for all $t \in [0,T]$*. Suppose further that* $\mathbb{E}_{\mathbb{P}}[L(T)] = 1$*.*

Then under the probability measure $\mathbb{Q}$ *on* $(\Omega, \mathcal{F})$ *equivalent to* $\mathbb{P}$ *defined by* $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T} = L(T)$*, the following hold:*

1. $\mathbf{B}^{\mathbb{Q}}(t) = \int_0^t \mathbf{L}^{-1}\boldsymbol{\theta}(s)\,\mathrm{d}s + \mathbf{B}(t)$ *is a standard 4-dimensional* $\mathbb{Q}$*-Wiener process;*
2. *the compound Poisson process* $Q_{i,t} = \sum_{n=1}^{N_{1,t}} Y_{i,n}$ *(and consequently the random measure* $p_i(\mathrm{d}y, \mathrm{d}t)$*) has a new intensity rate* $\tilde{\lambda}_i = \lambda_i e^{\nu_i}\mathbb{E}_{\mathbb{P}}[e^{\gamma_i Y_i}]$*,* $i = 1, 2$ *under* $\mathbb{Q}$*; and*
3. *the moment generating function of the jump size random variable* Y_i *under* $\mathbb{Q}$ *is given by*

$$M_{\mathbb{Q},Y_i}(u) = \frac{M_{\mathbb{P},Y_i}(u+\gamma_i)}{M_{\mathbb{P},Y_i}(\gamma_i)}, \qquad i = 1, 2. \tag{2.9}$$

Since it is assumed that the jump components in the model and in the Radon-Nikodým derivative are independent of the Wiener components, the new distributions of these components can be obtained separately from the corresponding components in equation (2.8) (Cheang and Teh, 2014). That is, the first factor of equation (2.8) facilitates the change of measure in the Wiener processes, whereas the last two factors handle the change in distribution of the jump components in the transition to a new probability measure. As such, $L(t)$ as given in equation (2.8) may be written as

$$L(t) = L^D(t) \times L^{J_1}(t) \times L^{J_2}(t),$$

where $L^D(t)$ facilitates the change of drift in in the diffusion part of the model (i.e. the Wiener processes) and $L^{J_1}(t)$ and $L^{J_2}(t)$ handle the changes in the distribution of jumps in stock 1 and stock 2, respectively.

Proposition 2.3 is a specialized form of the Girsanov Theorem for multidimensional Lévy processes (see for example, Øksendal and Sulem, 2019, Theorem 1.33). We state and prove the Girsanov Theorem for Lévy processes and discuss how exactly Proposition 2.3 relates to the Girsanov Theorem in Appendix A.3. In the following, we discuss how the intensity rate and the moment generating function corresponding to $p_i(\mathrm{d}y, \mathrm{d}t)$ under $\mathbb{Q}$ are obtained.

To determine the new distribution of the compound Poisson processes under $\mathbb{Q}$, we consider the joint moment generating function of $\mathbf{W}$, Q_1, and Q_2 under $\mathbb{Q}$. Let $\mathbf{u}$ be a constant vector (with length 4) and let u_1 and u_2 be constants. We then have

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[e^{\mathbf{u}^\top \mathbf{W}(t)+u_1 Q_1(t)+u_2 Q_2(t)}\right] &= \mathbb{E}_{\mathbb{P}}\left[L(t)\exp\left\{\mathbf{u}^\top \mathbf{W}(t)+u_1 Q_1(t)+u_2 Q_2(t)\right\}\right] \\ &= \mathbb{E}_{\mathbb{P}}\left[L^D(t)e^{\mathbf{u}^\top \mathbf{W}(t)}\right]\mathbb{E}_{\mathbb{P}}\left[L^{J_1}(t)\exp\left\{u_1\sum_{n=1}^{N_1(t)}Y_{1,n}\right\}\right] \\ &\quad \times \mathbb{E}_{\mathbb{P}}\left[L^{J_2}(t)\exp\left\{u_2\sum_{n=1}^{N_2(t)}Y_{2,n}\right\}\right],\end{aligned}$$

where the last equality follows from the mutual independence of the jump parts and the diffusion part and the definition of $Q_1(t)$ and $Q_2(t)$. We now focus on the last two factors of the last line above. Observe further that the second and third factors are the moment generating functions of $Q_1(t)$ and $Q_2(t)$ under $\mathbb{Q}$, respectively. Expanding the second factor, we have

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}}\left[L^{J_1}(t)\exp\left\{u_1\sum_{n=1}^{N_1(t)}Y_{1,n}\right\}\right] \\ &\quad= \mathbb{E}_{\mathbb{P}}\left[\exp\left\{\sum_{n=1}^{N_1(t)}(\gamma_1 Y_{1,n}+\nu_1)-\lambda_1 t\left(e^{\nu_1}\mathbb{E}_{\mathbb{P}}(e^{\gamma_1 Y_1})-1\right)\right\}\times\exp\left\{u_1\sum_{n=1}^{N_1(t)}Y_{1,n}\right\}\right] \\ &\quad= \mathbb{E}_{\mathbb{P}}\left[\exp\left\{\sum_{n=1}^{N_1(t)}\left((\gamma_1+u_1)Y_{1,n}+\nu_1\right)\right\}\right]\times\exp\left\{-\lambda_1 t\left(e^{\nu_1}\mathbb{E}_{\mathbb{P}}(e^{\gamma_1 Y_1})-1\right)\right\}\end{aligned}$$

Observe that the expectation is the moment generating function (evaluated at 1) of some compound Poisson process $\bar{Q}_1(t)=\sum_{n=1}^{N_1(t)}C_{1,n}$, where $C_{1,n}=(\gamma_1+u_1)Y_{1,n}+\nu_1$. It follows that

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}}\left[L^{J_1}(t)\exp\left\{u_1\sum_{n=1}^{N_1(t)}Y_{1,n}\right\}\right] \\ &\quad= \exp\left\{\lambda_1 t\left(\mathbb{E}_{\mathbb{P}}\left[e^{(\gamma_1+u_1)Y_1+\nu_1}\right]-1\right)\right\}\times\exp\left\{-\lambda_1 t\left(e^{\nu_1}\mathbb{E}_{\mathbb{P}}(e^{\gamma_1 Y_1})-1\right)\right\} \\ &\quad= \exp\left\{\lambda_1 t e^{\nu_1}\mathbb{E}_{\mathbb{P}}\left[e^{(\gamma_1+u_1)Y_1}\right]-\lambda_1 t-\lambda_1 t e^{\nu_1}\mathbb{E}_{\mathbb{P}}\left[e^{\gamma_1 Y_1}\right]+\lambda_1 t\right\}\end{aligned}$$

$$= \exp\left\{\lambda_1 e^{\nu_1}\mathbb{E}_{\mathbb{P}}[e^{\gamma_1 Y_1}]t\left(\frac{\mathbb{E}_{\mathbb{P}}[e^{(\gamma_1+u_1)Y_1}]}{\mathbb{E}_{\mathbb{P}}[e^{\gamma_1 Y_1}]}-1\right)\right\}.$$

Now suppose the intensity of $Q_1(t)$ is $\tilde{\lambda}_1$. Then by definition, the moment generating function of $Q_1(t)$ under $\mathbb{Q}$ is

$$\mathbb{E}_{\mathbb{Q}}\left[e^{u_1 Q_1(t)}\right] = \exp\left\{\tilde{\lambda}_1 t\left(\mathbb{E}_{\mathbb{Q}}\left[e^{u_1 Y_1}\right]-1\right)\right\}.$$

Comparing the right hand sides of this last equation and the last expression for $\mathbb{E}_{\mathbb{P}}[L^{J_1}(t)e^{u_1 Q_1(t)}]$, we find that the new intensity of the Poisson process and the moment generating function of Y_1 under $\mathbb{Q}$ are given by

$$\tilde{\lambda}_1 = \lambda_1 e^{\nu_1}\mathbb{E}_{\mathbb{P}}[e^{\gamma_1 Y_1}], \qquad \mathbb{E}_{\mathbb{Q}}[e^{u_1 Y_1}] = \frac{\mathbb{E}_{\mathbb{P}}[e^{(\gamma_1+u_1)Y_1}]}{\mathbb{E}_{\mathbb{P}}[e^{\gamma_1 Y_1}]},$$

respectively. The second expression is equivalent to $M_{\mathbb{Q},Y_1}(u) = M_{\mathbb{P},Y_1}(\gamma_1+u)/M_{\mathbb{P},Y_1}(\gamma_1)$, which is equation (2.9). The same calculations can be performed on $\mathbb{E}_{\mathbb{P}}[L^{J_1}(t)e^{u_2 Q_2(t)}]$ to determine the new intensity of $N_2(t)$ and the new moment generating function of Y_2 under $\mathbb{Q}$. The proposition also implies that the $\mathbb{Q}$-local characteristics of the counting measure $p_i(\mathrm{d}y,\mathrm{d}t)$ are given by $(\tilde{\lambda}_i, m^{(i)}_{\mathbb{Q}}(\mathrm{d}y))$, where

$$m^{(i)}_{\mathbb{Q}}(\mathrm{d}y) = \frac{e^{\gamma_i y}}{\mathbb{E}_{\mathbb{P}}[e^{\gamma_i Y_i}]}m^{(i)}_{\mathbb{P}}(\mathrm{d}y) \tag{2.10}$$

(see also Appendix A.3). This also shows that Proposition 2.3 induces an Esscher-type transform (Gerber and Shiu, 1994) on the distribution of jump sizes as we move from $\mathbb{P}$ to $\mathbb{Q}$.

Remark 2.4. *This remark addresses the martingale property of the Radon-Nikodým derivative process $\{L(t)\}$ in equation (2.8). As the three factors are independent, we have*

$$\mathbb{E}_{\mathbb{P}}[L(t)] = \mathbb{E}_{\mathbb{P}}\left[L^D(t)\right]\times\mathbb{E}_{\mathbb{P}}\left[L^{J_1}(t)\right]\times\mathbb{E}_{\mathbb{P}}\left[L^{J_2}(t)\right].$$

Note that $L^D(t)$ has a form similar to the stochastic exponential process discussed by Wong and Heyde (2004); in fact, it is the stochastic exponential process generated by $\mathbf{X}(t) = -\mathbf{L}^{-1}\boldsymbol{\theta}(t)$ (see Definition A.5). To see this, we note that for some 4-dimensional $\mathbb{P}$-Brownian motion $\mathbf{B}(t)$ with increment $\mathrm{d}\mathbf{B}(t)$ (see Remark 2.1), we have

$$\begin{aligned}
\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s)\right)^{\top}\mathrm{d}\mathbf{W}(s) &= \left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s)\right)^{\top}\mathbf{L}\,\mathrm{d}\mathbf{B}(s) = \boldsymbol{\theta}^{\top}(s)\left(\boldsymbol{\Sigma}^{-1}\right)^{\top}\mathbf{L}\,\mathrm{d}\mathbf{B}(s)\\
&= \boldsymbol{\theta}^{\top}(s)\left[\left(\mathbf{L}\mathbf{L}^{\top}\right)^{-1}\right]^{\top}\mathbf{L}\,\mathrm{d}\mathbf{B}(s) = \boldsymbol{\theta}^{\top}(s)\left[\left(\mathbf{L}^{\top}\right)^{-1}\mathbf{L}^{-1}\right]^{\top}\mathbf{L}\,\mathrm{d}\mathbf{B}(s)\\
&= \boldsymbol{\theta}^{\top}(s)\left(\mathbf{L}^{-1}\right)^{\top}\left[\left(\mathbf{L}^{\top}\right)^{-1}\right]^{\top}\mathbf{L}\,\mathrm{d}\mathbf{B}(s) = \boldsymbol{\theta}^{\top}(s)\left(\mathbf{L}^{-1}\right)^{\top}\mathbf{L}^{-1}\mathbf{L}\,\mathrm{d}\mathbf{B}(s)\\
&= \boldsymbol{\theta}^{\top}(s)(\mathbf{L}^{-1})^{\top}\,\mathrm{d}\mathbf{B}(s) = \left(\mathbf{L}^{-1}\boldsymbol{\theta}(s)\right)^{\top}\mathrm{d}\mathbf{B}(s).
\end{aligned}$$

Furthermore, we observe that

$$\begin{aligned}\left\|-\mathbf{L}^{-1}\boldsymbol{\theta}(s)\right\|^2 &= \left(-\mathbf{L}^{-1}\boldsymbol{\theta}(s)\right)^\top \left(-\mathbf{L}^{-1}\boldsymbol{\theta}(s)\right) = \boldsymbol{\theta}^\top(s)\left(\mathbf{L}^{-1}\right)^\top \mathbf{L}^{-1}\boldsymbol{\theta}(s)\\ &= \boldsymbol{\theta}^\top(s)\left(\mathbf{L}\mathbf{L}^\top\right)^{-1}\boldsymbol{\theta}(s) = \boldsymbol{\theta}^\top(s)\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s).\end{aligned}$$

Therefore, choosing $\mathbf{X}(t) = -\mathbf{L}^{-1}\boldsymbol{\theta}(t)$ *in Definition A.5 will generate the exponential process* $L^D(t)$*. After determining the components of* $\boldsymbol{\theta}(t)$*, we shall discuss some properties of* $L^D(t)$ *and the calculation of* $\mathbb{E}_{\mathbb{P}}[L^D(t)]$*. However, it can easily be seen that*

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[L^{J_1}(t)] &= \mathbb{E}_{\mathbb{P}}\left[\exp\left\{\sum_{n=1}^{N_1(t)}(\gamma_1 Y_{1,n}+\nu_1)\right\}\right] \times \exp\left\{-\lambda_1 t\left(e^{\nu_1}\mathbb{E}_{\mathbb{P}}\left[e^{\gamma_1 Y_1}\right]-1\right)\right\}\\ &= \exp\left\{\lambda_1 t\left(\mathbb{E}_{\mathbb{P}}\left[e^{\gamma_1 Y_1+\nu_1}\right]-1\right)\right\} \times \exp\left\{-\lambda_1 t\left(e^{\nu_1}\mathbb{E}_{\mathbb{P}}\left[e^{\gamma_1 Y_1}\right]-1\right)\right\} = 1,\end{aligned}$$

where, in the second line, we use the moment generating function the compound Poisson process $Q_1(t) = \sum_{n=1}^{N_1(t)}(\gamma_1 Y_{1,n}+\nu_1)$*. A similar calculation can be done to show that* $\mathbb{E}_{\mathbb{P}}[L^{J_2}(t)] = 1$*. Therefore, whether or not the process* L *defined by equation (2.8) has the martingale property on* $[0,T]$ *depends on the whether the exponential process* L^D *generated by* $-\mathbf{L}^{-1}\boldsymbol{\theta}(t)$ *is a martingale. In particular, in view of the discussion in Appendix A.2, it is sufficient for* $\boldsymbol{\theta}$ *to have components which are non-explosive on* $[0,T]$*. This shall be addressed in Remark 2.5 below.* △

Instead of using the standard four-dimensional $\mathbb{Q}$ Brownian motion $\mathbf{B}^{\mathbb{Q}}$, it is more convenient notationally to express the dynamics of the asset prices and variances in terms of correlated Wiener processes. One of the consequences of Proposition 2.3, $\mathbf{B}^{\mathbb{Q}}(t) = \int_0^t \mathbf{L}^{-1}\boldsymbol{\theta}(s)\,\mathrm{d}s + \mathbf{B}(t)$ is a standard 4-dimensional $\mathbb{Q}$-Wiener process. We can pre-multiply both sides of the equation by $\mathbf{L}$ to get

$$\mathbf{L}\mathbf{B}^{\mathbb{Q}}(t) = \mathbf{L}\mathbf{L}^{-1}\int_0^t \boldsymbol{\theta}(s)\,\mathrm{d}s + \mathbf{L}\mathbf{B}(t) = \int_0^t \boldsymbol{\theta}(s)\,\mathrm{d}s + \mathbf{W}(t).$$

Taking $\tilde{\mathbf{W}} = \mathbf{L}\mathbf{B}^{\mathbb{Q}}(t)$, we see from the equation above that the components of $\tilde{\mathbf{W}}$ are Wiener processes that inherit the same correlation matrix $\boldsymbol{\Sigma}$ as that of $\mathbf{W}$. Furthermore, the components of $\tilde{\mathbf{W}}$ can be written as

$$\begin{aligned}\mathrm{d}\tilde{W}_i(t) &= \psi_i(t)\,\mathrm{d}t + \mathrm{d}W_i(t), \qquad i=1,2\\ \mathrm{d}\tilde{Z}_i(t) &= \zeta_i(t)\,\mathrm{d}t + \mathrm{d}Z_i(t), \qquad i=1,2.\end{aligned} \tag{2.11}$$

Throughout the remainder of this chapter, any change of measure will be facilitated by a Radon-Nikodým derivative of the form given in equation (2.8) and the properties of the new probability measure will be reflected in the choice of parameters $\boldsymbol{\theta}(t)$, γ_1, γ_2, ν_1, and ν_2. We assume that the parameters γ_1, γ_2, ν_1, and ν_2 are constant so that the Poisson processes N_1 and N_2 remain

homogenous and that the jump sizes remain identically distributed under $\mathbb{Q}$.

The parameters γ_1, γ_2, ν_1, and ν_2 that appear in the Radon-Nikodým derivative process influence the properties the jumps may have under the equivalent probability measure $\mathbb{Q}$ (Cheang and Chiarella, 2011). For example:

1. one can choose not to price the jump risk of the model, following the original analysis of Merton (1976). This happens when we set $\gamma_1 = \gamma_2 = \nu_1 = \nu_2 = 0$. In doing this, there will be no change in the local characteristics of the counting measures in the transition from $\mathbb{P}$ to $\mathbb{Q}$.

2. if $\gamma_i = 0$ $(i = 1, 2)$ but $\nu_i \neq 0$ $(i = 1, 2)$, then the intensity of the Poisson processes change in the transition to $\mathbb{Q}$ without changing the distribution of jump-sizes. Likewise, if $\nu_i = -\ln \mathbb{E}_{\mathbb{P}}[e^{\gamma_i Y_i}]$ and $\gamma_i \neq 0$ $(i = 1, 2)$, then only the jump size distribution changes.

Our goal is to be able to characterize an equivalent martingale measure $\mathbb{Q}$ by imposing additional constraints on the parameters of the Radon-Nikodým process $L(t)$. In general, because the market in which asset prices exhibit jumps and stochastic volatility is incomplete in the Harrison and Pliska (1981) sense, there exist many equivalent martingale measures that are consistent with the no-arbitrage condition. Denote by $\mathcal{M}_E$ the collection of martingale measures on $(\Omega, \mathcal{F})$ equivalent to $\mathbb{P}$. Therefore, one must select, with respect to some other principle, the "most suitable" equivalent martingale measure, referred to now as the pricing measure, from those in $\mathcal{M}_E$ under which option prices are computed. One such example is the *minimum entropy martingale measure (MEMM)*, which is the measure $\mathbb{Q}^E$ that minimizes the relative entropy $\mathcal{D}(\mathbb{Q}|\mathbb{P})$ with respect to $\mathbb{P}$, i.e.

$$\mathcal{D}(\mathbb{Q}^E|\mathbb{P}) = \min_{\mathbb{Q}\in\mathcal{M}_E} \mathcal{D}(\mathbb{Q}|\mathbb{P}), \tag{2.12}$$

where the relative entropy is given by

$$\mathcal{D}(\mathbb{Q}|\mathbb{P}) = \begin{cases} \mathbb{E}_{\mathbb{P}}\left[\dfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \ln\left(\dfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)\right], & \text{if } \mathbb{Q} \text{ is absolutely continuous w.r.t. } \mathbb{P} \\ +\infty, & \text{otherwise} \end{cases}.$$

The construction of the MEMM corresponding to our SVJD model is discussed in detail in Appendix A.4.

The parameters of the Radon-Nikodým derivative may also be selected using a model calibration procedure that uses available option price data. For an overview of model calibration and how it can be implemented to our model, see Appendix E.

In our analysis, we assume that γ_1, γ_2, ν_1, and ν_2 have already been chosen through any of the possible methods. However, the processes $\{\psi_i(t)\}$ and $\{\zeta_i(t)\}$ must be determined such that

the discounted asset yield processes are $\mathbb{Q}$-martingales. How to choose these processes will now be discussed.

We now investigate the $\mathbb{Q}$-dynamics of the discounted yield processes. As in Runggaldier (2003), define the $\mathbb{Q}$-compensated counting measure $q(\mathrm{d}y_i, \mathrm{d}t)$ as

$$q_i(\mathrm{d}y, \mathrm{d}t) = p_i(\mathrm{d}y, \mathrm{d}t) - \tilde{\lambda}_i m_{\mathbb{Q}}^{(i)}(\mathrm{d}y)\,\mathrm{d}t. \tag{2.13}$$

As was shown in the prior section, the $\mathbb{P}$-dynamics of the discounted yield process $\{\tilde{S}_i(t)\}$ is given by the SDE

$$\mathrm{d}\tilde{S}_i(t) = \tilde{S}_i(t-)\left\{(\mu_i + q_i - r - \lambda_i\kappa_i)\,\mathrm{d}t + \sqrt{v_i(t)}\,\mathrm{d}W_i(t) + \int_{\mathbb{R}}(e^y - 1)p_i(\mathrm{d}y, \mathrm{d}t)\right\}.$$

Substituting the expressions for $\mathrm{d}W_i(t)$ from equation (2.11) and $q_i(\mathrm{d}y, \mathrm{d}t)$ from equation (2.13), we obtain

$$\begin{aligned}\mathrm{d}\tilde{S}_i(t) = {} & \tilde{S}_i(t-)(\mu_i + q_i - r - \lambda_i\kappa_i)\,\mathrm{d}t + \sqrt{v_i(t)}\tilde{S}_i(t-)(-\psi_i(t)\,\mathrm{d}t + \mathrm{d}\tilde{W}_i(t)) \\ & + \tilde{S}_i(t-)\int_{\mathbb{R}}(e^y - 1)\left[q_i(\mathrm{d}y, \mathrm{d}t) + \tilde{\lambda}_i m_{\mathbb{Q}}^{(i)}(\mathrm{d}y)\,\mathrm{d}t\right]\end{aligned}$$

If we let $\tilde{\kappa}_i = \int_{\mathbb{R}}(e^y - 1)m_{\mathbb{Q}}^{(i)}(\mathrm{d}y) = \mathbb{E}_{\mathbb{Q}}[e^{Y_i} - 1]$ be the mean relative jump size increment under $\mathbb{Q}$, then $\mathrm{d}\tilde{S}_i(t)$ can be simplified to

$$\begin{aligned}\mathrm{d}\tilde{S}_i(t) = {} & \tilde{S}_i(t-)\left(\mu_i + q_i - r - \lambda_i\kappa_i + \tilde{\lambda}_i\tilde{\kappa}_i - \sqrt{v_i(t)}\psi_i(t)\right)\mathrm{d}t + \sqrt{v_i(t)}\tilde{S}_i(t-)\,\mathrm{d}\tilde{W}_i(t) \\ & + \tilde{S}_i(t-)\int_{\mathbb{R}}(e^y - 1)q_i(\mathrm{d}y, \mathrm{d}t).\end{aligned}$$

From here, we can choose the market price of $W_i(t)$ risk as

$$\psi_i(t) = \frac{\mu_i + q_i - r - \lambda_i\kappa_i + \tilde{\lambda}_i\tilde{\kappa}_i}{\sqrt{v_i(t)}},$$

which is the risk premium of the stock $\mu_i + q_i - r$ less the jump risk $\lambda_i\kappa_i - \tilde{\lambda}_i\tilde{\kappa}_i$ per unit volatility $\sqrt{v_i(t)}$. It follows that

$$\mathrm{d}\tilde{S}_i(t) = \tilde{S}_i(t-)\left\{\sqrt{v_i(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}}(e^y - 1)q_i(\mathrm{d}y, \mathrm{d}t)\right\}.$$

and so with this selection for $\psi_i(t)$, by the Martingale Representation Theorem (see for example, Runggaldier (2003, Theorem 2.3) or Øksendal and Sulem (2019, Theorem 1.35)) the discounted yield process $\{e^{-(r-q_i)t}S_i(t)\}$ is a martingale under $\mathbb{Q}$.

From the above equation, we can recover the dynamics of the stock price $S_i(t)$ under $\mathbb{Q}$. Noting

that $S_i(t) = \tilde{S}_i(t)e^{(r-q_i)t}$, stochastic integration by parts yields

$$\begin{aligned}
\mathrm{d}S_i(t) &= \tilde{S}_i(t-)(r-q_i)e^{(r-q_i)t}\,\mathrm{d}t + e^{(r-q_i)t}\,\mathrm{d}\tilde{S}_i(t) \\
&= S_i(t-)(r-q_i)\,\mathrm{d}t + e^{(r-q_i)t}\tilde{S}_i(t-)\left\{\sqrt{v_i(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}}(e^y-1)q_i(\mathrm{d}y,\mathrm{d}t)\right\} \\
&= S_i(t-)\left\{(r-q_i)\,\mathrm{d}t + \sqrt{v_i(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}}(e^y-1)q_i(\mathrm{d}y,\mathrm{d}t)\right\}.
\end{aligned}$$

Equivalently, we can write

$$\mathrm{d}S_i(t) = S_i(t-)\left\{(r-q_i-\tilde{\lambda}_i\tilde{\kappa}_i)\,\mathrm{d}t + \sqrt{v_i(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}}(e^y-1)p_i(\mathrm{d}y,\mathrm{d}t)\right\}. \tag{2.14}$$

From here, the $\mathbb{Q}$-dynamics of the log-price $X_i(t) = \ln S_i(t)$ is given by

$$\mathrm{d}X_i(t) = \left(r-q_i-\tilde{\lambda}\tilde{\kappa}_i - \frac{1}{2}v_i(t)\right)\mathrm{d}t + \sqrt{v_i(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}} y p_i(\mathrm{d}y,\mathrm{d}t).$$

It follows that the solution $S_i(t)$, for $0 < t \le T$, to equation (2.14) is given by

$$S_i(t) = S_i(0)\exp\left\{(r-q_i-\tilde{\lambda}_i\tilde{\kappa}_i)t - \frac{1}{2}\int_0^t v_i(s)\,\mathrm{d}s + \int_0^t \sqrt{v_i(s)}\,\mathrm{d}\tilde{W}_i(s) + \sum_{n=1}^{N_i(t)} Y_{i,n}\right\}. \tag{2.15}$$

With equation (2.11) the $\mathbb{Q}$-dynamics of the volatility processes are given by

$$\begin{aligned}
\mathrm{d}v_i(t) &= \xi_i(\eta_i - v_i(t))\,\mathrm{d}t + \sigma_i\sqrt{v_i(t)}(-\zeta_i(t)\,\mathrm{d}t + \mathrm{d}\tilde{Z}_i(t)) \\
&= \left[\xi_i(\eta_i - v_i(t)) - \zeta_i(t)\sigma_i\sqrt{v_i(t)}\right]\mathrm{d}t + \sigma_i\sqrt{v_i(t)}\,\mathrm{d}\tilde{Z}_i(t).
\end{aligned}$$

The quantity $\zeta_i(t)\sigma_i\sqrt{v_i(t)}$ is interpreted as the market price of volatility risk and is assumed to be independent of the asset price and proportional to the current variance level $v_i(t)$ (Heston, 1993). That is, for some constant Λ_i, we can write the market price of risk as $\zeta_i(t)\sigma_i\sqrt{v_i(t)} = \Lambda_i v_i(t)$. The constant of proportionality Λ_i must be nonnegative to keep consistent with the financial argument that investors demand a positive premium for volatility risk (Heston, 1993; Cheang et al., 2013). Thus, we can re-express the $\mathbb{Q}$-dynamics of $v_i(t)$ as

$$\begin{aligned}
\mathrm{d}v_i(t) &= \left[\xi_i\eta_i - (\xi_i+\Lambda_i)v_i(t)\right]\mathrm{d}t + \sigma_i\sqrt{v_i(t)}\,\mathrm{d}\tilde{Z}_i(t) \\
&= (\xi_i+\Lambda_i)\left[\frac{\xi_i\eta_i}{\xi_i+\Lambda_i} - v_i(t)\right]\mathrm{d}t + \sigma_i\sqrt{v_i(t)}\,\mathrm{d}\tilde{Z}_i(t).
\end{aligned} \tag{2.16}$$

The form of $\psi_i(t)$ and $\zeta_i(t)$ requires that $v_i(t)$ is nonzero and finite. These are guaranteed by

Assumption 2.2. Indeed, if $\xi_i' = \xi_i + \Lambda_i$ and $\eta_i' = \xi_i(\eta_i/(\xi_i + \Lambda_i)$, we find that

$$2\xi_i'\eta_i' = 2(\xi_i + \Lambda_i)\frac{\xi_i\eta_i}{\xi_i + \Lambda_i} = 2\xi_i\eta_i \geq \sigma_i^2,$$

where the last inequality is due to Assumption 2.2.

Remark 2.5. *The preceding calculations show that the equivalent martingale measure $\mathbb{Q}$ is obtained by taking $\boldsymbol{\theta}$ as the predictable process given by*

$$\boldsymbol{\theta}(t) = \begin{bmatrix} (\mu_1 + q_1 - r - \lambda_1\kappa_1 + \tilde{\lambda}_1\tilde{\kappa}_1)(v_1(t))^{-1/2} \\ (\mu_2 + q_2 - r - \lambda_2\kappa_2 + \tilde{\lambda}_2\tilde{\kappa}_2)(v_2(t))^{-1/2} \\ \Lambda_1\sqrt{v_1(t)}/\sigma_1 \\ \Lambda_2\sqrt{v_2(t)}/\sigma_2 \end{bmatrix}. \tag{2.17}$$

As stated before, Assumption 2.2 ensures that v_1 and v_2 do not explode or become zero in finite time under $\mathbb{P}$, and we have demonstrated above that this also holds under the new measure $\mathbb{Q}$. This implies that the individual components of $\boldsymbol{\theta}(t)$ also do not explode in finite time. Therefore $\{L^D(t)\}$, which is the exponential process generated by $\mathbf{X}(t) = -\mathbf{L}^{-1}\boldsymbol{\theta}(t)$ (see Definition A.5), is a strictly positive $\mathbb{P}$-martingale and $\mathbb{E}_{\mathbb{P}}[L^D(t)] = 1$, following the discussion in Appendix A.2. △

The following proposition summarizes the results we have obtained above in relation to the choice of $\boldsymbol{\theta}$, γ_1, γ_2, ν_1, and ν_2 that parameterizes equivalent martingale measures for our SVJD model.

Proposition 2.6. *Suppose $\boldsymbol{\theta}$ is given by equation (2.17) and γ_1, γ_2, ν_1, and ν_2 are constants such that $\tilde{\lambda}_i$ and $M_{\mathbb{P},Y_i}(u + \gamma_i)/M_{\mathbb{P},Y_i}(\gamma_i)$, $i = 1, 2$, in Proposition 2.3 are well-defined. Define the probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ by $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T} = L(T)$, where $L(t)$ is given by equation (2.8). Then $\mathbb{E}_{\mathbb{P}}[L(T)] = 1$ and $\mathbb{Q}$ is an equivalent martingale measure, i.e. $\mathbb{Q}$ is a probability measure, equivalent to $\mathbb{P}$, under which the discounted yield processes $\{e^{-(r-q_i)t}S_i(t)\}_{0\leq t\leq T}$ (where S_i satisfies equation (2.4)), $i = 1, 2$, are martingales.*

2.4 Exchange Option Pricing IPDE

To explicitly denote the time dependence of the quantities and stochastic processes involved while reducing the number of delimiters present in the mathematical statements, dependence on time shall be expressed using subscripts for the remainder of this chapter.

Consider a European and an American exchange option written on two assets with prices $S_{1,t}$ and $S_{2,t}$. If both options expire at time T, then we have

$$C_T^E(S_{1,T}, S_{2,T}, v_{1,T}, v_{2,T}) = C_T^A(S_{1,T}, S_{2,T}, v_{1,T}, v_{2,T}) = (S_{1,T} - S_{2,T})^+.$$

Let $\mathbb{Q}$ be an equivalent martingale measure determined by the Radon-Nikodým derivative in Proposition 2.3. Then, the price of the European exchange option is given by

$$\begin{aligned} C_t^E(S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t}) &= e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}\left[(S_{1,T} - S_{2,T})^+ \middle| \mathcal{F}_t\right] \\ &= e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}\left[(S_{1,T} - S_{2,T})^+ \middle| S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t}\right]. \end{aligned} \tag{2.18}$$

Furthermore, let $\mathscr{T}_{t,T}$ denote the collection of all $\{\mathcal{F}_t\}$-stopping times τ in the interval $[t, T]$. Then the price of the American exchange option is given by

$$C_t^A(S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t}) = \sup_{\tau \in \mathscr{T}_{t,T}} \mathbb{E}_{\mathbb{Q}}\left[e^{-r(\tau-t)}(S_{1,\tau} - S_{2,\tau})^+ \middle| S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t}\right] \tag{2.19}$$

(see for example, Bjerskund and Stensland, 1993).

In order to apply Itô's formula for jump-diffusion processes, we require the following assumption on the European and American option price formulas.

Assumption 2.7. For $t \in [0, T]$, the functions $C_t^E(s_1, s_2, v_1, v_2)$ and $C_t^A(s_1, s_2, v_1, v_2)$ are twice-differentiable in s_1, s_2, v_1, v_2 with continuous second-order partial derivatives and they have a continuous first-order partial derivative with respect to t.

For now, we do not impose the assumption $\rho_z = 0$, as this is not needed in the derivation of the IPDE. We do, however, require that the other conditions in Assumption 2.2 hold.

Given the $\mathbb{Q}$-dynamics of asset prices and volatility processes in equations (2.14) and (2.16), respectively, we can now solve for the stochastic differential equation for exchange option prices $C_t(S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t})$ under $\mathbb{Q}$. Let

$$C_{t-} = C_t(S_{1,t-}, S_{2,t-}, v_{1,t}, v_{2,t})$$

denote the pre-jump price of the exchange option in the event that at time t there is a jump in either $S_{1,t}$ or $S_{2,t}$. From the dynamics of the exchange option price, we can then derive the corresponding pricing IPDE, as shown in the next proposition. Note that different boundary conditions will be used to characterize the American exchange option price from the derived IPDE.

Proposition 2.8. *Suppose Assumption 2.7 holds. Given the that asset prices and volatilities have $\mathbb{Q}$-dynamics given by equations (2.14) and (2.16), the exchange option price $C_t(S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t})$ satisfies the IPDE*

$$rC_{t-} = \frac{\partial C_{t-}}{\partial t} + \mathcal{L}_{\mathbf{s},\mathbf{v}}[C_{t-}] \tag{2.20}$$

where the integro-partial differential operator $\mathcal{L}_{\mathbf{s},\mathbf{v}}$ is defined such that

$$\begin{aligned}\mathcal{L}_{\mathbf{s},\mathbf{v}}[C_{t-}] = &\sum_{i=1}^{2}(r-q_i-\tilde{\lambda}_i\tilde{\kappa}_i)S_{i,t-}\frac{\partial C_{t-}}{\partial s_i}+\sum_{i=1}^{2}[\xi_i\eta_i-(\xi_i+\Lambda_i)v_{i,t}]\frac{\partial C_{t-}}{\partial v_i}\\ &+\frac{1}{2}\sum_{i=1}^{2}v_{i,t}S_{i,t-}^2\frac{\partial^2 C_{t-}}{\partial s_i^2}+\frac{1}{2}\sum_{i=1}^{2}\sigma_i^2 v_{i,t}\frac{\partial^2 C_{t-}}{\partial v_i^2}+\rho_w\sqrt{v_{1,t}v_{2,t}}S_{1,t-}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_1\partial s_2}\\ &+\sum_{i=1}^{2}\rho_{wz_i}\sigma_i v_{i,t}S_{i,t-}\frac{\partial^2 C_{t-}}{\partial s_i\partial v_i}+\rho_Z\sigma_1\sigma_2\sqrt{v_{1,t}v_{2,t}}\frac{\partial^2 f}{\partial v_1\partial v_2}\\ &+\tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_{t-}\left(e^{Y_1}S_{1,t-},S_{2,t-},v_{1,t},v_{2,t}\right)-C_{t-}\right]\\ &+\tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_{t-}\left(S_{1,t-},e^{Y_2}S_{2,t-},v_{1,t},v_{2,t}\right)-C_{t-}\right].\end{aligned}\tag{2.21}$$

Note that the above proposition lacks terminal and boundary conditions to specify the solution of the IPDE. A remark on these conditions will be provided after the proof of the proposition.

Proof. Using Itô's formula for jump-diffusion processes, we find that C_t satisfies the stochastic differential equation

$$\begin{aligned}\mathrm{d}C_t = &\left\{\frac{\partial C_{t-}}{\partial t}+\sum_{i=1}^{2}(r-q_i-\tilde{\lambda}_i\tilde{\kappa}_i)S_{i,t-}\frac{\partial C_{t-}}{\partial s_i}+\sum_{i=1}^{2}[\xi_i\eta_i-(\xi_i+\Lambda_i)v_{i,t}]\frac{\partial C_{t-}}{\partial v_i}\right.\\ &+\frac{1}{2}\sum_{i=1}^{2}v_{i,t}S_{i,t-}^2\frac{\partial^2 C_{t-}}{\partial s_i^2}+\frac{1}{2}\sum_{i=1}^{2}\sigma_i^2 v_{i,t}\frac{\partial^2 C_{t-}}{\partial v_i^2}+\rho_w\sqrt{v_{1,t}v_{2,t}}S_{1,t-}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_1\partial s_2}\\ &\left.+\rho_{wz_1}\sigma_1 v_{1,t}S_{1,t-}\frac{\partial^2 C_{t-}}{\partial s_1\partial v_1}+\rho_{wz_2}\sigma_2 v_{2,t}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_2\partial v_2}+\rho_Z\sigma_1\sigma_2\sqrt{v_{1,t}v_{2,t}}\frac{\partial^2 C_{t-}}{\partial v_1\partial v_2}\right\}\mathrm{d}t\\ &+\sum_{i=1}^{2}\sqrt{v_{i,t}}S_{i,t-}\frac{\partial C_{t-}}{\partial s_i}\,\mathrm{d}\tilde{W}_{i,t}+\sum_{i=1}^{2}\sigma_i\sqrt{v_{i,t}}\frac{\partial C_{t-}}{\partial v_i}\,\mathrm{d}\tilde{Z}_{i,t}\\ &+\int_{\mathbb{R}}\left[C_t(S_{1,t-}e^y,S_{2,t-},v_{1,t},v_{2,t})-C_{t-}\right]p_1(\mathrm{d}y,\mathrm{d}t)\\ &+\int_{\mathbb{R}}\left[C_t(S_{1,t-},S_{2,t-}e^y,v_{1,t},v_{2,t})-C_{t-}\right]p_2(\mathrm{d}y,\mathrm{d}t).\end{aligned}$$

The counting measure $p_i(\mathrm{d}y,\mathrm{d}t)$ can be replaced by the $\mathbb{Q}$-compensated counting measure $q_i(\mathrm{d}y,\mathrm{d}t)$, giving us

$$\begin{aligned}&\int_{\mathbb{R}}\left[C_t(S_{1,t-}e^y,S_{2,t-},v_{1,t},v_{2,t})-C_{t-}\right]p_1(\mathrm{d}y,\mathrm{d}t)\\ &\quad=\int_{\mathbb{R}}\left[C_t(S_{1,t-}e^y,S_{2,t-},v_{1,t},v_{2,t})-C_{t-}\right]\left[q_1(\mathrm{d}y,\mathrm{d}t)+\tilde{\lambda}_1 m_{\mathbb{Q}}^{(1)}(\mathrm{d}y)\,\mathrm{d}t\right]\\ &\quad=\tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_t\left(S_{1,t-}e^{Y_1},S_{2,t-},v_{1,t},v_{2,t}\right)-C_{t-}\right]\mathrm{d}t\\ &\qquad+\int_{\mathbb{R}}\left[C_t(S_{1,t-}e^y,S_{2,t-},v_{1,t},v_{2,t})-C_{t-}\right]q_1(\mathrm{d}y,\mathrm{d}t),\end{aligned}$$

where

$$\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_t(S_{1,t-}e^{Y_1}, S_{2,t-}, v_{1,t}, v_{2,t}) - C_{t-}\right] = \int_{\mathbb{R}}\left[C_t(S_{1,t-}e^{y}, S_{2,t-}, v_{1,t}, v_{2,t}) - C_{t-}\right] m_{\mathbb{Q}}^{(1)}(\mathrm{d}y)$$

represents the expected change in the price of the exchange option due to jumps in the price of stock 1. Likewise, we can write

$$\begin{aligned}
&\int_{\mathbb{R}}\left[C_t(S_{1,t-}, S_{2,t-}e^{y}, v_{1,t}, v_{2,t}) - C_{t-}\right] p_2(\mathrm{d}y, \mathrm{d}t) \\
&\quad = \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_t\left(S_{1,t-}, S_{2,t-}e^{Y_2}, v_{1,t}, v_{2,t}\right) - C_{t-}\right]\mathrm{d}t \\
&\qquad + \int_{\mathbb{R}}\left[C_t(S_{1,t-}, S_{2,t-}e^{y}, v_{1,t}, v_{2,t}) - C_{t-}\right] q_2(\mathrm{d}y, \mathrm{d}t)
\end{aligned}$$

We can thus rewrite $\mathrm{d}C_t$ as

$$\begin{aligned}
\mathrm{d}C_t = \Bigg\{ & \frac{\partial C_{t-}}{\partial t} + \sum_{i=1}^{2}(r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i)S_{i,t-}\frac{\partial C_{t-}}{\partial s_i} + \sum_{i=1}^{2}[\xi_i\eta_i - (\xi_i + \Lambda_i)v_{i,t}]\frac{\partial C_{t-}}{\partial v_i} \\
& + \frac{1}{2}\sum_{i=1}^{2} v_{i,t}S_{i,t-}^2\frac{\partial^2 C_{t-}}{\partial s_i^2} + \frac{1}{2}\sum_{i=1}^{2}\sigma_i^2 v_{i,t}\frac{\partial^2 C_{t-}}{\partial v_i^2} + \rho_w\sqrt{v_{1,t}v_{2,t}}S_{1,t-}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_1 \partial s_2} \\
& + \rho_{wz_1}\sigma_1 v_{1,t}S_{1,t-}\frac{\partial^2 C_{t-}}{\partial s_1 \partial v_1} + \rho_{wz_2}\sigma_2 v_{2,t}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_2 \partial v_2} + \rho_Z\sigma_1\sigma_2\sqrt{v_{1,t}v_{2,t}}\frac{\partial^2 C_{t-}}{\partial v_1 \partial v_2} \\
& + \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_t\left(S_{1,t-}e^{Y_1}, S_{2,t-}, v_{1,t}, v_{2,t}\right) - C_{t-}\right] \\
& + \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_t\left(S_{1,t-}, S_{2,t-}e^{Y_2}, v_{1,t}, v_{2,t}\right) - C_{t-}\right]\Bigg\}\mathrm{d}t \\
& + \sum_{i=1}^{2}\sqrt{v_{i,t}}S_{i,t-}\frac{\partial C_{t-}}{\partial s_i}\,\mathrm{d}\tilde{W}_{i,t} + \sum_{i=1}^{2}\sigma_i\sqrt{v_{i,t}}\frac{\partial C_{t-}}{\partial v_i}\,\mathrm{d}\tilde{Z}_{i,t} \\
& + \int_{\mathbb{R}}\left[C_t(S_{1,t-}e^{y}, S_{2,t-}, v_{1,t}, v_{2,t}) - C_{t-}\right] q_1(\mathrm{d}y, \mathrm{d}t) \\
& + \int_{\mathbb{R}}\left[C_t(S_{1,t-}, S_{2,t-}e^{y}, v_{1,t}, v_{2,t}) - C_{t-}\right] q_2(\mathrm{d}y, \mathrm{d}t).
\end{aligned}$$

If $\tilde{C}_t = e^{-rt}C_t$ represents the discounted exchange option price, then we find that $\tilde{C}_t$ satisfies

the SDE

$$
\begin{aligned}
\mathrm{d}\tilde{C}_t &= -rC_{t-}e^{-rt}\,\mathrm{d}t + e^{-rt}\,\mathrm{d}C_t \\
&= e^{-rt}\left\{ \frac{\partial C_{t-}}{\partial t} + \sum_{i=1}^{2}(r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i)S_{i,t-}\frac{\partial C_{t-}}{\partial s_i} + \sum_{i=1}^{2}[\xi_i\eta_i - (\xi_i + \Lambda_i)v_{i,t}]\frac{\partial C_{t-}}{\partial v_i} \right. \\
&\quad + \frac{1}{2}\sum_{i=1}^{2} v_{i,t}S_{i,t-}^2\frac{\partial^2 C_{t-}}{\partial s_i^2} + \frac{1}{2}\sum_{i=1}^{2}\sigma_i^2 v_{i,t}\frac{\partial^2 C_{t-}}{\partial v_i^2} + \rho_w\sqrt{v_{1,t}v_{2,t}}S_{1,t-}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_1\partial s_2} \\
&\quad + \rho_{wz_1}\sigma_1 v_{1,t}S_{1,t-}\frac{\partial^2 C_{t-}}{\partial s_1\partial v_1} + \rho_{wz_2}\sigma_2 v_{2,t}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_2\partial v_2} + \rho_Z\sigma_1\sigma_2\sqrt{v_{1,t}v_{2,t}}\frac{\partial^2 C_{t-}}{\partial v_1\partial v_2} \\
&\quad + \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_t\left(S_{1,t-}e^{Y_1}, S_{2,t-}, v_{1,t}, v_{2,t}\right) - C_{t-}\right] \\
&\quad \left. + \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_t\left(S_{1,t-}, S_{2,t-}e^{Y_2}, v_{1,t}, v_{2,t}\right) - C_{t-}\right] - rC_{t-} \right\}\mathrm{d}t \\
&\quad + e^{-rt}\sum_{i=1}^{2}\sqrt{v_{i,t}}S_{i,t-}\frac{\partial C_{t-}}{\partial s_i}\,\mathrm{d}\tilde{W}_{i,t} + e^{-rt}\sum_{i=1}^{2}\sigma_i\sqrt{v_{i,t}}\frac{\partial C_{t-}}{\partial v_i}\,\mathrm{d}\tilde{Z}_{i,t} \\
&\quad + e^{-rt}\int_{\mathbb{R}}\left[C_t(S_{1,t-}e^{y}, S_{2,t-}, v_{1,t}, v_{2,t}) - C_{t-}\right] q_1(\mathrm{d}y, \mathrm{d}t) \\
&\quad + e^{-rt}\int_{\mathbb{R}}\left[C_t(S_{1,t-}, S_{2,t-}e^{y}, v_{1,t}, v_{2,t}) - C_{t-}\right] q_2(\mathrm{d}y, \mathrm{d}t)
\end{aligned}
\tag{2.22}
$$

The non-explosion of the volatility processes (implied by Assumption 2.2) and the differentiability of the option price (Assumption 2.7) ensure that

$$
\mathbb{E}_{\mathbb{Q}}\left[\int_0^t \left| e^{-rs}\sqrt{v_{i,s}}S_{i,s-}\frac{\partial C_{s-}}{\partial s_i}\right|^2 \mathrm{d}s\right] < \infty \quad \text{and} \quad \mathbb{E}_{\mathbb{Q}}\left[\int_0^t \left| e^{-rs}\sigma_i\sqrt{v_{i,s}}\frac{\partial C_{s-}}{\partial v_i}\right|^2 \mathrm{d}s\right] < \infty
$$

hold for $i = 1, 2$, and so processes whose stochastic differentials correspond to the $\mathrm{d}\tilde{W}_{i,t}$ and $\mathrm{d}\tilde{Z}_{i,t}$ terms are $\mathbb{Q}$-martingales (see Kuo, 2006, Theorem 4.6.1). Furthermore, the last two terms of the right-hand side of the above equation correspond to $\mathbb{Q}$-martingales, provided integrability conditions hold for the option price increments (see for example, Runggaldier, 2003, Theorem 2.2).

Under $\mathbb{Q}$, we require $\tilde{C}_t$ to have no drift. This ensures that $\tilde{C}_t$ is a $\mathbb{Q}$ martingale since, in the absence of a drift term, $\mathrm{d}\tilde{C}_t$ will be expressed in terms of only integrals with respect to standard $\mathbb{Q}$-Wiener processes and integrals with respect to $\mathbb{Q}$-compensated counting measures.[3] Setting the

[3] Note that this is an application of a two-dimensional version of the Martingale Representation Theorem for jump-diffusion processes (see Runggaldier (2003, Theorem 2.2) for the one-dimensional case and Øksendal and Sulem (2019, Theorem 1.35) for the multi-dimensional case).

coefficient of dt in equation (2.22) to zero, we find that the exchange option price satisfies the IPDE

$$\begin{aligned} rC_{t-} = & \frac{\partial C_{t-}}{\partial t} + \sum_{i=1}^{2}(r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i)S_{i,t-}\frac{\partial C_{t-}}{\partial s_i} + \sum_{i=1}^{2}[\xi_i\eta_i - (\xi_i + \Lambda_i)v_{i,t}]\frac{\partial C_{t-}}{\partial v_i} \\ & + \frac{1}{2}\sum_{i=1}^{2} v_{i,t}S_{i,t-}^2\frac{\partial^2 C_{t-}}{\partial s_i^2} + \frac{1}{2}\sum_{i=1}^{2}\sigma_i^2 v_{i,t}\frac{\partial^2 C_{t-}}{\partial v_i^2} + \rho_w\sqrt{v_{1,t}v_{2,t}}S_{1,t-}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_1 \partial s_2} \\ & + \rho_{wz_1}\sigma_1 v_{1,t}S_{1,t-}\frac{\partial^2 C_{t-}}{\partial s_1 \partial v_1} + \rho_{wz_2}\sigma_2 v_{2,t}S_{2,t-}\frac{\partial^2 C_{t-}}{\partial s_2 \partial v_2} + \rho_Z\sigma_1\sigma_2\sqrt{v_{1,t}v_{2,t}}\frac{\partial^2 C_{t-}}{\partial v_1 \partial v_2} \\ & + \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_t\left(S_{1,t-}e^{Y_1}, S_{2,t-}, v_{1,t}, v_{2,t}\right) - C_{t-}\right] \\ & + \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_t\left(S_{1,t-}, S_{2,t-}e^{Y_2}, v_{1,t}, v_{2,t}\right) - C_{t-}\right] \end{aligned} \tag{2.23}$$

The result in the proposition is obtained by writing the above IPDE in terms of the operator $\mathcal{L}_{\mathbf{s},\mathbf{v}}$. ■

The IPDE derived above extends the result obtained by Cheang and Chiarella (2011) for exchange options under jump-diffusion dynamics to the case of stochastic volatility and jump-diffusion dynamics. It is also an extension of the IPDE derived by Cheang et al. (2013) for the one-asset option under SVJD dynamics to the case of two risky assets.

For the European exchange option, the terminal condition for the IPDE is

$$C_T^E = (S_{1,T} - S_{2,T})^+.$$

In the case of the American exchange option, additional conditions, namely the early exercise boundary condition and smooth-pasting conditions, must be specified given the early exercise boundary of the option (Chiarella et al., 2009; Cheang and Chiarella, 2011; Cheang et al., 2013; Chiarella et al., 2015). These will be discussed in Section 2.6.

2.5 A Representation of the European Exchange Option Price

In this section, we now consider analytical representations of the price of the European exchange option. To this end, we employ two methods: the change-of-numéraire technique of Geman et al. (1995), which was applied by Cheang and Chiarella (2011) to the exchange option, and the Fourier transform approach by Dempster and Hong (2002) and Caldana and Fusai (2013). For the first approach, we assume that $\rho_z = 0$ and that Assumption 2.2 hold. In the latter approach, we require additional restrictions on the correlation structure of the Wiener processes.

2.5.1 A Change-of-Numéraire Approach

In this section, we employ the change of numéraire technique of Geman et al. (1995) to evaluate the $\mathbb{Q}$-expectation that gives the price of the European exchange option. This technique has been employed in Cheang and Chiarella (2011) and Caldana et al. (2015) to price exchange options under jump-diffusion dynamics, Cheang and Teh (2014) to price single-asset options under jump-diffusion dynamics with stochastic interest rate, and Cheang et al. (2011) and Cheang et al. (2013) to price single-asset European options under SVJD dynamics. Here, we derive a representation for European exchange option prices under the two-asset SVJD model.

Without loss of generality, we analyze the European exchange option price at time $t = 0$, which is given by

$$C_0^E(S_{1,0}, S_{2,0}, v_{1,0}, v_{2,0}) = e^{-rT}\mathbb{E}_{\mathbb{Q}}\left[(S_{1,T} - S_{2,T})^+\right].$$

Define the event $\mathcal{A}_0 = \{S_{1,T} > S_{2,T}\}$ (the event that the option is in-the-money), so that we can write

$$C_0^E = e^{-rT}\mathbb{E}_{\mathbb{Q}}[S_{1,T}\mathbf{1}_{\mathcal{A}_0}] - e^{-rT}\mathbb{E}_{\mathbb{Q}}[S_{2,T}\mathbf{1}_{\mathcal{A}_0}]. \tag{2.24}$$

Using equation (2.15), we may substitute expressions for $S_{1,T}$ and $S_{2,T}$, giving us

$$C_0^E = S_{1,0}e^{-q_1T}\mathbb{E}_{\mathbb{Q}}[U_{1,T}\mathbf{1}_{\mathcal{A}_0}] - S_{2,0}e^{-q_2T}\mathbb{E}_{\mathbb{Q}}[U_{2,T}\mathbf{1}_{\mathcal{A}_0}], \tag{2.25}$$

where

$$U_{i,T} = \exp\left\{-\frac{1}{2}\int_0^T v_{i,t}\,\mathrm{d}t + \int_0^T \sqrt{v_{i,t}}\,\mathrm{d}\tilde{W}_{i,t} - \tilde{\lambda}_i\tilde{\kappa}_i T + \sum_{n=1}^{N_{i,T}} Y_{i,n}\right\}. \tag{2.26}$$

Therefore, the expectations may be viewed as the probability that the option is in the money at time T under two new probability measures $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$ whose relative densities with respect to $\mathbb{Q}$ are $U_{1,T}$ and $U_{2,T}$, respectively. The probability measures $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$ are those that result from using $S_{1,t}$ and $S_{2,t}$, respectively, as the numéraire.

To show that $U_{1,T}$ and $U_{2,T}$ are sensible Radon-Nikodým derivatives in the sense of Proposition 2.3, we must show that $U_{1,t}$ and $U_{2,t}$ are $\mathbb{Q}$-martingales and $\mathbb{E}_{\mathbb{Q}}[U_{i,t}] = 1$, $i = 1, 2$. This implies that we must ensure that the volatility processes do not explode under the new probability measures $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$ (which in effect guarantees the existence of an expression for asset prices in the new probability measures).

The subsequent analysis is for $\{U_{1,t}\}$ and is also applicable to $\{U_{2,t}\}$. As shown in Section 2.3, the condition $2\xi_1\eta_1 \geq \sigma_1^2$ on the $\mathbb{P}$-dynamics of $\{v_{1,t}\}$ is sufficient to ensure that the process neither explodes nor makes excursions to the origin under $\mathbb{Q}$. As such, $\int_0^t v_{1,s}\,\mathrm{d}s < \infty$ $\mathbb{Q}$-a.s., which satisfies

the Novikov condition to ensure that $\exp\{-\frac{1}{2}\int_0^t v_{1,s}\,\mathrm{d}s + \int_0^t \sqrt{v_{1,s}}\,\mathrm{d}W_{1,s}\}$ is a $\mathbb{Q}$-martingale and

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{-\frac{1}{2}\int_0^t v_{1,s}\,\mathrm{d}s + \int_0^t \sqrt{v_{1,s}}\,\mathrm{d}W_{1,s}\right\}\right] = 1.$$

We also note that

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\sum_{i=1}^{N_{1,t}} Y_{1,n}\right\}\right] = \exp\left\{\tilde{\lambda}_1 t(\mathbb{E}_{\mathbb{Q}}(e^{Y_i}) - 1)\right\} = e^{\tilde{\lambda}_1\tilde{\kappa}_1 t}.$$

Independence of the Wiener and jump components imply that

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[U_{1,t}] &= \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{-\frac{1}{2}\int_0^t v_{1,s}\,\mathrm{d}s + \int_0^t \sqrt{v_{1,s}}\,\mathrm{d}W_{1,s}\right\}\right]\cdot e^{-\tilde{\lambda}_1\tilde{\kappa}_1 t}\cdot \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\sum_{i=1}^{N_{1,t}} Y_{1,n}\right\}\right] \\ &= 1 = U_{1,0}.\end{aligned}$$

This also shows that $\{U_{1,t}\}$ is a $\mathbb{Q}$-martingale.

By Proposition 2.3, $U_{1,T}$ defines a Radon-Nikodým derivative that facilitates a change of measure from $\mathbb{Q}$ to some equivalent measure $\hat{\mathbb{Q}}_1$. To determine the drift of the Wiener processes and the distributional properties of the jumps under $\hat{\mathbb{Q}}_1$, we compare equation (2.26) (with $i = 1$) with equation (2.8) to determine the change of measure parameters.

We first investigate changes in the jump components arising from the shift to $\hat{\mathbb{Q}}_1$. Note that the jump component of $S_{2,t}$ does not appear in $U_{1,T}$, implying that the change of measure has no effect on the jumps in stock 2. From the comparison, we find that $\gamma_1 = 1$ and $\nu_1 = 0$, which implies that the Poisson process $\{N_{1,t}\}$ has $\hat{\mathbb{Q}}_1$-intensity

$$\hat{\lambda}_1^{[1]} = \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}(e^{Y_1}) = \tilde{\lambda}_1(1 + \tilde{\kappa}_1)$$

and the new distribution of the jump random variables $Y_{1,n}$ is given by the moment generating function

$$M_{\hat{\mathbb{Q}}_1,Y_1}(u) = \frac{M_{\mathbb{Q},Y_1}(u+1)}{M_{\mathbb{Q},Y_1}(1)}.$$

This relation between the moment generating functions also implies that the $\hat{\mathbb{Q}}_1$-density of Y_1 is given by

$$m^{(1)}_{\hat{\mathbb{Q}}_1}(\mathrm{d}y) = \frac{e^y}{\mathbb{E}_{\mathbb{Q}}(e^{Y_1})} m^{(1)}_{\mathbb{Q}}(\mathrm{d}y).$$

This analysis therefore implies that the compensated counting measures under $\hat{\mathbb{Q}}_1$ corresponding

to the original counting measure $p(\mathrm{d}y_i, \mathrm{d}t)$ are given by

$$q_i^{[1]}(\mathrm{d}y, \mathrm{d}t) = p_i(\mathrm{d}y, \mathrm{d}t) - \hat{\lambda}_i^{[1]} m_{\hat{\mathbb{Q}}_1}^{(i)}(\mathrm{d}y)\,\mathrm{d}t, \qquad i = 1, 2,$$

where $\hat{\lambda}_1^{[1]}$ and $m_{\hat{\mathbb{Q}}_1}^{(1)}(\mathrm{d}y)$ are given above, and $\hat{\lambda}_2^{[1]} = \tilde{\lambda}_2$ and $m_{\hat{\mathbb{Q}}_1}^{(2)}(\mathrm{d}y) = m_{\mathbb{Q}}^{(2)}(\mathrm{d}y)$ as no changes are introduced to the jump components of stock 2.

Now we consider the diffusion components. The parameter $\hat{\boldsymbol{\theta}}_t^{[1]}$ for this change of measure (analogous to $\boldsymbol{\theta}_t$ in equation (2.8)) is defined such that

$$\left(\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\theta}}_t^{[1]}\right)^\top \mathrm{d}\tilde{\mathbf{B}}_t = \sqrt{v_{1,t}}\,\mathrm{d}\tilde{W}_{1,t},$$

where $\mathrm{d}\tilde{\mathbf{B}}_t = (\mathrm{d}\tilde{W}_{1,t}, \mathrm{d}\tilde{W}_{2,t}, \mathrm{d}\tilde{Z}_{1,t}, \mathrm{d}\tilde{Z}_{2,t})^\top$ is the vector of the $\mathbb{Q}$-Wiener increments and $\boldsymbol{\Sigma}$ is the original correlation matrix of the Wiener processes. The above equation, in matrix form, can also be written as

$$\begin{bmatrix} 1 & \rho_w & \rho_{wz_1} & 0 \\ \rho_w & 1 & 0 & \rho_{wz_2} \\ \rho_{wz_1} & 0 & 1 & \rho_z \\ 0 & \rho_{wz_2} & \rho_z & 1 \end{bmatrix}^{-1} \hat{\boldsymbol{\theta}}_t^{(1)} = \begin{bmatrix} \sqrt{v_{1,t}} \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

This implies that

$$\hat{\boldsymbol{\theta}}_t^{[1]} = \left(\sqrt{v_{1,t}}, \rho_w\sqrt{v_{1,t}}, \rho_{wz_1}\sqrt{v_{1,t}}, 0\right)^\top.$$

Thus, if $\hat{W}_{1,t}^{[1]}$, $\hat{W}_{2,t}^{[1]}$, $\hat{Z}_{1,t}^{[1]}$, and $\hat{Z}_{2,t}^{[1]}$ are standard $\hat{\mathbb{Q}}_1$-Wiener processes, then the $\hat{\mathbb{Q}}_1$-dynamics of the $\mathbb{Q}$-Wiener processes are given by

$$\begin{aligned} \mathrm{d}\tilde{W}_{1,t} &= -\sqrt{v_{1,t}}\,\mathrm{d}t + \mathrm{d}\hat{W}_{1,t}^{[1]}, \\ \mathrm{d}\tilde{W}_{2,t} &= -\rho_w\sqrt{v_{1,t}}\,\mathrm{d}t + \mathrm{d}\hat{W}_{2,t}^{[1]}, \\ \mathrm{d}\tilde{Z}_{1,t} &= -\rho_{wz_1}\sqrt{v_{1,t}}\,\mathrm{d}t + \mathrm{d}\hat{Z}_{1,t}^{[1]}, \\ \mathrm{d}\tilde{Z}_{2,t} &= \mathrm{d}\hat{Z}_{2,t}^{[1]}. \end{aligned} \tag{2.27}$$

Under $\hat{\mathbb{Q}}_1$, the variance process $v_{1,t}$ satisfies the equation

$$\mathrm{d}v_{1,t} = (\xi_1 + \Lambda_1 - \sigma_1\rho_{wz_1})\left[\frac{\xi_1\eta_1}{\xi_1 + \Lambda_1 - \sigma_1\rho_{wz_1}} - v_{1,t}\right]\mathrm{d}t + \sigma_1\sqrt{v_{1,t}}\,\mathrm{d}\hat{Z}_{1,t}^{[1]}, \tag{2.28}$$

which is obtained by substituting $\mathrm{d}\tilde{Z}_{1,t} = -\rho_{wz_1}\sqrt{v_{1,t}}\,\mathrm{d}t + \mathrm{d}\hat{Z}_{1,t}^{[1]}$ into equation (2.16). In light of Assumption 2.2, we require that $\xi_1 + \Lambda_1 - \sigma_1\rho_{wz_1} > 0$. From here and from the fact that Λ_1 is

chosen to be a nonnegative constant, we see that the condition

$$-1 < \rho_{wz_1} < \min\left\{\frac{\xi_1}{\sigma_1}, 1\right\} \tag{2.29}$$

is sufficient to ensure that the coefficients of the volatility process under $\hat{\mathbb{Q}}_1$ are positive. Furthermore, we observe that

$$2\left(\xi_1 + \Lambda_1 - \sigma_1\rho_{wz_1}\right) \cdot \frac{\xi_1\eta_1}{\xi_1 + \Lambda_1 - \sigma_1\rho_{wz_1}} = 2\xi_1\eta_1 \geq \sigma_1^2,$$

which means that $v_{1,t}$ neither explodes in finite time or makes excursions to 0 under $\hat{\mathbb{Q}}_1$.

Similarly, the $\hat{\mathbb{Q}}_1$ dynamics of $v_{2,t}$ is given by

$$\mathrm{d}v_{2,t} = (\xi_2 + \Lambda_2)\left[\frac{\xi_2\eta_2}{\xi_2 + \Lambda_2} - v_{2,t}\right]\mathrm{d}t + \sigma_2\sqrt{v_{2,t}}\,\mathrm{d}\hat{Z}_{2,t}^{[1]},$$

which we note to be identical to equation (2.16) except for the change in the Wiener process. At this point, no further parameter assumptions are required aside from those in Assumption 2.2 to ensure that $v_{2,t}$ does not vanish or explode in finite time under $\hat{\mathbb{Q}}_1$.

An analysis similar to that above also shows that $\{U_{2,t}\}$ is a $\mathbb{Q}$-martingale, allowing us to define a new probability measure $\hat{\mathbb{Q}}_2$ equivalent to $\mathbb{Q}$ via the Radon-Nikodým derivative $U_{2,T}$. Under $\hat{\mathbb{Q}}_2$, the local characteristics $(\hat{\lambda}_1^{[2]}, m_{\hat{\mathbb{Q}}_2}^{(1)}(\mathrm{d}y))$ of the counting measure $p_1(\mathrm{d}y, \mathrm{d}t)$ are given by

$$\hat{\lambda}_1^{[2]} = \tilde{\lambda}_1, \qquad m_{\hat{\mathbb{Q}}_2}^{(1)}(\mathrm{d}y) = m_{\mathbb{Q}}^{(1)}(\mathrm{d}y),$$

as $U_{2,T}$ is parameterized such that no changes are induced on the distributional properties of the jump component of stock 1. Analogous to the above analysis, we find that the $\hat{\mathbb{Q}}_2$-local characteristics $(\hat{\lambda}_2^{[2]}, m_{\hat{\mathbb{Q}}_2}^{(2)}(\mathrm{d}y))$ of $p_2(\mathrm{d}y, \mathrm{d}t)$ are given by

$$\begin{aligned}\hat{\lambda}_2^{[2]} &= \tilde{\lambda}_2(1 + \tilde{\kappa}_2),\\ m_{\hat{\mathbb{Q}}_2}^{(2)}(\mathrm{d}y) &= \frac{e^y}{\mathbb{E}_{\mathbb{Q}}(e^{Y_2})} m_{\mathbb{Q}}^{(2)}(\mathrm{d}y).\end{aligned}$$

A comparison between equation (2.26) (with $i = 2$) with equation (2.8) produces

$$\hat{\boldsymbol{\theta}}_t^{[2]} = \left(\rho_w\sqrt{v_{2,t}}, \sqrt{v_{2,t}}, 0, \rho_{wz_2}\sqrt{v_{2,t}}\right)^\top,$$

to facilitate the change of drift upon shifting to $\hat{\mathbb{Q}}_2$. Let $\hat{W}_{1,t}^{[2]}$, $\hat{W}_{2,t}^{[2]}$, $\hat{Z}_{1,t}^{[2]}$, and $\hat{Z}_{2,t}^{[2]}$ be standard

$\hat{\mathbb{Q}}_2$-Wiener processes; then the $\hat{\mathbb{Q}}_2$-dynamics of the $\mathbb{Q}$-Wiener processes are given by

$$\begin{aligned}
\mathrm{d}\tilde{W}_{1,t} &= -\rho_w\sqrt{v_{2,t}}\,\mathrm{d}t + \mathrm{d}\hat{W}^{[2]}_{1,t},\\
\mathrm{d}\tilde{W}_{2,t} &= -\sqrt{v_{2,t}}\,\mathrm{d}t + \mathrm{d}\hat{W}^{[2]}_{2,t},\\
\mathrm{d}\tilde{Z}_{1,t} &= \mathrm{d}\hat{Z}^{[2]}_{1,t},\\
\mathrm{d}\tilde{Z}_{2,t} &= -\rho_{wz_2}\sqrt{v_{2,t}}\,\mathrm{d}t + \mathrm{d}\hat{Z}^{[2]}_{2,t}.
\end{aligned} \tag{2.30}$$

Consequently, the $\hat{\mathbb{Q}}_2$-dynamics of $v_{2,t}$ is given by the equation

$$\mathrm{d}v_{2,t} = \left(\xi_2 + \Lambda_2 - \sigma_2\rho_{wz_2}\right)\left[\frac{\xi_2\eta_2}{\xi_2 + \Lambda_2 - \sigma_2\rho_{wz_2}} - v_{2,t}\right]\mathrm{d}t + \sigma_2\sqrt{v_{2,t}}\,\mathrm{d}\hat{Z}^{[2]}_{2,t}, \tag{2.31}$$

which is guaranteed to neither explode in finite time nor make excursions to 0 under $\hat{\mathbb{Q}}_2$ by enforcing the condition

$$-1 < \rho_{wz_2} < \min\left\{\frac{\xi_2}{\sigma_2}, 1\right\}. \tag{2.32}$$

No additional restrictions need to be made to ensure that $v_{1,t}$, aside from those in Assumption 2.2, does not explode nor go to zero under $\hat{\mathbb{Q}}_2$.

The preceding discussion indeed show that Assumption 2.2, in the case that $\rho_z = 0$, is sufficient to ensure that the volatility processes $\{v_{1,t}\}$ and $\{v_{2,t}\}$ neither explode in finite time nor hit zero under the risk-neutral measure $\mathbb{Q}$ and the probability measures $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$ equivalent to $\mathbb{Q}$ determined by the Radon-Nikodým derivatives $U_{1,T}$ and $U_{2,T}$. This extends the discussion of Cheang et al. (2011) to the case of two assets modelled with stochastic volatility and jump-diffusion dynamics.

In terms of the new probability measures $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$, the price of the European exchange option may be written as

$$C_0^E = S_{1,0}e^{-q_1T}\hat{\mathbb{Q}}_1(\mathcal{A}_0) - S_{2,0}e^{-q_2T}\hat{\mathbb{Q}}_2(\mathcal{A}_0). \tag{2.33}$$

Using equation (2.15), the event $\mathcal{A}_0$ may be rewritten as

$$\left\{\mathfrak{A}_{0,T} > \ln\left(\frac{S_{2,0}}{S_{1,0}}\right) - \left(q_2 - q_1 - \tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2\right)T\right\}, \tag{2.34}$$

where $\mathfrak{A}_{0,T}$ is the random variable

$$\mathfrak{A}_{0,T} = -\frac{1}{2}\int_0^T (v_{1,t} - v_{2,t})\,\mathrm{d}t + \int_0^T \sqrt{v_{1,t}}\,\mathrm{d}\tilde{W}_{1,t} - \int_0^T \sqrt{v_{2,t}}\,\mathrm{d}\tilde{W}_{2,t} + \sum_{n=1}^{N_{1,T}} Y_{1,n} - \sum_{n=1}^{N_{2,T}} Y_{2,n}. \tag{2.35}$$

Similar to the original Margrabe (1978) formula, our characterization of the European exchange option price under SVJD dynamics is also independent of the risk-free interest rate r.

In general, for any $0 \leq t < T$, the time-t price of the European exchange option is given by

$$C_t^E = S_{1,t}e^{-q_1(T-t)}\hat{\mathbb{Q}}_1(\mathcal{A}_t) - S_{2,t}e^{-q_2(T-t)}\hat{\mathbb{Q}}_2(\mathcal{A}_t), \tag{2.36}$$

where $\mathcal{A}_t$ is the event

$$\left\{\mathfrak{A}_{t,T} > \ln\left(\frac{S_{2,t}}{S_{1,t}}\right) - \left(q_2 - q_1 - \tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2\right)(T-t)\right\}$$

and $\mathfrak{A}_{t,T}$ is the random variable

$$\mathfrak{A}_{t,T} = -\frac{1}{2}\int_t^T (v_{1,s} - v_{2,s})\,\mathrm{d}s + \int_t^T \sqrt{v_{1,s}}\,\mathrm{d}\tilde{W}_{1,s} - \int_t^T \sqrt{v_{2,s}}\,\mathrm{d}\tilde{W}_{2,s} + \sum_{n=N_{1,t}}^{N_{1,T}} Y_{1,n} - \sum_{n=N_{2,t}}^{N_{2,T}} Y_{2,n}.$$

Here, the summations count the number of jumps that occur in the period $(t, T]$. By construction, C_t^E as given above is a solution to the IPDE in Proposition 2.8 subject to the terminal condition $C_T^E = (S_{1,T} - S_{2,T})^+$.

In the presence of stochastic volatilities $v_{1,t}$ and $v_{2,t}$, a series expression for the European exchange option price similar to those obtained in Cheang and Chiarella (2011) and Caldana et al. (2015) cannot be obtained. In light of our general correlation structure for the Wiener processes, the probabilities above may be computed via simulation methods. Alternatively, a solution via characteristic functions can be made possible by making additional assumptions on the relationship between the stock prices and the volatilties, as will be shown in the next section. In the next section, we will derive a representation of the European exchange option price in terms of Fourier inversion formulas whose form is comparable to equations (2.24) and (2.33).

2.5.2 A Fourier Transform Approach

One may derive a Fourier inversion formula for the price of the European exchange option in terms of the joint characteristic function of the log-prices of the stocks. However, an explicit formula for the characteristic function is available if we make additional simplifying assumptions to the correlation structure of the Wiener processes. The derivation of the characteristic function and the application of the Caldana and Fusai (2013) result to derive the exact price of the European exchange option is the main topic of this section. Here, we show that the Caldana and Fusai (2013) result also lends itself to a decomposition similar to equation (2.33) in the case of the European exchange option.

Recall that the risk-neutral dynamics of the log-price of the stock is given by

$$\mathrm{d}X_{i,t} = \left(r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i - \frac{1}{2}v_{i,t}\right)\mathrm{d}t + \sqrt{v_{i,t}}\,\mathrm{d}\tilde{W}_{i,t} + \mathrm{d}\tilde{Q}_{i,t}, \quad i = 1, 2 \tag{2.37}$$

where $\tilde{Q}_{i,t}$ is a compound Poisson process with intensity $\tilde{\lambda}_i$ under $\mathbb{Q}$ whose jump components $Y_{i,1}, Y_{i,2}, \ldots$ independent and identically distributed with common characteristic function $\phi_{Y_i}(u)$. We also recall that under $\mathbb{Q}$, the volatility processes $v_{i,t}$ satisfy the equation

$$\mathrm{d}v_{i,t} = \left[\xi_i\eta_i - (\xi_i + \Lambda_i)v_{i,t}\right]\mathrm{d}t + \sigma_i\sqrt{v_{i,t}}\,\mathrm{d}\tilde{Z}_{i,t}, \tag{2.38}$$

where the model parameters follow Assumption 2.2.

Denote by $X^c_{i,t}$ the continuous part of $X_{i,t}$ which satisfies the equation

$$\mathrm{d}X^c_{i,t} = \left(r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i - \frac{1}{2}v_{i,t}\right)\mathrm{d}t + \sqrt{v_{i,t}}\,\mathrm{d}\tilde{W}_{i,t}. \tag{2.39}$$

At this juncture, we derive the joint characteristic function of $X^c_{1,t}$ and $X^c_{2,t}$ and, eventually, the joint characteristic function of the log-prices.

Lemma 2.9. *Suppose $\rho_w = \rho_z = 0$. The joint characteristic function $\phi_{X_t}(u_1, u_2)$ of the log-price vector $X_t = (X_{1,t}, X_{2,t})^\top$ is given by*

$$\begin{aligned}\phi_{X_t}(u_1, u_2) = \exp\left\{i(u_1x_1 + u_2x_2) + C(t; u_1, u_2) + D_1(t; u_1)v_1 + D_2(t; u_2)v_2\right\} \\ \times \exp\left\{\tilde{\lambda}_1 t\left(\phi_{Y_1}(u_1) - 1\right) + \tilde{\lambda}_2 t\left(\phi_{Y_2}(u_2) - 1\right)\right\}\end{aligned} \tag{2.40}$$

where

$$D_j(s; u_j) = -\frac{u_j^2 + iu_j}{\gamma_j\coth(\gamma_j s/2) + (\xi_j + \lambda_j) - iu_j\rho_{wz_j}\sigma_j}, \qquad j = 1, 2, \tag{2.41}$$

where γ_j is given by

$$\gamma_j = \sqrt{\sigma_j^2(u_j^2 + iu_j) + (\xi_j + \Lambda_j - iu_j\rho_{wz_j}\sigma_j)^2}, \qquad j = 1, 2,$$

and

$$\begin{aligned}C(s; u_1, u_2) = \sum_{j=1}^{2}\Bigg\{iu_j s(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j) + \frac{\xi_j\eta_j s(\xi_j + \lambda_j - iu_j\rho_{wz_j}\sigma_j)}{\sigma_j^2} \\ -\frac{2\xi_j\eta_j}{\sigma_j^2}\ln\left(\cosh\frac{\gamma_j s}{2} + \frac{\xi_j + \Lambda_j - iu_j\rho_{wz_j}\sigma_j}{\gamma}\sinh\frac{\gamma_j s}{2}\right)\Bigg\}.\end{aligned} \tag{2.42}$$

Proof. We first determine the characteristic function of the continuous parts of the log-prices. Define the function f as

$$f(t, x_1, x_2, v_1, v_2) = \mathbb{E}_{\mathbb{Q}}\left\{e^{i(u_1X^c_{1,T} + u_2X^c_{2,T})}|X^c_{1,t} = x_1, X^c_{2,t} = x_2, v_{1,t} = v_1, v_{2,t} = v_2\right\}, \tag{2.43}$$

and let $\mathfrak{M}_t = f(t, X^c_{1,t}, X^c_{2,t}, v_{1,t}, v_{2,t})$. In the next calculation, we do not yet invoke the assumption

that $\rho_w = \rho_z = 0$. By Itô's Lemma, $\mathfrak{M}_t$ satisfies

$$\begin{aligned}
\mathrm{d}\mathfrak{M}_t = \Bigg\{ & \frac{\partial f}{\partial t} + \sum_{i=1}^{2} \left(r - q_i - \tilde{\lambda}_i \tilde{\kappa}_i - \frac{1}{2} v_{i,t} \right) \frac{\partial f}{\partial x_i} + \sum_{i=1}^{2} \left[\xi_i \eta_i - (\xi_i + \Lambda_i) v_{i,t} \right] \frac{\partial f}{\partial v_i} \\
& + \frac{1}{2} \sum_{i=1}^{2} v_{i,t} \frac{\partial^2 f}{\partial x_i^2} + \frac{1}{2} \sum_{i=1}^{2} \sigma_i^2 v_{i,t} \frac{\partial^2 f}{\partial v_i^2} + \rho_w \sqrt{v_{1,t} v_{2,t}} \frac{\partial^2 f}{\partial x_1 \partial x_2} \\
& + \sum_{i=1}^{2} \rho_{w z_i} \sigma_i v_{i,t} \frac{\partial^2 f}{\partial x_i \partial v_i} + \rho_z \sigma_1 \sigma_2 \sqrt{v_{1,t} v_{2,t}} \frac{\partial^2 f}{\partial v_1 \partial v_2} \Bigg\} \mathrm{d}t \\
& + \sum_{i=1}^{2} \sqrt{v_{i,t}} \frac{\partial f}{\partial x_i} \mathrm{d}\tilde{W}_{i,t} + \sum_{i=1}^{2} \sigma_i \sqrt{v_{i,t}} \frac{\partial f}{\partial v_i} \mathrm{d}\tilde{Z}_{i,t}.
\end{aligned}$$

Since $\mathfrak{M}_t$ is a martingale, the drift coefficient is equal to zero, which leads to the equation

$$\begin{aligned}
0 = & \frac{\partial f}{\partial t} + \sum_{i=1}^{2} \left(r - q_i - \tilde{\lambda}_i \tilde{\kappa}_i - \frac{1}{2} v_{i,t} \right) \frac{\partial f}{\partial x_i} + \sum_{i=1}^{2} \left[\xi_i \eta_i - (\xi_i + \Lambda_i) v_{i,t} \right] \frac{\partial f}{\partial v_i} \\
& + \frac{1}{2} \sum_{i=1}^{2} v_{i,t} \frac{\partial^2 f}{\partial x_i^2} + \frac{1}{2} \sum_{i=1}^{2} \sigma_i^2 v_{i,t} \frac{\partial^2 f}{\partial v_i^2} + \rho_w \sqrt{v_{1,t} v_{2,t}} \frac{\partial^2 f}{\partial x_1 \partial x_2} \\
& + \sum_{i=1}^{2} \rho_{w z_i} \sigma_i v_{i,t} \frac{\partial^2 f}{\partial x_i \partial v_i} + \rho_z \sigma_1 \sigma_2 \sqrt{v_{1,t} v_{2,t}} \frac{\partial^2 f}{\partial v_1 \partial v_2}.
\end{aligned}$$

The corresponding terminal condition is given by

$$f(T, x_1, x_2, v_1, v_2) = e^{i(u_1 x_1 + u_2 x_2)}.$$

This PDE is to be solved for $t \in [0, T]$, $(x_1, x_2) \in \mathbb{R}^2$, and $(v_1, v_2) \in \mathbb{R}_+^2$, where $\mathbb{R}_+^2 = (0, \infty) \times (0, \infty)$.

To be able to solve the above PDE explicitly, we now assume that $\rho_w = \rho_z = 0$. Making this assumption simplifies our model (2.37) and (2.38) to the class of affine jump-diffusions (Duffie et al., 2000). The resulting equation given the the simplifying assumptions is

$$\begin{aligned}
0 = & \frac{\partial f}{\partial t} + \sum_{i=1}^{2} \left(r - q_i - \tilde{\lambda}_i \tilde{\kappa}_i - \frac{1}{2} v_{i,t} \right) \frac{\partial f}{\partial x_i} + \sum_{i=1}^{2} \left[\xi_i \eta_i - (\xi_i + \Lambda_i) v_{i,t} \right] \frac{\partial f}{\partial v_i} && (2.44) \\
& + \frac{1}{2} \sum_{i=1}^{2} v_{i,t} \frac{\partial^2 f}{\partial x_i^2} + \frac{1}{2} \sum_{i=1}^{2} \sigma_i^2 v_{i,t} \frac{\partial^2 f}{\partial v_i^2} + \sum_{i=1}^{2} \rho_{w z_i} \sigma_i v_{i,t} \frac{\partial^2 f}{\partial x_i \partial v_i}, && (2.45)
\end{aligned}$$

and to solve the equation, we guess a solution of the form

$$f(t, x_1, x_2, v_1, v_2) = \exp\left\{ i(u_1 x_1 + u_2 x_2) + C(T - t) + D_1(T - t) v_1 + D_2(T - t) v_2 \right\}, \tag{2.46}$$

for some functions $C(s)$, $D_1(s)$, and $D_2(s)$ of one variable evaluated at $s = T - t$ (Heston, 1993; Cont and Tankov, 2004a; Cane and Olivares, 2014). This representation is consistent with the

general form of the characteristic function of affine jump-diffusions (Duffie et al., 2000). At this point, we suppress the t subscript and introduce the notation $f_t = f(t, x_1, x_2, v_1, v_2)$. With this specification, the PDE becomes

$$\begin{aligned}
0 = f_t &\left[-C'(T-t) - D_1'(T-t)v_1 - D_2'(T-t)v_2\right] \\
&+ \left(r - q_1 - \tilde{\lambda}_1\tilde{\kappa}_1 - \frac{1}{2}v_1\right) f_t i u_1 + \left(r - q_2 - \tilde{\lambda}_2\tilde{\kappa}_2 - \frac{1}{2}v_2\right) f_t i u_2 \\
&+ [\xi_1\eta_1 - (\xi_1 + \Lambda_1)v_1] f_t D_1(T-t) + [\xi_2\eta_2 - (\xi_2 + \Lambda_2)v_2] f_t D_2(T-t) \\
&- \frac{1}{2}v_1 u_1^2 f_t - \frac{1}{2}v_2 u_2^2 f_t + \frac{1}{2}\sigma_1^2 v_1 D_1^2(T-t) f_t + \frac{1}{2}\sigma_2^2 v_2 D_2^2(T-t) f_t \\
&+ \rho_{wz_1}\sigma_1 v_1 i u_1 D_1(T-t) f_t + \rho_{wz_2}\sigma_2 v_2 i u_2 D_2(T-t) f_t,
\end{aligned}$$

where $C'(T-t)$, $D_1'(T-t)$, and $D_2'(T-t)$ are the first derivatives of $C(s)$, $D_1(s)$, and $D_2(s)$ evaluated at $s = T - t$. Simplifying and collecting the coefficients of v_1 and v_2, we have

$$\begin{aligned}
0 = v_1 &\left\{-D_1'(T-t) - \frac{iu_1}{2} - \frac{u_1^2}{2} + \frac{1}{2}\sigma_1^2 D_1^2(T-t) + (iu_1\rho_{wz_1}\sigma_1 - \xi_1 - \Lambda_1)D_1(T-t)\right\} \\
&+ v_2 \left\{-D_2'(T-t) - \frac{iu_2}{2} - \frac{u_2^2}{2} + \frac{1}{2}\sigma_2^2 D_2^2(T-t) + (iu_2\rho_{wz_2}\sigma_2 - \xi_2 - \Lambda_2)D_2(T-t)\right\} \\
&- C'(T-t) + (r - q_1 - \tilde{\lambda}_1\tilde{\kappa}_1)iu_1 + (r - q_2 - \tilde{\lambda}_2\tilde{\kappa}_2)iu_2 + \xi_1\eta_1 D_1(T-t) + \xi_2\eta_2 D_2(T-t).
\end{aligned}$$

Since v_1 and v_2 are nonzero, it follows that $C(s)$, $D_1(s)$, and $D_2(s)$ must satisfy the equations

$$D_1'(s) = \frac{\sigma_1^2}{2}D_1^2(s) + (iu_1\rho_{wz_1}\sigma_1 - \xi_1 - \Lambda_1)D_1(s) - \frac{1}{2}(iu_1 + u_1^2), \tag{2.47}$$

$$D_2'(s) = \frac{\sigma_2^2}{2}D_2^2(s) + (iu_2\rho_{wz_2}\sigma_2 - \xi_2 - \Lambda_2)D_2(s) - \frac{1}{2}(iu_2 + u_2^2), \tag{2.48}$$

$$C'(s) = (r - q_1 - \tilde{\lambda}_1\tilde{\kappa}_1)iu_1 + (r - q_2 - \tilde{\lambda}_2\tilde{\kappa}_2)iu_2 + \xi_1\eta_1 D_1(s) + \xi_2\eta_2 D_2(s), \tag{2.49}$$

subject to the initial condition $D_1(0) = D_2(0) = C(0) = 0$. Adapting the results of Heston (1993), Bates (1996), and Cont and Tankov (2004a), $D_1(s)$ and $D_2(s)$ are given by

$$D_j(s) = -\frac{u_j^2 + iu_j}{\gamma_j \coth(\gamma_j s/2) + (\xi_j + \lambda_j) - iu_j\rho_{wz_j}\sigma_j}, \qquad j = 1, 2,$$

where γ_j is given by

$$\gamma_j = \sqrt{\sigma_j^2(u_j^2 + iu_j) + (\xi_j + \Lambda_j - iu_j\rho_{wz_j}\sigma_j)^2}, \qquad j = 1, 2.$$

Integration for $C(s)$ yields

$$C(s) = \sum_{j=1}^{2} \left\{ iu_j s(r - q_j - \tilde{\lambda}_j \tilde{\kappa}_j) + \frac{\xi_j \eta_j s(\xi_j + \lambda_j - iu_j \rho_{wz_j} \sigma_j)}{\sigma_j^2} \right.$$
$$\left. - \frac{2\xi_j \eta_j}{\sigma_j^2} \ln \left(\cosh \frac{\gamma_j s}{2} + \frac{\xi_j + \Lambda_j - iu_j \rho_{wz_j} \sigma_j}{\gamma} \sinh \frac{\gamma_j s}{2} \right) \right\}.$$

Note that

$$f(0, x_1, x_2, v_1, v_2) = \mathbb{E}_{\mathbb{Q}} \left[e^{i(u_1 X_{1,T}^c + u_2 X_{2,T}^c)} \middle| X_{1,0}^c = x_1, X_{2,0}^c = x_2, v_{1,0} = v_1, v_{2,0} = v_2 \right]$$
$$= \mathbb{E}_{\mathbb{Q}} \left[e^{i(u_1 X_{1,T}^c + u_2 X_{2,T}^c)} \right].$$

If x_1 and x_2 represent the log of initial stock prices and v_1 and v_2 represent initial volatility levels, then from the assumed functional form of f, it follows that

$$\mathbb{E}_{\mathbb{Q}} \left[e^{i(u_1 X_{1,T}^c + u_2 X_{2,T}^c)} \right] = \exp \left\{ i(u_1 x_1 + u_2 x_2) + C(T) + D_1(T) v_1 + D_2(T) v_2 \right\},$$

where $C(T)$, $D_1(T)$, and $D_2(T)$ are given as above. Thus for any $0 < t \leq T$, the joint characteristic function of the continuous part of the log-asset prices, which we denote by $\phi_{X_t^c}(u_1, u_2)$ is given by

$$\phi_{X_t^c}(u_1, u_2) = \exp \left\{ i(u_1 x_1 + u_2 x_2) + C(t; u_1, u_2) + D_1(t; u_1, u_2) v_1 + D_2(t; u_1, u_2) v_2 \right\} \tag{2.50}$$

We now turn to the joint characteristic function $\phi_{Q_t}(u_1, u_2)$ of the jump parts. From the assumptions on the SVJD model, we note that the jump components of stock 1 and stock 2 are independent, hence

$$\phi_{Q_t}(u_1, u_2) = \mathbb{E}_{\mathbb{Q}} \left[\exp \left(iu_1 \sum_{n=1}^{N_{1,t}} Y_{1,n} + iu_2 \sum_{n=1}^{N_{2,t}} Y_{2,n} \right) \right]$$
$$= \mathbb{E}_{\mathbb{Q}} \left[\exp \left(iu_1 \sum_{n=1}^{N_{1,t}} Y_{1,n} \right) \right] \mathbb{E}_{\mathbb{Q}} \left\{ \exp \left(iu_2 \sum_{n=1}^{N_{2,t}} Y_{2,n} \right) \right]$$

If $\phi_{Y_i}(\cdot)$ denotes the common characteristic function of the $Y_{i,n}$'s under the risk-neutral measure $\mathbb{Q}$, then we have

$$\phi_{Q_t}(u_1, u_2) = \exp \left\{ \tilde{\lambda}_1 t \left(\phi_{Y_1}(u_1) - 1 \right) + \tilde{\lambda}_2 t \left(\phi_{Y_2}(u_2) - 1 \right) \right\}, \tag{2.51}$$

where $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ are the intensities of $N_{1,t}$ and $N_{2,t}$, respectively, under $\mathbb{Q}$.

Following the independence of the continuous and jump parts of the log-price processes, the joint

characteristic function of the log-prices $X_{1,t}$ and $X_{2,t}$, which we denote by $\phi_{X_t}(u_1, u_2)$ is therefore given by

$$\begin{aligned}\phi_{X_t}(u_1,u_2) = \exp\left\{i(u_1x_1+u_2x_2) + C(t;u_1,u_2) + D_1(t;u_1,u_2)v_1 + D_2(t;u_1,u_2)v_2\right\}\\ \times \exp\left\{\tilde{\lambda}_1 t\left(\phi_{Y_1}(u_1)-1\right) + \tilde{\lambda}_2 t\left(\phi_{Y_2}(u_2)-1\right)\right\}.\end{aligned}$$

This is defined for all $(x_1, x_2) \in \mathbb{R}^2$, $(v_1, v_2) \in \mathbb{R}^2_+$, $t \in [0, T]$, and $u_1, u_2 \in \mathbb{C}$. ■

Let $\phi_{X_T|t}(u_1, u_2) = \mathbb{E}_{\mathbb{Q}}\left[e^{i(u_1X_{1,T}+u_2X_{2,T})}\middle|\mathcal{F}_t\right]$. From the above calculations, we have

$$\begin{aligned}&\phi_{X_T|t}(u_1,u_2)\\ &\quad= \exp\left\{i(u_1x_1+u_2x_2) + C(T-t;u_1,u_2) + D_1(T-t;u_1,u_2)v_1 + D_2(T-t;u_1,u_2)v_2\right\}\\ &\qquad\times \exp\left\{\tilde{\lambda}_1(T-t)\left(\phi_{Y_1}(u_1)-1\right) + \tilde{\lambda}_2(T-t)\left(\phi_{Y_2}(u_2)-1\right)\right\},\end{aligned} \tag{2.52}$$

where x_1, x_2, v_1, and v_2 now denote time t values of the processes $\{X_{1,t}\}$, $\{X_{2,t}\}$, $\{v_{1,t}\}$, and $\{v_{2,t}\}$, respectively.

At this point, we briefly state the results of Caldana and Fusai (2013) on the approximate pricing of European spread options under general stock price dynamics, as this will be used to derive European exchange option prices under the assumed SVJD model. Consider a European option with strike price K written on the spread $S_{1,T} - S_{2,T}$. Then the price today of the spread option is $C_{K,0} = e^{-rT}\mathbb{E}_{\mathbb{Q}}\left[(S_{1,T} - S_{2,T} - K)^+\right]$. For some parameters α and k, define the event

$$A = \left\{S_{1,T} > \frac{e^k S_{2,T}}{\mathbb{E}_{\mathbb{Q}}[S^{\alpha}_{2,T}]}\right\}.$$

This event, as noted by Bjerskund and Stensland (2011), is a sub-optimal exercise strategy and produces the following lower bound on the European spread option payoff,

$$(S_{1,T} - S_{2,T} - K)^+ \geq (S_{1,T} - S_{2,T} - K)\mathbf{1}_A. \tag{2.53}$$

It follows therefore that a lower bound for the spread option price is given by

$$C^{k,\alpha}_{K,0} = e^{-rT}\mathbb{E}_{\mathbb{Q}}\left[(S_{1,T} - S_{2,T} - K)\mathbf{1}_A\right]. \tag{2.54}$$

The following theorem (Caldana and Fusai, 2013, Proposition 1) provides an expression for $C^{k,\alpha}_{K,0}$ in terms of the joint characteristic function $\phi_{X_T}(u_1, u_2)$of the log-price vector $X_T = (X_{1,T}, X_{2,T})^\top$.

Theorem 2.10. *The lower bound $C_{K,0}^{k,\alpha}$ for the European spread option price is given by*

$$C_{K,0}^{k,\alpha} = \left\{ \frac{e^{-\delta k - rT}}{\pi} \int_0^\infty e^{-izk} \Psi_T(z, \delta, \alpha)\, \mathrm{d}z \right\}^+, \tag{2.55}$$

where

$$\begin{aligned} \psi_T(z, \delta, \alpha) = \frac{e^{i(z-i\delta)\ln \phi_{X_T}(0,-i\alpha)}}{i(z-i\delta)} \Big[& \phi_{X_T}\left(z - i\delta - i, -\alpha(z-i\delta)\right) \\ & - \phi_{X_T}\left(z - i\delta, -\alpha(z-i\delta) - i\right) - K\phi_{X_T}\left(z - i\delta, -\alpha(z-i\delta)\right) \Big], \end{aligned} \tag{2.56}$$

and

$$\alpha = \frac{\mathbb{E}_{\mathbb{Q}}[S_{2,T}]}{\mathbb{E}_{\mathbb{Q}}[S_{2,T}] + K}, \qquad k = \ln\left(\mathbb{E}_{\mathbb{Q}}[S_{2,T}] + K\right). \tag{2.57}$$

The preceding theorem requires the existence of the joint characteristic function of the log-prices of the two stocks. In our model, we have shown that the characteristic function ϕ_{X_t} is given by equation (2.40), and hence the above result can be applied in approximating European spread option prices and pricing the European exchange option under our SVJD model.

We make some remarks on the other quantities that appear in the formula. First, δ is associated to the exponentially decaying term $e^{-\delta k}$ that must be multiplied to $C_{K,0}^{k,\alpha}$ to produce a square-integrable term in the negative k-axis. A discussion of the choice of δ can be found in Carr and Madan (1999) and Dempster and Hong (2002). Next, since the discounted yield process $\{e^{-(r-q_2)t} S_{2,t}\}$ is a martingale under $\mathbb{Q}$, $\mathbb{E}_{\mathbb{Q}}[S_{2,T}]$ represents the forward price today of asset 2 for delivery at time T. In terms of the characteristic function, the forward price can also be expressed as $\phi_{X_T}(0, -i)$. The positive part $(\cdot)^+$ is required since, without it, the formula (as well as the original Bjerskund and Stensland (2011) result) may produce negative values for deeply out-of-the-money options. Thus for practical purposes, out-of-the-money exchange options are assigned a value of 0 (Caldana and Fusai, 2013). Note that this consideration is consistent with equation (2.24), since if the option is out-of-the-money, then $\mathbf{1}_{\mathcal{A}} = 0$, which results to $C_0^E = 0$.

Caldana and Fusai (2013) note that this result is an improvement from the Hurd and Zhou (2010) Fourier inversion formula since the exchange option case ($K = 0$) can be handled here without complications. In this regard, the Caldana and Fusai (2013) formula also gives an *exact* price for the exchange option through an appropriate choice of the parameters k and α (which will be discussed next). Furthermore, the integration in the formula above involves a univariate Fourier inversion in contrast to the bivariate inversion of Hurd and Zhou (2010), implying that the Caldana and Fusai (2013) result requires less computation time. However, the result of Caldana and Fusai (2013) is a lower bound approximation for the spread option price, in contrast to the exact price derived by Hurd and Zhou (2010). The Hurd and Zhou (2010) formula also does not depend on the decay parameter δ.

In the following proposition we extend the Caldana and Fusai (2013) result in Theorem 2.10 to a version that provides a lower bound for the European spread option price at any time $t \in [0, T)$.

Proposition 2.11. *A lower bound for the time $t \in [0, T)$ price of a European spread option with strike price K is given by*

$$C_{K,t}^{\alpha,k} = \left\{ \frac{e^{-r(T-t)-\delta k}}{\pi} \int_0^\infty e^{-izk} \psi_{T|t}(z, \delta, \alpha)\, \mathrm{d}z \right\}^+, \tag{2.58}$$

where

$$\begin{aligned} \psi_{T|t}(z, \delta, \alpha) = \frac{e^{i(z-i\delta)\ln \phi_{X_T|t}(0,-i\alpha)}}{iz+\delta} \Big[& \phi_{X_T|t}\left(z - i\delta - i, -\alpha(z - i\delta)\right) \\ & - \phi_{X_T|t}\left(z - i\delta, -\alpha(z - i\delta) - i\right) - K\phi_{X_T|t}\left(z - i\delta, -\alpha(z - i\delta)\right) \Big], \end{aligned} \tag{2.59}$$

$\phi_{X_T|t} = \mathbb{E}_{\mathbb{Q}}[e^{iu_1 X_{1,T} + iu_2 X_{2,T}} | \mathcal{F}_t]$, *$\alpha$ and k are given by*

$$\alpha = \frac{F_2(t,T)}{F_2(t,T)+K} \qquad \text{and} \qquad k = \ln\left[F_2(t,T) + K\right], \tag{2.60}$$

and $F_2(t, T)$ is the time t forward price of asset 2 with for delivery at time T.

Proof. The proof we present follows the outline of Theorem 2.10 in Appendix A of Caldana and Fusai (2013). Recall that the time t price of the European spread option is given by

$$e^{-r(T-t)} \mathbb{E}_{\mathbb{Q}} \left[(S_{1,T} - S_{2,T} - K)^+ \middle| \mathcal{F}_t \right].$$

To obtain an approximation for the time t price, we define the event A_t as

$$A_t = \left\{ S_{1,T} \geq \frac{e^k S_{2,T}^\alpha}{\mathbb{E}_{\mathbb{Q}}[S_{2,T}^\alpha | \mathcal{F}_t]} \right\}, \tag{2.61}$$

where α and k are defined as above. This is similar to the event A defined by Caldana and Fusai (2013) (see discussion preceding Theorem 2.10), except that the expectation is conditional on $\mathcal{F}_t$ and forward prices are taken at time t. Following the argument of Bjerskund and Stensland (2011), the quantity

$$C_{K,t}^{\alpha,k} = e^{-r(T-t)} \mathbb{E}_{\mathbb{Q}} \left[(S_{1,T} - S_{2,T} - K) \mathbf{1}_{A_t} \middle| \mathcal{F}_t \right] \tag{2.62}$$

is a lower bound for the true option price at time t.

At this point, we rewrite some of the quantities above in terms of the notation established before. First, we note that since the discounted yield process $\{e^{-(r-q_2)t} S_{2,t}\}$ is a $\mathbb{Q}$-martingale, the forward price $F_2(t, T)$ can be expressed as $\mathbb{E}_{\mathbb{Q}}[S_{2,T} | \mathcal{F}_t]$. In terms of the conditional characteristic function $\phi_{X_T|t}$ (which, under the SVJD model, is given by equation (2.52)), the quantity $\mathbb{E}_{\mathbb{Q}}[S_{2,T}^\alpha | \mathcal{F}_t]$ is given

by $\phi_{X_T|t}(0,-i\alpha)$. Furthermore, we may rewrite the exercise strategy A_t in terms of log-prices as

$$A_t = \left\{ X_{1,T} - \alpha X_{2,T} + \ln \phi_{X_T|t}(0,-i\alpha) > k \right\}. \tag{2.63}$$

Now, we seek to express $C_K^{\alpha,k}(t)$ as a Fourier inversion formula following the proof in Caldana and Fusai (2013) (see Appendix A of their paper). To this end, let δ be a positive number and define $\psi_{T|t}$ as

$$\psi_{T|t}(z,\delta,\alpha) = \int_{-\infty}^{\infty} e^{izk+\delta k} \mathbb{E}_{\mathbb{Q}} \left[(S_{1,T} - S_{2,T} - K)\mathbf{1}_{A_t} \middle| \mathcal{F}_t \right] \mathrm{d}k.$$

In other words, $\psi_{T|t}$ is the Fourier transform of $e^{\delta k}\mathbb{E}_{\mathbb{Q}}[(S_{1,T} - S_{2,T} - K)\mathbf{1}_{A_t}|\mathcal{F}_t]$ in the k-variable. Let $g(x_1,x_2)$ be the transition density function of the log-price vector X_T. Thus, the $\psi_{T|t}$ may be evaluated as the triple integral

$$\psi_{T|t}(z,\delta,\alpha) = \int_{-\infty}^{\infty} e^{izk+\delta k} \left[\int_{-\infty}^{\infty} \int_{\beta}^{\infty} \left(e^{x_1} - e^{x_2} - K\right) g(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \right] \mathrm{d}k,$$

where $\beta = k + \alpha x_2 - \ln \phi_{X_T|t}(0,-i\alpha)$. Let $\beta' = x_1 - \alpha x_2 + \ln \phi_{X_T|t}(0,-i\alpha)$. Then, after changing the order of integration, we have the following:

$$\begin{aligned}
\psi_{T|t}(z,\delta,\alpha) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\int_{-\infty}^{\beta'} e^{izk+\delta k}\,\mathrm{d}k \right] \left(e^{x_1} - e^{x_2} - K\right) g(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{e^{i(z-i\delta)\beta'}}{iz+\delta} \left(e^{x_1} - e^{x_2} - K\right) g(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\
&= \frac{e^{i(z-i\delta)\ln \phi_{X_T|t}(0,-i\alpha)}}{iz+\delta} \\
&\quad \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(z-i\delta)(x_1-\alpha x_2)} \left(e^{x_1} - e^{x_2} - K\right) g(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\
&= \frac{e^{i(z-i\delta)\ln \phi_{X_T|t}(0,-i\alpha)}}{iz+\delta} \\
&\quad \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{ i(z-i\delta-i)x_1 - i\alpha(z-i\delta)x_2 \right\} g(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \right. \\
&\quad - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{ i(z-i\delta)x_1 - i(\alpha z - i\alpha\delta + i)x_2 \right\} g(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\
&\quad \left. - K \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{ i(z-i\delta)x_1 - i\alpha(z-i\delta)x_2 \right\} g(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \right]
\end{aligned}$$

The preceding double integrals may be written in terms of $\phi_{X_T|t}$, which thus results to

$$\begin{aligned}
\psi_{T|t}(z,\delta,\alpha) = \frac{e^{i(z-i\delta)\ln \phi_{X_T|t}(0,-i\alpha)}}{iz+\delta} &\left[\phi_{X_T|t}\left(z-i\delta-i, -\alpha(z-i\delta)\right) \right. \\
&\left. - \phi_{X_T|t}\left(z-i\delta, -\alpha(z-i\delta)-i\right) - K\phi_{X_T|t}\left(z-i\delta, -\alpha(z-i\delta)\right) \right]
\end{aligned}$$

We then invert the Fourier transform to obtain

$$e^{\delta k}\mathbb{E}_{\mathbb{Q}}[(S_{1,T}-S_{2,T}-K)\mathbf{1}_{A_t}|\mathcal{F}_t] = \frac{1}{\pi}\int_0^\infty e^{-izk}\psi_{T|t}(z,\delta,\alpha)\,\mathrm{d}z.$$

It follows therefore that

$$C_K^{\alpha,k}(t) = \left\{\frac{e^{-r(T-t)-\delta k}}{\pi}\int_0^\infty e^{-izk}\psi_{T|t}(z,\delta,\alpha)\,\mathrm{d}z\right\}^+.$$

We require the positive part in the preceding formula to avoid negative prices for deeply-out-of-money options (Caldana and Fusai, 2013). As in Bjerskund and Stensland (2011) and Caldana and Fusai (2013) α and k may be chosen as

$$\alpha = \frac{F_2(t,T)}{F_2(t,T)+K} \qquad \text{and} \qquad k = \ln\left[F_2(t,T)+K\right].$$

■

Theorem 2.10 follows naturally by setting $t=0$.

We recall the approximate strategy A_t introduced in the proof of Proposition 2.11. If we set $K=0$ (as is the case for the exchange option) in the expressions for α and k, we obtain $\alpha=1$ and $k=\ln F_2(t,T)$. Since $F_2(t,T)=\mathbb{E}_{\mathbb{Q}}[S_{2,T}|\mathcal{F}_t]$, setting $K=0$ causes A_t to coincide with the true exercise strategy $B=\{S_{1,T}\geq S_{2,T}\}$ for the European exchange option. Thus, similar to the findings of Caldana and Fusai (2013), $C_{K,t}^{\alpha,k}$ (or $C_{K,0}^{\alpha,k}$ in Theorem 2.10) is exact for the European exchange option. The following corollary provides the time t price of the European exchange option.

Corollary 2.12. *The time t price of the European exchange option is given by*

$$C_t^E = \left\{\frac{e^{-r(T-t)-\delta k}}{\pi}\int_0^\infty e^{-izk}\psi_{T|t}(z,\delta,1)\,\mathrm{d}z\right\}^+, \tag{2.64}$$

where

$$\psi_{T|t}(z,\delta,1) = \frac{e^{i(z-i\delta)k}}{iz+\delta}\left[\phi_{X_T|t}\left(z-i\delta-i,-z+i\delta\right)-\phi_{X_T|t}\left(z-i\delta,-z+i\delta-i\right)\right].$$

In the analysis that follows, we focus on the price of the European exchange option at time 0. Suppose the exchange option is *not* deeply out-of-the-money (i.e. $C_0^E\geq 0$). Then the following proposition expresses the Caldana and Fusai (2013) result into a form that is consistent with what we obtained in equation (2.33) from the change-of-numéraire technique.

Proposition 2.13. *The time* 0 *price of the European exchange option is given by*

$$\begin{aligned} C_0^E = S_{1,0}e^{-q_1 T}\int_0^\infty \frac{1}{\pi(iz+\delta)}\phi_{X_T}^{(1)}(iz+\delta,-iz-\delta)\,\mathrm{d}z \\ - S_{2,0}e^{-q_2 T}\int_0^\infty \frac{1}{\pi(iz+\delta)}\phi_{X_T}^{(2)}(iz+\delta,-iz-\delta)\,\mathrm{d}z. \end{aligned} \tag{2.65}$$

where $\phi_{X_T}^{(1)}$ *and* $\phi_{X_T}^{(2)}$ *denote the joint characteristic function of* $X_T = (X_{1,T}, X_{2,T})^\top$ *under* $\hat{\mathbb{Q}}_1$ *and* $\hat{\mathbb{Q}}_2$ *(defined in Section 2.5.1), respectively.*

Proof. Since we assume that the option price is nonnegative, we may remove the positive part in equation (2.64) (with $t = 0$) and write

$$C_0^E = e^{-rT}\int_0^\infty \frac{e^{-(iz+\delta)k}}{\pi}\psi_T(z,\delta,1)\,\mathrm{d}z.$$

Substituting the expression for ψ_T, we have

$$\begin{aligned} C_0^E = e^{-rT}\Bigg\{\int_0^\infty \frac{e^{-(iz+\delta)k+(iz+\delta)k}}{\pi(iz+\delta)}\phi_{X_T}\left(z-i\delta-i,-z+i\delta\right)\mathrm{d}z \\ - \int_0^\infty \frac{e^{-(iz+\delta)k+(iz+\delta)k}}{\pi(iz+\delta)}\phi_{X_T}\left(z-i\delta,-z+i\delta-i\right)\mathrm{d}z\Bigg\}. \end{aligned}$$

We also note that

$$\begin{aligned} \phi_{X_T}(z-i\delta-i,-z+i\delta) &= \mathbb{E}_{\mathbb{Q}}\left[S_{1,T}\exp\left\{(iz+\delta)X_{1,T}+(-iz-\delta)X_{2,T}\right\}\right], \\ \phi_{X_T}(z-i\delta,-z+i\delta-i) &= \mathbb{E}_{\mathbb{Q}}\left[S_{2,T}\exp\left\{(iz+\delta)X_{1,T}+(-iz-\delta)X_{2,T}\right\}\right]. \end{aligned}$$

But note that for $i = 1, 2$, the terminal stock prices can be written as $S_{i,T} = S_{i,0}e^{(r-q_i)T}U_{i,T}$, where $U_{i,T}$ is defined as in equation (2.26). We can thus write

$$\begin{aligned} \phi_{X_T}(z-i\delta-i,-z+i\delta) &= S_{1,0}e^{(r-q_1)T}\mathbb{E}_{\mathbb{Q}}\left[U_{1,T}\exp\left\{(iz+\delta)X_{1,T}+(-iz-\delta)X_{2,T}\right\}\right], \\ \phi_{X_T}(z-i\delta,-z+i\delta-i) &= S_{2,0}e^{(r-q_2)T}\mathbb{E}_{\mathbb{Q}}\left[U_{2,T}\exp\left\{(iz+\delta)X_{1,T}+(-iz-\delta)X_{2,T}\right\}\right]. \end{aligned}$$

Using these expressions for ϕ_{X_T}, the European exchange option price is therefore given by

$$\begin{aligned} C_0^E = S_{1,0}e^{-q_1 T}\int_0^\infty \frac{1}{\pi(iz+\delta)}\mathbb{E}_{\mathbb{Q}}\left[U_{1,T}\exp\left\{(iz+\delta)X_{1,T}+(-iz-\delta)X_{2,T}\right\}\right]\mathrm{d}z \\ - S_{2,0}e^{-q_2 T}\int_0^\infty \frac{1}{\pi(iz+\delta)}\mathbb{E}_{\mathbb{Q}}\left[U_{2,T}\exp\left\{(iz+\delta)X_{1,T}+(-iz-\delta)X_{2,T}\right\}\right]\mathrm{d}z. \end{aligned}$$

Recall from Section 2.5.1 that $U_{1,T}$ and $U_{2,T}$ are Radon-Nikodým derivatives that define new probability measures $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$ equivalent to $\mathbb{Q}$. Therefore, the expectations that appear above define joint characteristic functions of the log-prices under these new probability measures. The

desired result thus follows from replacing the expectations with $\phi^{(1)}_{X_T}$ and $\phi^{(2)}_{X_T}$, the characteristic function of $X_T = (X_{1,T}, X_{2,T})^\top$ under $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$ respectively. ■

Under the SVJD specification, the joint characteristic functions $\phi^{(1)}_{X_T}$ and $\phi^{(2)}_{X_T}$ have forms similar to ϕ_{X_T} as provided in equation (2.40), with some slight changes in the values of the parameters due to the change of measure. Properties of log-prices after the change of measure follow from the discussion in Section 2.5.1, under the assumption that $\rho_w = \rho_z = 0$ in addition to Assumption 2.2. As such, $\phi^{(1)}_{X_T}$ and $\phi^{(2)}_{X_T}$ may be determined in a manner similar to the calculations presented in the first part of this section.

The above calculations therefore show that the Caldana and Fusai (2013) result, when applied to European exchange options, allows for a decomposition similar to equation (2.33), which was obtained via the change-of-numéraire technique. This therefore presents the possibility that the probabilities $\hat{\mathbb{Q}}_1(\mathcal{A}_0)$ and $\hat{\mathbb{Q}}_2(\mathcal{A}_0)$ (the probability of the option being in-the-money under the alternative measures $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$) may be computed using Fourier inversion in equation (2.65).

The same analysis can be applied to produce a similar representation for the time t price of the European exchange option, as shown in the next proposition.

Proposition 2.14. *The time t European exchange option price is given by*

$$\begin{aligned} C^E_t = S_{1,t} e^{-q_1(T-t)} \int_0^\infty \frac{1}{\pi(iz+\delta)} \phi^{(1)}_{X_T|t}(iz+\delta, -iz-\delta)\,\mathrm{d}z \\ - S_{2,T} e^{-q_2(T-t)} \int_0^\infty \frac{1}{\pi(iz+\delta)} \phi^{(2)}_{X_T|t}(iz+\delta, -iz-\delta)\,\mathrm{d}z, \end{aligned} \tag{2.66}$$

where

$$\phi^{(i)}_{X_T|t}(u_1, u_2) = \mathbb{E}_{\hat{\mathbb{Q}}_i}\left[e^{iu_1 X_{1,T} + iu_2 X_{2,T}} \middle| \mathcal{F}_t \right], \qquad i = 1, 2$$

gives the joint conditional characteristic function of $X_T = (X_{1,T}, X_{2,T})^\top$ under $\hat{\mathbb{Q}}_i$, and $\hat{\mathbb{Q}}_1$ and $\hat{\mathbb{Q}}_2$ are the probability measures equivalent to $\mathbb{Q}$ discussed in Section 2.5.1.

Proof. The proof of this proposition is similar to that of Proposition 2.13, except that the expectations involved are replaced by their conditional counterparts (see Corollary 2.12). ■

Because of the equivalence between equations (2.36) and (2.66), it follows that equation (2.66) characterizes a solution to the IPDE in Proposition 2.8 with terminal condition $C^E_T = (S_{1,T} - S_{2,T})^+$.

In summary, the European exchange option price under the SVJD model may be represented by equation (2.33), which is a formula that resembles the original Margrabe (1978) result under the Black-Scholes framework. It is also notable that our results in equations (2.33) and (2.65) do not contain the risk-free rate r, similar to the Margrabe (1978) formula. Due to the addition of stochastic volatilities, a closed formula for the probabilities in equation (2.33) cannot be obtained. However, by making additional assumptions on the correlation structure of the Wiener processes

involved, we were able to obtain equation (2.65), a Fourier inversion formula for the European exchange option price (following the work of Caldana and Fusai (2013) and Cane and Olivares (2014)). Lastly, by extending the analysis of Caldana and Fusai (2013) to obtain time-t option prices, we were able to obtain a solution C_t^E of IPDE (2.8), for the case of the European exchange option, in terms of Fourier inversion formulas.

2.6 A Representation of the American Exchange Option Price

In this section, we show that the price of the American exchange option, under our SVJD model for underlying stock prices, can be decomposed into a sum of the European exchange option price and the early exercise premium. As will be shown, the early exercise premium can be decomposed into a premium arising from the diffusion part of the asset dynamics and premia arising from the possibility of sudden jumps in the asset prices.

Before we derive the early exercise representation of the American exchange option price under the SVJD model, we first discuss several monotonicity and convexity properties of the American exchange option price and the geometry of the so-called early exercise or stopping region. Discussing these properties also leads to the characterization of the American exchange option price as the solution of a free boundary problem or as the solution of a variational inequality.

2.6.1 Properties of American Exchange Options under SVJD Dynamics

To better depict the dependence of the American exchange option price on the current values of the asset prices and variances, we denote by $S_i^{(t,s_i,v_i)}(u)$ the value of the solution of (2.14) at time $u \geq t$ when $S_i(t) = s_i$ and $v_i(t) = v_i$, $i = 1, 2$, i.e.[4]

$$S_i^{(t,s_i,v_i)}(u) = s_i \exp\left\{(r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i)(u-t) - \frac{1}{2}\int_t^u V_i(s)\,\mathrm{d}s \right. \\ \left. + \int_t^u \sqrt{V_i(s)}\,\mathrm{d}W_i(s) + \sum_{n=N_i(t)+1}^{N_i(u)} Y_{i,n}\right\} \tag{2.67}$$

Furthermore, we denote by $C^A(t, s_1, s_2, v_1, v_2)$ the price at time t of the American exchange option, which we write as

$$C^A(t, s_1, s_2, v_1, v_2) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}}\left[e^{-r(\tau-t)}(S_1^{(t,s_1,v_1)}(\tau) - S_2^{(t,s_2,v_2)}(\tau))^+\right], \tag{2.68}$$

following the notation of Pham (1997).

The following proposition states monotonicty and convexity properties of C^A in t, s_1, and s_2.

[4] This is the notation typically used to denote the unique solution of stochastic differential equations (see for example Øksendal, 2013, Chapter 7).

Proposition 2.15. *The American exchange option price C^A, defined in (2.68), satisfies the following properties:*

1. *For fixed $(t, v_1, v_2) \in [0,T] \times \mathbb{R}^2_+$, $C^A(t, s_1, s_2, v_1, v_2)$ is nondecreasing in s_1 (for fixed $s_2 \in \mathbb{R}_+$) and non-increasing in s_2 (for fixed $s_1 \in \mathbb{R}_+$).*
2. *For fixed $(t, v_1, v_2) \in [0,T] \times \mathbb{R}^2_+$, $C^A(t, s_1, s_2, v_1, v_2)$ is convex in (s_1, s_2).*
3. *For fixed $(s_1, s_2, v_1, v_2) \in \mathbb{R}^4_+$, $C^A(t, s_1, s_2, v_1, v_2)$ is non-increasing in t.*

Proof. To show monotonicity in s_1, we fix $(t, v_1, v_2) \in [0,T] \times \mathbb{R}^2_+$ and let $s_1 \geq s_1$. From the pathwise solution (2.67), it follows that $S_1^{(t,s_1,v_1)}(\tau) \geq S_1^{(t,s_1',v_1)}(\tau)$ for any $\tau \in \mathcal{T}_{t,T}$. For fixed $s_2 \in \mathbb{R}_+$, we then have

$$e^{-r(\tau-t)}(S_1^{(t,s_1,v_1)}(\tau) - S_2^{(t,s_2,v_2)}(\tau))^+ \geq e^{-r(\tau-t)}(S_1^{(t,s_1',v_1)}(\tau) - S_2^{(t,s_2,v_2)}(\tau))^+.$$

Taking expectations and the supremum over $\tau \in \mathcal{T}_{t,T}$, we obtain

$$C^A(t, s_1, s_2, v_1, v_2) \geq C^A(t, s_1', s_2, v_1, v_2).$$

The same argument can be used to show that C^A is non-increasing in s_2.

To establish the convexity of the option price, we first show that the payoff function $g(s_1, s_2) = (s_1 - s_2)^+$ is convex. Let $\lambda \in [0,1]$ and consider two points $(s_1, s_2), (s_1', s_2') \in \mathbb{R}^2_+$. Using the property $(a+b)^+ \leq a^+ + b^+$ for any $a, b \in \mathbb{R}$, we find that

$$\begin{aligned}
g(\lambda s_1 + (1-\lambda)s_1', \lambda s_2 + (1-\lambda)s_2') &= \big[\lambda s_1 + (1-\lambda)s_1' - \lambda s_2 - (1-\lambda)s_2'\big]^+ \\
&= \big[\lambda(s_1 - s_2) + (1-\lambda)(s_1' - s_2')\big]^+ \\
&\leq \big[\lambda(s_1 - s_2)\big]^+ + \big[(1-\lambda)(s_1' - s_2')\big]^+ \\
&= \lambda(s_1 - s_2)^+ + (1-\lambda)(s_1' - s_2')^+ \\
&= \lambda g(s_1, s_2) + (1-\lambda) g(s_1', s_2').
\end{aligned}$$

From (2.67), we also note that

$$S_i^{(t,\lambda s_i + (1-\lambda)s_i', v_i)}(\tau) = \lambda S_i^{(t,s_i,v_i)}(\tau) + (1-\lambda) S_i^{(t,s_i',v_i)}(\tau).$$

Therefore,

$$\begin{aligned}
&C^A(t, \lambda s_1 + (1-\lambda)s_1', \lambda s_2 + (1-\lambda)s_2', v_1, v_2) \\
&\quad = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau-t)} (S_1^{(t,\lambda s_1 + (1-\lambda)s_1', v_1)}(\tau) - S_2^{(t,\lambda s_2 + (1-\lambda)s_2', v_2)}(\tau))^+ \right]
\end{aligned}$$

$$
\begin{aligned}
&= \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}}\Big[e^{-r(\tau-t)}(\lambda S_1^{(t,s_1,v_1)}(\tau) + (1-\lambda)S_1^{(t,s_1',v_1)}(\tau) \\
&\qquad\qquad -\lambda S_2^{(t,s_2,v_1)}(\tau) - (1-\lambda)S_2^{(t,s_2',v_1)}(\tau))^+\Big] \\
&\leq \lambda \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}}\left[e^{-r(\tau-t)}(S_1^{(t,s_1,v_1)}(\tau) - S_2^{(t,s_2,v_2)}(\tau))^+\right] \\
&\quad + (1-\lambda) \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}}\left[e^{-r(\tau-t)}(S_1^{(t,s_1',v_1)}(\tau) - S_2^{(t,s_2',v_2)}(\tau))^+\right] \\
&= \lambda C^A(t,s_1,s_2,v_1,v_2) + (1-\lambda)C^A(t,s_1',s_2',v_1,v_2).
\end{aligned}
$$

To show that the option price is non-increasing in t, we consider a change of variables to the time-to-maturity $u = T - t$. We can thus write the option price as

$$
\hat{C}^A(u,s_1,s_2,v_1,v_2) = \sup_{\tau\in\mathcal{T}_{0,u}} \mathbb{E}_{\mathbb{Q}}\left[e^{-r\tau}(S_1^{(0,s_1,v_1)}(\tau) - S_2^{(0,s_2,v_2)}(\tau))^+\right].
$$

We note that if $u \leq u'$, then $\mathcal{T}_{0,u} \subseteq \mathcal{T}_{0,u'}$, and so

$$
\begin{aligned}
\hat{C}^A(u,s_1,s_2,v_1,v_2) &\leq \sup_{\tau\in\mathcal{T}_{0,u'}} \mathbb{E}_{\mathbb{Q}}\left[e^{-r\tau}(S_1^{(0,s_1,v_1)}(\tau) - S_2^{(0,s_2,v_2)}(\tau))^+\right] \\
&= \hat{C}^A(u',s_1,s_2,v_1,v_2),
\end{aligned}
$$

that is, $\hat{C}^A$ is nondecreasing in u. Therefore, C^A is non-increasing in t. ■

Standard theory on American option pricing (Broadie and Detemple, 1997; Myneni, 1992; Pham, 1997, in particular) indicates that the domain $[0,T]\times\mathbb{R}_+^4$ for the American exchange option price C^A can be split into two regions: the *stopping (or early exercise) region* $\mathcal{S}$ and the *continuation region* $\mathcal{C}$. These are given by

$$
\begin{aligned}
\mathcal{S} &= \{(t,s_1,s_2,v_1,v_2) \in [0,T]\times\mathbb{R}_+^4 : C^A(t,s_1,s_2,v_1,v_2) = (s_1-s_2)^+\} \\
\mathcal{C} &= \{(t,s_1,s_2,v_1,v_2) \in [0,T]\times\mathbb{R}_+^4 : C^A(t,s_1,s_2,v_1,v_2) > (s_1-s_2)^+\}.
\end{aligned}
\tag{2.69}
$$

We refer to the shared boundary of the stopping and continuation regions, i.e. $\partial\mathcal{S}\cap\partial\mathcal{C}$, the *early exercise boundary* of the American exchange option. An optimal stopping time for the American exchange option problem is therefore the first time (on or after time t) the asset prices (and the variances) enter the stopping region, i.e.

$$
\begin{aligned}
\tau^* &= \tau^*(t,s_1,s_2,v_1,v_2) \\
&\equiv \inf\{u\in[t,T] : (u, S_1^{(t,s_1,v_1)}(u), S_2^{(t,s_2,v_2)}(u), V_1^{(t,v_1)}(u), V_2^{(t,v_2)}(u)) \in \mathcal{S}\},
\end{aligned}
$$

provided that the option has not yet been exercised at or before time t.

Under the risk-neutral measure $\mathbb{Q}$, the discounted American exchange option price is a martin-

gale in the continuation region, i.e. the process

$$\{e^{-r(u-t)}C^A(u, S_1^{(t,s_1,v_1)}(u), S_2^{(t,s_2,v_2)}(u), V_1^{(t,v_1)}(u), V_2^{(t,v_2)}(u))\}_{u\in[t,\tau^*]}$$

is a $\mathbb{Q}$-martingale. Therefore, the C^A satisfies the equation

$$\bar{\mathcal{L}}C^A(t, s_1, s_2, v_1, v_2) = 0, \qquad \forall (t, s_1, s_2, v_1, v_2) \in \mathcal{C}, \tag{2.70}$$

where $\bar{\mathcal{L}}$ is the integro-partial differential operator

$$\bar{\mathcal{L}}f = -rf + \frac{\partial f}{\partial u} + \mathcal{L}_{\mathbf{s},\mathbf{v}}f, \qquad \text{for } f \text{ satisfying Assumption 2.7, and } \mathcal{L}_{\mathbf{s},\mathbf{v}} \text{ in (2.21)}.$$

The function C^A is also smooth in the continuation region (Pham, 1997).

Following Broadie and Detemple (1997), we describe the geometry of the stopping region for the American exchange option. This in turn shows that for fixed $(t, v_1, v_2) \in [0, T] \times \mathbb{R}^2_+$ the early exercise boundary is a ray in the s_1s_2-plane. For the succeeding analysis, we define $R = s_1/s_2$ (or $R(t) = S_1(t)/S_2(t)$ when it is to be interpreted as a stochastic processes) to be the asset price ratio and note that the payoff can be written as $(s_1 - s_2)^+ = s_2(R - 1)^+$. We also note that (t, s_1, s_2, v_1, v_2) is short-hand for $(t, S_1(t), S_2(t), v_1(t), v_2(t))$.

Proposition 2.16. *The stopping region $\mathcal{S}$ satisfies the following properties:*

1. $(t, s_1, s_2, v_1, v_2) \in \mathcal{S}$ *implies* $R > 1$
2. $(t, s_1, s_2, v_1, v_2) \in \mathcal{S}$ *implies* $(t, \lambda s_1, s_2, v_1, v_2) \in \mathcal{S}$ *for any* $\lambda \geq 1$
3. $(t, s_1, s_2, v_1, v_2) \in \mathcal{S}$ *implies* $(t, \lambda s_1, \lambda s_2, v_1, v_2) \in \mathcal{S}$ *for any* $\lambda > 0$
4. $s_2 = 0$ *implies that immediate exercise is optimal at time t for all* $s_1 > 0$

The second assertion is referred to as the *right-connectedness* property of the stopping region, while the third assertion is the *ray connectedness* property. This terminology and its connection to the geometry of the stopping region shall be discussed after the proof of the proposition. The ray connectedness property is specific to the American exchange option, in contrast to American spread options (Broadie and Detemple, 1997).

Proof. For the first three assertions, we assume that $(t, s_1, s_2, v_1, v_2) \in \mathcal{S}$.

1. Suppose $R \leq 1$. Then the payoff from immediate exercise is $s_2(R - 1)^+ = 0$, so it is not optimal to exercise the option now in favor of a *waiting or continuation policy* with a positive probability of giving a strictly positive payoff at some fixed future date. This contradicts the assumption that $(t, s_1, s_2, v_1, v_2) \in \mathcal{S}$.

2. Let $\lambda \geq 1$. Suppose that $(t, \lambda s_1, s_2, v_1, v_2) \notin \mathcal{S}$, i.e.

$$C^A(t, \lambda s_1, s_2, v_1, v_2) > (\lambda s_1 - s_2)^+.$$

Starting from the definition of the American exchange option price and the pathwise solution (2.67), we have the following calculations:

$$\begin{aligned}
C^A(t, \lambda s_1, s_2, v_1, v_2) &= \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} (S_1^{(t, \lambda s_1, v_1)}(\tau) - S_2^{(t, s_2, v_2)}(\tau))^+ \right] \\
&= \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} (\lambda S_1^{(t, s_1, v_1)}(\tau) - S_2^{(t, s_2, v_2)}(\tau))^+ \right] \\
&= \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} S_2^{(t, s_2, v_2)}(\tau) (\lambda R(\tau) - 1)^+ \right] \\
&\qquad \text{(where } R(\tau) = S_1^{(t, s_1, v_1)}(\tau) / S_2^{(t, s_2, v_2)}(\tau)) \\
&= \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} S_2^{(t, s_2, v_2)}(\tau) (R(\tau) - 1 + (\lambda - 1) R(\tau) - 1)^+ \right] \\
&\leq \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} S_2^{(t, s_2, v_2)}(\tau) (R(\tau) - 1)^+ \right] \\
&\qquad + (\lambda - 1) \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} S_2^{(t, s_2, v_2)}(\tau) R(\tau) \right] \\
&= C^A(t, s_1, s_2, v_1, v_2) + (\lambda - 1) \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} S_1^{(t, s_1, v_1)}(\tau) \right].
\end{aligned}$$

Under $\mathbb{Q}$, the discounted asset yield process $\{e^{-(r - q_1)t} S_1(t)\}_{t \in [0,T]}$ is a martingale, so using the Markov property of $(S_1, v_1)^\top$,

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} S_1^{(t, s_1, v_1)}(\tau) \right] &= \mathbb{E}_{\mathbb{Q}} \left[e^{-(r - q_1)\tau} S_1^{(t, s_1, v_1)}(\tau) \right] e^{rt} e^{-q_1 \tau} \\
&= \mathbb{E}_{\mathbb{Q}} \left[e^{-(r - q_1)\tau} S_1(\tau) | S_1(t) = s_1, v_1(t) = v_1 \right] e^{rt} e^{-q_1 \tau} \\
&= e^{-(r - q_1)t} s_1 e^{rt} e^{-q_1 \tau} \\
&= e^{-q_1(\tau - t)} s_1 \\
&\leq s_1.
\end{aligned}$$

Therefore, we have

$$C^A(t, \lambda s_1, s_2, v_1, v_2) \leq C^A(t, s_1, s_2, v_1, v_2) + (\lambda - 1) s_1.$$

However, by assumption, $(t, s_1, s_2, v_1, v_2) \in \mathcal{S}$, so in conjunction with the first assertion of the proposition, we have

$$C^A(t, \lambda s_1, s_2, v_1, v_2) \leq (s_1 - s_2)^+ + (\lambda - 1) s_1$$

$$
\begin{aligned}
&= s_1 - s_2 + (\lambda - 1)s_1 \\
&= \lambda s_1 - s_2 \\
&\leq (\lambda s_1 - s_2)^+.
\end{aligned}
$$

This contradicts our assumption at the start of this proof.

3. Let $\lambda > 0$ and suppose $(t, \lambda s_1, \lambda s_2, v_1, v_2) \notin \mathcal{S}$, i.e.

$$C^A(t, \lambda s_1, \lambda s_2, v_1, v_2) > (\lambda s_1 - \lambda s_2)^+ = \lambda (s_1 - s_2)^+.$$

Since we have

$$
\begin{aligned}
C^A(t, \lambda s_1, \lambda s_2, v_1, v_2) &= \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} (\lambda S_1^{(t,s_1,v_1)}(\tau) - \lambda S_2^{(t,s_2,v_2)}(\tau))^+ \right] \\
&= \lambda \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} (S_1^{(t,s_1,v_1)}(\tau) - S_2^{(t,s_2,v_2)}(\tau))^+ \right] \\
&= \lambda C^A(t, s_1, s_2, v_1, v_2),
\end{aligned}
$$

it follows from the inequality above that

$$C^A(t, s_1, s_2, v_1, v_2) > (s_1 - s_2)^+,$$

which contradicts the assumption that $(t, s_1, s_2, v_1, v_2) \in \mathcal{S}$.

4. If $s_2 = 0$, then from (2.67) we have $S_2^{(t,s_2,v_2)}(\tau)$ for any $\tau \in \mathcal{T}_{t,T}$. Thus, if $s_1 > 0$, then

$$(S_1^{(t,s_1,v_1)}(\tau) - S_2^{(t,s_2,v_2)}(\tau))^+ = S_1^{(t,s_1,v_1))}(\tau) \qquad \forall \tau \in \mathcal{T}_{t,T}.$$

Using the supermartingale property of $\{e^{-r(u-t)} S_1^{(t,s_1,v_1)}(u)\}_{u \in [t,T]}$ (see the proof for the second assertion), we then have

$$C^A(t, s_1, 0, v_1, v_2) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{\mathbb{Q}} \left[e^{-r(\tau - t)} S_1^{(t,s_1,v_1)}(\tau) \right] \leq \sup_{\tau \in \mathcal{T}_{t,T}} s_1 = s_1.$$

Since $C^A(t, s_1, s_2, v_1, v_2) \geq (s_1 - s_2)^+$ for any $(t, s_1, s_2, v_1, v_2) \in [0, T] \times \mathbb{R}_+^4$, it follows that

$$C^A(t, s_1, 0, v_1, v_2) \geq (s_1 - 0)^+ = s_1.$$

Therefore,

$$C^A(t, s_1, 0, v_1, v_2) = s_1 = (s_1 - 0)^+,$$

i.e. $(t, s_1, 0, v_1, v_2) \in \mathcal{S}$ for any $s_1 > 0$.

■

Fix $(t, v_1, v_2) \in [0, T] \times \mathbb{R}^2_+$ and let $\mathcal{S}_{t,v_1,v_2}$ and $\mathcal{C}_{t,v_1,v_2}$ be the projection of the stopping and continuation regions, respectively, onto $\mathbb{R}^2_+$ where only (s_1, s_2) can vary (see Figure 2.1). The first assertion of Proposition 2.16 implies the inclusion

$$\mathcal{S}_{t,v_1,v_2} \subseteq \{(s_1, s_2) \in \mathbb{R}^2_+ : s_1 > s_2\} = \{(s_1, s_2) \in \mathbb{R}^2_+ : s_1/s_2 = R > 1\}.$$

The second and third assertions of the proposition implies that the (shared) boundary of $\mathcal{S}_{t,v_1,v_2}$ and $\mathcal{C}_{t,v_1,v_2}$ has the form

$$\partial\mathcal{S}_{t,v_1,v_2} \cap \partial\mathcal{C}_{t,v_1,v_2} = \{(s_1, s_2) \in \mathbb{R}^2_+ : s_1 = Bs_2\}$$

for some constant $B \geq 1$ (the critical price ratio) since:

1. if $(s_1, s_2) \in \mathcal{S}_{t,v_1,v_2}$, then any other point on the ray connecting the origin (in $\mathbb{R}^2$) to (s_1, s_2) must also be in $\mathcal{S}_{t,v_1,v_2}$ by the third assertion. Note that elements of this ray are given by $(\lambda s_1, \lambda s_2)$ for $\lambda > 0$.

2. if $(s_1, s_2) \in \mathcal{S}_{t,v_1,v_2}$, then any other point to the right, i.e. any point $(\lambda s_1, s_2)$ for $\lambda \geq 1$, must also be in $\mathcal{S}_{t,v_1,v_2}$ by the second assertion. Combining this with the previous item, any other point to the right of the ray described in the previous item must also be in $\mathcal{S}_{t,v_1,v_2}$.

These two points imply that $\mathcal{S}_{t,v_1,v_2}$ has the form $\mathcal{S}_{t,v_1,v_2} = \{(s_1, s_2) \in \mathbb{R}^2_+ : s_1 \geq Bs_2\}$.

Letting (t, v_1, v_2) vary, the critical price ratio may then depend on t, v_1, and v_2, so the stopping region can be expressed as

$$\mathcal{S} = \{(t, s_1, s_2, v_1, v_2) \in [0, T] \times \mathbb{R}^4_+ : s_1 \geq B(t, v_1, v_2)s_2\}$$

for some function $B : [0, T] \times \mathbb{R}^2_+$ satisfying $B(t, v_1, v_2) \geq 1$ for all (t, v_1, v_2). Since the decision to exercise an exchange option early ultimately depends on the ratio of the asset prices, it is reasonable to assume that the variances v_1 and v_2 only affect the early exercise decision via the the critical price ratio. This assumption is consistent with the properties proved by Touzi (1999) for single-asset American put options under stochastic volatility dynamics. The dependence of the early exercise boundary of the variance is also made clear by the numerical experiments carried out in Chapter 5 (see Figure 5.3).

From (2.70), it follows that C^A is the solution of the *free boundary problem*

$$rC^A = \frac{\partial C^A}{\partial t} + \mathcal{L}_{\mathbf{s},\mathbf{v}} C^A, \tag{2.71}$$

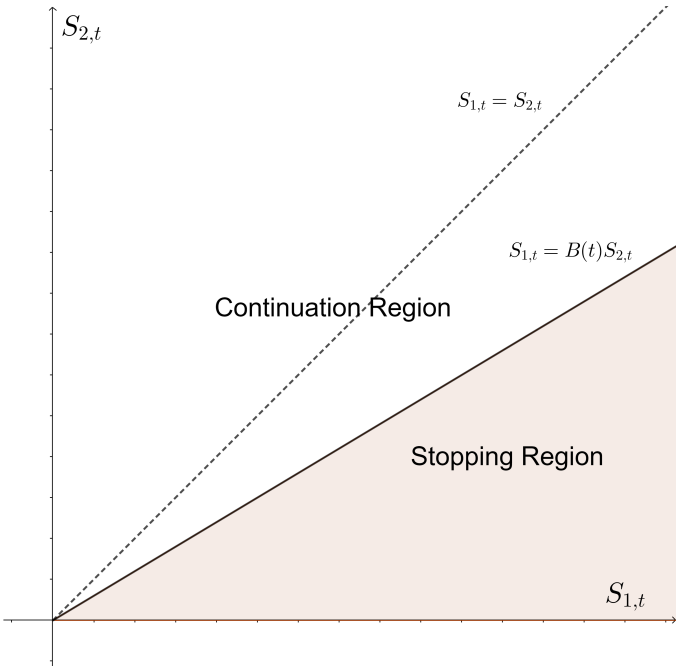


Figure 2.1: The early exercise boundary and the continuation and stopping regions for the American exchange option. Here, $B(t)$ is shorthand for $B(t, v_1, v_2)$.

which is to be solved for $0 < s_1 < B(t, v_1, v_2)s_2$, and $(t, s_2, v_1, v_2) \in [0, T] \times \mathbb{R}^3_+$. The associated terminal and boundary conditions for the free boundary problem are

$$\begin{aligned} C^A(T, s_1, s_2, v_1, v_2) &= (s_1 - s_2)^+, \qquad (s_1, s_2, v_1, v_2) \in \mathbb{R}^4_+ \\ C^A(t, 0, s_2, v_1, v_2) &= 0, \qquad (t, s_2, v_1, v_2) \in [0, T] \times \mathbb{R}^3_+ \\ C^A(t, s_1, 0, v_1, v_2) &= s_1, \qquad (t, s_1, v_1, v_2) \in [0, T] \times \mathbb{R}^3_+. \end{aligned} \tag{2.72}$$

To pin down the unknown early exercise boundary, represented here by $B(t, v_1, v_2)$, and to ensure the free boundary problem is well-posed, we impose additional conditions on the free boundary problem. These conditions, known as the *value-matching* and the *smooth-pasting* conditions, directly result from the differentiability assumptions on the American exchange option price made in Assumption 2.7.[5] The value-matching condition, given by

$$C^A(t, s_1, s_2, v_1, v_2) = s_1 - s_2, \qquad \text{for } s_1 \geq B(t, v_1, v_2)s_2,\ (t, s_2, v_1, v_2) \in [0, T] \times \mathbb{R}^3_+, \tag{2.73}$$

[5] Without making the assumption that the exchange option price is smooth (at least up until the second-order for the spatial variables), the continuity of the first-order partial derivatives can be established via a viscosity solution approach of the pricing IPDE, as was done by Pham (1997) and Touzi (1999) (among others) for the American put option under jump-diffusion and stochastic volatility, respectively. The American (exchange) option pricing problem is a special case of the optimal stopping problem considered by Pham (1998) for a controlled jump-diffusion process. Pham (1998) establishes that the value function of the optimal stopping problem is a viscosity solution of the corresponding Hamilton-Jacobi-Bellman variational inequality. Using the regularity properties of viscosity solutions, Pham (1997) and Touzi (1999) prove that the American option price has a continuous partial derivative in s and, when the value-matching condition is used, the smooth-pasting condition holds. See also Cheang and Chiarella (2011) for a discussion of the smooth-pasting conditions for the American exchange option under jump-diffusion dynamics.

is consistent with the assumption that C^A is continuous in s_1 and s_2 across the early exercise boundary. In addition, the smooth-pasting conditions, given by

$$\begin{aligned} \lim_{s_1 \to B(t,v_1,v_2)s_2} \frac{\partial C^A}{\partial s_1} &= 1, \\ \lim_{s_1 \to B(t,v_1,v_2)s_2} \frac{\partial C^A}{\partial s_2} &= -1, \end{aligned} \tag{2.74}$$

is consistent with the assumption that the first-order partial derivatives of C^A in s_1 and s_2 are also continuous across the early exercise boundary. We note that as a consequence of Assumption 2.7 and the structure of the payoff of the exchange option, we also have

$$\lim_{s_1 \to B(t,v_1,v_2)s_2} \frac{\partial C^A}{\partial v_1} = \lim_{s_1 \to B(t,v_1,v_2)s_2} \frac{\partial C^A}{\partial v_2} = \lim_{s_1 \to B(t,v_1,v_2)s_2} \frac{\partial C^A}{\partial t} = 0$$

although these are not necessary to close the free boundary problem.

The smooth-pasting conditions also have a financial interpretation related to the no-arbitrage assumption. One interpretation is that these conditions allow the investor to continuously adjust the hedging portfolio across the early exercise boundary without any additional costs for transitions through the boundary (Myneni, 1992, Remark 4.2). Another is that the smooth-pasting conditions are a consequence of investors selecting their exercise strategy so as to maximize the value of the American option (Chiarella et al., 2015, p. 5). As such, the smooth-pasting conditions do not allow investors to take advantage of any discontinuities in the option price or the hedge ratios by eliminating such discontinuities. The smooth-pasting conditions are therefore consistent with the no-arbitrage assumption.

The American exchange option problem may also be expressed in terms of a *variational inequality* involving the operator $\bar{\mathcal{L}}$. The advantage of this formulation is that it does not explicitly contain the early exercise boundary. The basis of the variational inequality formulation is the fact that in the stopping region $\mathcal{S}$, the discounted American exchange option price is a strict supermartingale (Broadie and Detemple, 1997; Cheang and Chiarella, 2011), which implies that

$$\bar{\mathcal{L}} C^A(t, s_1, s_2, v_1, v_2) < 0, \qquad (t, s_1, s_2, v_1, v_2) \in \mathcal{S},$$

while in the continuation region $\mathcal{C}$, we have (2.70). In addition, we also know from (2.69) that $C^A(t, s_1, s_2, v_1, v_2) = (s_1 - s_2)^+$ in $\mathcal{S}$. Since (t, s_1, s_2, v_1, v_2) is always either in $\mathcal{C}$ or $\mathcal{S}$, we therefore have

$$\min\left\{-\bar{\mathcal{L}} C^A(t, s_1, s_2, v_1, v_2), C^A(t, s_1, s_2, v_1, v_2) - (s_1 - s_2)^+\right\} = 0, \tag{2.75}$$

for all $(t, s_1, s_2, v_1, v_2) \in [0, T] \times \mathbb{R}^4_+$. The variational inequality (2.75) is supplemented by two

additional conditions

$$\bar{\mathcal{L}}C^A(t, s_1, s_2, v_1, v_2) \leq 0, \qquad C^A(t, s_1, s_2, v_1, v_2) \geq (s_1 - s_2)^+ \tag{2.76}$$

for all $(t, s_1, s_2, v_1, v_2) \in [0, T] \times \mathbb{R}^4_+$. Inequalities (2.75) and (2.76) are also collectively known as a *linear complementarity problem* and can be solved numerical using the projected successive overrelaxation (PSOR) method (see Section 5.3). Once the option price has been solved, the early exercise boundary can be obtained by finding the boundary of the set

$$\{(t, s_1, s_2, v_1, v_2) \in [0, T] \times \mathbb{R}^4_+ : C^A(t, s_1, s_2, v_1, v_2) > s_1 - s_2\}.$$

Our derivation of the early exercise representation for the American exchange option price in the following section does not directly deal with either the free boundary problem or the variational inequality for the American exchange option. Indeed, we shall revisit these formulations when we discuss various numerical methods to solve for exchange option prices under the SVJD model (see Chapter 5). However, our next analysis requires knowledge of the stopping and continuation regions and the behavior of the option price across the early exercise boundary.

2.6.2 The Early Exercise Representation under SVJD Dynamics

In this section, we revert to the notation where $C^A_t = C^A_t(S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t})$ denotes the price of the American exchange option price at time t and $C^A_{t-} = C^A_t(S_{1,t-}, S_{2,t-}, v_{1,t}, v_{2,t})$ is the American exchange option price prior to any jumps at time t. To derive the early exercise representation of C^A, we employ a series of probabilistic arguments in conjunction with the known behavior of the option price in the stopping and continuation regions defined in the previous section. The following proposition shows that the American exchange option can be decomposed as the sum of the price of the corresponding European exchange option and an early exercise premium.

Proposition 2.17. *The price of the American exchange option admits the representation*

$$C^A_t(S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t}) = C^E_t(S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t}) + C^P_t(S_{1,t}, S_{2,t}, v_{1,t}, v_{2,t}), \tag{2.77}$$

where C^E_t is the price of the corresponding European exchange option and C^P_t is the early exercise

premium of the American exchange option. The early exercise premium is given by

$$\begin{aligned} C_t^P = & \int_t^T e^{-r(u-t)}\mathbb{E}_{\mathbb{Q}}\left[\left(q_1 S_{1,u-} - q_2 S_{2,u-}\right)\mathbf{1}_{\mathcal{A}_u}\middle|\mathcal{F}_t\right]\mathrm{d}u \\ & - \int_t^T e^{-r(u-t)}\mathbb{E}_{\mathbb{Q}}\left\{\tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_u^A\left(S_{1,u-}e^{Y_1}, S_{2,u-}, v_{1,u}, v_{2,u}\right)\right.\right. \\ & \qquad \left.\left. - \left(S_{1,u-}e^{Y_1} - S_{2,u-}\right)\right]\mathbf{1}_{\mathcal{A}_{1,u}}\middle|\mathcal{F}_t\right\}\mathrm{d}u \\ & - \int_t^T e^{-r(u-t)}\mathbb{E}_{\mathbb{Q}}\left\{\tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_u^A\left(S_{1,u-}, S_{2,u-}e^{Y_2}, v_{1,u}, v_{2,u}\right)\right.\right. \\ & \qquad \left.\left. - \left(S_{1,u-} - S_{2,u-}e^{Y_2}\right)\right]\mathbf{1}_{\mathcal{A}_{2,u}}\middle|\mathcal{F}_t\right\}\mathrm{d}u, \end{aligned} \tag{2.78}$$

where the events $\mathcal{A}_u$, $\mathcal{A}_{1,u}$, and $\mathcal{A}_{2,u}$ are defined as

$$\begin{aligned} \mathcal{A}_u &= \left\{(u, S_{1,u-}, S_{2,u-}, v_{1,u}, v_{2,u}) \in \mathcal{S}\right\} = \left\{S_{1,u-} \geq B_u S_{2,u-}\right\}, \\ \mathcal{A}_{1,u} &= \mathcal{A}_u \cap \left\{S_{1,u-}e^{Y_1} < B_u S_{2,u-}\right\} = \left\{B_u \leq \frac{S_{1,u-}}{S_{2,u-}} < B_u e^{-Y_1}\right\}, \\ \mathcal{A}_{2,u} &= \mathcal{A}_u \cap \left\{S_{1,u-} < B_u S_{2,u-}e^{Y_2}\right\} = \left\{B_u \leq \frac{S_{1,u-}}{S_{2,u-}} < B_u e^{Y_2}\right\}. \end{aligned} \tag{2.79}$$

Here, $B_u = B(u, v_{1,u}, v_{2,u})$ is the slope of the early exercise boundary at time u.

Proof. Using Itô's formula, we obtain the following stochastic differential equation, written in terms of the operator $\mathcal{L}_{\mathbf{s},\mathbf{v}}$ in (2.21), satisfied by the discounted American exchange option price $\tilde{C}_t^A = e^{-rt}C_t^A$,

$$\begin{aligned} \mathrm{d}\tilde{C}_t^A = & e^{-rt}\left\{-rC_{t-}^A + \frac{\partial C_{t-}^A}{\partial t} + \mathcal{L}_{\mathbf{s},\mathbf{v}}\left[C_{t-}^A\right]\right\}\mathrm{d}t \\ & + e^{-rt}\sum_{i=1}^2 \sqrt{v_{i,t}}S_{i,t-}\frac{\partial C_{t-}^A}{\partial s_i}\,\mathrm{d}\tilde{W}_{i,t} + e^{-rt}\sum_{i=1}^2 \sigma_i\sqrt{v_{i,t}}\frac{\partial C_{t-}^A}{\partial v_i}\,\mathrm{d}\tilde{Z}_{i,t} \\ & + e^{-rt}\int_{\mathbb{R}}\left[C_t^A\left(S_{1,t-}e^y, S_{2,t-}, v_{1,t}, v_{2,t}\right) - C_{t-}^A\right]q_1(\mathrm{d}y, \mathrm{d}t) \\ & + e^{-rt}\int_{\mathbb{R}}\left[C_t^A\left(S_{1,t-}, S_{2,t-}e^y, v_{1,t}, v_{2,t}\right) - C_{t-}^A\right]q_2(\mathrm{d}y, \mathrm{d}t), \end{aligned} \tag{2.80}$$

Integrating over $[t, T]$, we obtain

$$\begin{aligned} \tilde{C}_T^A = \tilde{C}_t^A & + \int_t^T e^{-ru}\left\{-rC_{u-}^A + \frac{\partial C_{u-}^A}{\partial u} + \mathcal{L}_{\mathbf{s},\mathbf{v}}\left[C_{u-}^A\right]\right\}\mathrm{d}u \\ & + \int_t^T e^{-ru}\sum_{i=1}^2 \sqrt{v_{i,u}}S_{i,u-}\frac{\partial C_{u-}^A}{\partial s_i}\,\mathrm{d}\tilde{W}_{i,u} + \int_t^T e^{-ru}\sum_{i=1}^2 \sigma_i\sqrt{v_{i,u}}\frac{\partial C_{u-}^A}{\partial v_i}\,\mathrm{d}\tilde{Z}_{i,u} \\ & + \int_t^T e^{-ru}\int_{\mathbb{R}}\left[C_u^A\left(S_{1,u-}e^y, S_{2,u-}, v_{1,u}, v_{2,u}\right) - C_{u-}^A\right]q_1(\mathrm{d}y, \mathrm{d}u) \end{aligned}$$

$$+\int_t^T e^{-ru}\int_{\mathbb{R}}\left[C_u^A\left(S_{1,u-},S_{2,u-}e^y,v_{1,u},v_{2,u}\right)-C_{u-}^A\right]q_2(\mathrm{d}y,\mathrm{d}u).$$

A division by e^{-rt} produces the equation

$$\begin{aligned}\frac{C_T^A}{e^{r(T-t)}} &= C_t^A+\int_t^T e^{-r(u-t)}\left\{-rC_{u-}^A+\frac{\partial C_{u-}^A}{\partial u}+\mathcal{L}_{\mathbf{s},\mathbf{v}}\left[C_{u-}^A\right]\right\}\mathrm{d}u\\ &\quad+\int_t^T e^{-r(u-t)}\sum_{i=1}^2\sqrt{v_{i,u}}S_{i,u-}\frac{\partial C_{u-}^A}{\partial s_i}\,\mathrm{d}\tilde{W}_{i,u}\\ &\quad+\int_t^T e^{-r(u-t)}\sum_{i=1}^2\sigma_i\sqrt{v_{i,u}}\frac{\partial C_{u-}^A}{\partial v_i}\,\mathrm{d}\tilde{Z}_{i,u}\\ &\quad+\int_t^T e^{-r(u-t)}\int_{\mathbb{R}}\left[C_u^A\left(S_{1,u-}e^y,S_{2,u-},v_{1,u},v_{2,u}\right)-C_{u-}^A\right]q_1(\mathrm{d}y,\mathrm{d}u)\\ &\quad+\int_t^T e^{-r(u-t)}\int_{\mathbb{R}}\left[C_u^A\left(S_{1,u-},S_{2,u-}e^y,v_{1,u},v_{2,u}\right)-C_{u-}^A\right]q_2(\mathrm{d}y,\mathrm{d}u).\end{aligned}$$

In terms of the operator $\bar{\mathcal{L}}$, we can simplify the above equation as

$$\begin{aligned}\frac{C_T^A}{e^{r(T-t)}} &= C_t^A+\int_t^T e^{-r(u-t)}\bar{\mathcal{L}}[C_{u-}^A]\,\mathrm{d}u\\ &\quad+\int_t^T e^{-r(u-t)}\sum_{i=1}^2\sqrt{v_{i,u}}S_{i,u-}\frac{\partial C_{u-}^A}{\partial s_i}\,\mathrm{d}\tilde{W}_{i,u}+\int_t^T e^{-r(u-t)}\sum_{i=1}^2\sigma_i\sqrt{v_{i,u}}\frac{\partial C_{u-}^A}{\partial v_i}\,\mathrm{d}\tilde{Z}_{i,u}\\ &\quad+\int_t^T e^{-r(u-t)}\int_{\mathbb{R}}\left[C_u^A\left(S_{1,u-}e^y,S_{2,u-},v_{1,u},v_{2,u}\right)-C_{u-}^A\right]q_1(\mathrm{d}y,\mathrm{d}u)\\ &\quad+\int_t^T e^{-r(u-t)}\int_{\mathbb{R}}\left[C_u^A\left(S_{1,u-},S_{2,u-}e^y,v_{1,u},v_{2,u}\right)-C_{u-}^A\right]q_2(\mathrm{d}y,\mathrm{d}u).\end{aligned}$$

The terms involving the $\mathbb{Q}$-Wiener processes and the $\mathbb{Q}$-compensated counting measures are martingales with zero mean under $\mathbb{Q}$. Hence, taking the conditional expectation of the above equation under $\mathbb{Q}$ with respect to $\mathcal{F}_t$, we obtain

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}C_T^A\middle|\,\mathcal{F}_t\right]=C_t^A+\mathbb{E}_{\mathbb{Q}}\left[\int_t^T e^{-r(u-t)}\bar{\mathcal{L}}[C_{u-}^A]\,\mathrm{d}u\middle|\,\mathcal{F}_t\right]. \tag{2.81}$$

Since the terminal price for both the European and American exchange options are the same, we have the relation

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}C_T^A\middle|\,\mathcal{F}_t\right]=\mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}(S_{1,T}-S_{2,T})^+\middle|\,\mathcal{F}_t\right]=C_t^E,$$

and so therefore we have

$$C_t^A=C_t^E-\mathbb{E}_{\mathbb{Q}}\left[\int_t^T e^{-r(u-t)}\bar{\mathcal{L}}[C_{u-}^A]\,\mathrm{d}u\middle|\,\mathcal{F}_t\right]. \tag{2.82}$$

We then decompose the conditional expectation into integrals over the stopping and continuation regions. In terms of the event $\mathcal{A}_u$ defined in the proposition, we have

$$\begin{aligned} C_t^A &= C_t^E - \int_t^T e^{-r(u-t)}\mathbb{E}_{\mathbb{Q}}\left[\bar{\mathcal{L}}[C_{u-}^A]\mathbf{1}_{\mathcal{A}_u} + \bar{\mathcal{L}}[C_{u-}^A]\mathbf{1}_{\mathcal{A}_u^c}\middle|\mathcal{F}_t\right]\mathrm{d}u \\ &= C_t^E - \int_t^T e^{-r(u-t)}\mathbb{E}_{\mathbb{Q}}\left[\bar{\mathcal{L}}[C_{u-}^A]\mathbf{1}_{\mathcal{A}_u}\middle|\mathcal{F}_t\right]\mathrm{d}u, \end{aligned}$$

since $\bar{\mathcal{L}}C_{u-}^A = 0$ in the continuation region, i.e. when $\mathcal{A}_u$ holds. Suppose $(t, S_{1,t-}, S_{2,t-}, v_{1,t}, v_{2,t}) \in \mathcal{S}$ (i.e. the event $\mathcal{A}_t$ holds). It follows from the value-matching condition (2.73) that

$$C_{t-}^A = C_t^A(S_{1,t-}, S_{2,t-}, v_{1,t}, v_{2,t}) = S_{1,t-} - S_{2,t-},$$

and so

$$\frac{\partial C_{t-}^A}{\partial s_1} = 1, \quad \frac{\partial C_{t-}^A}{\partial s_2} = -1, \quad \frac{\partial C_{t-}^A}{\partial t} = \frac{\partial C_{t-}^A}{\partial v_1} = \frac{\partial C_{t-}^A}{\partial v_2} = 0$$

for $(t, S_{1,t-}, S_{2,t-}, v_{1,t}, v_{2,t}) \in \mathcal{S}$. Consequently, all second-order partial derivatives appearing in $\bar{\mathcal{L}}[C_{t-}^A]$ vanish in the stopping region. This implies that

$$\begin{aligned} \bar{\mathcal{L}}[C_{u-}^A]\mathbf{1}_{\mathcal{A}_u} = &\Big[(r - q_1 - \tilde{\lambda}_1\tilde{\kappa}_1)S_{1,u-}(1) + (r - q_2 - \tilde{\lambda}_2\tilde{\kappa}_2)S_{2,u-}(-1)\Big]\mathbf{1}_{\mathcal{A}_u} \\ &- r\left(S_{1,u-} - S_{2,u-}\right)\mathbf{1}_{\mathcal{A}_u} \\ &+ \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_u^A\left(S_{1,u-}e^{Y_1}, S_{2,u-}, v_{1,u}, v_{2,u}\right) - (S_{1,u-} - S_{2,u-})\right]\mathbf{1}_{\mathcal{A}_u} \\ &+ \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_u^A\left(S_{1,u-}, S_{2,u-}e^{Y_2}, v_{1,u}, v_{2,u}\right) - (S_{1,u-} - S_{2,u-})\right]\mathbf{1}_{\mathcal{A}_u}. \end{aligned}$$

Recalling that $\tilde{\kappa}_i = \mathbb{E}_{\mathbb{Q}}^{Y_i}\left[e^{Y_i} - 1\right]$, the last expression simplifies to

$$\begin{aligned} \bar{\mathcal{L}}[C_{u-}^A]\mathbf{1}_{\mathcal{A}_u} = &-\left[q_1S_{1,u-} - q_2S_{2,u-}\right]\mathbf{1}_{\mathcal{A}_u} \\ &+ \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_u^A\left(S_{1,u-}e^{Y_1}, S_{2,u-}, v_{1,u}, v_{2,u}\right) - \left(S_{1,u-}e^{Y_1} - S_{2,u-}\right)\right]\mathbf{1}_{\mathcal{A}_u} \\ &+ \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_u^A\left(S_{1,u-}, S_{2,u-}e^{Y_2}, v_{1,u}, v_{2,u}\right) - \left(S_{1,u-} - S_{2,u-}e^{Y_2}\right)\right]\mathbf{1}_{\mathcal{A}_u} \end{aligned}$$

It follows that the American exchange option price is given by

$$\begin{aligned} C_t^A = C_t^E &+ \int_t^T e^{-r(u-t)}\mathbb{E}_{\mathbb{Q}}\left[\left(q_1S_{1,u-} - q_2S_{2,u-}\right)\mathbf{1}_{\mathcal{A}_u}\middle|\mathcal{F}_t\right]\mathrm{d}u \\ &- \int_t^T e^{-r(u-t)}\mathbb{E}_{\mathbb{Q}}\left\{\tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C_u^A\left(S_{1,u-}e^{Y_1}, S_{2,u-}, v_{1,u}, v_{2,u}\right) - \left(S_{1,u-}e^{Y_1} - S_{2,u-}\right)\right]\mathbf{1}_{\mathcal{A}_u}\middle|\mathcal{F}_t\right\}\mathrm{d}u \\ &- \int_t^T e^{-r(u-t)}\mathbb{E}_{\mathbb{Q}}\left\{\tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C_u^A\left(S_{1,u-}, S_{2,u-}e^{Y_2}, v_{1,u}, v_{2,u}\right) - \left(S_{1,u-} - S_{2,u-}e^{Y_2}\right)\right]\mathbf{1}_{\mathcal{A}_u}\middle|\mathcal{F}_t\right\}\mathrm{d}u \end{aligned} \tag{2.83}$$

Note that the discounted price of the American exchange option $\tilde{C}_t^A$ is the Snell envelope (the smallest supermartingale majorant) of the discounted intrinsic value $e^{-rt}(S_{1,t} - S_{2,t})^+$, following the optimal stopping arguments for American options (Karatzas, 1988; Myneni, 1992; Cheang and Chiarella, 2011). Therefore, the discounted American exchange option price is always greater than or equal to $e^{-rt}(S_{1,t} - S_{2,t})$, with equality occurring only in the stopping region. In this light, we note that

- if $S_{1,u-}e^{Y_1} \geq B_u S_{2,u-}$ (i.e. when $(S_{1,u-}e^{Y_1}, S_{2,u-}) \in \mathcal{S}$), then

$$C_u^A\left(S_{1,u-}e^{Y_1}, S_{2,u-}, v_{1,u}, v_{2,u}\right) - \left(S_{1,u-}e^{Y_1} - S_{2,u-}\right) = 0$$

- if $S_{1,u-}e^{Y_1} < B_u S_{2,u-}$ (i.e. when $(S_{1,u-}e^{Y_1}, S_{2,u-}) \in \mathcal{C}$), then

$$C_u^A\left(S_{1,u-}e^{Y_1}, S_{2,u-}, v_{1,u}, v_{2,u}\right) - \left(S_{1,u-}e^{Y_1} - S_{2,u-}\right) > 0$$

- if $S_{1,u-} \geq B_u S_{2,u-}e^{Y_2}$ (i.e. when $(S_{1,u-}, S_{2,u-}e^{Y_2}) \in \mathcal{S}$), then

$$C_u^A\left(S_{1,u-}, S_{2,u-}e^{Y_2}, v_{1,u}, v_{2,u}\right) - \left(S_{1,u-} - S_{2,u-}e^{Y_2}\right) = 0$$

- if $S_{1,u-} < B_u S_{2,u-}e^{Y_2}$ (i.e. when $(S_{1,u-}, S_{2,u-}e^{Y_2}) \in \mathcal{C}$), then

$$C_u^A\left(S_{1,u-}, S_{2,u-}e^{Y_2}, v_{1,u}, v_{2,u}\right) - \left(S_{1,u-} - S_{2,u-}e^{Y_2}\right) > 0$$

Thus, in terms of the events $\mathcal{A}_{1,u}$ and $\mathcal{A}_{2,u}$ defined in the proposition, the integrals appearing in (2.83) can be simplified to obtain the early exercise premium in (2.78). ■

The Fourier inversion formulas we obtained for C_t^E in equations (2.64) and (2.66) may be used as an input in equation (2.77). Doing so allows equation (2.78) to be written solely in terms of the unknown early exercise premium C_t^P.

The event $\mathcal{A}_{i,u}$, $i = 1, 2$, represents the event that the pre-jump stock prices $S_{1,u-}$ and $S_{2,u-}$ were in the stopping/early exercise region, but due to a jump in stock i at time u, the post-jump stock prices were sent back to the continuation region. This interpretation is analogous to that offered in Pham (1997) for the single-asset jump-diffusion case and Cheang and Chiarella (2011) for the case of exchange options under jump-diffusion dynamics.

The decomposition offered in the previous proposition is also similar to that in Cheang and Chiarella (2011). We also note that the early exercise premium can also be decomposed into a premium arising from the diffusion of the dynamics (the positive term) and rebalancing costs arising from the possibility that stock prices suddenly jump back into the continuation region

(the negative terms), as was emphasized by Gukhal (2001) and Cheang and Chiarella (2011). In particular, the diffusion term captures the discounted expected value of cash flows due to dividends when asset prices are in the stopping region and the jump terms capture the rebalancing costs incurred by the holder of the American exchange option when a jump instantaneously occurs in the price of either asset, causing the asset prices to jump back into the continuation region immediately after the option is exercised. In this situation, the investor cannot reverse the decision to exercise, since the jumps occur instantaneously, and is therefore vulnerable to the rebalancing cost described earlier.[6] In these computations, we note that the exercise boundary, and consequently the events defined with respect to the exercise boundary, are all dependent on the volatility levels.

From the value-matching condition, we note that

$$C_t^A(B_t S_{2,t}, S_{2,t}, v_{1,t}, v_{2,t}) = S_{2,t}\,(B_t - 1) \tag{2.84}$$

when the stock prices are on the early exercise boundary (i.e. when $S_{1,t} = B_t S_{2,t}$). From the earlier proposition, we may therefore express the early exercise boundary $B_t = B(t, v_{1,t}, v_{2,t})$ as a solution to the equation

$$S_{2,t}(B_t - 1) = C_t^E(B_t S_{2,t}, S_{2,t}, v_{1,t}, v_{2,t}) + C_t^P(B_t S_{2,t}, S_{2,t}, v_{1,t}, v_{2,t}). \tag{2.85}$$

Note however that this equation must solved as a linked system in conjunction with equation (2.77), since the equation for the early exercise boundary involves the (yet unknown) American exchange option price C_t^A.

2.7 Chapter Summary and Conclusions

In this chapter, we have provided an extension to the results of Margrabe (1978) and Cheang and Chiarella (2011) to consider the case where, aside from the presence of jumps, asset prices are also driven by a stochastic volatility process. To facilitate changes of measure from the objective probability measure to other equivalent measures, we introduced a Radon-Nikodým derivative process, which requires some assumptions on the parameters of the volatility processes. From the construction of the Radon-Nikodým derivative, it was noted that equivalent martingale measures are not unique, which therefore can lead to multiple plausible option prices.

Representations for European exchange option prices were derived using two methods. The first method employs the change-of-numéraire technique to obtain a representation that is similar to the classical Margrabe (1978). Alternatively, we considered additional assumptions on the model's correlation structure to allow an explicit form of the joint-characteristic function of the log-prices

[6] A similar phenomenon in the context of consumption-investment problems with transaction costs in a Lévy-driven market is explored in greater technical detail by De Vallière et al. (2016).

of the stocks. This, in turn, enabled us to represent European exchange option prices using in terms of this characteristic function using the results of Caldana and Fusai (2013). We were able to show that the European exchange option price obtained via the Caldana and Fusai (2013) method can also be written in a form that is consistent with the characterization obtained via the change-of-numéraire procedure.

Finally, we demonstrated that the American exchange option price can also be represented as the sum of the price of the corresponding European exchange option price and an early exercise premium, similar to the findings of Broadie and Detemple (1997), Gukhal (2001), and Cheang and Chiarella (2011) in the case of jump-diffusion dynamics. We were also able to show that the early exercise premium can be decomposed into a premium on the diffusive component of asset prices and a premium owing to the possibility of jumps back into the continuation region right before exercise.

The use of a stochastic volatility jump-diffusion model for asset prices indeed allows for a more accurate characterization of asset prices but presents some complications in obtaining option prices. By making additional assumptions on the correlation structure of the market model, we were able to obtain a representation of the European exchange option price in terms of Fourier inversion formulas. The representations we obtained may be numerically evaluated via Monte Carlo simulation or fast Fourier transform methods (see Hurd and Zhou, 2010; Caldana and Fusai, 2013; Cane and Olivares, 2014, for example). Meanwhile, extensions to the numerical methods proposed by Chiarella et al. (2009) and Chiarella and Ziveyi (2011) may be considered in providing a numerical solution for the American exchange option pricing problem. Numerical approaches to pricing exchange options under SVJD dynamics are discussed in Chapter 5.

CHAPTER 3

An Integral Transform Approach to Exchange Option Pricing

3.1 Introduction

We derive integral representations for the price of European and American exchange options when asset prices are modelled as jump-diffusion processes share a single stochastic process. We refer to this model specification the *proportional stochastic volatility and jump-diffusion (SVJD)* model. The use of a single stochastic volatility process for two assets has been considered by Ziveyi (2011) in pricing American spread options under stochastic volatility and by Cane and Olivares (2014) in a Fourier transform approach to pricing European spread options under SVJD dynamics.[1] This model specification is motivated by the possibility that the two assets share a single source of systematic volatility risk, with the jumps capturing individual risk factors that may cause drastic changes in asset prices.

After deriving the option pricing integro-partial differential equation (IPDE) implied by the asset price dynamics, we determine the price of the European exchange option using Fourier and Laplace transforms. We then address the American case by first transforming the homogeneous IPDE, which to be solved over a stock price domain restricted by the early exercise boundary, into an inhomogeneous IPDE, to be solved over an unrestricted stock price domain, using the techniques presented by Jamshidian (1992) and Chiarella and Ziogas (2009). Given the inhomogeneous term of the transformed IPDE and the transition density function of the state variables, we are then able to obtain an integral representation of the early exercise premium associated to the American exchange option. We observe that the early exercise premium we obtain can also be decomposed into a sum of a diffusion part and two jump parts, with financial interpretations consistent with those offered by Gukhal (2001) and Cheang and Chiarella (2011).

[1] We note that Cane and Olivares (2014) take advantage of the affine structure of the model and use the results of Duffie et al. (2000) to derive the characteristic function of the underlying processes.

The methodology presented in this chapter builds on the work of McKean (1965) who used the incomplete Fourier transform to solve the free-boundary problem associated to the American call option. The integral transform approach was also taken by Chiarella and Ziogas (2009) to price American call options under jump-diffusion dynamics, by Chiarella et al. (2010) and Adolfsson et al. (2013) to price American options under the Heston (1993) stochastic volatility model, by Chiarella and Ziveyi (2013) to price American options involving two stochastic volatility processes, and by Cheang et al. (2013) to price European and American call options under stochastic volatility and jump-diffusion dynamics. This approach has also been extended to price multi-asset options. For example, Chiarella and Ziveyi (2014) applied integral transforms to price American spread options under pure-diffusion dynamics. Integral transform techniques have also been employed to solve actuarial valuation problems. For example, Ziveyi et al. (2013) used integral transforms to price European options on deferred annuities under an affine stochastic mortality model and Shen et al. (2016) followed the solution of Chiarella and Ziveyi (2014) for the valuation of guaranteed minimum maturity benefits in variable annuities with an early exercise option.

In the presence of stochastic volatility, a Laplace transform was also applied after a Fourier transform to further simplify the option pricing PDE. Taking the Laplace transform produces a first-order (non-homogenous) PDE in time and volatility, rather than a second-order PDE in time and volatility that results from just applying a Fourier transform. The first-order equation can then be solved for the Laplace transform function via the method of characteristics, and the solution can then be Laplace-inverted and Fourier-inverted to obtain the option pricing result.

The integral transform approach cannot be effectively applied to the full two-asset SVJD model described in Chapter 2 because of the terms involving $\sqrt{v_1 v_2}$ in the option pricing IPDE (2.20). One can bypass this technical complication by simply assuming that $\rho_w = \rho_z = 0$ (i.e. the random noise terms in the stock processes and the volatility processes are uncorrelated) in the correlation structure described in equation (2.3). With the proportional SVJD model specification, we find that such an assumption is not necessary since we have $v_1 = v_2 = v$ (the common variance process), implying further that the resulting IPDE will not contain terms with a square root.

This chapter is organized as follows. Section 3.2 specifies the proportional stochastic volatility jump-diffusion model for stock prices, the selection of an equivalent martingale measure $\mathbb{Q}$, and the dynamics of the state variables under $\mathbb{Q}$. We derive the exchange option pricing IPDE in Section 3.3 and reduce the IPDE to a first-order PDE via integral transforms in Section 3.4. In Section 3.5, we solve the first-order PDE using the method of characteristics and in Section 3.6, we recover the solution of the exchange option pricing IPDE by inverting the integral transforms. Section 3.7 and Section 3.8 contains the derivation of integral representations of the price of European and American exchange options, respectively. In Section 3.9, we consider the case where the jumps in asset prices follow a lognormal distribution, as in Merton (1976). Section 3.10 summarizes and

concludes the chapter. To streamline the presentation of this chapter, the proof of some of the results in this chpater are placed in Appendix B.

3.2 The Proportional SVJD Model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Let $T > 0$ be the maturity of the exchange options to be considered. Let $\{W_1(t)\}$, $\{W_2(t)\}$, and $\{Z(t)\}$ be standard $\mathbb{P}$-Brownian motions with instantaneous correlations given by

$$\mathrm{d}W_1(t)\,\mathrm{d}W_2(t) = \rho_w\,\mathrm{d}t \qquad \text{and} \qquad \mathrm{d}W_j(t)\,\mathrm{d}Z(t) = \rho_{wz_j}\,\mathrm{d}t, \quad j = 1, 2. \tag{3.1}$$

Equivalently, this means that the random vector $\mathbf{B}(t) = (W_1(t), W_2(t), Z(t))^\top$ has correlation matrix

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \rho_w & \rho_{wz_1} \\ \rho_w & 1 & \rho_{wz_2} \\ \rho_{wz_1} & \rho_{wz_2} & 1 \end{bmatrix}. \tag{3.2}$$

Let $p_j(\mathrm{d}y, \mathrm{d}t)$ $(j = 1, 2)$ be the counting measure, defined on the mark space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, associated to a marked Poisson process with $\mathbb{P}$-local characteristics $(\lambda_j, m_{\mathbb{P}}^{(j)}(\mathrm{d}y))$. Underlying $p_j(\mathrm{d}y, \mathrm{d}t)$ is a sequence of ordered pairs $\{(T_{j,n}, Y_{j,n})\}$ where $Y_{j,n}$ is the "mark" of the nth occurrence of an event that occurs at a non-explosive time $T_{j,n}$.The marks $Y_{j,1}, Y_{j,2}, \ldots$ are i.i.d. real-valued random variables with non-atomic density $m_{\mathbb{P}}^{(j)}(\mathrm{d}y)$. Associated to the event times, we define a Poisson counting process $\{N_j(t)\}$ given by

$$N_j(t) = \sum_{n=1}^{\infty} \mathbf{1}(T_{j,n} \le t)\mathbf{1}(Y_{j,n} \in \mathbb{R}),$$

where $\mathbf{1}(\cdot)$ is the indicator function. We assume that $\{N_j(t)\}$ has constant intensity λ_j under $\mathbb{P}$.

We assume that the Brownian motions and the counting measures are mutually independent. Henceforth, we equip the probability space with the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$, the natural filtration generated by $\mathbf{B}(t)$, $p_1(\mathrm{d}y, \mathrm{d}t)$, and $p_2(\mathrm{d}y, \mathrm{d}t)$, satisfying the usual conditions.

We now specify the proportional SVJD model. Denote by $\{S_1(t)\}$ and $\{S_2(t)\}$ the price processes of stocks 1 and 2 that pay a constant dividend yield of q_1 and q_2, respectively, per annum. As stock prices may jump, we note that $S_1(t)$ and $S_2(t)$ denote the stock prices *prior to any jumps occurring at time t*. Let $\{v(t)\}$ be the instantaneous variance process that governs the volatility of both stock price processes. We assume that the dynamics of the stock prices and the instantaneous variance satisfy the stochastic differential equations

$$\frac{\mathrm{d}S_j(t)}{S_j(t)} = (\mu_j - \lambda_j\kappa_j)\,\mathrm{d}t + \sigma_j\sqrt{v(t)}\,\mathrm{d}W_j(t) + \int_{\mathbb{R}} (e^y - 1)\,p_j(\mathrm{d}y, \mathrm{d}t), \qquad j = 1, 2, \tag{3.3}$$

$$\mathrm{d}v(t) = \xi\left(\eta - v(t)\right)\mathrm{d}t + \omega\sqrt{v(t)}\,\mathrm{d}Z(t). \tag{3.4}$$

Here, $\kappa_j \equiv \mathbb{E}_{\mathbb{P}}[e^{Y_j} - 1] = \int_{\mathbb{R}}(e^y - 1)m_{\mathbb{P}}^{(j)}(\mathrm{d}y)$ is the mean jump size of the price of asset j under $\mathbb{P}$, and μ_j, σ_j, ξ, η, and ω are positive constants. It is also assumed that the initial values $S_1(0)$, $S_2(0)$, and $v(0)$ of these stochastic processes are known positive values.

The constant $\sigma_i > 0$ can be interpreted to be the proportionality constant that defines the volatility of asset i with respect to the underlying variance process $\{v_t\}$. That is, the volatility of asset i is proportional, with coefficient σ_i, to the current volatility $\sqrt{v_t}$. A specification of a proportional stochastic volatility is useful in situations where the two underlying assets possibly share a systematic source of volatility (for example, stocks written on corporations belonging to the same industry).

We assume that the dividend yields, the risk-free rate, the parameters μ_j, σ_j, ξ, η, and ω, the jump the intensities λ_j and jump-size densities $m_{\mathbb{P}}(\mathrm{d}y_j)$ are constant through time, but the analysis can be extended to the case where these parameters are deterministic functions of time.

As described above, the model features a common instantaneous variance process and independent jump terms for each asset. The individual jump processes may be taken to model idiosyncratic risk factors in each asset (for example, firm-specific news, pending litigation) that cause sudden and drastic changes in returns. In addition, the common variance process models systematic market volatility or volatility at an industry or macroeconomic level. As such, individual asset prices may provide feedback to each other via the correlation between the diffusion components and the dependence on a common stochastic volatility.

It is possible that jumps in stock prices may be due to these firm-specific events or more widespread macroeconomic factors or industry-wide events (for example, market crashes). An analysis of exchange options when the underlying assets have idiosyncratic and systematic jump components has been conducted by Cheang and Chiarella (2011), although their model specification assumes constant volatility. With high-frequency data on stocks traded on the NYSE, the S&P500, and a constructed equi-weighted portfolio of the individual stocks, Bollerslev et al. (2008) found that high levels of idiosyncratic jumps in the individual stock returns tend to dominate modestly sized co-jumps or joint jumps which cause the index and portfolio values to jump. In addition, they have found that common jumps occur less frequently than individual jumps. As such, we assume only the presence of idiosyncratic jumps in our model. The extension of our results to the case where there is also a systemic jump component (specified as a separate compound Poisson process independent of the idiosyncratic jump components; see Cheang and Chiarella (2011)) is straightforward.

The following assumption is required of the parameters of the instantaneous variance process and the correlation parameters to ensure that the variance process remains strictly positive and

finite (in finite time) under $\mathbb{P}$ and any other probability measure equivalent to $\mathbb{P}$ (Andersen and Piterbarg, 2007).

Assumption 3.1. Assume that the parameters ξ, η, and ω and the correlation coefficients ρ_{wz_1} and ρ_{wz_2} satisfy $2\xi\eta \geq \omega^2$ and $-1 < \rho_{wz_j} < \min\{\xi/\omega, 1\}$, $j = 1, 2$.

We shall now construct the measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the discounted yield processes $\{e^{-(r-q_1)t}S_1(t)\}$ and $\{e^{-(r-q_2)t}S_2(t)\}$ are $\mathbb{Q}$-martingales. The following proposition is a version of Proposition 2.3 that is applicable to the underlying asset price dynamics stated above. The suitability of the Radon-Nikodým derivative specified in the following proposition has been discussed in Chapter 2.

Proposition 3.2. *Suppose* $\boldsymbol{\theta}(t) = \big(\psi_1(t), \psi_2(t), \zeta(t)\big)^\top$ *is a vector of* $\mathcal{F}_t$*-adapted processes and let* $\gamma_1, \gamma_2, \nu_1, \nu_2$ *be constants. Define the process* $\{L(t)\}$ *by*

$$\begin{aligned} L(t) = \exp\left\{ -\int_0^t \Big(\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s)\Big)^\top \mathrm{d}\mathbf{B}(s) - \frac{1}{2}\int_0^t \boldsymbol{\theta}(s)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s)\,\mathrm{d}s \right\} \\ \times \exp\left\{ \sum_{n=1}^{N_1(t)} (\gamma_1 Y_{1,n} + \nu_1) - \lambda_1 t\left(e^{\nu_1}\mathbb{E}_{\mathbb{P}}(e^{\gamma_1 Y_1}) - 1\right)\right\} \\ \times \exp\left\{ \sum_{n=1}^{N_2(t)} (\gamma_2 Y_{2,n} + \nu_2) - \lambda_2 t\left(e^{\nu_2}\mathbb{E}_{\mathbb{P}}(e^{\gamma_2 Y_2}) - 1\right)\right\} \end{aligned} \tag{3.5}$$

and suppose that $\{L(t)\}$ *is a strictly positive* $\mathbb{P}$*-martingale such that* $\mathbb{E}_{\mathbb{P}}[L(t)] = 1$ *for all* $t \in [0, T]$*. Define the probability measure* $\mathbb{Q}$ *on* $(\Omega, \mathcal{F})$ *equivalent to* $\mathbb{P}$ *by* $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T} = L(T)$*. Then:*

1. *Under* $\mathbb{Q}$*, the vector process* $\mathbf{B}(t)$ *has drift* $-\boldsymbol{\theta}(t)$*;*

2. *The Poisson process* $N_j(t)$ *has a new intensity*

$$\tilde{\lambda}_j = \lambda_j e^{\nu_j}\mathbb{E}_{\mathbb{P}}[e^{\gamma_j Y_j}], \qquad j = 1, 2 \tag{3.6}$$

   under $\mathbb{Q}$*; and*

3. *The moment generating function of jump sizes random variable* Y_j *under* $\mathbb{Q}$ *is given by*

$$M_{\mathbb{Q},Y_j}(u) = \frac{M_{\mathbb{P},Y_j}(u + \gamma_j)}{M_{\mathbb{P},Y_j}(\gamma_j)}, \qquad j = 1, 2. \tag{3.7}$$

We first investigate the properties of the variance process $\{v_t\}$ following the change of measure. Let $\{\tilde{Z}(t)\}$ be a standard $\mathbb{Q}$-Wiener process. Then by Proposition 3.2, we can write

$$\mathrm{d}Z(t) = -\zeta_t\,\mathrm{d}t + \mathrm{d}\tilde{Z}(t).$$

The stochastic differential $\mathrm{d}v(t)$ can then be written as

$$\mathrm{d}v(t) = \left[\xi(\eta - v(t)) - \zeta(t)\omega\sqrt{v(t)}\right]\mathrm{d}t + \omega\sqrt{v(t)}\,\mathrm{d}\tilde{Z}(t).$$

As before, we interpret the term $\zeta(t)\omega\sqrt{v(t)}$ as the market price of volatility risk and assume that it is proportional to current volatility $v(t)$ (Heston, 1993). In other words, assume that there exists a nonnegative constant Λ such that

$$\Lambda v(t) = \zeta(t)\omega\sqrt{v(t)}.$$

In terms of Λ, we thus have

$$\mathrm{d}v(t) = \left[\xi\eta - (\xi + \Lambda)v(t)\right]\mathrm{d}t + \omega\sqrt{v(t)}\,\mathrm{d}\tilde{Z}(t). \tag{3.8}$$

If we define $\xi' = \xi + \Lambda$ and $\eta' = \xi\eta/(\xi + \Lambda)$, then $\{v(t)\}$ does not explode in finite time or makes excursions to 0 under $\mathbb{Q}$ if $2\xi'\eta' \geq \omega^2$. Indeed, Assumption 3.1 ensures that this condition is met.

Next, we determine $\psi_1(t)$ and $\psi_2(t)$ to ensure that the discounted yield processes $\{\tilde{S}_1(t)\} = \{e^{-(r-q_1)t}S_1(t)\}$ and $\{\tilde{S}_2(t)\} = \{e^{-(r-q_2)t}S_2(t)\}$ are $\mathbb{Q}$-martingales. Let $\{\tilde{W}_i(t)\}$ be a standard $\mathbb{Q}$-Wiener process. Then by Proposition 3.2, we may write

$$\mathrm{d}W_i(t) = -\psi_i(t)\,\mathrm{d}t + \mathrm{d}\tilde{W}_i(t), \qquad i = 1, 2.$$

Let $m_{\mathbb{Q}}^{(i)}(\mathrm{d}y)$ be the density of the jump size random variable Y_i under $\mathbb{Q}$. By Proposition 3.2, we obtain the relation

$$m_{\mathbb{Q}}^{(i)}(\mathrm{d}y) = \frac{e^{\gamma_i y}}{\mathbb{E}_{\mathbb{P}}(e^{\gamma_i Y_i})} m_{\mathbb{P}}^{(i)}(\mathrm{d}y)$$

from the expression for $M_{\mathbb{Q},Y_i}(u)$. It also follows from the proposition that the $\mathbb{Q}$-local characteristics of the counting measure $p_i(\mathrm{d}y, \mathrm{d}t)$ is given by $(\tilde{\lambda}_i, m_{\mathbb{Q}}^{(i)}(\mathrm{d}y))$. From here, we can define the $\mathbb{Q}$-compensated counting measure $q_i(\mathrm{d}y, \mathrm{d}t)$ as

$$q_i(\mathrm{d}y, \mathrm{d}t) = p_i(\mathrm{d}y, \mathrm{d}t) - \tilde{\lambda}_i m_{\mathbb{Q}}^{(i)}(\mathrm{d}y)\,\mathrm{d}t.$$

Using Itô's Lemma for jump-diffusion processes (Theorem A.1), $\tilde{S}_i(t)$ satisfies the stochastic differential equation

$$\begin{aligned}\frac{\mathrm{d}\tilde{S}_i(t)}{\tilde{S}_i(t-)} &= (\mu_i + q_i - r - \lambda_i\kappa_i)\,\mathrm{d}t + \sigma_i\sqrt{v(t)}\left(-\psi_i(t)\,\mathrm{d}t + \mathrm{d}\tilde{W}_i(t)\right) \\ &\quad + \int_{\mathbb{R}} (e^y - 1)\left[q_i(\mathrm{d}y, \mathrm{d}t) + \tilde{\lambda}_i m_{\mathbb{Q}}^{(i)}(\mathrm{d}y)\,\mathrm{d}t\right]\end{aligned}$$

$$= \left(\mu_i + q_i - r - \lambda_i\kappa_i + \tilde{\lambda}_i\tilde{\kappa}_i - \sigma_i\sqrt{v(t)}\psi_{i(t)}\right)\mathrm{d}t + \sigma_i\sqrt{v(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}}(e^y - 1)q_i(\mathrm{d}y, \mathrm{d}t).$$

In the last expression, $\tilde{\kappa}_i \equiv \mathbb{E}_{\mathbb{Q}}[e^{Y_i} - 1] = \int_{\mathbb{R}}(e^y - 1)m_{\mathbb{Q}}^{(i)}(\mathrm{d}y)$ denotes the mean jump size of asset i under $\mathbb{Q}$. Taking

$$\psi_i(t) = \frac{\mu_i + q_i - r - \lambda_i\kappa_i + \tilde{\lambda}_i\tilde{\kappa}_i}{\sigma_i\sqrt{v(t)}}$$

eliminates the drift term. With this selection, we obtain

$$\mathrm{d}\tilde{S}_i(t) = \tilde{S}_i(t-)\left\{\sigma_i\sqrt{v(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}}(e^y - 1)q_i(\mathrm{d}y, \mathrm{d}t)\right\}.$$

The last equation verifies that the discounted yield process is a martingale under $\mathbb{Q}$ given our choice of $\psi_{i,t}$.

Using stochastic integration by parts, we find that S_i satisfies the stochastic differential equation

$$\frac{\mathrm{d}S_i(t)}{S_i(t-)} = (r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i)\,\mathrm{d}t + \sigma_i\sqrt{v(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}}(e^y - 1)p_i(\mathrm{d}y, \mathrm{d}t). \tag{3.9}$$

By Itô's Lemma, the log-price $X_i(t) = \ln S_i(t)$ satisfies the equation

$$\mathrm{d}X_i(t) = \left(r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i - \frac{1}{2}\sigma_i^2 v(t)\right)\mathrm{d}t + \sigma_i\sqrt{v(t)}\,\mathrm{d}\tilde{W}_i(t) + \int_{\mathbb{R}} yp_i(\mathrm{d}y, \mathrm{d}t).$$

It follows therefore that the solution $S_i(t)$ is given by

$$S_i(t) = S_i(0)\exp\left\{(r - q_i - \tilde{\lambda}_i\tilde{\kappa}_i)t - \frac{1}{2}\sigma_i^2\int_0^t v(s)\,\mathrm{d}s + \sigma_i\int_0^t\sqrt{v(s)}\,\mathrm{d}\tilde{W}_i(s) + \sum_{n=1}^{N_i(t)} Y_{i,n}\right\}.$$

3.3 Deriving the IPDE for Option Prices

Let $C \equiv C(t, S_1, S_2, v)$ denote the price of an exchange option written on the two stocks where $S_1 = S_1(t-)$, $S_2 = S_2(t-)$, and $v = v(t)$. It is understood that S_1 and S_2 denote the *pre-jump* price of the two stocks at time t and that $C(t, S_1, S_2, v)$ denotes the price of the option prior to any jumps in the stocks at time t. At this point, we do not specify whether the option is European or American, but the type of option will dictate the terminal and boundary conditions that we impose on the IPDE.

A technical assumption on the differentiability of the option price is required to proceed.

Assumption 3.3. Assume that $C(t, s_1, s_2, v)$ is twice differentiable in s_1, s_2, and v and that its second-order partial derivatives with respect to these variables are continuous for all $(s_1, s_2, v) \in \mathbb{R}^3_+$. Assume further that $C(t, s_1, s_2, v)$ is continuously differentiable in t (up to the first order) for all $t > 0$.

Using Itô's Lemma for jump-diffusion processes (Theorem A.1), we find that C satisfies the stochastic differential equation

$$\begin{aligned}
\mathrm{d}C = \Bigg\{ & \frac{\partial C}{\partial t} + \sum_{j=1}^{2}(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j)S_j\frac{\partial C}{\partial s_j} + [\xi\eta - (\xi+\Lambda)v]\frac{\partial C}{\partial v} \\
& + \frac{1}{2}\sum_{j=1}^{2}\sigma_j^2 v S_j^2\frac{\partial^2 C}{\partial s_j^2} + \frac{1}{2}\omega^2 v\frac{\partial^2 C}{\partial v^2} + \rho_w\sigma_1\sigma_2 v S_1 S_2\frac{\partial^2 C}{\partial s_1\partial s_2} \\
& + \sum_{j=1}^{2}\rho_{wz_j}\sigma_j\omega v S_j\frac{\partial^2 C}{\partial s_j\partial v}\Bigg\}\,\mathrm{d}t + \sum_{j=1}^{2}\sigma_j\sqrt{v}S_j\frac{\partial C}{\partial s_j}\,\mathrm{d}\tilde{W}_{j,t} + \omega\sqrt{v}\frac{\partial C}{\partial v}\,\mathrm{d}\tilde{Z}_t \\
& + \int_{-\infty}^{\infty}\left[C\left(t, e^y S_1, S_2, v\right) - C\right]p_1(\mathrm{d}y, \mathrm{d}t) \\
& + \int_{-\infty}^{\infty}\left[C\left(t, S_1, e^y S_2, v\right) - C\right]p_2(\mathrm{d}y, \mathrm{d}t).
\end{aligned}$$

It is understood that all partial derivatives are evaluated at (t, S_1, S_2, v). Note further that the integral terms may be rewritten as the sum of an expectation and an integral with respect to the $\mathbb{Q}$-compensated measure $q_i(\mathrm{d}y, \mathrm{d}t)$. Given $\mathrm{d}C$, we can derive an expression for the stochastic differential $\mathrm{d}\tilde{C}$, where $\tilde{C}(t, S_1, S_2, v) \equiv e^{-rt}C(t, S_1, S_2, v)$ is the discounted option price. As we require $\{\tilde{C}\}$ to be a martingale under $\mathbb{Q}$ (Cheang and Chiarella, 2011), we need to eliminate its drift term. Doing so results to the IPDE

$$\begin{aligned}
rC = & \frac{\partial C}{\partial t} + \sum_{j=1}^{2}(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j)S_j\frac{\partial C}{\partial s_j} + [\xi\eta - (\xi+\Lambda)v]\frac{\partial C}{\partial v} \\
& + \frac{1}{2}\sum_{j=1}^{2}\sigma_j^2 v S_j^2\frac{\partial^2 C}{\partial s_j^2} + \frac{1}{2}\omega^2 v\frac{\partial^2 C}{\partial v^2} + \rho_w\sigma_1\sigma_2 v S_1 S_2\frac{\partial^2 C}{\partial s_1\partial s_2} \\
& + \sum_{j=1}^{2}\rho_{wz_j}\sigma_j\omega v S_j\frac{\partial^2 C}{\partial s_j\partial v} + \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C\left(t, e^{Y_1}S_1, S_2, v\right) - C\right] \\
& + \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C\left(t, S_1, e^{Y_2}S_2, v\right) - C\right]
\end{aligned} \tag{3.10}$$

To simplify notation, define the integro-differential operator $\mathcal{L}_{\mathbf{s},v}$ on the spatial variables s_1, s_2,

and v, such that $\mathcal{L}_{\mathbf{s},v}[C]$, for any function $C(t, s_1, s_2, v)$, is given by

$$\begin{aligned}\mathcal{L}_{\mathbf{s},v}[C] \equiv & \sum_{j=1}^{2}(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j)S_j\frac{\partial C}{\partial s_j} + [\xi\eta - (\xi + \Lambda)v]\frac{\partial C}{\partial v} \\ & + \frac{1}{2}\sum_{j=1}^{2}\sigma_j^2 v S_j^2\frac{\partial^2 C}{\partial s_j^2} + \frac{1}{2}\omega^2 v\frac{\partial^2 C}{\partial v^2} + \rho_w\sigma_1\sigma_2 v S_1 S_2\frac{\partial^2 C}{\partial s_1\partial s_2} \\ & + \sum_{j=1}^{2}\rho_{wz_j}\sigma_j\omega v S_j\frac{\partial^2 C}{\partial s_j\partial v} + \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[C\left(t, e^{Y_1}S_1, S_2, v\right) - C\right] \\ & + \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[C\left(t, S_1, e^{Y_2}S_2, v\right) - C\right].\end{aligned} \tag{3.11}$$

Then equation (3.10) can be written as

$$rC = \frac{\partial C}{\partial t} + \mathcal{L}_{\mathbf{s},v}[C].$$

In the case of the European exchange option, the above IPDE must be solved for $(t, S_1, S_2, v) \in [0, T] \times \mathbb{R}_+^3$ with the terminal condition $C(T, S_1, S_2, v) = (S_1 - S_2)^+$.

To articulate the boundary and terminal conditions for the preceding IPDE for the American exchange option, we first need to define the associated stopping and continuation regions. Similar to the case of two volatility processes (see Section 2.6.1), the stopping and continuation regions for the American exchange option under the proportional volatility case are given by

$$\begin{aligned}\mathcal{S} &= \left\{(t, S_1, S_2, v) \in \mathbb{R}_+^2 : S_1 \geq B(t, v)S_2\right\}, \\ \mathcal{C} &= \left\{(t, S_1, S_2, v) \in \mathbb{R}_+^2 : S_1 < B(t, v)S_2\right\},\end{aligned}$$

where $B(t, v) \geq 1$ is the critical ratio of stock prices S_1 and S_2 above which it is optimal to exercise the American exchange option. Note that the critical ratio is dependent on v based on the results of Touzi (1999) for the single-asset American put option. The quantity $B(t, v)$ is also known as the slope of the early exercise boundary for American exchange options (see Figure 2.1) (Broadie and Detemple, 1997).

Thus the American exchange option is the solution of a free boundary version of IPDE (3.10), which must be solved for $S_1 < B(t, v)S_2$ and $(t, S_2, v) \in [0, T] \times \mathbb{R}_+^2$. The terminal and boundary conditions are given by

$$\begin{aligned}C(T, S_1, S_2, v) &= (S_1 - S_2)^+, && (S_1, S_2, v) \in \mathbb{R}_+^3 \\ C(t, 0, S_2, v) &= 0, && (t, S_2, v) \in [0, T] \times \mathbb{R}_+^2 \\ C(t, S_1, 0, v) &= S_1, && (t, S_2, v) \in [0, T] \times \mathbb{R}_+^2.\end{aligned}$$

As the critical price ratio $B(t,v)$ is also unknown, we require the value-matching and smooth-pasting conditions to ensure the free-boundary problem is well-posed. As in Section 2.6.1, the value-matching condition is

$$C(t,S_1,S_2,v) = S_1 - S_2, \qquad \text{whenever} \qquad S_1 \geq B(t,v)S_2, \quad (t,S_2,v) \in [0,T] \times \mathbb{R}_+^2,$$

and, because by Assumption 3.3 C^A, $\frac{\partial C^A}{\partial s_1}$, and $\frac{\partial C^A}{\partial s_1}$ are continuous across the early exercise boundary, the smooth-pasting conditions are

$$\lim_{S_1 \to B(t,v)S_2} \frac{\partial C}{\partial s_1}(t,S_1,S_2,v) = 1, \qquad \lim_{S_1 \to B(t,v)S_2} \frac{\partial C}{\partial s_2}(t,S_1,S_2,v) = -1.$$

Since Assumption 3.3 also implies the continuity of $\frac{\partial C^A}{\partial \tau}$ and $\frac{\partial C^A}{\partial v}$ across the early exercise boundary, we also have

$$\lim_{S_1 \to B(\tau,v)S_2} \frac{\partial C^A}{\partial \tau}(\tau,S_1,S_2,v) = \lim_{S_1 \to B(\tau,v)S_2} \frac{\partial C^A}{\partial v}(\tau,S_1,S_2,v) = 0.$$

3.4 Fourier and Laplace Transform of the Pricing IPDE

We now use Fourier and Laplace transforms to solve IPDE (3.10) for the European exchange option, proceeding similar to Ziveyi (2011), Cheang et al. (2013), and Chiarella et al. (2015). As a preliminary step, we consider a change of variables from S_i to $X_i = \ln S_i$ (the log-price variable) via the equation

$$C(t,S_1,S_2,v) = e^{rt}U(t,X_1,X_2,v). \tag{3.12}$$

Here we take U to denote the value of the function $U(t,X_1,X_2,v)$ prior to any jumps in X_1 and X_2 at time t. The partial derivatives of C in terms of the partial derivatives of U and the log-price variables are given as follows:

$$\begin{aligned}
\frac{\partial C}{\partial t} &= e^{rt}\frac{\partial U}{\partial t} + rC, \\
\frac{\partial C}{\partial v} &= e^{rt}\frac{\partial U}{\partial v}, \\
\frac{\partial^2 C}{\partial v^2} &= e^{rt}\frac{\partial^2 U}{\partial v^2}, \\
\frac{\partial C}{\partial s_j} &= \frac{e^{rt}}{S_j}\frac{\partial U}{\partial x_j}, \qquad j = 1,2, \\
\frac{\partial^2 C}{\partial s_j^2} &= \frac{e^{rt}}{S_j^2}\left[\frac{\partial^2 U}{\partial x_j^2} - \frac{\partial U}{\partial x_j}\right], \qquad j = 1,2, \\
\frac{\partial^2 C}{\partial s_1 \partial s_2} &= \frac{e^{rt}}{S_1 S_2}\frac{\partial^2 U}{\partial x_1 \partial x_2},
\end{aligned}$$

$$\frac{\partial^2 C}{\partial s_j \partial v} = \frac{e^{rt}}{S_j}\frac{\partial^2 U}{\partial s_j \partial v}, \qquad j = 1, 2.$$

The partial derivatives of U are understood to be evaluated at (t, X_1, X_2, v). We also note that

$$\begin{aligned} C\left(t, e^{Y_1}S_1, S_2, v\right) &= e^{rt}U\left(t, X_1 + Y_1, X_2, v\right), \\ C\left(t, S_1, e^{Y_2}S_2, v\right) &= e^{rt}U\left(t, X_1, X_2 + Y_2, v\right). \end{aligned}$$

As such, IPDE (3.10) thus becomes

$$\begin{aligned} 0 = \frac{\partial U}{\partial t} &+ \sum_{j=1}^{2}\left(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j - \frac{1}{2}\sigma_j^2 v\right)\frac{\partial U}{\partial x_j} + \left[\xi\eta - (\xi + \Lambda)v\right]\frac{\partial U}{\partial v} \\ &+ \frac{1}{2}\sum_{j=1}^{2}\sigma_j^2 v\frac{\partial^2 U}{\partial x_j^2} + \frac{1}{2}\omega^2 v\frac{\partial^2 U}{\partial v^2} + \rho_w\sigma_1\sigma_2 v\frac{\partial^2 U}{\partial x_1 \partial x_2} \\ &+ \sum_{j=1}^{2}\rho_{wz_j}\sigma_j\omega v\frac{\partial^2 U}{\partial x_j \partial v} + \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[U\left(t, X_1 + Y_1, X_2, v\right) - U\right] \\ &+ \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[U\left(t, X_1, X_2 + Y_2, v\right) - U\right]. \end{aligned} \tag{3.13}$$

This IPDE must then be solved for $X_1, X_2 \in \mathbb{R}$, $v > 0$, and $0 \leq t < T$, and the associated terminal condition is given by

$$U(T, X_1, X_2, v) = \left(e^{X_1} - e^{X_2}\right)^+. \tag{3.14}$$

At this point, we introduce another change of variable to reflect the time to maturity $\tau = T - t$ instead of current time t. We do so by introducing the function $\tilde{U}$ given by

$$\tilde{U}(\tau, X_1, X_2, v) = U(T - t, X_1, X_2, v), \qquad \tau = T - t. \tag{3.15}$$

Note that all partial derivatives of $\tilde{U}$ are equal to those of U, except for the time derivative. In this case, we note that $\partial U/\partial t = -\partial\tilde{U}/\partial\tau$. Also, we point out that X_1, X_2, and v now represent their respective values (pre-jump values for X_1 and X_2) at time-to-maturity τ. Therefore, IPDE (3.13) can now be rewritten as

$$\begin{aligned} 0 = -\frac{\partial\tilde{U}}{\partial\tau} &+ \sum_{j=1}^{2}\left(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j - \frac{1}{2}\sigma_j^2 v\right)\frac{\partial\tilde{U}}{\partial x_j} + \left[\xi\eta - (\xi + \Lambda)v\right]\frac{\partial\tilde{U}}{\partial v} \\ &+ \frac{1}{2}\sum_{j=1}^{2}\sigma_j^2 v\frac{\partial^2\tilde{U}}{\partial x_j^2} + \frac{1}{2}\omega^2 v\frac{\partial^2\tilde{U}}{\partial v^2} + \rho_w\sigma_1\sigma_2 v\frac{\partial^2\tilde{U}}{\partial x_1 \partial x_2} \\ &+ \sum_{j=1}^{2}\rho_{wz_j}\sigma_j\omega v\frac{\partial^2\tilde{U}}{\partial x_j \partial v} + \tilde{\lambda}_1\mathbb{E}_{\mathbb{Q}}^{Y_1}\left[\tilde{U}\left(\tau, X_1 + Y_1, X_2, v\right) - \tilde{U}\right] \\ &+ \tilde{\lambda}_2\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[\tilde{U}\left(\tau, X_1, X_2 + Y_2, v\right) - \tilde{U}\right]. \end{aligned} \tag{3.16}$$

The terminal condition for IPDE (3.13) will now be replaced by an *initial* condition

$$\tilde{U}(0, X_1, X_2, v) = \left(e^{X_1} - e^{X_2}\right)^+, \tag{3.17}$$

where, here, X_1 and X_2 represent their respective values at maturity ($\tau = 0$).

As the coefficients of the IPDE (3.16) are free of the log-price variables, it may be simplified by taking the Fourier transform of all terms with respect to x_1 and x_2.

Although, by construction, $\tilde{U}$ is not absolutely integrable with respect to x_1 and x_2, we may still proceed with taking the Fourier transform of equation (3.16) by treating $\tilde{U}$ as a *tempered distribution* and using a generalized form of the Fourier transform. This approach to solving PDEs via Fourier transform is standard (see for example, Davies, 1985; Debnath, 2016; Strichartz, 2003; Torchinsky, 2011). In particular, Chiarella and Ziogas (2005, 2009) approach the solution of American option pricing I/PDEs in this manner.

By Assumption 3.3 on C, it follows that $\tilde{U}$ is also twice continuously differentiable for all x_1 and x_2, so it is locally integrable on $\mathbb{R}^2$ (with respect to x_1 and x_2). Furthermore, $\tilde{U} \to 0$ as $x_1 \to -\infty$ or $x_2 \to \infty$ since the exchange option becomes worthless for very small values of S_1 or very large values of S_2. It is also reasonable to assume that $\tilde{U}$ exhibits at most polynomial growth as $x_1 \to \infty$ or $x_2 \to -\infty$ since the option price behaves asymptotically similar to the exchange option payoff, which is linear in S_1 and S_2 (see also Figures 5.2 and 5.5). Therefore, for fixed τ and v, $\tilde{U}$ can be identified as a tempered distribution on $\mathbb{R}^2$ (see for example, Folland, 1999; Rudin, 1991; Torchinsky, 2011), which is a continuous linear functional defined by the action

$$\langle \tilde{U}, g \rangle \equiv \int_{\mathbb{R}^2} \tilde{U}(\tau, x_1, x_2, v) g(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2, \tag{3.18}$$

for every g belonging to the class $\mathscr{S}(\mathbb{R}^2)$ of $C^\infty(\mathbb{R}^2)$ functions which, along with their partial derivatives, decrease faster than any power of $\|x\|$ ($x \in \mathbb{R}^2$, $\|x\|$ is the Euclidean norm) as $\|x\| \to \infty$ (see for example, Folland, 1999, Section 8.1).[2] We denote by $\mathscr{S}'(\mathbb{R}^2)$ the space of all tempered distributions on $\mathbb{R}^2$. For a comprehensive discussion of tempered distributions and Fourier analysis for distributions, see for example, Davies (1985), Richards and Youn (1990), Strichartz (2003), Rudin (1991), or van Dijk (2013).

[2] A tempered distribution on $\mathbb{R}^2$ is a continuous linear functional defined on $\mathscr{S}(\mathbb{R}^2)$. More generally, the space $\mathscr{S}(\mathbb{R}^n)$ is referred to as the *Schwartz space* on $\mathbb{R}^n$, which is defined as a collection of functions $g \in C^\infty(\mathbb{R}^n)$ such that

$$\|g\|_{(N,\alpha)} \equiv \sup_{x \in \mathbb{R}^n} (1 + \|x\|^N) |\partial^\alpha g(x)| < \infty$$

for all nonnegative integers N and multi-indices α. Here, a multi-index is an n-tuple of nonnegative integers $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\|\cdot\|$ denotes the usual Euclidean norm on $\mathbb{R}^n$, and ∂^α is defined as

$$\partial^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}.$$

For $g \in \mathscr{S}(\mathbb{R}^2)$, the (classical) Fourier transform $\hat{g}$ of g is defined as

$$\hat{g}(\boldsymbol{\phi}) = \int_{\mathbb{R}^2} g(\mathbf{x}) e^{i\boldsymbol{\phi}\cdot\mathbf{x}} \, \mathrm{d}\mathbf{x},$$

and the corresponding Fourier inversion formula is

$$\check{g}(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} g(\boldsymbol{\phi}) e^{-i\boldsymbol{\phi}\cdot\mathbf{x}} \, \mathrm{d}\boldsymbol{\phi}.$$

It is known that $\hat{\check{g}} = \check{\hat{g}} = g$ and the Fourier transform maps $\mathscr{S}(\mathbb{R}^2)$ onto itself (Folland, 1999, Corollary 8.23).

However, the Fourier transform can be generalized into a continuous linear map from $\mathscr{S}'(\mathbb{R}^2)$ onto itself by defining the Fourier transform of $f \in \mathscr{S}'(\mathbb{R}^2)$ to be the tempered distribution $\hat{f}$ that satisfies

$$\langle \hat{f}, g \rangle = \langle f, \hat{g} \rangle, \qquad \text{for any } g \in \mathscr{S}(\mathbb{R}^2). \tag{3.19}$$

Thus, if $f \in \mathscr{S}'(\mathbb{R}^2)$ is defined by the action (3.18) on $\mathscr{S}(\mathbb{R}^2)$, then its Fourier transform, $\hat{f}$ (with respect to x_1 and x_2), is characterized by the action

$$\begin{aligned}
\langle \hat{f}, g \rangle &= \langle f, \hat{g} \rangle \\
&= \int_{\mathbb{R}^2} f(x_1, x_2,) \hat{g}(x_1, x_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \\
&= \int_{\mathbb{R}^2} f(x_1, x_2, v) \int_{\mathbb{R}^2} g(y_1, y_2) e^{i(x_1 y_1 + x_2 y_2)} \, \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}x_1 \, \mathrm{d}x_2 \\
&= \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} g(y_1, y_2) f(x_1, x_2) e^{i(x_1 y_1 + x_2 y_2)} \, \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}x_1 \, \mathrm{d}x_2 \\
&= \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} f(x_1, x_2) e^{i(x_1 y_1 + x_2 y_2)} \, \mathrm{d}x_1 \, \mathrm{d}x_2 \right) g(y_1, y_2) \, \mathrm{d}y_1 \, \mathrm{d}y_2,
\end{aligned}$$

for any $g \in \mathscr{S}(\mathbb{R}^2)$.

The inverse Fourier transform $\check{f}$ of $f \in \mathscr{S}'(\mathbb{R}^2)$ is defined similarly,

$$\langle \check{f}, g \rangle = \langle f, \check{g} \rangle, \qquad \text{for any } g \in \mathscr{S}(\mathbb{R}^2), \tag{3.20}$$

where the right-hand side represents the action (3.18). Thus, from the definition of the inverse Fourier transform for $g \in \mathscr{S}(\mathbb{R}^2)$, we have

$$\begin{aligned}
\langle \check{f}, g \rangle &= \langle f, \check{g} \rangle \\
&= \int_{\mathbb{R}^2} f(y_1, y_2) \check{g}(y_1, y_2) \, \mathrm{d}y_1 \, \mathrm{d}y_2 \\
&= \int_{\mathbb{R}^2} f(y_1, y_2) \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} g(x_1, x_2) e^{-i(x_1 y_1 + x_2 y_2)} \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}y_1 \, \mathrm{d}y_2
\end{aligned}$$

$$= \int_{\mathbb{R}^2} \left(\frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} f(y_1, y_2) e^{-i(x_1 y_1 + x_2 y_2)} \, \mathrm{d}y_1 \, \mathrm{d}y_2 \right) g(x_1, x_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2.$$

Furthermore, equations (3.19) and (3.20) imply that the Fourier Inversion Theorem (see for example, Folland, 1999, Theorem 8.26) extends to $\mathscr{S}'(\mathbb{R}^2)$ via

$$\langle \check{\hat{f}}, g \rangle = \langle \hat{f}, \check{g} \rangle = \langle f, \hat{\check{g}} \rangle = \langle f, g \rangle, \tag{3.21}$$

for any $f \in \mathscr{S}'(\mathbb{R}^2)$ and $g \in \mathscr{S}(\mathbb{R}^2)$. This implies that $\check{\hat{f}} = f$ and $\hat{\check{f}} = f$. Treating $\tilde{U}(\tau, \cdot, \cdot, v)$ (for fixed τ and v) as a tempered distribution on $\mathbb{R}^2$, equations (3.19) and (3.20) lead to our definition of the Fourier transform (3.22) and inverse Fourier transform (3.37) for $\tilde{U}$.

Furthermore, by treating $\tilde{U}$ as a tempered distribution and solving (3.16) using the (generalized) Fourier transform, we are implicitly solving for a solution that does not grow too rapidly at infinity (Strichartz, 2003). However, our numerical analysis in Chapter 5 shows that a purely numerical solution of a similar IPDE yields an option price profile that is asymptotically linear. Thus, this implicit restriction is reasonable in solving (3.16) using Fourier transforms.

The fact that $\tilde{U}$ exhibits at most polynomial growth at the extremes of the domain is tied to the properties of the exchange option payoff function. Thus, for general payoff functions which may not have at most polynomial growth, one can adapt the Fourier transform methods of Carr and Madan (1999) or Lewis (2000) for option pricing.

The following proposition shows the second-order partial differential equation obtained by taking the Fourier transform of $\tilde{U}$ with respect to x_1 and x_2.

Proposition 3.4. *Let $\hat{U} = \hat{U}(\tau, \phi_1, \phi_2, v)$ denote the Fourier transform of $\tilde{U}(\tau, X_1, X_2, v)$ with respect to the x-variables,*

$$\hat{U}(\tau, \phi_1, \phi_2, v) = \iint_{\mathbb{R}^2} \tilde{U}(\tau, x_1, x_2, v) e^{i(\phi_1 x_1 + \phi_2 x_2)} \, \mathrm{d}x_1 \, \mathrm{d}x_2. \tag{3.22}$$

Then $\hat{U}$ satisfies the PDE

$$0 = -\frac{\partial \hat{U}}{\partial \tau} + (\xi\eta - \Theta v) \frac{\partial \hat{U}}{\partial v} + \frac{1}{2}\omega^2 v \frac{\partial^2 \hat{U}}{\partial v^2} + \left(\Psi + \frac{1}{2}\varepsilon v \right) \hat{U}. \tag{3.23}$$

where

$$\begin{aligned}
\Theta \equiv \Theta(\phi_1, \phi_2) &= \xi + \Lambda + i \sum_{j=1}^{2} \rho_{w z_j} \sigma_j \omega \phi_j, \\
\varepsilon \equiv \varepsilon(\phi_1, \phi_2) &= \sum_{j=1}^{2} \sigma_j^2 \left(i\phi_j - \phi_j^2 \right) - 2\rho_w \sigma_1 \sigma_2 \phi_1 \phi_2, \\
\Psi \equiv \Psi(\phi_1, \phi_2) &= \sum_{j=1}^{2} i\phi_j \left[\frac{\tilde{\lambda}_j}{i\phi_j} \left(\varphi_j(\phi_j) - 1 \right) - r + q_j + \tilde{\lambda}_j \tilde{\kappa}_j \right],
\end{aligned} \tag{3.24}$$

where $\varphi_j(\phi)$ is the characteristic function of the jump size random variable Y_j $(j = 1, 2)$ under $\mathbb{Q}$. The corresponding initial condition is

$$\hat{u}(\phi_1, \phi_2, v) \equiv \hat{U}(0, \phi_1, \phi_2, v) = \iint_{\mathbb{R}^2} e^{i\phi_1 x_1 + i\phi_2 x_2} \left(e^{x_1} - e^{x_2} \right)^+ \mathrm{d}x_1 \, \mathrm{d}x_2. \tag{3.25}$$

Proof. Using the properties of the (generalized) Fourier transform (see for example, Davies, 1985, Section 9.5), the Fourier transform of the partial derivatives are given below:

$$\begin{aligned}
\mathcal{F}_x \left\{ \frac{\partial \tilde{U}}{\partial \tau} \right\} &= \frac{\partial \hat{U}}{\partial \tau}, \\
\mathcal{F}_x \left\{ \frac{\partial \tilde{U}}{\partial v} \right\} &= \frac{\partial \hat{U}}{\partial v}, \\
\mathcal{F}_x \left\{ \frac{\partial^2 \tilde{U}}{\partial v^2} \right\} &= \frac{\partial^2 \hat{U}}{\partial v^2}, \\
\mathcal{F}_x \left\{ \frac{\partial \tilde{U}}{\partial x_j} \right\} &= -i\phi_j \hat{U}, j = 1, 2, \\
\mathcal{F}_x \left\{ \frac{\partial^2 \tilde{U}}{\partial x_j^2} \right\} &= -\phi_j^2 \hat{U}, j = 1, 2, \\
\mathcal{F}_x \left\{ \frac{\partial^2 \tilde{U}}{\partial x_1 \partial x_2} \right\} &= -\phi_1 \phi_2 \hat{U}, \\
\mathcal{F}_x \left\{ \frac{\partial^2 \tilde{U}}{\partial x_j \partial v} \right\} &= -i\phi_j \frac{\partial \hat{U}}{\partial v}.
\end{aligned}$$

Let $g_j(y)$ denote the probability density function of Y_j under $\mathbb{Q}$. If

$$\varphi_j(\phi) = \mathbb{E}_{\mathbb{Q}}(e^{-i\phi Y_j}) = \int_{\mathbb{R}} e^{-i\phi y} g_j(y) \, \mathrm{d}y$$

denotes the characteristic function of Y_j under $\mathbb{Q}$ (i.e. the Fourier transform of g_j), then the Fourier transform of the expectation terms can be written as

$$\mathcal{F}_x \left\{ \mathbb{E}_{\mathbb{Q}}^{Y_1} \left[\tilde{U}(\tau, X_1 + Y_1, X_2, v) - \tilde{U} \right] \right\} = \left(\varphi_1(\phi_1) - 1 \right) \hat{U},$$

$$\mathcal{F}_x\left\{\mathbb{E}_{\mathbb{Q}}^{Y_2}\left[\tilde{U}\left(\tau, X_1, X_2+Y_2, v\right)-\tilde{U}\right]\right\}=\left(\varphi_2(\phi_2)-1\right)\hat{U},$$

following the convolution property for Fourier transforms (see for example, Debnath (2016, Theorem 2.5.5) or Folland (1999, p. 295)). The Fourier transform of IPDE (3.13) is therefore given by the PDE

$$\begin{aligned}0=&-\frac{\partial\hat{U}}{\partial\tau}+\sum_{j=1}^{2}\left(r-q_j-\tilde{\lambda}_j\tilde{\kappa}_j-\frac{1}{2}\sigma_j^2 v\right)\left(-i\phi_j\hat{U}\right)+\left[\xi\eta-(\xi+\Lambda)v\right]\frac{\partial\hat{U}}{\partial v}\\&+\frac{1}{2}\sum_{j=1}^{2}\sigma_j^2 v\left(-\phi_j^2\hat{U}\right)+\frac{1}{2}\omega^2 v\frac{\partial^2\hat{U}}{\partial v^2}+\rho_w\sigma_1\sigma_2 v\left(-\phi_1\phi_2\hat{U}\right)\\&+\sum_{j=1}^{2}\phi_{wz_j}\sigma_j\omega v\left(-i\phi_j\frac{\partial\hat{U}}{\partial v}\right)+\left[\tilde{\lambda}_1\left(\varphi_1(\phi_1)-1\right)+\tilde{\lambda}_2\left(\varphi_2(\phi_2)-1\right)\right]\hat{U}.\end{aligned}$$

To simplify the PDE, we introduce the functions Θ, ε, and Ψ of ϕ_1 and ϕ_2 as given in the statement of the proposition. The simplified PDE for $\hat{U}$ is therefore given by

$$0=-\frac{\partial\hat{U}}{\partial\tau}+(\xi\eta-\Theta v)\frac{\partial\hat{U}}{\partial v}+\frac{1}{2}\omega^2 v\frac{\partial^2\hat{U}}{\partial v^2}+\left(\Psi+\frac{1}{2}\varepsilon v\right)\hat{U}.$$

The initial condition, which we denote by $\hat{u}(\phi_1,\phi_2,v)$, is obtained by taking the Fourier transform of $\tilde{U}(0,X_1,X_2,v)=(e^{X_1}-e^{X_2})^+$, and so

$$\hat{u}(\phi_1,\phi_2,v)\equiv\hat{U}(0,\phi_1,\phi_2,v)=\iint_{\mathbb{R}^2}e^{i\phi_1x_1+i\phi_2x_2}\left(e^{x_1}-e^{x_2}\right)^+\mathrm{d}x_1\,\mathrm{d}x_2.$$

At this point, we do not seek to evaluate this Fourier transform. ■

Note now that equation (3.23) is first-order in τ and second-order in v with coefficients that are linear in τ and v. The equation may be reduced to a first-order PDE in τ and v by taking the Laplace transform with respect to v.

To ensure that all required Laplace transforms exist, we make the following assumption.

Assumption 3.5. Assume that $e^{-\vartheta v}\hat{U}$ and $e^{-\vartheta v}\partial\hat{U}/\partial v$ vanish as $v\to\infty$.

The following proposition presents the PDE that results from taking the Laplace transform of $\hat{U}$ and its associated PDE (3.23).

Proposition 3.6. *Define $\bar{U}=\bar{U}(\tau,\phi_1,\phi_2,\vartheta)$ to be the Laplace transform of $\hat{U}$ in the v-direction,*

$$\bar{U}(\tau,\phi_1,\phi_2,\vartheta)=\mathscr{L}_v\left\{\hat{U}(\tau,\phi_1,\phi_2,v)\right\}(\vartheta)=\int_0^\infty e^{-\vartheta v}\hat{U}(\tau,\phi_1,\phi_2,v)\,\mathrm{d}v. \tag{3.26}$$

Then $\bar{U}$ satisfies the equation

$$\frac{\partial \bar{U}}{\partial \tau} + \left(\frac{1}{2}\omega^2\vartheta^2 - \Theta\vartheta + \frac{1}{2}\varepsilon\right)\frac{\partial \bar{U}}{\partial \vartheta} = \left[(\xi\eta - \omega^2)\vartheta + \Theta + \Psi\right]\bar{U} + f(\tau), \tag{3.27}$$

for $\vartheta > 0$ and $0 \leq \tau \leq T$ and where $f(\tau)$ is a function determined such that

$$\lim_{\vartheta\to\infty} \bar{U}(\tau, \phi_1, \phi_2, \vartheta) = 0. \tag{3.28}$$

The associated initial condition is given by

$$\bar{u}(\phi_1, \phi_2, \vartheta) \equiv \bar{U}(0, \phi_1, \phi_2, \vartheta) = \mathscr{L}_v\left\{\hat{u}(\phi_1, \phi_2, v)\right\}, \tag{3.29}$$

Proof. By the properties of the Laplace transform for derivatives (see for example, Debnath, 2016, Theorem 3.4.3), the Laplace transforms relevant to PDE (3.23) are given by

$$\begin{aligned}
\mathscr{L}_v\left\{\frac{\partial \hat{U}}{\partial \tau}\right\} &= \frac{\partial \bar{U}}{\partial \tau}, \\
\mathscr{L}_v\left\{\left(\Psi + \frac{1}{2}\varepsilon v\right)\hat{U}\right\} &= \Psi\bar{U} - \frac{1}{2}\varepsilon\frac{\partial \bar{U}}{\partial \vartheta}, \\
\mathscr{L}_v\left\{(\xi\eta - \Theta v)\frac{\partial \hat{U}}{\partial v}\right\} &= -\xi\eta\hat{U}(\tau, \phi_1, \phi_2, 0) + (\xi\eta\vartheta + \Theta)\bar{U}_{t-} + \Theta\vartheta\frac{\partial \bar{U}}{\partial \vartheta}, \\
\mathscr{L}_v\left\{\frac{1}{2}\omega^2 v\frac{\partial^2 \hat{U}}{\partial v^2}\right\} &= \frac{1}{2}\omega^2\left(\hat{U}(\tau, \phi_1, \phi_2, 0) - \vartheta^2\frac{\partial \bar{U}}{\partial \vartheta} - 2\vartheta\bar{U}\right).
\end{aligned}$$

Thus, the Laplace transform of PDE (3.23) is given by

$$\frac{\partial \bar{U}}{\partial \tau} = \left(\frac{1}{2}\omega^2 - \xi\eta\right)\hat{U}(\tau, \phi_1, \phi_2, 0) + \left[(\xi\eta - \omega^2)\vartheta + \Theta + \Psi\right]\bar{U} - \left(\frac{1}{2}\omega^2\vartheta^2 - \Theta\vartheta + \frac{1}{2}\varepsilon\right)\frac{\partial \bar{U}}{\partial \vartheta}.$$

Note that $\hat{U}(\tau, \phi_1, \phi_2, 0)$ is unknown. To this end, we define the function $f(\tau)$ given by

$$f(\tau) = \left(\frac{1}{2}\omega^2 - \xi\eta\right)\hat{U}(\tau, \phi_1, \phi_2, 0). \tag{3.30}$$

In the next set of calculations, the function f will be determined such that

$$\lim_{\vartheta\to\infty} \bar{U}(\tau, \phi_1, \phi_2, \vartheta) = 0.$$

This condition ensures that the Laplace transform $\bar{U}(\tau, \phi_1, \phi_2, \vartheta)$ is finite for all $\vartheta > 0$. Therefore, the Laplace transformed PDE is given by

$$\frac{\partial \bar{U}}{\partial \tau} + \left(\frac{1}{2}\omega^2\vartheta^2 - \Theta\vartheta + \frac{1}{2}\varepsilon\right)\frac{\partial \bar{U}}{\partial \vartheta} = \left[(\xi\eta - \omega^2)\vartheta + \Theta + \Psi\right]\bar{U} + f(\tau).$$

The associated initial condition is obtained by taking the Laplace transform of $\hat{u}(\phi_1, \phi_2, v)$ with respect to v. ■

3.5 Solving the First-Order PDE

As equation (3.27) is a first-order PDE, we will use the method of characteristics to solve for $\bar{U}$.[3] The solution $\bar{U}$ is given in the following proposition.

Proposition 3.7. *The solution $\bar{U}(\tau, \phi_1, \phi_2, \vartheta)$ of equation (3.27) is given by*

$$\begin{aligned}\bar{U}(\tau, \phi_1, \phi_2, \vartheta) = \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right]\tau\right\}\left[\frac{2\delta}{(\omega^2\vartheta - \Theta + \delta)(e^{\delta\tau}-1)+2\delta}\right]^{-\frac{2\alpha}{\omega^2}} \\ \times \int_0^\infty \hat{u}(\phi_1, \phi_2, b)\exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\} \\ \times \exp\left\{-\frac{2b\delta(\omega^2\vartheta - \Theta + \delta)e^{\delta\tau}}{\omega^2\left[(\omega^2\vartheta - \Theta + \delta)(e^{\delta\tau}-1)+2\delta\right]}\right\} \\ \times \Gamma\left(\frac{2\alpha}{\omega^2}+1; \beta^*(b)\right)\mathrm{d}b,\end{aligned} \tag{3.31}$$

where

$$\begin{aligned}\alpha &= \xi\eta - \omega^2, \\ \delta = \delta(\phi_1, \phi_2) &= \sqrt{\Theta^2(\phi_1, \phi_2) - \varepsilon(\phi_1, \phi_2)\omega^2}, \\ \beta^*(b) &= \frac{2\delta b e^{\delta\tau}}{\omega^2(e^{\delta\tau}-1)} \times \frac{2\delta}{(\omega^2\vartheta - \Theta + \delta)(e^{\delta\tau}-1)+2\delta},\end{aligned} \tag{3.32}$$

$\Gamma(u;\beta)$ is the lower incomplete gamma function given by

$$\Gamma(u;\beta) = \frac{1}{\Gamma(u)}\int_0^\beta e^{-x}x^{u-1}\,\mathrm{d}x, \tag{3.33}$$

and $\Gamma(u)$ is the gamma function given by

$$\Gamma(u) = \int_0^\infty e^{-x}x^{u-1}\,\mathrm{d}x. \tag{3.34}$$

Proof. See Appendix B.1. ■

3.6 Inverting the Laplace Transform

The next step is to invert the Laplace transform $\bar{U}(\tau, \phi_1, \phi_2, \vartheta)$ to recover the Fourier transform $\hat{U}(\tau, \phi_1, \phi_2, v)$. This result is shown in the following proposition.

[3] For a more detailed discussion of the method of characteristics, see for example Zauderer (2006).

Proposition 3.8. *The inverse of the Laplace transform $\bar{U}(\tau,\phi_1,\phi_2,\vartheta)$ is the Fourier transform $\hat{U}(\tau,\phi_1,\phi_2,v)$ given by*

$$\begin{aligned}\hat{U}(\tau,\phi_1,\phi_2,v) = \int_0^\infty \hat{u}(\phi_1,\phi_2,b)\exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)(\xi\eta\tau+v-b)\right\}e^{\Psi\tau}\frac{2\delta e^{\delta\tau}}{\omega^2(e^{\delta\tau}-1)} \\ \times\left(\frac{be^{\delta\tau}}{v}\right)^{\frac{\xi\eta}{\omega^2}-\frac{1}{2}}\exp\left\{-\frac{2\delta(be^{\delta\tau}+v)}{\omega^2(e^{\delta\tau}-1)}\right\}\times I_{\left(\frac{2\xi\eta}{\omega^2}-1\right)}\left(\frac{4\delta\sqrt{bve^{\delta\tau}}}{\omega^2(e^{\delta\tau}-1)}\right)\mathrm{d}b,\end{aligned} \tag{3.35}$$

where $I_k(u)$ is the modified Bessel function of the first kind, which is given by

$$I_k(u) = \sum_{n=0}^{\infty}\frac{(u/2)^{2n+k}}{n!\Gamma(n+k+1)}. \tag{3.36}$$

Proof. See Appendix B.2. ∎

3.7 The European Exchange Option Price

Having found $\hat{U}(\tau,\phi_1,\phi_2,v)$, we can now apply the Fourier inversion formula

$$\tilde{U}(\tau,X_1,X_2,v) = \frac{1}{4\pi^2}\iint_{\mathbb{R}^2} e^{-i\phi_1X_1-i\phi_2X_2}\hat{U}(\tau,\phi_1,\phi_2,v)\,\mathrm{d}\phi_1\phi_2 \tag{3.37}$$

to recover the function $\tilde{U}(\tau,X_1,X_2,v)$. Given $\tilde{U}(\tau,X_1,X_2,v)$, the price of the European exchange option $C^E(\tau,S_1,S_2,v)$ at time to maturity τ can be obtained from equations (3.12) and (3.15). This result is shown in the following proposition.

Proposition 3.9. *Let S_1 and S_2 denote the price of asset 1 and 2 and let v be the variance, all observed at time to maturity τ. The price of the European exchange option, denoted by $C^E(\tau,S_1,S_2,v)$ at time to maturity τ is given by*

$$C^E(\tau,S_1,S_2,v) = e^{-r\tau}\int_0^\infty\iint_{\mathbb{R}^2}\left[(e^{y_1}-e^{y_2})^+\right]G(\tau,y_1,y_2,b;S_1,S_2,v)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}b, \tag{3.38}$$

where $G(\tau,y_1,y_2,b;S_1,S_2,v)$ is defined as

$$\begin{aligned}&G(\tau,y_1,y_2,b;S_1,S_2,v)\\ &\quad=\frac{1}{4\pi^2}\iint_{\mathbb{R}^2}e^{i\phi_1y_1+i\phi_2y_2}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau}\exp\{-i\phi_1\ln S_1-i\phi_2\ln S_2\}\\ &\qquad\times\exp\left\{-\sum_{j=1}^{2}i\phi_j(r-q_j-\tilde{\lambda}_j\tilde{\kappa}_j)\tau\right\}\\ &\qquad\times\left[\sum_{m,n=0}^{\infty}\frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!}\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{-i\phi_1\Upsilon_{1,m}-i\phi_2\Upsilon_{2,n}\right\}\right]\right]F(b)\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2.\end{aligned} \tag{3.39}$$

Here, $F(b) \equiv F(\tau, \phi_1, \phi_2, b; v)$ *is given by*

$$\begin{aligned} F(b) = \frac{2\delta e^{\delta\tau}}{\omega^2(e^{\delta\tau}-1)} \left(\frac{be^{\delta\tau}}{v}\right)^{\frac{\xi\eta}{\omega^2}-\frac{1}{2}} \exp\left\{-\frac{2\delta(be^{\delta\tau}+v)}{\omega^2(e^{\delta\tau}-1)}\right\} \\ \times \exp\left\{\left(\frac{\Theta-\delta}{\omega^2}(\xi\eta\tau+v-b)\right)\right\} I_{\left(\frac{2\xi\eta}{\omega^2}-1\right)}\left(\frac{4\delta\sqrt{bve^{\delta\tau}}}{\omega^2(e^{\delta\tau}-1)}\right) \end{aligned} \tag{3.40}$$

and the random variables $\Upsilon_{1,m}$ *and* $\Upsilon_{2,n}$ *are defined as*

$$\Upsilon_{1,m} = \sum_{k=1}^{m} Y_{1,k} \qquad \Upsilon_{2,n} = \sum_{l=1}^{n} Y_{2,l}, \tag{3.41}$$

where $\{Y_{1,1}, \ldots, Y_{1,m}\}$ *and* $\{Y_{2,1}, \ldots, Y_{2,n}\}$ *are collections of i.i.d. random variables sampled from the* $\mathbb{Q}$*-densities* $g_1(y)$ *and* $g_2(y)$*, respectively, of the jump-size random variables* Y_1 *and* Y_2 *(see proof of Proposition 3.4). Note that the expectation operator acts only on the random variables* $\Upsilon_{1,m}$ *and* $\Upsilon_{2,n}$*.*

PROOF. See Appendix B.3. ∎

The next two results are useful in simplifying the representation of C^E in (3.38). First, the following lemma provides the integral of $F(b)$, as defined in (3.40), with respect to b on $(0, \infty)$.

Lemma 3.10. *The integral* $\int_0^\infty F(b)\,\mathrm{d}b$*, where* F *is given by* (3.40)*, is*

$$\begin{aligned} \int_0^\infty F(b)\,\mathrm{d}b = \exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)(\xi\eta\tau+v)\right\} \exp\left\{-\frac{2\delta v}{\omega^2(e^{\delta\tau}-1}\right\} \\ \times \left[\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right]^{\frac{2\xi\eta}{\omega^2}} \\ \times \exp\left\{\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta} \cdot \frac{2\delta v}{\omega^2(e^{\delta\tau}-1)}\right\}. \end{aligned} \tag{3.42}$$

PROOF. See Appendix B.4. ∎

With the results in Lemma 3.10, we can then evaluate $\int_0^\infty G(\tau, y_1, y_2, b; S_1, S_2, v)\,\mathrm{d}b$. This integral is given in the following lemma.

Lemma 3.11. *The integral* $\int_0^\infty G(\tau, y_1, y_2, b; S_1, S_2, v)\,\mathrm{d}b$*, where* G *is given by* (3.39)*, is*

$$\begin{aligned}
&\int_0^\infty G(\tau, y_1, y_2, b; S_1, S_2, v)\,\mathrm{d}b \\
&\quad = \sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau} \\
&\qquad \times \mathbb{E}_{\mathbb{Q}}\Bigg[\frac{1}{4\pi^2}\iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} \exp\Big\{ -i\phi_1\Big(\ln S_1 + \Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1\tau\Big) \\
&\qquad - i\phi_2\Big(\ln S_1 + \Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2\tau\Big) + B_1(\tau, -\phi_1, -\phi_2) \\
&\qquad + B_2(\tau, -\phi_1, -\phi_2)v\Big\}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2\Bigg],
\end{aligned} \tag{3.43}$$

where B_1 *and* B_2 *are the functions*

$$\begin{aligned}
B_1(\tau, \phi_1, \phi_2) &= i\phi_1(r - q_1)\tau + i\phi_2(r - q_2)\tau \\
&\quad + \frac{\xi\eta}{\omega^2}\left\{(\Theta_2 + \delta_2)\tau - 2\ln\left[\frac{1 - \chi_2 e^{\delta_2\tau}}{1 - \chi_2}\right]\right\}
\end{aligned} \tag{3.44}$$

and

$$B_2(\tau, \phi_1, \phi_2) = \frac{\Theta_2 + \delta_2}{\omega^2}\left[\frac{1 - e^{\delta_2\tau}}{1 - \chi_2 e^{\delta_2\tau}}\right]. \tag{3.45}$$

with

$$\begin{aligned}
\Theta_2 &= \Theta_2(\phi_1, \phi_2) \equiv \Theta(-\phi_1, -\phi_2), \\
\delta_2 &= \delta_2(\phi_1, \phi_2) \equiv \delta(-\phi_1, -\phi_2), \\
\chi_2 &= \chi_2(\phi_1, \phi_2) \equiv \frac{\Theta_2 + \delta_2}{\Theta_2 - \delta_2}.
\end{aligned} \tag{3.46}$$

Proof. See Appendix B.5. ∎

Having simplified $\int_0^\infty G(\tau, y_1, y_2, b; S_1, S_2, v)\,\mathrm{d}b$, we can rewrite $C^E(\tau, S_1, S_2, v)$ as

$$\begin{aligned}
C^E(\tau, S_1, S_2, v) &= e^{-r\tau}\iint_{\mathbb{R}^2}(e^{y_1} - e^{y_2})^+\left[\int_0^\infty G(\tau, y_1, y_2, b; S_1, S_2, v)\,\mathrm{d}b\right]\mathrm{d}y_1\,\mathrm{d}y_2 \\
&= \frac{e^{-r\tau}}{4\pi^2}\int_{-\infty}^{\infty}\int_{y_2}^{\infty}(e^{y_1} - e^{y_2})\sum_{m,n=0}^{\infty}\frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!}e^{-\left(\tilde{\lambda}_1+\tilde{\lambda}_2\right)\tau} \\
&\qquad \times \mathbb{E}_{\mathbb{Q}}\Bigg[\iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2}\exp\Big\{ -i\phi_1\Big(\ln S_1 + \Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1\tau\Big) \\
&\qquad - i\phi_2\Big(\ln S_2 + \Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2\tau\Big) + B_1(\tau, -\phi_1, -\phi_2) \\
&\qquad + B_2(\tau, -\phi_1, -\phi_2)v\Big\}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2\Bigg]\mathrm{d}y_1\,\mathrm{d}y_2.
\end{aligned}$$

Note that the rearrangement of the order of integration is allowed since the pay-off function is independent of the variance v. If we define the function

$$f_2(\tau, z_1, z_2, v; \phi_1, \phi_2) = \exp\left\{i\phi_1 \ln z_1 + i\phi_2 \ln z_2 + B_1(\tau, \phi_1, \phi_2) + B_2(\tau, \phi_1, \phi_2)v\right\}, \tag{3.47}$$

then we can write

$$\begin{aligned} C^E(\tau, S_1, S_2, v) = \frac{e^{-r\tau}}{4\pi^2} \sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!} e^{-\left(\tilde{\lambda}_1+\tilde{\lambda}_2\right)\tau} \\ \mathbb{E}_{\mathbb{Q}}\left[\iint_{\mathbb{R}^2} f_2\left(\tau, S_1 e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau}, v; -\phi_1, -\phi_2\right)\right. \\ \left. \times \int_{-\infty}^{\infty}\int_{y_2}^{\infty} (e^{y_1} - e^{y_2}) e^{i\phi_1 y_1 + i\phi_2 y_2}\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}\phi_1\, \mathrm{d}\phi_2\right]. \end{aligned} \tag{3.48}$$

Equation (3.38) can also be used to evaluate other European-type claims provided that the terminal payoff function is dependent only the asset prices observed at maturity. In fact, this structure of the payoff function is what allows us to consider only the integral of the transition density function G with respect to v in the derivation of the European exchange option price. If the payoff function is also dependent on the instantaneous variance, then it has to be included in the integration with respect to v, which may complicate the required calculations.

We take note of the following result in Shephard (1991a) and Shephard (1991b) to assist in evaluating the integrals in equation (3.48).

Lemma 3.12. *Suppose $f(x_1, x_2)$ is a Lebesgue-integrable bivariate density function for which a mean exists and let $\hat{f}(\phi_1, \phi_2)$ be the corresponding characteristic function defined as*

$$\hat{f}(\phi_1, \phi_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\phi_1 x_1 + i\phi_2 x_2} f(x_1, x_2)\, \mathrm{d}x_1, \mathrm{d}x_2.$$

Furthermore, define the joint distribution function $F(x_1, x_2)$ by

$$F(x_1, x_2) = \int_{-\infty}^{x_2}\int_{-\infty}^{x_1} f(u_1, u_2)\, \mathrm{d}u_1\, \mathrm{d}u_2$$

and the marginal distribution functions $F_1(x)$ and $F_2(x)$ by

$$F_1(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x_1, x_2)\, \mathrm{d}x_2\, \mathrm{d}x_1 \quad \textit{and} \quad F_2(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2.$$

Then it holds that

$$\begin{aligned}
&4F(x_1,x_2) - 2[F_1(x_1) + F_2(x_2)] + 1 \\
&\quad = \frac{2^2}{(2\pi)^2}\int_0^\infty\int_0^\infty \Delta_{\phi_1}\Delta_{\phi_2}\left[\frac{\hat{f}(\phi_1,\phi_2)e^{-i\phi_1x_1-i\phi_2x_2}}{i\phi_1 i\phi_2}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\
&\quad = \frac{(-2)^3}{(2\pi)^2}\int_0^\infty\int_0^\infty \Delta_{\phi_2}\Re\left[\frac{\hat{f}(\phi_1,\phi_2)e^{-i\phi_1x_1-i\phi_2x_2}}{\phi_1\phi_2}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2,
\end{aligned} \tag{3.49}$$

where $\Delta_t g(t) = g(t) + g(-t)$

When expanded, the integrand in the second line of equation (3.49) becomes

$$\begin{aligned}
&\Delta_{\phi_1}\Delta_{\phi_2}\left[\frac{\hat{f}(\phi_1,\phi_2)e^{-i\phi_1x_1-i\phi_2x_2}}{i\phi_1 i\phi_2}\right] \\
&\quad = \Delta_{\phi_1}\left[\frac{\hat{f}(\phi_1,\phi_2)e^{-i\phi_1x_1-i\phi_2x_2}}{i\phi_1 i\phi_2} + \frac{\hat{f}(\phi_1,-\phi_2)e^{-i\phi_1x_1+i\phi_2x_2}}{i\phi_1(-i\phi_2)}\right] \\
&\quad = \frac{\hat{f}(\phi_1,\phi_2)e^{-i\phi_1x_1-i\phi_2x_2}}{i\phi_1 i\phi_2} + \frac{\hat{f}(\phi_1,-\phi_2)e^{-i\phi_1x_1+i\phi_2x_2}}{i\phi_1(-i\phi_2)} + \frac{\hat{f}(-\phi_1,\phi_2)e^{i\phi_1x_1-i\phi_2x_2}}{(-i\phi_1)i\phi_2} \\
&\qquad + \frac{\hat{f}(-\phi_1,-\phi_2)e^{i\phi_1x_1+i\phi_2x_2}}{(-i\phi_1)(-i\phi_2)} \\
&\quad = \frac{1}{i\phi_1 i\phi_2}\Big[\hat{f}(\phi_1,\phi_2)e^{-i\phi_1x_1-i\phi_2x_2} - \hat{f}(\phi_1,-\phi_2)e^{-i\phi_1x_1+i\phi_2x_2} - \hat{f}(-\phi_1,\phi_2)e^{i\phi_1x_1-i\phi_2x_2} \\
&\qquad + \hat{f}(-\phi_1,-\phi_2)e^{i\phi_1x_1+i\phi_2x_2}\Big].
\end{aligned}$$

Before we continue to simplify the integrals appearing in equation (3.48), we require some additional notations that will facilitate the use the inversion formulas in Lemma 3.12.

Definition 3.13. In connection to the function $f_2(\tau, z_1, z_2, v; \phi_1, \phi_2)$ that appears in equation (3.48), define the functions $f_{21}(\phi_1,\phi_2)$ and $f_{22}(\phi_1,\phi_2)$ as

$$\begin{aligned}
f_{2j}(\phi_1,\phi_2) &= f_{2j}(\tau, z_1, z_2, v; \phi_1, \phi_2) \\
&= \exp\left\{i\phi_1 \ln z_1 + i\phi_2 \ln z_2 + B_{1j}(\tau,\phi_1,\phi_2) + B_{2j}(\tau,\phi_1,\phi_2)v\right\},
\end{aligned} \tag{3.50}$$

where

$$\begin{aligned}
B_{1j}(\tau,\phi_1,\phi_2) &= i\phi_1(r-q_1)\tau + i\phi_2(r-q_2)\tau \\
&\qquad + \frac{\xi\eta}{\omega^2}\left\{\left(\Theta_{2j}+\delta_{2j}\right)\tau - 2\ln\left[\frac{1-\chi_{2j}e^{\delta_{2j}\tau}}{1-\chi_{2j}}\right]\right\} \\
B_{2j}(\tau,\phi_2,\phi_2) &= \frac{\Theta_{2j}+\delta_{2j}}{\omega^2}\left[\frac{1-e^{\delta_{2j}\tau}}{1-\chi_{2j}e^{\delta_{2j}\tau}}\right]
\end{aligned} \tag{3.51}$$

for $j = 1, 2$, and

$$\begin{aligned}
\Theta_{21} &= \Theta_{21}(\phi_1, \phi_2) \equiv \Theta_2(\phi_1 - i, \phi_2), \\
\Theta_{22} &= \Theta_{22}(\phi_1, \phi_2) \equiv \Theta_2(\phi_1, \phi_2 - i), \\
\delta_{21} &= \delta_{21}(\phi_1, \phi_2) \equiv \delta_2(\phi_1 - i, \phi_2), \\
\delta_{22} &= \delta_{22}(\phi_1, \phi_2) \equiv \delta_2(\phi_1, \phi_2 - i), \\
\chi_{21} &= \chi_{21}(\phi_1, \phi_2) \equiv \chi_2(\phi_1 - i, \phi_2), \\
\chi_{22} &= \chi_{22}(\phi_1, \phi_2) \equiv \chi_2(\phi_1, \phi_2 - i).
\end{aligned} \tag{3.52}$$

The functions Θ_2, δ_2, and χ_2 are defined in equation (3.46).

Having established the above notation, we now present and prove the following proposition for the price of the European exchange option.

Proposition 3.14. *The price of the European exchange option $C^E(\tau, S_1, S_2, v)$ can be written as*

$$\begin{aligned}
C^E(\tau, S_1, S_2, v) = \sum_{m,n=0}^{\infty} &\frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau} \\
&\times \mathbb{E}_{\mathbb{Q}}\Big[S_1 e^{-q_1\tau} e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau} P_1\left(\tau, S_1 e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau}, v\right) \\
&\quad - S_2 e^{-q_2\tau} e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau} P_2\left(\tau, S_1 e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau}, v\right)\Big].
\end{aligned} \tag{3.53}$$

where the functions P_1 and P_2 are defined as

$$\begin{aligned}
P_1(\tau, z_1, z_2, v) &= \frac{1}{2} - \frac{1}{4\pi^2} \lim_{c_1,c_2\to\infty} \iint_{\mathbb{R}^2} f_{21}(\tau, z_1, z_2, v; -\phi_1, -\phi_2) \\
&\qquad \times \frac{e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2}}{i\phi_1(i\phi_1 + i\phi_2)} \,\mathrm{d}\phi_1 \,\mathrm{d}\phi_2, \\
P_2(\tau, z_1, z_2, v) &= \frac{1}{2} - \frac{1}{4\pi^2} \lim_{d_1,d_2\to\infty} \iint_{\mathbb{R}^2} f_{22}(\tau, z_1, z_2, v; -\phi_1, -\phi_2) \\
&\qquad \times \frac{e^{i\phi_1 d_1 + i\phi_2 d_1} - e^{-i\phi_1 d_2 - i\phi_2 d_2}}{i\phi_1(i\phi_1 + i\phi_2)} \,\mathrm{d}\phi_1 \,\mathrm{d}\phi_2,
\end{aligned} \tag{3.54}$$

and f_{21} and f_{22} (and their associated functions) are defined in Definition 3.13.

Proof. See Appendix B.6. ■

Equation (3.53) expresses the price of the European exchange option as an infinite sum of Poisson-weighted expectations of Margrabe-type formulas (referring to the expressions in the square brackets), similar to the findings of Cheang and Chiarella (2011). In particular, the right-hand side of equation (3.53) may be seen as an expectation conditioned on the number of jumps observed in stocks 1 and 2, given by m and n respectively, over the life of the option. The quantity that occurs in the Margrabe-type formulas are dependent on the stock prices, their dividend yields, the jump

parameters, the accumulation of jumps (given by $\Upsilon_{1,m}$ and $\Upsilon_{2,n}$), and the instantaneous variance level v. At this point, we have left unspecified the distribution of the jump sizes. As such, one may provide a specific jump size distribution (for example, a log-normal distribution (Merton, 1976) or an asymmetric double-exponential distribution (Kou, 2002)) replace the expectation operators with the appropriate integrals with respect to the joint density function of $\Upsilon_{1,m}$ and $\Upsilon_{2,n}$.

We note further that the European exchange option may also be presented in the form

$$C^E(\tau, S_1, S_2, v) = S_1 e^{-q_1\tau} Q_1(\tau, S_1, S_2, v) - S_2 e^{-q_2\tau} Q_1(\tau, S_1, S_2, v), \tag{3.55}$$

where

$$\begin{aligned} Q_1(\tau, S_1, S_2, v) = &\sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau} \\ &\times \mathbb{E}_{\mathbb{Q}}\left[e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1} P_1\left(\tau, S_1 e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau}, v\right)\right] \\ Q_2(\tau, S_1, S_2, v) = &\sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau} \\ &\times \mathbb{E}_{\mathbb{Q}}\left[e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2} P_2\left(\tau, S_1 e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau}, v\right)\right]. \end{aligned} \tag{3.56}$$

This expression emphasizes that our representation is similar to the result we obtained in Section 2.5.1 using the change-of-numéraire technique. When written in the above form, Q_1 and Q_2 may then be interpreted as probabilities that the exchange option will be in the money at maturity under probability measures $\mathbb{Q}_1$ and $\mathbb{Q}_2$ (equivalent to the risk-neutral measure $\mathbb{Q}$), respectively, which are obtained by taking S_1 and S_2, respectively, as the numéraire process. Similar to the original Margrabe (1978) result, our representation for the European exchange option price is independent of the risk-free rate r. We further note that equation (3.55) resembles the single-asset European option price representations obtained by Heston (1993) and Bates (1996).

3.8 The American Exchange Option Price

Having derived an integral representation for the price of the European exchange option, we now characterize the American exchange option price and the associated early exercise premium.

We recall that the free-boundary problem for the American exchange option price $C^A \equiv C^A(\tau, S_1, S_2, v)$, *now to be written in terms of the time-to-maturity variable* $\tau \equiv T - t$, is given by

$$\frac{\partial C^A}{\partial \tau} = -rC^A + \mathcal{L}_{\mathbf{s},v} C^A, \tag{3.57}$$

where $\mathcal{L}_{\mathbf{s},v}$ is the IPDE operator defined by equation (3.11).[4] This IPDE is to be solved for

[4] Similar to the notation used in Section 3.7, it is understood from this point forward that S_1 and S_2 are the

$0 \leq \tau \leq T$, $0 < S_1 < B(\tau, v)S_2$, $0 < S_2 < \infty$, and $0 < v < \infty$, subject to the initial and boundary conditions

$$\begin{aligned} C^A(0, S_1, S_2, v) &= (S_1 - S_2)^+, \\ C^A(\tau, 0, S_2, v) &= 0, \\ C^A(\tau, S_1, 0, v) &= S_1, \end{aligned} \tag{3.58}$$

the value-matching condition

$$C^A(\tau, S_1, S_2, v) = S_1 - S_2, \qquad S_1 \geq B(\tau, v)S_2, (\tau, S_2, v) \in [0, T] \times \mathbb{R}^2_+, \tag{3.59}$$

and the smooth-pasting conditions

$$\begin{aligned} \lim_{S_1 \to B(\tau,v)S_2} \frac{\partial C^A}{\partial s_1}(\tau, S_1, S_2, v) &= 1, \\ \lim_{S_1 \to B(\tau,v)S_2} \frac{\partial C^A}{\partial s_2}(\tau, S_1, S_2, v) &= -1. \end{aligned} \tag{3.60}$$

As emphasized in Section 3.3, the value-matching and the smooth-pasting conditions follow from the continuity of the option price and its first-order partial derivatives in S_1 and S_2 implied by Assumption 3.3. Since Assumption 3.3 also implies the continuity of $\frac{\partial C^A}{\partial \tau}$ and $\frac{\partial C^A}{\partial v}$ across the early exercise boundary, we also have

$$\lim_{S_1 \to B(\tau,v)S_2} \frac{\partial C^A}{\partial \tau}(\tau, S_1, S_2, v) = \lim_{S_1 \to B(\tau,v)S_2} \frac{\partial C^A}{\partial v}(\tau, S_1, S_2, v) = 0.$$

As in Jamshidian (1992), we derive an inhomogeneous IPDE for C^A that will be solved over an unrestricted stock price domain.

Proposition 3.15. *The American exchange option price $C^A \equiv C^A(\tau, S_1, S_2, v)$ satisfies the inhomogeneous IPDE*

$$\frac{\partial C^A}{\partial \tau} = -rC^A + \mathcal{L}_{\mathbf{s},v} C^A + \Xi(\tau, S_1, S_2, v), \tag{3.61}$$

prices of stocks 1 and 2 (immediately before any jumps) at time to maturity τ. The instantaneous variance level at time to maturity τ is denoted by v. All partial derivatives involved are evaluated at (τ, S_1, S_2, v).

where the inhomogeneous term Ξ *is given by*[5]

$$\begin{aligned}\Xi(\tau, S_1, S_2, v) \equiv \mathbf{1}_{\mathcal{A}_\tau} \Bigg\{ & q_1 S_1 - q_2 S_2 \\ & - \tilde{\lambda}_1 \int_{-\infty}^{\ln[B(\tau,v)S_2/S_1]} \left[C^A\left(\tau, S_1 e^y, S_2, v\right) - (S_1 e^y - S_2) \right] g_1(y)\, \mathrm{d}y \\ & - \tilde{\lambda}_2 \int_{-\ln[B(\tau,v)S_2/S_1]}^{\infty} \left[C^A\left(\tau, S_1, S_2 e^y, v\right) - (S_1 - S_2 e^y) \right] g_2(y)\, \mathrm{d}y \Bigg\},\end{aligned} \tag{3.62}$$

$\mathcal{A}_\tau = \{S_1 \geq B(\tau, v)S_2\}$ *is the event that* (S_1, S_2) *is in the stopping region at time-to-maturity* τ, $\mathbf{1}_{(\cdot)}$ *is the indicator function, and* $g_j(y)$ *is the probability density function of the jump random variable* Y_j $(j = 1, 2)$ *under the equivalent martingale measure* $\mathbb{Q}$.

This IPDE is to be solved over the domain $0 \leq \tau \leq T$ *and* $(S_1, S_2, v) \in \mathbb{R}^3_+$, *subject to boundary conditions* (3.58), *value-matching condition* (3.59), *and smooth-pasting conditions* (3.60).

Proof. IPDE 3.57 is a homogeneous equation defined for S_1 and S_2 on a *bounded* domain for which the boundary (i.e. the early exercise boundary) has to be determined as part of the solution. In this proposition, we transform the preceding IPDE into an inhomogeneous equation defined on an *unrestricted* domain in the stock prices, following the method of Jamshidian (1992).[6]

When S_1 and S_2 are in the stopping region, i.e. $S_1 \geq B(\tau, v)S_2$, then $C^A = S_1 - S_2$. Evaluating equation (3.57) in this region with this expression for C^A gives

$$\begin{aligned} r(S_1 - S_2) = (r - q_1 - \tilde{\lambda}_1\tilde{\kappa}_1)S_1 - (r - q_2 - \tilde{\lambda}_2\tilde{\kappa}_2)S_2 \\ + \tilde{\lambda}_1 \mathbb{E}^{Y_1}_{\mathbb{Q}} \left[C^A\left(\tau, S_1 e^{Y_1}, S_2, v\right) - (S_1 - S_2) \right] \\ + \tilde{\lambda}_2 \mathbb{E}^{Y_2}_{\mathbb{Q}} \left[C^A\left(\tau, S_1, S_2 e^{Y_2}, v\right) - (S_1 - S_2) \right]. \end{aligned}$$

The smooth-pasting conditions (3.60) ensure that no singularities in the form of Dirac-delta functions appear when differentiating the known analytical form of the American option price in the stopping region (Jamshidian, 1992; Myneni, 1992). Recalling that $\tilde{\kappa}_i \equiv \mathbb{E}^{Y_i}_{\mathbb{Q}}[e^{Y_i} - 1]$, the last equation simplifies to

$$\begin{aligned} 0 = -q_1 S_1 + q_2 S_2 + \tilde{\lambda}_1 \mathbb{E}^{Y_1}_{\mathbb{Q}} \left[C^A\left(\tau, S_1 e^{Y_1}, S_2, v\right) - \left(S_1 e^{Y_1} - S_2\right) \right] \\ + \tilde{\lambda}_2 \mathbb{E}^{Y_2}_{\mathbb{Q}} \left[C^A\left(\tau, S_1, S_2 e^{Y_2}, v\right) - \left(S_1 - S_2 e^{Y_2}\right) \right]. \end{aligned}$$

[5] Recall that, for $i = 1, 2$, q_i is the constant dividend yield for stock i and g_i is the probability density function of the jump random variable Y_i under $\mathbb{Q}$.

[6] This approach has been applied to the analysis of single-asset American call options Chiarella and Ziogas (2009) (under jump-diffusion dynamics), Adolfsson et al. (2013) (under Heston's stochastic volatility model), and Chiarella and Ziveyi (2013) (under a two stochastic volatility processes). The discussion here also extends, to the case of SVJD dynamics, the work of Chiarella and Ziveyi (2014) who used this technique to derive an inhomogeneous equation for the price of American spread options under pure-diffusion dynamics.

Suppose Y_i has a probability density function $g_i(y)$ under $\mathbb{Q}$ $(i = 1, 2)$. Note further that if $S_1 e^{Y_1} \geq B(\tau, v)S_2$, or if $Y_1 \geq \ln[B(\tau, v)S_2/S_1]$, then

$$C^A\left(\tau, S_1 e^{Y_1}, S_2, v\right) = S_1 e^{Y_1} - S_2.$$

Likewise, if $S_1 \geq B(\tau, v)S_2 e^{Y_2}$, or if $Y_2 \leq -\ln[B(\tau, v)S_2/S_1]$, then

$$C^A\left(\tau, S_1, S_2 e^{Y_2}, v\right) = S_1 - S_2 e^{Y_2}.$$

Therefore, the expectations may be written in terms of integrals as

$$\begin{aligned} 0 = -q_1 S_1 + q_2 S_2 &+ \tilde{\lambda}_1 \int_{-\infty}^{\ln[B(\tau,v)S_2/S_1]} \left[C^A\left(\tau, S_1 e^y, S_2, v\right) - (S_1 e^y - S_2)\right] g_1(y)\,\mathrm{d}y \\ &+ \tilde{\lambda}_2 \int_{-\ln[B(\tau,v)S_2/S_1]}^{\infty} \left[C^A\left(\tau, S_1, S_2 e^y, v\right) - (S_1 - S_2 e^y)\right] g_2(y)\,\mathrm{d}y. \end{aligned}$$

Given this analysis, we may now define the inhomogeneous term as

$$\begin{aligned} \Xi(\tau, S_1, S_2, v) \equiv \mathbf{1}_{\mathcal{A}_\tau} \Bigg\{ & q_1 S_1 - q_2 S_2 \\ & - \tilde{\lambda}_1 \int_{-\infty}^{\ln[B(\tau,v)S_2/S_1]} \left[C^A\left(\tau, S_1 e^y, S_2, v\right) - (S_1 e^y - S_2)\right] g_1(y)\,\mathrm{d}y \\ & - \tilde{\lambda}_2 \int_{-\ln[B(\tau,v)S_2/S_1]}^{\infty} \left[C^A\left(\tau, S_1, S_2 e^y, \mathbf{v}\right) - (S_1 - S_2 e^y)\right] g_2(y)\,\mathrm{d}y \Bigg\}, \end{aligned}$$

where $\mathcal{A}_\tau = \{S_1 \geq B(\tau, v)S_2\}$ is the event where S_1 and S_2 are in the stopping region at time τ (see also the events defined in Proposition 2.17) and $\mathbf{1}_{(\cdot)}$ is the indicator function. Therefore, the inhomogeneous equation corresponding to IPDE (3.57) is given by equation (3.61) which can now be solved in an unrestricted stock price domain $(S_1, S_2) \in \mathbb{R}^2_+$ and for $v > 0$ and $0 \leq \tau < T$. ■

Using Duhamel's principle, one can infer the solution of the inhomogeneous IPDE from the solution of its homogeneous counterpart.[7] Note that the solution of the homogeneous version of equation (3.61) (i.e. if $\Xi \equiv 0$) is the price of the European exchange option $C^E \equiv C^E(\tau, S_1, S_2, v)$, which is given is Proposition 3.9. The following proposition provides a representation of C^A, the solution of the inhomogeneous equation, as the sum of C^E and the early exercise premium denoted by C^P.

The derivation of the inhomogeneous equation requires the following lemma on an IPDE satisfied by the function G defined in (3.39). This result uses the link between G and the transition density

[7] Duhamel's principle has been applied by a number of studies on American options, particularly those by Adolfsson et al. (2013), Chiarella et al. (2015), and Chiarella and Ziveyi (2014), to derive the early exercise representation for the American option price. A general discussion of Duhamel's principle for ordinary and partial differential equations can be found in Logan (2015).

function $\wp$ of the vector stochastic process $(S_1, S_2, v)^\top$ whose dynamics under $\mathbb{Q}$ are given by SDEs (3.8) and (3.9). In particular, this link arises from the probabilistic definition of C^E as an expected value. In Appendix A.5, we review the definition of the transition density function $\wp$ and derive the corresponding Kolmogorov backward equation corresponding to our proportional SVJD model.

Lemma 3.16. *The function $G(\tau, y_1, y_2, b; S_1, S_2, v)$ in (3.38) satisfies the IPDE*

$$\frac{\partial G}{\partial \tau} = \mathcal{L}_{\mathbf{s},v} G \tag{3.63}$$

for every $(y_1, y_2, b) \in \mathbb{R}^2 \times \mathbb{R}_+$.

Proof. We recall the probabilistic definition of the European exchange option, i.e.[8]

$$\begin{aligned}\hat{C}^E(t, S_1, S_2, v) &= e^{-r(T-t)} \mathbb{E}^{(t,S_1,S_2,v)} \left[(S_1(T) - S_2(T))^+ \right] \\ &= e^{-r(T-t)} \int_{\mathbb{R}^3_+} (u_1 - u_2)^+ \wp(T, u_1, u_2, b; t, S_1, S_2, v) \, \mathrm{d}u_1 \, \mathrm{d}u_2 \, \mathrm{d}b,\end{aligned}$$

where the second equality follows from (A.27). Introducing the change of variables $\tau = T - t$, $u_1 = e^{y_1}$ and $u_2 = e^{y_2}$, we obtain

$$C^E(\tau, S_1, S_2, v) = e^{-r\tau} \int_0^\infty \int_{\mathbb{R}^2} (e^{y_1} - e^{y_2})^+ \wp(0, e^{y_1}, e^{y_2}, b; T - \tau, S_1, S_2, v) e^{y_1} e^{y_2} \, \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}b. \tag{3.64}$$

Comparing equations (3.38) and (3.64), we deduce that

$$G(\tau, y_1, y_2, b; S_1, S_2, v) = e^{y_1 + y_2} \wp(0, e^{y_1}, e^{y_2}, b; T - \tau, S_1, S_2, v)$$

for every $\tau \in [0, T]$, $(y_1, y_2, b) \in \mathbb{R}^2 \times \mathbb{R}_+$ and $(S_1, S_2, v) \in \mathbb{R}^3$. Computing the requisite derivatives, we find that

$$\mathcal{L}_{\mathbf{s},v} G(\tau, y_1, y_2, b; S_1, S_2, v) = e^{y_1 + y_2} \mathcal{L}_{\mathbf{s},v} \wp(0, e^{y_1}, e^{y_2}, b; T - \tau, S_1, S_2, v)$$

and

$$\frac{\partial G}{\partial \tau} = \frac{\partial \wp}{\partial \tau} = -\frac{\partial \wp}{\partial t}.$$

Thus,

$$\frac{\partial G}{\partial \tau} - \mathcal{L}_{\mathbf{s},v} G = -e^{y_1 + y_2} \left(\frac{\partial \wp}{\partial t} + \mathcal{L}_{\mathbf{s},v} \wp \right) = 0$$

by Lemma A.13. Therefore, for fixed $(y_1, y_2, b) \in \mathbb{R}^2 \times \mathbb{R}_+$, G satisfies the IPDE (3.63). ∎

[8] We use $\hat{C}^E$ instead of C^E since the definition of C^E in (3.38) expresses the European exchange option price in terms of the time to maturity τ.

Proposition 3.17. *The solution C^A to the inhomogeneous IPDE in Proposition 3.15 is given by*

$$C^A(\tau, S_1, S_2, v) = C^E(\tau, S_1, S_2, v) + C^P(\tau, S_1, S_2, v), \tag{3.65}$$

where C^E is the price of the European exchange option given in equation 3.38 and C^P is the early exercise premium given by

$$\begin{aligned} C^P(\tau, S_1, S_2, v) = \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \iint_{\mathbb{R}^2} \Xi(u, e^{y_1}, e^{y_2}, b) \\ \times G(\tau - u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u, \end{aligned} \tag{3.66}$$

where $G(\tau, y_1, y_2, b; S_1, S_2, b)$ is given by equation (3.39).

Proof. We need to show that C^A, as given by equation (3.65), is a solution to the inhomogeneous equation (3.61). Based on the analysis in Sections 3.3 to 3.7, it was shown that C^E, as given by equation (3.38), is a solution to the IPDE

$$\frac{\partial C^E}{\partial \tau} = -rC^E + \mathcal{L}_{\mathbf{s},v} C^E.$$

Substituting the early exercise representation (3.65) to the inhomogeneous IPDE (3.61), we find that

$$\begin{aligned} &\frac{\partial C^A}{\partial \tau} - \mathcal{L}_{\mathbf{s},v} C^A + rC^A - \Xi(\tau, S_1, S_2, v) \\ &\quad = \left(\frac{\partial C^E}{\partial \tau} - \mathcal{L}_{\mathbf{s},v} C^E + rC^E \right) + \left(\frac{\partial C^P}{\partial \tau} - \mathcal{L}_{\mathbf{s},v} C^P + rC^P \right) - \Xi(\tau, S_1, S_2, v) \\ &\quad = \frac{\partial C^P}{\partial \tau} - \mathcal{L}_{\mathbf{s},v} C^P + rC^P - \Xi(\tau, S_1, S_2, v) \end{aligned}$$

(these follow from the linearity of the operator $\mathcal{L}_{\mathbf{s},v}$). We also have

$$\mathcal{L}_{\mathbf{s},v} C^P = \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \iint_{\mathbb{R}^2} \Xi(u, e^{y_1}, e^{y_2}, b) \mathcal{L}_{\mathbf{s},v} G(\tau - u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u.$$

Using Lemma 3.16 and the Leibniz rule for differentiation under the integral sign[9], we obtain

$$\begin{aligned} \frac{\partial C^P}{\partial \tau} = -rC^P(\tau, S_1, S_2, v) + e^{-r\tau} \frac{\partial}{\partial \tau} \Bigg[\int_0^\tau e^{ru} \int_0^\infty \iint_{\mathbb{R}_2} \Xi(u, e^{y_1}, e^{y_2}, b) \\ \times G(\tau - u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u \Bigg] \end{aligned}$$

[9] Suppose $a(x)$ and $b(x)$ are continuously differentiable for $x_0 \le x \le x_1$ and suppose $f(t, x)$ and $\partial f(t, x)/\partial x$ are continuous on some region in the tx-plane including the region $\{(t, x) : a(x) \le t \le b(x), x_0 \le x \le x_1\}$. Then for $x_0 \le x \le x_1$,

$$\frac{\partial}{\partial x} \int_{a(x)}^{b(x)} f(t, x)\, \mathrm{d}t = f(b(x), x) b'(x) - f(a(x), x) a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f}{\partial x}(t, x)\, \mathrm{d}t.$$

$$
\begin{aligned}
&= -rC^P(\tau, S_1, S_2, v) + \int_0^\infty \iint_{\mathbb{R}^2} \Xi(\tau, e^{y_1}, e^{y_2}, b) G(0, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b \\
&\quad + \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \iint_{\mathbb{R}^2} \Xi(u, e^{y_1}, e^{y_2}, b) \frac{\partial G}{\partial \tau}(\tau - u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u \\
&= -rC^P(\tau, S_1, S_2, v) \\
&\quad + \int_0^\infty \iint_{\mathbb{R}^2} \Xi(\tau, e^{y_1}, e^{y_2}, b) \delta(y_1 - \ln s_1) \delta(y_2 - \ln s_2) \delta(b - v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b \\
&\quad + \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \iint_{\mathbb{R}^2} \Xi(u, e^{y_1}, e^{y_2}, b) \frac{\partial G}{\partial \tau}(\tau - u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u \\
&= -rC^P(\tau, S_1, S_2, v) + \Xi(\tau, S_1, S_2, v) \\
&\quad + \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \iint_{\mathbb{R}^2} \Xi(u, e^{y_1}, e^{y_2}, b) \frac{\partial G}{\partial \tau}(\tau - u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\frac{\partial C^A}{\partial \tau} - \mathcal{L}_{\mathbf{s},v} C^A + rC^A - \Xi(\tau, S_1, S_2, v) \\
&\quad = \frac{\partial C^P}{\partial \tau} - \mathcal{L}_{\mathbf{s},v} C^P + rC^P - \Xi(\tau, S_1, S_2, v) \\
&\quad = -rC^P(\tau, S_1, S_2, v) + \Xi(\tau, S_1, S_2, v) \\
&\qquad + \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \iint_{\mathbb{R}^2} \Xi(u, e^{y_1}, e^{y_2}, b) \\
&\qquad \times \left(\frac{\partial G}{\partial \tau}(\tau - u, y_1, y_2, b; S_1, S_2, v) - \mathcal{L}_{\mathbf{s},v} G(\tau - u, y_1, y_2, b; S_1, S_2, v) \right) \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u \\
&\qquad + rC^P(\tau, S_1, S_2, v) - \Xi(\tau, S_1, S_2, v) \\
&\quad = 0.
\end{aligned}
$$

Therefore, C^A, as given by the equation (3.65), is a solution of the inhomogeneous IPDE and is thus an appropriate representation of the price of the American exchange option. Furthermore, equation (3.66) gives us a representation for the early exercise premium associated to the American exchange option. ■

Note that equation (3.65) is consistent with the early exercise representations obtained by Kim (1990), Jacka (1991), Carr et al. (1992), and Gukhal (2001) for American options written on a single underlying asset. Once the early exercise premium C^P in equation (3.66) has been evaluated, we see their interpretations of its components (particularly those discussed by Gukhal (2001) for the jump-diffusion case) also hold under our model.

To facilitate the following calculations for the early exercise premium, we introduce some nota-

tion. First, it is helpful to rewrite the function $G(\tau, y_1, y_2, b; S_1, S_2, v)$ (see equation (3.39)) as

$$\begin{aligned} G(\tau, y_1, y_2, b; S_1, S_2, v) = \sum_{m,n=0}^{\infty} & \frac{(\tilde{\lambda}_1\tau)^m e^{-\tilde{\lambda}_1\tau}}{m!} \frac{(\tilde{\lambda}_2\tau)^n e^{-\tilde{\lambda}_2\tau}}{n!} \\ & \mathbb{E}_{\mathbb{Q}}\Bigg[\frac{1}{4\pi^2}\iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} \exp\Big\{-i\phi_1 \ln\Big(S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1\tau}\Big) \\ & - i\phi_2 \ln\Big(S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2\tau}\Big)\Big\} \exp\Bigg\{-\sum_{j=1}^{2} i\phi_j(r - q_j)\tau\Bigg\} \\ & \times F(\tau, \phi_1, \phi_2, b; v)\, \mathrm{d}\phi_1\, \mathrm{d}\phi_2\Bigg], \end{aligned} \tag{3.67}$$

where $F(\tau, \phi_1, \phi_2, b; v)$ is given by equation (3.40). Next, we define some functions associated to the functions G and F.

Definition 3.18. Given the function $F(\tau, \phi_1, \phi_2, b; v)$ from equation (3.40), define the function $F_1(\tau, \phi_1, \phi_2, b; v)$ by

$$\begin{aligned} F_1(\tau, \phi_1, \phi_2, b; v) &\equiv F(\tau, -\phi_1, -\phi_2, b; v) \\ &= \frac{2\delta_2 e^{\delta_2\tau}}{\omega^2(e^{\delta_2\tau} - 1)} \left(\frac{be^{\delta_2\tau}}{v}\right)^{\frac{\xi\eta}{\omega^2} - \frac{1}{2}} \exp\left\{-\frac{2\delta_2(be^{\delta_2\tau} + v)}{\omega^2(e^{\delta_2\tau} - 1)}\right\} \\ &\quad \times \exp\left\{\left(\frac{\Theta_2 - \delta_2}{\omega^2}\right)(\xi\eta\tau + v - b)\right\} I_{\left(\frac{2\xi\eta}{\omega^2} - 1\right)}\left(\frac{4\delta_2\sqrt{bve^{\delta_2\tau}}}{\omega^2(e^{\delta_2\tau} - 1)}\right), \end{aligned} \tag{3.68}$$

where Θ_2 and δ_2 are given by equation (3.46) and $I_k(u)$ is the modified Bessel function given by equation (3.36). Define the function h by

$$\begin{aligned} h(\tau, z_1, z_2, v; \phi_1, \phi_2, b) \equiv & \exp\left\{i\phi_1 \ln z_1 + i\phi_2 \ln z_2 + i\phi_1(r - q_1)\tau + i\phi_2(r - q_2)\tau\right\} \\ & \times F_1(\tau, \phi_1, \phi_2, b; v). \end{aligned} \tag{3.69}$$

Associated to the function h, define the function $h_j(\tau, z_1, z_2, v; \phi_1, \phi_2, b)$, for $j = 1, 2$, by

$$\begin{aligned} h_j(\tau, z_1, z_2, v; \phi_1, \phi_2, b) \equiv & \exp\left\{i\phi_1 \ln z_1 + i\phi_2 \ln z_2 + i\phi_1(r - q_1)\tau + i\phi_2(r - q_2)\tau\right\} \\ & \times F_{1j}(\tau, \phi_1, \phi_2, b; v), \end{aligned} \tag{3.70}$$

where

$$\begin{aligned} F_{1j}(\tau, \phi_1, \phi_2, b; v) = & \frac{2\delta_{2j} e^{\delta_{2j}\tau}}{\omega^2(e^{\delta_{2j}\tau} - 1)} \left(\frac{be^{\delta_{2j}\tau}}{v}\right)^{\frac{\xi\eta}{\omega^2} - \frac{1}{2}} \exp\left\{-\frac{2\delta_2(be^{\delta_{2j}\tau} + v)}{\omega^2(e^{\delta_{2j}\tau} - 1)}\right\} \\ & \times \exp\left\{\left(\frac{\Theta_{2j} - \delta_{2j}}{\omega^2}\right)(\xi\eta\tau + v - b)\right\} I_{\left(\frac{2\xi\eta}{\omega^2} - 1\right)}\left(\frac{4\delta_{2j}\sqrt{bve^{\delta_{2j}\tau}}}{\omega^2(e^{\delta_{2j}\tau} - 1)}\right), \end{aligned} \tag{3.71}$$

where Θ_{2j} and δ_{2j}, $j = 1, 2$, are given in Definition (3.13).

We first make some remarks regarding our notation. Note that F_1, F_{11}, and F_{12} are related via the equations

$$\begin{aligned} F_{11}(\tau, \phi_1, \phi_2, b; v) &= F_1(\tau, \phi_1 - i, \phi_2, b; v), \\ F_{12}(\tau, \phi_1, \phi_2, b; v) &= F_1(\tau, \phi_1, \phi_2 - i, b; v). \end{aligned}$$

Furthermore, in terms of the function h (equation (3.69)), we may write G (equation (3.67)) as

$$\begin{aligned} G(\tau, y_1, y_2, b; S_1, S_2, v) = \sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1\tau)^m e^{-\tilde{\lambda}_1\tau}}{m!} \frac{(\tilde{\lambda}_2\tau)^n e^{-\tilde{\lambda}_2\tau}}{n!} \mathbb{E}_{\mathbb{Q}} \Bigg[\frac{1}{4\pi^2} \iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} \\ \times h\left(\tau, S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2\tau}, v; -\phi_1, -\phi_2, b\right) \mathrm{d}\phi_1 \, \mathrm{d}\phi_2 \Bigg]. \end{aligned} \tag{3.72}$$

We now derive and present an expression for the early exercise premium of the American exchange option that is obtained by evaluating the integral in equation (3.66).

Proposition 3.19. *The early exercise premium $C^P(\tau, S_1, S_2, v)$ for the American exchange option can be expressed as*

$$C^P(\tau, S_1, S_2, v) = C_D^P(\tau, S_1, S_2, v) - \tilde{\lambda}_1 C_{J_1}^P(\tau, S_1, S_2, v) - \tilde{\lambda}_2 C_{J_2}^P(\tau, S_1, S_2, v), \tag{3.73}$$

where C_D^P is given by

$$\begin{aligned} &C_D^P(\tau, S_1, S_2, v) \\ &\quad = \sum_{m,n=0}^{\infty} \frac{\tilde{\lambda}_1^m}{m!} \frac{\tilde{\lambda}_2^n}{n!} \mathbb{E}_{\mathbb{Q}} \Bigg[\int_0^{\tau} \int_0^{\infty} (\tau - u)^{m+n} e^{-(\tilde{\lambda}_1 + \tilde{\lambda}_2)(\tau - u)} \\ &\qquad \times \Big\{ q_1 S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)} e^{-q_1(\tau-u)} \\ &\qquad\quad \times P_1^D\left(\tau - u, S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; b, B(u, b)\right) \\ &\qquad - q_2 S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)} e^{-q_2(\tau-u)} \\ &\qquad\quad \times P_2^D\left(\tau - u, S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; b, B(u, b)\right) \Big\} \, \mathrm{d}b \, \mathrm{d}u \Bigg], \end{aligned} \tag{3.74}$$

with

$$
\begin{aligned}
P_1^D\left(\tau, z_1, z_2, v; b, B(u,b)\right) &= \frac{1}{2} - \frac{1}{4\pi^2}\lim_{c_1,c_2\to\infty}\iint_{\mathbb{R}^2} e^{i\phi_1 \ln B(u,b)} h_1(\tau, z_1, z_2, v; -\phi_1, -\phi_2, b) \\
&\qquad \times \left(\frac{e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2}}{i\phi_1(i\phi_1 + i\phi_2)}\right) \mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\
P_2^D\left(\tau, z_1, z_2, v; b, B(u,b)\right) &= \frac{1}{2} - \frac{1}{4\pi^2}\lim_{d_1,d_2\to\infty}\iint_{\mathbb{R}^2} e^{i\phi_1 \ln B(u,b)} h_2(\tau, z_1, z_2, v; -\phi_1, -\phi_2, b) \\
&\qquad \times \left(\frac{e^{i\phi_1 d_1 + i\phi_2 d_1} - e^{-i\phi_1 d_2 - i\phi_2 d_2}}{i\phi_1(i\phi_1 + i\phi_2)}\right) \mathrm{d}\phi_1\,\mathrm{d}\phi_2,
\end{aligned}
\tag{3.75}
$$

and $C_{J_1}^P$ *and* $C_{J_2}^P$ *are given by*

$$
\begin{aligned}
&C_{J_1}^P(\tau, S_1, S_2, v) \\
&\quad = \sum_{m,n=0}^{\infty} \frac{\tilde{\lambda}_1^m}{m!}\frac{\tilde{\lambda}_2^n}{n!}\mathbb{E}_{\mathbb{Q}}\Bigg[\frac{1}{4\pi^2}\int_0^\tau\int_0^\infty (\tau-u)^{m+n} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(\tau-u)} e^{-r(\tau-u)} \\
&\qquad \times \int_{-\infty}^0 g_1(y)\int_{-\infty}^{\infty}\int_{\ln B(u,b)+y_2}^{\ln B(u,b)+y_2-y}\left[C^A\left(u, e^{y_1+y}, e^{y_2}, b\right) - \left(e^{y_1+y} - e^{y_2}\right)\right] \\
&\qquad \times \iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} h\left(\tau-u, S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; -\phi_1, -\phi_2, b\right) \\
&\qquad \mathrm{d}\phi_1\,\mathrm{d}\phi_2\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}y\,\mathrm{d}b\,\mathrm{d}u\Bigg], \\
&C_{J_2}^P(\tau, S_1, S_2, v) \\
&\quad = \sum_{m,n=0}^{\infty} \frac{\tilde{\lambda}_1^m}{m!}\frac{\tilde{\lambda}_2^n}{n!}\mathbb{E}_{\mathbb{Q}}\Bigg[\frac{1}{4\pi^2}\int_0^\tau\int_0^\infty (\tau-u)^{m+n} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(\tau-u)} e^{-r(\tau-u)} \\
&\qquad \times \int_{0}^{\infty} g_2(y)\int_{-\infty}^{\infty}\int_{\ln B(u,b)+y_2+y}^{\ln B(u,b)+y_2}\left[C^A\left(u, e^{y_1}, e^{y_2+y}, b\right) - \left(e^{y_1} - e^{y_2+y}\right)\right] \\
&\qquad \times \iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} h\left(\tau-u, S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; -\phi_1, -\phi_2, b\right) \\
&\qquad \mathrm{d}\phi_1\,\mathrm{d}\phi_2\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}y\,\mathrm{d}b\,\mathrm{d}u\Bigg].
\end{aligned}
\tag{3.76}
$$

The functions h, h_1, *and* h_2 *that appear in the above representations are defined in Definition 3.18. Note that the expectation operator acts on the random variables* $\Upsilon_{1,m}$ *and* $\Upsilon_{2,n}$ *only.*

PROOF. See Appendix B.7. ■

The expression for the early exercise boundary in equation (3.73) is consistent with its the probabilistic representation given in equation (2.78). Comparing the two equations, we may interpret C_P^D as the discounted expected value of the cash flow $q_1 S_1 - q_2 S_2$ whenever the asset prices are in the stopping region. Furthermore, $C_{J_i}^P$ $(i = 1, 2)$ represents the discounted expected value of the cost, referred to as the *rebalancing cost* by Gukhal (2001), incurred by the holder of the option when asset prices suddenly move back into the continuation region as a result of jumps

in the price of asset i at time τ. [10] When such sudden movements occur, the amount obtained when the option is exercised has less value than the live/unexercised American exchange option. Furthermore, if jumps are not present (i.e. $\tilde{\lambda}_1 = \tilde{\lambda}_2 = 0$), then $C^P_{J_i}$ no longer contributes to the early exercise premium, and so the premium arises only from the diffusion part of the asset prices. These interpretations are similar to those given by Gukhal (2001), Cheang and Chiarella (2011), and Cheang et al. (2013).

The value-matching condition (3.59) allows us to evaluate $C^A(\tau, S_1, S_2, v)$ at the early exercise boundary $S_1 = B(\tau, v)S_2$, giving us an integral equation for the unknown slope of the boundary $B(\tau, v)$. This leads to the following corollary.

Corollary 3.20. *The critical stock price ratio $B(\tau, v)$ that determines the early exercise boundary $S_1 = B(\tau, v)S_2$ is the solution to the integral equation*

$$S_2\left(B(\tau,v)-1\right) = C^E\left(\tau, B(\tau,v)S_2, S_2, v\right) + C^P\left(\tau, B(\tau,v)S_2, S_2, v\right), \tag{3.77}$$

where C^E and C^P are given in Propositions 3.14 and 3.19, respectively.

Proof. The integral equation is obtained by evaluating equations (3.59) and (3.65) at the early exercise boundary $S_1 = B(\tau, v)S_2$ and equating the results as they are both equal to the price of the American exchange option. ■

The integral equation (3.77) cannot be solved sequentially since C^P is dependent on the unknown American exchange option value C^A. This dependence is present due to the rebalancing costs incurred from the possibility of jumps in the asset prices. Equations (3.65) and (3.77) therefore comprise a *linked* system of (integral) equations that characterizes the American exchange option price and the critical stock price ratio in the corresponding early exercise boundary.

3.9 Exchange Option Prices under Log-Normal Jumps

In this section, we consider the exchange option price representations that result from assuming that the jumps in the asset prices have a log-normal distribution, following the analysis of Merton (1976). Determining exchange option prices under this assumption is important when we consider a numerical implementation of our exchange option pricing formulas.

From the way we specified our SVJD stock price model, the log-normal assumption is equivalent to assuming that the random variables Y_1 and Y_2 are normally distributed. Suppose that under the equivalent martingale measure $\mathbb{Q}$, $Y_1 \sim N(\mu_{Y_1}, \sigma^2_{Y_1})$ and $Y_2 \sim N(\mu_{Y_2}, \sigma^2_{Y_2})$.[11] The corresponding

[10] In Chapter 2, these events are denoted by $\mathcal{A}_{1,t}$ and $\mathcal{A}_{2,t}$, as defined in Proposition 2.17.

[11] As seen from the discussion in Section 3.2, the parameters μ_{Y_i} and $\sigma^2_{Y_i}$ may be derived from the parameters of the distribution of Y_i under the objective market measure $\mathbb{P}$ via Proposition 2.3. In our analysis in this section, we assume that such computations have already been performed to obtain μ_{Y_i} and $\sigma^2_{Y_i}$.

probability density functions $g_1(y)$ and $g_2(y)$ are therefore

$$\begin{aligned} g_1(y) &= \frac{1}{\sqrt{2\pi\sigma_{Y_1}^2}} \exp\left\{-\frac{1}{2}\left(\frac{y-\mu_{Y_1}}{\sigma_{Y_1}}\right)^2\right\}, \qquad y \in \mathbb{R} \\ g_2(y) &= \frac{1}{\sqrt{2\pi\sigma_{Y_2}^2}} \exp\left\{-\frac{1}{2}\left(\frac{y-\mu_{Y_2}}{\sigma_{Y_2}}\right)^2\right\}, \qquad y \in \mathbb{R} \end{aligned} \tag{3.78}$$

This also implies that

$$\tilde{\kappa}_j \equiv \mathbb{E}_{\mathbb{Q}}[e^{Y_j}] - 1 = \exp\left\{\mu_{Y_j} + \frac{1}{2}\sigma_{Y_j}^2\right\} - 1, \qquad j = 1, 2, \tag{3.79}$$

(see Section 3.2). Furthermore, let $Y_{1,1}, \ldots, Y_{1,m} \overset{i.i.d.}{\sim} N(\mu_{Y_1}, \sigma_{Y_1}^2)$ and $Y_{2,1}, \ldots, Y_{2,n} \overset{i.i.d.}{\sim} N(\mu_{Y_2}, \sigma_{Y_2}^2)$, where all random variables are mutually independent.

Given the distribution of $\{Y_{1,k}\}_{k=1}^m$ and $\{Y_{2,l}\}_{l=1}^n$, it follows that

$$\begin{aligned} \Upsilon_{1,m} &\equiv \sum_{k=1}^{m} Y_{1,k} \sim N(m\mu_{Y_1}, m\sigma_{Y_1}^2) \\ \Upsilon_{2,n} &\equiv \sum_{l=1}^{n} Y_{2,l} \sim N(n\mu_{Y_2}, n\sigma_{Y_2}^2). \end{aligned}$$

Since $\Upsilon_{1,m}$ and $\Upsilon_{2,n}$ are independent, their joint density function $g_{\boldsymbol{\Upsilon}}(x_1, x_2)$ is given by

$$\begin{aligned} g_{\boldsymbol{\Upsilon}}(x_1, x_2) = &\frac{1}{\sqrt{2\pi m\sigma_{Y_1}^2}} \exp\left\{-\frac{1}{2}\left(\frac{x_1 - m\mu_{Y_1}}{\sqrt{m}\sigma_{Y_1}}\right)^2\right\} \\ &\times \frac{1}{\sqrt{2\pi n\sigma_{Y_2}^2}} \exp\left\{-\frac{1}{2}\left(\frac{x_1 - n\mu_{Y_2}}{\sqrt{n}\sigma_{Y_2}}\right)^2\right\}, \qquad (x_1, x_2) \in \mathbb{R}^2. \end{aligned} \tag{3.80}$$

The expected value appearing in the transition density function G, the European exchange option price C^E, and the early exercise premium C^P (and its components) may now be further specified. With $g_{\boldsymbol{\Upsilon}}$, the transition density function G in equation (3.72) can now be written as

$$\begin{aligned} &G(\tau, y_1, y_2, b; S_1, S_2, v) \\ &\quad = \sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1\tau)^m e^{-\tilde{\lambda}_1\tau}}{m!} \frac{(\tilde{\lambda}_2\tau)^n e^{-\tilde{\lambda}_2\tau}}{n!} \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} \iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} \\ &\qquad \times h\left(\tau, S_1 e^{x_1 - \tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2 e^{x_2 - \tilde{\lambda}_2\tilde{\kappa}_2\tau}, v; -\phi_1, -\phi_2, b\right) g_{\boldsymbol{\Upsilon}}(x_1, x_2)\, \mathrm{d}\phi_1\, \mathrm{d}\phi_2\, \mathrm{d}x_1\, \mathrm{d}x_2, \end{aligned} \tag{3.81}$$

where $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$ are given by equation (3.79). Similarly, the price of the European exchange option

is now given by

$$\begin{aligned} C^E(\tau, S_1, S_2, v) = \sum_{m,n=0}^{\infty} & \frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau} \\ & \times \iint_{\mathbb{R}^2}\Big[S_1e^{-q_1\tau}e^{x_1-\tilde{\lambda}_1\tilde{\kappa}_1\tau}P_1\left(\tau, S_1e^{x_1-\tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2e^{x_2-\tilde{\lambda}_2\tilde{\kappa}_2\tau}, v\right) \\ & \quad - S_2e^{-q_2\tau}e^{x_2-\tilde{\lambda}_2\tilde{\kappa}_2\tau}P_2\left(\tau, S_1e^{x_1-\tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2e^{x_2-\tilde{\lambda}_2\tilde{\kappa}_2\tau}, v\right)\Big] \\ & \times g_{\boldsymbol{\Upsilon}}(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2, \end{aligned} \tag{3.82}$$

where P_1 and P_2 are defined in Proposition 3.14. Likewise, the diffusion part C_D^P of the early exercise premium may now be written as

$$\begin{aligned} & C_D^P(\tau, S_1, S_2, v) \\ & \quad = \sum_{m,n=0}^{\infty}\frac{\tilde{\lambda}_1^m}{m!}\frac{\tilde{\lambda}_2^n}{n!}\iint_{\mathbb{R}^2}\int_0^{\tau}\int_0^{\infty}(\tau-u)^{m+n}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(\tau-u)} \\ & \qquad \times\Big\{q_1S_1e^{x_1-\tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}e^{-q_1(\tau-u)} \\ & \qquad\quad \times P_1^D\left(\tau-u, S_1e^{x_1-\tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2e^{x_2-\tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; b, B(u,b)\right) \\ & \qquad - q_2S_2e^{x_2-\tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}e^{-q_2(\tau-u)} \\ & \qquad\quad \times P_2^D\left(\tau-u, S_1e^{x_2-\tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2e^{x_2-\tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; b, B(u,b)\right)\Big\} \\ & \qquad \times g_{\boldsymbol{\Upsilon}}(x_1, x_2)\,\mathrm{d}b\,\mathrm{d}u\,\mathrm{d}x_1\,\mathrm{d}x_2, \end{aligned}$$

and the jump components $C_{J_1}^P$ and $C_{J_2}^P$ may be expressed as

$$\begin{aligned} & C_{J_1}^P(\tau, S_1, S_2, v) \\ & \quad = \sum_{m,n=0}^{\infty}\frac{\tilde{\lambda}_1^m}{m!}\frac{\tilde{\lambda}_2^n}{n!}\frac{1}{4\pi^2}\iint_{\mathbb{R}^2}\int_0^{\tau}\int_0^{\infty}(\tau-u)^{m+n}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(\tau-u)}e^{-r(\tau-u)} \\ & \qquad \times\int_{-\infty}^{0}\frac{1}{\sqrt{2\pi\sigma_{Y_1}^2}}\exp\left\{-\frac{1}{2}\left(\frac{y-\mu_{Y_1}}{\sigma_{Y_1}}\right)^2\right\} \\ & \qquad \times\int_{-\infty}^{\infty}\int_{\ln B(u,b)+y_2}^{\ln B(u,b)+y_2-y}\left[C^A\left(u, e^{y_1+y}, e^{y_2}, b\right)-\left(e^{y_1+y}-e^{y_2}\right)\right] \\ & \qquad \times\iint_{\mathbb{R}^2}e^{i\phi_1y_1+i\phi_2y_2}h\left(\tau-u, S_1e^{x_1-\tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2e^{x_2-\tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; -\phi_1, -\phi_2, b\right) \\ & \qquad \times g_{\boldsymbol{\Upsilon}}(x_1, x_2)\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}y\,\mathrm{d}b\,\mathrm{d}u\,\mathrm{d}x_1\,\mathrm{d}x_2 \\ & C_{J_2}^P(\tau, S_1, S_2, v) \\ & \quad = \sum_{m,n=0}^{\infty}\frac{\tilde{\lambda}_1^m}{m!}\frac{\tilde{\lambda}_2^n}{n!}\frac{1}{4\pi^2}\iint_{\mathbb{R}^2}\int_0^{\tau}\int_0^{\infty}(\tau-u)^{m+n}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(\tau-u)}e^{-r(\tau-u)} \end{aligned}$$

$$\times \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_{Y_2}^2}} \exp\left\{-\frac{1}{2}\left(\frac{y-\mu_{Y_2}}{\sigma_{Y_2}}\right)^2\right\}$$
$$\times \int_{-\infty}^\infty \int_{\ln B(u,b)+y_2+y}^{\ln B(u,b)+y_2} \left[C^A\left(u, e^{y_1}, e^{y_2+y}, b\right) - \left(e^{y_1} - e^{y_2+y}\right)\right]$$
$$\times \iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} h\left(\tau - u, S_1 e^{x_1 - \tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2 e^{x_2 - \tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; -\phi_1, -\phi_2, b\right)$$
$$\times g_{\boldsymbol{\Upsilon}}(x_1, x_2)\, \mathrm{d}\phi_1\, \mathrm{d}\phi_2\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}y\, \mathrm{d}b\, \mathrm{d}u\, \mathrm{d}x_1\, \mathrm{d}x_2.$$

The transition density function and the European exchange option price, as given by equations (3.81) and (3.82) can then be evaluated using numerical integration. However, the components of the early exercise premium still require knowledge of the critical stock price ratio $B(\tau, v)$ before the integrals can be fully evaluated. Once the European exchange option and the components of the early exercise premium have been computed, we can then use equation (3.73) and (3.65) to solve for the early exercise premium and the American exchange option price, respectively.

3.10 Chapter Summary and Conclusion

This chapter considered the price of European and American exchange options under the assumption that the asset prices are driven by *proportional* stochastic volatility jump-diffusion (SVJD) processes. By assuming that the volatilities of the two stock prices are driven by a single instantaneous variance process, we derived an integro-partial differential equation for exchange option prices that can reduced to a first-order partial differential equation in the variance variable using Fourier and Laplace integral transforms.

We then showed the extensive calculations involved in deriving an integral representation of the price of the European exchange option. We note that the representation we obtained for the European exchange option price can be seen as a sum of Poisson-weighted Margrabe-type formulas and can be seen as consistent with the expressions obtained by Heston (1993), Bates (1996), Cheang and Chiarella (2011), and Cheang et al. (2013). Similar to the original Margrabe (1978) result, our European exchange option price representation is free of the risk-free rate r.

We then derived the early exercise premium representation for the American exchange option price by considering an inhomogeneous IPDE, to be solved over an unrestricted stock price domain, that is equivalent to the homogeneous IPDE whose stock price domain is restricted by the early exercise boundary. The inhomogeneous term was derived by adapting the techniques discussed by Jamshidian (1992) and Chiarella and Ziogas (2009). With the inhomogeneous term and the transition density function of the underlying state variables, we were able to derive an integral representation for the early exercise premium. Similar to the findings of Gukhal (2001), Chiarella and Ziogas (2009), and Cheang and Chiarella (2011), the early exercise premium can be further

decomposed into a diffusion premium and two rebalancing cost terms that arise from the possibility of jumps in asset prices. Following this analysis, we characterize the American exchange option price and the unknown critical stock price ratio (associated to the early exercise boundary) as the solution of a linked system of integral equations. Lastly, we derived representations of option prices and the early exercise premium that will result from assuming that the jump sizes have a lognormal distribution, following the assumption of Merton (1973). The computations presented in this chapter can be extended to obtain integral representation of prices of options whose payoff depends on the underlying asset prices but not the variance level.

The representations that we obtained in this chapter express the exchange option prices as multidimensional integrals which can be evaluated numerically using standard computing software, although the efficiency of doing so has yet to be considered. However in this thesis, we do not provide details of the numerical integration of the integral representations for the option prices and instead we considered numerical methods to solve the option pricing IPDEs (see Chapter 5). We leave the implementation of a numerical integration scheme to future research, with the implementation of Chiarella and Ziogas (2009), Adolfsson et al. (2013), and Chiarella and Ziveyi (2013) as starting points.

CHAPTER 4

A Put-Call Transformation of the Exchange Option Problem

4.1 Introduction

In this chapter, we consider an alternative strategy to deriving option pricing formulas for the European and American exchange options under the assumption that underlying asset prices are driven by the proportional stochastic volatility and jump-diffusion model considered in Chapter 3.

In the original analysis of European exchange options, Margrabe (1978) transformed the problem to the classical European call option pricing problem, assuming that the exchange option price is linear homogeneous in the underlying asset prices, and computed the exchange option price using the solution of Black and Scholes (1973) (see also (Jamshidian, 2010; Poulsen, 2010)). Bjerskund and Stensland (1993) considered the pricing of American exchange options as an optimal stopping problem in a pure diffusion setting. They suggested that by choosing one of the stocks as the numéraire and by a change of measure to the corresponding equivalent martingale measure, the American exchange option pricing problem may be simplified to the problem of pricing an American call or put option. Bjerskund and Stensland (1993) refer to this technique as the "put-call transformation."

In this chapter, we derive integral representations of the price of European and American exchange options following the suggested approach of Bjerskund and Stensland (1993) and assigning the yield process of the second stock as the numéraire asset. This is in contrast to the usual approach where the money market account is used as the numéraire. Doing so simplifies the original two-dimensional problem to a one-dimensional problem of pricing an ordinary call option written on the ratio of the yield processes of the two underlying assets.[1] This allows us to follow the techniques

[1] While the title of this chapter mentions a "put-call transformation", our main analysis consists of transforming the exchange option pricing problem to the problem of pricing a call option on the asset yield ratio by choosing the second asset yield process as the numéraire. The exchange option problem can be reduced to a put option

of Cheang et al. (2013) who analyzed single-asset American call options under the Bates (1996) stochastic volatility and jump-diffusion dynamics.

Using the put-call transformation approach to price European and American exchange options when both stochastic volatility and jumps are present in underlying asset price dynamics has received little to no attention in the literature. For the separate case where jumps and stochastic volatility are present, Cheang et al. (2006) and Alòs and Rheinlander (2017), respectively, priced exchange options using the put-call transformation approach. Fajardo and Mordecki (2006) also used a similar transformation, which they called the "dual market method", to price options (including perpetual exchange options) when underlying asset prices are driven by Lévy processes. Others (Antonelli and Scarlatti, 2010; Cheang and Chiarella, 2011; Caldana et al., 2015; Kim and Park, 2017; Cufaro-Petroni and Sabino, 2020; Ma et al., 2020; Parischa and Goel, 2021) priced exchange options under the equivalent martingale measure (EMM) corresponding to the money market account when either stochastic volatility or jumps are present (but not both at the same time).

The main contribution of this chapter is the extension of the Bjerskund and Stensland (1993) strategy for valuing American exchange options in a pure-diffusion setting into the SJVD framework. In contrast to previous work that use the technique to simplify the European exchange option price (Alòs and Rheinlander, 2017; Antonelli and Scarlatti, 2010; Cheang et al., 2006; Margrabe, 1978; Parischa and Goel, 2021), we place significant added attention to pricing American exchange options using the put-call transformation. Our analysis also extends the study of Cheang et al. (2006) and Cheang and Chiarella (2011) to consider American exchange options and European exchange options when assets are driven by a stochastic volatility *and* jump-diffusion model. In this analysis, pricing takes place under the equivalent martingale measure $\hat{\mathbb{Q}}$ corresponding to setting the second asset yield process as the numéraire. Under $\hat{\mathbb{Q}}$, we find that the the no-arbitrage price of the European exchange option can be written as a function of only the asset yield ratio $\tilde{s}$ and the instantaneous variance v. Furthermore, we verify an early exercise representation of the discounted American exchange option price, which can also be written as a function of only $\tilde{s}$ and v.

To evaluate expectations under $\hat{\mathbb{Q}}$, we require the joint transition density function of $\tilde{s}$ and v. To do so, we determine the Kolmogorov equation for the transition density and solve this using Fourier and Laplace transforms.[2] With the joint transition density function, we then provide integral representations for the European exchange option price and the early exercise premium and present a linked system of integral equations that characterize the price of the American exchange

on the asset yield ratio if one alternatively chooses the yield process of the first asset as the numéraire.

[2] This method has been used by Chiarella and Ziogas (2009) to price American call options under jump-diffusion dynamics, by Chiarella et al. (2010) to price American call options under stochastic volatility, by Cheang et al. (2013) to price American call options under SVJD dynamics, and by Chiarella and Ziveyi (2014) to price American spread call options under pure diffusion dynamics.

option and the unknown early exercise boundary. We also note that the joint transition density function we obtain in this analysis enables us to price any European-type option written on the two assets provided that its payoff function can be written in terms of the terminal asset yield ratio and instantaneous variance.

Our analysis also serves as an alternative of that in Chapters 2 and 3 in the following aspects.

1. We take one of the asset yield processes as the numéraire instead of the money market account used in Chapters 2 and 3.

2. We express option prices in terms of the transition density function of the underlying stochastic processes under $\hat{\mathbb{Q}}$.

3. We provide a more in-depth analysis of the early exercise boundary (particularly its behavior near maturity) and the early exercise premium for the American exchange option, and thus extending the earlier work of Chiarella and Ziogas (2009) who investigated the limit of the early exercise boundary for American call options under jump-diffusion dynamics.

We are primarily concerned with obtaining analytical representations for exchange option prices under our SVJD model. We shall discuss the numerical implementation of our results here in the next chapter.

The rest of the chapter is organized as follows. Section 4.2 specifies the stochastic volatility and jump-diffusion model for underlying asset prices, discusses the construction of the measure $\hat{\mathbb{Q}}$, and presents the dynamics of the asset yield ratio and variance processes under $\hat{\mathbb{Q}}$. Section 4.3 discusses the integro-partial differential equation (IPDE) for the discounted European exchange option price and the free-boundary IPDE for the American exchange option. Section 4.4 uses probabilistic arguments to verify the early exercise representation of the American exchange option price and to express the American exchange option price as a solution of an inhomogeneous IPDE to be solved over a domain unrestricted by the early exercise boundary. Section 4.5 discusses some properties of the early exercise boundary near the maturity of the option. Section 4.6 shows the solution of the Kolmogorov equation for the joint transition density function using integral transform methods. With the transition density function, Sections 4.7 and 4.8 present the integral representations for the European and American exchange option prices, respectively. Section 4.9 concludes the chapter. To streamline the presentation of this chapter, the proof of some of the results in this chpater are placed in Appendix C.

4.2 The Proportional SVJD Model and a Change of Measure

In this analysis, we will use the proportional stochastic volatility and jump-diffusion model for the underlying asset price dynamics. This model has been specified in Chapter 3, but for convenience

we shall restate the important details here. We also discuss here the construction of the equivalent martingale measure corresponding to the choice of the second asset yield process as the numéraire.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the historical or market probability space. Consider two risky assets S_1 and S_2 which pay continuously-compounded dividend yields q_1 and q_2 and whose prices satisfy the stochastic differential equations

$$\begin{aligned} \frac{\mathrm{d}S_j(t)}{S_j(t)} &= (\mu_j - \lambda_j \kappa_j)\,\mathrm{d}t + \sigma_j \sqrt{v(t)}\,\mathrm{d}W_j(t) + \int_{\mathbb{R}} (e^y - 1) p_j(\mathrm{d}y, \mathrm{d}t), \qquad j = 1, 2 \\ \mathrm{d}v(t) &= \xi(\eta - v(t))\,\mathrm{d}t + \omega \sqrt{v(t)}\,\mathrm{d}Z(t). \end{aligned} \tag{4.1}$$

Here, we assume that μ_j, σ_j $(j = 1, 2)$, ξ, η, and ω are positive constants, $\mathbf{B}(t) = (W_1(t), W_2(t), Z(t))^\top$ is a vector of $\mathbb{P}$-Brownian motions with correlation matrix

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \rho_w & \rho_{wz_1} \\ \rho_w & 1 & \rho_{wz_2} \\ \rho_{wz_1} & \rho_{wz_2} & 1 \end{bmatrix},$$

$p_j(\mathrm{d}y, \mathrm{d}t)$ is a non-explosive counting measure with $\mathbb{P}$-local characteristics $(\lambda_j, m_{\mathbb{P}}^{(j)}(\mathrm{d}y))$ $(j = 1, 2)$, and $\kappa_j = \mathbb{E}_{\mathbb{P}}[e^{Y_j} - 1] = \int_{\mathbb{R}} (e^y - 1) m_{\mathbb{P}}^{(j)}(\mathrm{d}y)$. For notational convenience, it is understood that $S_j(t)$ denotes the *pre-jump* price of asset j at time t, if a jump should occur at that time. We also assume that $S_1(0)$, $S_2(0)$, and $v(0)$ are known positive values. As in Assumption 3.1, we assume that the coefficients of the variance process satisfy the Feller condition, i.e. $2\xi\eta \geq \omega^2$, and the asset-variance correlations satisfy $-1 < \rho_{wz_j} < \min\{\xi/\omega, 1\}$, $j = 1, 2$.

We assume that the counting measures and the jump size random variables are mutually independent and are independent of the Brownian motions. As before, we assume that $\{\mathcal{F}_t\}$ is the natural filtration generated by $\mathbf{B}(t)$, $p_1(\mathrm{d}y, \mathrm{d}t)$, and $p_2(\mathrm{d}y, \mathrm{d}t)$, and that the filtration satisfies the usual conditions.

We also assume the existence of a money market account whose value process is denoted by $\{B(t)\}$. We assume that it deterministically grows according to the equation $\mathrm{d}B(t) = rB(t)\,\mathrm{d}t$, with $B(0) = 1$, where $r > 0$ is the (constant) risk-free interest rate. In other words, $B(t) = e^{rt}$ for $t > 0$.

As before, we assume that the dividend yields, the risk-free rate, the parameters μ_j, σ_j, ξ, η, and ω, the jump the intensities λ_j and jump-size densities $m_{\mathbb{P}}^{(j)}(\mathrm{d}y)$ are constant through time, but the analysis can be extended to the case where these parameters are deterministic functions of time.

In the previous chapters, we priced exchange options as expected values under a probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, which was constructed such that the discounted asset yield processes $\{S_1(t)e^{q_1 t}/B(t)\}$ and $\{S_2(t)e^{q_2 t}/B(t)\}$ are martingales. Here, the money market account is taken to

be the numéraire asset and the discounted asset yield processes represent the asset yields measured in units of the money market account. The probability measure $\mathbb{Q}$, an equivalent martingale measure, constructed in this manner yields the well-known risk-neutral probability measure.

In this chapter, we take $\{S_2(t)e^{q_2 t}\}$, the second asset yield process, as the numéraire and define the probability measure $\hat{\mathbb{Q}}$, equivalent to $\mathbb{P}$, such that the first asset yield process and the money market account, when discounted by $S_2(t)e^{q_2 t}$ are martingales under $\hat{\mathbb{Q}}$. In doing so, we show that the exchange option pricing problem, which is a multi-dimensional problem under the risk-neutral measure, can be reduced to the problem of pricing a call option under $\hat{\mathbb{Q}}$.

With the second asset yield process as the numéraire, the discounted price of any other asset with price process $\{\mathcal{X}(t)\}$ is defined by

$$\tilde{\mathcal{X}}(t) = \frac{\mathcal{X}(t)}{S_2(t)e^{q_2 t}}. \tag{4.2}$$

We will sometimes refer to the quantity $[S_2(t)e^{q_2 t}]^{-1}$ as the discounting factor for the period $[0,t]$ and denote it by $\mathsf{DF}(0,t)$. In the absence of arbitrage opportunities, the appropriate discounting factor for the period $[t,t']$ for any $t' \in [t,T]$ is given by

$$\mathsf{DF}(t,t') = \frac{\mathsf{DF}(0,t')}{\mathsf{DF}(0,t)} = \frac{S_2(t)e^{q_2 t}}{S_2(t')e^{q_2 t'}}. \tag{4.3}$$

Before we conclude this section, we establish the suitability of the process $\{S_2(t)e^{q_2 t}\}$ to be considered a numéraire; that is, we need to show that it is almost surely strictly positive for all $t \in [0,T]$ (Geman et al., 1995). Indeed, the solution to equation (3.3) (for $j=2$) is given by

$$S_2(t) = S_2(0)\exp\left\{(\mu_2 - \lambda_2\kappa_2)\,t - \frac{1}{2}\sigma_2^2\int_0^t v(s)\,\mathrm{d}s + \sigma_2\int_0^t \sqrt{v(s)}\,\mathrm{d}W_2(s) + \sum_{n=1}^{N_2(t)} Y_{2,n}\right\} \tag{4.4}$$

for all $0 < t \leq T$ (see Section 3.2). Here, $N_2(t)$ is the Poisson counting process corresponding to $p(\mathrm{d}y_2, \mathrm{d}t)$ and $\{Y_{2,1}, Y_{2,2}, \dots\}$ is an i.i.d. sequence from the jump size density $m_{\mathbb{P}}^{(2)}(\mathrm{d}y)$. Given Assumption 3.1, the variance process $\{v(t)\}_{0\leq t\leq T}$ is $\mathbb{P}$-a.s. positive and finite (Andersen and Piterbarg, 2007) and so $\int_0^t v(s)\,\mathrm{d}s < \infty$ $\mathbb{P}$-a.s. This also implies that the stochastic integral $\int_0^t \sqrt{v(s)}\,\mathrm{d}W_2(s)$ is properly defined (Kuo, 2006, Chapter 5). Since the point process underlying $N_2(t)$ is non-explosive, then $N_2(t) < \infty$ $\mathbb{P}$-a.s. for all $t \in [0,T]$ (Brémaud, 1981). This observation, coupled with the fact that the $Y_{2,n}$'s are real-valued random variables, implies that the sum $\sum_{n=1}^{N_2(t)} Y_{2,n}$ is finite $\mathbb{P}$-a.s. Therefore, $S_2(t) > 0$ and, consequently, $S_2(t)e^{q_2 t} > 0$ $\mathbb{P}$-a.s. for all $t \in [0,T]$.

We now discuss the construction of the equivalent probability measure $\hat{\mathbb{Q}}$ under which the yield process of stock 1 and the money market account, when measured in units of the second stock's yield process, are martingales. The basis of our construction of $\hat{\mathbb{Q}}$ is Proposition 3.2 which states the

construction of probability measures $\mathbb{Q}$ equivalent to $\mathbb{P}$. Note that the Radon-Nikodým derivative $L(T) = \mathrm{d}\hat{\mathbb{Q}}/\,\mathrm{d}\mathbb{P}$ provided in Proposition 3.2 can be used to characterize any probability measure $\hat{\mathbb{Q}}$ equivalent to $\mathbb{P}$ as parameterized by the vector process $\{\boldsymbol{\theta}(t)\}$ and the constants $\gamma_1, \gamma_2, \nu_1, \nu_2$. In the succeeding analysis, these parameters will be specified so that $\hat{\mathbb{Q}}$ becomes an equivalent martingale measure corresponding to the numéraire $\{S_2(t)e^{q_2(t)}\}$.

Consider the processes $\{\tilde{S}(t)\}$ and $\{\tilde{B}(t)\}$ where

$$\tilde{S}(t) = \frac{S_1(t)e^{q_1 t}}{S_2(t)e^{q_2 t}} \qquad \text{and} \qquad \tilde{B}(t) = \frac{e^{rt}}{S_2(t)e^{q_2 t}}. \tag{4.5}$$

These two processes are, respectively, the first stock's yield process and the money market account discounted using the second stock's yield process. In particular, we will refer to $\{\tilde{S}(t)\}$ as the asset yield ratio process. *The equivalent martingale measure $\hat{\mathbb{Q}}$ corresponding to this choice of numéraire is thus determined by choosing the parameters $\{\boldsymbol{\theta}(t)\}$ and $\gamma_1, \gamma_2, \nu_1, \nu_2$ such that $\{\tilde{S}(t)\}$ and $\{\tilde{B}(t)\}$ are $\hat{\mathbb{Q}}$-martingales.* The following proposition presents some restrictions on the choice of parameters of the Radon-Nikodým derivative in constructing the equivalent martingale measure $\hat{\mathbb{Q}}$.

Proposition 4.1. *Let $\hat{\mathbb{Q}}$ be a probability measure equivalent to $\mathbb{P}$ such that $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T} = L(T)$, where $L(t)$ is given by equation (3.5) and is parameterized by $\mathcal{F}_t$-adapted processes $\{\psi_1(t)\}$, $\{\psi_2(t)\}$, and $\{\zeta(t)\}$ and constants $\gamma_1, \gamma_2, \nu_1, \nu_2$. Suppose $\{\psi_1(t)\}$, $\{\psi_2(t)\}$, and $\{\zeta(t)\}$ are defined as*

$$\psi_1(t) = \frac{\mu_1 + q_1 - r - \rho_w\sigma_1\sigma_1 v(t) - \lambda_1\kappa_1 + \tilde{\lambda}_1\tilde{\kappa}_1}{\sigma_1\sqrt{v(t)}}, \tag{4.6}$$

$$\psi_2(t) = \frac{\mu_2 + q_2 - r - \sigma_2^2 v(t) - \lambda_2\kappa_2 - \tilde{\lambda}_2\tilde{\kappa}_2^-}{\sigma_2\sqrt{v(t)}}, \tag{4.7}$$

$$\zeta(t) = \frac{\Lambda}{\omega}\sqrt{v(t)} \qquad \textit{for some } \Lambda \geq 0, \tag{4.8}$$

respectively, where $\tilde{\kappa}_1 = \mathbb{E}_{\hat{\mathbb{Q}}}[e^{Y_1} - 1]$ and $\tilde{\kappa}_2^- = \mathbb{E}_{\hat{\mathbb{Q}}}[e^{-Y_2} - 1]$. Suppose further that γ_1, γ_2, ν_1, and ν_2 are determined such that $\tilde{\lambda}_j$ and $M_{\hat{\mathbb{Q}},Y_j}$, $j = 1, 2$, are well defined (see equations (3.6) and (3.7)). Then $\hat{\mathbb{Q}}$ is an equivalent martingale measure under which the processes $\{\tilde{S}(t)\}$ and $\{\tilde{B}(t)\}$ (given by equation (4.5)) are martingales on $[0, T]$.

Proof. We start with $\tilde{B}(t)$. By Itô's Lemma for jump-diffusion processes, we find that $\tilde{B}(t)$ satisfies the stochastic differential equation

$$\frac{\mathrm{d}\tilde{B}(t)}{\tilde{B}(t)} = \Big(r - q_2 - \mu_2 + \lambda_2\kappa_2 + \sigma_2^2 v(t)\Big)\,\mathrm{d}t - \sigma_2\sqrt{v(t)}\,\mathrm{d}W_2(t) + \int_{\mathbb{R}}\Big(e^{-y} - 1\Big)\,p_2(\mathrm{d}y, \mathrm{d}t).$$

By Proposition 3.2, there exists a standard $\hat{\mathbb{Q}}$-Brownian motion $\{\bar{W}_2(t)\}$ such that

$$\mathrm{d}W_2(t) = -\psi_2(t)\,\mathrm{d}t + \mathrm{d}\bar{W}_2(t).$$

Furthermore, we can replace $p_2(\mathrm{d}y, \mathrm{d}t)$ using the expression

$$p_2(\mathrm{d}y, \mathrm{d}t) = q_2(\mathrm{d}y, \mathrm{d}t) + \tilde{\lambda}_2 m_{\hat{\mathbb{Q}}}^{(2)}(\mathrm{d}y)\,\mathrm{d}t,$$

where $q_2(\mathrm{d}y, \mathrm{d}t)$ is the $\hat{\mathbb{Q}}$-compensated counting measure corresponding to $p_2(\mathrm{d}y, \mathrm{d}t)$. With these, we can write $\mathrm{d}\tilde{B}(t)$ as

$$\begin{aligned}\mathrm{d}\tilde{B}(t) = \tilde{B}(t)\left[r - \mu_2 - q_2 + \lambda_2\kappa_2 + \sigma_2^2 v(t) + \sigma_2\sqrt{v(t)}\psi_2(t) + \tilde{\lambda}_2\tilde{\kappa}_2^-\right]\mathrm{d}t \\ - \sigma_2\sqrt{v(t)}\tilde{B}(t)\,\mathrm{d}\bar{W}_2(t) + \int_{\mathbb{R}}\left(e^{-y} - 1\right)\tilde{B}(t)q_2(\mathrm{d}y, \mathrm{d}t),\end{aligned}$$

where $\tilde{\kappa}_2^-$ is defined in the proposition. The process $\psi_2(t)$ must be defined such that $\tilde{B}(t)$ is a $\hat{\mathbb{Q}}$-martingale. In line with the martingale representation result for jump-diffusion processes (see Runggaldier (2003, Theorem 2.2) or Øksendal and Sulem (2019, Theorem 1.35)), $\psi_2(t)$ must be defined as equation (4.7) so that the drift term of $\mathrm{d}\tilde{B}(t)$ vanishes. With this choice of $\psi_2(t)$, $\tilde{B}(t)$ is a martingale under $\hat{\mathbb{Q}}$.

We next determine $\psi_1(t)$ such that $\tilde{S}(t)$ is also a martingale under $\hat{\mathbb{Q}}$. By the Itô formula for jump-diffusion processes, $\tilde{S}(t)$ satisfies the SDE

$$\begin{aligned}\frac{\mathrm{d}\tilde{S}(t)}{\tilde{S}(t)} = \left(q_1 - q_2 + \mu_1 - \mu_2 - \lambda_1\kappa_1 + \lambda_2\kappa_2 + \sigma_2^2 v(t) - \rho_1\sigma_1\sigma_2 v(t)\right)\mathrm{d}t \\ + \sigma_1\sqrt{v(t)}\,\mathrm{d}W_1(t) - \sigma_2\sqrt{v(t)}\,\mathrm{d}W_2(t) + \int_{\mathbb{R}}\left(e^{y} - 1\right)p_1(\mathrm{d}y, \mathrm{d}t) \\ + \int_{\mathbb{R}}\left(e^{-y} - 1\right)p_2(\mathrm{d}y, \mathrm{d}t).\end{aligned}$$

Using Proposition 3.2, we may make the substitutions

$$\mathrm{d}W_j(t) = -\psi_j(t)\,\mathrm{d}t + \mathrm{d}\bar{W}_j(t),$$

where $\{\bar{W}_j(t)\}$ is a standard $\hat{\mathbb{Q}}$-Brownian motion, and

$$p_j(\mathrm{d}y, \mathrm{d}t) = q_j(\mathrm{d}y, \mathrm{d}t) + \tilde{\lambda}_j m_{\hat{\mathbb{Q}}}^{(j)}(\mathrm{d}y)\,\mathrm{d}t,$$

for $j = 1, 2$. We note that $\psi_2(t)$ is already given by equation (4.7). We can thus write $\mathrm{d}\tilde{S}(t)$ as

$$\begin{aligned}\mathrm{d}\tilde{S}(t) = \tilde{S}(t)\left[\mu_1 + q_1 - r - \rho_w\sigma_1\sigma_2 v(t) - \lambda_1\kappa_1 + \tilde{\lambda}_1\tilde{\kappa}_1 - \sigma_1\sqrt{v(t)}\psi_1(t)\right]\mathrm{d}t \\ + \sigma_1\sqrt{v(t)}\tilde{S}(t)\,\mathrm{d}\bar{W}_1(t) - \sigma_2\sqrt{v(t)}\tilde{S}(t)\,\mathrm{d}\bar{W}_2(t) \\ + \int_{\mathbb{R}}\left(e^{y} - 1\right)\tilde{S}(t)q_1(\mathrm{d}y, \mathrm{d}t) + \int_{\mathbb{R}}\left(e^{-y} - 1\right)\tilde{S}(t)q_2(\mathrm{d}y, \mathrm{d}t),\end{aligned}$$

where $\tilde{\kappa}_1 \equiv \mathbb{E}_{\hat{\mathbb{Q}}}[e^{Y_1} - 1] = \int_{\mathbb{R}}(e^y - 1)m^{(1)}_{\hat{\mathbb{Q}}}(\mathrm{d}y)$. The drift term vanishes if we choose $\psi_1(t)$ as in equation (4.6). With these choices for $\psi_1(t)$ and $\psi_2(t)$, $\{\tilde{S}(t)\}$ is a $\hat{\mathbb{Q}}$-martingale.

The dynamics of the variance process will also be modified as a result of the change in measure from $\mathbb{P}$ to $\hat{\mathbb{Q}}$. By Proposition 3.2, there exists a standard $\hat{\mathbb{Q}}$-Brownian motion $\{\bar{Z}(t)\}$ such that $\mathrm{d}Z(t) = -\zeta(t)\,\mathrm{d}t + \mathrm{d}\bar{Z}(t)$, and so $\mathrm{d}v(t)$ may be written as

$$\mathrm{d}v(t) = \left[\xi(\eta - v(t)) - \zeta(t)\omega\sqrt{v(t)}\right]\mathrm{d}t + \omega\sqrt{v(t)}\,\mathrm{d}\bar{Z}(t).$$

Note that unlike $\psi_1(t)$ and $\psi_2(t)$, $\zeta(t)$ is not directly involved in ensuring that $\tilde{B}(t)$ and $\tilde{S}(t)$ are martingales. However, it has to be defined such that $\{v(t)\}$ is positive and finite $\hat{\mathbb{Q}}$-almost surely on $[0, T]$, which in turn ensures that $\psi_1(t)$ and $\psi_2(t)$ are properly defined.

Following Heston (1993), we assume that the market price of volatility risk, given here by $\zeta(t)\omega\sqrt{v(t)}$ is proportional to the current variance $v(t)$; that is, we assume that there exists a nonnegative constant Λ such that

$$\Lambda v(t) = \zeta(t)\omega\sqrt{v(t)}. \tag{4.9}$$

Therefore, we have

$$\mathrm{d}v(t) = \left[\xi\eta - (\xi + \Lambda)v(t)\right]\mathrm{d}t + \omega\sqrt{v(t)}\,\mathrm{d}\bar{Z}(t). \tag{4.10}$$

From this equation, we find that the new rate of mean reversion and the new long-run mean of the process under $\hat{\mathbb{Q}}$ are given by $\xi_{\hat{\mathbb{Q}}} = \xi + \Lambda$ and $\eta_{\hat{\mathbb{Q}}} = \xi\eta/(\xi + \Lambda)$, respectively. The variance process $\{v(t)\}$ does not go to zero or explode in finite time under $\hat{\mathbb{Q}}$ provided that $2\xi_{\hat{\mathbb{Q}}}\eta_{\hat{\mathbb{Q}}} \geq \omega^2$ (Andersen and Piterbarg, 2007). Indeed, given Assumption 3.1, this condition is satisfied since

$$2\xi_{\hat{\mathbb{Q}}}\eta_{\hat{\mathbb{Q}}} = 2(\xi + \Lambda)\cdot\frac{\xi\eta}{\xi + \Lambda} = 2\xi\eta \geq \omega^2.$$

Note that no additional restrictions are needed for the constant parameters γ_1,γ_2,ν_1, and ν_2 as long as they are chosen such that the $\hat{\mathbb{Q}}$ intensities $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ and the moment generating functions $M_{\hat{\mathbb{Q}},Y_1}$ and $M_{\hat{\mathbb{Q}},Y_2}$ of the jump-size r.v.'s under $\mathbb{Q}$ are well defined (see equations (3.6) and (3.7)). ■

From the risk premia given by $\psi_1(t)$ and $\psi_2(t)$, we can find the market prices of jump risk for stocks 1 and 2 by considering the terms $\lambda_1\kappa_1 - \tilde{\lambda}_1\tilde{\kappa}_1$ and $\lambda_2\kappa_2 + \tilde{\lambda}_2\tilde{\kappa}_2^-$. From Proposition 3.2 and the definitions of κ_1 and $\tilde{\kappa}_1$, we note that

$$\begin{aligned}
\lambda_1\kappa_1 - \tilde{\lambda}_1\tilde{\kappa}_1 &= \lambda_1\left(M_{\mathbb{P},Y_1}(1) - 1\right) - \lambda_1 e^{\nu_1}\mathbb{E}_{\mathbb{P}}\left[e^{\gamma_1 Y_1}\right]\mathbb{E}_{\hat{\mathbb{Q}}}\left[e^{Y_1} - 1\right] \\
&= \lambda_1\left(M_{\mathbb{P},Y_1}(1) - 1\right) - \lambda_1 e^{\nu_1}M_{\mathbb{P},Y_1}(\gamma_1)\left(M_{\hat{\mathbb{Q}},Y_1}(1) - 1\right) \\
&= \lambda_1\left(M_{\mathbb{P},Y_1}(1) - 1\right) - \lambda_1 e^{\nu_1}M_{\mathbb{P},Y_1}(\gamma_1)\left[\frac{M_{\mathbb{P},Y_1}(\gamma_1 + 1)}{M_{\mathbb{P},Y_1}(\gamma_1)} - 1\right]
\end{aligned}$$

$$= \lambda_1 \left(M_{\mathbb{P},Y_1}(1) - 1\right) \left[1 - e^{\nu_1} \frac{M_{\mathbb{P},Y_1}(\gamma_1 + 1) - M_{\mathbb{P},Y_1}(\gamma_1)}{M_{\mathbb{P},Y_1}(1) - 1}\right]$$
$$= \lambda_1 \kappa_1 \left[1 - e^{\nu_1} \frac{M_{\mathbb{P},Y_1}(\gamma_1 + 1) - M_{\mathbb{P},Y_1}(\gamma_1)}{M_{\mathbb{P},Y_1}(1) - 1}\right].$$

Following Cheang et al. (2006), we may interpret the quantity

$$1 - e^{\nu_1} \frac{M_{\mathbb{P},Y_1}(\gamma_1 + 1) - M_{\mathbb{P},Y_1}(\gamma_1)}{M_{\mathbb{P},Y_1}(1) - 1}$$

as the market price of jump risk for the first stock. Likewise, we find that

$$\lambda_2 \kappa_2 + \tilde{\lambda}_2 \tilde{\kappa}_2^- = \lambda_2 \left(M_{\mathbb{P},Y_2}(1) - 1\right) + \lambda_2 e^{\nu_2} M_{\mathbb{P},Y_2}(\gamma_2) \left(M_{\hat{\mathbb{Q}},Y_2}(-1) - 1\right)$$
$$= \lambda_2 \kappa_2 \left[1 - e^{\nu_2} \frac{M_{\mathbb{P},Y_2}(\gamma_2) - M_{\mathbb{P},Y_2}(\gamma_2 - 1)}{M_{\mathbb{P},Y_2}(1) - 1}\right],$$

and here we may also interpret the expression

$$1 - e^{\nu_2} \frac{M_{\mathbb{P},Y_2}(\gamma_2) - M_{\mathbb{P},Y_2}(\gamma_2 - 1)}{M_{\mathbb{P},Y_2}(1) - 1}$$

as the market price of jump risk for the second stock.

Note that the parameters γ_1, γ_2, ν_1, and ν_2 may be selected depending on the properties we elect to preserve in the transition from $\mathbb{P}$ to $\hat{\mathbb{Q}}$. For example, the trivial choice of setting all these parameters to zero entails the preservation of the intensity rate of the Poisson processes and the jump size distribution. This also implies that the jump risk in both assets are not priced (as in the analysis of Merton (1976)). However, we at least restrict these parameters to be constants to preserve the time-homogeneity of the intensity rate and the jump size distributions. If the market prices of jump risk are calibrated from market data and have been specified by Π_1 and Π_2, then the risk premia processes $\psi_1(t)$ and $\psi_2(t)$ will be given by

$$\psi_1(t) = \frac{\mu_1 + q_1 - r - \rho_w \sigma_1 \sigma_1 v(t) - \lambda_1 \kappa_1 \Pi_1}{\sigma_1 \sqrt{v(t)}},$$
$$\psi_2(t) = \frac{\mu_2 + q_2 - r - \sigma_2^2 v(t) - \lambda_2 \kappa_2 \Pi_2}{\sigma_2 \sqrt{v(t)}}.$$

The parameters γ_1, γ_2, ν_1, and ν_2 that specify the change in intensity and jump size distribution may then be solved from Π_1 and Π_2 via the equations

$$\Pi_1 = 1 - e^{\nu_1} \frac{M_{\mathbb{P},Y_1}(\gamma_1 + 1) - M_{\mathbb{P},Y_1}(\gamma_1)}{M_{\mathbb{P},Y_1}(1) - 1}, \tag{4.11}$$
$$\Pi_2 = 1 - e^{\nu_2} \frac{M_{\mathbb{P},Y_2}(\gamma_2) - M_{\mathbb{P},Y_2}(\gamma_2 - 1)}{M_{\mathbb{P},Y_2}(1) - 1}. \tag{4.12}$$

Before we address the option pricing problem, we first rewrite the stochastic differential equation for $\tilde{S}(t)$ in terms of a single standard $\hat{\mathbb{Q}}$-Brownian motion and the random counting measures $p(\mathrm{d}y_1, \mathrm{d}t)$ and $p(\mathrm{d}y_2, \mathrm{d}t)$. Using the definition of the $\hat{\mathbb{Q}}$-compensated counting measures, we find that

$$\begin{aligned}\mathrm{d}\tilde{S}(t) = -\tilde{S}(t)\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)\mathrm{d}t + \sqrt{v(t)}\tilde{S}(t)\left(\sigma_1\,\mathrm{d}\bar{W}_1(t) - \sigma_2\,\mathrm{d}\bar{W}_2(t)\right)\\ + \int_{\mathbb{R}}\left(e^y - 1\right)\tilde{S}(t)p_1(\mathrm{d}y, \mathrm{d}t) + \int_{\mathbb{R}}\left(e^{-y} - 1\right)\tilde{S}(t)p_2(\mathrm{d}y, \mathrm{d}t).\end{aligned}$$

Consider the $\hat{\mathbb{Q}}$-Brownian motion $\{\bar{W}(t)\}$ and define $\sigma\,\mathrm{d}\bar{W}(t) \equiv \sigma_1\,\mathrm{d}\bar{W}_1(t) - \sigma_2\,\mathrm{d}\bar{W}_2(t)$, where[3] $\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_w\sigma_1\sigma_2$. It follows therefore that $\tilde{S}(t)$ satisfies

$$\begin{aligned}\mathrm{d}\tilde{S}(t) = -\tilde{S}(t)\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)\mathrm{d}t + \sigma\sqrt{v(t)}\tilde{S}(t)\,\mathrm{d}\bar{W}(t)\\ + \int_{\mathbb{R}}\left(e^y - 1\right)\tilde{S}(t)p_1(\mathrm{d}y, \mathrm{d}t) + \int_{\mathbb{R}}\left(e^{-y} - 1\right)\tilde{S}(t)p_2(\mathrm{d}y, \mathrm{d}t).\end{aligned} \tag{4.13}$$

If $\tilde{X}(t) = \ln\tilde{S}(t)$ represents the log-asset yield ratio at time t, then by the Itô formula, we find that $\tilde{X}(t)$ satisfies the stochastic differential equation

$$\begin{aligned}\mathrm{d}\tilde{X}(t) = -\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \frac{1}{2}\sigma^2 v(t)\right)\mathrm{d}t + \sigma\sqrt{v(t)}\,\mathrm{d}\bar{W}(t)\\ + \int_{\mathbb{R}} y p_1(\mathrm{d}y, \mathrm{d}t) - \int_{\mathbb{R}} y p_2(\mathrm{d}y, \mathrm{d}t).\end{aligned}$$

Integrating over $[0, t]$ and taking the exponential of both sides yield a solution $\tilde{S}(t)$ of equation (4.13) given by

$$\begin{aligned}\tilde{S}(t) = \tilde{S}(0)\exp\left\{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)t - \frac{1}{2}\sigma^2\int_0^t v(s)\,\mathrm{d}s + \sigma\int_0^t\sqrt{v(s)}\,\mathrm{d}\bar{W}(s)\right.\\ \left. + \sum_{m=1}^{N_1(t)} Y_{1,m} - \sum_{n=1}^{N_2(t)} Y_{2,n}\right\}\end{aligned}$$

In our succeeding calculations, we will need the instantaneous correlation between the $\hat{\mathbb{Q}}$-Brownian motions $\{\bar{W}(t)\}$ and $\{\bar{Z}(t)\}$. From the above analysis, we note that

$$\begin{aligned}\mathbb{E}_{\hat{\mathbb{Q}}}\left[\sigma\,\mathrm{d}\bar{W}(t)\,\mathrm{d}\bar{Z}(t)\right] &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\left(\sigma_1\,\mathrm{d}\bar{W}_1(t) - \sigma_2\,\mathrm{d}\bar{W}_2(t)\right)\mathrm{d}\bar{Z}(t)\right]\\ &= \sigma_1\mathbb{E}_{\hat{\mathbb{Q}}}\left[\mathrm{d}\bar{W}_1(t)\,\mathrm{d}\bar{Z}(t)\right] - \sigma_2\mathbb{E}_{\hat{\mathbb{Q}}}\left[\mathrm{d}\bar{W}_2(t)\,\mathrm{d}\bar{Z}(t)\right]\\ &= \left(\sigma_1\rho_{wz_1} - \sigma_2\rho_{wz_2}\right)\mathrm{d}t.\end{aligned}$$

[3] Here, σ is chosen such that the both sides of the equation $\sigma\,\mathrm{d}\bar{W}(t) \equiv \sigma_1\,\mathrm{d}\bar{W}_1(t) - \sigma_2\,\mathrm{d}\bar{W}_2(t)$ have the same variance under $\hat{\mathbb{Q}}$.

It follows that

$$\mathbb{E}_{\hat{\mathbb{Q}}}\left[\mathrm{d}\bar{W}(t)\,\mathrm{d}\bar{Z}(t)\right] = \frac{1}{\sigma}\left(\sigma_1\rho_{wz_1} - \sigma_2\rho_{wz_2}\right)\mathrm{d}t = \frac{\sigma_1\rho_{wz_1} - \sigma_2\rho_{wz_2}}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho_w\sigma_1\sigma_2}}\,\mathrm{d}t.$$

For brevity, we let $\bar{\rho}_{wz} = \frac{\sigma_1\rho_{wz_1} - \sigma_2\rho_{wz_2}}{\sqrt{\sigma_1^2+\sigma_2^2-2\rho_w\sigma_1\sigma_2}}$.

The mathematical structure of the asset yield ratio (4.13) is similar to that of the equity price model in Li (2020), which is based on a market factor structure. The equity price model in Li (2020, equations 3 and 4) assumes that the diffusion/Wiener risk and the jump risk each have an idiosyncratic and a systematic component, the latter being captured by an underlying market factor modelled as a separate SVJD process. In contrast, the diffusion risk and the jump risk in the asset yield ratio process (4.13) comes from the idiosyncratic diffusion and idiosyncratic jump components of the underlying assets S_1 and S_2. We assume that the only systematic risk source is the common stochastic volatility process defined by (4.10). Nonetheless, given the similarity in mathematical structures (i.e. the affine structure) one can derive the joint conditional characteristic function of $\tilde{S}$ and v using the techniques described by Li (2020), although this task is not within the scope of this chapter.[4]

4.3 The Exchange Option Pricing IPDE

Now we derive the integro-partial differential equation (IPDE) for the price of an exchange option written on S_1 and S_2. Denote by $C(t, S_1(t), S_2(t), v(t))$ the price of a European exchange option whose terminal payoff is given by

$$C\left(T, S_1(T), S_2(T), v(T)\right) = \left(S_1(T) - S_2(T)\right)^+,$$

where $x^+ \equiv \max\{x, 0\}$. Observe that the terminal payoff can be written as

$$\begin{aligned} C\left(T, S_1(T), S_2(T), v(T)\right) &= \left[\frac{S_1(T)e^{q_1T}}{e^{q_1T}} - \frac{S_2(T)e^{q_2T}}{e^{q_2T}}\right]^+ \\ &= e^{-q_1T}S_2(T)e^{q_2T}\left[\frac{S_1(T)e^{q_1T}}{S_2(T)e^{q_2T}} - e^{(q_1-q_2)T}\right]^+. \end{aligned}$$

Thus, the terminal payoff discounted with respect to the numéraire $\{S_2(t)e^{q_2t}\}$ is given by

$$\frac{C\left(T, S_1(T), S_2(T), v(T)\right)}{S_2(T)e^{q_2T}} = e^{-q_1T}\left(\tilde{S}(T) - e^{(q_1-q_2)T}\right)^+. \tag{4.14}$$

[4] See Chapter 2 for a similar derivation of the joint conditional characteristic function for a pair of correlated Bates (1996) SVJD processes.

Let

$$\tilde{C}(t, S_1(t), S_2(t), v(t)) \equiv \frac{C\left(t, S_1(t), S_2(t), v(t)\right)}{S_2(t)e^{q_2 t}}$$

denote the price at time t of the European exchange option price discounted by the second asset yield process. Then, assuming that no arbitrage opportunities exist, $\tilde{C}(t, S_1(t), S_2(t), v(t))$ is given by

$$\begin{aligned}\tilde{C}\left(t, S_1(t), S_2(t), v(t)\right) &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{C}\left(T, S_1(T), S_2(T), v(T)\right)\middle|\mathcal{F}_t\right] \\ &= e^{-q_1 T}\mathbb{E}_{\hat{\mathbb{Q}}}\left[\left(\tilde{S}(T) - e^{(q_1-q_2)T}\right)^+\middle|\mathcal{F}_t\right].\end{aligned} \tag{4.15}$$

In other words, the no-arbitrage price at any time $t < T$ of the European exchange option measured in units of the second asset yield process is the expected value, under the probability measure $\hat{\mathbb{Q}}$, of the option's terminal payoff measured in units of the second asset yield process (Geman et al., 1995; Benninga et al., 2002). From the last equation, we also note that the terminal payoff is variable only in the asset yield ratio $\tilde{S}(t)$. Thus, we assume that the discounted European exchange option price is represented by the process $\tilde{V}(t, \tilde{S}(t), v(t)) \equiv \tilde{C}\left(t, S_1(t), S_2(t), v(t)\right)$, and so

$$\tilde{V}(t, \tilde{S}(t), v(t)) = e^{-q_1 T}\mathbb{E}_{\hat{\mathbb{Q}}}\left[\left(\tilde{S}(T) - e^{(q_1-q_2)T}\right)^+\middle|\mathcal{F}_t\right]. \tag{4.16}$$

At this point, we have shown that, by taking the second stock's yield process as the numéraire asset, the exchange option pricing problem is simplified to pricing a European call option on the asset yield price ratio $\tilde{S}(t)$ with maturity date T and strike price $e^{(q_1-q_2)T}$. In the succeeding analysis, we shall take advantage of this dimension in reduction and employ techniques in pricing European call options under stochastic volatility and jump-diffusion dynamics, such as those presented by Bates (1996) and Cheang et al. (2013). An analysis similar to the preceding arguments can be carried out if one instead chooses to set the first stock's yield process $\{S_1(t)e^{q_1 t}\}$ as the numéraire asset to define the equivalent martingale measure. In doing so, the exchange option pricing problem simplifies to the valuation of a *put option* written on the asset yield ratio $(S_2(t)e^{q_2 t})/(S_1(t)e^{q_1 t})$.

As before, we assume that $\tilde{V}$ satisfies some differentiability conditions.

Assumption 4.2. The function $\tilde{V}$ is continuously differentiable in t and twice-continuously differentiable in $\tilde{s}$ and v

In the following proposition, we derive the IPDE that characterizes the discounted European exchange option price.

Proposition 4.3. *The price at time $t \in [0, T)$ of the European exchange option, denoted by $C(t, S_1(t), S_2(t), v(t))$, is given by*

$$C(t, S_1(t), S_2(t), v(t)) = S_2(t)e^{q_2 t}\tilde{V}(t, \tilde{S}(t), v(t)), \tag{4.17}$$

where $\tilde{V}$, which satisfies Assumption 4.2, is the solution of the terminal value problem

$$0 = \frac{\partial \tilde{V}}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}(t,\tilde{S}(t),v(t))\right], \qquad (t,\tilde{S}(t),\tilde{v}(t)) \in [0,T]\times\mathbb{R}^2_+ \tag{4.18}$$

$$\tilde{V}(T) = e^{-q_1 T}\left(\tilde{S}(T) - e^{(q_1-q_2)T}\right)^+, \tag{4.19}$$

with the IPDE operator $\mathcal{L}_{\tilde{s},v}$ defined as

$$\begin{aligned}\mathcal{L}_{\tilde{s},v}\left[\tilde{V}(t,\tilde{S},v)\right] = &-\tilde{S}\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)\frac{\partial\tilde{V}}{\partial\tilde{s}} + \left[\xi\eta - (\xi+\Lambda)v\right]\frac{\partial\tilde{V}}{\partial v}\\ &+ \frac{1}{2}\sigma^2 v\tilde{S}^2\frac{\partial^2\tilde{V}}{\partial\tilde{s}^2} + \frac{1}{2}\omega^2 v\frac{\partial^2\tilde{V}}{\partial v^2} + \omega(\sigma_1\rho_{wz_1} - \sigma_2\rho_{wz_2})v\tilde{S}\frac{\partial^2\tilde{V}}{\partial\tilde{s}\partial v}\\ &+ \tilde{\lambda}_1\mathbb{E}^{Y_1}_{\hat{\mathbb{Q}}}\left[\tilde{V}\left(t,\tilde{S}e^{Y_1},v\right) - \tilde{V}(t,\tilde{S},v)\right]\\ &+ \tilde{\lambda}_2\mathbb{E}^{Y_2}_{\hat{\mathbb{Q}}}\left[\tilde{V}\left(t,\tilde{S}e^{-Y_2},v\right) - \tilde{V}(t,\tilde{S},v)\right].\end{aligned} \tag{4.20}$$

(Note that all partial derivatives are evaluated at $(t,\tilde{S}(t),v(t))$.)

Proof. We first establish the martingale property of $\{\tilde{V}(t)\}$ under $\hat{\mathbb{Q}}$. Let $0 \le t_0 \le t \le T$. We note from equation (4.15) that

$$\begin{aligned}\tilde{C}\left(t_0,S_1(t_0),S_2(t_0),v(t_0)\right) &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{C}\left(T,S_1(T),S_2(T),v(T)\right)\middle|\mathcal{F}_{t_0}\right],\\ \tilde{C}\left(t,S_1(t),S_2(t),v(t)\right) &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{C}\left(T,S_1(T),S_2(T),v(T)\right)\middle|\mathcal{F}_{t}\right].\end{aligned}$$

Taking the expected value under $\hat{\mathbb{Q}}$ of both sides of the second equation conditional on $\mathcal{F}_{t_0}$ and using the tower property of conditional expectations, we find that

$$\begin{aligned}\mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{C}\left(t,S_1(t),S_2(t),v(t)\right)\middle|\mathcal{F}_{t_0}\right] &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{C}\left(T,S_1(T),S_2(T),v(T)\right)\middle|\mathcal{F}_{t}\right]\middle|\mathcal{F}_{t_0}\right]\\ &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{C}\left(T,S_1(T),S_2(T),v(T)\right)\middle|\mathcal{F}_{t_0}\right]\\ &= \tilde{C}\left(t_0,S_1(t_0),S_2(t_0),v(t_0)\right).\end{aligned}$$

Thus, the discounted European exchange option price process $\{\tilde{V}(t,\tilde{S}(t),v(t))\}_{0\le t\le T}$ is a martingale under $\hat{\mathbb{Q}}$.

Given equations (4.10) and (4.13), an application of Itô's Lemma for jump-diffusion processes yields the following stochastic differential equation for $\tilde{V}(t)$,

$$\begin{aligned}\mathrm{d}\tilde{V}(t) = \Bigg\{&\frac{\partial\tilde{V}}{\partial t} - \tilde{S}(t)\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)\frac{\partial\tilde{V}}{\partial\tilde{s}} + \left[\xi\eta - (\xi+\Lambda)v(t)\right]\frac{\partial\tilde{V}}{\partial v}\\ &+ \frac{1}{2}\sigma^2 v(t)\tilde{S}^2(t)\frac{\partial^2\tilde{V}}{\partial\tilde{s}^2} + \frac{1}{2}\omega^2 v(t)\frac{\partial^2\tilde{V}}{\partial v^2} + \sigma\omega\bar{\rho}_{wz}v(t)\tilde{S}(t)\frac{\partial^2\tilde{V}}{\partial\tilde{s}\partial v}\Bigg\}\,\mathrm{d}t\end{aligned}$$

$$
\begin{aligned}
&+ \sigma\sqrt{v(t)}\tilde{S}(t)\frac{\partial \tilde{V}}{\partial \tilde{s}}\,\mathrm{d}\bar{W}(t) + \omega\sqrt{v(t)}\frac{\partial \tilde{V}}{\partial v}\,\mathrm{d}\bar{Z}(t) \\
&+ \int_{\mathbb{R}}\left[\tilde{V}\left(t, \tilde{S}(t)e^{y}, v(t)\right) - \tilde{V}(t, \tilde{S}(t), v(t))\right] p_1(\mathrm{d}y, \mathrm{d}t) \\
&+ \int_{\mathbb{R}}\left[\tilde{V}\left(t, \tilde{S}(t)e^{-y}, v(t)\right) - \tilde{V}(t, \tilde{S}(t), v(t))\right] p_2(\mathrm{d}y, \mathrm{d}t).
\end{aligned}
$$

Replacing the counting measures $p(\mathrm{d}y_1, \mathrm{d}t)$ and $p(\mathrm{d}y_2, \mathrm{d}t)$ by their $\hat{\mathbb{Q}}$-compensated counterparts, we find that

$$
\begin{aligned}
\mathrm{d}\tilde{V}(t) = &\left\{\frac{\partial \tilde{V}}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}(t, \tilde{S}(t), v(t))\right]\right\}\mathrm{d}t + \sigma\sqrt{v(t)}\tilde{S}(t)\frac{\partial \tilde{V}}{\partial \tilde{s}}\,\mathrm{d}\bar{W}(t) + \omega\sqrt{v(t)}\frac{\partial \tilde{V}}{\partial v}\,\mathrm{d}\bar{Z}(t) \\
&+ \int_{\mathbb{R}}\left[\tilde{V}\left(t, \tilde{S}(t)e^{y}, v(t)\right) - \tilde{V}(t, \tilde{S}(t), v(t))\right] q_1(\mathrm{d}y, \mathrm{d}t) \\
&+ \int_{\mathbb{R}}\left[\tilde{V}\left(t, \tilde{S}(t)e^{-y}, v(t)\right) - \tilde{V}(t, \tilde{S}(t), v(t))\right] q_2(\mathrm{d}y, \mathrm{d}t),
\end{aligned}
\tag{4.21}
$$

where $\mathcal{L}_{\tilde{s},v}$ is the IPDE operator defined by equation (4.20). Since $\{\tilde{V}(t)\}$ is a $\hat{\mathbb{Q}}$-martingale, the drift must be equal to zero, giving us the equation

$$
0 = \frac{\partial \tilde{V}}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}(t, \tilde{S}(t), v(t))\right].
$$

Note that all partial derivatives are evaluated at $(t, \tilde{S}(t), v(t))$.

In the case of the European exchange option, IPDE (4.18) must be solved for $0 \le t \le T$, $\tilde{S}(t) > 0$, and $v(t) > 0$, subject to the terminal condition

$$
\tilde{V}(T) \equiv \tilde{V}(T, \tilde{S}(T), v(T)) = e^{-q_1 T}\left(\tilde{S}(T) - e^{(q_1 - q_1)T}\right)^{+}.
$$

Once the IPDE has been solved for $\tilde{V}(t)$, the price of the European exchange option may be obtained by multiplying $\tilde{V}$ by $S_2(t)e^{q_2 t}$. ■

Let $C^A(t, S_1(t), S_2(t), v(t))$ be the price at time t of an American exchange option written on S_1 and S_2. If the option is exercised at any stopping time $u \in [t, T]$, then the reward realized by the option holder is

$$
\text{reward} = \left(S_1(u) - S_2(u)\right)^{+} = e^{-q_1 u}S_2(u)e^{q_2 u}\left(\tilde{S}(u) - e^{(q_1 - q_2)u}\right)^{+}.
$$

Applying the discounting factor $\mathsf{DF}(t, u)$ on this quantity gives the discounted (or present value of the) reward at time t,

$$
\text{discounted reward} = S_2(t)e^{q_2 t}e^{-q_1 u}\left(\tilde{S}(u) - e^{(q_1 - q_2)u}\right)^{+}.
$$

The holder of the American exchange option may exercise the option at any time before T and will exercise the option at the (stopping) time at which the $\hat{\mathbb{Q}}$-expectation of the discounted option reward is maximal. Thus, the price of the American exchange option is given by

$$C^A(t, S_1(t), S_2(t), v(t)) = S_2(t)e^{q_2 t} \sup_{u \in \mathscr{T}_{t,T}} e^{-q_1 u} \mathbb{E}_{\hat{\mathbb{Q}}} \left[\left(\tilde{S}(u) - e^{(q_1 - q_2)u} \right)^+ \middle| \mathcal{F}_t \right], \tag{4.22}$$

where the supremum is taken over $\mathscr{T}_{t,T}$, the collection of all $\{\mathcal{F}_t\}$-stopping times u in $[t, T]$. In terms of the discounted American exchange option price $\tilde{V}^A(t, \tilde{S}(t), v(t))$, we can write

$$\tilde{V}^A(t, \tilde{S}(t), v(t)) = \sup_{u \in \mathscr{T}_{t,T}} e^{-q_1 u} \mathbb{E}_{\hat{\mathbb{Q}}} \left[\left(\tilde{S}(u) - e^{(q_1 - q_2)u} \right)^+ \middle| \mathcal{F}_t \right]. \tag{4.23}$$

From here, we see that the change of numéraire reduces the problem to pricing an American call option on the asset yield price ratio $\tilde{S}(t)$ with maturity date T and strike price $e^{(q_1-q_2)T}$, similar to our observation for the European exchange option. The price of the American exchange option also hedges against the exchange option's payoff in the sense that

$$\begin{aligned} C^A(t, S_1(t), S_2(t), v(t)) &\geq \left(S_1(t) - S_2(t) \right)^+, \qquad \forall t \in [0, T) \\ C^A(T, S_1(T), S_2(T), v(T)) &= \left(S_1(T) - S_2(T) \right)^+. \end{aligned}$$

In terms of the $\tilde{V}^A(t, \tilde{S}(t), v(t))$ these properties may be written as

$$\begin{aligned} \tilde{V}^A(t, \tilde{S}(t), v(t)) &\geq e^{-q_1 t} \left(\tilde{S}(t) - e^{(q_1 - q_2)t} \right)^+, \qquad \forall t \in [0, T) \\ \tilde{V}^A(T, \tilde{S}(T), v(T)) &= e^{-q_1 T} \left(\tilde{S}(T) - e^{(q_1 - q_2)T} \right)^+. \end{aligned}$$

Before prescribing additional boundary conditions to IPDE (4.18) for the American exchange option, we first define the continuation and stopping regions, denoted by $\mathcal{C}$ and $\mathcal{S}$, respectively, that divide the domain $[0, T] \times \mathbb{R}^2_+$ of IPDE (4.18). These regions are given by

$$\begin{aligned} \mathcal{S} &= \left\{ (t, \tilde{S}, v) \in [0, T] \times \mathbb{R}^2_+ : \tilde{V}^A(t, \tilde{S}, v) = e^{-q_1 t} \left(\tilde{S} - e^{(q_1 - q_2)t} \right)^+ \right\}, \\ \mathcal{C} &= \left\{ (t, \tilde{S}, v) \in [0, T] \times \mathbb{R}^2_+ : \tilde{V}^A(t, \tilde{S}, v) > e^{-q_1 t} \left(\tilde{S} - e^{(q_1 - q_2)t} \right)^+ \right\}. \end{aligned} \tag{4.24}$$

Denote by $\mathcal{S}(t)$ and $\mathcal{C}(t)$ the t-cross-sections of the stopping and continuation regions at any time $t \in [0, T]$.

Following the discussion in Section 2.6.1, there exists a critical stock price ratio $B(t, v) \geq 1$, dependent on the current variance level, such that the stopping and continuation regions can be

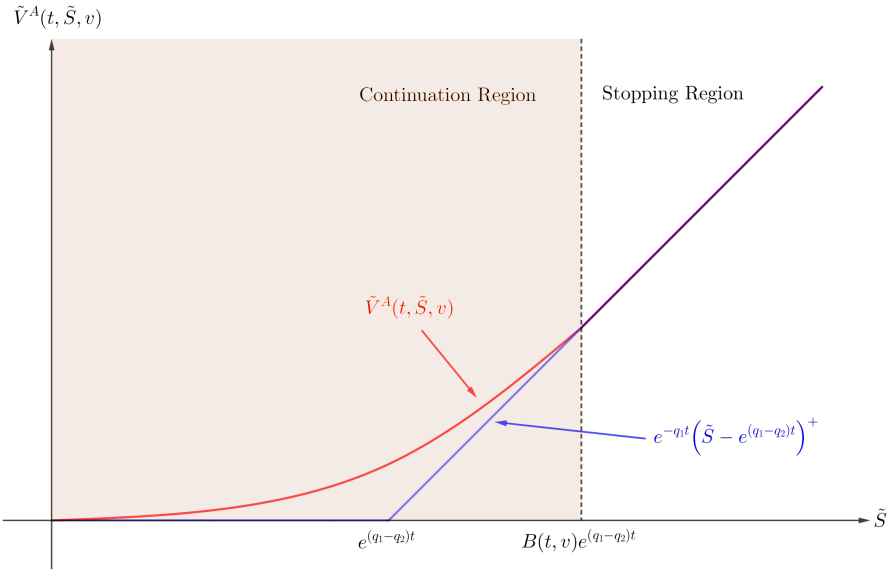


Figure 4.1: The early exercise boundary and the continuation and stopping regions for the American exchange option in terms of the asset yield ratio $\tilde{S}$.

written as

$$\mathcal{S} = \left\{(t, S_1, S_2, v) \in [0,T] \times \mathbb{R}^3_+ : S_1 \geq B(t,v) S_2\right\},$$
$$\mathcal{C} = \left\{(t, S_1, S_2, v) \in [0,T] \times \mathbb{R}^3_+ : S_1 < B(t,v) S_2\right\}.$$

The line $s_1 = B(t,v)s_2$ on the $s_1 s_2$-plane is known as the early exercise boundary. In terms of the asset yield ratio, we have

$$\begin{aligned} \mathcal{S} &= \left\{(t, \tilde{S}, v) \in [0,T] \times \mathbb{R}^2_+ : \tilde{S} \geq B(t,v) e^{(q_1-q_2)t}\right\}, \\ \mathcal{C} &= \left\{(t, \tilde{S}, v) \in [0,T] \times \mathbb{R}^2_+ : \tilde{S} < B(t,v) e^{(q_1-q_2)t}\right\}. \end{aligned} \tag{4.25}$$

For a fixed $t \in [0,T]$ and $v \in (0,\infty)$, the early exercise boundary and the continuation and stopping regions are illustrated in Figure 4.1. It is known that in the continuation region the American exchange option behaves like its live European counterpart, and so $\tilde{V}^A$ satisfies IPDE (4.18) for $(t, \tilde{S}, v) \in \mathcal{C}$.[5] For simplicity, we let $A(t,v) = B(t,v)e^{(q_1-q_2)t}$ denote the early exercise boundary on $\tilde{S}$.

As in Section 2.6.1, to close the free boundary problem for $\tilde{V}^A$, we require the value-matching

[5] Mishura and Shevchenko (2009) analyzed, in further detail, the properties of the exercise region of the finite-maturity American exchange option in a pure diffusion setting. In the same setting, Villeneuve (1999) established the nonemptiness of exercise regions of American rainbow options, which include spread and exchange options as special cases. Still in the pure-diffusion setting, Rapuch (2005) extended the results of Villeneuve (1999) to consider more general payoff functions and the case of local stochastic volatility.

and smooth-pasting conditions. Specifically, the required value-matching condition is

$$\tilde{V}^A(t, A(t,v), v(t)) = e^{-q_1 t}\left(A(t,v) - e^{(q_1-q_2)t}\right), \tag{4.26}$$

and the smooth-pasting condition is

$$\lim_{\tilde{S}\to A(t,v)} \frac{\partial \tilde{V}^A}{\partial \tilde{s}}(t, \tilde{S}(t), v(t)) = e^{-q_1 t}. \tag{4.27}$$

This results from Assumption 4.2 and the known analytical form of $\tilde{V}^A$

$$\tilde{V}^A(t, \tilde{S}(t), v(t)) = e^{-q_1 t}\tilde{S}(t) - e^{-q_2 t}, \qquad \text{whenever } \tilde{S}(t) \geq A(t,v).$$

Furthermore, Assumption 4.2 also implies that

$$\begin{aligned}
\lim_{\tilde{S}\to A(t,v)} \frac{\partial \tilde{V}^A}{\partial v}(t, \tilde{S}(t), v(t)) &= 0, \\
\lim_{\tilde{S}\to A(t,v)} \frac{\partial \tilde{V}^A}{\partial t}(t, \tilde{S}(t), v(t)) &= -q_1 e^{-q_1 t}\tilde{S}(t) + q_2 e^{-q_2 t}.
\end{aligned}$$

Therefore, the discounted American exchange option price is a solution to IPDE (4.18) over the restricted domain $0 \leq t \leq T$, $0 < \tilde{S} < A(t,v)$, $0 < v < \infty$. The IPDE has terminal and boundary conditions

$$\begin{aligned}
\tilde{V}(T, \tilde{S}(T), v(T)) &= e^{-q_1 T}\left(\tilde{S}(T) - e^{(q_1-q_2)T}\right)^+, \\
\tilde{V}(t, 0, v(t)) &= 0,
\end{aligned} \tag{4.28}$$

value-matching condition (4.26) and smooth-pasting condition (4.27).

4.4 An Early Exercise Representation

In this section, we derive the early exercise representation of the discounted American exchange option price $\tilde{V}^A(t, \tilde{S}(t), v(t))$.

Proposition 4.4. *Suppose $\tilde{V}^A(t, \tilde{S}, v)$ satisfies Assumption 4.2. Assume further that the smooth pasting conditions (4.27) across the early exercise boundary hold. Then $\tilde{V}^A(t, \tilde{S}(t), v(t))$ can be expressed as*

$$\tilde{V}^A(t, \tilde{S}(t), v(t)) = \tilde{V}(t, \tilde{S}(t), v(t)) + \tilde{V}^P(t, \tilde{S}(t), v(t)), \tag{4.29}$$

where $\tilde{V}(t, \tilde{S}(t), v(t))$ is the discounted European exchange option price given by equation (4.16)

and $\tilde{V}^P(t,\tilde{S}(t),v(t))$ *is the early exercise premium given by*

$$\tilde{V}^P(t,\tilde{S}(t),v(t)) = -\mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T \left\{\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right\} \mathrm{d}s \middle| \mathcal{F}_t\right]. \tag{4.30}$$

Note that all partial derivatives in $\partial\tilde{V}^A/\partial t + \mathcal{L}_{\tilde{s},v}[\tilde{V}^A(s,\tilde{S}(s),v(s))]$ *are all evaluated at* $(s,\tilde{S}(s),v(s))$.

Proof. An application of Itô's Lemma for jump-diffusion processes verifies that $\tilde{V}^A(t) \equiv \tilde{V}^A(t,\tilde{S}(t),v(t))$ satisfies SDE (4.21). Integrating the SDE over the time interval $[t_0,t]$, where $0 \le t_0 \le t \le T$, we find that

$$\begin{aligned}
\tilde{V}^A(t) = \tilde{V}^A(t_0) &+ \int_{t_0}^t \left\{\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right\} \mathrm{d}s \\
&+ \int_{t_0}^t \sigma\sqrt{v(s)}\tilde{S}(s)\frac{\partial \tilde{V}^A}{\partial \tilde{s}}\,\mathrm{d}\bar{W}(s) + \int_{t_0}^t \omega\sqrt{v(s)}\frac{\partial \tilde{V}^A}{\partial v}\,\mathrm{d}\bar{Z}(s) \\
&+ \int_{t_0}^t \int_{\mathbb{R}} \left[\tilde{V}^A\left(s,\tilde{S}(s)e^{y},v(s)\right) - \tilde{V}^A(s,\tilde{S}(s),v(s))\right] q_1(\mathrm{d}y,\mathrm{d}s) \\
&+ \int_{t_0}^t \int_{\mathbb{R}} \left[\tilde{V}^A\left(s,\tilde{S}(s)e^{-y},v(s)\right) - \tilde{V}^A(s,\tilde{S}(s),v(s))\right] q_2(\mathrm{d}y,\mathrm{d}s).
\end{aligned}$$

Next, we take the $\hat{\mathbb{Q}}$-expectation of the above equation conditional on $\mathcal{F}_{t_0}$. Noting that the (stochastic) integrals with respect to the Wiener processes and the compensated counting measures are all independent of $\mathcal{F}_{t_0}$ (Brémaud, 1981; Runggaldier, 2003) and that the unconditional expectation of the integrals with respect to the Wiener processes is zero, we have

$$\begin{aligned}
\mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{V}^A(t)\middle|\mathcal{F}_{t_0}\right] = \tilde{V}^A(t_0) &+ \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_{t_0}^t \left\{\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right\} \mathrm{d}s \middle| \mathcal{F}_{t_0}\right] \\
&+ \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_{t_0}^t \int_{\mathbb{R}} \left[\tilde{V}^A\left(s,\tilde{S}(s)e^{y},v(s)\right) - \tilde{V}^A(s,\tilde{S}(s),v(s))\right] q_1(\mathrm{d}y,\mathrm{d}s)\right] \\
&+ \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_{t_0}^t \int_{\mathbb{R}} \left[\tilde{V}^A\left(s,\tilde{S}(s)e^{-y},v(s)\right) - \tilde{V}^A(s,\tilde{S}(s),v(s))\right] q_2(\mathrm{d}y,\mathrm{d}s)\right].
\end{aligned}$$

The martingale representation theorem for marked point processes (see Brémaud, 1981; Runggaldier, 2003) states that there are $\hat{\mathbb{Q}}$-martingales $\{\mathcal{M}_1(t)\}$ and $\{\mathcal{M}_2(t)\}$ such that

$$\begin{aligned}
\mathcal{M}_1(t) &= \mathcal{M}_1(0) + \int_0^t \int_{\mathbb{R}} \left[\tilde{V}^A\left(s,\tilde{S}(s)e^{y},v(s)\right) - \tilde{V}^A(s,\tilde{S}(s),v(s))\right] q_1(\mathrm{d}y,\mathrm{d}s), \\
\mathcal{M}_2(t) &= \mathcal{M}_2(0) + \int_0^t \int_{\mathbb{R}} \left[\tilde{V}^A\left(s,\tilde{S}(s)e^{-y},v(s)\right) - \tilde{V}^A(s,\tilde{S}(s),v(s))\right] q_2(\mathrm{d}y,\mathrm{d}s).
\end{aligned}$$

Subtracting $\mathcal{M}_1(t_0)$ from both sides of the first equation yields

$$\mathcal{M}_1(t) = \mathcal{M}_1(t_0) + \int_{t_0}^t \int_{\mathbb{R}} \left[\tilde{V}^A\left(s,\tilde{S}(s)e^{y},v(s)\right) - \tilde{V}^A(s,\tilde{S}(s),v(s))\right] q_1(\mathrm{d}y,\mathrm{d}s).$$

Taking the expectation of both sides conditional on $\mathcal{F}_{t_0}$, using the martingale property of $\{\mathcal{M}_1(t)\}$, and noting that the integral with respect to $q_1(\mathrm{d}y, \mathrm{d}s)$ is independent of $\mathcal{F}_0$ allow us to conclude that

$$\mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_{t_0}^{t}\int_{\mathbb{R}}\left[\tilde{V}^A\left(s, \tilde{S}(s)e^y, v(s)\right) - \tilde{V}^A(s, \tilde{S}(s), v(s))\right] q_1(\mathrm{d}y, \mathrm{d}s)\right] = 0.$$

An analogous argument applies to the integral with respect to $q(\mathrm{d}y_2, \mathrm{d}s)$. Therefore, the conditional expectation of $\tilde{V}^A(t)$ simplifies to

$$\mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{V}^A(t)\middle|\mathcal{F}_{t_0}\right] = \tilde{V}^A(t_0) + \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_{t_0}^{t}\left\{\frac{\partial\tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s, \tilde{S}(s), v(s))\right]\right\}\mathrm{d}s\middle|\mathcal{F}_{t_0}\right]. \tag{4.31}$$

In equation (4.31), set $t = T$ and $t_0 = t$. We then have

$$\mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{V}^A(T)\middle|\mathcal{F}_{t}\right] = \tilde{V}^A(t) + \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_{t}^{T}\left\{\frac{\partial\tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s, \tilde{S}(s), v(s))\right]\right\}\mathrm{d}s\middle|\mathcal{F}_{t}\right]. \tag{4.32}$$

Terminal condition (4.28) and equation (4.16) imply that the left-hand side of this equation is equal to $\tilde{V}(t, \tilde{S}(t), v(t))$, the price of the European exchange option. Rearranging yields the result stated in the proposition. ■

For a fixed $t \in [0, T]$, let $\mathcal{A}(t)$ be the event that the state variables $\tilde{S}(t)$ and $v(t)$ are in the t-cross-section of the stopping region $\mathcal{S}(t)$; that is,

$$\mathcal{A}(t) \equiv \left\{(\tilde{S}(t), v(t)) \in \mathcal{S}(t)\right\}.$$

The complement event $\mathcal{A}^c(t)$ denotes the event that $\tilde{S}(t)$ and $v(t)$ are in the t-cross-section of the continuation region $\mathcal{C}(t)$, since at any time t, $(\tilde{S}(t), v(t))$ is in either the stopping or the continuation region (but never both at the same time).

We now seek to evaluate the expectation defining the early exercise premium in equation (4.30). To this end, we have the following proposition.

Proposition 4.5. *The early exercise premium* $\tilde{V}^P(t,\tilde{S}(t),v(t))$ *is given by*

$$\begin{aligned}
&\tilde{V}^P(t,\tilde{S}(t),v(t)) \\
&\quad= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T \left(q_1 e^{-q_1 s}\tilde{S}(s) - q_2 e^{-q_2 s}\right)\mathbf{1}(\mathcal{A}(s))\,\mathrm{d}s \,\middle|\, \mathcal{F}_t\right] \\
&\qquad - \tilde{\lambda}_1 \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T \mathbb{E}^{Y_1}_{\hat{\mathbb{Q}}}\left[\left(\tilde{V}^A\left(s,\tilde{S}(s)e^{Y_1},v(s)\right)\right.\right.\right. \\
&\qquad\qquad \left.\left.\left. - \left(e^{-q_1 s}\tilde{S}(s)e^{Y_1} - e^{-q_2 s}\right)\right)\mathbf{1}(\mathcal{A}_1(s))\right]\mathrm{d}s \,\middle|\, \mathcal{F}_t\right] \\
&\qquad - \tilde{\lambda}_2 \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T \mathbb{E}^{Y_2}_{\hat{\mathbb{Q}}}\left[\left(\tilde{V}^A\left(s,\tilde{S}(s)e^{-Y_2},v(s)\right)\right.\right.\right. \\
&\qquad\qquad \left.\left.\left. - \left(e^{-q_1 s}\tilde{S}(s)e^{-Y_2} - e^{-q_2 s}\right)\right)\mathbf{1}(\mathcal{A}_2(s))\right]\mathrm{d}s \,\middle|\, \mathcal{F}_t\right],
\end{aligned} \tag{4.33}$$

where $\mathbf{1}(\cdot)$ *is the indicator function,* $\mathcal{A}(s)$ *is the event* $\{(\tilde{S}(t),v(t)) \in \mathcal{S}(t)\}$ *and*

$$\mathcal{A}_1(s) = \left\{B(s,v)e^{(q_1-q_2)s} \le \tilde{S}(s) < B(s,v)e^{(q_1-q_2)s}e^{-Y_1}\right\}, \tag{4.34}$$

$$\mathcal{A}_2(t) = \left\{B(s,v)e^{(q_1-q_2)s} \le \tilde{S}(s) < B(s,v)e^{(q_1-q_2)s}e^{Y_2}\right\}. \tag{4.35}$$

Proof. Note first that for any time $s \in [t,T]$,

$$\begin{aligned}
\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right] &= \left(\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right)\mathbf{1}(\mathcal{A}(t)) \\
&\quad + \left(\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right)\mathbf{1}(\mathcal{A}^c(t)) \\
&= \left(\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right)\mathbf{1}(\mathcal{A}(t)),
\end{aligned}$$

since in the continuation region, the American exchange option behaves like its live European counterpart and the integro-partial differential terms vanish. In the stopping region (i.e. if $\mathbf{1}(\mathcal{A}(s)) = 1$), we note that

$$\tilde{V}^A(s,\tilde{S}(s),v(s)) = e^{-q_1 s}\tilde{S}(s) - e^{-q_2 s},$$

(see equation (4.24)). Applying the integro-partial differential operators, we find that

$$\begin{aligned}
&\left(\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right)\mathbf{1}(\mathcal{A}(s)) \\
&\quad= \left[-q_1 e^{-q_1 s}\tilde{S}(s) + q_2 e^{-q_2 s} - e^{-q_1 s}\tilde{S}(s)\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)\right]\mathbf{1}(\mathcal{A}(s))
\end{aligned}$$

$$+\tilde{\lambda}_1\mathbb{E}^{Y_1}_{\hat{\mathbb{Q}}}\left[\tilde{V}^A\left(s,\tilde{S}(s)e^{Y_1},v(s)\right)-\left(e^{-q_1 s}\tilde{S}(s)-e^{-q_2 s}\right)\right]\mathbf{1}(\mathcal{A}(s))$$
$$+\tilde{\lambda}_2\mathbb{E}^{Y_2}_{\hat{\mathbb{Q}}}\left[\tilde{V}^A\left(s,\tilde{S}(s)e^{-Y_2},v(s)\right)-\left(e^{-q_1 s}\tilde{S}(s)-e^{-q_2 s}\right)\right]\mathbf{1}(\mathcal{A}(s)).$$

Recalling the definition of $\tilde{\kappa}_1$ and $\tilde{\kappa}_2^-$ and rearranging the terms, we find that

$$\begin{aligned}
&\left(\frac{\partial\tilde{V}^A}{\partial t}+\mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right)\mathbf{1}(\mathcal{A}(s))\\
&\quad=\left[-q_1e^{-q_1 s}\tilde{S}(s)+q_2e^{-q_2 s}\right]\mathbf{1}(\mathcal{A}(s))\\
&\qquad+\tilde{\lambda}_1\mathbb{E}^{Y_1}_{\hat{\mathbb{Q}}}\left[\tilde{V}^A\left(s,\tilde{S}(s)e^{Y_1},v(s)\right)-\left(e^{-q_1 s}\tilde{S}(s)e^{Y_1}-e^{-q_2 s}\right)\right]\mathbf{1}(\mathcal{A}(s))\\
&\qquad+\tilde{\lambda}_2\mathbb{E}^{Y_2}_{\hat{\mathbb{Q}}}\left[\tilde{V}^A\left(s,\tilde{S}(s)e^{-Y_2},v(s)\right)-\left(e^{-q_1 s}\tilde{S}(s)e^{-Y_2}-e^{-q_2 s}\right)\right]\mathbf{1}(\mathcal{A}(s)).
\end{aligned}$$

Note that the expectations above contain option prices determined *after the jump* in $\tilde{S}(s)$ occurring at time s. These jumps can cause the asset price ratio to go back into the continuation region, which may result to losses on the option holder's part should he or she choose to exercise at time s.

Define $\mathcal{A}_1(s)$ and $\mathcal{A}_2(s)$ as the events in which the asset price ratio is initially in the stopping region but is sent back into the continuation region after a jump by a factor e^{Y_1} and e^{-Y_2}, respectively, at time s. That is,

$$\begin{aligned}
\mathcal{A}_1(s)&=\mathcal{A}(s)\cap\left\{\tilde{S}(s)e^{Y_1}<B(s,v)e^{(q_1-q_2)s}\right\}\\
&=\left\{B(s,v)e^{(q_1-q_2)s}\leq\tilde{S}(s)<B(s,v)e^{(q_1-q_2)s}e^{-Y_1}\right\},\\
\mathcal{A}_2(t)&=\mathcal{A}(s)\cap\left\{\tilde{S}(s)e^{-Y_2}<B(s,v)e^{(q_1-q_2)s}\right\}\\
&=\left\{B(s,v)e^{(q_1-q_2)s}\leq\tilde{S}(s)<B(s,v)e^{(q_1-q_2)s}e^{Y_2}\right\}.
\end{aligned}$$

Following the stopping and continuation criteria in equation (4.24), we note that

$$\tilde{V}^A\left(s,\tilde{S}(s)e^{Y_1},v(s)\right)\geq e^{-q_1 s}\tilde{S}(s)e^{Y_1}-e^{-q_2 s},\tag{4.36}$$

with the strict inequality occurring when $\mathcal{A}_1(s)$ is true, and

$$\tilde{V}^A\left(s,\tilde{S}(s)e^{-Y_2},v(s)\right)\geq e^{-q_1 s}\tilde{S}(s)e^{-Y_2}-e^{-q_2 s},\tag{4.37}$$

with the strict inequality occurring when $\mathcal{A}_2(t)$ is true. Therefore, we have

$$\left(\frac{\partial\tilde{V}^A}{\partial t}+\mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(s,\tilde{S}(s),v(s))\right]\right)\mathbf{1}(\mathcal{A}(s))$$

$$
\begin{aligned}
&= -\left[q_1 e^{-q_1 s}\tilde{S}(s) - q_2 e^{-q_2 s}\right]\mathbf{1}(\mathcal{A}(s)) \\
&\quad + \tilde{\lambda}_1 \mathbb{E}^{Y_1}_{\hat{\mathbb{Q}}}\left[\left(\tilde{V}^A\left(s, \tilde{S}(s)e^{Y_1}, v(s)\right) - \left(e^{-q_1 s}\tilde{S}(s)e^{Y_1} - e^{-q_2 s}\right)\right)\mathbf{1}(\mathcal{A}_1(s))\right] \\
&\quad + \tilde{\lambda}_2 \mathbb{E}^{Y_2}_{\hat{\mathbb{Q}}}\left[\left(\tilde{V}^A\left(s, \tilde{S}(s)e^{-Y_2}, v(s)\right) - \left(e^{-q_1 s}\tilde{S}(s)e^{-Y_2} - e^{-q_2 s}\right)\right)\mathbf{1}(\mathcal{A}_2(s))\right].
\end{aligned}
$$

Using the above expression in the early exercise premium given in equation (4.30), we obtain (4.33). ■

If $G_1(y)$ and $G_2(y)$ are the probability density functions of Y_1 and Y_2, respectively under $\hat{\mathbb{Q}}$, then the early exercise premium may be written as

$$
\begin{aligned}
&\tilde{V}^P(t, \tilde{S}(t), v(t)) \\
&\quad = \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T \left(q_1 e^{-q_1 s}\tilde{S}(s) - q_2 e^{-q_2 s}\right)\mathbf{1}(\mathcal{A}(s))\,\mathrm{d}s \,\middle|\, \mathcal{F}_t\right] \\
&\qquad - \tilde{\lambda}_1 \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T \int_{-\infty}^{b(s,\tilde{S}(s),v(s))} \left(\tilde{V}^A\left(s, \tilde{S}(s)e^{y}, v(s)\right) - \left(e^{-q_1 s}\tilde{S}(s)e^{y} - e^{-q_2 s}\right)\right)\right. \\
&\qquad\qquad \left. \times\, G_1(y)\mathbf{1}(\mathcal{A}(s))\,\mathrm{d}y\,\mathrm{d}s \,\middle|\, \mathcal{F}_t\right] \\
&\qquad - \tilde{\lambda}_2 \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T \int_{-b(s,\tilde{S}(s),v(s))}^{\infty} \left(\tilde{V}^A\left(s, \tilde{S}(s)e^{-y}, v(s)\right) - \left(e^{-q_1 s}\tilde{S}(s)e^{-y} - e^{-q_2 s}\right)\right)\right. \\
&\qquad\qquad \left. \times\, G_2(y)\mathbf{1}(\mathcal{A}(s))\,\mathrm{d}y\,\mathrm{d}s \,\middle|\, \mathcal{F}_t\right],
\end{aligned}
\tag{4.38}
$$

where $b(s, \tilde{S}(s), v(s)) \equiv \ln\left(B(s, v(s))e^{(q_1-q_2)s}/\tilde{S}(s)\right)$.

The early exercise premium can be decomposed into a diffusion component (the positive term) and a jump component (the negative terms). However, unlike the early exercise premium derived by Cheang et al. (2013) for an American call option under SVJD dynamics, our early exercise premium representation contains two jump terms. This is because $\tilde{S}(t)$ has two sources of jumps: the jumps in the price of the first asset and the jumps in the numéraire. Nonetheless, the interpretation remains the same: the diffusion term captures the discounted expected value of cash flows due to dividends when asset prices are in the stopping region and the jump terms capture the rebalancing costs incurred by the holder of the American exchange option when a jump instantaneously occurs in the price of either asset, causing $\tilde{S}(t)$ to jump back into the continuation region immediately after the option is exercised. In this situation, the investor is unable to adjust the decision to exercise in response to the instantaneous jump in asset prices and is therefore vulnerable to the rebalancing cost described earlier. Figure 4.2 illustrates the loss (captured by the difference in option value and the exercise value) incurred by the option holder when the asset yield ratio jumps back into the continuation region due to a jump, by a factor e^{Y_1}, in the price of the first asset. A similar

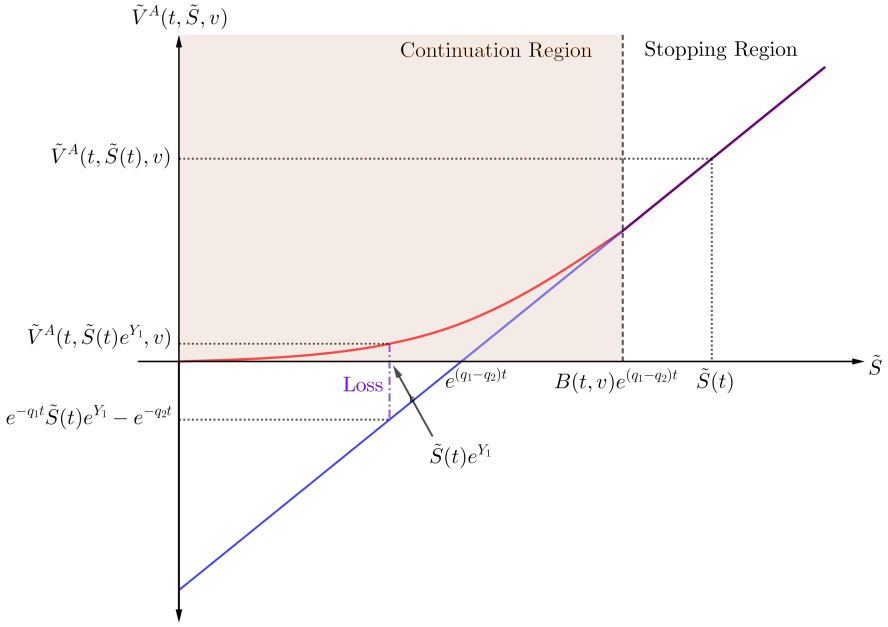


Figure 4.2: Loss incurred by the option holder when the asset ratio instantaneously jumps from $\tilde{S}(t)$ in the stopping region to $\tilde{S}(t)e^{Y_1}$ back in the continuation region.

graphical analysis holds if the price of the numéraire asset instantaneously jumps instead of the price of the first asset (in this case, the new asset yield ratio is $\tilde{S}(t)e^{-Y_2}$).

Now suppose the asset price ratio stays within the stopping region after the jumps have occurred; that is, suppose $\tilde{S}(t)e^{Y_1} \geq B(t,v)e^{(q_1-q_2)t}$ and $\tilde{S}(t)e^{-Y_2} \geq B(t,v)e^{(q_1-q_2)t}$.[6] Then equality will hold in conditions (4.36) and (4.37) and so the early exercise premium from equation (4.33) will just be

$$\tilde{V}^P(t,\tilde{S}(t),v(t)) = \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T \left(q_1 e^{-q_1 s}\tilde{S}(s) - q_2 e^{-q_2 s}\right)\mathbf{1}(\mathcal{A}(s))\,\mathrm{d}s \middle| \mathcal{F}_t\right].$$

Note that, provided that $q_1 > q_2$, the integrand is strictly positive $\hat{\mathbb{Q}}$-a.s. for all $s \in [t,T]$ and so $\tilde{V}^P(t,\tilde{S}(t),v(t))$ is $\hat{\mathbb{Q}}$-a.s. decreasing as t increases to T. Therefore, $\tilde{V}^A(t,\tilde{S}(t),v(t))$ has a Doob-Meyer decomposition[7] given by equation (4.29) and is thus a supermartingale provided that pre- and post-jump asset price ratios are in the stopping region.

Recall that the discounted American exchange option $\tilde{V}^A(t,\tilde{S}(t),v(t))$ is a solution of the homogeneous IPDE

$$\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(t,\tilde{S}(t),v(t))\right] = 0$$

over the restricted domain $0 \leq t \leq T$, $0 < \tilde{S}(t) < B(t,v)e^{(q_1-q_2)t}$, and $0 < v < \infty$ subject to

[6] Equivalently, these events may be written as

$$Y_1 \geq \ln\left(\frac{B(t,v)e^{(q_1-q_2)t}}{\tilde{S}(t)}\right) \quad \text{and} \quad Y_2 \leq -\ln\left(\frac{B(t,v)e^{(q_1-q_2)t}}{\tilde{S}(t)}\right).$$

[7] See Theorem 1.2.1.6 of Jeanblanc et al. (2009).

the value-matching condition (4.26), the smooth-pasting condition (4.27), and boundary conditions (4.28). Following Jamshidian (1992) and Chiarella and Ziogas (2004), the restriction on the domain can be lifted by adding an appropriate inhomogeneous term to the IPDE such that the equation holds for all values of the asset yield ratio.[8] The inhomogeneous IPDE corresponding to our analysis is presented in the following proposition.

Proposition 4.6. *Suppose Assumption 4.2 holds and let G_1 and G_2 denote the jump size density functions under $\hat{\mathbb{Q}}$. The discounted American exchange option price $\tilde{V}^A(t,\tilde{S},v)$ is a solution to the inhomogeneous IPDE*

$$0 = \frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(t,\tilde{S}(t),v(t))\right] + \Xi(t,\tilde{S}(t),v(t)), \tag{4.39}$$

where the inhomogeneous term Ξ is given by

$$\begin{aligned}
&\Xi(t,\tilde{S}(t),v(t)) \\
&\quad = \left(q_1 e^{-q_1 t}\tilde{S}(t) - q_2 e^{-q_2 t}\right)\mathbf{1}(\mathcal{A}(t)) \\
&\qquad - \tilde{\lambda}_1 \mathbf{1}(\mathcal{A}(t)) \int_{-\infty}^{b(t,\tilde{S}(t),v(t))} \left[\tilde{V}^A\left(t,\tilde{S}(t)e^{y},v(t)\right) - \left(e^{-q_1 t}\tilde{S}(t)e^{y} - e^{-q_2 t}\right)\right] G_1(y)\,\mathrm{d}y \\
&\qquad - \tilde{\lambda}_2 \mathbf{1}(\mathcal{A}(t)) \int_{-b(t,\tilde{S}(t),v(t))}^{\infty} \left[\tilde{V}^A\left(t,\tilde{S}(t)e^{-y},v(t)\right) - \left(e^{-q_1 t}\tilde{S}(t)e^{-y} - e^{-q_2 t}\right)\right] G_2(y)\,\mathrm{d}y,
\end{aligned} \tag{4.40}$$

with $b(t,\tilde{S}(t),v(t)) \equiv \ln[B(t,v(t))e^{(q_1-q_2)t}/\tilde{S}(t)]$. This equation is to be solved over the domain $(t,\tilde{S}(t),v(t)) \in [0,T]\times\mathbb{R}^2_+$, subject to terminal and boundary conditions (4.28).

Proof. The inhomogeneous term arises from the observation that, for all $(t,\tilde{S}(t),v(t)) \in [0,T]\times\mathbb{R}^2_+$, the equation

$$0 = \frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(t,\tilde{S}(t),v(t))\right] - \left(\frac{\partial \tilde{V}^A}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[\tilde{V}^A(t,\tilde{S}(t),v(t))\right]\right)\mathbf{1}(\mathcal{A}(t))$$

holds. Equation (4.40) is obtained by expanding the negative term in the above equation, as was done in the proof of Proposition 4.5, taking into account the probability density functions G_1 and G_2 for the jumps. The steps of the proof are similar to those taken in the proof of Proposition 3.15. ∎

[8] Similar to what was noted in the proof of Proposition 3.15, provided $\tilde{V}^A(t,\tilde{S}(t),v(t))$ is continuously differentiable, degenerate Dirac-delta densities will not appear in the inhomogeneous term. Assumption 4.2 guarantees these differentiability conditions are met.

4.5 Limit of the Early Exercise Boundary at Maturity

It is important to know the behavior of the critical asset price ratio $B(t,v)$ associated to the early exercise boundary of the American exchange option immediately before the option matures (i.e. as $t \to T^-$) as it helps investors decide whether to exercise the exchange option right before maturity or if to hold until maturity. The behavior of the early exercise boundary near maturity can also be used to formulate an appropriate terminal condition $B(T,v)$ for the unknown early exercise boundary that can be used in conjunction with the terminal condition on the discounted American exchange option price when implementing numerical schemes to solve the inhomogeneous IPDE in the American case.

The next proposition presents the limit of the early exercise boundary, which we obtain following the method of Chiarella and Ziogas (2009).[9]

Proposition 4.7. *The limit $B(T^-,v)$ of the critical asset price ratio $B(t,v)$ at maturity (i.e. as $t \to T^-$) is a solution of the equation*

$$B(T^-,v) = \max\left\{1, \frac{q_2 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln B(T^-,v)} G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln B(T^-,v)}^{\infty} G_2(y)\,\mathrm{d}y}{q_1 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln B(T^-,v)} e^{y} G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln B(T^-,v)}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y}\right\}. \tag{4.41}$$

Proof. The method of Chiarella and Ziogas (2009), adapted to our situation, is as follows.[10] First, we set the inhomogeneous term $\Xi(t,\tilde{S},v)$ (given by equation (4.40)) to zero and evaluate the result at $t = T$ and $\tilde{S} = B(T^-,v)e^{(q_1-q_2)T}$. The resulting expression is then rearranged to yield equation (4.41).

Performing the first step yields the equation

$$\begin{aligned} 0 = {} & e^{-q_2 T}\left(q_1 B(T^-,v) - q_2\right) \\ & - \tilde{\lambda}_1 \int_{-\infty}^{-\ln\left[\frac{B(T^-,v)e^{(q_1-q_2)T}}{B(T^-,v)e^{(q_1-q_2)T}}\right]} \Big[\tilde{V}^A\left(T, B(T^-,v)e^{(q_1-q_2)T}e^{y}, v(T)\right) \\ & \qquad - e^{-q_2 T}\left(B(T^-,v)e^{y} - 1\right)\Big] G_1(y)\,\mathrm{d}y \\ & - \tilde{\lambda}_2 \int_{\ln\left[\frac{B(T^-,v)e^{(q_1-q_2)T}}{B(T^-,v)e^{(q_1-q_2)T}}\right]}^{\infty} \Big[\tilde{V}^A\left(T, B(T^-,v)e^{(q_1-q_2)T}e^{-y}, v(T)\right) \\ & \qquad - e^{-q_2 T}\left(B(T^-,v)e^{-y} - 1\right)\Big] G_2(y)\,\mathrm{d}y. \end{aligned} \tag{4.42}$$

(Note that in the above calculation, setting $\tilde{S} = B(T^-,v)e^{(q_1-q_2)T}$ induces the stopping criterion in equation (4.25), implying that $\mathbf{1}(\mathcal{A}(T)) = 1$ in the inhomogeneous term (4.40)). Terminal condition

[9] Nyström (2008) provide a formal asymptotic analysis and the requisite regularity conditions for the behavior of the early exercise boundary near maturity for general multi-asset American options where underlying asset prices are driven by a general Itô diffusion.

[10] Chiarella and Ziogas (2009) proposed this method as an alternative to the local analysis of the option PDE for small time-to-maturity options as was done by Wilmott et al. (1993) in the pure diffusion case.

(4.28) implies that

$$\tilde{V}^A\left(T, B(T^-,v)e^{(q_1-q_2)T}e^y, v(T)\right) = \max\left\{0, e^{-q_2T}\left(B(T^-,v)e^y - 1\right)\right\},$$
$$\tilde{V}^A\left(T, B(T^-,v)e^{(q_1-q_2)T}e^{-y}, v(T)\right) = \max\left\{0, e^{-q_2T}\left(B(T^-,v)e^{-y} - 1\right)\right\}.$$

Thus, the first integral in equation (4.42) will be zero if $B(T^-,v)e^y - 1 \geq 0$ or if $y \geq -\ln B(T^-,v)$. Likewise, the second integral will be zero if $B(T^-,v)e^{-y} - 1 \geq 0$ or if $y \leq \ln B(T^-,v)$. (In other words, the integral terms will vanish if the maximum functions yield the nonzero alternative.) Following this analysis, equation (4.42) simplifies to

$$\begin{aligned} 0 = q_1 B(T^-,v) - q_2 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln B(T^-,v)} \left[B(T^-,v)e^y - 1\right] G_1(y)\,\mathrm{d}y \\ + \tilde{\lambda}_2 \int_{\ln B(T^-,v)}^{\infty} \left[B(T^-,v)e^{-y} - 1\right] G_2(y)\,\mathrm{d}y. \end{aligned}$$

Rearranging the terms yields the equation

$$B(T^-,v) = \frac{q_2 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln B(T^-,v)} G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln B(T^-,v)}^{\infty} G_2(y)\,\mathrm{d}y}{q_1 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln B(T^-,v)} e^y G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln B(T^-,v)}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y}.$$

We note lastly from Broadie and Detemple (1997) that $B(t,v) \geq 1$ for any $t \in [0,T]$ and $v \in (0,\infty)$. Therefore, we must enforce a lower bound of 1 on $B(T^-,v)$ via the maximum function. The result stated in the proposition thus holds. ■

Equation (4.41) must be solved implicitly for $B(T^-,v)$, which can be done using any standard root-finding technique. In equation (4.41), we see that the limit is dependent on the asset dividend yields q_1 and q_2, the jump intensities $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ under $\hat{\mathbb{Q}}$, and the jump size density functions G_1 and G_2 under $\hat{\mathbb{Q}}$. These dependencies highlight the influence of jumps in asset prices on the limiting behavior of the early exercise boundary.[11] In the succeeding discussions, we shall investigate the more specific effects of these parameters on the limit of the early exercise boundary. We note further that equation (4.41) does not depend on the instantaneous variance v since the option payoff is independent of v (however, equation (4.41) is true for all $v \in (0,\infty)$).

In the absence of jumps (i.e. when $\tilde{\lambda}_1 = \tilde{\lambda}_2 = 0$), the limit reduces to $\max\{1, q_2/q_1\}$. This is consistent with the result of Broadie and Detemple (1997) for American exchange options under a pure diffusion specification. In the pure diffusion case it is known that $B(T^-,v) = q_2/q_1 > 1$ if $q_2 > q_1$, implying that the early exercise boundary many not be continuous in t at maturity. When jumps are present, the analysis of continuity becomes more complicated, as shall be seen below.

[11] This is in contrast to the proposition of Carr and Hirsa (2003), in their analysis of the one-asset American put option where the log-price is driven by a Lévy process, that the limit of the early exercise boundary is only dependent on the dividend yield and the risk-free rate.

We present below some conditions under which equation (4.41) has a solution.

Proposition 4.8. *Suppose $q_1, q_2 \geq 0$ and $\tilde{\lambda}_1, \tilde{\lambda}_2 > 0$ are given and let G_1 and G_2 be continuous probability density functions. The equation*

$$x = \frac{q_2 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln x} G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln x}^{\infty} G_2(y)\,\mathrm{d}y}{q_1 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln x} e^y G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln x}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y} \tag{4.43}$$

has a unique solution $x^* \in (0, \infty)$ if and only if $q_1 > 0$. Furthermore, $x^* > 1$ if and only if

$$q_2 - q_1 + \tilde{\lambda}_1 \int_{-\infty}^{0} (1 - e^y) G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{0}^{\infty} (1 - e^{-y}) G_2(y)\,\mathrm{d}y > 0.$$

Proof. Our proof adapts the arguments made by Chiarella et al. (2015, pp. 34-35). For $x \in (0, \infty)$, define the function

$$\begin{aligned} f(x) = q_2 &+ \tilde{\lambda}_1 \int_{-\infty}^{-\ln x} G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln x}^{\infty} G_2(y)\,\mathrm{d}y \\ &- x \left(q_1 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln x} e^y G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln x}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y \right). \end{aligned}$$

Denote by x^* a zero of f (i.e. $f(x^*) = 0$) on $(0, \infty)$, if any exist.

Differentiating f with respect to x yields

$$\begin{aligned} f'(x) &= \tilde{\lambda}_1 G_1(-\ln x) \left(-\frac{1}{x}\right) - \tilde{\lambda}_2 G_2(\ln x) \left(\frac{1}{x}\right) \\ &\quad - x \left(\tilde{\lambda}_1 e^{-\ln x} G_1(-\ln x)(-\frac{1}{x}) - \tilde{\lambda}_2 e^{-\ln x} G_2(\ln x)(\frac{1}{x}) \right) \\ &\quad - \left(q_1 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln x} e^y G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln x}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y \right) \\ &= - \left(q_1 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln x} e^y G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln x}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y \right). \end{aligned}$$

Note that the integrals that appear above are nonnegative. Hence, $f'(x) < 0$ if and only if $q_1 > 0$. That is, f is strictly decreasing on $(0, \infty)$ if and only if $q_1 > 0$. We also observe that

$$\lim_{x \to 0^+} f(x) = q_2 + \tilde{\lambda}_1 \int_{-\infty}^{\infty} G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{-\infty}^{\infty} G_2(y)\,\mathrm{d}y = q_2 + \tilde{\lambda}_1 + \tilde{\lambda}_2 > 0.$$

Furthermore, we note that for a fixed $x > 0$, $0 < x e^y G_1(y) \leq G_1(y)$ for all $y \leq -\ln x$. Thus,

$$0 < x \int_{-\infty}^{-\ln x} e^y G_1(y)\,\mathrm{d}y \leq \int_{-\infty}^{-\ln x} G_1(y)\,\mathrm{d}y \to 0 \quad \text{as } x \to \infty.$$

A similar argument can be used to show that

$$0 < x \int_{\ln x}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y \leq \int_{\ln x}^{\infty} G_2(y)\,\mathrm{d}y \to 0 \quad \text{as } x \to \infty.$$

As such, we find that

$$\begin{aligned}\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \Bigg(& q_2 - q_1 x + \tilde{\lambda}_1 \left[\int_{-\infty}^{-\ln x} G_1(y)\,\mathrm{d}y - x \int_{-\infty}^{-\ln x} e^{y} G_1(y)\,\mathrm{d}y \right] \\ & + \tilde{\lambda}_2 \left[\int_{\ln x}^{\infty} G_2(y)\,\mathrm{d}y - x \int_{\ln x}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y \right] \Bigg) \\ = \lim_{x\to\infty} & (q_2 - q_1 x).\end{aligned}$$

Thus, if $q_1 > 0$, then $\lim_{x\to\infty} f(x) < 0$ and so f strictly decreases from positive to negative values as x increases on $(0, \infty)$. Therefore, there exists a unique $x^* \in (0, \infty)$ such that $f(x^*) = 0$.

Now suppose $q_1 > 0$. Evaluating f at $x = 1$ gives us

$$f(1) = q_2 - q_1 + \tilde{\lambda}_1 \int_{-\infty}^{0} (1 - e^{y}) G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{0}^{\infty} (1 - e^{-y}) G_2(y)\,\mathrm{d}y.$$

If $f(1) \leq 0$, then x^* must be in the interval $(0, 1]$ since f is strictly decreasing. Otherwise, $x^* > 1$ if and only if $f(1) > 0$, which is the condition stated in the proposition. Lastly, by Proposition 4.7, we have $B(T^-, v) = \max\{1, x^*\}$. ∎

If $q_1 = 0$ and $q_2 > 0$, then $f'(x) \leq 0$ for all $x \in (0, \infty)$ and $\lim_{x\to\infty} f(x) = q_2 > 0$. That is, f is non-increasing and remains positive as x increases in $(0, \infty)$. Thus, f does not have any zeros on $(0, \infty)$. If a solution x^* of equation (4.43) exists, then the limit of the early exercise boundary is $B(T^-, v) = \max\{1, x^*\}$.

An immediate result from Proposition 4.8 is a condition for the continuity of $B(t, v)$ at maturity. We present this in the following corollary.

Corollary 4.9. *Suppose $q_1 > 0$. For any fixed $v \in (0, \infty)$, $B(t, v)$ is continuous at maturity $t = T$ if*

$$q_1 \geq q_2 + \tilde{\lambda}_1 \int_{-\infty}^{0} (1 - e^{y}) G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{0}^{\infty} (1 - e^{-y}) G_2(y)\,\mathrm{d}y. \tag{4.44}$$

PROOF. Suppose $q_1 > 0$ and condition (4.44) holds. Then from the discussion at the end of the proof of Proposition 4.8, the solution x^* to equation (4.43) lies in the interval $(0, 1]$. It follows that $B(T^-, v) = 1$, which is also the value of $B(T, v)$. Thus, $B(t, v)$ is continuous at the option's maturity. ∎

We briefly discuss the behavior of $B(T^-, v)$ in response to changes in q_1. We note that $\partial f / \partial q_1 = -x < 0$, so when q_1 decreases, $f(x)$ increases. In particular, for a given $q_1 > 0$, there exists

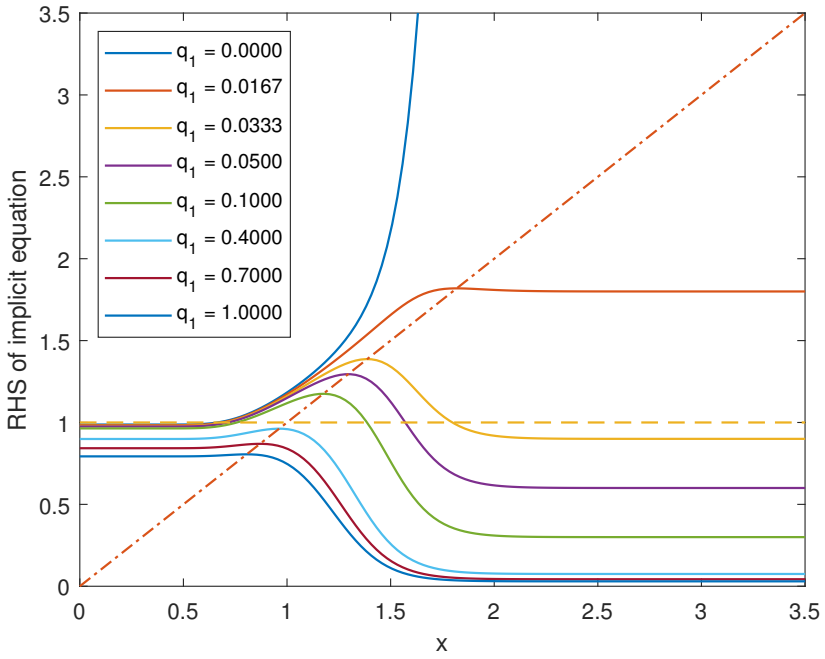
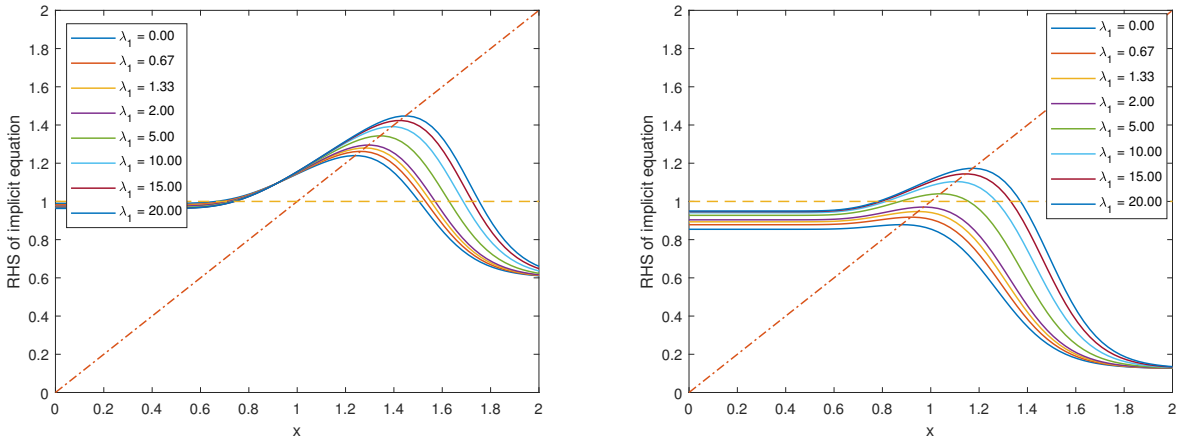


Figure 4.3: Behavior of equation (4.43) with respect to q_1 when jumps are normally distributed. Solid lines represent the right-hand side of (4.43) and their intersection with the dash-dotted 45° line represents the solution x^* of (4.43). The horizontal dashed line indicates the position x^* relative to unity.

$x^* \in (0, \infty)$ such that $f(x^*) = 0$. If q_1 decreases, then $f(x^*)$ increases away from zero, thereby moving the unique zero of f to some other number $x' \in (0, \infty)$ such that $x' > x^*$. In other words, the unique solution x^* of equation (4.43) increases without bound, and consequently $B(T^-, v) \to \infty$, as $q_1 \to 0^+$. *Thus, when the first asset bears no dividend yield, it is not optimal to exercise the American exchange option early or at least immediately prior to the option's maturity.*

We give a numerical illustration of the aforementioned properties in the case of when the jump size random variables Y_1 and Y_2 have normal distributions $N(\alpha_{j_1}, \beta_{j_1}^2)$ and $N(\alpha_{j_2}, \beta_{j_2}^2)$, respectively, under $\hat{\mathbb{Q}}$. The integrals in equation (4.43) can then be expressed in terms of the standard normal distribution. Figure 4.3 shows the behavior of the right-hand side of equation (4.43), which we denote by $R(x)$, for increasing values of q_1 given this specific distribution of jump sizes and fixed parameters $q_2 = 0.03$, $\tilde{\lambda}_1 = \tilde{\lambda}_2 = 2$, $\alpha_{j_1} = \alpha_{j_2} = 0.02$, and $\beta_{j_1} = \beta_{j_2} = 0.2$. We observe that, except for when $q_1 = 0$, the behavior of $R(x)$ is generally the same for any value q_1. That is, it increases up until some value of x, then eventually decreases to q_2/q_1. The graph of $R(x)$ eventually crosses the 45° line, indicating the solution x^* of equation (4.43). However, if $q_1 = 0$, then the graph of $R(x)$ never crosses the line, implying that there is no solution exists for (4.43) and that the equation will only be satisfied by taking the limit as $x \to 0$, giving an infinite early exercise boundary $B(T^-, v)$. We also note that as q_1 increases further away from q_2, x^* eventually falls below unity. One can also formulate the difference $q_1 - q_2$ in (4.44) in terms of the standard normal distribution

A comparative static analysis of $B(T^-, v)$ with respect to the jump intensities $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ can



(a) $q_1 = 0.05$, $q_2 = 0.03$ (b) $q_1 = 0.40$, $q_2 = 0.05$

Figure 4.4: Behavior of equation (4.43) with respect to $\tilde{\lambda}_1$ when jumps are normally distributed for two configurations of the dividend yields. The dashed and dash-dotted lines function in a way similar to that for Figure 4.3.

also be conducted. We note that

$$\frac{\partial f(x)}{\partial \tilde{\lambda}_1} = \int_{-\infty}^{-\ln x} G_1(y)\,\mathrm{d}y - x\int_{-\infty}^{-\ln x} e^y G_1(y)\,\mathrm{d}y \geq 0,$$
$$\frac{\partial f(x)}{\partial \tilde{\lambda}_2} = \int_{\ln x}^{\infty} G_2(y)\,\mathrm{d}y - x\int_{\ln x}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y \geq 0,$$

as was established in the proof of Proposition 4.8. With the same arguments as above, these inequalities imply that $B(T^-, v)$ are non-decreasing with $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$. This is graphically illustrated by Figure 4.4, where we once again plot the right-hand side of equation (4.43) for various values of $\tilde{\lambda}_1$. In this example, we exhibit the behavior of the boundary for $q_1 = 0.05$ and $q_2 = 0.03$ and for $q_1 = 0.40$ and $q_2 = 0.05$ (the second set of dividend yields were chosen such that x^* below and above unity are illustrated), with the value of all other jump parameters chosen similarly to those for Figure 4.3. A similar illustration can also be generated for a comparative static analysis with respect to $\tilde{\lambda}_2$.

The response of the early exercise boundary limit at maturity with respect to the jump intensities is indeed expected. As discussed in the previous section, there is a nonzero probability that asset prices may jump back into the continuation region immediately after exercise, resulting to losses for the option holder. The option holder is thus more conservative in opting for early exercise and is willing to exercise early only when the asset yield ratios are higher compared to the no-jump case. This is so that even if the asset yield jumps downward, it is less likely to move all the way back in to the continuation region thereby reducing risk on the investor's end.

4.6 The Transition Density Function

To determine the price of the American exchange option using equation (4.29), we need to solve for the price of the European exchange option $\tilde{V}$ and evaluate the early exercise premium $\tilde{V}^P$ in equation (4.33). To do so, we need to solve for the joint transition density function of the asset yield ratio process $\tilde{S}$ and the variance process v under $\hat{\mathbb{Q}}$. The solution presented in this chapter uses an integral transform approach similar to that used in Chapter 3. However, we perform these calculations directly on the Kolmogorov backward equation corresponding to the joint transition density function of $\tilde{S}$ and v. ty With the joint transition density function, we may evaluate the expectation in equation (4.16) for the European exchange option price (instead of solving IPDE (4.18)) and the expectations in (4.33) for the early exercise premium. For the remainder of this chapter, we assume that $G_1(y)$ and $G_2(y)$ are the $\hat{\mathbb{Q}}$-probability density functions of the jump size variables Y_1 and Y_2, respectively.

Let $Q(T, \tilde{s}_T, v_T; t, \tilde{s}, v)$ denote the joint transition density function of $(\tilde{S}, v)$ under the probability measure $\hat{\mathbb{Q}}$. That is, for fixed $0 \leq t < T$, $Q(T, \tilde{s}_T, v_T; t, \tilde{s}, v)$ is the nonnegative $\mathcal{B}(\mathbb{R}^2_+) \otimes \mathcal{B}(\mathbb{R}^2_+)$-measurable function of $(\tilde{s}_T, v_T, \tilde{s}, v) \in \mathbb{R}^2_+ \times \mathbb{R}^2_+$ such that

$$\hat{\mathbb{Q}}[(\tilde{S}(T), v(T)) \in B | \tilde{S}(t) = \tilde{s}, v(t) = v] = \int_B Q(T, \tilde{s}_T, v_T; t, \tilde{s}, v)\, \mathrm{d}\tilde{s}_T\, \mathrm{d}v_T \tag{4.45}$$

for any $B \in \mathcal{B}(\mathbb{R}^2_+)$ (see Appendix A.5 for additional details). Following calculations similar to those in the proof of Lemma A.13, the Kolmogorov backward equation associated to Q is given by

$$\frac{\partial Q}{\partial t} + \mathcal{L}_{\tilde{s},v}\left[Q(T, \tilde{s}_T, v_T; t, \tilde{s}, v)\right] = 0, \tag{4.46}$$

which is to be solved for $t \in [0, T]$ and $(\tilde{s}, v) \in \mathbb{R}^2_+$ subject to the terminal condition

$$Q(T, \tilde{s}_T, v_T; T, \tilde{s}, v) = \delta(\tilde{s} - \tilde{s}_T)\delta(v - v_T), \tag{4.47}$$

where $\delta(\cdot)$ is the Dirac-delta function Since T, $\tilde{s}_T$, and v_T are fixed in (4.46), we will denote $Q(T, \tilde{s}_T, v_T; t, \tilde{s}, v)$ by $Q(t, \tilde{s}, v)$ in the succeeding calculations for notational brevity.

To solve equation (4.46), we introduce a change of variables by letting $x = \ln \tilde{s}$ and defining the function H by

$$H(T, x_T, v_T; t, x, v) = Q(T, e^{x_T}, v_t; t, e^x, v). \tag{4.48}$$

In terms of partial derivatives of H, the partial derivatives of Q involved in equation (4.46) are given by

$$\frac{\partial H}{\partial t} = \frac{\partial Q}{\partial t}, \qquad \frac{\partial H}{\partial v} = \frac{\partial Q}{\partial v}, \qquad \frac{\partial^2 H}{\partial v^2} = \frac{\partial^2 Q}{\partial v^2}$$

$$\frac{\partial H}{\partial x} = \tilde{s}\frac{\partial Q}{\partial \tilde{s}}, \qquad \frac{\partial^2 H}{\partial x^2} - \frac{\partial H}{\partial x} = \tilde{s}^2\frac{\partial^2 Q}{\partial \tilde{s}^2}, \qquad \frac{\partial^2 H}{\partial x \partial v} = \tilde{s}\frac{\partial^2 Q}{\partial \tilde{s}\partial v}.$$

We also note that

$$Q(t, \tilde{s}e^{Y_1}, v) = H(t, x + Y_1, v) \quad \text{and} \quad Q(t, \tilde{s}e^{-Y_2}, v) = H(t, x - Y_2, v).$$

Thus, when written in terms of H and its partial derivatives, equation (4.46) becomes

$$\begin{aligned} 0 = \frac{\partial H}{\partial t} &- \left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \frac{1}{2}\sigma^2 v\right)\frac{\partial H}{\partial x} + \left[\xi\eta - (\xi + \Lambda)v\right]\frac{\partial H}{\partial v} \\ &+ \frac{1}{2}\sigma^2 v\frac{\partial^2 H}{\partial x^2} + \frac{1}{2}\omega^2 v\frac{\partial^2 H}{\partial v^2} + \omega\left(\sigma_1\rho_{wz_1} - \sigma_2\rho_{wz_2}\right)v\frac{\partial^2 H}{\partial x \partial v} \\ &+ \tilde{\lambda}_1\mathbb{E}_{\hat{\mathbb{Q}}}^{Y_1}\left[H(t, x + Y_1, v) - H(t, x, v)\right] \\ &+ \tilde{\lambda}_2\mathbb{E}_{\hat{\mathbb{Q}}}^{Y_2}\left[H(t, x - Y_2, v) - H(t, x, v)\right]. \end{aligned} \tag{4.49}$$

From (4.47), the associated terminal condition for H is

$$H(T, x_T, v_T; T, x, v) = \delta(x - x_T)\delta(v - v_T). \tag{4.50}$$

Since the transformation $x = \ln \tilde{s}$ is one-to-one, the function $H(T, x_T, v_T; t, x, v)$ may be interpreted as the joint transition density function of the process $(X(t), v(t))$ under $\hat{\mathbb{Q}}$, i.e. for fixed $0 \le t < T$, $H(T, x_T, v_T; t, x, v)$ is the nonnegative $\mathcal{B}(\mathbb{R} \times \mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R} \times \mathbb{R}_+)$-measurable function of (x_T, v_T, x, v) such that

$$\hat{\mathbb{Q}}[(X(T), v(T)) \in B | X(t) = x, v(t) = v] = \int_B H(T, x_T, v_T; t, x, v)\,\mathrm{d}x_T\,\mathrm{d}v_T, \qquad B \in \mathcal{B}(\mathbb{R} \times \mathbb{R}_+). \tag{4.51}$$

Moving forward, we shall denote $H(T, x_T, v_T; T, x, v)$ by $H(t, x, v)$ to emphasize that we are solving equation (4.49) in t, x, and v and that x_T and v_T are given terminal values of these variables.

Since the coefficients of equation (4.49) do not contain x, we can take its Fourier transform with respect to x to further simplify the equation. Since H is a transition density function, it is absolutely integrable with respect to x. Therefore its Fourier transform exists and satisfies the second-order PDE stated in the following proposition.

Proposition 4.10. *Let $\hat{H}(t, \phi, v)$ denote the Fourier transform of $H(t, x, v)$ with respect to x,*

$$\hat{H}(t, \phi, v) = \mathscr{F}_x\left\{H(t, x, v)\right\}(\phi) = \int_{-\infty}^{\infty} e^{i\phi x}H(t, x, v)\,\mathrm{d}x. \tag{4.52}$$

Then $\hat{H}$ is the solution of the PDE

$$0 = \frac{\partial \hat{H}}{\partial t} + \left(\frac{1}{2}\varepsilon v - i\phi\Psi\right)\hat{H} + (\alpha - \Theta v)\frac{\partial \hat{H}}{\partial v} + \frac{1}{2}\omega^2 v \frac{\partial^2 \hat{H}}{\partial v^2}, \tag{4.53}$$

where

$$\begin{aligned} \alpha &\equiv \xi\eta, \\ \Theta = \Theta(\phi) &\equiv \xi + \Lambda + i\phi\omega\left(\sigma_1\rho_{wz_1} - \sigma_2\rho_{wz_2}\right), \\ \varepsilon = \varepsilon(\phi) &\equiv \sigma^2\left(i\phi - \phi^2\right), \\ \Psi = \Psi(\phi) &\equiv -\tilde{\lambda}_1\tilde{\kappa}_1 - \tilde{\lambda}_2\tilde{\kappa}_2^- - \frac{\tilde{\lambda}_1}{i\phi}\left(\varphi_1(\phi) - 1\right) - \frac{\tilde{\lambda}_2}{i\phi}\left(\varphi_2(-\phi) - 1\right), \end{aligned} \tag{4.54}$$

and $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ are the characteristic functions of Y_1 and Y_2, respectively, under $\hat{\mathbb{Q}}$.[12] *The associated terminal condition is*

$$\hat{H}(T, \phi, v) = e^{i\phi x_T}\delta(v - v_T). \tag{4.55}$$

PROOF. A straightforward application of equation (4.52) yields

$$\mathscr{F}_x\left\{\frac{\partial H}{\partial t}\right\} = \frac{\partial \hat{H}}{\partial t}, \qquad \mathscr{F}_x\left\{\frac{\partial H}{\partial v}\right\} = \frac{\partial \hat{H}}{\partial v}, \qquad \mathscr{F}_x\left\{\frac{\partial^2 H}{\partial v^2}\right\} = \frac{\partial^2 \hat{H}}{\partial v^2}.$$

By the properties of the Fourier transform (see for example, Davies, 1985, Section 9.5), we obtain

$$\mathscr{F}_x\left\{\frac{\partial H}{\partial x}\right\} = -i\phi\hat{H}, \qquad \mathscr{F}_x\left\{\frac{\partial^2 H}{\partial x^2}\right\} = -\phi^2\hat{H}, \qquad \mathscr{F}_x\left\{\frac{\partial^2 H}{\partial x \partial v}\right\} = -i\phi\frac{\partial \hat{H}}{\partial v}.$$

The Fourier transform of the first expectation in equation (4.49) is the Fourier transform of a convolution-type integral and is calculated as follows:

$$\begin{aligned} \mathscr{F}_x\left\{\mathbb{E}_{\hat{\mathbb{Q}}}^{Y_1}\left[H(t, x + Y_1, v)\right]\right\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\phi x} H(t, x + y, v) G_1(y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} e^{-i\phi y} G_1(y)\left(\int_{-\infty}^{\infty} e^{i\phi z} H(t, z, v)\,\mathrm{d}z\right)\mathrm{d}y \\ &= \left(\int_{-\infty}^{\infty} e^{-i\phi y} G_1(y)\,\mathrm{d}y\right)\hat{H}(t, \phi, v) \\ &= \varphi_1(\phi)\hat{H}(t, \phi, v). \end{aligned}$$

[12] If $G_1(y)$ and $G_2(y)$ are the probability density functions of Y_1 and Y_2, respectively, under $\hat{\mathbb{Q}}$, then

$$\varphi_1(\phi) = \int_{-\infty}^{\infty} e^{i\phi y} G_1(y)\,\mathrm{d}y \qquad \text{and} \qquad \varphi_2(-\phi) = \int_{-\infty}^{\infty} e^{i\phi y} G_2(y)\,\mathrm{d}y.$$

A similar calculation shows that

$$\mathscr{F}_x\left\{\mathbb{E}^{Y_2}_{\hat{\mathbb{Q}}}\left[H(t,x-Y_2,v)\right]\right\} = \varphi_2(-\phi)\hat{H}(t,\phi,v).$$

Thus, the Fourier transform of equation (4.49) is

$$\begin{aligned}0 = \frac{\partial \hat{H}}{\partial t} &+ i\phi\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \frac{1}{2}\sigma^2 v\right)\hat{H} + \left[\xi\eta - (\xi+\Lambda)v\right]\frac{\partial \hat{H}}{\partial v}\\ &- \frac{1}{2}\sigma^2 v\phi^2\hat{H} + \frac{1}{2}\omega^2 v\frac{\partial^2\hat{H}}{\partial v^2} - \omega\left(\sigma_1\rho_{wz_1} - \sigma_2\rho_{wz_2}\right)vi\phi\frac{\partial\hat{H}}{\partial v}\\ &+ \tilde{\lambda}_1\left(\varphi_1(\phi)-1\right)\hat{H} + \tilde{\lambda}_2\left(\varphi_2(-\phi)-1\right)\hat{H}.\end{aligned}$$

Factorization yields

$$\begin{aligned}0 = \frac{\partial\hat{H}}{\partial t} &+ \hat{H}\left[\frac{1}{2}\sigma^2(i\phi-\phi^2)v - i\phi\left(-\tilde{\lambda}_1\tilde{\kappa}_1 - \tilde{\kappa}_2\tilde{\kappa}_2^- - \frac{\tilde{\lambda}_1}{i\phi}\left(\varphi_1(\phi)-1\right) - \frac{\tilde{\lambda}_2}{i\phi}\left(\varphi_2(-\phi)-1\right)\right)\right]\\ &+ \frac{\partial\hat{H}}{\partial v}\left[\xi\eta - v\left(\xi+\lambda+i\phi\omega(\sigma_1\rho_{wz_1}-\sigma_2\rho_{wz_2})\right)\right] + \frac{1}{2}\omega^2 v\frac{\partial^2\hat{H}}{\partial v^2}.\end{aligned}$$

Using the notation introduced in equation (4.54) yields equation (4.53).

The associated terminal condition is obtained by a straightforward application of equation (4.52) and the definition of the Dirac-delta function,

$$\hat{H}(T,\phi,v) = \mathscr{F}_x\left\{H(T,x,v)\right\} = \int_{-\infty}^{\infty} e^{i\phi x}\delta(x-x_T)\delta(v-v_T)\,\mathrm{d}x = e^{i\phi x_T}\delta(v-v_T).$$

■

Save for some minor notational differences, the simplified PDE (4.53) that we obtained is identical to the Fourier-transformed PDE presented in Cheang et al. (2013, Proposition 4.1) for the transition density function of a single-asset stochastic volatility jump-diffusion model. This resemblance is expected since the underlying asset yield ratio process $\tilde{S}(t)$ is modelled similarly with a Heston-type stochastic volatility process but with two jump components (one for each stock price process in consideration). The notational difference is pronounced in the definition of $\Psi(\phi)$ in equation (4.54) where we have two terms corresponding to the jump size variables Y_1 and Y_2, in contrast to the representation of Cheang et al. (2013) who only have one jump term. Due to these similarities, we will follow the method of Cheang et al. (2013) in the succeeding calculations to determine the solution of equation (4.53).[13]

At this point, we have reduced IPDE (4.49) to the PDE (4.53) which is first-order in t and

[13] Note also that our calculations moving forward are similar to those presented in Sections 3.4 and 3.5, although in those sections, the Fourier and Laplace transforms were applied on the IPDE for the European exchange option price, not the transition density function of the underlying state variables.

second-order in v. Solutions of second-order PDEs such as equation (4.53) have been obtained by Feller (1951) using a Laplace transform with respect to v.[14] Following Chiarella et al. (2010) and Cheang et al. (2013), we further simplify equation (4.53) by taking its Laplace transform with respect to v. First, additional technical assumptions must be imposed on $\hat{H}(t, \phi, v)$ to ensure that all required Laplace transforms are well-defined.

Assumption 4.11. Assume that $e^{-\vartheta v}\hat{H}(t, \phi, v) \to 0$ and $e^{-\vartheta v}\partial\hat{H}/\partial v \to 0$ as $v \to +\infty$.

As before, these technical conditions can be reasonably assumed since H is a transition density function. In particular, the Fourier transform $\hat{H}$ defined in equation (4.52) and its partial derivative in v both decay to zero since $H(t, x, v)$ decays to zero as $v \to \infty$. Assumption 4.11 also requires that the growth of $\hat{H}$ and $\partial\hat{H}/\partial v$ is dominated by the growth of the exponential term $e^{\vartheta v}$ as $v \to \infty$ for any $\vartheta > 0$. Again, these growth restrictions can be reasonably assumed. Taking Assumption 4.11 to be true, we now reduce equation (4.53) to a first-order PDE in the following proposition.

Proposition 4.12. *Let $\bar{H}(t, \phi, \vartheta)$ be the Laplace transform of $\hat{H}(t, \phi, v)$ with respect to v,*

$$\bar{H}(t, \phi, \vartheta) = \mathscr{L}_v\left\{\hat{H}(t, \phi, v)\right\}(\vartheta) = \int_0^\infty e^{-\vartheta v}\hat{H}(t, \phi, v)\,\mathrm{d}v. \tag{4.56}$$

Then $\bar{H}$ satisfies the equation

$$-\frac{\partial\bar{H}}{\partial t} + \left[\frac{1}{2}\omega^2\vartheta^2 - \Theta\vartheta + \frac{1}{2}\varepsilon\right]\frac{\partial\bar{H}}{\partial\vartheta} = \left[(\alpha - \omega^2) + \Theta - i\phi\Psi\right]\bar{H} + f(t), \tag{4.57}$$

where $f(t) \equiv (\omega^2/2 - \alpha)\hat{H}(t, \phi, 0)$ must be determined such that

$$\lim_{\vartheta\to\infty}\bar{H}(t, \phi, \vartheta) = 0. \tag{4.58}$$

Equation (4.57) has terminal condition

$$\bar{H}(T, \phi, \vartheta) = \exp\left\{i\phi x_T - \vartheta v_T\right\}. \tag{4.59}$$

Proof. Given that Assumption 4.11 holds, we can use the definition of the Laplace transform (4.56) and integration by parts to determine the Laplace transform of each term in equation (4.53). Using the properties of the Laplace transform (see for example, Debnath, 2016, Theorem 3.4.3), we obtain

$$\mathscr{L}_v\left\{\frac{\partial\hat{H}}{\partial t}\right\} = \frac{\partial\bar{H}}{\partial t},$$

[14] Feller (1951) solved, using this method, the Fokker-Planck equation associated to a particular diffusion problem. The equation in his analysis is more general than equation (4.53).

$$\mathscr{L}_v\left\{\left(\frac{1}{2}\varepsilon v - i\phi\Psi\right)\hat{H}\right\} = -\frac{1}{2}\varepsilon\frac{\partial\bar{H}}{\partial\vartheta} - i\phi\Psi\bar{H},$$
$$\mathscr{L}_v\left\{(\alpha - \Theta v)\frac{\partial\hat{H}}{\partial v}\right\} = -\alpha\hat{H}(t,\phi,0) + (\alpha\vartheta + \Theta)\bar{H} + \Theta\vartheta\frac{\partial\bar{H}}{\partial\vartheta},$$
$$\mathscr{L}_v\left\{\frac{1}{2}\omega^2 v\frac{\partial^2\hat{H}}{\partial v^2}\right\} = \frac{1}{2}\omega^2\hat{H}(t,\phi,0) - \frac{1}{2}\omega^2\vartheta^2\frac{\partial\bar{H}}{\partial\vartheta} - \omega^2\vartheta\bar{H}.$$

The Laplace transform of equation (4.53) is therefore

$$0 = \frac{\partial\bar{H}}{\partial t} - \frac{1}{2}\varepsilon\frac{\partial\bar{H}}{\partial\vartheta} - i\phi\Psi\bar{H} - \alpha\hat{H}(t,\phi,0) + (\alpha\vartheta + \Theta)\bar{H} + \Theta\vartheta\frac{\partial\bar{H}}{\partial\vartheta}$$
$$+ \frac{1}{2}\omega^2\hat{H}(t,\phi,0) - \frac{1}{2}\omega^2\vartheta^2\frac{\partial\bar{H}}{\partial\vartheta} - \omega^2\vartheta\bar{H}.$$

Factoring and defining $f(t) \equiv (\omega^2/2 - \alpha)\hat{H}(t,\phi,0)$ yields equation (4.57).

Note that, at this point, $\hat{H}(t,\phi,0)$ is unknown but it has to be determined so that the solution $\bar{H}$ to equation (4.57) is finite for all $\vartheta > 0$. A sufficient condition for this is equation (4.58) (Chiarella et al., 2010; Cheang et al., 2013).

The terminal condition (4.59) is a consequence of the definition (4.56) and the Dirac-delta function,

$$\bar{H}(T,\phi,\vartheta) = \mathscr{L}_v\left\{\hat{H}(T,\phi,v)\right\} = \int_0^\infty e^{-\vartheta v}e^{i\phi x_T}\delta(v - v_T)\,\mathrm{d}v = \exp\left\{i\phi x_T - \vartheta v_T\right\}.$$

■

The Laplace-transformed PDE (4.57) is an inhomogeneous first-order equation which can be solved via the method of characteristics and the method of variation of parameters (see for example, Zauderer, 2006; Duffy, 2011). This procedure will yield a solution $\bar{H}$ containing the unknown function $f(t)$. Condition (4.58) will then be applied to determine $f(t)$ and characterize $\bar{H}$ completely in terms of the parameters introduced in Propositions 4.10 and 4.12. In the next proposition, we present the solution of equation (4.57).

Proposition 4.13. *The solution $\bar{H}(t,\phi,\vartheta)$ of equation (4.57) is given by*

$$\bar{H}(t,\phi,\vartheta) = \exp\left\{\left[\frac{(\alpha-\omega^2)(\Theta-\digamma)}{\omega^2} + \Theta - i\phi\Psi\right](T-t)\right\}$$
$$\times\left[\frac{2\digamma}{(\omega^2\vartheta - \Theta + \digamma)(e^{\digamma(T-t)} - 1) + 2\digamma}\right]^{2-\frac{2\alpha}{\omega^2}}\exp\left\{i\phi x_T - \left(\frac{\Theta-\digamma}{\omega^2}\right)v_T\right\} \quad (4.60)$$
$$\times\exp\left\{\frac{-2\digamma v_T(\omega^2\vartheta - \Theta + \digamma)e^{\digamma(T-t)}}{\omega^2\left[(\omega^2\vartheta - \Theta + \digamma)(e^{\digamma(T-t)} - 1) + 2\digamma\right]}\right\}$$
$$\times\Gamma\left(\frac{2\alpha}{\omega^2} - 1; \beta(\phi,\vartheta;v_T)\right),$$

Equation (4.57)	**Chiarella et al. (2010)**	**Cheang et al. (2013)**
ω^2	σ^2	σ^2
ε	Λ	Λ
ϑ	s	s
$-i\phi\Psi$	$-i\phi(r-q)$	$-i\phi\Psi$
$T-t$	τ	$T-t$

Table 4.1: Comparison of notation and coefficients of equation (4.57), Chiarella et al. (2010, Equation 29) and Cheang et al. (2013, Equation 46).

where

$$\begin{aligned} \digamma &= \digamma(\phi) \equiv \sqrt{\Theta^2(\phi) - \omega^2\varepsilon(\phi)}, \\ \beta(\phi, \vartheta; v_T) &\equiv \frac{2\digamma v_T e^{\digamma(T-t)}}{\omega^2(e^{\digamma(T-t)}-1)} \times \frac{2\digamma}{(\omega^2\vartheta - \Theta + \digamma)(e^{\digamma(T-t)}-1) + 2\digamma}, \end{aligned} \tag{4.61}$$

$\Gamma(u;\beta)$ *is the (lower) incomplete gamma function* $\Gamma(u;\beta) = \frac{1}{\Gamma(u)}\int_0^\beta e^{-x}x^{u-1}\,\mathrm{d}x$, *and* $\Gamma(u)$ *is the gamma function* $\Gamma(u) = \int_0^\infty e^{-x}x^{u-1}\,\mathrm{d}x$.

Proof. Details of a similar calculation can also be found in the proof of Proposition 3.7 of this thesis.[15] Since our Laplace-transformed PDE is identical to those presented in Chiarella et al. (2010) and Cheang et al. (2013), see also Chiarella et al. (2010, Appendix 4) for the derivation of equation (4.60) for $\hat{H}$.[16] For easy comparison between our analysis and the proofs presented in Chiarella et al. (2010) and Cheang et al. (2013), we show in Table 4.1 the equivalence of notations used in the PDE. ∎

At this point, we now recover the original transition density function $Q(t, \tilde{s}, v)$ by inverting the Laplace and Fourier transforms on $\bar{H}(t, \phi, \vartheta)$ given in equation (4.60). To this end, we first solve for the inverse Laplace transform of $\bar{H}$.

Proposition 4.14. *The inverse Laplace transform of* $\bar{H}(t, \phi, \vartheta)$ *in equation* (4.60) *is*

$$\begin{aligned} \hat{H}(t, \phi, v) = \exp&\left\{\frac{(\Theta - \digamma)}{\omega^2}(v - v_T + \alpha(T-t))\right\}\exp\left\{i\phi x_T - i\phi\Psi(T-t)\right\} \\ &\times \frac{2\digamma e^{\digamma(T-t)}}{\omega^2(e^{\digamma(T-t)}-1)}\left[\frac{v_T e^{\digamma(T-t)}}{v}\right]^{\frac{\alpha}{\omega^2}-\frac{1}{2}}\exp\left\{-\frac{2\digamma(v_T e^{\digamma(T-t)} + v)}{\omega^2(e^{\digamma(T-t)}-1)}\right\} \\ &\times I_{\frac{2\alpha}{\omega^2}-1}\left(\frac{4\digamma\sqrt{v_T v e^{\digamma(T-t)}}}{\omega^2(e^{\digamma(T-t)}-1)}\right), \end{aligned} \tag{4.62}$$

[15] Proposition 3.7 solves the Laplace-transformed PDE corresponding to the European exchange option price, similar to the analysis of Chiarella et al. (2015) who consider the Laplace-transformed PDE corresponding to the European call option price. In contrast, the current proposition provides the solution of the Laplace-transformed PDE for the transition density function. However, the outline of the proof and the main techniques used do not vary drastically.

[16] Note however that Chiarella et al. (2010) uses time-to-maturity $\tau \equiv T - t$ instead of calendar time t, as what was done in our analysis and by Cheang et al. (2013).

where $I_k(u)$ is the modified Bessel function of the first kind

$$I_k(u) = \sum_{n=0}^{\infty} \frac{(u/2)^{2n+k}}{n!\Gamma(n+k+1)}. \tag{4.63}$$

PROOF. Details of a similar calculation can be found in the proof of Proposition 3.8 of this thesis. See also Chiarella et al. (2010, Appendix 5), keeping in mind the notational equivalence established in Table 4.1. ■

Having determined $\hat{H}(t,\phi,v)$, we now perform an inverse Fourier transform to recover the original function $H(t,x,v)$ using the inversion formula

$$H(t,x,v) = \mathscr{F}_x^{-1}\left\{\hat{H}(t,\phi,v)\right\}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\phi x}\hat{H}(t,\phi,v)\,\mathrm{d}\phi \tag{4.64}$$

that corresponds to the Fourier transform defined by equation (4.52). The original transition density function $Q(t,\tilde{s},v)$ can then be determined by reversing the substitution $x = \ln\tilde{s}$ that we made in equation (4.48). The result is presented in the proposition below.

Proposition 4.15. *The transition density function $Q(t,\tilde{s},v) \equiv Q(T,\tilde{s}_T,v_T;t,\tilde{s},v)$ defined by equation (4.45) is given by*

$$\begin{aligned} Q(t,\tilde{s},v) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} &\frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!} \\ &\times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left\{-i\phi\ln\left(\frac{\tilde{s}}{\tilde{s}_T}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\right)\right\}\right. \\ &\left.\times\, h(T-t,\phi,v;v_T)\,\mathrm{d}\phi\right], \end{aligned} \tag{4.65}$$

where

$$\begin{aligned} h(\tau,\phi,v;v_T) = \exp&\left\{\frac{(\Theta-\digamma)}{\omega^2}(v-v_T+\alpha\tau)\right\}\frac{2\digamma e^{\digamma\tau}}{\omega^2(e^{\digamma\tau}-1)}\left[\frac{v_T e^{\digamma\tau}}{v}\right]^{\frac{\alpha}{\omega^2}-\frac{1}{2}} \\ &\times\exp\left\{-\frac{2\digamma(v_T e^{\digamma\tau}+v)}{\omega^2(e^{\digamma\tau}-1)}\right\}\times I_{\frac{2\alpha}{\omega^2}-1}\left(\frac{4\digamma\sqrt{v_T v e^{\digamma\tau}}}{\omega^2(e^{\digamma\tau}-1)}\right). \end{aligned} \tag{4.66}$$

Here, $\Upsilon_{1,m}$ and $\Upsilon_{1,n}$ are given by

$$\Upsilon_{1,m} = \sum_{k=1}^{m} Y_{1,k} \qquad \text{and} \qquad \Upsilon_{2,n} = \sum_{l=1}^{n} Y_{2,l},$$

where $\{Y_{1,1},\ldots,Y_{1,m}\}$ and $\{Y_{2,1},\ldots,Y_{2,n}\}$ are collections of i.i.d. random variables sampled from populations with $\hat{Q}$-density functions $G_1(y)$ and $G_2(y)$, respectively, of the jump-size random variables Y_1 and Y_2 (c.f. Proposition 4.10), and $\mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}[\cdot]$ is the expectation operator with respect to

$\Upsilon_{1,m}$ *and* $\Upsilon_{2,n}$ *only.*

PROOF. See Appendix C.1. ■

4.7 The European Exchange Option Price

With knowledge of the transition density function $Q(T, \tilde{s}_T, v_T; t, \tilde{s}, v)$, we can now compute the price of any European-style option on the two assets which matures at time T and whose payoff, when discounted by the second asset yield process, can be written as a function $\tilde{F}(T, \tilde{s}_T, v_T)$ of the terminal asset yield ratio $\tilde{S}(T) = \tilde{s}_T$ and the terminal instantaneous variance $v(T) = v_T$. Following equation (4.15) and the results stated in Proposition 4.3, the no-arbitrage price at time $t \in [0, T)$ of such a claim is given by

$$\begin{aligned} P(t, \tilde{s}, v) &= S_2(t) e^{q_2 t} \mathbb{E}_{\hat{\mathbb{Q}}} \left[\tilde{F}(T, \tilde{S}(T), v(T)) \middle| \tilde{S}(t) = \tilde{s}, v(t) = v \right] \\ &= S_2(t) e^{q_2 t} \int_0^\infty \int_0^\infty \tilde{F}(T, \tilde{s}_T, v_T) Q(T, \tilde{s}_T, v_T; t, \tilde{s}, v) \, \mathrm{d}v_T \, \mathrm{d}x_T. \end{aligned} \tag{4.67}$$

In this section, we calculate the price of the European exchange option following the valuation formula stated above.

For our calculations for the European exchange option price, it is more convenient to use the log-price variable $x_T = \ln \tilde{s}_T$ and the corresponding transition density function $H(T, x_T, v_T; t, \tilde{s}_T, v_T)$. This implies that the option price $\tilde{V}(t, \tilde{s}, v)$, as given by equation (4.16), can be written as[17]

$$\tilde{V}(t, \tilde{s}, v) = e^{-q_1 T} \int_{-\infty}^\infty \int_0^\infty \left(e^{x_T} - e^{(q_1 - q_2)T} \right)^+ H(T, x_T, v_T; t, \ln \tilde{s}, v) \, \mathrm{d}v_T \, \mathrm{d}x_T. \tag{4.68}$$

Since the payoff is independent of the terminal variance level v_T, we may evaluate the innermost integral by evaluating $\int_0^\infty H(T, x_T, v_T; t, x, v) \, \mathrm{d}v_T$. We present this integral in the following lemma.

Lemma 4.16. *The integral* $\int_0^\infty H(T, x_T, v_T; t, \ln \tilde{s}, v) \, \mathrm{d}v_T$ *is given by*

$$\begin{aligned} &\int_0^\infty H(T, x_T, v_T; t, \ln \tilde{s}, v) \, \mathrm{d}v_T \\ &\quad = \sum_{m=0}^\infty \sum_{n=0}^\infty \frac{(\tilde{\lambda}_1 (T-t))^m (\tilde{\lambda}_2 (T-t))^n e^{-(\tilde{\lambda}_1 + \tilde{\lambda}_2)(T-t)}}{m! n!} \\ &\qquad \times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)} \left[\frac{1}{2\pi} \int_{-\infty}^\infty e^{i\phi x_T} f\left(T - t, \tilde{s} e^{-(\tilde{\lambda}_1 \tilde{\kappa}_1 + \tilde{\lambda}_2 \tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}}, v; -\phi \right) \mathrm{d}\phi \right], \end{aligned} \tag{4.69}$$

[17] Note that hereafter, we make the simplifying notation $\tilde{S}(t) = \tilde{s}$ and $v(t) = v$ to be consistent with our notation for the transition density function.

where

$$
\begin{aligned}
f(\tau, z, v; \phi) &= \exp\left\{ i\phi \ln z + B(\tau, -\phi) + D(\tau, -\phi)v \right\}, \\
B(\tau, \phi) &= \frac{\alpha}{\omega^2} \left\{ (\Theta + \digamma)\tau - 2\ln\left(\frac{1 - \chi e^{\digamma \tau}}{1 - \chi} \right) \right\}, \\
D(\tau, \phi) &= \frac{\Theta + \digamma}{\omega^2} \left(\frac{1 - e^{\digamma \tau}}{1 - \chi e^{\digamma \tau}} \right),
\end{aligned}
\tag{4.70}
$$

with $\chi = (\Theta + \digamma)/(\Theta - \digamma)$, α *and* Θ *as defined in Proposition 4.10, and* $\digamma$ *as defined in Proposition 4.13.*

PROOF. See Appendix C.2. ■

Having evaluated the integral of the transition density function with respect to the terminal value of the variance process, we are now ready to solve for the European exchange option price. Our results are shown in the next proposition.

Proposition 4.17. *The discounted price* $\tilde{V}(t, \tilde{s}, v)$ *of the European exchange option at time* t *is given by*

$$
\begin{aligned}
&\tilde{V}(t, \tilde{s}, v) \\
&\quad = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(\tilde{\lambda}_1(T-t))^m (\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1 + \tilde{\lambda}_2)(T-t)}}{m! n!} \\
&\qquad \times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)} \Big[e^{-q_1 T} \tilde{s} e^{-(\tilde{\lambda}_1 \tilde{\kappa}_1 + \tilde{\lambda}_2 \tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}} \\
&\qquad\quad \times P_1^E \left(T - t, \tilde{s} e^{-(\tilde{\lambda}_1 \tilde{\kappa}_1 + \tilde{\lambda}_2 \tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}}, v; (q_1 - q_2)T \right) \\
&\qquad\quad - e^{-q_2 T} P_2^E \left(T - t, \tilde{s} e^{-(\tilde{\lambda}_1 \tilde{\kappa}_1 + \tilde{\lambda}_2 \tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}}, v; (q_1 - q_2)T \right) \Big]
\end{aligned}
\tag{4.71}
$$

where P_1^E *and* P_2^E *are defined as*

$$
\begin{aligned}
P_1^E(\tau, z, v; K) &= \frac{1}{2} + \frac{1}{2\pi} \int_0^{\infty} \frac{f_1(\tau, z, v; \phi) e^{-i\phi K} - f_1(\tau, z, v; -\phi) e^{i\phi K}}{i\phi} \, \mathrm{d}\phi, \\
P_2^E(\tau, z, v; K) &= \frac{1}{2} + \frac{1}{2\pi} \int_0^{\infty} \frac{f(\tau, z, v; \phi) e^{-i\phi K} - f(\tau, z, v; -\phi) e^{i\phi K}}{i\phi} \, \mathrm{d}\phi,
\end{aligned}
\tag{4.72}
$$

with $f(\tau, z, v; \phi)$ *given in equation (4.70),* $f_1(\tau, z, v; \phi)$ *given by*

$$
\begin{aligned}
f_1(\tau, z, v; \phi) &= \exp\left\{ i\phi \ln z + B_1(\tau, -\phi) + D_1(\tau, -\phi) \right\}, \\
B_1(\tau, \phi) &= \frac{\alpha}{\omega^2} \left\{ (\Theta_1 + \digamma_1)\tau - 2\ln\left[\frac{1 - \chi_1 e^{\digamma_1 \tau}}{1 - \chi_1} \right] \right\}, \\
D_1(\tau, \phi) &= \frac{\Theta_1 + \digamma_1}{\omega^2} \left[\frac{1 - e^{\digamma_1 \tau}}{1 - \chi_1 e^{\digamma_1 \tau}} \right],
\end{aligned}
\tag{4.73}
$$

$\Theta_1(\phi) \equiv \Theta(\phi - i)$, $\digamma_1(\phi) \equiv \digamma(\phi - i)$, *and* $\chi_1(\phi) = \chi(\phi - i)$.

PROOF. See Appendix C.3. ■

Equation (4.71) expresses the discounted European exchange option price as a sum of Poisson-weighted expectations of Heston-Bates-type formulas which arise when asset price dynamics have stochastic volatility. Our results are similar to those of Cheang et al. (2013), except that in the present analysis, there are two jump components considered. Equation (4.71) may also be read as an expected value of these Heston-Bates-type expressions conditional on the number of jumps in the prices of the two assets (denoted by m and n) observed over the remaining life of the option. Given the number of jumps m and n, the Heston-Bates expressions are then averaged with respect to the accumulated jump sizes $\Upsilon_{1,m}$ and $\Upsilon_{2,n}$ in the prices of stocks 1 and 2, respectively. Our representation for the discounted European exchange option price depends on the current asset yield ratio $\tilde{s}$, the dividend yields of the two assets, the jump parameters under $\hat{\mathbb{Q}}$, the accumulation of jumps $\Upsilon_{1,m}$ and $\Upsilon_{2,n}$, and the instantaneous variance level v. Notably absent is the risk-free rate r, which is consistent with the original result of Margrabe (1978). The expectation $\mathbb{E}^{(m,n)}_{\hat{\mathbb{Q}}}$ may be evaluated further by specifying the jump-size distribution G_1 and G_2 (for example, a log-normal distribution as in Merton (1976) or an asymmetric double exponential distribution as in Kou (2002)), but we leave this unspecified in this analysis.

The time t price of the European exchange option price $C(t, S_1, S_2, v)$ may be obtained by multiplying $\tilde{V}(t, \tilde{s}, v)$ by $S_2 e^{q_2 t}$. Doing so, we find that

$$\begin{aligned}
&C(t, S_1, S_2, v) \\
&\quad = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(\tilde{\lambda}_1(T-t))^m (\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!} \\
&\qquad \times \mathbb{E}^{(m,n)}_{\hat{\mathbb{Q}}} \Big[S_1 e^{-q_1(T-t)} e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}} \\
&\qquad\quad \times P_1^E \left(T-t, \tilde{s} e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}}, v; (q_1 - q_2)T \right) \\
&\qquad\quad - S_2 e^{-q_2(T-t)} P_2^E \left(T-t, \tilde{s} e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}}, v; (q_1 - q_2)T \right) \Big].
\end{aligned} \tag{4.74}$$

This result resembles our analysis in Proposition 3.14. The main difference is the dependence of P_1^E and P_2^E on the asset yield ratio $\tilde{s}$, in contrast to the separate dependence of P_1 and P_2 on the asset prices S_1 and S_2 in Proposition 3.14. This difference in functional dependence is offset by the multiplication of jump-size parameters pertaining to the second asset, namely $e^{-\tilde{\lambda}_2\tilde{\kappa}_2^-(T-t)}$ and $e^{-\Upsilon_{2,n}}$, to the first term rather than the second term. The overall Poisson-weighted infinite sum structure is nonetheless preserved in our current analysis.

The similarity of our result to the original Margrabe (1978) result and to the Cheang and

Chiarella (2011) extension to the jump-diffusion case can be further emphasized by writing equation (4.74) as

$$C(t,S_1,S_2,v) = S_1 e^{-q_1(T-t)}\hat{Q}_1 - S_2 e^{-q_2(T-t)}\hat{Q}_2, \tag{4.75}$$

where

$$\begin{aligned}
\hat{Q}_1 = &\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!}\\
&\times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\Bigg[e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\\
&\qquad\times P_1^E\left(T-t,\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}},v;(q_1-q_2)T\right)\Bigg],\\
\hat{Q}_2 = &\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!}\\
&\times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\left[P_2^E\left(T-t,\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}},v;(q_1-q_2)T\right)\right].
\end{aligned} \tag{4.76}$$

This alternative representation is also consistent with the form obtained in our analysis in Section 2.5.1 where we used the change-of-numéraire technique. Viewed from this perspective, $\hat{Q}_1$ and $\hat{Q}_2$ may be interpreted as probabilities that the exchange option is in-the-money at maturity, taken under probability measures $\mathbb{Q}_1$ and $\mathbb{Q}_2$ equivalent to the risk-neutral probability measure $\mathbb{Q}$ obtained when using the money-market account as the numéraire. The alternative measures $\mathbb{Q}_1$ and $\mathbb{Q}_2$ are obtained by taking the asset price processes $S_1(t)$ and $S_2(t)$, respectively, as the numéraire process.[18]

4.8 The Early Exercise Premium

Knowledge of the transition density function will also allow us to evaluate the expectations occurring in the early exercise premium $\tilde{V}^P(t,\tilde{s},v)$ given in equation (4.33). To this end, the starting point of our calculations is equation (4.38) which expresses the early exercise premium in terms of the jump size density functions. Furthermore, for any $u \in [t,T]$, let $\tilde{S}(u) = \tilde{s}_u$, $v(u) = v_u$, and $x_u = \ln \tilde{s}_u$. The following proposition provides an integral representation for the early exercise premium.

Proposition 4.18. *The early exercise premium $\tilde{V}^P(t,\tilde{s},v)$ for the American exchange option is given by*

$$\tilde{V}^P(t,\tilde{s},v) = \tilde{V}_D^P(t,\tilde{s},v) - \tilde{\lambda}_1\tilde{V}_{J_1}^P(t,\tilde{s},v) - \tilde{\lambda}_2\tilde{V}_{J_2}^P(t,\tilde{s},v). \tag{4.77}$$

[18] Note that the probability measure $\hat{\mathbb{Q}}$ used in this analysis is different from $\mathbb{Q}_2$ as the former corresponds to the asset *yield* process $S_2(t)e^{q_2 t}$ as the numéraire asset. However, we note that $\hat{\mathbb{Q}}$, $\mathbb{Q}_1$, and $\mathbb{Q}_2$ are equivalent probability measures as they are all equivalent to the objective market measure $\mathbb{P}$. In this chapter, we chose to perform our calculations under $\hat{\mathbb{Q}}$ as it reduces the exchange option pricing problem to the pricing of an ordinary call option on the asset yield process $\tilde{S}(t)$.

Here, $\tilde{V}_D^P$ is given by

$$\begin{aligned}
&\tilde{V}_D^P(t,\tilde{s},v) \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{\tilde{\lambda}_1^m \tilde{\lambda}_2^n}{m!n!} \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)} \Bigg[\int_0^T \int_0^{\infty} (u-t)^{m+n} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(u-t)} \\
&\quad \times \Bigg(q_1 e^{-q_1 u} \tilde{s} e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)} e^{\Upsilon_{1,m}-\Upsilon_{2,n}} \\
&\qquad \times P_1^A\left(u-t, \tilde{s} e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)} e^{\Upsilon_{1,m}-\Upsilon_{2,n}}, v; v_u, A(u,v_u)\right) \\
&\quad - q_2 e^{-q_2 u} \\
&\qquad \times P_2^A\left(u-t, \tilde{s} e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)} e^{\Upsilon_{1,m}-\Upsilon_{2,n}}, v; v_u, A(u,v_u)\right) \Bigg) \, \mathrm{d}v_u \, \mathrm{d}u \Bigg],
\end{aligned} \tag{4.78}$$

where $A(u,v_u) \equiv B(u,v_u)e^{(q_1-q_2)u}$ is the critical asset yield ratio associated to the stopping region $\mathcal{S}(u)$ (cf. equation (4.25)), and P_1^A and P_2^A are defined as

$$\begin{aligned}
P_1^A(\tau,z,v;v_u,K) &= \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty} \frac{f_{21}(\tau,z,v;\phi,v_u)e^{-i\phi\ln K} - f_{21}(\tau,z,v;-\phi,v_u)e^{i\phi\ln K}}{i\phi}\,\mathrm{d}\phi, \\
P_2^A(\tau,z,v;v_u,K) &= \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty} \frac{f_2(\tau,z,v;\phi,v_u)e^{-i\phi\ln K} - f_2(\tau,z,v;-\phi,v_u)e^{i\phi\ln K}}{i\phi}\,\mathrm{d}\phi,
\end{aligned} \tag{4.79}$$

with $f_2(\tau,z,v;\phi,v_u) = e^{i\phi\ln z}h(\tau,-\phi,v;v_u)$, f_{21} given by

$$f_{21}(\tau,z,v;\phi,v_u) = e^{i\phi\ln z}h(\tau,-(\phi-i),v;v_u),$$

and $h(\tau,\phi,v;v_u)$ given by equation (4.66). Furthermore, $\tilde{V}_{J_1}^P(t,\tilde{s},v)$ and $\tilde{V}_{J_1}^P(t,\tilde{s},v)$ are given by

$$\begin{aligned}
&\tilde{V}_{J_1}^P(t,\tilde{s},v) \\
&\quad = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{\tilde{\lambda}_1^m \tilde{\lambda}_2^n}{m!n!} \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)} \Bigg[\frac{1}{2\pi} \int_0^T \int_0^{\infty} (u-t)^{m+n} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(u-t)} \\
&\quad \times \int_{-\infty}^{0} G_1(y) \int_{\ln A(u,v_u)-y}^{\ln A(u,v_u)-y} \left(V^A(u, e^{x_u+y}, v_u) - \left(e^{-q_1 u} e^{x_u+y} - e^{-q_2 u} \right) \right) \\
&\quad \times \int_{-\infty}^{\infty} e^{i\phi x_u} \exp\left\{ -i\phi \ln\left(\tilde{s} e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)} e^{\Upsilon_{1,m}-\Upsilon_{2,n}} \right) \right\} \\
&\quad \times h(u-t,\phi,v;v_u)\,\mathrm{d}\phi\,\mathrm{d}x_u\,\mathrm{d}y\,\mathrm{d}v_u\,\mathrm{d}u \Bigg],
\end{aligned} \tag{4.80}$$

and

$$\begin{aligned}
&\tilde{V}_{J_2}^P(t,\tilde{s},v) \\
&\quad= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\tilde{\lambda}_1^m\tilde{\lambda}_2^n}{m!n!}\mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\Bigg[\frac{1}{2\pi}\int_0^T\int_0^{\infty}(u-t)^{m+n}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(u-t)} \\
&\quad\times\int_0^{\infty}G_2(y)\int_{\ln A(u,v_u)}^{\ln A(u,v_u)+y}\left(V^A(u,e^{x_u-y},v_u)-\left(e^{-q_1u}e^{x_u-y}-e^{-q_2u}\right)\right) \\
&\quad\times\int_{-\infty}^{\infty}e^{i\phi x_u}\exp\left\{-i\phi\ln\left(\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\right)\right\} \\
&\quad\times h(u-t,\phi,v;v_u)\,\mathrm{d}\phi\,\mathrm{d}x_u\,\mathrm{d}y\,\mathrm{d}v_u\,\mathrm{d}u\Bigg].
\end{aligned} \tag{4.81}$$

PROOF. See Appendix C.4. ∎

Because of the possibility of jumps in the prices of the two underlying assets, the early exercise premium, as seen from its jump components, remains dependent on the (yet unknown) discounted American exchange option price $\tilde{V}^A(t,\tilde{s},v)$. The early exercise premium components also require knowledge of the unknown critical asset yield ratio $B(t,v)$ for the term $A(t,v)=B(t,v)e^{(q_1-q_2)t}$ that appears in equations (4.78), (4.80), and (4.81). In this regard, we require another equation that, when coupled with equation (4.29), characterizes both $\tilde{V}^A$ and $B(t,v)$.

Proposition 4.19. *The discounted American exchange option $\tilde{V}^A(t,\tilde{s},v)$ and the critical asset yield ratio $B(t,v)$ are the solution of the linked system of integral equations*

$$\begin{aligned}
\tilde{V}^A(t,\tilde{s},v) &= \tilde{V}(t,\tilde{s},v)+\tilde{V}^P(t,\tilde{s},v), \\
e^{-q_1t}\left(A(t,v)-e^{(q_1-q_2)t}\right) &= \tilde{V}(t,A(t,v),v)+\tilde{V}^P(t,A(t,v),v),
\end{aligned} \tag{4.82}$$

where $A(t,v)=B(t,v)e^{(q_1-q_2)t}$, $\tilde{V}(t,\tilde{s},v)$ is the price of the European exchange option given in Proposition 4.17, and $\tilde{V}^P(t,\tilde{s},v)$ is the early exercise premium given in Proposition 4.18.

PROOF. The first equation is a restatement of the main result in Proposition 4.4, whereas the second equation is obtained by evaluating the first equation at $\tilde{s}=A(t,v)$ and invoking the value-matching condition (4.26) to rewrite the left-hand side. The linked system is supplemented with expressions for $\tilde{V}(t,\tilde{s},v)$ and $\tilde{V}^P(t,\tilde{s},v)$ given by Propositions 4.17 and 4.18, respectively. ∎

The discounted American exchange option price $\tilde{V}^A(t,\tilde{s},v)$, once determined, is measured in units of the second asset yield process. To obtain the nominal price, $\tilde{V}^A(t,\tilde{s},v)$ is simply multipled by $S_2(t)e^{-q_2t}$.

4.9 Chapter Summary and Conclusion

We considered the problem of pricing European and American exchange options when the underlying asset prices are modelled using the proportional SVJD model discussed in Chapter 3. Our results extend those obtained by Margrabe (1978) and Bjerskund and Stensland (1993) for European and American exchange options under pure diffusion dynamics and those by Cheang and Chiarella (2011) who analyzed exchange options under jump-diffusion dynamics.

The results of this chapter complement those presented in Chapter 2. In the SVJD model in Chapter 2, the assumption that asset prices were uncorrelated (among other simplifying assumptions on the correlation structure of the model) had to be enforced to obtain analytical representation of exchange option prices. In this chapter, we used a proportional stochastic volatility jump-diffusion model to include the possibility of correlation between asset returns in our representation of exchange option prices. We showed that the representations we obtained in this analysis are of similar form to those obtained in Chapter 2, as well as those obtained in earlier studies by Margrabe (1978), Fischer (1978), and Cheang and Chiarella (2011).

Given the proportional SVJD model specification, the two-dimensional exchange option pricing problem was reduced to a one-dimensional call option pricing problem by taking the second asset yield process as the numéraire, following the put-call transformation method described by Bjerskund and Stensland (1993). We defined the equivalent martingale measure $\hat{\mathbb{Q}}$, corresponding to our choice of numéraire, which was then used to express the exchange option prices as expectations under $\hat{\mathbb{Q}}$. Then, with usual martingale arguments, we derived the IPDE for the European option price and the Kolmogorov backward equation for the joint transition density function of $\tilde{S}(t)$ and $v(t)$.

Having reduced the problem to pricing a call option on the asset yield ratio $\tilde{S}(t)$, which also has stochastic volatility and jump-diffusion dynamics, we then obtained an early exercise representation for the American exchange option via probabilistic arguments and to solve for the joint transition density function via Fourier and Laplace transforms. Equipped with the joint transition density function, we then obtained integral representations for the price of the European exchange option and the early exercise premium and a linked system of integral equations for the American exchange option price and the unknown early exercise boundary associated to the American option. We also analyzed the limiting behavior of the early exercise boundary immediately before maturity and how the dividend yields and the jump parameters affected this limit.

The methods presented in this chapter can also be applied to pricing options whose payoff is homogeneous in the asset prices and is independent of the variance leve. Furthermore, with the simplification of the exchange option pricing problem, our results can then be linked to a number of existing numerical and simulation techniques for single-asset options (see for example, Adolfsson et al., 2013; Chiarella et al., 2009) to obtain option prices given the model parameters. In particular,

numerical algorithms to solve the simplified IPDE arising from the put-call transformation are discussed in the next chapter.

CHAPTER 5

A Method of Lines (MOL) Approach

5.1 Introduction

Our analysis in this chapter starts with the exchange option pricing IPDE derived in the previous chapter (Proposition 4.3), for which we now provide a numerical solution using the method of lines (MOL).

The method of lines is a numerical method to solve PDEs which consists of discretizing the equation in all but one variable resulting to a sequence or system of ODEs in the remaining continuous variable. Schiesser and Griffiths (2009) provide a general exposition on the method, but for applications in option and fixed income instrument pricing, the time-discrete MOL, expertly discussed by Meyer (2015), has gained particular traction.[1] This type of MOL has been applied to pricing American put options in the Black-Scholes framework (Meyer and van der Hoek, 1997), put options under jump-diffusion dynamics (Meyer, 1998), call options under stochastic volatility (Adolfsson et al., 2013; Chiarella and Ziveyi, 2013), call options under Bates SVJD (Chiarella et al., 2009), American options with SV and stochastic interest rates (Kang and Meyer, 2014), American options under a regime-switching GBM (Chiarella et al., 2016), American options with jumps in both the asset price and the volatility (Taruvinga, 2019), and spread options under pure-diffusion dynamics (Chiarella and Ziveyi, 2014). It is particularly useful for approximating American option prices as the algorithm can be easily adjusted to accommodate unknown free boundaries. It is also attractive for financial applications as the option delta and gamma are calculated as part of the algorithm with no additional computational cost.

However as with any numerical technique for solving PDEs, the MOL becomes highly complex

[1] The time-discrete MOL involves discretizing the PDE in all but one *spatial variable*, as opposed to most applications where time is left as the continuous variable (see for example, Schiesser and Griffiths, 2009). This approach is also known as Rothe's method or the horizontal MOL.

the more spatial variables are involved. In this chapter, we take advantage of the put-call transformation technique used in the previous chapter in an effort to simplify the MOL approximation of the exchange option price. To this end, this chapter makes the following contributions:

1. We give a detailed discussion of a method of lines (MOL) scheme to numerically determine the price of exchange options, expressed in units of the second asset yield process. We also present a detailed structure (in the form of a pseudo-code) of the implementation of the MOL to the exchange option pricing IPDE resulting from our proportional SVJD model specification (see Algorithms 1 and 2 in Section 5.2).

2. As alternatives to the often assumed boundary condition $\lim_{v\to\infty} \frac{\partial V}{\partial v} = 0$ for the behavior of option prices at high volatility levels (Clarke and Parrott, 1999; Chiarella et al., 2009), we consider *Venttsel boundary conditions* for the far-but-finite limits of the computational domain and their impact on option prices and the performance of the MOL algorithm. As discussed in later sections, these Venttsel boundary conditions also encode specific assumptions about the option price behavior at high levels of volatility.

3. Using MOL-generated prices, we investigate how stochastic volatility and jumps in the asset prices affect option prices. Furthermore, we conduct an extensive numerical comparative static analysis to see how key model parameters affect the exchange option prices and the early exercise boundary. In addition, we illustrate how the presence of jumps and stochastic volatility affects the American exchange option holders' decision to exercise early, supplementing our numerical investigation with an analytical approach to understanding the behavior of the early exercise boundary near maturity. This analysis extends that of Chiarella and Ziogas (2009) who studied the early exercise boundary of American call options under a jump-diffusion model. Furthermore, our numerical experiments aim to understand the interplay of the dynamics of the two asset prices, in the presence of stochastic volatility and jumps, and how this relationship impacts the exercise value of the exchange options.

4. To assess the performance and accuracy of the MOL, we compare our results to simulations of American exchange option prices obtained using the Least Squares Monte Carlo (LSMC) method proposed by Longstaff and Schwartz (2001) and the numerical solution of the corresponding linear complementarity problem via the projected successive overrelaxation (PSOR) algorithm.

As such, this chapter serves as a numerical complement to the analytical work featured in the previous chapter. For convenience, some of the notation from the previous chapter have been simplified here.

The analysis in this chapter starts *after* the put-call transformation has been implemented (refer to Section 4.2 for details). That is, we shall work under the equivalent martingale measure $\hat{\mathbb{Q}}$ in which the second asset yield ratio is the numéraire and under which the asset yield ratio process $\tilde{S}(t) = (S_1(t)e^{q_1 t})/(S_2(t)e^{q_2 t})$ and the instantaneous variance process $v(t)$ satisfy the stochastic differential equations

$$\begin{aligned} \mathrm{d}\tilde{S}(t) &= -\tilde{S}(t)\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)\mathrm{d}t + \sigma\sqrt{v(t)}\tilde{S}(t)\,\mathrm{d}\bar{W}(t) \\ &\quad + \int_{\mathbb{R}} (e^y - 1)\,\tilde{S}(t)p_1(\mathrm{d}y, \mathrm{d}t) + \int_{\mathbb{R}}\left(e^{-y} - 1\right)\tilde{S}(t)p_2(\mathrm{d}y, \mathrm{d}t), \end{aligned} \tag{5.1}$$

$$\mathrm{d}v(t) = \left[\xi\eta - (\xi + \Lambda)v(t)\right]\mathrm{d}t + \omega\sqrt{v(t)}\,\mathrm{d}\bar{Z}(t). \tag{5.2}$$

Here, $\bar{W}$ and $\bar{Z}$ are $\hat{\mathbb{Q}}$-Wiener processes with

$$\mathbb{E}_{\hat{\mathbb{Q}}}[\mathrm{d}\bar{W}(t)\,\mathrm{d}\bar{Z}(t)] = \frac{\sigma_1\rho_1 - \sigma_2\rho_2}{\sigma}, \qquad \sigma^2 \equiv \sigma_1^2 + \sigma_2^2 - 2\rho_w\sigma_1\sigma_2,$$

$p_j(\mathrm{d}y, \mathrm{d}t)$ is a Poisson random measure with $\hat{\mathbb{Q}}$-local characteristics $(\tilde{\lambda}_j, m_{\hat{\mathbb{Q}}}^{(j)}(\mathrm{d}y))$, $j = 1, 2$, $\tilde{\kappa}_1 \equiv \mathbb{E}_{\hat{\mathbb{Q}}}[e^{Y_1} - 1]$, and $\tilde{\kappa}_2^- \equiv \mathbb{E}_{\hat{\mathbb{Q}}}[e^{-Y_2} - 1]$, and Y_j, $j = 1, 2$, is a random variable with density function $m_{\hat{\mathbb{Q}}}^{(j)}(\mathrm{d}y)$ modelling the jumps in the price of asset j. Model parameters are assumed to satisfy Assumption 3.1, with further details provided in Section 4.2.

We also recall the associated IPDEs for the European and American exchange option prices. As established in the previous chapter, the *discounted price* of the European exchange option $V = V(t, \tilde{S}, v)$, where $\tilde{S} = \tilde{S}(t)$ and $v = v(t)$, satisfies the IPDE

$$\begin{aligned} 0 = \frac{\partial V}{\partial t} &+ \frac{1}{2}\sigma^2 v\tilde{S}^2\frac{\partial^2 V}{\partial \tilde{S}^2} + \frac{1}{2}\omega^2 v\frac{\partial^2 V}{\partial v^2} + \omega(\sigma_1\rho_1 - \sigma_2\rho_2)\frac{\partial^2 V}{\partial \tilde{S}\partial v} \\ &- \tilde{S}\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)\frac{\partial V}{\partial \tilde{S}} + [\xi\eta - (\xi + \Lambda)v]\frac{\partial V}{\partial v} \\ &+ \tilde{\lambda}_1\mathbb{E}_{\hat{\mathbb{Q}}}^{Y_1}\left[V(t, \tilde{S}e^{Y_1}, v) - V(t, \tilde{S}, v)\right] + \tilde{\lambda}_2\mathbb{E}_{\hat{\mathbb{Q}}}^{Y_2}\left[V(t, \tilde{S}e^{-Y_2}, v) - V(t, \tilde{S}, v)\right], \end{aligned} \tag{5.3}$$

where $(t, \tilde{S}, v) \in [0, T] \times (0, \infty) \times (0, \infty)$, subject to the terminal condition

$$V(T, \tilde{S}, v) = e^{-q_1 T}\left(\tilde{S} - e^{(q_1 - q_2)T}\right)^+.$$

The *discounted price* of the American exchange option $V^A = V^A(t, \tilde{S}, v)$, on the other hand, satisfies equation (5.3) but on the restricted domain

$$(t, \tilde{S}, v) \in \mathcal{C}, \qquad \mathcal{C} = \left\{(t, \tilde{S}, v) \in [0, T] \times (0, \infty) \times (0, \infty) : \tilde{S} < A(t, v)\right\},$$

where $A(t, v) \equiv B(t, v)e^{(q_1 - q_2)t}$, with $B(t, v) \geq 1$, for all $t \in [0, T]$, $v \in (0, \infty)$, the critical asset price ratio that determines the early exercise boundary. Outside the restricted domain $\mathcal{C}$, also

known as the continuation region, the American exchange option price is known analytically as

$$V^A(t,\tilde{S},v) = e^{-q_1 t}\left(\tilde{S} - e^{(q_1-q_2)t}\right), \qquad (t,\tilde{S},v) \in [0,T]\times[A(t,v),\infty)\times(0,\infty).$$

Since by Assumption 4.2 V^A and $\frac{\partial V^A}{\partial \tilde{S}}$ are continuous across the early exercise boundary, we have the following value-matching and smooth-pasting conditions,

$$\begin{aligned} V^A(t,A(t,v),v) &= e^{-q_1 t}\left(A(t,v) - e^{(q_1-q_2)t}\right), \\ \frac{\partial V^A}{\partial \tilde{S}}(t,A(t,v),v) &= e^{-q_1 t}. \end{aligned}$$

The associated terminal and boundary conditions are

$$\begin{aligned} V^A(T,\tilde{S},v) &= e^{-q_1 t}\left(\tilde{S} - e^{(q_1-q_2)t}\right), && \forall(\tilde{S},v)\in(0,\infty)\times(0,\infty) \\ V^A(t,0,v) &= 0, && \forall(t,v)\in[0,T]\times(0,\infty). \end{aligned}$$

The American exchange option IPDE is formulated as a *free boundary problem*, where the unknown free boundary is given by $A(t,v)$. As shown in the previous chapter, a suitable terminal condition for the free boundary is given by $A(T^-,v) = B(T^-,v)e^{(q_1-q_2)T}$, where $B(T^-,v) = \max\{1,x^*\}$, with x^* the solution of the implicit equation

$$x = \frac{q_2 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln x} G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln x}^{\infty} G_2(y)\,\mathrm{d}y}{q_1 + \tilde{\lambda}_1 \int_{-\infty}^{-\ln x} e^{y} G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\ln x}^{\infty} e^{-y} G_2(y)\,\mathrm{d}y}.$$

In the equation above, G_1 and G_2 are the (continuous) probability density functions of the jump size random variables Y_1 and Y_2.

The results recalled above serve as the starting point of the implementation of the MOL to approximate exchange option prices. The algorithm and its implementation are discussed in the next section.

5.2 Method of Lines Implementation

Here, we discuss the MOL algorithm to numerically solve the IPDEs for the discounted European and American exchange option prices. For convenience, the terminal value problem is transformed into an initial value problem involving the time to maturity $\tau = T - t$. We also denote the discounted exchange option prices by $V(\tau,s,v)$, where $s = \tilde{S}(T-\tau)$ and $v = v(T-\tau)$. We first discuss the MOL for the European exchange option and present the required adjustments for the approximation of the unknown exercise boundary for the American exchange option. We conclude this section with a discussion of alternative boundary conditions at the far boundary of v.

When solving for American option prices, it has been observed that MOL performs as competitively, if not more efficiently, compared to other methods such as componentwise splitting methods, numerical integration with the joint tdf, Monte Carlo techniques, and full finite difference schemes (Chiarella et al., 2009; Chiarella and Ziveyi, 2013; Kang and Meyer, 2014, among others). This is mainly due to the ability of the MOL to calculate the option delta and gamma to any desired accuracy with no considerable additional computational effort. To illustrate the accuracy and efficiency of the MOL, we compare the MOL American exchange option prices to prices generated by a finite difference approximation of the linear complementarity problem (LCP) solved using the projected successive overrelaxation (PSOR) method and the Longstaff and Schwartz (2001) least-squares Monte Carlo (LSMC) algorithm, a simple and popular simulation-based method for valuing early exercise options.[2]

European exchange option prices can also be efficiently approximated using Fourier inversion techniques (see for example, Dempster and Hong, 2002; Hurd and Zhou, 2010; Caldana and Fusai, 2013) since the joint characteristic function of the underlying state variables s and v is available in semi-analytical form given our affine model specification (see for example, Duffie et al., 2000). In this chapter, we present the MOL approach to pricing European exchange options as a starting point for the MOL approximation of American exchange option prices. Indeed, one of the key merits of the MOL is the relative ease with which the process for the European case can be adjusted to accommodate free boundary nature of the American case, as seen in the succeeding sections.

5.2.1 MOL Approximations of the Pricing IPDEs

With the new notation introduced in this section, the IPDE for the discounted European exchange option price can be written as

$$\begin{aligned} 0 = &\frac{1}{2}\sigma^2 v s^2 \frac{\partial^2 V}{\partial s^2} + \frac{1}{2}\omega^2 v \frac{\partial^2 V}{\partial v^2} + \omega(\sigma_1\rho_1 - \sigma_2\rho_2) v s \frac{\partial^2 V}{\partial s \partial v} \\ &- \left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right) s\frac{\partial V}{\partial s} + \left[\xi\eta - (\xi+\Lambda)v\right]\frac{\partial V}{\partial v} - (\tilde{\lambda}_1 + \tilde{\lambda}_2)V \\ &- \frac{\partial V}{\partial \tau} + \tilde{\lambda}_1 \int_{\mathbb{R}} V(\tau, se^y, v) G_1(y)\,\mathrm{d}y + \tilde{\lambda}_2 \int_{\mathbb{R}} V(\tau, se^{-y}, v) G_2(y)\,\mathrm{d}y. \end{aligned} \tag{5.4}$$

This equation is to be solved for $s \in (0,\infty)$, $v \in (0,\infty)$, and $\tau \in [0,T]$ subject to the initial condition

$$V(0,s,v) = e^{-q_1 T}\left(s - e^{(q_1-q_2)T}\right)^+. \tag{5.5}$$

[2] The LSMC algorithm only simulates the optimal exercise strategy and is unable to estimate the early exercise boundary. For a simulation-based method for approximating the early exercise frontier, see Ibáñez and Zapatero (2004). More recently, Bayer et al. (2020) proposed a new simulation-based method that estimates exercise rates of randomized exercise strategies, an optimization problem which they show is equivalent to the original optimal stopping formulation.

We also impose the following (asymptotic) boundary conditions on (5.4):

$$\lim_{s\to 0^+} V(\tau,s,v)=0, \quad \lim_{s\to\infty}\frac{\partial^2 V}{\partial s^2}=0, \quad \lim_{v\to\infty}\frac{\partial V}{\partial v}=0. \tag{5.6}$$

The first boundary condition implies that the option becomes worthless for very small values of s (i.e. when the price of the second asset dominates that of the first asset). The second boundary condition implies that the option delta becomes insensitive to changes in the asset yield ratio when it is sufficiently large. The third boundary condition means that the option price is insensitive to changes in the instantaneous variance for sufficiently large values of v.

For computational purposes, the infinite domains for s and v are truncated to $[0, s_J]$ and $[0, v_M]$, respectively, with some suitably chosen s_J and v_M. We consider the partition $0=\tau_0<\tau_1<\cdots<\tau_N=T$ of $[0,T]$ with uniform width $\Delta\tau$ and the partition $0=v_0<v_1<\cdots<v_M$ of $[0,v_M]$, also with uniform size Δv. Denote by $V_{n,m}(s)$ the approximate solution of (5.4) at the time line $\tau=\tau_n$ and variance line $v=v_m$, and let $\mathscr{V}_{n,m}(s)=V'_{n,m}(s)$ be the approximation of the option delta.

We then use finite difference quotients to discretize equation (5.4) in all variables but s. For the time derivative, a backward difference approximation is used following Meyer and van der Hoek (1997),

$$\frac{\partial V_{n,m}}{\partial\tau}\approx\begin{cases}\frac{1}{\Delta\tau}(V_{n,m}(s)-V_{n-1,m}(s))+O(\Delta\tau) & \text{if } n=1,2,\\ \frac{3}{2\Delta\tau}(V_{n,m}(s)-V_{n-1,m}(s))-\frac{1}{2\Delta\tau}(V_{n-1,m}(s)-V_{n-2,m}(s))+O((\Delta\tau)^2) & \text{if } n\geq 3.\end{cases}$$

The approximation is only first-order accurate with respect to τ for the first two time steps, but for all subsequent time steps, it is second-order accurate. Using such a scheme improves the accuracy of the method of lines algorithm. The second-order partial derivatives in v are approximated using central difference quotients

$$\begin{aligned}\frac{\partial^2 V_{n,m}}{\partial v^2}&\approx\frac{V_{n,m+1}(s)-2V_{n,m}(s)+V_{n,m-1}(s)}{(\Delta v)^2}+O((\Delta v)^2),\\ \frac{\partial^2 V_{n,m}}{\partial s\partial v}&\approx\frac{\mathscr{V}_{n,m+1}(s)-\mathscr{V}_{n,m-1}(s)}{2\Delta v}+O((\Delta v)^2).\end{aligned}$$

We use an upwinding difference approximation for $\partial V/\partial v$,

$$\begin{aligned}\big[\xi\eta-(\xi+\Lambda)v\big]\frac{\partial V_{n,m}}{\partial v}\approx{}&\max\big\{\xi\eta-(\xi+\Lambda)v_m,0\big\}\left[\frac{V_{n,m+1}(s)-V_{n,m}(s)}{\Delta v}\right]\\ &+\min\big\{\xi\eta-(\xi+\Lambda)v_m,0\big\}\left[\frac{V_{n,m}(s)-V_{n,m-1}(s)}{\Delta v}\right]+O(\Delta v).\end{aligned}$$

This approximation is only first-order accurate with respect to v, but it helps keep the finite difference approximation in v stable as the coefficients of the second-order partial derivatives vanish

as $v \to 0^+$ and it improves the performance of the iterative method used in the method of lines algorithm (Chiarella et al., 2009; Meyer, 2015).

At the last variance line $v = v_M$, we invoke the asymptotic boundary condition (5.6) in v. The boundary condition can be approximated by

$$\frac{\partial V(\tau_n, s, v_M)}{\partial v} = 0, \tag{5.7}$$

and a forward difference approximation of this implies that $V_{n,M+1}(s) = V_{n,m}(s)$. Differentiating with respect to s, we also find that $\mathscr{V}_{n,M+1}(s) = \mathscr{V}_{n,M}(s)$.

The integral terms must also be approximated at each time and variance line. Assuming that $Y_1 \sim N(\alpha_{j_1}, \beta_{j_1}^2)$ and $Y_2 \sim N(\alpha_{j_2}, \beta_{j_2}^2)$ under $\hat{\mathbb{Q}}$, the integrals, which we denote by $I_1(\tau_n, s, v_m)$ and $I_2(\tau_n, s, v_m)$, can be approximated using Gauss-Hermite quadrature,

$$\begin{aligned}
I_1(\tau_n, s, v_m) &= -\tilde{\lambda}_1 \int_{\mathbb{R}} V(\tau_n, se^y, v_m) G_1(y)\, \mathrm{d}y \approx -\frac{\tilde{\lambda}_1}{\sqrt{\pi}} \sum_{l=1}^{L} \varrho_l^H V_{n,m}(se^{\sqrt{2}\beta_{j_1} z_l^H + \alpha_{j_1}}), \\
I_2(\tau_n, s, v_m) &= -\tilde{\lambda}_2 \int_{\mathbb{R}} V(\tau_n, se^{-y}, v_m) G_2(y)\, \mathrm{d}y \approx -\frac{\tilde{\lambda}_2}{\sqrt{\pi}} \sum_{l=1}^{L} \varrho_l^H V_{n,m}(se^{-\sqrt{2}\beta_{j_2} z_l^H - \alpha_{j_2}}).
\end{aligned}$$

where ϱ_l^H and z_l^H are the weights and abscissas of the Gauss-Hermite quadrature scheme with L integration points.[3] The parameters $\tilde{\kappa}_1$ and $\tilde{\kappa}_2^-$ are also approximated using Gauss-Hermite quadrature.

The MOL approximation of equation (5.4) at $\tau = \tau_n$ and $v = v_m$ is therefore

$$\begin{aligned}
&a(s, v_m) V''_{n,m}(s) + b(s, v_m) V'_{n,m}(s) - c(\tau_n, s, v_m) V_{n,m}(s) \\
&\quad = F(\tau_n, s, v_m) + I_1(\tau_n, s, v_m) + I_2(\tau_n, s, v_m),
\end{aligned} \tag{5.8}$$

[3] The implementation of the Gauss-Hermite quadrature in this chapter uses the MATLAB® function developed by Van Damme (2010).

where the coefficients are given by

$$
\begin{aligned}
a(s,v_m) &= \max\left\{\frac{1}{2}\sigma^2 v_m s^2, 10^{-4}\right\}, \qquad \text{(regularized coefficient)}\\
b(s,v_m) &= -\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)s,\\
c(\tau_n,s,v_m) &= \frac{\omega^2 v_m}{(\Delta v)^2} + (\tilde{\lambda}_1+\tilde{\lambda}_2) + \frac{\max\left\{\xi\eta-(\xi+\Lambda)v_m,0\right\}}{\Delta v}\\
&\quad - \frac{\min\left\{\xi\eta-(\xi+\Lambda)v_m,0\right\}}{\Delta v} + \begin{cases} 1/\Delta\tau & \text{if } n=1,2\\ 3/(2\Delta\tau) & \text{if } n\geq 3,\end{cases}\\
F(\tau_n,s,v_m) &= -\frac{\omega^2 v_m}{2(\Delta v)^2}\left[V_{n,m+1}(s)+V_{n,m-1}(s)\right]\\
&\quad - \frac{\omega(\sigma_1\rho_1-\sigma_2\rho_2)}{2\Delta v} v_m s\left[\mathscr{V}_{n,m+1}(s)-\mathscr{V}_{n,m-1}(s)\right]\\
&\quad - \frac{\max\left\{\xi\eta-(\xi+\Lambda)v_m,0\right\}}{\Delta v}V_{n,m+1}(s)\\
&\quad + \frac{\min\left\{\xi\eta-(\xi+\Lambda)v_m,0\right\}}{\Delta v}V_{n,m-1}(s)\\
&\quad - \begin{cases} V_{n-1,m}(s)/\Delta\tau & \text{if } n=1,2\\ \left[4V_{n-1,m}(s)-V_{n-2,m}(s)\right]/(2\Delta\tau) & \text{if } n\geq 3,\end{cases}
\end{aligned}
\tag{5.9}
$$

for $m=1,\ldots,M-1$. At the last variance line $m=M$, boundary condition (5.7) implies that some of the coefficients above have to be redefined as

$$
\begin{aligned}
b(s,v_M) &= \frac{\omega(\sigma_1\rho_1-\sigma_2\rho_2)}{2\Delta v} v_M s - \left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right)s,\\
c(\tau_n,s,v_M) &= \frac{\omega^2 v_M}{2(\Delta v)^2} + (\tilde{\lambda}_1+\tilde{\lambda}_2)\\
&\quad - \frac{\min\left\{\xi\eta-(\xi+\Lambda)v_M,0\right\}}{\Delta v} + \begin{cases} 1/\Delta\tau & \text{if } n=1,2\\ 3/(2\Delta\tau) & \text{if } n\geq 3,\end{cases}\\
F(\tau_n,s,v_M) &= -\frac{\omega^2 v_m}{2(\Delta v)^2}V_{n,M-1}(s) - \frac{\omega(\sigma_1\rho_1-\sigma_2\rho_2)}{2\Delta v} v_M s\mathscr{V}_{n,M-1}(s)\\
&\quad + \frac{\min\left\{\xi\eta-(\xi+\Lambda)v_M,0\right\}}{\Delta v}V_{n,M-1}(s)\\
&\quad - \begin{cases} V_{n-1,M}(s)/\Delta\tau & \text{if } n=1,2\\ \left[4V_{n-1,M}(s)-V_{n-2,M}(s)\right]/(2\Delta\tau) & \text{if } n\geq 3,\end{cases}
\end{aligned}
\tag{5.10}
$$

These are obtained by comparing (5.8) to (5.4) after substituting the finite difference and integral approximations.

5.2.2 Solution via the Riccati Transform

As the source term of equation (5.8) at each $m = 1, \ldots, M-1$ depends on $V_{n,m+1}(s)$ and $\mathscr{V}_{n,m+1}(s)$, none of which are available at the current variance line, and involves integrals based on $V_{n,m}(s)$, the sequence of equations (5.8) represents a system of coupled second-order integral-ordinary differential equations. To this end, we employ a nested two-level iterative scheme in conjunction with the Riccati transform approach of Meyer and van der Hoek (1997), as was done by Chiarella et al. (2009). In the first iteration level, referred to as *integral term iterations*, the integral terms are approximated using the latest available values of the option price. This enables us to treat (5.8) as a second-order ODE which can then be solved using the Riccati transform method in the second iteration level, referred to as the *variance line iterations*.

The two-stage iterative process and the Riccati transform solution of equation (5.8) is shown in Algorithm 1. The succeeding discussions add details to the structure of the algorithm presented in Algorithm 1.

Let k' and k be the iteration counters for the integral term and variance line iterations, respectively, and let $V^k_{n,m}(s)$ and $\mathscr{V}^k_{n,m}(s)$ be the kth iterates of V and $\mathscr{V}$, replacing k with k' when we refer to their counterparts in the integral term iterations. "Previous" iterates will be denoted with a superscript $k-1$ or $k'-1$. The Riccati transform method requires that the computational s-domain be partitioned with a mesh $0 = s_0 < s_1 < \cdots < s_J$. Note that this mesh may not necessarily have a uniform width. In view of boundary conditions (5.6), it is known that $V^k_{n,m}(s_0) = 0$ and $\mathscr{V}'^k_{n,m}(s_J) = 0$.

In each integral term iteration, I_1 and I_2 are computed using previous iteration values of V. But since $V^{k'-1}_{n,m}(s)$ is available only at the mesh points for s, the summands required for I_1 and I_2 are linearly interpolated or extrapolated. The integrals are computed only once per integral term iteration and are not recomputed during the variance line iterations.

We now discuss what transpires in the general kth variance line iteration.

At $m = 0$, equation (5.8) is not solved directly. Instead, we apply the quadratic extrapolation formulas

$$\begin{aligned} V^k_{n,0}(s_j) &\approx 3V^{k-1}_{n,1}(s_j) - 3V^{k-1}_{n,2}(s_j) + V^{k-1}_{n,3}(s_j), \\ \mathscr{V}^k_{n,0}(s_j) &\approx 3\mathscr{V}^{k-1}_{n,1}(s_j) - 3\mathscr{V}^{k-1}_{n,2}(s_j) + \mathscr{V}^{k-1}_{n,3}(s_j) \end{aligned} \tag{5.11}$$

to approximate the option price and delta at the first variance line.[4]

[4] Provided that the Feller condition is satisfied, the quadratic extrapolation combined with the MOL approach produces a consistent approximation of the pricing equation as $v \to 0$ (Chiarella et al., 2009, Appendix). See Section 5.2.4 for additional remarks on boundary conditions at $v = 0$ when the Feller condition is violated.

Algorithm 1: MOL for the discounted European exchange option price

Result: Option price V and delta $\mathscr{V}$
Inputs: SVJD model parameters, jump distribution parameters, partitions for $[0,T]$, $[0,s_J]$, and $[0,v_M]$;
Compute Gauss-Hermite parameters ϱ_l^H and z_l^H, $l=1,\ldots,L$ and compute κ using Gauss-Hermite quadrature;
Initialize $V_{0,m}(s_j)=e^{-q_1T}(s_j-e^{(q_1-q_2)T})^+$, $\mathscr{V}_{0,m}(s_j)=e^{-q_1T}\text{Heaviside}(s_j-e^{(q_1-q_2)T})$ for all j,m;
for $n=1$ **to** N **do**
  Set $V_{n,m}^{k'-1}(s_j)=V_{n-1,m}(s_j)$ and $\mathscr{V}_{n,m}^{k'-1}(s_j)=\mathscr{V}_{n-1,m}(s_j)$;
  Initialize $\epsilon_{\text{integ}}=1$ and $k'=1$;
  /* Commence integral term iterations */
  while $\epsilon_{integ}>10^{-8}$ & $k'<k'_{max}$ **do**
    Compute $I_1(\tau_n,s_j,v_m)$, $I_2(\tau_n,s_j,v_m)$ using $V_{n,m}^{k'-1}(s_j)$ for all j,m;
    Set $V_{n,m}^{k-1}(s_j)=V_{n,m}^{k'-1}(s_j)$ and $\mathscr{V}_{n,m}^{k-1}(s_j)=\mathscr{V}_{n,m}^{k'-1}(s_j)$;
    Initialize $\epsilon_{\text{var}}=1$, $k=1$, and set $k_{max}=50$;
    /* Commence variance line iterations */
    while $\epsilon_{var}>10^{-8}$ & $k<k_{max}$ **do**
      for $m=0$ **to** M **do**
        if $m=1$ **then**
          /* Quadratic extrapolation formulas */
          Compute $V_{n,0}^k$ and $\mathscr{V}_{n,0}^k$ using equations (5.11);
        end
        else if $1\le m\le M-1$ **then**
          /* Forward sweep */
          for $j=1$ **to** J **do**
            Compute $R(s_j)$ and $w(s_j)$ using the trapezoidal rule on (5.14);
          end
          /* Reverse sweep over $[0,s_J]$ */
          Compute $\mathscr{V}_{n,m}^k(s_J)$;
          for $j=J$ **to** 1 **do**
            Compute $\mathscr{V}_{n,m}^k(s_j)$ using the trapezoidal rule;
            Compute $V_{n,m}^k(s_j)$ using the Riccati transform (5.13);
          end
        end
        else
          Recalculate the coefficients of (5.14) with (5.10)// At $v=v_M$
          Perform the forward sweep and reverse sweep as in when $1\le m\le M-1$ to compute $V_{n,M}^k(s_j)$ and $\mathscr{V}_{n,M}^k(s_j)$ for all j;
        end
      end
      Update $\epsilon_{\text{var}}=\max_m\{\max_j\{|V_{n,M}^k(s_j)-V_{n,m}^{k-1}(s_j)|\}$;
      Update $V_{n,m}^{k-1}=V_{n,m}^k$ and $\mathscr{V}_{n,m}^{k-1}=\mathscr{V}_{n,m}^k$;
      Update $k=k+1$;
    end
    Update $V_{n,m}^{k'}=V_{n,m}^k$ and $\mathscr{V}_{n,m}^{k'}=\mathscr{V}_{n,m}^k$;
    Update $\epsilon_{\text{integ}}=\max_m\{\max_j\{|V_{n,M}^{k'}(s_j)-V_{n,m}^{k'-1}(s_j)|\}$;
    Update $V_{n,m}^{k'-1}=V_{n,m}^{k'}$ and $\mathscr{V}_{n,m}^{k'-1}=\mathscr{V}_{n,m}^{k'}$;
    Update $k'=k'+1$
  end
  Update $V_{n,m}=V_{n,m}^{k'}$ and $\mathscr{V}_{n,m}=\mathscr{V}_{n,m}^{k'}$;
end

For $m = 1, \ldots, M-1$, equation (5.8) is transformed into a system of first-order ODEs

$$\begin{aligned} V'^{k}_{n,m}(s) &= \mathscr{V}^{k}_{n,m}(s), \\ \mathscr{V}'^{k}_{n,m}(s) &= C(\tau_n, s, v_m)V^{k}_{n,m}(s) + D(\tau_n, s, v_m)\mathscr{V}^{k}_{n,m}(s) + g(\tau_n, s, v_m), \end{aligned} \tag{5.12}$$

where $C = c/a$, $D = -b/a$, and $g = (F + I_1 + I_2)/a$. The corresponding Riccati transformation is given by

$$V^{k}_{n,m}(s) = R(s)\mathscr{V}^{k}_{n,m}(s) + w(s), \tag{5.13}$$

where R and w are solutions of the initial value problems

$$\begin{aligned} R'(s) &= 1 - D(\tau_n, s, v_m)R(s) - C(\tau_n, s, v_m)R^2(s), \\ w'(s) &= -C(\tau_n, s, v_m)R(s)w(s) - g(\tau_n, s, v_m)R(s), \end{aligned} \tag{5.14}$$

with $R(0) = w(0) = 0$. System (5.14) is integrated using the trapezoidal rule starting at $s_0 = 0$ until s_J. Details on the trapezoidal rule are provided in Appendix D.1. This process is referred to as the *forward sweep*. Once R and w have been determined, the equation

$$\begin{aligned} \mathscr{V}'^{k}_{n,m}(s) = &\left[C(\tau_n, s, v_m)R(s) + D(\tau_n, s, v_m)\right]\mathscr{V}^{k}_{n,m}(s) \\ &+ C(\tau_n, s, v_m)w(s) + g(\tau_n, s, v_m) \end{aligned}$$

is integrated using the trapezoidal rule starting at $s = s_J$. The starting point $\mathscr{V}^{k}_{n,m}(s_J)$ is calculated using the boundary condition $\mathscr{V}'^{k}_{n,m}(s_J) = 0$ and the above equation. The option price $V^{k}_{n,m}(s)$ can be recovered using the Riccati transform (5.13). The *reverse sweep* concludes after calculating $V^{k}_{n,m}(s)$ and $\mathscr{V}^{k}_{n,m}(s)$ for all s.

The process at the last variance line $m = M$ is similar to that for $1 \leq m \leq M-1$, except that the coefficients C, D, and g must be adjusted in view of equations (5.10).

The variance line iterations terminate once the convergence criterion

$$\max_{0 \leq m \leq M} \left\{ \max_{0 \leq j \leq J} \left| V^{k}_{n,m}(s_j) - V^{k-1}_{n,m}(s_j) \right| \right\} < 10^{-8} \tag{5.15}$$

is satisfied. Otherwise, the current iterates are stored as previous iterates and the next variance line iteration commences. If (5.15) is met, then $V^{k}_{n,m}(s)$ and $\mathscr{V}^{k}_{n,m}(s)$ are stored as $V^{k'}_{n,m}(s)$ and $\mathscr{V}^{k'}_{n,m}(s)$, respectively, and the current integral term iteration continues. The convergence criterion may also include the option delta, if a more stringent criterion is desired.

Now back at the current integral term iteration, criterion (5.15) is checked again using $V^{k'}_{n,m}(s)$ and $\mathscr{V}^{k'}_{n,m}(s)$. If it is not satisfied, then the current iterates are stored as previous iterates and the next integral term iteration commences. At this point, I_1 and I_2 are recalculated and another pass

of variance line iterations commences. Otherwise, the integral term iterations terminate and the latest iterates $V_{n,m}^{k'}(s)$ and $\mathscr{V}_{n,m}^{k'}(s)$ are taken to be the solutions $V_{n,m}(s)$ and $\mathscr{V}_{n,m}(s)$ at the nth time step. The algorithm then proceeds to the next time step.

As seen above, the MOL algorithm naturally computes the option price $V_{n,m}(s)$, the option delta $\mathscr{V}_{n,m}(s)$, and the option gamma $\mathscr{V}'_{n,m}(s)$. The option gamma profile can be stored by computing $\mathscr{V}'^{k}_{n,m}(s_j)$ with the second equation in (5.12) after computing $V^{k}_{n,m}(s_j)$ and $\mathscr{V}^{k}_{n,m}(s_j)$ in the reverse sweep.

The MOL algorithm discussed above can also be used to approximate the joint transition density function of the log-asset yield ratio $X(t) = \ln S(t)$ and the variance $v(t)$. The joint tdf of $(X(t), v(t))$ is the function $H(T, x_T, v_T; t, x, v) = H(t, x, v)$ defined in (4.51) and, for every fixed $(x_T, v_T) \in \mathbb{R} \times \mathbb{R}_+$, is known to satisfy equation (4.49) for $(t, x, v) \in [0, T] \times \mathbb{R} \times \mathbb{R}_+$ subject to terminal condition (4.50) (see also Appendix A.5). For the MOL, we consider a change of variables $\tau = T - t$, so the terminal value problem (4.49) and (4.50) becomes the initial value problem

$$\begin{aligned}\frac{\partial H}{\partial \tau} &= \frac{1}{2}\sigma^2 v \frac{\partial^2 H}{\partial x^2} + \frac{1}{2}\omega^2 v \frac{\partial^2 H}{\partial v^2} + \omega(\sigma_1\rho_1 - \sigma_2\rho_2) v \frac{\partial^2 H}{\partial x \partial v} \\ &\quad - \left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \frac{1}{2}\sigma^2 v\right)\frac{\partial H}{\partial x} + \left[\xi\eta - (\xi + \Lambda)v\right]\frac{\partial H}{\partial v} - (\tilde{\lambda}_1 + \tilde{\lambda}_2)H \\ &\quad + \tilde{\lambda}_1 \int_{\mathbb{R}} H(\tau, x + y, v) G_1(y)\, \mathrm{d}y + \tilde{\lambda}_2 \int_{\mathbb{R}} H(\tau, x - y, v) G_2(y)\, \mathrm{d}y,\end{aligned} \tag{5.16}$$

for $(\tau, x, v) \in [0, T] \times \mathbb{R} \times \mathbb{R}_+$ with initial condition

$$H(0, x, v; 0, x_0, v_0) = \delta(x - x_0)\delta(v - v_0).$$

Here, $(x_0, v_0) \in \mathbb{R} \times \mathbb{R}_+$ are given constants and $\delta(\cdot)$ is the Dirac delta function. Equation (5.16) must also be solved subject to the boundary conditions

$$\lim_{x \to \pm\infty} H(\tau, x, v) = 0, \qquad \lim_{v \to \infty} H(\tau, x, v) = 0, \qquad \lim_{v \to 0^+} H(\tau, x, v) = 0.$$

5.2.3 Adjustments for the American Exchange Option Price

The MOL algorithm discussed above has to be slightly adjusted to determine the unknown early exercise boundary during the forward sweep. For convenience, we continue to denote the discounted American exchange option price by $V(\tau, s, v)$. Then V also satisfies (5.4), subject to the same initial condition. However, this equation must be solved over the restricted domain $0 < v < \infty$, $0 \le \tau \le T$, and $0 < s < A(\tau, v) \equiv B(\tau, v)e^{(q_1 - q_2)(T-\tau)}$, where $B(\tau, v)$ us the unknown early exercise boundary. For $s \ge A(\tau, v)$, the option price is known analytically as $V(\tau, s, v) = e^{-q_1(T-\tau)}(s - e^{(q_1-q_2)(T-\tau)})$. In addition to the usual boundary conditions $\lim_{s\to 0^+} V(\tau, s, v) = 0$ and $\lim_{v\to\infty} \partial V/\partial v = 0$, we also impose smooth-pasting condition $\lim_{s \to A(\tau,v)} \partial V/\partial s = e^{-q_1(T-\tau)}$. A suitable initial condition

for the early exercise boundary is $A(0^+, v) = B(0^+, v)e^{(q_1-q_2)T}$, where $B(0^+, v)$ is a solution of equation (4.41), and let $A_{n,m} = A(\tau_n, v_m)$ be the approximation of the boundary at $\tau = \tau_n$ and $v = v_m$.

The MOL approximation of the American exchange option at $\tau = \tau_n$ and $v = v_m$ is also given by equation (5.8) with coefficients (5.9) for $m = 1, \ldots, M-1$ and (5.10) for $m = M$. Thus, in this section, we highlight how the Riccati transform method in the previous section is modified to solve for the early exercise boundary. We assume that s_J is chosen such that $A_{n,m} < s_J$ for all n, m. Algorithm 2 shows the modified MOL algorithm.

In each of the k variance line iterations, the $V^k_{n,0}(s)$ and $\mathscr{V}^k_{n,0}(s)$ are estimated using the quadratic extrapolation formulas (5.11). For each $m = 1, \ldots, M-1$, the forward sweep of system (5.14) via the trapezoidal rule starts at s_0 in the direction of increasing s. The main difference is that for the American exchange option, the sign of the function

$$\phi(s_j) = R(s_j)e^{-q_1(T-\tau_n)} + w(s_j) - e^{-q_1(T-\tau_n)}\left(s_j - e^{(q_1-q_2)(T-\tau_n)}\right) \tag{5.17}$$

is monitored at each j. Equation (5.17) arises from combining the value-matching condition and the Riccati transform equation. Indeed, since the option price and delta are known analytically as

$$V^k_{n,m}(s) = e^{-q_1(T-\tau_n)}(s - e^{(q_1-q_2)(T-\tau_n)}), \qquad \mathscr{V}^k_{n,m}(s) = e^{-q_1(T-\tau_n)}$$

for $s \geq A(\tau_n, v_m) = B(\tau_n, v_m)e^{(q_1-q_2)(T-\tau_n)}$, courtesy of the value-matching condition, we substitute these known forms into (5.13) to obtain (5.17). The initial value is $\phi(s_0) = e^{-q_2(T-\tau_n)} > 0$. The forward sweep stops at the index j^* at which ϕ first becomes negative; i.e. j^* is the first index such that $\phi(s_{j^*-1})\phi(s_{j^*}) < 0$. The approximation of the early exercise boundary at this iteration, which we denote by $A^k_{n,m}$, is estimated to be the zero of the cubic spline interpolant through the points $\{(s_i, \phi(s_i))\}_{i=j^*-2}^{j^*+1}$ that occurs in between s_{j^*-1} and s_{j^*}. This zero is numerically solved using any standard root finding algorithm.

Once $A^k_{n,m}$ has been determined, the reverse sweep for $\mathscr{V}^k_{n,m}(s)$ is then carried out over the interval $[0, A^k_{n,m}]$. Since $A^k_{n,m}$ is not part of the regular mesh, the trapezoidal rule is first applied over $[s_{j^*-1}, A^k_{n,m}]$, where the required values of C, D, g, R, and w are linearly interpolated from values at s_{j^*-1} and s_{j^*}, which were determined as part of the forward sweep. Once $\mathscr{V}^k_{n,m}(s_{j^*-1})$ has been calculated, the reverse sweep can continue over $[0, s_{j^*-1}]$ along the regular mesh. For $j \geq j^*$, it is known that $\mathscr{V}^k_{n,m}(s_j) = e^{-q_1(T-\tau_n)}$. The value of $V^k_{n,m}(s_j)$ can be calculated using the Riccati transform equation.

The same process applies at the last variance line, except that a recalculation of the coefficients of the Riccati system according to equation (5.10) is required before commencing the forward sweep.

Once the convergence criterion (5.15) is satisfied at both iteration levels, the final iterates are

Algorithm 2: MOL for the discounted American exchange option price

Result: Option price V and delta $\mathscr{V}$
Inputs: SVJD model parameters, jump distribution parameters, partitions for $[0,T]$, $[0,s_J]$, and $[0,v_M]$;
Compute Gauss-Hermite parameters ϱ_l^H and z_l^H, $l=1,\ldots,L$ and compute κ using Gauss-Hermite quadrature;
Compute $A_{0,m}=B(0^+,v_m)e^{(q_1-q_2)T}$ using (4.41);
Initialize $V_{0,m}(s_j)=e^{-q_1T}(s_j-e^{(q_1-q_2)T})^+$, $\mathscr{V}_{0,m}(s_j)=e^{-q_1T}\text{Heaviside}(s_j-e^{(q_1-q_2)T})$ for all j,m;
for $n=1$ **to** N **do**
  Set $V_{n,m}^{k'-1}(s_j)=V_{n-1,m}(s_j)$ and $\mathscr{V}_{n,m}^{k'-1}(s_j)=\mathscr{V}_{n-1,m}(s_j)$;
  Initialize $\epsilon_{\text{integ}}=1$ and $k'=1$;
  /* Commence integral term iterations */
  while $\epsilon_{integ}>10^{-8}$ *&* $k'<k'_{max}$ **do**
    Compute $I_1(\tau_n,s_j,v_m)$, $I_2(\tau_n,s_j,v_m)$ using $V_{n,m}^{k'-1}(s_j)$ for all j,m;
    Set $V_{n,m}^{k-1}(s_j)=V_{n,m}^{k'-1}(s_j)$ and $\mathscr{V}_{n,m}^{k-1}(s_j)=\mathscr{V}_{n,m}^{k'-1}(s_j)$;
    Initialize $\epsilon_{\text{var}}=1$, $k=1$, and set $k_{max}=50$;
    /* Commence variance line iterations */
    while $\epsilon_{var}>10^{-8}$ *&* $k<k_{max}$ **do**
      for $m=0$ **to** M **do**
        if $m=1$ **then**
          Compute $V_{n,0}^k$ and $\mathscr{V}_{n,0}^k$ using equations (5.11);
        end
        else if $1\le m\le M-1$ **then**
          /* Forward sweep and computation of $A_{n,m}^k$ */
          Initialize forward sweep while-loop parameters: $i=1$, $\phi(s_0)=e^{-q_2(T-\tau_n}$, $\Pi=1$;
          while $\Pi>0$ **do**
            Compute $R(s_i)$, $w(s_i)$, and $\phi(s_i)$;
            Update $\Pi=\phi(s_i)\phi(s_{i-1})$ and assign $j^*=i$;
            Update $i=i+1$;
          end
          Compute $\phi(s_{j^*+1})$ and fit a cubic spline through $\{(s_i,\phi(s_i))\}_{i=j^*-2}^{j^*+1}$;
          Take $A_{n,m}^k$ as the zero of the cubic spline in (s_{j^*-1},s_{j^*});
          /* Reverse sweep over $[s_{j^*-1},A_{n,m}^k]$ */
          Set $\mathscr{V}(A_{n,m}^k)=e^{-q_1(T-\tau_n)}$ and linearly interpolate values of C, D, g, R, and w at $s=A_{n,m}^k$;
          Calculate $\mathscr{V}_{n,m}^k(s_{j^*-1})$ using the trapezoidal rule and $V_{n,m}^k(s_{j^*-1})$ using (5.13);
          /* Reverse sweep over $[0,s_{j^*-1}]$ */
          for $j=j^*-2$ **to** 1 **do**
            Compute $\mathscr{V}_{n,m}^k(s_j)$ using the trapezoidal rule;
            Compute $V_{n,m}^k(s_j)$ using the Riccati transform (5.13);
          end
          Calculate $V_{n,m}^k(s_j)$ and $\mathscr{V}_{n,m}^k(s_j)$ using the analytical form in the stopping region, $j=j^*,\ldots,J$;
        end
        else
          Recalculate the coefficients of (5.14) with (5.10)// At $v=v_M$
          Perform the forward sweep and reverse sweep as in when $1\le m\le M-1$;
        end
      end
      Update $\epsilon_{\text{var}}=\max_m\{\max_j\{|V_{n,M}^k(s_j)-V_{n,m}^{k-1}(s_j)|\}$;
      Update $V_{n,m}^{k-1}=V_{n,m}^k$ and $\mathscr{V}_{n,m}^{k-1}=\mathscr{V}_{n,m}^k$;
      Update $k=k+1$;
    end
    Update $V_{n,m}^{k'}=V_{n,m}^k$ and $\mathscr{V}_{n,m}^{k'}=\mathscr{V}_{n,m}^k$;
    Update $\epsilon_{\text{integ}}=\max_m\{\max_j\{|V_{n,M}^{k'}(s_j)-V_{n,m}^{k'-1}(s_j)|\}$;
    Update $V_{n,m}^{k'-1}=V_{n,m}^{k'}$ and $\mathscr{V}_{n,m}^{k'-1}=\mathscr{V}_{n,m}^{k'}$;
    Update $k'=k'+1$
  end
  Update $V_{n,m}=V_{n,m}^{k'}$ and $\mathscr{V}_{n,m}=\mathscr{V}_{n,m}^{k'}$;
end

then stored as the solution at the nth time step. The algorithm then proceeds to the next time step. In the American case, the convergence criterion may also include the early exercise boundary

5.2.4 Venttsel Boundary Conditions at $v = v_M$

The choice of boundary conditions for equation (5.4) affects the quality of the MOL approximation obtained using the algorithms discussed above. In most financial applications, the initial condition is usually dictated by the type of instrument being priced. However, articulating boundary conditions in the spatial variables is not as straightforward since the equation may be degenerate at certain points of the boundary or the domain may be unbounded. The problem considered in this chapter struggles with both issues.

The pricing equation (5.4) is degenerate when $s = 0$ or $v = 0$. As such, the structure of the IPDE determines whether conditions at these boundaries should be independently imposed or if the equation itself should naturally hold at these boundaries. Chiarella et al. (2009) consider this issue in more detail, with the aid of the Fichera function for degenerate equations, for the closely related American call option under SVJD dynamics and justify the use of quadratic extrapolation for the solution of the pricing equation along the first variance line $v = v_0$.[5] They showed that equation (5.4) holds when $v = 0$ provided that the Feller condition (i.e. Assumption 3.1) is satisfied. This means that no independent condition is required at that point.

If the Feller condition is violated, then one can impose an independent Dirichlet condition along $v = 0$. However, independent Dirichlet conditions may not be consistent with the solution of the pricing equation and would result to discontinuities with respect to the parameters of the variance processes (Meyer, 2015, Example 1.15). For the specific case of pricing European options under a Heston (1993) stochastic volatility process, Meyer (2015, Example 1.22) shows that, regardless of whether the Feller condition is satisfied, one can regularize the pricing equation (i.e. instead of just $\frac{1}{2}vs^2$ and $\frac{1}{2}\sigma^2 v$ as the coefficients of $\frac{\partial^2 V}{\partial s^2}$ and $\frac{\partial^2 V}{\partial v^2}$, respectively, one has $\max\{\frac{1}{2}vs^2, \epsilon\}$ and $\max\{\frac{1}{2}\sigma^2 v, \epsilon\}$ for $0 < \epsilon \ll 1$) and use the regularized pricing equation evaluated at $v = 0$ as a (Venttsel) boundary condition. In the implementation here, we regularize the coefficient of $V''_{n,m}(s)$ in (5.8) (see $a(s, v_m)$ in (5.9)) even if the parameter values chosen for the numerical experiments do satisfy the Feller condition.[6]

In contrast, we allow the dynamics of $\tilde{S}(t)$ to dictate the boundary condition at $s = 0$. From (4.13), if the process starts at zero, then it will stay at zero, implying that the option becomes worthless throughout its life. This motivates the choice of the first boundary condition in (5.6).

[5] Such an analysis was also done by Kang and Meyer (2014) for American calls under stochastic volatility and stochastic interest rates. Meyer (2015) provides a discussion of the Fichera theory applied to common financial problems.

[6] In this regard, Assumption 3.1 is imposed mainly as a technical requirement for the construction of the equivalent martingale measure or the pricing measure $\mathbb{Q}$ (see also the discussion below the statement of Assumption 2.2 and the discussion in Section 2.3).

For the MOL algorithm, we truncated the infinite s and v domains. As such, boundaries must be prescribed at the far boundaries $s = s_J$ and $v = v_M$. For the Riccati transform method to work in the European case, the boundary condition at s_J must be equivalent to a scalar equation in V and $\mathscr{V}$, and this is attained through the second condition in (5.6) when approximated as $\mathscr{V}'_{n,m}(s_J) = 0$ (although there may be alternatives to this). In the American case, no such condition is required since we already know the American exchange option price in analytical form for $s_j \geq A(\tau_n, v_m)$ and s_J is chosen to be large enough such that $0 < A(\tau_n, v_m) < s_J$ for all n and m. Hence, we only need to consider boundary conditions at v_M.

As an alternative to the boundary condition $\partial V(\tau, s, v_M)/\partial v = 0$, we consider Venttsel boundary conditions derived from the pricing equation.[7] Following the theory of Venttsel boundary conditions in Meyer (2015), we find that the equation

$$\begin{aligned} 0 = \frac{1}{2}\sigma^2 v_M s^2 \frac{\partial^2 V}{\partial s^2} - \left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right) s\frac{\partial V}{\partial s} + \left[\xi\eta - (\xi+\Lambda)v_M\right]\frac{\partial V}{\partial v} \\ - (\tilde{\lambda}_1 + \tilde{\lambda}_2)V - \frac{\partial V}{\partial \tau} - I_1(\tau, s, v_M) - I_2(\tau, s, v_M) \end{aligned} \tag{5.18}$$

is an admissible Venttsel condition at $v = v_M$ provided that $(\xi + \Lambda)v_M \geq \xi\eta$. In practical terms, equation (5.18) is equivalent to the assumption that the option delta $\partial V/\partial s$ and vega $\partial V/\partial v$ are insensitive to the instantaneous variance for when the variance level is sufficiently large, since the second-order terms with respect to v are absent from the equation. Another Venttsel condition is given by

$$\begin{aligned} 0 = \frac{1}{2}\sigma^2 v_M s^2 \frac{\partial^2 V}{\partial s^2} - \left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-\right) s\frac{\partial V}{\partial s} - (\tilde{\lambda}_1 + \tilde{\lambda}_2)V - \frac{\partial V}{\partial \tau} \\ - I_1(\tau, s, v_M) - I_2(\tau, s, v_M), \end{aligned} \tag{5.19}$$

which resembles the pricing equation when volatility is constant at $\sigma\sqrt{v_M}$. Equation (5.19) also implies that when volatility is high, the option price, delta, and vega are no longer sensitive to changes in volatility and the randomness of the volatility no longer affects the option price. Whichever Venttsel boundary condition is chosen, the equation must be solved subject to boundary conditions in τ and s.

In the MOL implementation, the adoption of either Venttsel condition implies a change in the coefficients of equation (5.8) at $v = v_M$. For equation (5.18), we employ a backward difference approximation of $\partial V/\partial v$. This is appropriate in view of the upwinding difference approximation for the other variance lines as this Venttsel condition requires that $(\xi + \Lambda)v_M \geq \xi\eta$. The usual discretization is employed for $\partial V/\partial\tau$. Once the coefficients of equation (5.8) have been updated, the Riccati solution for the last variance line can proceed as outlined in Section 5.2.

[7] We do not formally define Venttsel boundary conditions in this thesis. We refer to Meyer (2015, Section 1.2.3) for a discussion of Venttsel boundary conditions in the context of financial pricing problems.

The impact of these Venttsel conditions on the performance of the MOL algorithm is explored in the succeeding sections.

5.3 Finite Difference with PSOR

As an alternative to the LSMC and the MOL approaches, we also consider an approximation of American exchange option prices using a full finite difference scheme that is then solved using the projected successive overrelaxation (PSOR) algorithm. The American exchange option price can be characterized as a solution of the linear complementarity problem (LCP)

$$\begin{aligned}(\mathscr{L}V(\tau,s,v))(V(\tau,s,v)-\phi(\tau,s,v))&=0,\\ V(\tau,s,v)-\phi(\tau,s,v)&\geq 0,\\ \mathscr{L}V(\tau,s,v)&\leq 0,\end{aligned} \tag{5.20}$$

where $\phi(\tau,s,v)=(se^{-q_1(T-\tau)}-e^{-q_2(T-\tau)})^+$ is the payoff function and $\mathscr{L}$ is the IPDE operator corresponding to the right-hand side of equation (5.4). This problem is solved for all $(\tau,s,v)\in[0,T]\times\mathbb{R}^2_+$ subject to the initial condition (5.5) and boundary conditions (5.6). Compared to the MOL approach, the LCP formulation notably does not involve the early exercise boundary. It is instead determined after-the-fact as the boundary of the set $\{(\tau,s,v)\in[0,T]\times\mathbb{R}^2_+ : V(\tau,s,v)>\phi(\tau,s,v)\}$.

5.3.1 Spatial Discretization

We then use finite difference approximations to discretize the LCP in the spatial variables s and v. As in the MOL, we consider bounded computational domains $[0,s_J]$ and $[0,v_M]$ for s and v, respectively, that are partitioned as $0=s_0<s_1<\cdots<s_J$ and $0=v_0<v_1<\cdots<v_M$ with uniform mesh sizes Δs and Δv, respectively. We let $V_{m,j}(\tau)=V(\tau,s_j,v_m)$ denote the approximation of the discounted American exchange option price at $s=s_j$ and $v=v_m$. We momentarily keep τ continuous, although the derivative with respect to τ will also be approximated via finite differences later. We follow Chiarella et al. (2009), Ekström et al. (2009) and Kang and Meyer (2014) in the choices of finite difference approximations of the spatial derivatives.

Interior of the Computational Domain: We retain the central difference approximation for $\frac{\partial^2 V}{\partial v^2}$ and the upwinding difference approximation for $\frac{\partial V}{\partial v}$ at the interior points of the mesh (i.e. for $1\leq m\leq M-1$ and $1\leq j\leq J-1$). We also use central differences to approximate partial derivatives in s,

$$\frac{\partial V_{m,j}(\tau)}{\partial s}\approx\frac{V_{m,j+1}(\tau)-V_{m,j-1}(\tau)}{2\Delta s}+O((\Delta s)^2),$$

$$\frac{\partial^2 V_{m,j}(\tau)}{\partial s^2} \approx \frac{V_{m,j+1}(\tau) - 2V_{m,j}(\tau) + V_{m,j-1}(\tau)}{(\Delta s)^2} + O((\Delta s)^2).$$

To approximate the cross partial derivative $\frac{\partial^2 V}{\partial s \partial v}$, we use a non-standard seven-point stencil dependent on the sign of $\bar{\rho} = (\sigma_1\rho_1 - \sigma_2\rho_2)/\sigma$, the instantaneous correlation coefficient between the Wiener processes $\bar{W}$ and $\bar{Z}$ in equations (4.10) and (4.13). Specifically, if $\bar{\rho} \le 0$, then we use the approximation

$$\frac{\partial^2 V_{m,j}(\tau)}{\partial s \partial v} \approx \frac{1}{2}\left[\frac{V_{m+1,j+1}(\tau) - V_{m+1,j}(\tau) - (V_{m,j+1}(\tau) - V_{m,j}(\tau))}{\Delta s \Delta v} + \frac{V_{m,j}(\tau) - V_{m,j-1}(\tau) - (V_{m-1,j}(\tau) - V_{m-1,j-1}(\tau))}{\Delta s \Delta v}\right],$$

and if $\bar{\rho} > 0$, we use

$$\frac{\partial^2 V_{m,j}(\tau)}{\partial s \partial v} \approx \frac{1}{2}\left[\frac{V_{m+1,j}(\tau) - V_{m+1,j-1}(\tau) - (V_{m,j}(\tau) - V_{m,j-1}(\tau))}{\Delta s \Delta v} + \frac{V_{m,j+1}(\tau) - V_{m,j}(\tau) - (V_{m-1,j+1}(\tau) - V_{m-1,j}(\tau))}{\Delta s \Delta v}\right].$$

This choice of this approximation for the cross partial derivative helps to produce non-positive off-diagonal entries in the coefficient matrix corresponding to the finite difference approximation of the IPDE, approximating the M-matrix property as much as possible.[8] Similar to the upwinding difference approximation for $\frac{\partial V}{\partial v}$, both cases of the approximation of the cross partial derivative term can be written compactly as

$$\begin{aligned}
&\omega(\sigma_1\rho_1 - \sigma_2\rho_2)v_m s_j \frac{\partial^2 V_{m,j}(\tau)}{\partial s \partial v} = \omega\sigma\bar{\rho} v_m s_j \frac{\partial^2 V_{m,j}(\tau)}{\partial s \partial v} \\
&\quad \approx \omega\sigma v_m s_j \max\{\bar{\rho}, 0\}\frac{1}{2}\left[\frac{V_{m+1,j}(\tau) - V_{m+1,j-1}(\tau) - (V_{m,j}(\tau) - V_{m,j-1}(\tau))}{\Delta s \Delta v} + \frac{V_{m,j+1}(\tau) - V_{m,j}(\tau) - (V_{m-1,j+1}(\tau) - V_{m-1,j}(\tau))}{\Delta s \Delta v}\right] \\
&\quad + \omega\sigma v_m s_j \min\{\bar{\rho}, 0\}\frac{1}{2}\left[\frac{V_{m+1,j+1}(\tau) - V_{m+1,j}(\tau) - (V_{m,j+1}(\tau) - V_{m,j}(\tau))}{\Delta s \Delta v} + \frac{V_{m,j}(\tau) - V_{m,j-1}(\tau) - (V_{m-1,j}(\tau) - V_{m-1,j-1}(\tau))}{\Delta s \Delta v}\right]
\end{aligned} \tag{5.21}$$

The integral terms are also approximated using the same Gauss-Hermite quadrature scheme in

[8] A matrix is an M-matrix if and only if it has non-positive off-diagonal elements and all of its real eigenvalues are positive. It suffices for a matrix to be diagonally dominant with positive diagonal elements and non-positive off-diagonal elements for it to be an M-matrix. In the component-wise splitting method of Ikonen and Toivanen (2007), the M-matrix property is important to prevent oscillations in the numerical solution due to the spatial discretization. However, as inferred by Kang and Meyer (2014), this property does not seem to be as crucial for the PSOR method.

the MOL. They shall also be treated as part of the source term in the discretized equation.

With these approximations for the partial derivatives, the spatial discretization of the IPDE (5.4) at $s = s_j$, $1 \le j \le J-1$, and $v = v_m$, $1 \le m \le M-1$, is given by

$$
\begin{aligned}
0 = V_{m,j}(\tau)\Bigg[& -\frac{\sigma^2 v_m s_j^2}{(\Delta s)^2} - \frac{\omega^2 v_m}{(\Delta v)^2} - \frac{\omega\sigma v_m s_j \max\{\bar\rho,0\}}{\Delta s \Delta v} + \frac{\omega\sigma v_m s_j \min\{\bar\rho,0\}}{\Delta s \Delta v} \\
& - \frac{\max\{\xi\eta - (\xi+\Lambda)v_m, 0\}}{\Delta v} + \frac{\min\{\xi\eta - (\xi+\Lambda)v_m, 0\}}{\Delta v} - (\tilde\lambda_1 + \tilde\lambda_2)\Bigg] \\
& + V_{m,j+1}(\tau)\left[\frac{\sigma^2 v_m s_j^2}{2(\Delta s)^2} + \frac{\omega\sigma v_m s_j \max\{\bar\rho,0\}}{2\Delta s\Delta v} - \frac{\omega\sigma v_m s_j \min\{\bar\rho,0\}}{2\Delta s\Delta v} - \frac{(\tilde\lambda_1\tilde\kappa_1 + \tilde\lambda_2\tilde\kappa_2^-)s_j}{2\Delta s}\right] \\
& + V_{m,j-1}(\tau)\left[\frac{\sigma^2 v_m s_j^2}{2(\Delta s)^2} + \frac{\omega\sigma v_m s_j \max\{\bar\rho,0\}}{2\Delta s\Delta v} - \frac{\omega\sigma v_m s_j \min\{\bar\rho,0\}}{2\Delta s\Delta v} + \frac{(\tilde\lambda_1\tilde\kappa_1 + \tilde\lambda_2\tilde\kappa_2^-)s_j}{2\Delta s}\right] \\
& + V_{m+1,j}(\tau)\left[\frac{\omega^2 v_m}{2(\Delta v)^2} + \frac{\omega\sigma v_m s_j \max\{\bar\rho,0\}}{2\Delta s\Delta v} - \frac{\omega\sigma v_m s_j \min\{\bar\rho,0\}}{2\Delta s\Delta v} + \frac{\max\{\xi\eta - (\xi+\Lambda)v_m,0\}}{\Delta v}\right] \\
& + V_{m-1,j}(\tau)\left[\frac{\omega^2 v_m}{2(\Delta v)^2} + \frac{\omega\sigma v_m s_j \max\{\bar\rho,0\}}{2\Delta s\Delta v} - \frac{\omega\sigma v_m s_j \min\{\bar\rho,0\}}{2\Delta s\Delta v} - \frac{\min\{\xi\eta - (\xi+\Lambda)v_m,0\}}{\Delta v}\right] \\
& + V_{m+1,j-1}(\tau)\left[-\frac{\omega\sigma v_m s_j \max\{\bar\rho,0\}}{2\Delta s\Delta v}\right] + V_{m-1,j+1}(\tau)\left[-\frac{\omega\sigma v_m s_j \max\{\bar\rho,0\}}{2\Delta s\Delta v}\right] \\
& + V_{m+1,j+1}(\tau)\left[\frac{\omega\sigma v_m s_j \min\{\bar\rho,0\}}{2\Delta s\Delta v}\right] + V_{m-1,j-1}(\tau)\left[\frac{\omega\sigma v_m s_j \min\{\bar\rho,0\}}{2\Delta s\Delta v}\right] \\
& - \frac{\partial V_{m,j}(\tau)}{\partial\tau} - I_1(\tau, s_j, v_m) - I_2(\tau, s_j, v_m).
\end{aligned}
\tag{5.22}
$$

Boundaries of the Computational Domain: At the boundaries $s = s_0 = 0$ and $s = s_J$, the discounted option price is known analytically as

$$
\begin{aligned}
& V_{m,0}(\tau) = \phi(\tau, 0, v_m) = 0, \\
& V_{m,J}(\tau) = \phi(\tau, s_J, v_m) = s_J e^{-q_1(T-\tau)} - e^{-q_2(T-\tau)}, \qquad \forall m = 0, 1, \ldots, M.
\end{aligned}
$$

For the second condition above to hold, we assume that s_J is chosen sufficiently large. With respect to v, we have the Neumann boundary condition

$$
\frac{\partial V(\tau_n, s, v_M)}{\partial v} = 0.
$$

The known option prices at $s = s_0 = 0$ and $s = s_J$ are incorporated into the source term for equations corresponding to $j = 1$ and $j = J-1$, whereas the approximation has to be reworked to accommodate the boundary data at $v = v_0$ and $v = v_M$. The modifications to equation (5.22) at $v = v_0$ and $v = v_M$ are detailed below.

- At $v = v_0 = 0$, $1 \leq j \leq J-1$: Since the Feller condition is assumed to hold, Fichera theory (see Meyer, 2015, Example 1.15) implies that the IPDE should still hold when $v = v_0 = 0$. This results to the equation

$$\begin{aligned} 0 = &-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)s_j\frac{\partial V_{0,j}(\tau)}{\partial s} + \xi\eta\frac{\partial V_{0,j}(\tau)}{\partial v} - (\tilde{\lambda}_1 + \tilde{\lambda}_2)V_{0,j}(\tau) \\ &- \frac{\partial V_{0,j}(\tau)}{\partial \tau} - I_1(\tau, s_j, v_0) - I_2(\tau, s_j, v_0). \end{aligned} \tag{5.23}$$

  With the usual central difference approximation for $\frac{\partial V_{0,j}(\tau)}{\partial s}$ and a forward difference approximation for $\frac{\partial V_{0,j}(\tau)}{\partial v}$, the corresponding semi-discrete equation is

$$\begin{aligned} 0 = &V_{0,j}(\tau)\left[-(\tilde{\lambda}_1 + \tilde{\lambda}_2) - \frac{\xi\eta}{\Delta v}\right] + V_{0,j-1}(\tau)\left[\frac{(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)s_j}{2\Delta s}\right] \\ &+ V_{0,j+1}(\tau)\left[-\frac{(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)s_j}{2\Delta s}\right] + V_{1,j}(\tau)\left[\frac{\xi\eta}{\Delta v}\right] \\ &- \frac{\partial V_{0,j}(\tau)}{\partial \tau} - I_1(\tau, s_j, v_0) - I_2(\tau, s_j, v_0). \end{aligned} \tag{5.24}$$

- At $v = v_M$, $1 \leq j \leq J-1$: We approximate the Neumann boundary condition (5.7) with a central difference (instead of the backward difference approximation used in the MOL implementation), implying that $V_{M+1,j}(\tau) = V_{M-1,j}(\tau)$ for $1 \leq j \leq J-1$. This allows us to eliminate from (5.22) any references to $V_{M+1,j}(\tau)$ which lies outside the computational domain.

5.3.2 The Semi-Discrete IPDE

The sequence of $(M+1)(J-1)$ equations corresponding to the semi-discrete IPDE can be written more compactly in matrix form as

$$\frac{\partial \mathbf{V}(\tau)}{\partial \tau} - \mathbf{F}\mathbf{V}(\tau) = \mathbf{b}(\tau), \tag{5.25}$$

where $\mathbf{V}(\tau)$ is an $(M+1)(J-1)$ vector containing the unknown option price approximations at each $s = s_j$ and $v = v_m$ at time-to-maturity τ,

$$\mathbf{V}(\tau) = \begin{bmatrix} \mathbf{V}_0(\tau) \\ \mathbf{V}_1(\tau) \\ \vdots \\ \mathbf{V}_M(\tau) \end{bmatrix} = \begin{bmatrix} V_{0,1}(\tau) \\ \vdots \\ V_{0,J-1}(\tau) \\ V_{1,1}(\tau) \\ \vdots \\ V_{1,J-1}(\tau) \\ \vdots \\ V_{M,1}(\tau) \\ \vdots \\ V_{M,J-1}(\tau) \end{bmatrix},$$

$\mathbf{F}$ is an $(M+1)(J-1) \times (M+1)(J-1)$ block tridiagonal matrix, and $\mathbf{b}(\tau)$ is an $(M+1)(J-1)$ vector dependent on τ (indexed in the same manner as $\mathbf{V}(\tau)$). Option prices at $s = s_0$ and $s = s_J$ are excluded from the vector $\mathbf{V}$ as they are already known.

The coefficient matrix $\mathbf{F}$ has the form

$$\mathbf{F} = \begin{bmatrix} \mathbf{B}^0 & \mathbf{C}^1 & \mathbf{O} & \cdots & & & & \\ \mathbf{A}^0 & \mathbf{B}^1 & \mathbf{C}^2 & \mathbf{O} & \cdots & & & \\ \mathbf{O} & \mathbf{A}^1 & \mathbf{B}^2 & \mathbf{C}^3 & \mathbf{O} & \cdots & & \\ & & \ddots & \ddots & \ddots & & & \\ & & \cdots & \mathbf{O} & \mathbf{A}^{M-3} & \mathbf{B}^{M-2} & \mathbf{C}^{M-1} & \mathbf{O} \\ & & & \cdots & \mathbf{O} & \mathbf{A}^{M-2} & \mathbf{B}^{M-1} & \mathbf{C}^M \\ & & & & \cdots & \mathbf{O} & \mathbf{A}^{M-1} & \mathbf{B}^M \end{bmatrix}, \tag{5.26}$$

where $\mathbf{A}^m = [A^m_{j,k}]$, $\mathbf{B}^m = [B^m_{j,k}]$, and $\mathbf{C}^m = [C^m_{j,k}]$ are $(J-1) \times (J-1)$ matrices and $\mathbf{O}$ is the $(J-1) \times (J-1)$ zero matrix. With this notation, in the row of $V_{m,j}(\tau)$, $A^{m-1}_{j,k}$ is the coefficient of $V_{m-1,k}(\tau)$, $B^m_{j,k}$ is the coefficient of $V_{m,k}(\tau)$, and $C^{m+1}_{j,k}$ is the coefficient of $V_{m+1,k}(\tau)$, $k = 0, 1, \ldots, J$. As evident in the spatial discretization of the IPDE, each sub-matrix is tridiagonal, so for $0 \le m \le M$ and $0 \le j \le J$,

$$A^m_{j,k} = B^m_{j,k} = C^m_{j,k} = 0 \qquad \text{whenever } |j-k| \ge 2.$$

The next step is to identify the entries of each of the sub-matrices appearing in $\mathbf{F}$. To this end, we discuss separately the matrices corresponding to the interior and the boundaries of the computational domain.

Interior of the Computational Domain: From approximation (5.22) of the IPDE in the interior of the domain, we identify the coefficients as

$$\begin{aligned} A^m_{j,j} &= \frac{\omega^2 v_{m+1}}{2(\Delta v)^2} + \frac{\omega\sigma v_{m+1} s_j \max\{\bar\rho, 0\}}{2\Delta s \Delta v} - \frac{\omega\sigma v_{m+1} s_j \min\{\bar\rho, 0\}}{2\Delta s \Delta v} \\ &\quad - \frac{\min\{\xi\eta - (\xi+\Lambda) v_{m+1}, 0\}}{\Delta v}, \\ A^m_{j,j+1} &= -\frac{\omega\sigma v_{m+1} s_j \max\{\bar\rho, 0\}}{2\Delta s \Delta v}, \\ A^m_{j,j-1} &= \frac{\omega\sigma v_{m+1} s_j \min\{\bar\rho, 0\}}{2\Delta s \Delta v}, \end{aligned} \tag{5.27}$$

for $0 \le m \le M-2$ and $1 \le j \le J-1$,

$$\begin{aligned} B^m_{j,j} &= -\frac{\sigma^2 v_m s_j^2}{(\Delta s)^2} - \frac{\omega^2 v_m}{(\Delta v)^2} - \frac{\omega\sigma v_m s_j \max\{\bar\rho, 0\}}{\Delta s \Delta v} + \frac{\omega\sigma v_m s_j \min\{\bar\rho, 0\}}{\Delta s \Delta v} \\ &\quad - \frac{\max\{\xi\eta - (\xi+\Lambda) v_m, 0\}}{\Delta v} + \frac{\min\{\xi\eta - (\xi+\Lambda) v_m, 0\}}{\Delta v} - (\tilde\lambda_1 + \tilde\lambda_2), \\ B^m_{j,j+1} &= \frac{\sigma^2 v_m s_j^2}{2(\Delta s)^2} + \frac{\omega\sigma v_m s_j \max\{\bar\rho, 0\}}{2\Delta s \Delta v} - \frac{\omega\sigma v_m s_j \min\{\bar\rho, 0\}}{2\Delta s \Delta v} - \frac{(\tilde\lambda_1\tilde\kappa_1 + \tilde\lambda_2\tilde\kappa_2^-) s_j}{2\Delta s}, \\ B^m_{j,j-1} &= \frac{\sigma^2 v_m s_j^2}{2(\Delta s)^2} + \frac{\omega\sigma v_m s_j \max\{\bar\rho, 0\}}{2\Delta s \Delta v} - \frac{\omega\sigma v_m s_j \min\{\bar\rho, 0\}}{2\Delta s \Delta v} + \frac{(\tilde\lambda_1\tilde\kappa_1 + \tilde\lambda_2\tilde\kappa_2^-) s_j}{2\Delta s}, \end{aligned} \tag{5.28}$$

for $1 \le m \le M-1$ and $1 \le j \le J-1$, and

$$\begin{aligned} C^m_{j,j} &= \frac{\omega^2 v_{m-1}}{2(\Delta v)^2} + \frac{\omega\sigma v_{m-1} s_j \max\{\bar\rho, 0\}}{2\Delta s \Delta v} - \frac{\omega\sigma v_{m-1} s_j \min\{\bar\rho, 0\}}{2\Delta s \Delta v} \\ &\quad + \frac{\max\{\xi\eta - (\xi+\Lambda) v_{m-1}, 0\}}{\Delta v}, \\ C^m_{j,j+1} &= \frac{\omega\sigma v_{m-1} s_j \min\{\bar\rho, 0\}}{2\Delta s \Delta v}, \\ C^m_{j,j-1} &= -\frac{\omega\sigma v_{m-1} s_j \max\{\bar\rho, 0\}}{2\Delta s \Delta v}, \end{aligned} \tag{5.29}$$

for $2 \le m \le M$ and $1 \le j \le J-1$.

At the Boundaries of v: As before, we consider each part of the boundary separately.

- At $v = v_0 = 0$, $1 \le j \le J-1$: The semi-discrete equation (5.24) implies that $\mathbf{B}^0$ has entries given by

$$B^0_{j,j} = -(\tilde\lambda_1 + \tilde\lambda_2) - \frac{\xi\eta}{\Delta v}, \quad B^0_{j,j-1} = \frac{(\tilde\lambda_1\tilde\kappa_1 + \tilde\lambda_2\tilde\kappa_2^-) s_j}{2\Delta s}, \quad B^0_{j,j+1} = -\frac{(\tilde\lambda_1\tilde\kappa_1 + \tilde\lambda_2\tilde\kappa_2^-) s_j}{2\Delta s},$$

for all $1 \leq j \leq J-1$, and that

$$C^1_{j,j} = \frac{\xi\eta}{\Delta v}, \qquad 1 \leq j \leq J-1$$

with all other coefficients of $\mathbf{C}^1$ being zero.

- At $v = v_M$, $1 \leq j \leq J-1$: The central difference approximation of the Neumann boundary condition at $v = v_M$ implies that the coefficients of $V_{M+1,j}(\tau)$ are then combined to those of $V_{M-1,j}(\tau)$, giving us

$$\begin{aligned}
A^{M-1}_{j,j} &= \frac{\omega^2 v_M}{(\Delta v)^2} + \frac{\omega\sigma v_M s_j \max\{\bar{\rho},0\}}{\Delta s \Delta v} - \frac{\omega\sigma v_M s_j \min\{\bar{\rho},0\}}{\Delta s \Delta v} \\
&\quad + \frac{\max\{\xi\eta - (\xi+\Lambda)v_M, 0\}}{\Delta v} - \frac{\min\{\xi\eta - (\xi+\Lambda)v_M, 0\}}{\Delta v}, \\
A^{M-1}_{j,j-1} &= \frac{\omega\sigma v_M s_j \min\{\bar{\rho},0\}}{2\Delta s \Delta v} - \frac{\omega\sigma v_M s_j \max\{\bar{\rho},0\}}{2\Delta s \Delta v}, \\
A^{M-1}_{j,j+1} &= \frac{\omega\sigma v_M s_j \min\{\bar{\rho},0\}}{2\Delta s \Delta v} - \frac{\omega\sigma v_M s_j \max\{\bar{\rho},0\}}{2\Delta s \Delta v},
\end{aligned}$$

  for $1 \leq j \leq J-1$. The coefficients $B^M_{j,j}$, $B^M_{j,j-1}$, and $B^M_{j,j+1}$ are defined in the same manner as for $\mathbf{B}^m$, $1 \leq m \leq M-1$, as given by equation (5.28).

In defining the sub-matrices $\mathbf{A}^m$, $\mathbf{B}^m$, and $\mathbf{C}^m$ appearing in equation (5.26), we only use the entries $A^m_{j,k}$, $B^m_{j,k}$, and $C^m_{j,k}$ for which $1 \leq k \leq J-1$. Entries with $k = 0$ (when $j = 1$) or $k = J$ (when $j = J-1$) are not used in the coefficient matrix $\mathbf{F}$ itself, but are used in defining the inhomogenous vector terms $b_{m,1}(\tau)$ and $b_{m,J-1}(\tau)$, as discussed below.

Inhomogenous Term: The elements of the inhomogenous vector term $\mathbf{b}(\tau)$ consist of the integral terms and, in the case where $j = J-1$, some terms dependent on the payoff function.[9] When the semi-discrete equation does not involve the $V_{m,0}(\tau)$ and $V_{m,J}(\tau)$, i.e. for $2 \leq j \leq J-2$, the elements of $\mathbf{b}$ are given by

$$b_{m,j}(\tau) = -I_1(\tau, s_j, v_m) - I_2(\tau, s_j, v_m), \qquad \text{for } 2 \leq j \leq J-2 \text{ and } 0 \leq m \leq M.$$

We make a separate analysis of the coefficients at $j = 1$ and $j = J-1$ as the known option prices at these boundaries are to be incorporated in $b_{m,1}(\tau)$ and $b_{m,J-1}(\tau)$.

- $j = 1$, $1 \leq m \leq M-1$: In this case, equation (5.22) becomes

$$\begin{aligned}
0 = {} & V_{m,1}(\tau)B^m_{1,1} + V_{m,2}(\tau)B^m_{1,2} + V_{m,0}(\tau)B^m_{1,0} \\
& + V_{m+1,1}(\tau)C^{m+1}_{1,1} + V_{m+1,2}(\tau)C^{m+1}_{1,2} + V_{m+1,0}(\tau)C^{m+1}_{1,0}
\end{aligned}$$

[9] If the payoff function is nonzero at $j = 0$, then $b_{m,1}(\tau)$ will also be defined differently from the case where $2 \leq j \leq J-2$. Nonetheless, we detail here how the inhomogenous term is characterized when $j = 0$.

$$+ V_{m-1,1}(\tau)A_{1,1}^{m-1} + V_{m-1,2}(\tau)A_{1,2}^{m-1} + V_{m-1,0}(\tau)A_{1,0}^{m-1}$$
$$- \frac{\partial V_{m,1}(\tau)}{\partial \tau} - I_1(\tau, s_1, v_m) - I_2(\tau, s_1, v_m).$$

Since $V_{m,0}(\tau)$, $V_{m+1,0}(\tau)$, and $V_{m-1,0}(\tau)$ are known and are equal to zero, it follows that

$$b_{m,1}(\tau) = -I_1(\tau, s_1, v_m) - I_2(\tau, s_1, v_m), \qquad 1 \le m \le M-1.$$

- $j = 1$, $m = 0$: A similar calculation, this time on equation (5.24), gives

$$0 = V_{0,1}(\tau)B_{1,1}^0 + V_{0,2}(\tau)B_{1,2}^0 + V_{0,0}(\tau)B_{1,0}^0 + V_{1,1}(\tau)C_{1,1}^1$$
$$- \frac{\partial V_{0,1}(\tau)}{\partial \tau} - I_1(\tau, s_1, v_0) - I_2(\tau, s_1, v_0).$$

After substituting $V_{0,0}(\tau) = 0$, we obtain the inhomogenous term

$$b_{0,1}(\tau) = -I_1(\tau, s_1, v_0) - I_2(\tau, s_1, v_0).$$

- $j = 1$, $m = M$: Equation (5.22), with $V_{M+1,1}(\tau) = V_{M-1,1}(\tau)$, becomes

$$0 = V_{M,1}(\tau)B_{1,1}^M + V_{M,2}(\tau)B_{1,2}^M + V_{M,0}(\tau)B_{1,0}^M$$
$$+ V_{M-1,1}(\tau)A_{1,1}^{M-1} + V_{M-1,2}(\tau)A_{1,2}^{M-1} + V_{M-1,0}(\tau)A_{1,0}^{M-1}$$
$$- \frac{\partial V_{M,1}(\tau)}{\partial \tau} - I_1(\tau, s_1, v_M) - I_2(\tau, s_1, v_M).$$

Using the boundary data at $s = s_0$ to eliminate $V_{M,0}(\tau)$ and $V_{M-1,0}(\tau)$ from above, we obtain

$$b_{M,1}(\tau) = -I_1(\tau, s_1, v_M) - I_2(\tau, s_1, v_M).$$

- $j = J-1$, $1 \le m \le M-1$: Equation (5.22) becomes

$$0 = V_{m,J-1}(\tau)B_{J-1,J-1}^m + V_{m,J}(\tau)B_{J-1,J}^m + V_{m,J-2}(\tau)B_{J-1,J-2}^m$$
$$+ V_{m+1,J-1}(\tau)C_{J-1,J-1}^{m+1} + V_{m+1,J}(\tau)C_{J-1,J}^{m+1} + V_{m+1,J-2}(\tau)C_{J-1,J-2}^{m+1}$$
$$+ V_{m-1,J-1}(\tau)A_{J-1,J-1}^{m-1} + V_{m-1,J}(\tau)A_{J-1,J}^{m-1} + V_{m-1,J-2}(\tau)A_{J-1,J-2}^{m-1}$$
$$- \frac{\partial V_{m,J-1}(\tau)}{\partial \tau} - I_1(\tau, s_{J-1}, v_m) - I_2(\tau, s_{J-1}, v_m).$$

Since $V_{m,J}(\tau)$, $V_{m+1,J}(\tau)$, and $V_{m-1,J}(\tau)$ are all equal to $\phi(\tau, s_J, v_m)$, the inhomogenous term

is given by

$$\begin{aligned} b_{m,J-1}(\tau) &= \phi(\tau, s_J, v_m)\left[A^{m-1}_{J-1,J} + B^{m}_{J-1,J} + C^{m+1}_{J-1,J}\right] - I_1(\tau, s_{J-1}, v_m) - I_2(\tau, s_{J-1}, v_m) \\ &= \phi(\tau, s_J, v_m)\left[\frac{\sigma^2 v_m s^2_{J-1}}{2(\Delta s)^2} - \frac{(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)s_{J-1}}{2\Delta s}\right] \\ &\quad - I_1(\tau, s_{J-1}, v_m) - I_2(\tau, s_{J-1}, v_m), \qquad 1 \le m \le M-1. \end{aligned}$$

- $\underline{j = J-1,\ m = 0}$: Replacing $V_{0,J}(\tau)$ with $\phi(\tau, s_J, v_0)$ in equation (5.24) when $j = J-1$ produces the equation

$$\begin{aligned} 0 = {} & V_{0,J-1}(\tau)B^0_{J-1,J-1} + V_{0,J-2}(\tau)B^0_{J-1,J-2} + V_{1,J-1}(\tau)C^1_{J-1,J-1} \\ & - \frac{\partial V_{0,J-1}(\tau)}{\partial \tau} + b_{0,J-1}(\tau), \end{aligned}$$

where the inhomogenous term is given by

$$\begin{aligned} b_{0,J-1}(\tau) &= \phi(\tau, s_J, v_0)B^0_{J-1,J} - I_1(\tau, s_{J-1}, v_0) - I_2(\tau, s_{J-1}, v_0) \\ &= -\phi(\tau, s_J, v_0)\frac{(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)s_{J-1}}{2\Delta s} - I_1(\tau, s_{J-1}, v_0) - I_2(\tau, s_{J-1}, v_0) \end{aligned}$$

- $\underline{j = J-1,\ m = M}$: Applying the boundary data $V_{M+1,j}(\tau) = V_{M-1,j}(\tau)$ and $V_{M,J}(\tau) = V_{M-1,J}(\tau) = \phi(\tau, s_J, v_M)$ to equation (5.22) gives us

$$\begin{aligned} 0 = {} & V_{M,J-1}(\tau)B^M_{J-1,J-1} + V_{M,J-2}(\tau)B^M_{J-1,J-2} + V_{M-1,J-1}(\tau)A^{M-1}_{J-1,J-1} \\ & + V_{M-1,J-2}(\tau)A^{M-1}_{J-1,J-2} - \frac{\partial V_{M,J-1}(\tau)}{\partial \tau} + b_{M,J-1}(\tau), \end{aligned}$$

where $b_{M,J-1}(\tau)$ is given by

$$\begin{aligned} b_{M,J-1}(\tau) &= \phi(\tau, s_J, v_m)\left[A^{M-1}_{J-1,J} + B^{M}_{J-1,J}\right] - I_1(\tau, s_{J-1}, v_M) - I_2(\tau, s_{J-1}, v_M) \\ &= \phi(\tau, s_J, v_M)\left[\frac{\sigma^2 v_M s^2_{J-1}}{2(\Delta s)^2} - \frac{(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)s_{J-1}}{2\Delta s}\right] \\ &\quad - I_1(\tau, s_{J-1}, v_M) - I_2(\tau, s_{J-1}, v_M). \end{aligned}$$

5.3.3 Completing the Finite Difference Approximation

Finally, we complete the finite difference approximation of the semi-discrete IPDE (5.25) by using a θ-scheme for $\frac{\partial \mathbf{V}(\tau)}{\partial \tau}$ over the partition $0 = \tau_0 < \tau_1 < \cdots < \tau_N = T$ of $[0, T]$ with uniform width

$\Delta\tau$. Specifically, the approximation over the sub-interval $[\tau_{n-1}, \tau_n]$ is

$$\frac{\mathbf{V}(\tau_n) - \mathbf{V}(\tau_{n-1})}{\Delta\tau} = \theta\left[\mathbf{F}\mathbf{V}(\tau_n) + \mathbf{b}(\tau_n)\right] + (1-\theta)\left[\mathbf{F}\mathbf{V}(\tau_{n-1}) + \mathbf{b}(\tau_{n-1})\right],$$

where $\theta \in [0,1]$. This can be rearranged to yield

$$(\mathbf{I} - \theta\Delta\tau\mathbf{F})\mathbf{V}(\tau_n) - \theta\Delta\tau\mathbf{b}(\tau_n) = (\mathbf{I} + (1-\theta)\Delta\tau\mathbf{F})\mathbf{V}(\tau_{n-1}) + (1-\theta)\Delta\tau\mathbf{b}(\tau_{n-1}), \tag{5.30}$$

where I is the $(M+1)(J-1) \times (M+1)(J-1)$ identity matrix.

The discrete form of the LCP (5.20) is therefore given by[10]

$$\begin{aligned} \mathbf{G}\mathbf{V}(\tau_n) + \mathbf{g}(\tau_n) - \mathbf{H}\mathbf{V}(\tau_{n-1}) - \mathbf{h}(\tau_{n-1}) &\geq 0, \\ \mathbf{V}(\tau_n) - \boldsymbol{\phi}(\tau_n) &\geq 0, \\ \left(\mathbf{G}\mathbf{V}(\tau_n) + \mathbf{g}(\tau_n) - \mathbf{H}\mathbf{V}(\tau_{n-1}) - \mathbf{h}(\tau_{n-1})\right)^\top \left(\mathbf{V}(\tau_n) - \boldsymbol{\phi}(\tau_n)\right) &= 0 \end{aligned} \tag{5.31}$$

where $\mathbf{G} = \mathbf{I} - \theta\Delta\tau\mathbf{F}$, $\mathbf{g}(\tau_n) = -\theta\Delta\tau\mathbf{b}(\tau_n)$, $\mathbf{H} = \mathbf{I} + (1-\theta)\Delta\tau\mathbf{F}$, $\mathbf{h}(\tau_{n-1}) = (1-\theta)\Delta\tau\mathbf{b}(\tau_{n-1})$, and $\boldsymbol{\phi}(\tau_n)$ is a $(M+1)(J-1)$-vector (indexed in the same manner as $\mathbf{V}(\tau_n)$) whose entries are equal to the value of the payoff function evaluated at $\tau = \tau_n$, $s = s_j$ for $0 \leq j \leq J$, and $v = v_m$ for $0 \leq m \leq M$. The inequalities for vectors are understood componentwise.

Using a pure Crank-Nicholson scheme ($\theta = 1/2$) does not damp out fast enough the oscillations in the solution due to the non-smooth payoff function dictating the initial value of the solution. One possible remedy is to use an implicit scheme ($\theta = 1$) for the first three time steps and then switch to a Crank-Nicholson scheme thereafter. This method, called the Rannacher time stepping scheme, reduces the oscillations while maintaining the second-order accuracy with respect to τ afforded by the Crank-Nicholson scheme (Rannacher, 1984; Ikonen and Toivanen, 2007). With the discretized version (5.31) of the LCP (5.20), we can now implement the PSOR algorithm to approximate the solution of the LCP.

[10] The semi-discrete approximation of the IPDE operation $\mathscr{L}V(\tau, s, v)$ at $s = s_j$ and $v = v_m$ is

$$\mathscr{L}V(\tau, s, v) \approx \mathbf{F}\mathbf{V}(\tau) - \frac{\partial\mathbf{V}(\tau)}{\partial\tau} + \mathbf{b}(\tau),$$

and so the semi-discrete approximation of the integro-partial differential inequality $\mathscr{L}V(\tau, s, v) \leq 0$ is

$$\frac{\partial\mathbf{V}(\tau)}{\partial\tau} \geq \mathbf{F}\mathbf{V}(\tau) + \mathbf{b}(\tau).$$

Upon discretizing with respect to τ, the last inequality above is equivalent to

$$(\mathbf{I} - \theta\Delta\tau\mathbf{F})\mathbf{V}(\tau_n) - \theta\Delta\tau\mathbf{b}(\tau_n) \geq (\mathbf{I} + (1-\theta)\Delta\tau\mathbf{F})\mathbf{V}(\tau_{n-1}) + (1-\theta)\Delta\tau\mathbf{b}(\tau_{n-1})$$

5.3.4 Overview of the PSOR Algorithm for General LCPs

We first review the basic PSOR algorithm for general LCPs. The PSOR algorithm is an iterative method to solve linear complementarity problems of the form

$$\mathbf{A}\mathbf{x} \geq \mathbf{b}, \qquad \mathbf{x} \geq \mathbf{c}, \qquad (\mathbf{A}\mathbf{x} - \mathbf{b})^\top (\mathbf{x} - \mathbf{c}) = 0, \tag{5.32}$$

where $\mathbf{A} = [A_{ij}] \in \mathbb{R}^{L \times L}$, $\mathbf{x}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^L$. The succeeding discussion is based mostly on Wilmott et al. (1993).

The algorithm proceeds as follows. Let k be the iteration counter. We start with an initial iterate $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_L^0)^\top \geq \mathbf{c}$. In the kth iteration, $k \geq 1$, we form the vector $\mathbf{x}^k = (x_1^k, x_2^k, \ldots, x_L^k)^\top$ *componentwise* from the vector $\mathbf{x}^{k-1}$ from the previous iteration. For each $l = 1, 2, \ldots, L$, we compute the intermediate value

$$y_l^k = \frac{1}{A_{ll}} \left[b_l - \sum_{j=1}^{l-1} A_{lj} x_j^k - \sum_{j=l+1}^{L} A_{lj} x_j^{k-1} \right]. \tag{5.33}$$

Then we compute x_l^k as

$$x_l^k = \max \left\{ c_l, x_l^{k-1} + w(y_l^k - x_l^{k-1}) \right\}, \tag{5.34}$$

where $w > 0$ is a constant. The parameter w is known as the *relaxation parameter* and it must be chosen to accelerate the convergence of the iterative process. The iterations stop at the iteration K at which

$$\max_l |x_l^K - x_l^{K-1}| < \epsilon$$

is satisfied for some error tolerance $\epsilon > 0$. We then take $\mathbf{x}^K$ as the solution of the LCP. This process is summarized in Algorithm 3.

Algorithm 3: PSOR for LCP (5.32)

Result: Solution $\mathbf{x}$ of LCP (5.32)
Inputs: Matrix $\mathbf{A}$, vectors $\mathbf{b}, \mathbf{c}$, error $e > 0$, relaxation parameter w, maximum iteration count k_{max};
Choose initial iterate $\mathbf{x}^{k-1} \geq \mathbf{c}$;
while $\epsilon > 10^{-8}$ *&* $k < k_{max}$ **do**
  for $l = 1$ **to** L **do**
    Compute y_l^k using equation (5.33);
    Compute x_l^k using equation (5.34);
  end
  Update $e = \max_l |x_l^k - x_l^{k-1}|$;
  Update $k = k + 1$;
end

The following theorem provides sufficient conditions for the convergence of the PSOR (Elliott and Ockendon, 1982, Theorem IV.7).

Theorem 5.1. *If A is symmetric positive definite and $0 < w < 2$, then the PSOR algorithm converges.*

5.3.5 The PSOR for LCP (5.31)

At each time step, we solve LCP (5.31) using the PSOR outline in Algorithm 3. The solution of the discrete LCP (5.31) is stored in the $(M+1)(J-1) \times (N+1)$ matrix $\mathbf{V}$,

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}(\tau_0) & \mathbf{V}(\tau_1) & \cdots & \mathbf{V}(\tau_{N-1}) & \mathbf{V}(\tau_N) \end{bmatrix},$$

where $\mathbf{V}(\tau_n)$ is the $(M+1)(J-1)$-vector containing the solution of the LCP at the nth time step. The LCP also requires the values of the payoff function at each time step, so we store these values in the $(M+1)(J-1) \times (N+1)$ matrix $\mathbf{\Phi}$,

$$\mathbf{\Phi} = \begin{bmatrix} \boldsymbol{\phi}(\tau_0) & \boldsymbol{\phi}(\tau_1) & \cdots & \boldsymbol{\phi}(\tau_{N-1}) & \boldsymbol{\phi}(\tau_N) \end{bmatrix},$$

where $\boldsymbol{\phi}(\tau_n)$ is defined following LCP (5.31). The matrix $\mathbf{\Phi}$ can be calculated outside the for-loop in n. Additionally, $\boldsymbol{\phi}(\tau_0)$ is stored as the first column of $\mathbf{V}$ representing the initial condition of the LCP. We outline in Algorithm 4 the modified implementation of the PSOR.

The coefficient matrix $\mathbf{G}$ is not necessarily symmetric, implying that Theorem 5.1 does not apply to the discretized LCP (5.31). However, Huang and Pang (1998) have established an alternative set of sufficient conditions that guarantee the converges of the PSOR for LCPs with non-symmetric coefficient matrices. Due to the use of upwind difference approximations and the non-standard approximation of $\frac{\partial^2 V}{\partial s \partial v}$, $\mathbf{G}$ has positive diagonal entries. Furthermore, for sufficiently small time steps $\Delta\tau$, $\mathbf{G}$ is also strictly diagonally dominant. Therefore, the LCP (5.31) has a unique solution at each time step and the PSOR method, when applied to (5.31), converges for all $w \in (0, \bar{w})$, where $\bar{w} \in (1, 2]$ is computed based on the entries of $\mathbf{G}$. To compute the upper bound $\bar{w}$ of the relaxation parameter, we first construct $\bar{\mathbf{G}} = [\bar{G}_{ij}]$, the *comparison matrix* of $\mathbf{G}$ defined as

$$\bar{G}_{ij} = \begin{cases} |G_{ij}| & \text{if } i = j \\ -|G_{ij}| & \text{otherwise.} \end{cases}$$

Then, $\bar{w}$ is calculated as

$$\bar{w} = 2 \min_i \left\{ \frac{G_{ii} d_i}{\sum_j |G_{ij}| d_j} \right\}, \tag{5.35}$$

where $\mathbf{d}$ is an $(M+1)(J+1)$ vector such that $\bar{\mathbf{G}}\mathbf{d} > 0$. Such a vector $\mathbf{d}$ exists due to the strict diagonal dominance of $\mathbf{G}$.

Once the spatial discretization coefficient matrix $\mathbf{F}$ (5.26) has been computed, we calculate two

versions of the coefficient matrices $\mathbf{G}$ and $\mathbf{H}$ in view of the Rannacher time stepping scheme. The coefficient matrices under the implicit scheme for $n \leq 3$ are given by

$$\mathbf{G}_1 = \mathbf{I} - \Delta\tau\mathbf{F} \qquad \mathbf{H}_1 = \mathbf{I},$$

while, under the Crank-Nicholson scheme for $4 \leq n \leq N$, the coefficient matrices are

$$\mathbf{G}_2 = \mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{F} \qquad \mathbf{H}_2 = \mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{F}.$$

We solve for the matrices $\mathbf{G}_1$, $\mathbf{H}_1$, $\mathbf{G}_2$, and $\mathbf{H}_2$ outside the for-loop in n since the coefficient matrices only change at $n = 4$. Given $\mathbf{G}_1$ and $\mathbf{G}_2$, we can then calculate the upper bounds of the relaxation parameters for each phase using the process described in the previous paragraph. The implementation of the PSOR algorithm to solve the discretized LCP (5.31) is detailed in Algorithm 4.

The iterative nature of the PSOR also makes in an appropriate method to solve LCP (5.31), in view of the integral terms involved in $\mathbf{b}(\tau_n)$ (present in $\mathbf{g}(\tau_n)$) that are dependent on the yet-unknown option prices $\mathbf{V}(\tau_n)$. In our implementation of the PSOR, the vector $\mathbf{b}(\tau_n)$ is iteration-dependent. Its initial iterate $\mathbf{b}^0(\tau_n)$ is calculated using the initial iterate of the LCP solution then, before the next iteration commences, we recalculate the integral terms using the recently-obtained iterate of the LCP solution. We thus favor the PSOR method over the direct elimination method for the usual LCP (5.32) (see for example, Seydel, 2017, Section 4.6.4).

Once the option price matrix $\mathbf{V}$ has been fully determined by the algorithm, we then solve for the early exercise boundary for each time step τ_n and variance line v_m. Fixing n and m, we first determine the index $k^* > 1$ at which the numerical approximation of the option price $V_{m,j}(\tau_n)$ first coincides with the payoff $\phi(\tau_n, s_j, v_m)$. Next, we fit a cubic spline interpolant $S(\tau_n, x, v_m)$ through the four nearest option prices around k^*, i.e. $S(\tau_n, x, v_m)$ is the cubic spline fitted through the points $\{(s_i, V_{m,i}(\tau_n))\}_{i=k^*-2}^{k^*+1}$. We then define the function

$$D(x) = S(\tau_n, x, v_m) - \left[xe^{-q_1(T-\tau_n)} - e^{-q_2(T-\tau_n)}\right],$$

to be the difference between the cubic spline interpolant and the intrinsic value of the option. The early exercise boundary is then approximated as the zero of $D(x)$ in the interval $[s_{k^*-1}, s_{k^*}]$, which can be found using any standard root-finding procedure.

5.4 Numerical Results and Discussion

In this section, we present numerical approximations of the discounted European and American exchange option prices, the early exercise boundary, and the discounted early exercise premium

Algorithm 4: PSOR for LCP (5.31)

Result: Solution $\mathbf{V}(\tau_n)$ of LCP (5.31) for each $n = 1, 2, \ldots, N$
Inputs: Model and discretization parameters, maximum iteration count k_{max};
/* Compute the coefficient matrices of the LCP */
Set up the matrix $\mathbf{F}$ in (5.26);
Set up the matrix $\mathbf{V}$ as an $(M+1)(J-1) \times (N+1)$ zero matrix;
Compute the payoff matrix $\mathbf{\Phi}$ and the LCP coefficient matrices $\mathbf{G}_1$, $\mathbf{H}_1$, $\mathbf{G}_2$, and $\mathbf{H}_2$;
Let $L_{\text{psor}} = (M+1)(J-1)$;
/* Determine the relaxation parameters */
Compute the comparison matrices $\bar{\mathbf{G}}_1$ and $\bar{\mathbf{G}}_2$;
Compute for $(M+1)(J-1)$-vectors $\mathbf{d}_1$ and $\mathbf{d}_2$ such that $\bar{\mathbf{G}}_1\mathbf{d}_1 > 0$ and $\bar{\mathbf{G}}_2\mathbf{d}_2 > 0$;
Calculate using (5.35) the upper bounds $\bar{w}_1$ and $\bar{w}_2$ for the implicit and CN schemes;
Choose the relaxation parameters $w_1 \in (1, \bar{w}_1)$ and $w_2 \in (1, \bar{w}_2)$;
/* Start for-loop in n */
for $n = 1, 2, \ldots, N$ **do**
  Initialize error $e = 1$ and iteration counter $k = 1$;
  Initialize $\mathbf{x}^k$ as the zero vector with length $(M+1)(J-1)$ // x is temporary storage for the solution at this time step
  Choose initial iterate $\mathbf{x}^{k-1} = \mathbf{V}(\tau_{n-1})$;
  Compute $\mathbf{b}(\tau_{n-1})$ with the values $\mathbf{V}(\tau_{n-1})$ using Gauss-Hermite quadrature;
  Set $\mathbf{c} = \boldsymbol{\phi}(\tau_n)$;
  if $n \leq 3$ **then**
    Set $\mathbf{G} = \mathbf{G}_1$, $\mathbf{H} = \mathbf{H}_1$, $\theta = 1$, and $w = w_1$;
    /* PSOR while-loop */
    while $e > 10^{-8}$ *and* $k < k_{max}$ **do**
      Calculate $\mathbf{b}^{k-1}(\tau_n)$ based on the previous iterate $\mathbf{x}^{k-1}$ using Gauss-Hermite quadrature;
      Set $\mathbf{b}' = \mathbf{H}\mathbf{V}(\tau_{n-1}) + \mathbf{h}(\tau_{n-1}) - \mathbf{g}^k(\tau_n)$, where $\mathbf{h}(\tau_{n-1}) = (1-\theta)\Delta\tau\mathbf{b}(\tau_{n-1})$ and $\mathbf{g}^k(\tau_n) = -\theta\Delta\tau\mathbf{b}^{k-1}(\tau_n)$;
      for $l = 1, 2, \ldots, L_{psor}$ **do**
        Compute $y_l^k = \frac{1}{G_{ll}}\left[\mathbf{b}'_l - \sum_{j=1}^{l-1} G_{lj}x_j^k - \sum_{j=l+1}^{L} G_{lj}x_j^{k-1}\right]$;
        Compute $x_l^k = \max\left\{c_l, x_l^{k-1} + w(y_l^k - x_l^{k-1})\right\}$;
      end
      Update $e \leftarrow \max_l |x_l^k - x_l^{k-1}|$;
      Update $\mathbf{x}^{k-1} \leftarrow \mathbf{x}^k$;
      Update $k \leftarrow k + 1$;
    end
    Store $\mathbf{x}^k$ in $\mathbf{V}(\tau_n)$;
  end
  else
    Set $\mathbf{G} = \mathbf{G}_2$, $\mathbf{H} = \mathbf{H}_2$, $\theta = \frac{1}{2}$ and $w = w_2$;
    /* Same PSOR while-loop as in the case $n \leq 3$ */
    Store $\mathbf{x}^k$ in $\mathbf{V}(\tau_n)$;
  end
end

Table 5.1: Values of parameters in the SVJD model (4.10) and (4.13), upper bounds s_J and v_M and mesh sizes for the spatial computational domain used for all numerical experiments in Section 5.4.

Asset Price		**Stoch. Vol.**		**Jumps**		**Mesh Sizes**	
T	0.50	ξ	2.00	$\tilde{\lambda}_1$	5.00	s_J	4.00
q_1	0.05	η	0.56	β_{j_1}	0.20	v_M	2.00
q_2	0.03	Λ	0.00	α_{j_1}	0.00	J	140
σ_1	0.50	ω	0.40	$\tilde{\lambda}_2$	2.00	M	25
σ_2	0.30	ρ_1	-0.50	β_{j_2}	0.20		
ρ_w	0.50	ρ_2	-0.50	α_{j_2}	0.00		
				L	20		

generated using the MOL algorithm discussed in the previous section. These results were obtained using the parameter values enumerated in Table 5.1, although in Section 5.5, we investigate how the option prices and the early exercise boundary change in response to alternative model parameter values.[11] The values of q_1 and q_2 were chosen so that the exercise boundary is discontinuous at maturity, thereby highlighting the impact of jumps in asset prices. The chosen means of the jump size variables Y_1 and Y_2 are zero, indicating that upward and downward jumps in asset prices are equally likely to occur. We also assume that they have equal standard deviations.[12] Correlations among Wiener processes are initially assumed to be positive, although the effect of negative correlations is also be explored in Section 5.5. The parameter values in Table 5.1, particularly the long-run mean of the variance process η and the jump intensities $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$, are chosen mainly to highlight the effect of stochastic volatility and jumps on exchange option prices.[13]

All source codes were implemented using MATLAB® on an x64-based personal computer with an Intel® Core™i7-10710U CPU with 1.10GHz, 1608 MHz, 6 cores, 12 logical processors, and 8GB RAM. Due to hardware constraints, modest mesh sizes for τ, v, and s (or x) are used, although as seen in Table 5.2 below, the MOL does not require too many time steps to converge. The computational domain $[0, 4]$ for s was divided into three intervals, $[0, 0.5]$, $[0.5, 3]$, and $[3, 4]$, which was then subdivided using 20, 80, and 40 mesh points, respectively. This allows for a more precise approximation of the early exercise boundary as it tends to occur near $e^{(q_1-q_2)T}$. Despite the use of relatively few mesh points, convergence was attained for all MOL implementations with parameters reported in Table 5.1 and the alternative values explored in Section 5.5. The MATLAB codes used in this analysis can be found here: https://drive.google.com/drive/folders/1KrLDaiXiXPfaJ1JulRyDVKFgwJae82_v?usp=sharing.

[11] Model parameter values used in this chapter are similar to those used by Chiarella et al. (2009), although additional parameters are used to accommodate the second asset.

[12] Less emphasis on the jump size distribution concentrates the analysis on the effect of the jump intensities. Alternative jump size distributions may be considered as well, although the form of the density function dictates the quadrature formula for approximating the integral terms.

[13] As such, these values does not particularly reflect market scenarios determined by real data, since the calibration of the asset price models used in this work goes beyond the scope of the thesis.

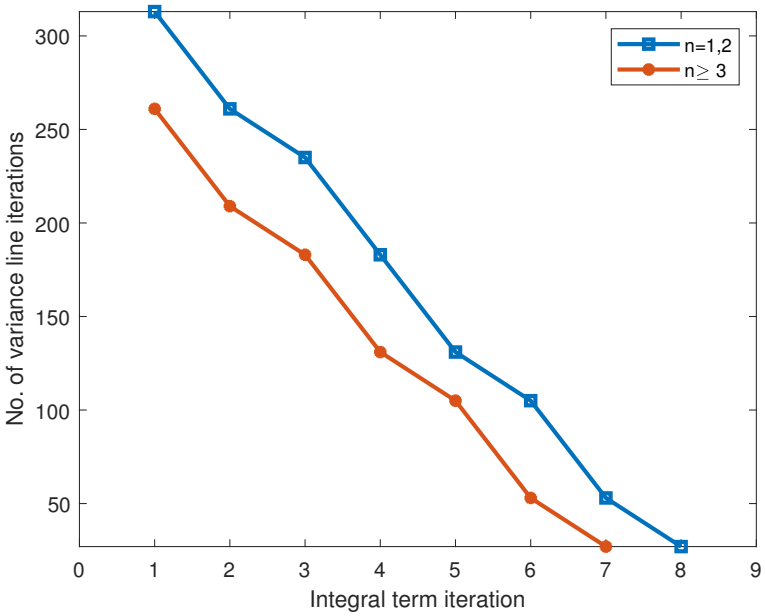


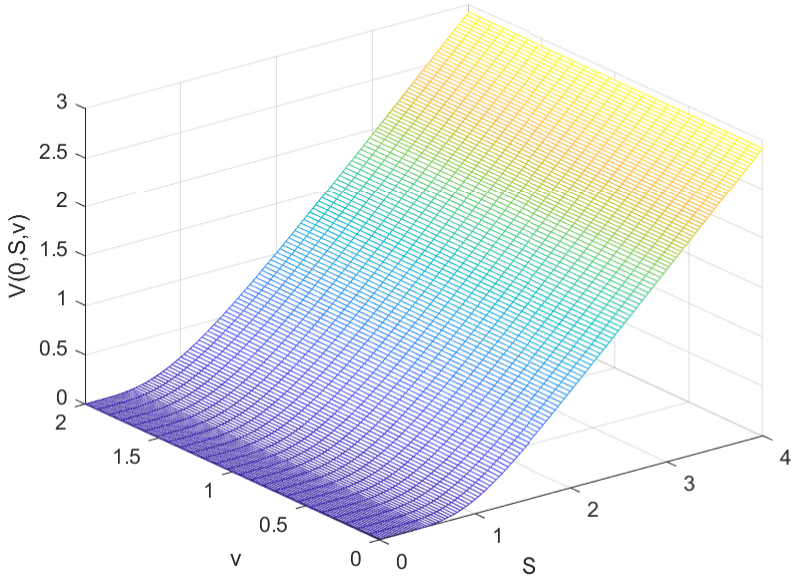
Figure 5.1: Typical number of variance line iterations per integral term iteration at the nth time step for the MOL for European and American exchange options.

All option prices, option deltas, and price differences presented in this section are in their *discounted form*, i.e. they are expressed in units of the second asset yield. As such, prices and price differences may be more pronounced if the exchange option is written on highly priced underlying assets.

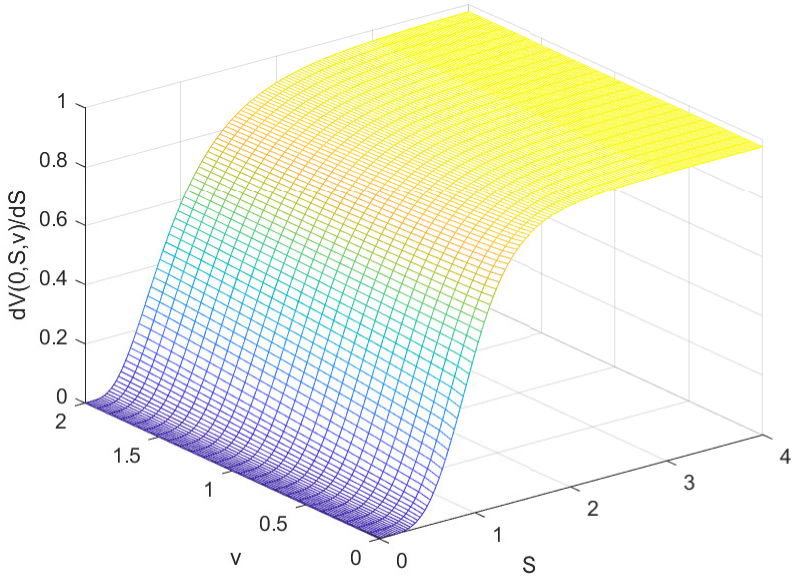
5.4.1 MOL Approximation of Exchange Option Prices

Figure 5.1 shows the typical convergence pattern of the integral term and variance line iterations for the MOL approximation of the European and American exchange option in the nth time step. It has been observed that the convergence behavior is the same for both types of exchange options. However, we note a decrease in the number of integral term iterations required on and after the third time step, signifying the effect of adopting a second-order backward difference approximation for the time derivative for $n \geq 3$ on the efficiency of the computation. The number of variance line iterations generally decreases as the integral line iterations continue. Chiarella et al. (2009) reported that the number of variance line iterations required drastically increases as the number of mesh points in v increases, but we were unable to replicate this phenomenon due to hardware constraints. While the convergence behavior is the same for both European and American cases, a typical MOL pass yields a computing time for the American exchange option that is almost double that required for the European case. In a sample implementation, the European and American cases were completed in 74.45 and 175.85 seconds, respectively (see also Table 5.2).

Figure 5.2 shows the approximation of the discounted European exchange option price and delta against s and v at $\tau = T$ (or $t = 0$), the start of the life of the option. Similar to ordinary



(a) Discounted European exchange option price



(b) Discounted European delta

Figure 5.2: Method of lines approximation of the discounted European exchange option price $V(\tau, s, v)$ and delta $\partial V(\tau, s, v)/\partial s$ at $\tau = T$.

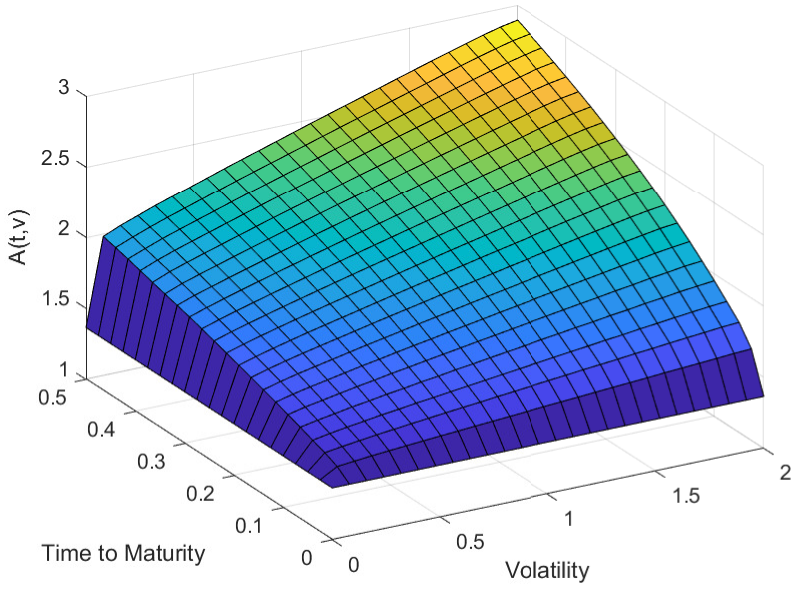


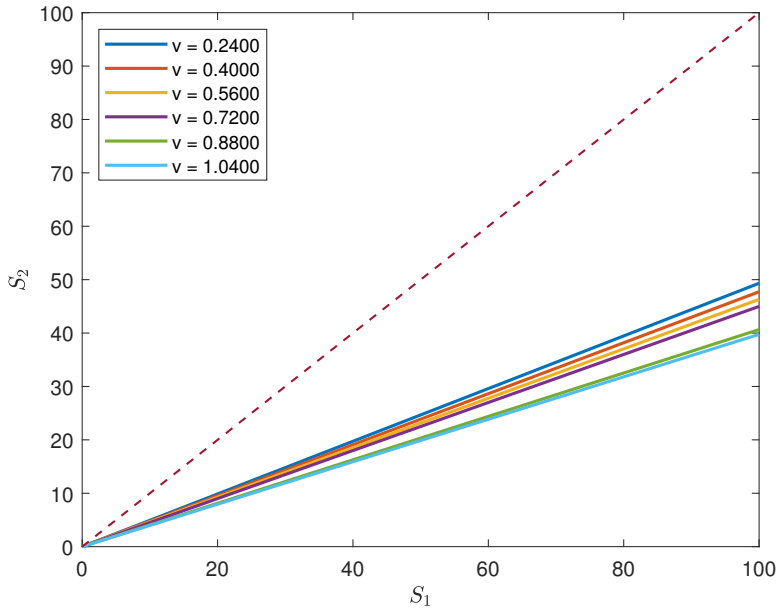
Figure 5.3: Method of lines approximation of the early exercise boundary surface $A(\tau, v)$.

European call options, the delta of the European exchange option is steepest when the asset yield ratio is close to $e^{(q_1-q_2)T}$.

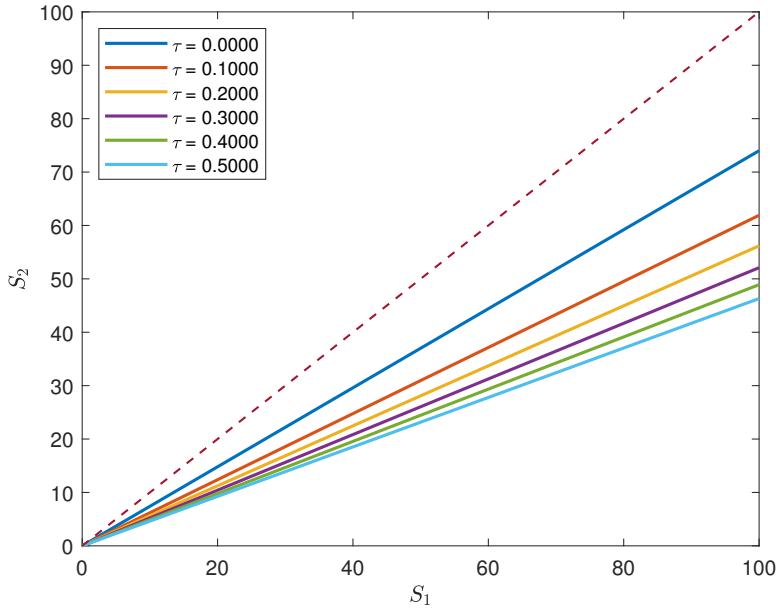
Figures 5.3 and 5.5 show the MOL approximations of the early exercise surface, the discounted American exchange option price, and the American exchange option delta at time-to-maturity $\tau = T$ (or time $t = 0$). Figure 5.3 shows that the early exercise boundary at $\tau = 0$ is constant with respect to v since the option payoff is independent of v as discussed in Section 4.5. The figure also illustrates that the $A(t, v)$ is an increasing function of v, thus supporting our assumption in the previous chapters that the critical price ratio must be dependent on v, as well as τ. As expected from the put-call transformation technique, the American and European exchange option price profiles, when expressed in terms of the asset yield ratio, behave in a similar manner to their ordinary call option counterparts.

We also show the early exercise boundary $s_1 = B(\tau, v)s_2$ on the s_1s_2-plane in Figure 5.4 for various values of τ and v. Note that the MOL produces values of $A(\tau, v)$, so to retrieve the critical asset price ratio $B(\tau, v)$, we use the formula $B(\tau, v) = A(\tau, v)e^{-(q_1-q_2)(T-\tau)}$. The region to the right of the boundary is the stopping region for the American exchange option. Similar to the conclusions from Figure 5.3, we also see in Figure 5.4 that for fixed τ (resp. v), the boundary becomes flatter as v (resp. τ) increases.

The price and delta profiles of the European and American exchange options appear to behave similarly in s and v, but the difference is emphasized when we compute the early exercise premium V^P. Figure 5.6 shows the discounted early exercise premium surface in the asset yield ratio and variance at $\tau = T$. For a fixed v, the premium is increasing in s. On closer inspection, the early

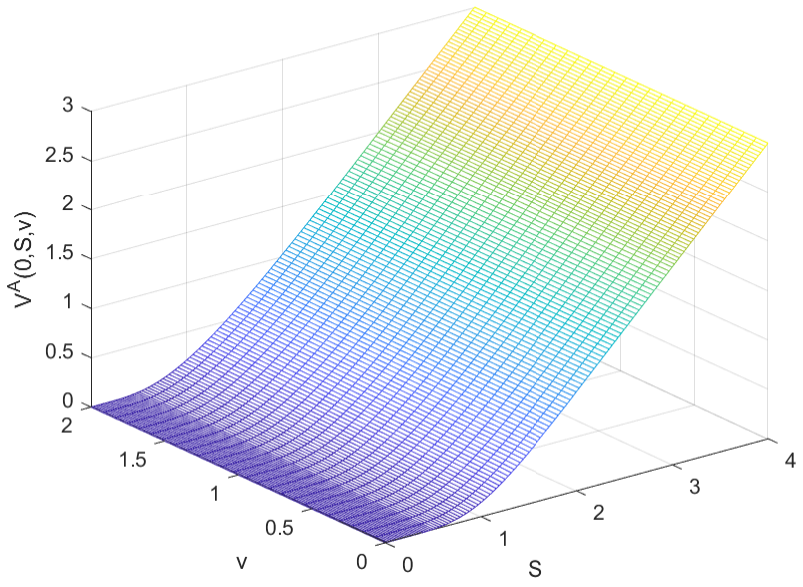


(a) Fixed $\tau = 0.5$

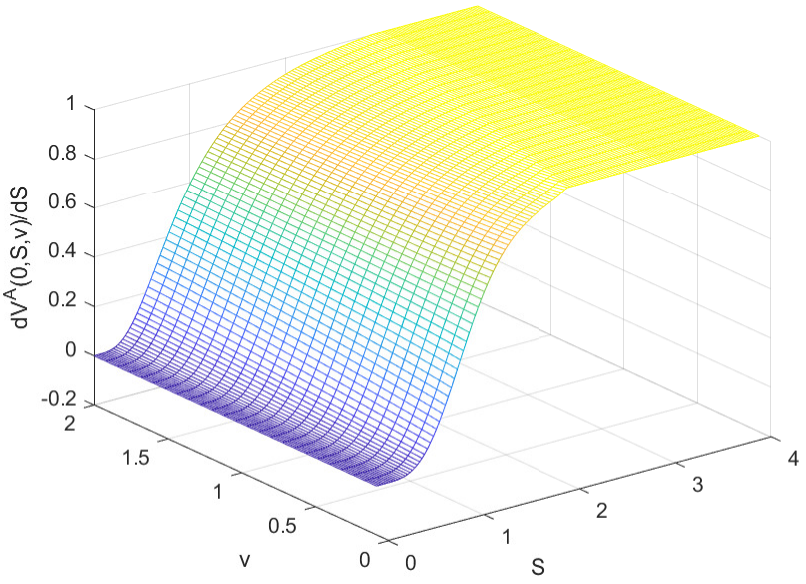


(b) Fixed $v = 0.56$

Figure 5.4: The early exercise boundary $s_1 = B(\tau, v)s_2$ on the s_1s_2-plane for various values of τ and v.

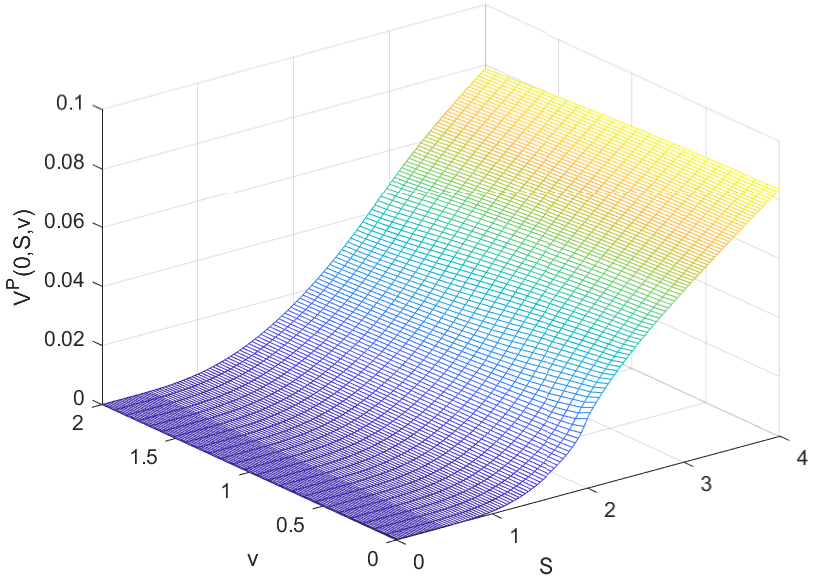


(a) Discounted American exchange option price

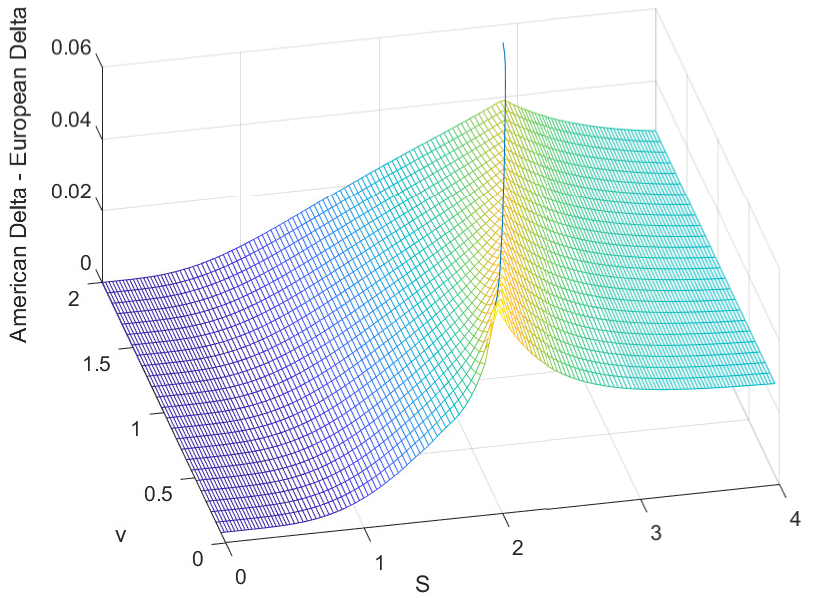


(b) Discounted American delta

Figure 5.5: Method of lines approximation of the discounted American exchange option price $V^A(\tau, s, v)$ and delta $\partial V^A(\tau, s, v)/\partial s$ at $\tau = T$.



(a) Discounted early exercise premium



(b) Difference between American and European deltas

Figure 5.6: Method of lines approximation of the discounted early exercise premium $V^P(\tau, s, v)$ and the difference between the deltas of the American and European exchange options at time $\tau = T$. The blue curve in (b) indicates the early exercise boundary at $\tau = T$.

exercise premium has an inflection which occurs at the early exercise boundary $A(T, v)$. This is also where the difference in American and European deltas peaks. This implies that the premium increases the fastest as the asset yield ratio approaches the exercise boundary from the left.

5.4.2 Comparison with Option Prices Approximated with the LSMC and PSOR Algorithms

As a benchmark for comparing the American option prices approximated using the MOL, we also estimated prices using the projective successive overrelaxation (PSOR) algorithm. This method involves articulating the linear complementarity problem (LCP) associated to the American exchange option, discretizing the problem using finite differences with Rannacher time-stepping, and solving the resulting algebraic complementary problem using the PSOR algorithm (see Section 5.3). We treat the prices solved using the PSOR with a reasonably high discretization resolution as the reference prices for comparison in the absence of data on actual exchange option prices (Ikonen and Toivanen, 2007; Chiarella et al., 2009). Given our hardware constraints, we chose $N = 500$, $J = 250$, and $M = 50$ as the mesh sizes for the reference prices. The spatial mesh for the reference prices consists of at least twice as many points in the s- and v-direction, so the spatial mesh is at least four times finer than the one used in the MOL. We also used the PSOR to solve for option prices with spatial mesh sizes identical to those used in the MOL and with three time discretization schemes: the Rannacher scheme, the implicit Euler (IE) scheme, and the Crank-Nicholson (CN) scheme.[14]

We also compare the price estimates from the MOL to those obtained using the least-squares Monte Carlo (LSMC) method, closely following the implementation of Longstaff and Schwartz (2001). We do so to provide insights on how the numerical IPDE solution techniques (MOL and PSOR) compare to a pure simulation approach whose starting point is the stochastic dynamics of the underlying state variables. We implement the LSMC using 10,000 scenarios (half of which are antithetic variates) and $N = 1000$ time steps, necessary to improve the accuracy of the results. Further details on the implementation of the LSMC are provided in Appendix D.2.

In Table 5.2, we compare estimates of the discounted American option price and the early exercise boundary obtained using the MOL, the PSOR algorithm, and the LSMC method for $\tau = T$ (or calendar time $t = 0$) and $v = 0.56$. Since we assume in Table 5.1 that $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$, a value of $v = 0.56$ corresponds to an instantaneous volatility of $0.5\sqrt{0.56} \approx 0.3742$ and $0.3\sqrt{0.56} \approx 0.2245$ for the first and second assets, respectively. Pricing errors with respect to the reference prices are measured using the root mean-square relative difference (RMSRD), whereas for the early exercise boundary, we measure the error using a simple percentage difference with respect to the reference boundary estimate. We also provide the total run time for each implementation

[14] To ensure the convergence of the PSOR algorithm, a larger value for N ($N = 150$) was used in the time-discretization of the LCP.

(i.e. the total time to run the algorithm for all time steps and values of s and v) and the average run time per time step.

From the results in Table 5.2, we immediately notice that the MOL produces price estimates that are closest to the reference prices, compared to the secondary PSOR implementations. Increasing the number of time steps in the MOL does not significantly reduce the error measurement, indicating that one can use the MOL with relatively fewer time steps (compared to the PSOR and LSMC approaches) and still achieve accurate results. Prices from the PSOR with Rannacher time stepping and the Crank-Nicholson scheme are identical since at $\tau = T$, both approaches use the Crank-Nicholson scheme. Using the implicit Euler scheme for all time steps (as opposed to just the first three in the Rannacher scheme) produces a slightly higher error. However, we report that no significant oscillations occurred in using a pure Crank-Nicholson scheme.

The LSMC method produced the least accurate set of prices compared to all methods considered in the analysis. Discrepancies are particularly larger when the option is deeply out-of-the-money, most likely due to the slow convergence of the algorithm, similar to what was observed by Chiarella and Ziveyi (2014). In addition, the regression step in the LSMC tends to be ill-conditioned for smaller values of s. Furthermore, price estimates from the LSMC when the option is in-the-money do not coincide with the intrinsic value of the option possibly due to errors in the regression step, but LSMC prices for higher values of s tend to agree more with prices from the MOL and PSOR compared to the out-of-the-money scenarios. Additional runs of the LSMC with larger values of N and with more basis functions require significantly longer processing time but were found to not drastically improve the price estimates.

It also takes less time to estimate option prices using the MOL compared to the PSOR and the LSMC methods. While the PSOR runs require less computation time per time step, the method requires a large value of N to ensure convergence. The computation times indicated for the reference and PSOR prices only involve the calculation of the price profile, since the method does not require the simultaneous solution of the early exercise boundary unlike the MOL. In addition, the option delta and gamma (among other price sensitivities) have to be calculated separately from the price profile when using the PSOR, similar to other finite difference approaches such as the componentwise splitting method (Ikonen and Toivanen, 2007; Toivanen, 2010a) and the sparse grid approach (Kang and Meyer, 2014). As such, the MOL is more efficient in that the calculations for the early exercise boundary, option delta, and option gamma are integrated into the solution of the price profile.

As with any numerical method, the usual sources of error include (but not are limited to) the discretization scheme for the simulation of state variables or the discretization of partial derivatives, truncation of infinite domains, the choice of mesh sizes and partition points in all variables, boundary conditions for the IPDE, numerical integration scheme (including the interpolation and

Table 5.2: Comparison of discounted American option price and early exercise boundary estimates obtained using the MOL algorithm, the projective successive overrelaxation (PSOR) algorithm, and the least squares Monte Carlo (LSMC) method. Shown are estimates at time $t = 0$ (or $\tau = T$) and $v = 0.56$ for a selection of values for $\tilde{S}$. Parameter values are given in Table 5.1.

		MOL			**PSOR**			**LSMC**		
$\tilde{S}$	Ref	$N = 20$	$N = 50$	$N = 100$	Ran	IE	CN	Price	95% CI	
0.5000	0.0084	0.0089	0.0089	0.0089	0.0100	0.0100	0.0100	0.0015	0.0012	0.0018
0.6250	0.0249	0.0255	0.0255	0.0255	0.0256	0.0257	0.0256	0.0101	0.0091	0.0111
0.7500	0.0559	0.0558	0.0559	0.0559	0.0529	0.0528	0.0529	0.0317	0.0298	0.0335
0.8750	0.1037	0.1022	0.1022	0.1022	0.1058	0.1057	0.1058	0.0723	0.0694	0.0752
1.0000	0.1684	0.1647	0.1647	0.1647	0.1634	0.1633	0.1634	0.1313	0.1278	0.1348
1.5000	0.5385	0.5355	0.5356	0.5356	0.5471	0.5471	0.5471	0.5144	0.5103	0.5185
2.0000	1.0017	1.0018	1.0018	1.0018	1.0015	1.0015	1.0015	1.0036	1.0013	1.0059
2.5000	1.5000	1.5000	1.5000	1.5000	1.5000	1.5000	1.5000	1.5061	1.5022	1.5099
3.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0099	2.0049	2.0149
3.5000	2.5000	2.5000	2.5000	2.5000	2.5000	2.5000	2.5000	2.5097	2.5046	2.5149
4.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0116	3.0053	3.0179
RMSRD		0.0224	0.0218	0.0217	0.0614	0.0617	0.0614	0.3502		
EEB	2.1760	2.1595	2.1553	2.1481	2.2130	2.2229	2.2132			
EEB error (%)		−0.7583	−0.9513	−1.2822	1.7004	2.1553	1.7096			
Total run time	10,112.00	175.83	345.88	519.64	527.89	588.00	497.35	324,168.04		
Run time per N	20.22	8.79	6.92	5.20	3.52	3.92	3.32	3,241.68		

Notes: In additional to total run time, we also provide the average run time per time step for each implementation. The average run time of the LSMC algorithm for one set of initial values of $\tilde{S}$ and $\tilde{v}$ is 92.6194 seconds. The computational time shown above is obtained by multiplying the average run time by $J = 140$ and $M = 25$. Total run times for the PSOR implementations, including the reference prices, exclude the time to compute the early exercise boundary.

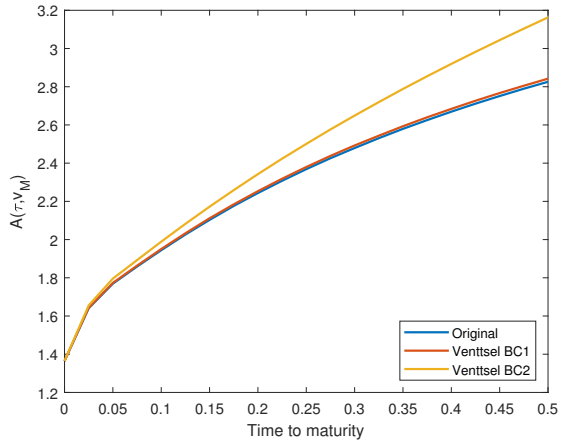
extrapolation involved), the errors involved in the iterative processes, and the choice of basis functions for the LSMC. Generally, these errors can be reasonably mitigated with finer discretization grids or imposing a more stringent tolerance in the iterative processes. However, a comprehensive error analysis and a formal proof of convergence for all numerical techniques considered goes beyond the scope of this thesis.

5.4.3 Effect of Venttsel Boundary Conditions at $v = v_M$

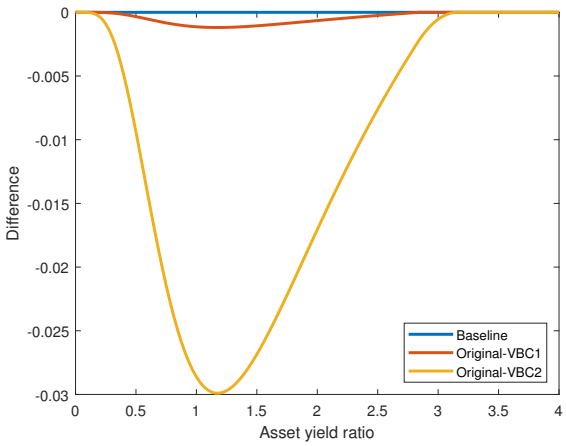
The final numerical experiment in this subsection explores how MOL approximations of option prices and the early exercise boundary are affected by the use of Venttsel boundary conditions. Figure 5.7 shows the early exercise boundary, American option price, and delta computed for each Venttsel boundary condition (5.18) and (5.19) using parameters values in Table 5.1. For the option price and delta, we exhibit the differences at $t = 0$ and $v = v_M$. As can be seen from the graphs, there is very little difference in the MOL approximations when either (5.7) or (5.18) is used. This may indicate that if the option price becomes insensitive to changes in volatility at large volatility levels, then the option delta and vega also become insensitive to further movements in volatility. We furthermore note that the computational time when (5.18) is active is shorter than when (5.7) is used, indicating that using the Venttsel boundary condition may be computationally more efficient while delivering the similar results. The differences in option price and delta are much more pronounced when (5.19) is used, particularly when the option is at-the-money. Using boundary condition (5.19) also results to a much larger approximation of the early exercise boundary at $t = 0$ compared to when either (5.7) or (5.18) is used. Thus it may not be appropriate to use the boundary condition (5.19) since, aside from substantial discrepancies from the usual price, it also somewhat omits the effect of stochastic volatility on the option price at higher levels of volatility. Since an actual calibration of the model may yield parameter values which violate the Feller condition, it is suggested that the coefficients of the second-order terms are regularized, as we did in this implementation, to ensure that Venttsel boundary conditions can still be used.

5.5 Impact of Jumps and Stochastic Volatility

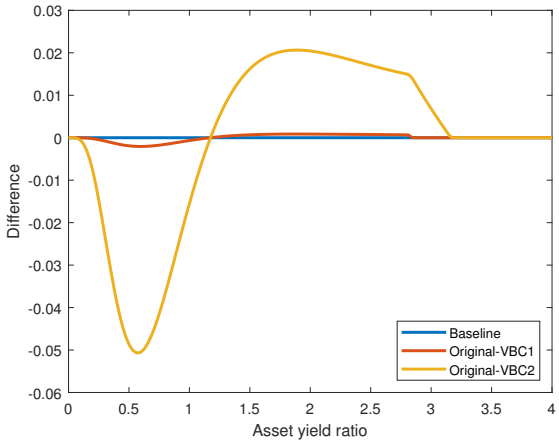
In this section, we investigate how the discounted American and European exchange option prices and the early exercise boundary time profile change in the presence of stochastic volatility and jumps in the asset prices. In all subsequent analyses, option prices and the early exercise boundary are shown for $v = 0.56$ (the long-term variance η) and parameter values in Table 5.1 except for those which were varied in the numerical experiment and for other parameters as noted in the figure descriptions. Prices and the early exercise boundary were generated using the method of lines discussed in the previous sections. The analysis of option price differences at $t = 0$ (or $\tau = T$)



(a) Early exercise boundary

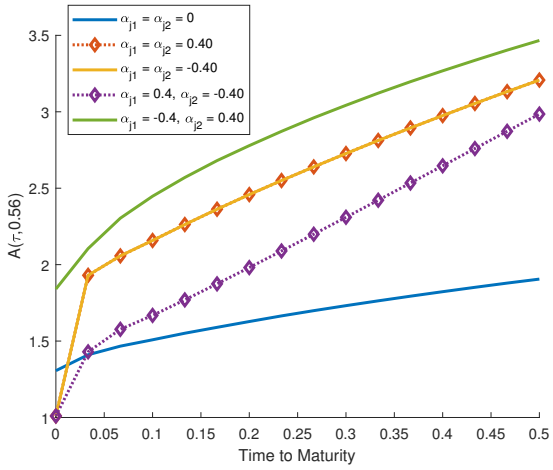


(b) Discounted prices

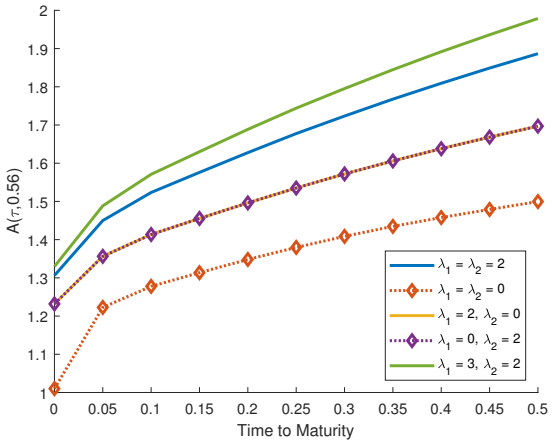


(c) Discounted deltas

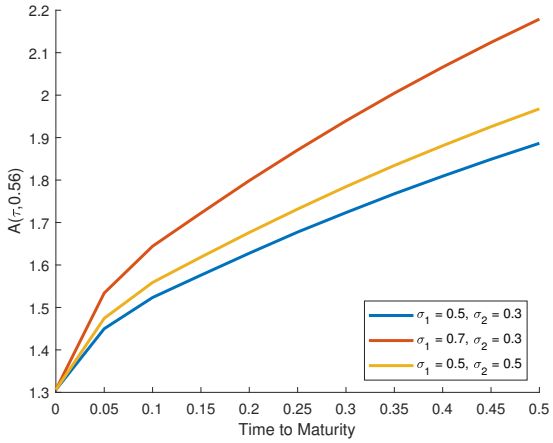
Figure 5.7: Comparison of early exercise boundaries $A(t, v)$, discounted option prices, and discounted deltas for the American exchange option for various choices of boundary conditions at $v = v_M$. VBC1 and VBC2 refer to (5.18) and (5.19), respectively.



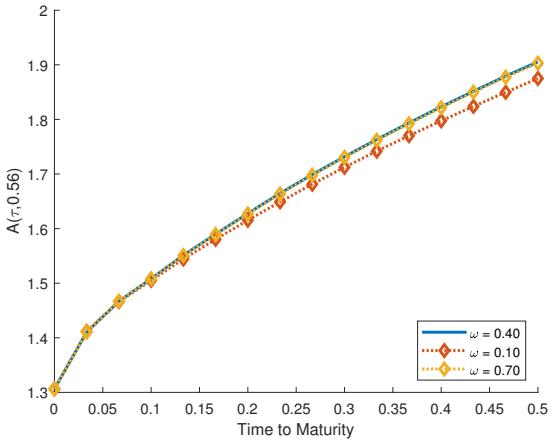
(a) α_{j_1} and α_{j_2}



(b) $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$



(c) σ_1 and σ_2



(d) ω

Figure 5.8: Effect of changes in key jump and stochastic volatility parameters on the early exercise boundary when $v = 0.56$.

are shown with respect to the asset yield ratio $\tilde{S}$ as this highlights how the early exercise boundary influences the comparative statics of the American exchange option.

5.5.1 Impact on the Early Exercise Boundary

Figure 5.8 shows the effect of the model parameters on the early exercise boundary. In particular, we focus on the effect of the jump intensities $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$, the jump size means α_{j_1} and α_{j_2}, the coefficients σ_1 and σ_2 of instantaneous volatility in the asset prices, and the volatility of volatility ω.

As expected from the results on the limit of the early exercise boundary in Section 4.5, changing the jump intensity or the jump size mean leads to upward translations in the early exercise boundary profile. In addition, there are some configurations of the jump size means that lead to a continuous exercise boundary at maturity. In panel (a), the baseline refers to the case where positive and negative jumps are equally likely in both asset prices, whereas a positive (negative, resp.) value of

α_{j_i} means that upward (downward, resp.) jumps are more likely in the price of asset i, $i = 1, 2$. Here, the highest exercise boundaries are generated in the case where downward jumps are more likely in the first asset while upward jumps are more probably in the second asset. This presents the highest probability that the asset yield ratio will be sent back into the continuation region, making early exercise sub-optimal. Thus, the option holder is less likely to exercise and will only exercise for higher values of $\tilde{S}$ to avoid rebalancing costs. A similar interpretation can be cast for the other cases: a nonzero value of the jump size mean implies greater risk of incurring rebalancing costs upon exercising the option, so the boundary is higher than the baseline.

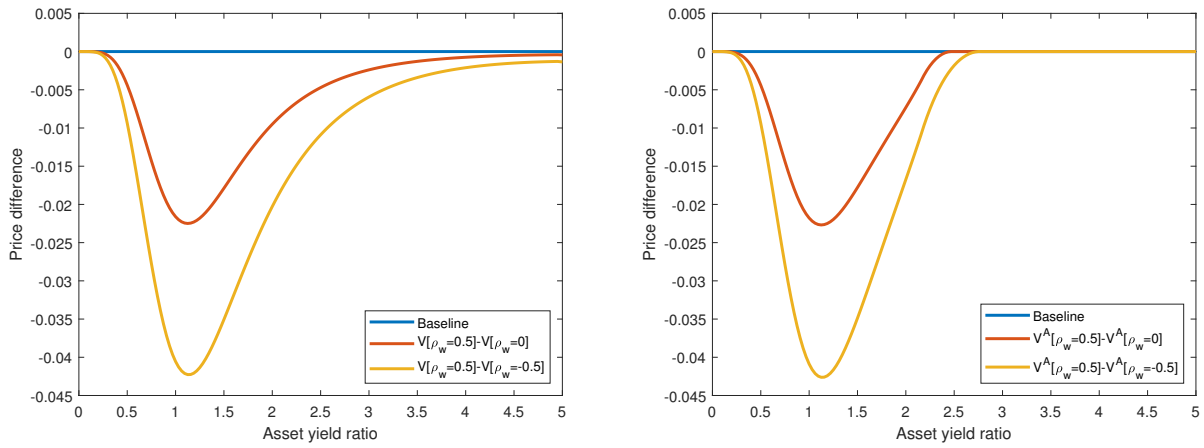
A similar analysis holds for the effect of the jump intensities on the early exercise boundary (see panel (b)). Larger values of $\tilde{\lambda}_1$ and/or $\tilde{\lambda}_2$ mean that the respective asset price is more likely to jump. The lowest values of the boundary are produced when there are no jumps since the likelihood of the asset yield ratio jumping back into the continuation region immediately upon exercising the option is less than the other cases. The results produced by assuming $\tilde{\lambda}_1 = 2$ and $\tilde{\lambda}_2 = 0$ (yellow line) and $\tilde{\lambda}_1 = 0$ and $\tilde{\lambda}_2 = 2$ (purple line) are similar because we assumed the same the jump size density parameters for both asset prices (see Table 5.1). Results similar to those shown in panel (a) are expected if the assumed jump parameters were different.

In contrast, changes in parameters related to the stochastic volatility case does not affect the exercise boundary at maturity. Instead, the effect becomes more pronounced the farther the option is from maturity. A higher value of σ_i, $i = 1, 2$, indicates that the ith asset price is more sensitive to movements in the stochastic volatility, whereas a higher value of ω means that the underlying volatility itself is more volatile.[15] The boundary does not seem to be significantly affected by the volatility of volatility, but as seen in panel (c) the boundary increases when the asset prices are more sensitive to volatility. Thus, for example, the configuration $\sigma_1 = 0.7$ and $\sigma_2 = 0.3$ generates a higher boundary than the base case (blue) since there is more risk of unfavorable movements in the first asset price especially in high volatility scenarios.

5.5.2 Impact of Model Parameters on Exchange Option Prices

We now proceed to numerical comparative statics on the discounted European and American prices. We display the difference when the price generated after modifying a given parameter is subtracted from the option price solved with the parameters in Table 5.1. A common observation is that price differences vanish for deeply in-the-money American exchange options because of the value-matching condition. The point at which the price differences meet the baseline is the exercise boundary, which is more clearly depicted in Figure 5.8. In contrast, the price differences for deeply

[15] This is not to be confused with the correlation parameters ρ_1 and ρ_2. Our implementations assume that these correlations are negative, reflecting the well-known leverage effect that is observed empirically between the diffusion components in the asset prices and in the underlying (stochastic) volatility (Cont, 2001). The parameters σ_1 and σ_2 govern the direct contemporaneous relationship between the asset price and the instantaneous volatility level.



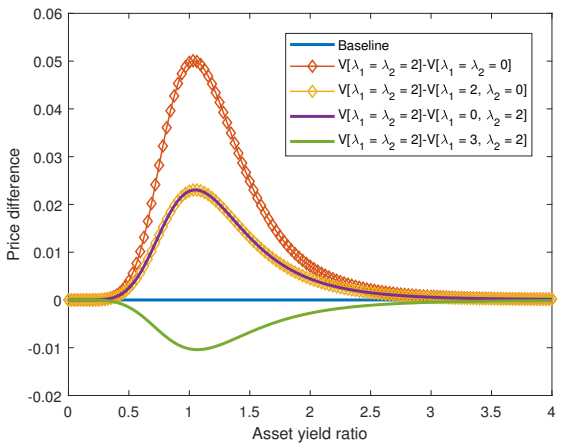
(a) European exchange option (b) American exchange option

Figure 5.9: Price differences in European and American exchange option prices for various values of the correlation ρ_w between the asset price processes at $t = 0$ and $v = 0.56$.

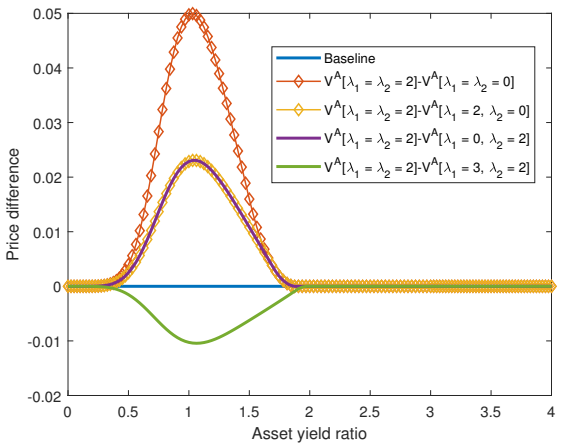
in-the-money European options dissipate more slowly, if at all as can be seen in some simulations.

The succeeding discussions are based on the principle that exchange option holders find it desirable when it is more likely that the price of the first asset is greater than the price of the second asset, or when the asset yield ratio achieves a value greater than 1. Favorable movements in $\tilde{S}$ are caused by, among other factors, an upward movement in S_1 and/or a downward movement in S_2. From a risk management perspective, exchange option prices tend to be higher when the likelihood of unfavorable movements occurring in $\tilde{S}$ are higher, since the exchange option provides added protection from losses compared to a simpler forward contract that forces the exchange of two assets regardless of their relative values. In addition, we tackle the discussion of the European case and the American cases from separate perspectives. For the European case, the investor is concerned whether $\tilde{S}$ will move towards a favorable position at maturity as viewed from information known at calendar time $t = 0$. In contrast, the holder of an American exchange option not only has to contend with the issue stated for the European case, they are also wary of potential adverse movements in $\tilde{S}$ immediately upon early exercise, as these adverse movements can lead to rebalancing costs. With these principles in mind, the analysis below investigates how the model parameters affect the likelihood of realizing favorable movements in the asset yield ratio and how the exchange option prices are affected by changes in these parameters.

Figure 5.9 shows how the exchange option prices are affected by changes with respect to the correlation between the diffusion components in the asset prices. Both options are more expensive when the asset prices are uncorrelated or negatively correlated to each other, since in both these cases there is a higher tendency for smaller spreads between the prices of the two assets. The price differences are highest in magnitude when the option is near at-the-money since when the asset yield ratio is within that range, the exchange option offers more risk protection than in the case

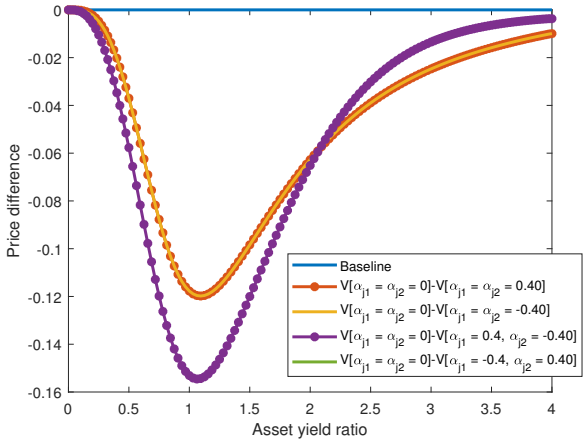


(a) European exchange option

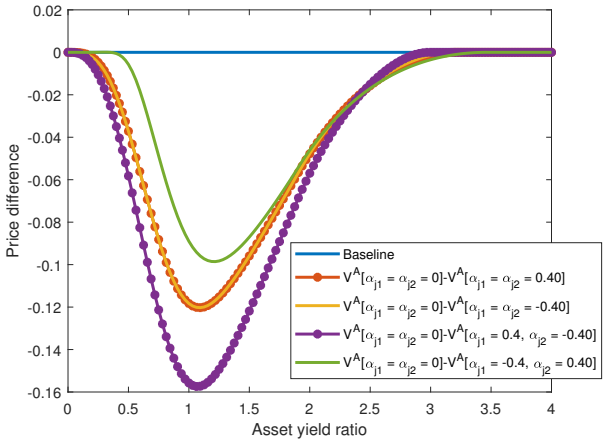


(b) American exchange option

Figure 5.10: Price differences in European and American exchange option prices for various values of the jump intensity parameters $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ at $t = 0$ and $v = 0.56$.



(a) European exchange option



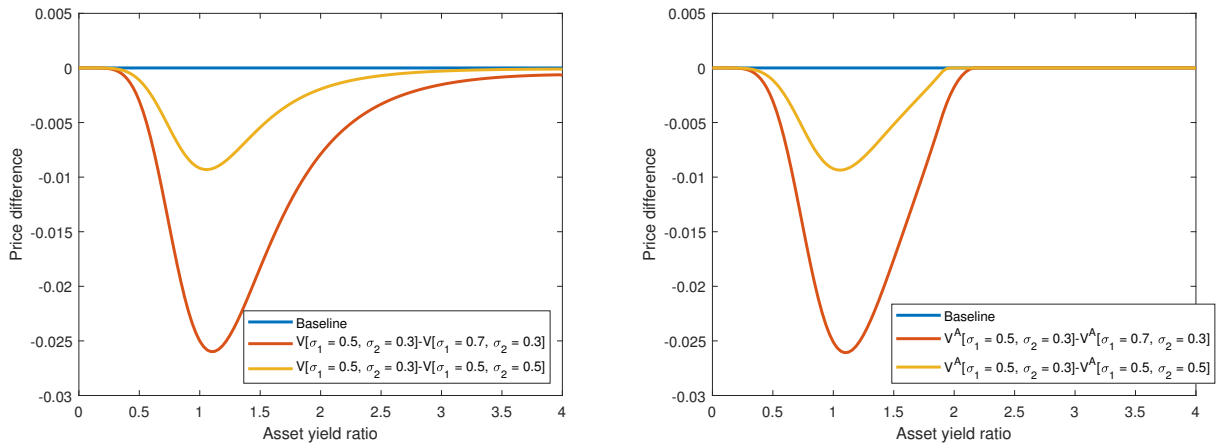
(b) American exchange option

Figure 5.11: Price differences in European and American exchange option prices for various values of the jump size means α_{j_1} and α_{j_2} at $t = 0$ and $v = 0.56$.

where the option is deeply in-the-money or out-of-the-money.

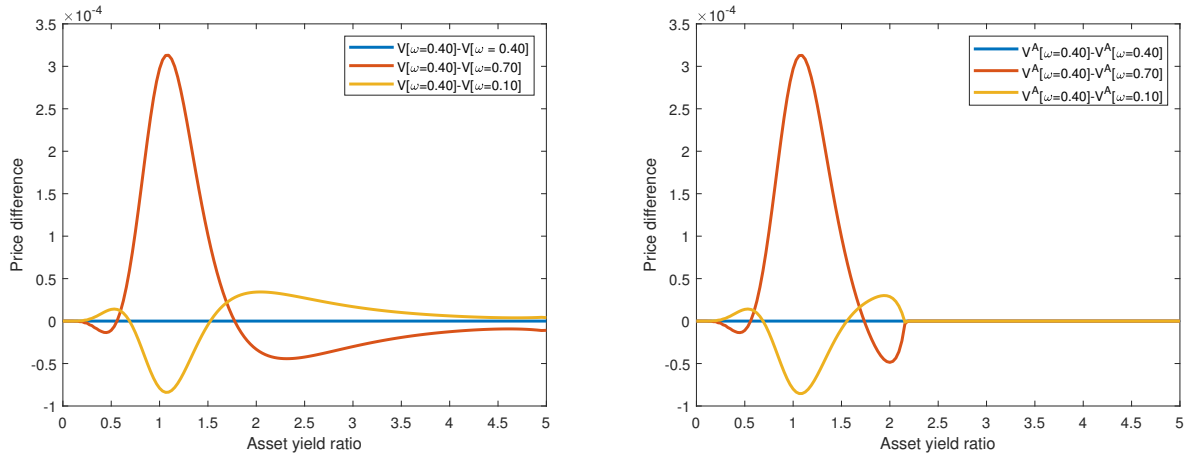
Figures 5.10 and 5.11 show how the jump intensities and mean jump sizes affect the discounted exchange option prices. From Figure 5.10, we see that option prices in the base case are higher than prices when one or both assets do not exhibit jumps (in which case they are driven only by stochastic volatility), while option prices are higher than the base case when jump intensities increase. This is consistent with the principle that higher values of the jump intensity leads to a higher likelihood of unfavorable movements in the asset yield ratio. The price differences are more pronounced when the option is near at-the-money, which can be again attributed to the risk management perspective described above.

A similar discussion can also be made on changes in option prices due to changes in the jump size distribution mean. For both European and American exchange options, option prices are



(a) European exchange option (b) American exchange option

Figure 5.12: Price differences in European and American exchange option prices for various values of the volatility coefficients σ_1 and σ_2 of the asset price processes at $t = 0$ and $v = 0.56$.



(a) European exchange option (b) American exchange option

Figure 5.13: Price differences in European and American exchange option prices for various values of the volatility of volatility ω at $t = 0$ and $v = 0.56$.

generally greater than the base case when asset prices tend to jump upward or downward more dominantly. Interestingly, near at-the-money option prices are highest when the first asset price tends to jump upward and when the second asset price tends to jump downward. The net effect is that the asset yield ratio tends to jump upward. Especially when the option at the current time is out-of-the-money, this case presents the greatest potential increase to the exercise value of the option, hence the large price differences even for deeply out-of-the-money options. As before, price differences are more pronounced when the option is near at-the-money.

Sudden short-term movements in the asset yield ratio can be attributed to jumps in the individual asset prices. Thus, it is also expected that the impact of jumps is significant when the option is near or at maturity since adverse movements in the asset prices can instantly diminish the exercise value of the option, even in the European case.

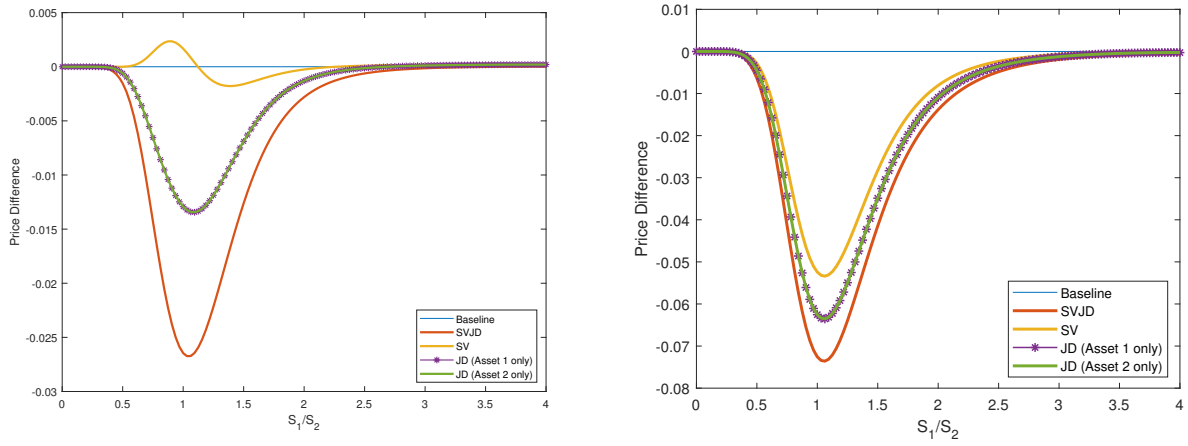
Finally, we consider the effect of volatility parameters on the discounted exchange option prices. In Figure 5.12, we show how exchange option prices change with respect to changes in the volatility coefficients σ_1 and σ_2. Option prices generally increase when σ_1 or σ_2 increases because higher values of these coefficients intensify the direct effect current volatility levels have on the asset prices, which in turn introduce a higher likelihood of adverse movements in the asset yield ratio. While the effect of the volatility coefficients is straightforward, the same cannot be said for that of the volatility of volatility ω as seen in Figure 5.13. The effect of the volatility of volatility is less straightforward. Increasing ω (relative to the default $\omega = 0.40$) results to prices which are lower when the option is near-the-money but higher when the option is either deeply out-of-the-money or in-the-money. We note however that the magnitude of the price differences observed in Figure 5.13 are substantially smaller those observed in the previous analyses. We expect that the effect of ω will be more pronounced for options with volatility-dependent payoffs.

In contrast to jumps, stochastic volatility is associated more to the long-term evolution in asset prices given its diffusive nature. This potentially allows stochastic volatility to offset movements in asset prices due to jumps for longer times-to-maturity (Chiarella et al., 2009). As such, the effect of jumps are more likely to impact the exercise value compared to the stochastic volatility, especially as the option approaches maturity. Thus, similar to the observations made from the analysis of the early exercise boundary, it is possible that the impact of the volatility parameters on the option prices diminishes as the option approaches maturity.

5.5.3 Impact of Jumps and Stochastic Volatility on Exchange Option Prices

This numerical experiment is concerned with assessing the impact of stochastic volatility and jumps to the discounted price of European exchange options. Here, we focus only on the European case, since as exhibited in the previous numerical experiments the American case behaves similarly, except for when the asset yield ratio exceeds the exercise boundary where the price differences vanish. Cases for a negative and a positive correlation between asset price processes are considered. For simplicity, we assume no dividend yields for both assets. As opposed to the per-parameter analysis in the previous section, here we consider the general effect of whether the underlying asset price dynamics exhibit stochastic volatility or jumps.

The base prices correspond to the pure diffusion case, for which a formula has been provided by Margrabe (1978). Furthermore, since the variance process is mean-reverting, we assume that the constant volatility for the pure diffusion case is $\sigma_i\sqrt{\eta_i}$ for asset $i = 1, 2$, as was done by Chiarella and Ziveyi (2013). The stochastic volatility case was simulated by setting the jump intensities to zero, but these were subsequently allowed to have nonzero values eventually building up to the default SVJD case. When exactly one of the jump intensities was equal to zero, it was assumed that this asset is driven by stochastic volatility dynamics whereas the other asset had both stochastic



(a) $\rho_w = 0.5$ (b) $\rho_w = -0.5$

Figure 5.14: Comparison of discounted European exchange option prices at $t = 0$ and $v = 0.56$ generated under pure diffusion, stochastic volatility, SVJD (jumps in asset 1 only), SVJD (jumps in asset 2 only), and SVJD (jumps in both assets). Results are presented for a positive and a negative correlation coefficient between the Brownian motions in the asset price processes. So that the effect of jumps does not dominate the effect of stochastic volatility, jump intensities in analysis were revised to $\tilde{\lambda}_1 = 1$ and $\tilde{\lambda}_2 = 1$ in the cases where at least one of the asset prices jump.

volatility and jumps.

As seen in Figure 5.14, stochastic volatility prices, when the correlation between the Brownian motions in the asset prices are positive, are slightly higher than the pure diffusion case when the option is out-of-the-money, but are lower when the option is in-the-money. In contrast, prices when volatility is stochastic are consistently higher than in the constant volatility case when the correlation is negative. Furthermore, the presence of jumps lead to higher option prices irrespective of the moneyness of the option, with the highest prices attained when both assets have SVJD dynamics. In both positive and negative correlation cases, the SV and SVJD option prices converge to the Margrabe price when the option is deeply out-of-the-money. The magnitude of the price differences are also more pronounced in the negative correlation case.

These results are unsurprising since the presence of stochastic volatility and/or jumps in asset prices increase the likelihood that it may not be optimal to exercise the option at its maturity. By purchasing the exchange option, the option holder (who holds shares of the second asset which will be exchanged for the first at maturity) gains additional protection from market movements that might result in the second asset having higher value than the first at the option's expiry. As a reflection of the added risk posed by stochastic volatility and jumps, the option price must therefore be higher than in the absence of these additional risk factors. This is especially true in the negative correlation case, since there is higher probability that the option will either be deeply in-the-money or deeply out-of-the-money due to opposite movements in the diffusive components of the asset prices. In addition, the price differences are larger in magnitude when the option is near at-the-money since the risk that the option becomes sub-optimal to exercise is higher within

that range.

The implications of jumps and stochastic volatility are also significant for the American case. In the American case, option holders face the risk that upon exercising the option early, asset prices may instantaneously and adversely jump back into the continuation region, making early exercise sub-optimal. In the SVJD case, such a situation becomes more probable since a jump from at least one of the asset prices (downward in the first and/or upward in the second) is sufficient to cause this phenomenon. The increased likelihood that the investor incurs these rebalancing costs (see Section 4.3) may contribute to higher exchange option prices in the presence of jumps.

We emphasize that these observations and conclusions were made for the specific parameter values we assumed in Table 5.1 and the modifications thereof introduced in each numerical experiment. The parameter values were chosen so as to best illustrate the effect that stochastic volatility and jumps on the exchange option prices and the early exercise boundary and are similar to those used by Chiarella et al. (2009). We anticipate that the insights explored here may still hold for other parameter values or when a full calibration exercise with respect to actual data is implemented.

5.6 Chapter Summary and Conclusion

This chapter discusses the pricing of European and American exchange options under stochastic volatility and jump-diffusion dynamics and its numerical implementation via a method-of-lines algorithm. With the use of Bjerskund and Stensland's (1993) put-call transformation technique, we are able to reduce the number of dimensions in the main problem and write the option price and the associated IPDE as a function of time, the asset yield ratio, and the instantaneous variance level. With the inhomogeneous form of the IPDE for the American exchange option, we were also able to analyze the behavior of the early exercise boundary near maturity. It was found that the limit of the boundary at maturity is strongly influenced by the magnitude of the difference between asset dividend yields and the jump component of each asset price process. The numerical implementation in this chapter complements the integral representations of the exchange option prices obtained using Fourier and Laplace transforms (see Chapter 4).

Given the reduction in dimensions, we then formulated a method of lines algorithm to numerically solve the option pricing IPDE, thereby detailing and extending the method presented by Chiarella et al. (2009). As noted in previous work that applied the MOL in option pricing, this method is particularly useful as it naturally computes for the option delta and gamma and minimal adjustments are required to calculate the free boundary associated with American options. While for simplicity our implementation uses constant parameter values, Algorithms 1 and 2 can easily be extended to parameters that are deterministic functions of time. However, adding stochastic elements will result to a higher number of dimensions in the MOL implementation, which will

then require additional iteration levels (see Kang and Meyer, 2014, for example). Our numerical analysis in Section 5.4 confirms that the MOL indeed performs more efficiently compared to a full finite difference and PSOR approach and the Longstaff and Schwartz (2001) Monte Carlo approach. In addition, the MOL also produced prices which were closest to the reference prices. As reported in previous work (Chiarella et al., 2009; Chiarella and Ziveyi, 2013; Kang and Meyer, 2014, among others), we expect that the MOL will perform just as efficiently, if not more, than other numerical solutions such as numerical integration, componentwise splitting methods, and sparse grid approaches. Algorithms 1 and 2 and the accompanying discussion in Section 5.2 can be easily modified and implemented to accommodate other payoff structures or underlying asset price dynamics.

Our analysis also showed that (5.18) at the far variance boundary is a plausible alternative to the often-used vanishing vega assumption, producing comparable results with less computational time. However, the choice of boundary conditions largely depends on the calibration of model parameters. Specifically, if calibration yields parameters which violate the Feller condition, which makes the Venttsel boundary condition (5.18) viable, then alternative conditions must be imposed on the boundary of the computational domain or the coefficients of the second-order partial derivatives need to be regularized.

Using the MOL approach, we were also able to assess the impact of the model parameters on the exchange option prices and the early exercise boundary. We find that the jump intensities and the jump size density mean have substantial impact on option prices as it is able to shift the early exercise boundary curves upward and generate the largest price differences relative to the default prices computed using parameters in Table 5.1. This is because the presence of jumps drastically increases the probability that the holder of an American exchange option will incur a rebalancing cost or the cost when the asset yield ratio instantaneously and unexpectedly jumps back into the continuation region upon exercise. We also find the correlation parameters and the volatility constants of proportionality have a considerable impact on the early exercise boundary and the option prices, while the volatility of volatility seems to have the least impact. We expect that the effect of stochastic volatility will be more pronounced for options with volatility-dependent payoffs. In addition, we find evidence that the general presence of stochastic volatility and or jumps in one or both assets also has a considerable effect on option prices when the asset prices are positively or negatively correlated with one another. These conclusions are based on the principle that the exchange option prices tend to be higher when there is a greater probability that the value of exercising the option decreases. However, these conclusions are true for the specific set of parameter values used in the numerical approximation and may not hold in full generality. A complete analytical comparative static analysis remains to be seen in literature, but nonetheless we have shown that jumps and stochastic volatility have considerable impact on option prices and

the early exercise boundary.

A formal convergence analysis of the MOL is also an issue to be investigated in future work, although the algorithm we presented converges for all reported parameter values. An application of the put-call transformation transformation and/or the MOL in pricing multi-asset derivatives under other asset price model specifications (for example, general Lévy processes, regime switching models, non-affine models) is also a topic that we aim to explore in future studies. In a future study, we also aim to see how the MOL can be extended when pricing takes place in the risk-neutral world (i.e. the second asset price then becomes a separate spatial variable) or when additional risk factors are included, such as multi-factor stochastic volatility (Christoffersen et al., 2009) and stochastic interest rates.

CHAPTER 6

The Affine Two-Asset SVJD Model and Volatility Indices

6.1 Introduction

In the previous chapters where the focus is the option pricing problem, we have assumed that all relevant parameters, including those for the change of probability measures, have been determined. However, the estimation of model parameters from historical asset price data and the calibration of models with respect to historical option price data are crucial issues that have to be resolved in order to apply the methods and techniques used in the preceding chapters.

The main difficulty in the estimation of the SVJD model is the unobservability of the volatilities. While several approaches for the estimation of stochastic differential equations with unobserved components have been proposed in the literature (see Appendix E.1), an alternative approach is to use volatility indices, such as the VIX, as a proxy for the unobserved instantaneous variance process. The use of volatility indices is justified for the class of affine jump-diffusion processes since it can be shown that the volatility index has an affine dependence on the instantaneous variance.

One crucial characteristic of the SVJD model (2.1)-(2.2) assumed in Chapter 2 is that, given the appropriate assumptions on the correlations between the Brownian motions that appear in the model, leads to an affine jump-diffusion model for the log-prices $X_1 = \ln S_1$ and $X_2 = \log S_2$ and the instantaneous variances v_1 and v_2 (see Section 2.5.2). In addition, the affine form holds for both the historical probability measure $\mathbb{P}$ and the risk-neutral probability measure $\mathbb{Q}$ as a consequence of the specific form of the Radon-Nikodým derivative $\frac{d\mathbb{Q}}{d\mathbb{P}} = L(T)$ we assumed in equation (2.8). As an affine jump-diffusion with observable and latent components, the model parameters can therefore be estimated using the approximate maximum likelihood approach of Bates (2006), the GMM method of Jiang and Knight (2002), or MCMC algorithms like that of Eraker et al. (2003).

However, if there exist a volatility index for the underlying asset and the underlying asset price

process is an affine jump-diffusion, then the volatility index can be written as an affine function of the unobservable instantaneous variance process with coefficients that depend on the parameters of the underlying asset price process (Lin, 2007; Duan and Yeh, 2010; Zhu and Lian, 2011; Aït-Sahalia et al., 2018; Bardgett et al., 2019). Therefore, we can use the volatility index as an observable proxy for the instantaneous variance, thus making the vector process $\mathbf{Y} = (X_1, v_1, X_2, v_2)^\top$ fully observable. As a further consequence, this allows the use of parameter estimation and calibration methods (see Appendix E) which require that all components of the price process be observable.

In this chapter, we shall consider an extension of the SVJD model (2.1)-(2.2) to the case where even the variance processes may jump and where common jumps may occur between the two assets, between asset and variance pair, and between the variances. Such a specification is a two-asset version of the affine models considered by Eraker et al. (2003), Aït-Sahalia et al. (2018), Bardgett et al. (2019), and Bakshi et al. (2021). We specify an affine jump-diffusion model with these characteristics and show that, given an appropriate Radon-Nikodým derivative, the affine structure is preserved. We then show how the the volatility index corresponding to each of the two assets can be written as an affine function of the unobservable instantaneous variance process.

We emphasize that we shall not discuss the implementation of methods for estimating the parameters or for calibrating the two-asset SVJD models used in this thesis, since this goes beyond the scope of the work. Instead, we shall lay the groundwork for future research in this matter by discussing the affine dependence of volatility indices under the two-asset SVJD model, which then allows one to use methods for estimating the parameters of fully-observed stochastic processes.

6.2 A Two-Asset Affine SVJD Model with Volatility Jumps and Common Jumps

Since volatility indices are defined in terms of expectations in the risk-neutral measure (Carr and Wu, 2006; Lin, 2007; Zhu and Lian, 2011; Aït-Sahalia et al., 2018), we shall work directly in the risk-neutral probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, where $\mathbb{Q}$ is the risk-neutral measure equivalent to the physical market measure $\mathbb{P}$ via a Radon-Nikodým density whose form is similar to $L(T)$ in Section 2.3.

We first establish our notation before specifying the SVJD model. Let $X_1 = \ln S_1$ and $X_2 = \ln S_2$ denote the log-prices of the underlying assets (stocks or indices) with price processes S_1 and S_2 and let v_1 and v_2 be the stochastic instantaneous variance corresponding to X_1 and X_2, respectively. We collectively refer to these stochastic processes by $\mathbf{Y} = (X_1, v_1, X_2, v_2)^\top$, also referred to as the underlying state process. *The indexing of the elements of* $\mathbf{Y}$ *are important*; in the notation of the stochastic processes that follows, subscripts or superscripts with 1, 2, 3, or 4 (or some pairwise combination) refer to X_1, v_1, X_2, and v_2, respectively. At each t, $\mathbf{Y}(t)$ takes values in the state space $\mathscr{D} = \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$, where $\mathbb{R}^+ = (0, \infty)$.

Suppose the diffusion components of $\mathbf{Y}$ are driven by the vector $\mathbf{W} = (W_1, W_2, W_3, W_4)^\top$ of

$\mathbb{Q}$-Wiener processes with correlation matrix $\mathbf{\Sigma} \in \mathbb{R}^{4\times 4}$,

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{bmatrix}.$$

Assuming that $\mathbf{\Sigma}$ is positive definite, by Cholesky decomposition we can write $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^\top$, where $\mathbf{L} \in \mathbb{R}^{4\times 4}$ is a lower triangular matrix.[1] Let L_{ij} denote the (ij)th element of $\mathbf{L}$, with $L_{ij} = 0$ for $j > i$, $i, j = 1, 2, 3, 4$. We can thus write $\mathbf{W} = \mathbf{L}\mathbf{B}$, where $\mathbf{B}$ is a four-dimensional standard $\mathbb{Q}$-Wiener process.

The jump components are specified as follows:

1. The Poisson counting processes N_1, N_2, N_3, and N_4 refer to the jumps that separately affect X_1, v_1, X_2, and v_2, respectively. The intensities are denoted by λ_1, λ_2, λ_3, and λ_4, respectively, and we assume that these are constants. The random jump sizes corresponding to these counting processes are denoted by J_1, J_2, J_3, and J_4, respectively.

2. The counting process $N_{1,2}$, whose intensity is denoted by $\lambda_{1,2}$, counts the number of jumps affecting both X_1 and v_1. The size of the common jump affecting the log-price is denoted by $J_{1,2}^{(1)}$, whereas that which affects the variance is denoted by $J_{1,2}^{(2)}$. Analogously, we denote by $N_{3,4}$ the Poisson counting process counting the jumps affecting both X_2 and v_2. Its intensity is $\lambda_{3,4}$ and the associated jump sizes affecting X_2 and v_2 are denoted by $J_{3,4}^{(3)}$ and $J_{3,4}^{(4)}$, respectively.

3. The counting process $N_{1,3}$ with intensity $\lambda_{1,3}$ counts the number of jumps affecting both X_1 and X_2. The size of the common jump affecting the X_1 is denoted by $J_{1,3}^{(1)}$, while that which affects X_2 is denoted by $J_{1,3}^{(3)}$.

4. The counting process $N_{2,4}$ with intensity $\lambda_{2,4}$ counts the number of jumps affecting both v_1 and v_2. The size of the common jump affecting the v_1 is denoted by $J_{2,4}^{(2)}$, while that which affects v_2 is denoted by $J_{2,4}^{(4)}$.

We exclude the possibility that the pairs (X_1, v_2) and (X_2, v_1) jump together. A tabular summary of the individual and common jump components is shown in Table 6.1. Whenever necessary, we let

$$\mathbf{N} = (N_1, N_2, N_3, N_4, N_{1,2}, N_{1,3}, N_{2,4}, N_{3,4})^\top \quad \boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_{1,2}, \lambda_{1,3}, \lambda_{2,4}, \lambda_{3,4})^\top$$

[1] At this point, no further assumptions are needed on $\mathbf{\Sigma}$ aside from it being positive definite.

Table 6.1: The Poisson counting processes and their intensity parameters corresponding to individual and common jump components affecting the log-prices X_1, X_2 and the variance processes v_1, v_2. This table is symmetric about the main diagonal.

Process (Index)	X_1 (1)	v_1 (2)	X_2 (3)	v_2 (4)
X_1 (1)	(N_1, λ_1)	$(N_{1,2}, \lambda_{1,2})$	$(N_{1,3}, \lambda_{1,3})$	N/A
v_1 (2)		(N_2, λ_2)	N/A	$(N_{2,4}, \lambda_{2,4})$
X_2 (3)			(N_3, λ_3)	$(N_{3,4}, \lambda_{3,4})$
v_2 (4)				(N_4, λ_4)

collectively refer to the counting processes and their respective intensity parameters. All Poisson counting processes are assumed to be independent of each other and of $\mathbf{B}(t)$.

We equip the space $(\Omega, \mathcal{F}, \mathbb{Q})$ with the right-continuous, complete filtration $\mathbb{F} = \{\mathcal{F}_t\}_{0 \le t \le T}$, where $T > 0$ is a constant and $\mathcal{F}_t$ is the σ-algebra generated by $\mathbf{B}(t)$ and $\mathbf{N}(t)$, $\mathcal{F}_0$ contains all $\mathbb{Q}$-null sets of $\mathcal{F}$, and $\mathcal{F}_T = \mathcal{F}$.

Under $\mathbb{Q}$, the discounted asset prices must be martingales, so we have to include the compensators for the jumps that occur in the log-price processes. To this end, we let $\kappa_1 = \mathbb{E}_{\mathbb{Q}}[\exp\{J_1\} - 1]$, $\kappa_3 = \mathbb{E}_{\mathbb{Q}}[\exp\{J_3\} - 1]$, $\kappa_{1,2}^{(1)} = \mathbb{E}_{\mathbb{Q}}[\exp\{J_{1,2}^{(1)}\} - 1]$, $\kappa_{1,3}^{(1)} = \mathbb{E}_{\mathbb{Q}}[\exp\{J_{1,3}^{(1)}\} - 1]$, $\kappa_{1,3}^{(3)} = \mathbb{E}_{\mathbb{Q}}[\exp\{J_{1,3}^{(3)}\} - 1]$, and $\kappa_{3,4}^{(3)} = \mathbb{E}_{\mathbb{Q}}[\exp\{J_{3,4}^{(3)}\} - 1]$ be the mean jump sizes in the asset prices under $\mathbb{Q}$. We emphasize that at each time t, each jump size variable is sampled from the same probability distribution that does not change with respect to time. Hence the κ's are constant through time.

The dynamics of X_1, v_1, X_2, and v_2 under $\mathbb{Q}$ are described by the following stochastic differential equations,

$$
\begin{aligned}
\mathrm{d}X_1(t) &= \left(r - q_1 - \frac{1}{2}v_1(t-)\right)\mathrm{d}t + \sqrt{v_1(t-)}\mathbf{L}_1\,\mathrm{d}\mathbf{B}(t) + J_1(t)\,\mathrm{d}N_1(t) - \lambda_1\kappa_1\,\mathrm{d}t \\
&\quad + J_{1,2}^{(1)}(t)\,\mathrm{d}N_{1,2}(t) - \lambda_{1,2}\kappa_{1,2}^{(1)}\,\mathrm{d}t + J_{1,3}^{(1)}(t)\,\mathrm{d}N_{1,3}(t) - \lambda_{1,3}\kappa_{1,3}^{(1)}\,\mathrm{d}t, \\
\mathrm{d}v_1(t) &= \xi_1(\eta_1 - v_1(t-))\,\mathrm{d}t + \sigma_1\sqrt{v_1(t-)}\mathbf{L}_2\,\mathrm{d}\mathbf{B}(t) + J_2(t)\,\mathrm{d}N_2(t) \\
&\quad + J_{1,2}^{(2)}(t)\,\mathrm{d}N_{1,2}(t) + J_{2,4}^{(2)}(t)\,\mathrm{d}N_{2,4}(t), \\
\mathrm{d}X_2(t) &= \left(r - q_2 - \frac{1}{2}v_2(t-)\right)\mathrm{d}t + \sqrt{v_2(t-)}\mathbf{L}_3\,\mathrm{d}\mathbf{B}(t) + J_3(t)\,\mathrm{d}N_3(t) - \lambda_3\kappa_3\,\mathrm{d}t \\
&\quad + J_{1,3}^{(3)}(t)\,\mathrm{d}N_{1,3}(t) - \lambda_{1,3}\kappa_{1,3}^{(3)}\,\mathrm{d}t + J_{3,4}^{(3)}(t)\,\mathrm{d}N_{3,4}(t) - \lambda_{3,4}\kappa_{3,4}^{(3)}\,\mathrm{d}t, \\
\mathrm{d}v_2(t) &= \xi_2(\eta_2 - v_2(t-))\,\mathrm{d}t + \sigma_2\sqrt{v_2(t-)}\mathbf{L}_4\,\mathrm{d}\mathbf{B}(t) + J_4(t)\,\mathrm{d}N_4(t) \\
&\quad + J_{2,4}^{(4)}(t)\,\mathrm{d}N_{2,4}(t) + J_{3,4}^{(4)}(t)\,\mathrm{d}N_{3,4}(t),
\end{aligned}
\tag{6.1}
$$

where $\mathbf{Y}(0) = (X_1(0), v_1(0), X_2(0), v_2(0))^\top \in \mathscr{D}$ is given. Here, $\mathbf{L}_i \in \mathbb{R}^{1\times 4}$ is the ith row of $\mathbf{L}$, $r > 0$ is the risk-free interest rate, q_1 and q_2 are the continuously-compounded dividend yields of the two assets, and ξ_1, ξ_2, η_1, η_2, σ_1, and σ_2 are positive constants. For the diffusion part of

the variance processes to remain strictly positive and finite almost surely under $\mathbb{Q}$, we impose the Feller condition $2\kappa_i\xi_i \geq \sigma_i^2$, $i = 1, 2$. At each time $t \in [0, T]$, the jump size variables J_1, J_3, $J_{1,2}^{(1)}$, $J_{1,3}^{(1)}$, $J_{1,3}^{(3)}$, and $J_{3,4}^{(3)}$ are assumed to have a normal distribution under $\mathbb{Q}$, although other jump size distributions may also be assumed. In contrast, the jumps occurring in the variances, namely J_2, J_4, $J_{1,2}^{(2)}$, $J_{2,4}^{(2)}$, $J_{2,4}^{(4)}$, and $J_{3,4}^{(4)}$, are required to be positive, as implied by the square-root structure. As such, the variance jumps are typically drawn from an exponential distribution (Aït-Sahalia et al., 2018; Bardgett et al., 2019; Da Fonseca and Ignatieva, 2019). All jump size random variables are independent of $\mathbf{B}(t)$ and $\mathbf{N}(t)$. The dynamics of the state variables can be expressed more succinctly in matrix form as

$$\mathrm{d}\mathbf{Y}(t) = \boldsymbol{\alpha}(\mathbf{Y}(t-))\,\mathrm{d}t + \boldsymbol{\beta}(\mathbf{Y}(t-))\,\mathrm{d}\mathbf{B}(t) + \boldsymbol{\gamma}(t)\,\mathrm{d}\mathbf{N}(t), \tag{6.2}$$

where

$$\boldsymbol{\alpha}(\mathbf{Y}(t-)) = \begin{bmatrix} r - q_1 - \lambda_1\kappa_1 - \lambda_{1,2}\kappa_{1,2}^{(1)} - \lambda_{1,3}\kappa_{1,3}^{(1)} - \frac{1}{2}v_1(t-) \\ \xi_1\eta_1 - \xi_1 v_1(t-) \\ r - q_2 - \lambda_3\kappa_3 - \lambda_{1,3}\kappa_{1,3}^{(3)} - \lambda_{3,4}\kappa_{3,4}^{(3)} - \frac{1}{2}v_2(t-) \\ \xi_2\eta_2 - \xi_2 v_2(t-) \end{bmatrix},$$

$$\boldsymbol{\beta}(t) = \mathbf{V}(\mathbf{Y}(t-))\mathbf{L}, \quad \text{where} \quad \mathbf{V} = \begin{bmatrix} \sqrt{v_1(t-)} & 0 & 0 & 0 \\ 0 & \sigma_1\sqrt{v_1(t-)} & 0 & 0 \\ 0 & 0 & \sqrt{v_2(t-)} & 0 \\ 0 & 0 & 0 & \sigma_2\sqrt{v_2(t-)} \end{bmatrix},$$

and

$$\boldsymbol{\gamma}(t) = \begin{bmatrix} J_1(t) & 0 & 0 & 0 & J_{1,2}^{(1)}(t) & J_{1,3}^{(1)}(t) & 0 & 0 \\ 0 & J_2(t) & 0 & 0 & J_{1,2}^{(2)}(t) & 0 & J_{2,4}^{(2)}(t) & 0 \\ 0 & 0 & J_3(t) & 0 & 0 & J_{1,3}^{(3)}(t) & 0 & J_{3,4}^{(3)}(t) \\ 0 & 0 & 0 & J_4(t) & 0 & 0 & J_{2,4}^{(4)}(t) & J_{3,4}^{(4)}(t) \end{bmatrix}.$$

The process $\mathbf{Y}$ is an affine jump-diffusion (AJD) if the elements of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}\boldsymbol{\beta}^\top$, $\boldsymbol{\lambda}$, and r are affine functions on $\mathscr{D}$ (i.e. affine functions of the components of $\mathbf{Y}$) (Duffie et al., 2000) (see also the notation introduced in equation (E.2)). From the model specification above, it is clear that $\boldsymbol{\alpha}$ satisfies this requirement, as do $\boldsymbol{\lambda}$ and r since they are assumed to be constant. However, we note that

$$\boldsymbol{\beta}(t)\boldsymbol{\beta}^\top(t) = \mathbf{V}\mathbf{L}(\mathbf{V}\mathbf{L})^\top = \mathbf{V}\mathbf{L}\mathbf{L}^\top\mathbf{V}^\top = \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}$$

$$= \begin{bmatrix} v_1(t-) & \sigma_1 v_1(t-)\rho_{12} & \sqrt{v_1(t-)v_2(t-)}\rho_{13} & \sigma_2\sqrt{v_1(t-)v_2(t-)}\rho_{14} \\ \sigma_1 v_1(t-)\rho_{12} & \sigma_1^2 v_1(t-) & \sigma_1\sqrt{v_1(t-)v_2(t-)}\rho_{23} & \sigma_1\sigma_2\sqrt{v_1(t-)v_2(t-)}\rho_{24} \\ \sqrt{v_1(t-)v_2(t-)}\rho_{13} & \sigma_1\sqrt{v_1(t-)v_2(t-)}\rho_{23} & v_2(t-) & \sigma_2 v_2(t-)\rho_{34} \\ \sigma_2\sqrt{v_1(t-)v_2(t-)}\rho_{14} & \sigma_1\sigma_2\sqrt{v_1(t-)v_2(t-)}\rho_{24} & \sigma_2 v_2(t-)\rho_{34} & \sigma_2^2 v_2(t-) \end{bmatrix}.$$

Thus, for $\mathbf{Y}$ to be an AJD, we impose further assumptions on the correlation matrix of $\mathbf{W}$, specifically

$$\rho_{13} = \rho_{23} = \rho_{14} = \rho_{24} = 0. \tag{6.3}$$

The condition $\rho_{13} = 0$ and $\rho_{24} = 0$ implies that the diffusive components of the pairs (X_1, X_2) and (v_1, v_2) are uncorrelated.[2] To make up for this restrictive correlation structure on the continuous parts of the pairs (X_1, X_2) and (v_1, v_2), we allow the jump size pairs $(Y^{(1)}_{1,3}, Y^{(3)}_{1,3})$ and $(Y^{(2)}_{2,4}, Y^{(4)}_{2,4})$ to exhibit joint dependence.

The risk-neutral measure $\mathbb{Q}$ is defined via a straightforward extension of the Radon-Nikodým derivative considered in Chapter 2. The Radon-Nikodým derivative is carefully constructed such that $\mathbf{Y}$ has an affine structure in both the historical and risk-neutral measures, just with different coefficients. We base our characterization of the risk-neutral measure on that of Cheang and Chiarella (2011) whose definition considers changes in the intensity and probability distribution of both individual and joint jumps.

We first introduce the relevant stochastic processes and parameters that are used to define the Radon-Nikodým process. Let $\mathbf{W}^{\mathbb{P}} = (W_1^{\mathbb{P}}, W_2^{\mathbb{P}}, W_3^{\mathbb{P}}, W_4^{\mathbb{P}})^\top$ be a vector of $\mathbb{P}$-Wiener processes with correlation matrix $\mathbf{\Sigma}$. Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4)^\top$ be a vector of real-valued predictable processes and for $k = 1, 2, 3, 4$ let $\nu_k, \varrho_k \in \mathbb{R}$ be constants. Let $\nu_{1,2}, \nu_{1,3}, \nu_{2,4}, \nu_{3,4} \in \mathbb{R}$ be constants as well. We also consider the two-dimensional vectors $\boldsymbol{\varrho}_{1,2} = (\varrho^{(1)}_{1,2}, \varrho^{(2)}_{1,2})^\top$, $\boldsymbol{\varrho}_{1,3} = (\varrho^{(1)}_{1,3}, \varrho^{(3)}_{1,3})^\top$, $\boldsymbol{\varrho}_{2,4} = (\varrho^{(2)}_{2,4}, \varrho^{(4)}_{2,4})^\top$, and $\boldsymbol{\varrho}_{3,4} = (\varrho^{(3)}_{3,4}, \varrho^{(4)}_{3,4})^\top$ whose entries are real constants. The joint jump size variables shall also be written in vector notation as $\mathbf{J}_{1,2} = (J^{(1)}_{1,2}, J^{(2)}_{1,2})^\top$, $\mathbf{J}_{1,3} = (J^{(1)}_{1,3}, J^{(3)}_{1,3})^\top$, $\mathbf{J}_{2,4} = (J^{(2)}_{2,4}, J^{(4)}_{2,4})^\top$, and $\mathbf{J}_{3,4} = (J^{(3)}_{3,4}, J^{(4)}_{3,4})^\top$. We introduce the index set $\mathcal{I} = \{(1,2), (1,3), (2,4), (3,4)\}$ containing the indices of the state variables that possibly have joint jumps as shorthand for the subscripts of the aforementioned parameters.

[2] These restrictions on the correlation matrix of $\mathbf{W}$ were also necessary to obtain a closed form of the joint characteristic function of $\mathbf{Y}$ (see Section 2.5).

We consider the process $L = \{L(t)\}_{t\in[0,T]}$ where

$$\begin{aligned} L(t) = \exp\left\{ -\int_0^t (\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s))^\top \,\mathrm{d}\mathbf{W}^{\mathbb{P}}(s) - \frac{1}{2}\int_0^t \boldsymbol{\theta}^\top(s)\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}^\top(s)\,\mathrm{d}s \right\} \\ \times \prod_{k=1}^{4} \exp\left\{ -\lambda_k^{\mathbb{P}} t\left(e^{\nu_k}\mathbb{E}_{\mathbb{P}}[e^{\varrho_k J_k}] - 1\right) + \sum_{n=1}^{N_k(t)} \left(\varrho_k J_k(n) + \nu_k\right) \right\} \\ \times \prod_{l\in\mathcal{I}} \exp\left\{ -\lambda_l^{\mathbb{P}} t\left(e^{\nu_l}\mathbb{E}_{\mathbb{P}}[e^{\boldsymbol{\varrho}_l^\top \mathbf{J}_l}] - 1\right) + \sum_{n=1}^{N_l(t)} \left(\boldsymbol{\varrho}_l^\top \mathbf{J}_l(n) + \nu_l\right) \right\}, \end{aligned} \tag{6.4}$$

where, for $k = 1, 2, 3, 4$ and $l \in \mathcal{I}$, $\lambda_k^{\mathbb{P}}$ and $\lambda_l^{\mathbb{P}}$ are the jump intensity of N_k and N_l, respectively, and $\{J_k(n)\}_{n\in\mathbb{N}}$ and $\{\mathbf{J}_l(n)\}_{n\in\mathbb{N}}$ are i.i.d. sequences of random variables sampled from the distribution of J_k and $\mathbf{J}_l$, respectively.

Supposing the process $L = \{L(t)\}_{t\in[0,T]}$, defined in equation (6.4), is a $\mathbb{P}$-martingale and satisfies $\mathbb{E}_{\mathbb{P}}[L(T)] = 1$ and defining the probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ by $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = L(t)$. Then:

1. For $k = 1, 2, 3, 4$, the process $W_k(t) = \int_0^t \theta_k(s)\,\mathrm{d}s + W_k^{\mathbb{P}}(t)$ is a Wiener process under $\mathbb{Q}$. Furthermore, the vector process $\mathbf{W} = (W_1, W_2, W_3, W_4)^\top$ has the same correlation matrix $\boldsymbol{\Sigma}$.

2. For $k = 1, 2, 3, 4$, the intensity of N_k under $\mathbb{Q}$ is

$$\lambda_k = \lambda_k^{\mathbb{P}} e^{\nu_k}\mathbb{E}_{\mathbb{P}}\left[e^{\varrho_k J_k}\right],$$

   and for $l \in \mathcal{I}$, the intensity of N_l under $\mathbb{Q}$ is

$$\lambda_l = \lambda_l^{\mathbb{P}} e^{\nu_l}\mathbb{E}_{\mathbb{P}}\left[e^{\boldsymbol{\varrho}_l^\top J_l}\right].$$

3. For $k = 1, 2, 3, 4$, the moment generating function (mgf) of J_k under $\mathbb{Q}$ is given by

$$M_{J_k}^{\mathbb{Q}}(u) = \frac{M_{J_k}^{\mathbb{P}}(u + \varrho_k)}{M_{J_k}^{\mathbb{P}}(\varrho_k)},$$

   and for $l \in \mathcal{I}$, the (joint) mgf of $\mathbf{J}_l$ under $\mathbb{Q}$ is

$$M_{\mathbf{J}_l}^{\mathbb{Q}}(\mathbf{u}) = \frac{M_{\mathbf{J}_l}^{\mathbb{P}}(\mathbf{u} + \boldsymbol{\varrho}_l)}{M_{\mathbf{J}_l}^{\mathbb{P}}(\boldsymbol{\varrho}_l)}, \qquad \mathbf{u} \in \mathbb{R}^2.$$

The discussion above describes a Girsanov-type theorem for the change of measure from $\mathbb{P}$ to $\mathbb{Q}$. A proof of this result for the case of no joint jumps is provided in Chapter 2. For a proof of the results pertaining to the joint jump moment generating function, see Cheang and Teh (2014).

The parameters of the Radon-Nikodým derivative (6.4), specifically the ν's and the ϱ's, can be chosen depending on the desired effect of the change of measure on the jump intensity or the jump

size distribution. In empirical studies that aim to extract the premia for jump and volatility risk (Pan, 2002; Eraker, 2004; Aït-Sahalia et al., 2018; Bardgett et al., 2019), it is assumed that the jump intensities do not change following the transition from $\mathbb{P}$ to $\mathbb{Q}$ so as to avoid introducing what is called a *jump timing risk premium*, which is separate from the premium arising from the jump size uncertainty. The main concern is that these two types of jump risk are difficult to disentangle from one another using historical stock or index prices (Pan, 2002). One way to reflect this assumption is to choose $\nu_k = -\ln \mathbb{E}_{\mathbb{P}}[e^{\varrho_k J_k}]$ for the individual jump intensities and $\nu_l = -\ln \mathbb{E}_{\mathbb{P}}[e^{\boldsymbol{\varrho}_l^\top \mathbf{J}_l}]$ for the joint jump intensities. In this situation, the jump risk premium computed from historical time series incorporates both jump timing and size uncertainty.

Once the parameters affecting the jumps have been determined, the parameter $\boldsymbol{\theta}$ is chosen so that the discounted asset yield processes $\{S_1(t)e^{-(r-q_1)t}\}$ and $\{S_2(t)e^{-(r-q_2)t}\}$ are martingales under $\mathbb{Q}$ (see the discussion following Proposition 2.3). This results to $\mathbb{Q}$ being an equivalent martingale measure or, specifically, the risk-neutral probability measure.

If the jump sizes are assumed to be sampled from distributions in the exponential family, the Esscher-type transformation in the moment generating functions result to jump sizes having the same probability distribution under $\mathbb{Q}$ but with different parameters (Gerber and Shiu, 1994).

As mentioned before, the key property of this specification of the Radon-Nikodým process is that it preserves the affine structure of the stochastic dynamics of the state variables. Related calculations reinforcing this observation have been presented in Sections 2.3 and 2.5.2, where the log-price and instantaneous variance processes with risk-neutral dynamics similar to (6.1) (see equations (2.37) and (2.38)) have been shown to originate from affine processes under $\mathbb{P}$ (see equations (2.2) and (2.5)) provided that $\mathbb{Q}$ is an equivalent martingale measure.

6.3 Affine Dependence of the Volatility Index

Having specified an affine stochastic volatility and jump-diffusion model for the log-prices and the instantaneous variances, we now consider the analytical characterization of the VIX and its connection to the integrals of the variance processes. A measure of the 30-day expected volatility of the S&P 500 index, the squared-VIX is calculated as the weighted sum out-of-the-money call and put options on the S&P 500 index with more than 23 but less than 37 days until expiration (details can be found in Carr and Wu, 2006; Cboe Global Markets Inc., 2019). The VIX calculation involves market prices of options, so its characterization is model-free. Alternatively, the squared-VIX can be characterized as the risk-neutral expectation of the log contract

$$\mathsf{VIX}^2(t) = -\frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\frac{F(t+\tau, t+\tau)}{F(t, t+\tau)}\right], \tag{6.5}$$

where τ is 30 days expressed in annual terms (using the appropriate day-count convention), $F(t, t+\tau)$ is the futures price, deliverable at time $t+\tau$, of the S&P 500 index observed at time t, and $\mathbb{E}_{\mathbb{Q}}^{(t)}[\cdot] = \mathbb{E}_{\mathbb{Q}}[\cdot|\mathcal{F}_t]$ is the risk-neutral $\mathcal{F}_t$-conditional expectation operator (Lin, 2007; Zhu and Lian, 2011; Aït-Sahalia et al., 2018). Since we assume that the risk-free rate is a known constant, the futures price can be calculated as the forward price on the S&P 500; that is,

$$F(t, t+\tau) = S(t)e^{(r-q)\tau} = S(t)e^{(r-q)(\bar{T}-t)},$$

where $\bar{T} = t+\tau$. As such, $F(t+\tau, t+\tau)$ is the same as the spot price observed at time $t+\tau$.

We further pursue equation (6.5) to obtain analytical representations of VIX-type indices for underlying assets whose respective price processes are given by (6.1). Since the calculations are identical for both assets, we present only the details for a VIX-type index corresponding to X_1. From the definition of the forward price, the log-futures price $\ln F_1(t, t+\tau)$ satisfies the differential equation

$$\begin{aligned}
\mathrm{d}\ln F_1(t,\bar{T}) &= \mathrm{d}X_1(t) - (r-q)\,\mathrm{d}t \\
&= -\frac{1}{2}v_1(t-)\,\mathrm{d}t + \sqrt{v_1(t-)}\mathbf{L}_1\,\mathrm{d}\mathbf{B}(t) + J_1(t)\,\mathrm{d}N_1(t) - \lambda_1\kappa_1\,\mathrm{d}t \\
&\quad + J_{1,2}^{(1)}(t)\,\mathrm{d}N_{1,2}(t) - \lambda_{1,2}\kappa_{1,2}^{(1)}\,\mathrm{d}t + J_{1,3}^{(1)}(t)\,\mathrm{d}N_{1,3}(t) - \lambda_{1,3}\kappa_{1,3}^{(1)}\,\mathrm{d}t.
\end{aligned}$$

The squared VIX process is then given by

$$\begin{aligned}
\mathsf{VIX}_1^2(t) &= -\frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\frac{F_1(\bar{T},\bar{T})}{F_1(t,\bar{T})}\right] = -\frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}}\mathrm{d}\ln F_1(u,\bar{T})\right] \\
&= \frac{1}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} v_1(u-)\,\mathrm{d}u\right] - \frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}}\sqrt{v_1(u-)}\mathbf{L}_1\,\mathrm{d}\mathbf{B}(u)\right] \\
&\quad - \frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} J_1(u)\,\mathrm{d}N_1(u) - \int_t^{\bar{T}}\lambda_1\kappa_1\,\mathrm{d}u\right] \\
&\quad - \frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} J_{1,2}^{(1)}(u)\,\mathrm{d}N_{1,2}(u) - \int_t^{\bar{T}}\lambda_{1,2}\kappa_{1,2}^{(1)}\,\mathrm{d}u\right] \\
&\quad - \frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} J_{1,3}^{(1)}(u)\,\mathrm{d}N_{1,3}(u) - \int_t^{\bar{T}}\lambda_{1,3}\kappa_{1,3}^{(1)}\,\mathrm{d}u\right].
\end{aligned}$$

By the martingale property of integrals with respect to Wiener processes, the second conditional expectation above reduces to 0. Furthermore, by the independence of increments of N_1 over $[t,\bar{T}]$ from $\mathcal{F}_t$ and the projection theorem for marked Poisson processes (see for example, Brémaud, 1981, Theorem T3), we have

$$\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} J_1(u)\,\mathrm{d}N_1(u)\right] = \mathbb{E}_{\mathbb{Q}}\left[\int_t^{\bar{T}} J_1(u)\,\mathrm{d}N_1(u)\right]$$

$$= \mathbb{E}_{\mathbb{Q}}\left[\int_t^{\bar{T}} \mathbb{E}_{\mathbb{Q}}[J_1(u)]\lambda_1 \,\mathrm{d}u\right]$$
$$= \mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} \mathbb{E}_{\mathbb{Q}}[J_1(u)].\lambda_1 \,\mathrm{d}u\right]$$

Therefore,

$$-\frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} J_1(u)\,\mathrm{d}N_1(u) - \int_t^{\bar{T}} \lambda_1\kappa_1 \,\mathrm{d}u\right] = -\frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} \mathbb{E}_{\mathbb{Q}}[J_1(u)]\lambda_1 \,\mathrm{d}u - \int_t^{\bar{T}} \lambda_1\kappa_1 \,\mathrm{d}u\right]$$
$$= \frac{2}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{\bar{T}} \mathbb{E}_{\mathbb{Q}}\left[e^{J_1(u)} - 1 - J_1(u)\right]\lambda_1 \,\mathrm{d}u\right]$$
$$= 2\lambda_1\zeta_1,$$

where $\zeta_1 = \mathbb{E}_{\mathbb{Q}}[e^{J_1(u)} - 1 - J_1(u)] = \mathbb{E}_{\mathbb{Q}}[e^{J_1} - 1 - J_1]$, which is constant through u. A similar analysis on the other terms allows us to write the VIX as

$$\mathsf{VIX}_1^2(t) = \frac{1}{\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}\left[\int_t^{t+\tau} v_1(u-)\,\mathrm{d}u\right] + 2\left(\lambda_1\zeta_1 + \lambda_{1,2}\zeta_{1,2}^{(1)} + \lambda_{1,3}\zeta_{1,3}^{(1)}\right), \tag{6.6}$$

where $\zeta_{1,2}^{(1)} = \mathbb{E}_{\mathbb{Q}}[e^{J_{1,2}^{(1)}} - 1 - J_{1,2}^{(1)}]$ and $\zeta_{1,3}^{(1)} = \mathbb{E}_{\mathbb{Q}}[e^{J_{1,3}^{(1)}} - 1 - J_{1,3}^{(1)}]$.

The above analysis reveals that the squared VIX has an affine dependence on the conditional risk-neutral expected integrated variance over $[t, t+\tau]$. In the absence of jumps in the log-price X_1, the squared VIX coincides with the conditional $\mathbb{Q}$-expectation of the annualized realized variance over the next 30 calendar days. The jump terms, when nonzero, represent the error with which the VIX approximates the 30-day realized variance (Carr and Wu, 2006).

The affine structure of the variance process v_1 allows us to evaluate the conditional expectation in equation (6.6). Given that $v_1(t)$ is a nonnegative measurable function with respect to the Lebesgue measure on $[t, t+\tau]$ and the risk-neutral measure $\mathbb{Q}$, Tonelli's Theorem (see for example, Cohen and Elliott, 2015, Theorem 1.4.6) allows us to interchange the order of integration. Thus, we can compute $\mathbb{E}_{\mathbb{Q}}^{(t)}[v_1(u)]$ first before integrating with respect to u over $[t, t+\tau]$. To do so, we first introduce the compensators of the jump terms in $\mathrm{d}v_1(t)$; that is, we write

$$\mathrm{d}v_1(t) = \xi_1(\eta_1 - v_1(t-))\,\mathrm{d}t + \lambda_2\kappa_2\,\mathrm{d}t + \lambda_{1,2}\kappa_{1,2}^{(2)}\,\mathrm{d}t + \lambda_{2,4}\kappa_{2,4}^{(2)}\,\mathrm{d}t + \sigma_1\sqrt{v_1(t-)}\mathbf{L}_2\,\mathrm{d}\mathbf{B}(t)$$
$$+ J_2(t)\,\mathrm{d}N_2(t) - \lambda_2\kappa_2\,\mathrm{d}t + J_{1,2}^{(2)}(t)\,\mathrm{d}N_{1,2}(t) - \lambda_{1,2}\kappa_{1,2}^{(2)}\,\mathrm{d}t + J_{2,4}^{(2)}(t)\,\mathrm{d}N_{2,4}(t) - \lambda_{2,4}\kappa_{2,4}^{(2)}\,\mathrm{d}t,$$

where $\kappa_2 = \mathbb{E}_{\mathbb{Q}}[e^{J_2} - 1]$, $\kappa_{1,2}^{(2)} = \mathbb{E}_{\mathbb{Q}}[e^{J_{1,2}^{(2)}} - 1]$, and $\kappa_{2,4}^{(2)} = \mathbb{E}_{\mathbb{Q}}[e^{J_{2,4}^{(2)}} - 1]$. Considering the $\mathbb{Q}$-martingale $M_{v_1}(t)$ whose increment is

$$\mathrm{d}M_{v_1}(t) = \sigma_1\sqrt{v_1(t-)}\mathbf{L}_2\,\mathrm{d}\mathbf{B}(t) + J_2(t)\,\mathrm{d}N_2(t) - \lambda_2\kappa_2\,\mathrm{d}t + J_{1,2}^{(2)}(t)\,\mathrm{d}N_{1,2}(t) - \lambda_{1,2}\kappa_{1,2}^{(2)}\,\mathrm{d}t$$

$$+ J_{2,4}^{(2)}(t)\,\mathrm{d}N_{2,4}(t) - \lambda_{2,4}\kappa_{2,4}^{(2)}\,\mathrm{d}t,$$

we can write

$$\mathrm{d}v_1(t) = \xi_1(\tilde{\eta}_1 - v_1(t-))\,\mathrm{d}t + \mathrm{d}M_{v_1}(t),$$

where $\tilde{\eta}_1 = \eta_1 + \frac{1}{\xi_1}(\lambda_2\kappa_2 + \lambda_{1,2}\kappa_{1,2}^{(2)} + \lambda_{2,4}\kappa_{2,4}^{(2)})$. The affine dependence of the jump intensities allows us to write the increment of v_1 as above.[3] Since in our model specification the intensity parameters are constant, this operation is straightforward. Applying the Itô formula to $e^{\xi_1 t}v_1(t)$, we obtain

$$\begin{aligned}\mathrm{d}(e^{\xi_1 t}v_1(t)) &= v_1(t-)\xi_1 e^{\xi_1 t}\,\mathrm{d}t + e^{\xi_1 t}\xi_1(\tilde{\eta}_1 - v_1(t-))\,\mathrm{d}t + e^{\xi_1 t}\,\mathrm{d}M_{v_1}(t)\\ &= e^{x_1 t}\xi_1\tilde{\eta}_1\,\mathrm{d}t + e^{\xi_1 t}\,\mathrm{d}M_{v_1}(t),\end{aligned}$$

and integrating over $[t,u]$ and rearranging the terms, we have

$$v_1(u) = e^{-\xi_1(u-t)}v_1(t) + \tilde{\eta}_1\left[1 - e^{-\xi_1(u-t)}\right] + \int_t^u e^{-\xi_1(u-s)}\,\mathrm{d}M_{v_1}(s).$$

Taking the conditional expectation $\mathbb{E}_{\mathbb{Q}}^{(t)}[\cdot]$ yields

$$\mathbb{E}_{\mathbb{Q}}^{(t)}[v_1(u)] = e^{-\xi_1(u-t)}v_1(t) + \tilde{\eta}_1\left[1 - e^{-\xi_1(u-t)}\right],$$

since the conditional expectation of the integral with respect to M_{v_1} is 0. Finally, integrating with respect to u over $[t,t+\tau]$ results to

$$\begin{aligned}\int_t^{t+\tau}\mathbb{E}_{\mathbb{Q}}^{(t)}[v_1(u)]\,\mathrm{d}u &= e^{\xi_1 t}v_1(t)\int_t^{t+\tau}e^{-\xi_1 u}\,\mathrm{d}u + \tilde{\eta}_1\tau - \tilde{\eta}_1 e^{\xi_1 t}\int_t^{t+\tau}e^{-\xi_1 u}\,\mathrm{d}u\\ &= e^{\xi_1 t}v_1(t)\left[-\frac{e^{-\xi_1(t+\tau)}}{\xi_1} + \frac{e^{-\xi_1 t}}{\xi_1}\right] + \tilde{\eta}_1\tau - \tilde{\eta}_1 e^{\xi_1 t}\left[-\frac{e^{-\xi_1(t+\tau)}}{\xi_1} + \frac{e^{-\xi_1 t}}{\xi_1}\right]\\ &= \left(\frac{1-e^{-\xi_1\tau}}{\xi_1}\right)v_1(t) + \frac{\tilde{\eta}_1}{\xi_1}\left(\xi_1\tau + e^{-\xi_1\tau} - 1\right).\end{aligned}$$

Thus, we have shown that, given our affine model specification, the squared VIX has an affine dependence on the instantaneous variance process,

$$\mathsf{VIX}_1^2(t) = \phi_1 + \varphi_1 v_1(t), \tag{6.7}$$

where

$$\phi_1 = \frac{1-e^{-\xi_1\tau}}{\xi_1\tau} \quad \text{and} \quad \varphi_1 = \frac{\tilde{\eta}_1}{\xi_1\tau}\left(\xi_1\tau + e^{-\xi_1\tau} - 1\right) + 2\left(\lambda_1\zeta_1 + \lambda_{1,2}\zeta_{1,2}^{(1)} + \lambda_{1,3}\zeta_{1,3}^{(1)}\right).$$

[3] For the case that the intensity is an affine function of the instantaneous variance, that is, $\lambda_2(t) = \alpha_0 + \alpha_1 v_1(t)$ for some $\alpha_0, \alpha_1 \in \mathbb{R}$, Lin (2007) and Aït-Sahalia et al. (2018) (among others) provide details of these calculations.

These coefficients are constants.

By similar calculations, it follows that the square of a VIX-type index corresponding to X_2 can be written as

$$\mathsf{VIX}_2^2(t) = \phi_2 + \varphi_2 v_2(t),$$

where

$$\phi_2 = \frac{1 - e^{-\xi_2 \tau}}{\xi_2 \tau} \quad \text{and} \quad \varphi_2 = \frac{\tilde{\eta}_2}{\xi_2 \tau}\left(\xi_2 \tau + e^{-\xi_2 \tau} - 1\right) + 2\left(\lambda_3 \zeta_3 + \lambda_{1,3}\zeta_{1,3}^{(3)} + \lambda_{3,4}\zeta_{3,4}^{(3)}\right),$$

with $\tilde{\eta}_2 = \eta_2 + \frac{1}{\xi_2}(\lambda_4 \kappa_4 + \lambda_{2,4}\kappa_{2,4}^{(4)} + \lambda_{3,4}\kappa_{3,4}^{(4)})$, $\zeta_3 = \mathbb{E}_{\mathbb{Q}}[e^{J_3} - 1 - J_3]$, $\zeta_{1,3}^{(3)} = \mathbb{E}_{\mathbb{Q}}[e^{J_{1,3}^3} - 1 - J_{1,3}^{(3)}]$, and $\zeta_{3,4}^{(3)} = \mathbb{E}_{\mathbb{Q}}[e^{J_{3,4}^{(3)}} - 1 - J_{3,4}^{(3)}]$.

The affine structure of the squared VIX (and by extension the assumption of an affine SVJD model for the state variables) is crucial for two reasons. Firstly, if the risk-neutral model parameters are known, the instantaneous variance, which is a latent variable, can be recovered from observations of the VIX via equation (6.7) (see for example, Duan and Yeh, 2010). In the same spirit, the market price of jump risk and the variance risk premium can also be estimated from the difference between the variance swap rate and the VIX (Aït-Sahalia et al., 2018; Bardgett et al., 2019). Secondly, equation (6.7) implies that any jumps detected from the squared VIX time series directly relates to jumps that occur in the instantaneous variance process. These points serve as the basis of a number of jump activity analyses on both empirical asset or commodity prices that exhibit stochastic volatility (see for example, Aït-Sahalia et al., 2018; Bardgett et al., 2019; Da Fonseca and Ignatieva, 2019)

6.4 Chapter Summary and Conclusion

The application of these parameter estimation and calibration methods (see Appendix E) to the two-asset stochastic volatility and jump-diffusion models used in this thesis is not straightforward for at least two reasons. First, the instantaneous variance is a latent or unobservable variable, and parameter estimation methods with unobserved variables are substantially more challenging to implement that with a fully-observed process. Second, it is typically the case that there is no sufficient market information available with respect to which the multi-asset model can be calibrated.

As an initial step towards a solution to these issues, we considered the affine property of the asset pricing models we used in this thesis. The affine property is crucial because there exist a number of parameter estimation methods that are able to estimate the asset pricing model despite some components being unobservable (see for example, Jiang and Knight, 2002; Eraker et al., 2003; Bates, 2006). In addition, it is well-known that if there exists a volatility index corresponding to

the underlying asset and if the underlying asset dynamics has an affine structure, then we can use the volatility index as a proxy for the unobservable instantaneous variance. In terms of model calibration, there exist a number of methods that can be used despite the lack of market data that can properly account for the dependency structures between asset price and instantaneous variance processes. These considerations can then serve as the basis for a complete parameter estimation and calibration analysis of multi-asset stochastic volatility and jump-diffusion models as future research.

CHAPTER 7

Conclusions and Directions for Future Research

7.1 Thesis Summary and Conclusions

This thesis tackles the problem of pricing European and American exchange options when the underlying asset prices exhibit both jumps and stochastic volatility. Numerous empirical analyses on time series of index and stock prices (see Section 1.2) have concluded that the classical Black and Scholes (1973) and Merton (1973) assumptions on asset price dynamics are insufficient to capture the stylized empirical characteristics of asset returns, thus reinforcing the need to consider models with jumps and/or stochastic volatility for more accurate asset pricing. The problem of pricing single-asset European options under more general asset price dynamics is a well-researched topic in financial mathematics, but extensions to the multi-asset case and to the case of American-style or early exercise options are not straightforward. Thus, this thesis contributes to the literature by showing how the European and American exchange option, one of the simplest two-asset options, can be priced under a stochastic volatility and jump-diffusion (SVJD) asset price model. In particular, we provided probabilistic and integral representations of the European and American exchange option prices, conducted an in-depth examination of the behavior of the early exercise boundary near maturity, proposed numerical methods to approximate exchange option prices from the option pricing IPDEs, and analyzed the effect of model parameters on the exchange option prices and the early exercise boudnary.

In Chapter 2, we presented a two-asset version of the Bates (1996) (SVJD) model, discussed the construction of equivalent martingale measures corresponding to this model, and derived the exchange option pricing IPDEs. We then derived representations of the European exchange option price in two ways. The first method uses a change-of-numéraire technique to express the European exchange option price in terms of probabilities that the option is in-the-money at maturity. The

second method expresses the European exchange option in terms of a Fourier inversion formula using the results obtained by Caldana and Fusai (2013) for European spread options. However, the second method hinges on the availability in closed form of the joint characteristic function of the asset price and instantaneous variances. We thus had to make additional assumptions on the correlation structure of the Brownian motions in the model to ensure an affine structure for the model, thereby allowing us to compute the joint characteristic function in closed form (Duffie et al., 2000). We also note that our results substantially differ from the jump-diffusion case (Cheang and Chiarella, 2011) due to the presence of stochastic volatility, preventing us from obtaining infinite series representations of the European exchange option prices. We then obtained an early exercise representation of the American exchange option price by taking advantage of the behavior of the American exchange option price in the stopping and continuation regions. We find, however, that an explicit formula for the American exchange option is not available due to the unknown early exercise boundary. Nonetheless, we were able to show that the early exercise premium is significantly affected by the presence of jumps (a phenomenon referred to as rebalancing costs by Chiarella and Ziogas (2009)), as was the case for the single-asset American option. We then expressed the American exchange option price and the early exercise boundary as the solution of a joint system of integral equations, which would require numerical methods to solve.

In Chapter 3, we considered an alternative SVJD model specification where the two asset prices depend on the same stochastic volatility process. We referred to this model as the proportional SVJD model. The main mathematical advantage of this model is that it is affine without having to impose additional simplifying assumptions on the correlation structure of the diffusion components. Taking advantage of this property, we detailed a solution of the pricing IPDEs using Fourier and Laplace integral transforms. This allowed us to obtain integral representations of the European and American exchange option prices which are similar in structure to the representations derived in Chapter 2. These integral representations can then be evaluated numerical using quadrature methods (see for example, Chiarella and Ziogas, 2009; Adolfsson et al., 2013; Chiarella and Ziveyi, 2013). The method shown in Chapter 3 can also be applied to obtain integral representations of prices of options with payoffs that do not depend on the variance level.

Our analysis in Chapter 4 further simplified the exchange option pricing problem, specifically the problem of pricing American exchange options, by implementing the put-call transformation technique proposed by Bjerskund and Stensland (1993), which we extended to the case of stochastic volatility and jump-diffusion. Because the exchange option payoff is homogeneous in the asset prices, the put-call transformation technique allowed us to reduce the problem to pricing an exchange option on the asset yield ratio (i.e. the ratio of the asset yield processes of the two assets). We then detailed the construction of the equivalent martingale measure corresponding to the second asset yield process as the numéraire and derived the corresponding pricing IPDE. In contrast

to Chapter 3, we derived and solved, using Fourier and Laplace integral transforms, the Kolmogorov backward equation for the joint transition density function for the asset yield ratio and the instantaneous variance process. With the transition density function, we were able to derive representations for the European and American exchange option prices and the early exercise premium similar to those obtained in Chapter 3. We also analyzed in further detail the behavior of the early exercise boundary near maturity, given the simplified IPDE resulting from the put-call transformation. We showed, with numerical illustrations for the case of normally distributed jump sizes, that the early exercise boundary limit is significantly affected by the dividend yields and the presence of jumps. In particular, we find that the early exercise boundary increases when the jump intensity for either asset increases. This is consistent with the diffusive and jump components of the early exercise premium obtained in Chapters 2 and 3, where American option holders only tend to exercise for higher values of the asset yield ratio, since instantaneous jumps may occur upon early exercise, resulting to a suboptimal pay-out. Since the payoff of the exchange option not dependent on volatility, the volatility parameters do not play a role in determining the limit of the early exercise boundary. The put-call transformation shown in this chapter can also be applied to price options whose payoff function is homogeneous in the asset prices and does not depend on the variance level. Furthermore, the integral representation of the transition density function obtained in this chapter can also be used to price other European-type options written on the asset yield ratio and the variance level.

Having obtained a simpler pricing IPDE as a result of the put-call transformation in Chapter 4, we proposed in Chapter 5 a numerical solution for the IPDE based on the method of lines (MOL). The MOL is computationally efficient in that it is able to approximate the exchange option price, the option delta, and the option gamma in a single implementation, is able to accommodate free boundary problems with minimal adjustments from the standard case, and is able to do so in less time than other numerical methods for American option pricing. In particular, we compared the performance of the MOL to a projected successive overrelaxation (PSOR) solution of the discretized IPDE and the least squares Monte Carlo (LSMC) algorithm, and found that the MOL produced prices which were closest to the reference prices (obtained using a PSOR algorithm on a significantly finer computational grid) within less computational time. In this chapter, we also provided detailed pseudo-algorithms for the implementation of the MOL and the PSOR methods. We also explored the use of alternative, Venttsel boundary conditions at the far variance boundary and how it can improve computational speed and accuracy. We noted however that the use of Venttsel boundary conditions is permissible if the Feller condition for the variance process is satisfied. Otherwise, one can regularize the coefficients of the second-order partial derivatives if the Feller condition is violated. We then conducted a series of numerical experiments highlighting the effect of specific model parameters on the exchange option prices and the early exercise boundary surface. We found

that jumps and stochastic volatility do significantly impact option prices, especially when near or at-the-money. This is again consistent with the issue of rebalancing costs with respect to sudden jumps at early exercise, as discussed in the previous chapters.

Finally in Chapter 6, we provided an initial analysis of volatility indices and how these can be used as a proxy for the unobservable instantaneous variance processes under an affine SVJD model. Parameter estimation and calibration, especially for the SVJD model, are further complicated by the fact that the variances are not observable, but estimation methods such as the approximate maximum likelihood method of Bates (2006) and the Markov Chain Monte Carlo algorithms of Eraker et al. (2003) have been proposed to address this exact complication. The calibration problem is also complicated by the unavailability of data on multi-asset option prices, so methods have been proposed to consolidate market information obtained from a calibration with respect to single-asset vanilla options. With the introduction of volatility indices for commonly traded assets, such as the VIX for the S&P 500 index, it has been found that the VIX has an affine dependence on the instantaneous variance process when the underlying asset pricing model is affine, which allows one to use the VIX as a proxy for the instantaneous variance for purposes of jump activity analysis or even parameter estimation. In this chapter, we confirmed this affine dependence for a general affine version of our two-asset SVJD model where there are jumps in both asset prices and volatility, similar to the specifications of Eraker et al. (2003) and Bakshi et al. (2021). The results of this chapter informs future research in which we develop the implementation of methods for parameter estimation and calibration (see Appendix E) for the two-asset SVJD model.

7.2 Directions for Future Research

There are various interesting directions in which we intend to extend the results of this thesis:

- From the results of Caldana and Fusai (2013), the exact pricing of European exchange options is possible if the joint characteristic function of the asset price and stochastic volatility processes are available. Within the affine framework of Duffie et al. (2000), this is possible. However, an excursion into American exchange option pricing within the general affine framework may prove challenging because of the generalized structure of jump intensities, but it is nonetheless an interesting extension of the results established here. Pricing European and American exchange option prices can also be investigated for non-affine models, for example, general Itô-Lévy processes, polynomial processes (Cuchiero et al., 2012), and regime-switching models.

- Pricing other two-asset options, either European or American, can also be considered under the SVJD model considered in this thesis. Properties of the early exercise boundary for some multi-asset American options have been resolved when assets are modelled as geometric

Brownian motions or general diffusion processes (Broadie and Detemple, 1997; Villeneuve, 1999; Rapuch, 2005; Nyström, 2008), but as established in this thesis, the possibility of jumps add an additional factor of complication to the analysis (see also Ekström and Tysk, 2007b). As such, it would also be interesting to formally establish properties of the early exercise boundaries of early exercise options when asset prices are driven by both stochastic volatility and jump-diffusion dynamics.

In this direction of pricing other two-asset options, we can also extend the put-call transformation technique to price other options whose payoff function has the applicable homogeneity structure (see for example, Fajardo and Mordecki, 2006). To this end, we can also investigate what conditions need to be imposed on the asset price dynamics so that the two-asset option can be reduced to an option on the ratio of the two asset prices. In addition, it is also worthwhile to see if the put-call transformation can be extended to options on three or more assets as an iterative process.

Furthermore, the exchange option pricing framework can be used to investigate other problems in finance, such as the pricing of real option, options written on foreign exchange, energy market options, and other problems that arise in corporate finance.

- Financial markets where asset prices exhibit SVJD dynamics are generally incomplete, so there are many equivalent martingale measures which are consistent with the no-arbitrage condition. In this thesis, we assumed that a pricing measure has already been chosen among the class of martingale measures equivalent to the market measure $\mathbb{P}$. As such, future research in this regard involves the construction of equivalent martingale measures for our SVJD model that are consistent with additional criteria, such as the minimum entropy martingale measure (MEMM) and related equivalent martingale measures satisfying "distance"-based criteria. In Appendix A.4, we have provided initial steps towards the construction of the MEMM corresponding to our SVJD model specification, but additional work has to be done to solve the corresponding entropy minimization problem and to provide sufficient conditions to guarantee the existence of the MEMM. In a broader sense, the implementation of other pricing frameworks in incomplete markets, such as utility indifference pricing (Biagini et al., 2010; Delbaen et al., 2002; Henderson and Hobson, 2009; Musiela and Zariphopoulou, 2004), superhedging (Bellamy and Jeanblanc, 1999; Eberlein and Jacod, 1997; El Karoui and Quenez, 1995), and pricing via risk measures (Artzner et al., 1999; Klöppel and Schweizer, 2007; Xu, 2006) to our SVJD context merits further investigation.

- Our main focus in this thesis was the pricing of exchange options, so the problem of hedging such options is a natural next step. Markets in which volatility is stochastic and asset prices jump are generally incomplete, so approximate hedging techniques must be considered. In

particular, it would be interesting to extend the approximate hedging results under jump-diffusion of Mina et al. (2015) to the case with stochastic volatility. The hedging problem may also be addressed using techniques in Malliavin calculus, as was done by Alòs and Rheinlander (2017). Within our specific SVJD framework, we do have to resort to the relevant Malliavin calculus for Itô-Lévy processes (see for example, Benth et al., 2003; Løkka, 2004; Solé et al., 2007; Suzuki, 2013, 2017).

- As established in Chapter 5, the method of lines is a powerful numerical algorithm for solving pricing PDEs and IPDEs in that it can also easily accommodate early exercise options and generate prices and hedge ratios efficiently. However, it remains to be seen how efficiently the MOL algorithm performs for in the general affine case or in the non-affine case, since it is still susceptible to the curse of dimensionality. As such, we intend to implement the MOL for the four-dimensional pricing IPDE obtained in Chapter 2 in future work, following the high-dimensional implementations presented by Kang and Meyer (2014) and Taruvinga (2019). It is also worthwhile extending the MOL solution for multi-asset options under regime-switching asset price dynamics (see Chiarella et al. (2016) for the MOL for single-asset regime-switching models).

  In addition, we also intend to consider other numerical methods for the SVJD model aside from those considered in this thesis. In particular, future work can be done towards a numerical solution of the linked system of integral equations for the early exercise premium and the American exchange option, extending similar numerical work by Chiarella and Ziogas (2009) and Chiarella and Ziveyi (2013) to the multi-asset case, or formulating an efficient numerical integration scheme for the integral representations obtained in Chapters 3 and 4. It would also be interesting to see how deep learning-based methods for solving high-dimensional PDEs can be extended to include integral terms as an alternative to the aforementioned numerical methods.

- Our work assumes that model parameters have already been estimated, but the parameter estimation and model calibration is anything but a trivial task. While parameter estimation for affine SVJD models is largely resolved in literature, the model calibration problem remains challenging due to the lack of relevant market data for multi-asset options. Nonetheless, the methods for parameter estimation and model calibration presented in Chapter 6 for the SVJD model, including the deep learning-based methods alluded to in Chapter 1, may serve as useful starting points for such an analysis.

APPENDIX A

Preliminary Concepts in Stochastic Calculus

This appendix contains relevant results in stochastic calculus for jump-diffusion processes and Lévy processes that are used throughout this thesis. We also discuss here some properties of stochastic exponentials, summarizing the key results obtained by Wong and Heyde (2004). The concepts and proofs introduced here are mainly based on the discussion by Øksendal and Sulem (2019). This appendix also discusses the construction of the minimum entropy martingale measure for the SVJD model in Chapter 2 and some concepts related to the transition probabilities of the proportional SVJD model introduced in Chapter 3.

A.1 Itô's Lemma for Lévy Processes

In this appendix, we state a form of Itô's Lemma that is applicable to the calculation of stochastic differential equations of functions of our stock price and instantaneous variance process, which collectively can be expressed as a four-dimensional Lévy process. We start with the general notation of an n-dimensional Itô-Lévy process and state Itô's Lemma thereafter.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a filtration $\{\mathcal{F}_t\}$ satisfying the usual conditions (Karatzas and Shreve, 1988, Definition 2.25):

1. $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets
2. $\{\mathcal{F}_t\}$ is right-continuous: $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $0 \leq t < T$.

We also assume that $\mathcal{F}_T = \mathcal{F}$.

On this probability space, suppose $\mathbf{B}$ is a standard m-dimensional $\mathbb{P}$-Wiener process and $\tilde{p}_k(\mathrm{d}z, \mathrm{d}t)$, $k = 1, \ldots, K$ are independent $\mathbb{P}$-compensated Poisson random measures with Lévy measure $\nu_k(\mathrm{d}z)$. Furthermore, assume that $\mathbf{W}$ is a vector of correlated scalar Wiener processes with correlation matrix $\mathbf{\Sigma}$ for which a Cholesky decomposition exists, i.e. there is a lower triangular matrix $\mathbf{L}$ such that $\mathbf{L}\mathbf{L}^\top = \mathbf{\Sigma}$, and that $\mathbf{W} = \mathbf{L}\mathbf{B}$.

The SVJD model (2.1) and (2.2) belongs to the class of Itô-Lévy processes $X : [0, t] \times \Omega \to \mathbb{R}^n$, $X(t) = (X_1(t), \ldots, X_n(t))^\top$, which have the general form

$$\mathrm{d}X_i(t) = \alpha_i(t)\,\mathrm{d}t + \beta_i^\top(t)\,\mathrm{d}\mathbf{W}(t) + \sum_{k=1}^{K} \int_{\mathbb{R}_0} \gamma_{ik}(t, z)\tilde{p}_k(\mathrm{d}z, \mathrm{d}t), \qquad i = 1, \ldots, n. \tag{A.1}$$

The coefficients $\alpha_i \in \mathbb{R}$, $\beta_i \in \mathbb{R}^m$, and $\gamma_{ik} \in \mathbb{R}$ are compiled into vectors/matrices as $\alpha(t) = (\alpha_1(t), \ldots, \alpha_n(t))^\top$, $\beta(t) = [\beta_{ij}(t)] \in \mathbb{R}^{n\times m}$, and $\gamma(t) = [\gamma_{ik}(t)] \in \mathbb{R}^{n\times K}$ and are assumed to be predictable processes. We denote by $\gamma^{(k)}$ the kth column of γ and $(\beta\beta^\top)_{ij}(t)$ the ij-th element of

$\beta(t)\beta^\top(t)$. In terms of the vector Brownian motion $\mathbf{B}$, we can write $\mathrm{d}X_i(t)$ as

$$\mathrm{d}X_i(t) = \alpha_i(t)\,\mathrm{d}t + \sigma_i^\top(t)\,\mathrm{d}\mathbf{B}(t) + \sum_{k=1}^{K}\int_{\mathbb{R}_0}\gamma_{ik}(t,z)\tilde{p}_k(\mathrm{d}z,\mathrm{d}t), \qquad i=1,\ldots,n, \tag{A.2}$$

where $\sigma(t) = \beta(t)\mathbf{L}$.

We now state a version of Itô's Lemma for computing the stochastic differential of a process $Y(t) = f(t, X(t))$, where $f \in C^{1,2}([0,T]\times\mathbb{R}^n)$ based on Øksendal and Sulem (2019, Theorem 1.16). For a proof, the reader is referred to standard texts on stochastic calculus for Lévy processes (see for example, Cont and Tankov, 2004a, Section 8.3.3.).

Theorem A.1 (Itô's Lemma for Itô-Lévy processes)**.** *Suppose X is an n-dimensional Itô-Lévy process whose ith component X_i satisfies the stochastic differential equation (A.1). Let $Y(t) = f(t, X(t))$, where $f : [0,T]\times\mathbb{R}^n \to \mathbb{R}$ is a function in $C^{1,2}([0,T]\times\mathbb{R}^n)$. Then Y is also an Itô-Lévy process satisfying the stochastic differential equation*

$$\begin{aligned}\mathrm{d}Y(t) = &\left[\frac{\partial f}{\partial t} + \sum_{i=1}^{n}\alpha_i(t)\frac{\partial f}{\partial x_i} - \sum_{i=1}^{n}\sum_{k=1}^{K}\int_{\mathbb{R}_0}\gamma_{ik}(t,z)\frac{\partial f}{\partial x_i}\upsilon_k(\mathrm{d}z) + \frac{1}{2}\sum_{i,j=1}^{n}(\beta\boldsymbol{\Sigma}\beta^\top)_{ij}(t)\frac{\partial^2 f}{\partial x_i\partial x_j}\right]\mathrm{d}t \\ &+ \sum_{i=1}^{n}\sum_{j=1}^{m}\beta_{ij}(t)\frac{\partial f}{\partial x_i}\,\mathrm{d}W_j(t) + \sum_{k=1}^{K}\int_{\mathbb{R}_0}\Big[f(t, X(t-)+\gamma^{(k)}(t,z)) - f(t,X(t-))\Big]\,p_k(\mathrm{d}z,\mathrm{d}t)\end{aligned} \tag{A.3}$$

All partial derivatives and coefficients (if applicable) are evaluated at $(t, X(t-))$.

Itô's Lemma is often written in terms of the coefficients of equation (A.2), where the standard Brownian motion $\mathbf{B}$ was used (see for example, Øksendal and Sulem, 2019, Theorem 1.16),

$$\begin{aligned}\mathrm{d}Y(t) = &\left[\frac{\partial f}{\partial t} + \sum_{i=1}^{n}\alpha_i(t)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n}(\sigma\sigma^\top)_{ij}(t)\frac{\partial^2 f}{\partial x_i\partial x_j}\right]\mathrm{d}t + \sum_{i=1}^{n}\sum_{j=1}^{m}\sigma_{ij}(t)\frac{\partial f}{\partial x_i}\,\mathrm{d}B_j(t) \\ &+ \sum_{k=1}^{K}\int_{\mathbb{R}_0}\left[f(t,X(t-)+\gamma^{(k)}(t,z)) - f(t,X(t-)) - \sum_{i=1}^{n}\gamma_{ik}(t,z)\frac{\partial f}{\partial x_i}\right]\upsilon_k(\mathrm{d}z)\,\mathrm{d}t \\ &+ \sum_{k=1}^{K}\int_{\mathbb{R}_0}\Big[f(t,X(t-)+\gamma^{(k)}(t,z)) - f(t,X(t-))\Big]\,\tilde{p}_k(\mathrm{d}z,\mathrm{d}t).\end{aligned}$$

Equation (A.3) is obtained by using the definition of the compensated random measure $\tilde{p}_k(\mathrm{d}z,\mathrm{d}t) = p(\mathrm{d}z,\mathrm{d}t) - \upsilon_k(\mathrm{d}z)\,\mathrm{d}t$ and by noting that $\sigma\sigma^\top = \beta\mathbf{L}\mathbf{L}^\top\beta^\top = \beta\boldsymbol{\Sigma}\beta^\top$ and

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{m}\sigma_{ij}(t)\frac{\partial f}{\partial x_i}\,\mathrm{d}B_j(t) &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\sum_{j=1}^{m}\left(\sum_{l=1}^{m}\beta_{il}(t)L_{lj}\right)\mathrm{d}B_j(t) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\sum_{l=1}^{m}\beta_{il}(t)\sum_{j=1}^{m}L_{lj}\,\mathrm{d}B_j(t) \\ &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\sum_{l=1}^{m}\beta_{il}(t)\,\mathrm{d}W_l(t) = \sum_{i=1}^{n}\sum_{j=1}^{m}\beta_{ij}(t)\frac{\partial f}{\partial x_i}\,\mathrm{d}W_j(t).\end{aligned}$$

The discussion in the following appendix about the Girsanov Theorem relies on the quadratic covariation process. Given two scalar Lévy processes X and Y, the quadratic covariation process $[X,Y]$ is the unique semimartingale process satisfying

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(s-)\,\mathrm{d}Y(s) + \int_0^t Y(s-)\,\mathrm{d}X(s) + [X,Y](t)$$

(Cont and Tankov, 2004a, Definition 8.4). The quadratic covariation process can be calculated with a stochastic product rule, which is stated as follows.

Lemma A.2 (Stochastic Product Rule and Quadratic Covariation)**.** *Suppose X and Y are two scalar Itô-Lévy processes satisfying*

$$\mathrm{d}X(t) = \alpha^x(t)\,\mathrm{d}t + \sigma^x(t)\cdot\mathrm{d}\mathbf{B}(t) + \sum_{k=1}^{K}\int_{\mathbb{R}_0}\gamma_k^x(t,z)\tilde{p}_k(\mathrm{d}z,\mathrm{d}t),$$
$$\mathrm{d}Y(t) = \alpha^y(t)\,\mathrm{d}t + \sigma^y(t)\cdot\mathrm{d}\mathbf{B}(t) + \sum_{k=1}^{K}\int_{\mathbb{R}_0}\gamma_k^y(t,z)\tilde{p}_k(\mathrm{d}z,\mathrm{d}t),$$

where $\alpha^x,\alpha^y\in\mathbb{R}$, $\sigma^x,\sigma^y\in\mathbb{R}^m$ and $\gamma^x,\gamma^y\in\mathbb{R}^K$ are predictable processes.

Then the product $\{X(t)Y(t)\}$ satisfies the equation

$$\begin{aligned}\mathrm{d}(X(t)Y(t)) &= Y(t-)\,\mathrm{d}X(t) + X(t-)\,\mathrm{d}Y(t) + \sigma^x(t)\cdot\sigma^y(t)\,\mathrm{d}t\\ &\quad+\sum_{k=1}^{K}\int_{\mathbb{R}_0}\gamma_k^x(t,z)\gamma_k^y(t,z)\left(\tilde{p}_k(\mathrm{d}z,\mathrm{d}t)+\upsilon_k(\mathrm{d}z)\,\mathrm{d}t\right)\end{aligned}\tag{A.4}$$

and the quadratic covariation process is given by

$$[X,Y](t) = \int_0^t\sigma^x(s)\cdot\sigma^y(s)\,\mathrm{d}s + \sum_{k=1}^{K}\int_0^t\int_{\mathbb{R}_0}\gamma_k^x(s,z)\gamma_k^y(s,z)\left(\tilde{p}_k(\mathrm{d}z,\mathrm{d}s)+\upsilon_k(\mathrm{d}z)\,\mathrm{d}s\right).\tag{A.5}$$

The quadratic variation of X follows from taking $X=Y$.

Equation (A.4) can be derived by applying Itô's Lemma (Theorem A.1) on the product process $V(t)=X(t)Y(t)$. Equation (A.5) follows from the definition of the quadratic covariation $[X,Y]$.

A.2 Martingale Property of Exponential Processes

In this section, we review salient results on exponential processes. As seen in this chapter and those that follow, exponential processes appear in the explicit formulas for stock price dynamics and in the Radon-Nikodým derivative processes that will be used to facilitate changes of measure and oftentimes these exponential processes are required to satisfy some martingale property. Thus in this section, we review some crucial definitions related to exponential processes and some necessary and sufficient conditions to determine whether a given exponential process is indeed a true martingale. Our discussion here is mainly based on the paper by Wong and Heyde (2004).

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space as in the previous section (Appendix A.1). However, we assume that the process $\mathbf{X}$ is an explicit function of t and a standard d-dimensional Brownian motion on $\mathbb{P}$ (different from the previous section). We also further assume that $\mathbf{X}_t$ is a progressively measurable process, defined as follows (see also Karatzas and Shreve, 1988, Definition 1.11).

Definition A.3 (Progressively measurable process)**.** The stochastic process $\mathbf{X}_t=(X_{1,t},\ldots,X_{d,t})^\top$ is **progressively measurable** with respect to the filtration $\{\mathcal{F}_t\}$ if for each $t\geq 0$ and $A\in\mathcal{B}(\mathbb{R}^d)$ (the σ-algebra of Borel sets on $\mathbb{R}^d$), the set

$$\left\{(s,w):0\leq s\leq t,\omega\in\Omega,\mathbf{X}_s(\omega)\in A\right\}$$

belongs to the product σ-algebra $\mathcal{B}([0,t])\otimes\mathcal{F}_t$. Any progressively measurable process is measurable and adapted.

Assumption A.4. Assume that $\mathbf{X}_t$ may be written as a function of t and $\mathbf{B}_t$; i.e. $\mathbf{X}_t=F(\mathbf{B}_t,t)$ for some function F.[1]

We now define the stochastic exponential process generated by a progressively measurable process $\mathbf{X}$.

[1] For $\mathbf{X}_t$ to be a progressively measurable process, the function F must be a progressively measurable functional. See Karatzas and Shreve (1988, pp. 199-200) for a definition of progressively measurable functionals.

Definition A.5 (Stochastic exponential process)**.** Let $\{\mathbf{X}_t\}$ be a progressively measurable process and $\{\mathbf{B}_t\}$ be a Brownian motion (in the sense of the preceding definitions). The **stochastic exponential process generated by $\mathbf{B}_t$**, denoted by $\{Z_t^X\}$, is defined as

$$Z_t^X = \exp\left\{\int_0^t \mathbf{X}_u \cdot \mathrm{d}\mathbf{B}_u - \frac{1}{2}\int_0^t \|\mathbf{X}_u\|^2 \,\mathrm{d}u\right\}, \qquad Z_0^X = 0, \tag{A.6}$$

where $\mathbf{X}_u \cdot \mathrm{d}\mathbf{B}_u = \sum_{i=1}^d X_{i,u}\,\mathrm{d}B_{i,u}$ and $\|\mathbf{X}_u\| \equiv \sqrt{X_{1,u}^2 + \cdots + X_{d,u}^2}$.

It is known that if $\mathbf{X}_t$ is a local martingale, then $\{Z_t^X\}$ is a nonnegative local martingale.[2] Furthermore, $\{Z_t^X\}$ is a $\mathbb{P}$-martingale if and only if

$$\mathbb{E}_{\mathbb{P}}[Z_t^X] = 1, \qquad \forall t \in [0,T]$$

(Cohen and Elliott, 2015, Lemma 15.3.2). This condition is often difficult to check, so alternative checks for the martingale property are given by Novikov's condition and Kamazaki's criterion.[3] In this section, we review the condition proposed by Wong and Heyde (2004) that relates the explosion time of $\mathbf{X}_t$ and the martingale property of $\{Z_t^X\}$.

Given the process $\mathbf{X}_t$, we define the process $\{M_t\}$ by

$$M_t = \int_0^t \|\mathbf{X}_u\|^2 \,\mathrm{d}u \tag{A.7}$$

and the *explosion time* τ^M given by

$$\tau^M = \lim_{N\to\infty} \tau_N^M, \tag{A.8}$$

$$\text{where} \qquad \tau_N^M = \inf\left\{t \in (0,\infty) : M_t(X) \geq N\right\}, \tag{A.9}$$

with $\inf \emptyset = \infty$. Whenever $\tau^M \leq T$, we set

$$\mathbf{X}_t = \mathbf{X}_{\tau^M}, \qquad \forall t \in [\tau^M, T].$$

We note that $\int_0^t \mathbf{X}_u \cdot \mathrm{d}\mathbf{B}_u$ is well-defined for all $t \in [0, \tau^M)$. Notice also that because $\mathbf{X}_t$ depends on $\mathbf{B}_t$, the probability measure we consider will affect the behavior of the explosion time.

The main idea behind the Wong and Heyde (2004) martingale condition is to consider the explosion time of $\mathbf{X}_t$ under another probability measure $\mathbb{P}^*$ equivalent to $\mathbb{P}$ that is constructed by taking Z_T^X as the Radon-Nikodým derivative $\mathrm{d}\mathbb{P}^*/\,\mathrm{d}\mathbb{P}$. Note that this is allowed if $\{Z_t^X\}$ has been determined to be a strictly positive $\mathbb{P}$-martingale. To this end, we consider the notion of *candidate measures* and state (without proof) the Wong and Heyde (2004) martingale criteria based on these candidate measures.

Definition A.6. A candidate measure $\mathbb{P}^*$, corresponding to the process $\mathbf{X}_t = F(\mathbf{B}_t, t)$ where $\mathbf{B}_t$ is a $\mathbb{P}$-Brownian motion, is a measure under which the process

$$\mathbf{X}_t = F\left(\mathbf{B}_t^* + \int_0^t \mathbf{X}_u \,\mathrm{d}u, t\right)$$

is defined, possibly up to the explosion time τ^M, where $\mathbf{B}_t^*$ is a $\mathbb{P}^*$-Brownian motion.[4] If $\{Z_t^X\}$ is

[2] An $\{\mathcal{F}_t\}$-adapted process $\{X_t\}_{t\in[0,T]}$ is called a **local martingale** with respect to $\{\mathcal{F}_t\}$ if there exists an increasing sequence of stopping times $\{\tau_n\}_{n=1}^\infty$ such that $\lim_{n\to\infty}\tau_n = T$ and for each n, the stopped process $\{X_{t\wedge\tau_n}\}$ is a martingale with respect to $\{\mathcal{F}_t\}$ (Kuo, 2006, Definition 5.5.1). Here, we note that $t\wedge\tau_n \equiv \min\{t,\tau_n\}$.

[3] See Cohen and Elliott (2015, Section 15.4) and Protter (2004, Chapter III Section 8) for a discussion of these conditions.

[4] The change in the first argument of the functional form of $\mathbf{X}_t$ represents a Girsanov transformation on the Brownian motion $\mathbf{B}_t$ from the original measure $\mathbb{P}$ to another probability measure $\mathbb{P}^*$.

a $\mathbb{P}$-martingale, then $\mathbb{P}^*$ is defined as

$$\mathbb{P}^*(\mathbf{X} \in A) = \mathbb{E}_{\mathbb{P}}\left[Z_T^X \times \mathbf{1}(\mathbf{X} \in A)\right] \tag{A.10}$$

for all $A \in \mathcal{F} = \mathcal{F}_T$.

Theorem A.7. *The stochastic exponential $\{Z_t^X\}$ is a $\mathbb{P}$-martingale for $t \in [0,T]$ if and only if the corresponding process $\mathbf{X}_t = F\left(\mathbf{B}_t^* + \int_0^t \mathbf{X}_u \,\mathrm{d}u, t\right)$ satisfies $\mathbb{P}^*\left(\tau^M > T\right) = 1$.*

In summary, suppose $\mathbb{P}^*$ can be defined by taking Z_T^X as the Radon-Nikodým derivative with respect to $\mathbb{P}$. If it can be shown that the process $\{M_t\}$ explodes under the new probability measure $\mathbb{P}^*$, then $\mathbb{P}^*$ cannot be equivalent to $\mathbb{P}$. *By the theorem above, $\{Z_t^X\}$ is a $\mathbb{P}$-martingale if and only if $\{M_t\}$ does not explode under $\mathbb{P}^*$ in the finite time interval $[0,T]$.*

A similar result also holds when $\mathbf{X}$ itself satisfies a given stochastic differential equation

$$\mathrm{d}\mathbf{X}(t) = \alpha(t, \mathbf{X}(t))\,\mathrm{d}t + \beta(t, \mathbf{X}(t))\,\mathrm{d}\mathbf{B}(t), \tag{A.11}$$

where $\mathbf{X}(0) = \mathbf{x}_0 \in \mathbb{R}^n$, $\alpha(t,x) \in \mathbb{R}^n$, $\beta(t,x) \in \mathbb{R}^{n\times n}$, $\mathbf{B}$ is an n-dimensional standard Brownian motion in $\mathbb{P}$, and α and β are progressively measurable functionals. In this case, the explosion time of $\mathbf{X}$, denoted by τ^X, is defined as

$$\tau^X = \lim_{N\to\infty} \tau_N^X, \qquad \text{where } \tau_N^X = \inf\left\{t \in (0,\infty) : \sup_{i=1,\ldots,n} |X_i(t)| \geq N\right\}.$$

Theorem A.8. *Suppose $\mathbf{X}$ is the unique weak solution of (A.11) and Z^X is the exponential process generated by $\mathbf{X}$ as in Definition A.5. Then*

$$\mathbb{E}_{\mathbb{P}}[Z^X(T)] = \mathbb{P}^*[\tau^X > T],$$

where

$$\mathrm{d}\mathbf{X}(t) = (\alpha(t, \mathbf{X}(t)) + \beta(t, \mathbf{X}(t))\mathbf{X}(t)) + \beta(t, \mathbf{X}(t))\,\mathrm{d}\mathbf{B}^*(t),$$

with $\mathbf{B}^*$ a standard n-dimensional Brownian motion under some candidate measure $\mathbb{P}^*$.

The same insight following Theorem A.7 also applies to Theorem A.8. That is, $\{Z^X\}_{t\in[0,T]}$ is a $\mathbb{P}$-martingale if the process $\mathbf{X}$ that generates Z^X has no component that is explosive prior to time T under the candidate measure $\mathbb{P}^*$, which can be defined as equivalent to $\mathbb{P}$ by taking $\mathrm{d}\mathbb{P}^* = Z^X(T)\,\mathrm{d}\mathbb{P}$. The main point of contention becomes testing $\mathbf{X}$ for explosive behavior over $[0,T]$. For specific types of processes, like the Cox et al. (1985) diffusion, we can refer to the Feller condition. For more general tests of explosion for one-dimensional diffusion processes, see Karatzas and Ruf (2015)

A.3 Girsanov Theorem for Lévy Processes

In this appendix, we state and prove a multidimensional version of the Girsanov Theorem for Itô-Lévy processes based on Øksendal and Sulem (2019, Theorem 1.33) and show that Proposition 2.3 is a specialized version of the general Girsanov Theorem. The notation in this appendix follows that of Appendix A.1, where we suppose $\mathbf{B}$ is a standard m-dimensional $\mathbb{P}$-Wiener process and $\tilde{p}_k(\mathrm{d}z, \mathrm{d}t)$, $k = 1,\ldots,K$ be independent $\mathbb{P}$-compensated Poisson random measures with Lévy measure $\upsilon_k(\mathrm{d}z)$.

Theorem A.9 (Girsanov Theorem for Lévy Processes)**.** *Let $g : [0,T]\times\Omega \to \mathbb{R}^m$ and $h : [0,T]\times\mathbb{R}_0\times\Omega \to \mathbb{R}^K$ (where $\mathbb{R}_0 = \mathbb{R}\backslash\{0\}$) be predictable processes, with $h_k(t,z) < 1$ for almost all*

$(\omega, t, z) \in \Omega \times [0,T] \times \mathbb{R}_0$ *for* $k = 1, \ldots, K$, *such that the process* $L = \{L(t)\}$ *given by*

$$\begin{aligned} L(t) = \exp\Bigg\{ &- \int_0^t g^\top(s)\,\mathrm{d}\mathbf{B}(s) - \frac{1}{2}\int_0^t \|g(s)\|^2\,\mathrm{d}s \\ &+ \sum_{k=1}^K \int_0^t \int_{\mathbb{R}_0} \ln(1 - h_k(s,z))\tilde{p}_k(\mathrm{d}z, \mathrm{d}s) \\ &+ \sum_{k=1}^K \int_0^t \int_{\mathbb{R}_0} [\ln(1 - h_k(s,z)) + h_k(s,z)]\upsilon_k(\mathrm{d}z)\,\mathrm{d}t \Bigg\}. \end{aligned} \tag{A.12}$$

exists. Suppose further that $\mathbb{E}_{\mathbb{P}}[L(T)] = 1$.

Let $\mathbb{Q}$ *be a probability measure on* $\mathcal{F}_T$ *such that* $\mathrm{d}\mathbb{Q} = L(T)\,\mathrm{d}\mathbb{P}$. *Define the process* $\mathbf{B}^{\mathbb{Q}} = \{\mathbf{B}^{\mathbb{Q}}(t)\}$ *and the compensated random measures* $\{\tilde{p}_k^{\mathbb{Q}}(\mathrm{d}z, \mathrm{d}t)\}$ *by*

$$\mathrm{d}\mathbf{B}^{\mathbb{Q}}(t) = g(t)\,\mathrm{d}t + \mathrm{d}\mathbf{B}(t), \tag{A.13}$$

$$\tilde{p}_k^{\mathbb{Q}}(\mathrm{d}z, \mathrm{d}t) = h_k(t,z)\upsilon_k(\mathrm{d}z)\,\mathrm{d}t + \tilde{p}_k(\mathrm{d}z, \mathrm{d}t), \qquad k = 1, \ldots, K. \tag{A.14}$$

Then $\mathbf{B}^{\mathbb{Q}}$ *is an* m*-dimensional standard* $\mathbb{Q}$*-Wiener process and* $\tilde{p}_k^{\mathbb{Q}}(\mathrm{d}z, \mathrm{d}t)$, $k = 1, \ldots, K$, *are* $\mathbb{Q}$*-compensated Poisson random measures (corresponding to* $p_k(\mathrm{d}z, \mathrm{d}t)$*) in the sense that the process*

$$\xi(t) = \sum_{k=1}^K \int_0^t \int_{\mathbb{R}_0} \zeta_k(s,z)\tilde{p}_k^{\mathbb{Q}}(\mathrm{d}z, \mathrm{d}s), \qquad 0 \le t \le T \tag{A.15}$$

is a $\mathbb{Q}$*-local martingale for all predictable processes* $\{\zeta_k(t,z)\}$ *such that*

$$\sum_{k=1}^K \int_0^T \int_{\mathbb{R}_0} \zeta_k^2(t,z)h_k^2(t,z)\upsilon_k(\mathrm{d}z)\,\mathrm{d}t < \infty, \qquad \textit{almost surely.} \tag{A.16}$$

We shall establish that $W^{\mathbb{Q}}$ is a Brownian motion under $\mathbb{Q}$ using the Lévy characterization of Brownian motion/Wiener processes, which is stated (without proof) below (statement adapted from Øksendal, 2013, Theorem 8.6.1).

Lemma A.10. *Let* $B(t) = (B_1(t), \ldots, B_m(t))$ *be a continuous* $\mathbb{R}^m$*-valued stochastic process on a probability space* $(\Omega, \mathcal{G}, \mu)$. *Then (1) and (2) below are equivalent:*

1. *B is a Brownian motion with respect to Q, i.e. the law of B with respect to Q is the same as the law of an m-dimensional Brownian motion.*
2. *(a) Each component of B is a local martingale with respect to Q and its own filtration; and*
   (b) The quadratic covariation process satisfies $[B_i, B_j](t) = \delta_{ij}t$ Q-almost surely for all $1 \le i, j \le n$. Here, $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

The statement of Lévy's characterization of Brownian motion in Øksendal (2013) requires that B is a martingale with respect to Q. The result is still true even if the components of B are only *local* martingales (see Durrett, 1996, Section 3.4).

We now present the proof of Theorem A.9

Proof. (Theorem A.9) Consider an m-dimensional process $\boldsymbol{\xi}(t) = (\xi_1(t), \ldots, \xi_m(t))^\top$ whose components are of the form (A.15), i.e.

$$\xi_j(t) = \sum_{k=1}^K \int_0^t \int_{\mathbb{R}_0} \zeta_{jk}(s,z)\tilde{p}_k^{\mathbb{Q}}(\mathrm{d}z, \mathrm{d}t), \qquad 0 \le t \le T, \quad j = 1, \ldots, m$$

where $\zeta(t,z) = [\zeta_{jk}(t,z)] \in \mathbb{R}^{m \times K}$ is a predictable process with whose rows satisfy condition (A.16).

Let $\epsilon \in [0,1]$. Define the m-dimensional process X^ϵ such that its jth component is given by

$$X_j^\epsilon(t) = \epsilon B_j^{\mathbb{Q}}(t) + \xi_j(t).$$

It follows that

$$\begin{aligned}
\mathrm{d}X_j^\epsilon(t) &= \epsilon\,\mathrm{d}B_j^{\mathbb{Q}}(t) + \mathrm{d}\xi_j(t) \\
&= \epsilon g_j(t)\,\mathrm{d}t + \epsilon\,\mathrm{d}W_j(t) + \sum_{k=1}^K \int_{\mathbb{R}_0} \zeta_{jk}(t,z)h_k(t,z)\upsilon_k(\mathrm{d}z)\,\mathrm{d}t + \sum_{k=1}^K \int_{\mathbb{R}_0} \zeta_{jk}(t,z)\tilde{p}_k(\mathrm{d}z,\mathrm{d}t) \\
&= \alpha_j^\epsilon(t)\,\mathrm{d}t + \epsilon\,\mathrm{d}B_j(t) + \sum_{k=1}^K \int_{\mathbb{R}_0} \zeta_{jk}(t,z)\tilde{p}_k(\mathrm{d}t,\mathrm{d}z),
\end{aligned}$$

where $\alpha_j^\epsilon(t)$ is defined as

$$\alpha_j^\epsilon(t) = \epsilon g_j(t) + \sum_{k=1}^K \int_{\mathbb{R}_0} \zeta_{jk}(t,z)h_k(t,z)\upsilon_k(\mathrm{d}z).$$

The stochastic differential $\mathrm{d}X_j^\epsilon(t)$ takes on the form

$$\mathrm{d}X_j(t) = \alpha_j(t)\,\mathrm{d}t + \sigma_j(t)\cdot\mathrm{d}\mathbf{B}(t) + \sum_{k=1}^K \int_{\mathbb{R}_0} \gamma_{jk}(t,z)\tilde{p}_k(\mathrm{d}z,\mathrm{d}t),$$

where $\alpha_j = \alpha_j^\epsilon$, $\sigma_j = (0,\ldots,0,\epsilon,0,\ldots,0)^\top$ (where the ϵ appears in the jth position), $\gamma_{jk}(t,z) = \zeta_{jk}(t,z)$, and α_j^ϵ satisfies

$$\alpha_i(t) = \sigma_i(t)\cdot g(t) + \sum_{k=1}^K \int_{\mathbb{R}_0} h_k(t,z)\gamma_{ik}(t,z)\upsilon_k(\mathrm{d}z) \tag{A.17}$$

for almost all t and z. Therefore, Øksendal and Sulem (2019, Theorem 1.31) implies that X_j^ϵ is a $\mathbb{Q}$-local martingale for all $j = 1,\ldots,m$ and for any choice of $\epsilon \in [0,1]$.

Choosing $\epsilon = 0$ means that $X_j^0(t) = \xi_j(t)$ is a $\mathbb{Q}$-local martingale. This establishes that $\tilde{p}_k^{\mathbb{Q}}$ is the $\mathbb{Q}$-compensated Poisson random measure corresponding to p_k.

Furthermore, choosing $\epsilon = 1$ implies that $B_j^{\mathbb{Q}}(t) = X_j^1(t) - \xi_j(t)$ is also a $\mathbb{Q}$-local martingale. To show that $\mathbf{B}^{\mathbb{Q}}$ is a $\mathbb{Q}$-Brownian motion, consider two arbitrary components $B_i^{\mathbb{Q}}$ and $B_j^{\mathbb{Q}}$ of $\mathbf{B}^{\mathbb{Q}}$ and the representation

$$\begin{bmatrix} \mathrm{d}B_i^{\mathbb{Q}}(t) \\ \mathrm{d}B_j^{\mathbb{Q}}(t) \end{bmatrix} = \begin{bmatrix} g_i(t) \\ g_j(t) \end{bmatrix}\mathrm{d}t + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \mathrm{d}B_i(t) \\ \mathrm{d}B_j(t) \end{bmatrix}.$$

In view of equation (A.5), if $i \neq j$, then $[W_i^{\mathbb{Q}}, W_j^{\mathbb{Q}}](t) = 0$. But if $i = j$, then $[W_i^{\mathbb{Q}}, W_j^{\mathbb{Q}}](t) = t$. Therefore by Lemma A.10, $\mathbf{B}^{\mathbb{Q}}$ is a Brownian motion with respect to $\mathbb{Q}$. ■

We now show that Proposition 2.3 is consistent with the Girsanov Theorem A.9 by specifying the processes g and h that parameterize the Radon-Nikodým derivative in Theorem A.9. In Proposition 2.3, we take $m = 4$ and $K = 2$ as our SVJD model has four Wiener components and two jump components. To further facilitate the comparison, as stated below the statement of Proposition 2.3, we factor $L(t)$ in equation (2.8) as $L(t) = L^D(t) \times L^{J_1}(t) \times L^{J_2}(t)$, where

$$L^D(t) = \exp\left\{ -\int_0^t (\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s))^\top\,\mathrm{d}\mathbf{W}(s) - \frac{1}{2}\int_0^t \boldsymbol{\theta}^\top(s)\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}(s)\,\mathrm{d}s \right\},$$

$$L^{J_1}(t) = \exp\left\{ \sum_{n=1}^{N_1(t)} (\gamma_1 Y_{1,n} + \nu_1) - \lambda_1 t(e^{\nu_1}\mathbb{E}_{\mathbb{P}}[e^{\gamma_1 Y_1}] - 1)\right\},$$
$$L^{J_2}(t) = \exp\left\{ \sum_{n=1}^{N_2(t)} (\gamma_2 Y_{2,n} + \nu_2) - \lambda_2 t(e^{\nu_2}\mathbb{E}_{\mathbb{P}}[e^{\gamma_2 Y_2}] - 1)\right\}.$$

We first look at $L^D(t)$. As was seen in Remark 2.4, we can write $L^D(t)$ as

$$L^D(t) = \exp\left\{ -\int_0^t (\mathbf{L}^{-1}\boldsymbol{\theta}(s))^\top \, \mathrm{d}\mathbf{B}(s) - \frac{1}{2}\int_0^t \|\mathbf{L}^{-1}\boldsymbol{\theta}(s)\|^2 \, \mathrm{d}s\right\}.$$

Thus, g in Theorem A.9 is chosen to be $g(t) = \mathbf{L}^{-1}\boldsymbol{\theta}(t)$.

We then rewrite the product $L^{J_1}(t)L^{J_2}(t)$ so that it is consistent with the form of the jump components of the Radon-Nikodým derivative (A.12). The third term in the exponent of $L(t)$ in equation (A.12) is written in terms of the Lévy measure $\upsilon_k(\mathrm{d}y)$ of the counting measure $p_i(\mathrm{d}y, \mathrm{d}t)$, which is just $\lambda_i m_{\mathbb{P}}^{(i)}(\mathrm{d}y)$. Furthermore, the sum $\sum_{n=1}^{N_i(t)}(\gamma_i Y_{i,n} + \nu_i)$ can be also written as the integral $\int_0^t \int_{\mathbb{R}_0}(\gamma_i y + \nu_i)p_i(\mathrm{d}y, \mathrm{d}t)$ (Runggaldier, 2003). Keeping these notations in mind and using the definition of $\mathbb{E}_{\mathbb{P}}[e^{\gamma_i Y_i}]$, we can write

$$\begin{aligned}
L^{J_1}(t)L^{J_2}(t) &= \exp\left\{ \sum_{k=1}^{2}\left[\int_0^t \int_{\mathbb{R}_0} (\gamma_k y + \nu_k)p_k(\mathrm{d}y, \mathrm{d}s) - \lambda_k t\left(e^{\nu_k}\int_{\mathbb{R}_0} e^{\gamma_k y} m_{\mathbb{P}}^{(k)}(\mathrm{d}y) - 1\right)\right]\right\} \\
&= \exp\left\{ \sum_{k=1}^{2}\left[\int_0^t \int_{\mathbb{R}_0} (\gamma_k y + \nu_k)(\tilde{p}_k(\mathrm{d}y, \mathrm{d}s) + \lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\,\mathrm{d}t)\right.\right. \\
&\qquad \left.\left. - \int_0^t \lambda_k \int_{\mathbb{R}_0} [e^{\gamma_k y + \nu_k} - 1] m_{\mathbb{P}}^{(i)}(\mathrm{d}y)\,\mathrm{d}s\right]\right\} \\
&= \exp\left\{ \sum_{k=1}^{2} \int_0^t \int_{\mathbb{R}_0} (\gamma_k y + \nu_k)\tilde{p}_k(\mathrm{d}y, \mathrm{d}t)\right. \\
&\qquad \left. + \sum_{k=1}^{2} \int_0^t \int_{\mathbb{R}_0} [(\gamma_k y + \nu_k) + (1 - e^{\gamma_k y + \nu_k})]\lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\,\mathrm{d}s\right\}.
\end{aligned}$$

Thus, setting $h_k(t, y) = h_k(y) = 1 - e^{\gamma_k y + \nu_k}$ in equation (A.12) yields the jump component $L^{J_1}(t)L^{J_2}(t)$. We note that $h_k(y) < 1$ for all $y \in \mathbb{R}_0$ for any choice of constants γ_k and ν_k as required by Theorem A.9.

Assuming that $g(t) = \mathbf{L}^{-1}\boldsymbol{\theta}(t)$ is a predictable process and $g(t)$ and $h_k(y) = 1 - e^{\gamma_k y + \nu_k}$ are defined such that $\mathbb{E}_{\mathbb{P}}[L(T)] = 1$, then consequently by Theorem A.9, the vector process $\mathbf{B}^{\mathbb{Q}}$ given by

$$\mathrm{d}\mathbf{B}^{\mathbb{Q}}(t) = g(t)\,\mathrm{d}t + \mathrm{d}\mathbf{B}(t) = \mathbf{L}^{-1}\boldsymbol{\theta}(t)\,\mathrm{d}t + \mathrm{d}\mathbf{B}(t)$$

is a standard 4-dimensional Brownian motion under $\mathbb{Q}$. In addition

$$\begin{aligned}
\tilde{p}_k^{\mathbb{Q}}(\mathrm{d}y, \mathrm{d}t) &= h_k(y)\upsilon_k(\mathrm{d}y)\,\mathrm{d}t + \tilde{p}_k(\mathrm{d}y, \mathrm{d}t) \\
&= h_k(y)\lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\,\mathrm{d}t + p_k(\mathrm{d}y, \mathrm{d}t) - \lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\,\mathrm{d}t \\
&= p_k(\mathrm{d}y, \mathrm{d}t) + (h_k(y) - 1)\lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\,\mathrm{d}t \\
&= p_k(\mathrm{d}y, \mathrm{d}t) - e^{\gamma_k y + \nu_k}\lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\,\mathrm{d}t
\end{aligned}$$

is the $\mathbb{Q}$-compensated random measure corresponding to $p_k(\mathrm{d}y, \mathrm{d}t)$. By the definition of compensated Poisson random measures and from the second term above, the Lévy measure of $p_k(\mathrm{d}y, \mathrm{d}t)$

under $\mathbb{Q}$, denoted by $\upsilon_k^{\mathbb{Q}}(\mathrm{d}y)$, can be written as

$$\begin{aligned}\upsilon_k^{\mathbb{Q}}(\mathrm{d}y) &= e^{\gamma_k y+\nu_k}\lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y) = \lambda_k e^{\nu_k} e^{\gamma_k y} m_{\mathbb{P}}^{(k)}(\mathrm{d}y) \\ &= \left(\lambda_k e^{\nu_k}\int_{\mathbb{R}_0} e^{\gamma_k y} m_{\mathbb{P}}^{k}(\mathrm{d}y)\right)\left(\frac{e^{\gamma_k y}}{\int_{\mathbb{R}_0} e^{\gamma_k y} m_{\mathbb{P}}^{(k)}(\mathrm{d}y)} m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\right) \\ &= \left(\lambda_k e^{\nu_k}\mathbb{E}_{\mathbb{P}}[e^{\gamma_k Y_k}]\right)\left(\frac{e^{\gamma_k y}}{\mathbb{E}_{\mathbb{P}}[e^{\gamma_k Y_k}]} m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\right)\end{aligned}$$

We can thus take $\tilde{\lambda}_k = \lambda_k e^{\nu_k}\mathbb{E}_{\mathbb{P}}[e^{\gamma_k Y_k}]$ and $m_{\mathbb{Q}}^{(k)}(\mathrm{d}y) = \frac{e^{\gamma_k y}}{\mathbb{E}_{\mathbb{P}}[e^{\gamma_k Y_k}]} m_{\mathbb{P}}^{(k)}(\mathrm{d}y)$ as the new intensity rate and jump size density corresponding to $p_k(\mathrm{d}y, \mathrm{d}t)$ under $\mathbb{Q}$. Using $m_{\mathbb{Q}}^{(k)}(\mathrm{d}y)$ to compute the mgf $M_{\mathbb{Q},Y_k}(u)$ gives us equation (2.9).

A.4 The Minimum Entropy Martingale Measure

Let X be an n-dimensional Itô-Lévy process whose components satisfy the stochastic differential equation (A.1). Let $\mathbb{Q}$ be the probability measure on $(\Omega, \mathcal{F})$ equivalent to $\mathbb{P}$ as defined in Theorem A.9. Then $\mathbb{Q}$ is said to be an equivalent martingale measure for X if the processes g and h in Theorem A.9 satisfy the additional condition that

$$\alpha_i(t) = \sigma_i^{\top}(t)g(t) + \sum_{k=1}^{K}\int_{\mathbb{R}_0}\gamma_{ik}(t,z)h_k(t,z)\upsilon_k(\mathrm{d}z) \tag{A.18}$$

for almost all $(\omega, t) \in \Omega \times [0, T]$, for each $i = 1, \ldots, n$.

It is well-known that a financial market whose asset prices are driven by Itô-Lévy processes is incomplete (Cont and Tankov, 2004a; Miyahara, 2012), so by the Second Fundamental Theorem of Asset Pricing (Harrison and Pliska, 1981; Björk, 2009), the equivalent martingale measure is not unique. As such, one can invoke additional criteria in the selection of an equivalent martingale measure that will be used as the pricing measure.

One such criterion is the minimization of the relative entropy with respect to $\mathbb{P}$, which we define below.

Definition A.11 (Relative Entropy)**.** Let $\mathbb{P}$ and $\mathbb{Q}$ be probability measures on $(\Omega, \mathcal{F})$. The relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$, denoted by $\mathcal{D}(\mathbb{Q}|\mathbb{P})$, is given by

$$\mathcal{D}(\mathbb{Q}|\mathbb{P}) = \begin{cases}\mathbb{E}_{\mathbb{P}}\left[\dfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\ln\left(\dfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)\right], & \text{if } \mathbb{Q} \text{ is absolutely continuous w.r.t. } \mathbb{P} \\ +\infty, & \text{otherwise.}\end{cases} \tag{A.19}$$

In particular, if $\mathbb{Q} \sim \mathbb{P}$ ($\mathbb{Q}$ is equivalent to $\mathbb{P}$), then

$$\mathcal{D}(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\ln\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)\right] = \mathbb{E}_{\mathbb{Q}}\left[\ln\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)\right].$$

We may also write $\mathcal{D}(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{P}}[f(\frac{d\mathbb{Q}}{d\mathbb{P}})]$, where f is the strictly convex function $f(x) = x\ln x$. For any probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$, $\mathcal{D}(\mathbb{Q}|\mathbb{P}) \geq 0$ and $\mathcal{D}(\mathbb{Q}|\mathbb{P}) = 0$ if and only if $\mathbb{Q} = \mathbb{P}$. As shown by Frittelli (2000); Fujiwara and Miyahara (2003); Fujiwara (2006), among others, the MEMM arises from the maximization of the expected exponential utility of terminal wealth (see also Miyahara, 2012, Chapter 5). The construction of MEMMs and related entropy-based martingale measures for exponential Lévy processes are discussed in depth by Miyahara (2012).

Definition A.12 (Minimal Entropy Martingale Measures)**.** Let $\mathcal{M}$ denote the collection of martingale measures for X on $(\Omega, \mathcal{F})$ that are absolutely continuous with respect to $\mathbb{P}$. The measure

$\mathbb{Q}^\star \in \mathcal{M}$ is called the **minimum entropy martingale measure** (MEMM) if it satisfies

$$\mathcal{D}(\mathbb{Q}^\star|\mathbb{P}) = \min_{\mathbb{Q}\in\mathcal{M}} \mathcal{D}(\mathbb{Q}|\mathbb{P}).$$

In our analysis, we are interested in finding the minimum entropy martingale measure among the equivalent probability measures $\mathbb{Q}$ as defined as in Theorem A.9 by the Radon-Nikodým derivative process $L(t)$ in (A.12) with parameters g and h satisfying (A.18). That is, defining $\mathcal{M}_e^*$ by

$$\mathcal{M}_e^* = \left\{ \mathbb{Q} \sim \mathbb{P} : \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\bigg|_{\mathcal{F}_T} = L(T),\ L(t) \text{ is given by (A.12) with } g, h \text{ satisfying (A.18)} \right\}$$

we wish to find the MEMM in $\mathcal{M}_e^*$.

We now calculate $\mathcal{D}(\mathbb{Q}|\mathbb{P})$ for $\mathbb{Q} \in \mathcal{M}_e^*$. Immediately from (A.12) and (A.19), we obtain

$$\begin{aligned}\mathcal{D}(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\Bigg[&- \int_0^T g^\top(s)\,\mathrm{d}\mathbf{B}(s) - \frac{1}{2}\int_0^T \|g(s)\|^2\,\mathrm{d}s + \sum_{k=1}^K \int_0^T \int_{\mathbb{R}_0} \ln(1-h_k(s,z))\tilde{p}_k(\mathrm{d}z,\mathrm{d}s) \\ &+ \sum_{k=1}^K \int_0^T \int_{\mathbb{R}_0} [\ln(1-h_k(s,z)) + h_k(s,z)]\upsilon_k(\mathrm{d}z)\,\mathrm{d}t\Bigg].\end{aligned}$$

We can use equations (A.13) and (A.14) to write $\mathcal{D}(\mathbb{Q}|\mathbb{P})$ in terms of $\mathbf{B}^{\mathbb{Q}}$ and $\tilde{p}_k^{\mathbb{Q}}(\mathrm{d}z,\mathrm{d}t)$ as

$$\begin{aligned}\mathcal{D}(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\Bigg[&\frac{1}{2}\int_0^T \|g(s)\|^2\,\mathrm{d}s - \int_0^T g^\top(s)\,\mathrm{d}\mathbf{B}^{\mathbb{Q}}(s) + \sum_{k=1}^K \int_0^T \int_{\mathbb{R}_0} h_k(s,z)\upsilon_k(\mathrm{d}z)\,\mathrm{d}s \\ &+ \sum_{k=1}^K \int_0^T \int_{\mathbb{R}_0} \ln(1-h_k(s,z))\left(\tilde{p}_k^{\mathbb{Q}}(\mathrm{d}z,\mathrm{d}s) + (1-h_k(s,z))\upsilon_k(\mathrm{d}z)\,\mathrm{d}s\right)\Bigg].\end{aligned}$$

Simplifying, we obtain the following expression for the relative entropy for $\mathbb{Q} \in \mathcal{M}_e^*$,

$$\mathcal{D}(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{2}\int_0^T \|g(s)\|^2\,\mathrm{d}s + \sum_{k=1}^K \int_0^T \int_{\mathbb{R}_0} \left(\ln(1-h_k(s,z))(1-h_k(s,z)) + h_k(s,z)\right)\upsilon_k(\mathrm{d}z)\,\mathrm{d}s\right]. \tag{A.20}$$

The MEMM $\mathbb{Q}^\star \in \mathcal{M}_e^*$ is then characterized by choosing the $g^\star$ and $h_k^\star$ $(k = 1,\ldots,K)$ which solve the minimization problem

$$\min_{g,h} \int_0^T \left(\frac{1}{2}\|g(s)\|^2 + \sum_{k=1}^K \int_{\mathbb{R}_0} \left(\ln(1-h_k(s,z))(1-h_k(s,z)) + h_k(s,z)\right)\upsilon_k(\mathrm{d}z)\right)\mathrm{d}s \tag{A.21}$$

subject to the constraint (A.18).

The minimum is computed over all predictable processes g and h such that $h_k(t,z) < 1$ for almost all (ω,t,z). We note that the objective function in problem (A.21) is convex in g and h_k. The corresponding density of $\mathbb{Q}^\star$ with respect to $\mathbb{P}$ is $L^\star$ computed using equation (A.12), taking $g = g^*$ and $h_k = h_k^*$.

We now discuss the initial steps toward the construction of the minimum entropy martingale measure (MEMM) for the SVJD model satisfied by our asset prices, particularly the derivation of additional constraints on the parameters of $L(t)$, given by (2.8), that will result to an MEMM. However, the complete construction of the MEMM, including the solution of the relative entropy minimization problem, is beyond the scope of this thesis.

In Appendix A.3, we established that the Radon-Nikodým derivative process $L(t)$ given by (2.8) is a special case of (A.12), where we have $K = 2$, $g(t) = \mathbf{L}^{-1}\boldsymbol{\theta}(t)$, $h_k(s,y) = 1 - e^{\gamma_k y + \nu_k}$, and $\upsilon_k(\mathrm{d}y) = \lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)$ for $k = 1,2$. By equation (A.20), the relative entropy of $\mathbb{Q}$, defined by

$\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T} = L(T)$, with respect to $\mathbb{P}$ is given by

$$\begin{aligned}\mathcal{D}(\mathbb{Q}|\mathbb{P}) &= \mathbb{E}_{\mathbb{Q}}\Bigg[\frac{1}{2}\int_0^T \|\mathbf{L}^{-1}\boldsymbol{\theta}(s)\|^2\,\mathrm{d}s \\ &\qquad + \sum_{k=1}^{2}\int_0^T\int_{\mathbb{R}_0}\left[\ln(e^{\gamma_k y+\nu_k})e^{\gamma_k y+\nu_k} + (1-e^{\gamma_k y+\nu_k})\right]\lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\,\mathrm{d}s\Bigg] \\ &= \mathbb{E}_{\mathbb{Q}}\Bigg[\frac{1}{2}\int_0^T \|\mathbf{L}^{-1}\boldsymbol{\theta}(s)\|^2\,\mathrm{d}s + T\sum_{k=1}^{2}\int_{\mathbb{R}_0}\left[\ln(e^{\gamma_k y+\nu_k})e^{\gamma_k y+\nu_k} + (1-e^{\gamma_k y+\nu_k})\right]\lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\Bigg].\end{aligned}$$

To clearly formulate the relative entropy minimization problem, we explicitly write $\psi_1(t)$ and $\psi_2(t)$ (the first two components of $\boldsymbol{\theta}(t)$ in (2.17)) in terms of the parameters γ_1, γ_2, ν_1, and ν_2. From the expression for $\tilde{\lambda}_k$ in Proposition 2.3 and $m_{\mathbb{Q}}^{(k)}(\mathrm{d}y)$ in equation (2.10) (resulting from equation (2.9)), we can write

$$\tilde{\kappa}_k = \int_{\mathbb{R}}(e^y-1)m_{\mathbb{Q}}^{(k)}(\mathrm{d}y) = \frac{1}{\mathbb{E}_{\mathbb{P}}[e^{\gamma_k Y_k}]}\int_{\mathbb{R}} e^{(\gamma_i+1)y}m_{\mathbb{P}}^{(k)}(\mathrm{d}y) - 1,$$

and so

$$\begin{aligned}\tilde{\lambda}_k\tilde{\kappa}_k &= \lambda_k e^{\nu_k}\mathbb{E}_{\mathbb{P}}[e^{\gamma_k Y_k}]\left(\frac{1}{\mathbb{E}_{\mathbb{P}}[e^{\gamma_k Y_k}]}\int_{\mathbb{R}} e^{(\gamma_i+1)y}m_{\mathbb{P}}^{(k)}(\mathrm{d}y) - 1\right) \\ &= \lambda_k e^{\nu_k}\left[\int_{\mathbb{R}} e^{(\gamma_k+1)y}m_{\mathbb{P}}^{(k)}(\mathrm{d}y) - \int_{\mathbb{R}} e^{\gamma_k y}m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\right] \\ &= \lambda_k e^{\nu_k}\left(\mathbb{E}_{\mathbb{P}}[e^{(\gamma_k+1)Y_k}] - \mathbb{E}_{\mathbb{P}}[e^{\gamma_k Y_k}]\right),\end{aligned}$$

for $k = 1, 2$.

Therefore, we need to solve the minimization problem

$$\begin{aligned}\min_{\boldsymbol{\theta}(t),\gamma_1,\gamma_2,\nu_1,\nu_2,\Lambda_1,\Lambda_2}\Bigg[&\frac{1}{2}\int_0^T \|\mathbf{L}^{-1}\boldsymbol{\theta}(s)\|^2\,\mathrm{d}s \\ &+ T\sum_{k=1}^{2}\int_{\mathbb{R}_0}\left[\ln(e^{\gamma_k y+\nu_k})e^{\gamma_k y+\nu_k} + (1-e^{\gamma_k y+\nu_k})\right]\lambda_k m_{\mathbb{P}}^{(k)}(\mathrm{d}y)\Bigg]\end{aligned} \tag{A.22}$$

where the predictable process $\boldsymbol{\theta}(t) = (\psi_1(t), \psi_2(t), \zeta_1(t), \zeta_2(t))^\top$ and the constants $\gamma_1, \gamma_2, \nu_1, \nu_2, \Lambda_1, \Lambda_2$ must jointly satisfy the constraints

$$\begin{aligned}\psi_1(t)\sqrt{v_1(t)} &= \mu_1 + q_1 - r - \lambda_1\kappa_1 + \lambda_1 e^{\nu_1}\left(\mathbb{E}_{\mathbb{P}}[e^{(\gamma_1+1)Y_1}] - \mathbb{E}_{\mathbb{P}}[e^{\gamma_1 Y_1}]\right), \\ \psi_2(t)\sqrt{v_2(t)} &= \mu_2 + q_2 - r - \lambda_2\kappa_2 + \lambda_2 e^{\nu_2}\left(\mathbb{E}_{\mathbb{P}}[e^{(\gamma_2+1)Y_2}] - \mathbb{E}_{\mathbb{P}}[e^{\gamma_2 Y_2}]\right), \\ \sigma_1\zeta_1(t) &= \Lambda_1\sqrt{v_1(t)}, \\ \sigma_2\zeta_2(t) &= \Lambda_2\sqrt{v_2(t)}, \\ &\text{where } \Lambda_1, \Lambda_2 \geq 0.\end{aligned} \tag{A.23}$$

The first two constraints in the (A.23) are required to ensure the discounted asset yield processes are martingales (i.e. such that the resulting probability measure is an EMM), while the last three constraints (including the nonnegativity constraint on Λ_1 and Λ_2) follow from the parameterization of the market price of volatility risk as assumed by Heston (1993). The constraints arising from the market price of volatility risk are not required for the martingale property of the discounted asset yield processes, but they do factor into the minimization problem since ζ_1 and ζ_2 are part of

the objective function.

If $\boldsymbol{\theta}^\star$, $\gamma_1^\star$, $\gamma_2^\star$, $\nu_1^\star$, $\nu_2^\star$, $\Lambda_1^\star$, and $\Lambda_2^\star$ denote the solution of the minimization problem (A.22)-(A.23), then the probability measure $\mathbb{Q}^\star$, defined by $\frac{d\mathbb{Q}^\star}{d\mathbb{P}} = L^\star(T)$, where $L^\star(t)$ is the process in equation (2.8) obtained by using the optimal parameters, is the MEMM equivalent to $\mathbb{P}$.

Establishing the existence of a minimum entropy martingale measure for our SVJD model lies outside the scope of this thesis. Theorem 2.1 of Frittelli (2000) implies that if the relative entropy is finite for all martingale measures that are absolutely continuous with respect to $\mathbb{P}$ and if the discounted asset yield processes are bounded, then there exists an MEMM. Fujiwara and Miyahara (2003) provide, for geometric Lévy processes, sufficient conditions for the existence of the MEMM. However, these conditions do not seem to apply directly to our SVJD context. As such, further work has to be done to provide conditions for the existence of an MEMM for an SVJD model.

A.5 Transition Probabilities for the Proportional SVJD Model

This appendix contains basic results regarding the transition probabilities of the proportional SVJD model $(S_1, S_2, v)^\top$ introduced in Chapter 3 and their connections to the price of European and American exchange options.

The representation for C^E in (3.38) suggests that the function G defined in (3.39) is closely related to the transition density function of the $\{\mathcal{F}_t\}$-Markov process $(S_1, S_2, v)^\top$ whose dynamics under $\mathbb{Q}$ are given by SDEs (3.8) and (3.9). To establish this connection, we recall some preliminary concepts on transition probabilities for Markov processes.

Recall that the $\mathbb{Q}$-*transition probability function* of $(S_1, S_2, v)^\top$ is the function $\mathcal{Q}$ given by

$$\mathcal{Q}(T, B; t, s_1, s_2, v) = \mathbb{Q}[(S_1(T), S_2(T), v(T)) \in B | S_1(t) = s_1, S_2(t) = s_2, v(t) = v],$$

which is defined for any $0 \le t \le T$, $(s_1, s_2, v) \in \mathbb{R}^3_+$, and $B \in \mathcal{B}(\mathbb{R}^3_+)$, the Borel σ-algebra on $\mathbb{R}^3_+$. This definition implies that $\mathcal{Q}(T, \cdot; t, s_1, s_2, v)$ is a probability on $\mathcal{B}(\mathbb{R}^3_+)$ and $\mathcal{Q}(t, B; t, s_1, s_2, v) = \mathbf{1}_B(s_1, s_2, v)$, which is equal to 1 if $(s_1, s_2, v) \in B$ and 0 otherwise (see for example Cohen and Elliott (2015, Definition 17.2.2) or Yan (2018, Section 4.1.3.1)). Furthermore, the transition probability function allows us to define the expectation $\mathbb{E}_{\mathbb{Q}}^{(t,s_1,s_2,v)}[\cdot]$ with respect to the $\mathcal{Q}$ as

$$\mathbb{E}_{\mathbb{Q}}^{(t,s_1,s_2,v)}[h(S_1(T), S_2(T), v(T))] = \int_{\mathbb{R}^3_+} h(u_1, u_2, b)\mathcal{Q}(T, \mathrm{d}u_1, \mathrm{d}u_2, \mathrm{d}b; t, s_1, s_2, v) \tag{A.24}$$

for any bounded Borel-measurable function $h : \mathbb{R}^3_+ \to \mathbb{R}$ and any fixed $(s_1, s_2, v) \in \mathbb{R}^3_+$ and $0 \le t < T$. The expectation taken with the operator $\mathbb{E}_{\mathbb{Q}}^{(t,s_1,s_2,v)}[\cdot]$ is sometimes written as

$$\mathbb{E}_{\mathbb{Q}}^{(t,s_1,s_2,v)}[h(S_1(T), S_2(T), v(T)] = \mathbb{E}_{\mathbb{Q}}[h(S_1^{(t,s_1,v)}(T), S_2^{(t,s_2,v)}(T), v^{(t,v)}(T))], \tag{A.25}$$

where $S_i^{(t,s_i,v)}(T)$ is the (unique) solution of (3.9) at time T with initial condition $S_i(t) = s_i$ and $v^{(t,v)}(T)$ is the (unique) solution of (3.8) at time T with initial condition $v(t) = v$. Given the Markov property of $(S_1, S_2, v)^\top$ (see Shreve (2004, Corollary 6.3.2) or Øksendal and Sulem (2019, Theorem 1.30)), we also have

$$\mathbb{E}_{\mathbb{Q}}[h(S_1(T), S_2(T), v(T))|\mathcal{F}_t] = \mathbb{E}_{\mathbb{Q}}^{(t,S_1(t),S_2(t),v(t))}[h(S_1(T), S_2(T), v(T))], \tag{A.26}$$

for any bounded Borel-measurable function h and any $0 \le t < T$ (see also Cohen and Elliott (2015, Theorem 17.2.3) or Øksendal (2013, Theorem 7.1.2) for a similar result).

Suppose there exists a nonnegative function $\wp(T, u_1, u_2, b; t, s_1, s_2, v)$ that is a $\mathcal{B}(\mathbb{R}^3_+) \otimes \mathcal{B}(\mathbb{R}^3_+)$-measurable function of $(u_1, u_2, b, s_1, s_2, v) \in \mathbb{R}^3_+ \times \mathbb{R}^3_+$ for all $0 \le t \le T$ such that

$$\mathcal{Q}(T, B; t, s_1, s_2, v) = \int_B \wp(T, u_1, u_2, b; t, s_1, s_2, v)\, \mathrm{d}u_1\, \mathrm{d}u_2\, \mathrm{d}b, \qquad B \in \mathcal{B}(\mathbb{R}^3_+).$$

Then $\wp(T, u_1, u_2, b; t, s_1, s_2, v)$ is called the *transition density function* of $(S_1, S_2, v)^\top$. In particular, $\wp$ is the Radon-Nikodým derivative or density of $\mathcal{Q}$ with respect to the Lebesgue measure on $\mathbb{R}^3_+$. In terms of $\wp$, we can write the expectation (A.24) as

$$\mathbb{E}_{\mathbb{Q}}^{(t,s_1,s_2,v)}[h(S_1(T), S_2(T), v(T))] = \int_{\mathbb{R}^3_+} h(u_1, u_2, b)\wp(T, u_1, u_2, b; t, s_1, s_2, v)\,\mathrm{d}u_1\,\mathrm{d}u_2\,\mathrm{d}b. \tag{A.27}$$

The following lemma derives the Kolmogorov backward equation satisfied by the transition density function $\wp$.

Lemma A.13. *For fixed $(u_1, u_2, b) \in \mathbb{R}^3_+$, the transition density function $\wp(T, u_1, u_2, b; t, s_1, s_2, v)$ of (S_1, S_2, v) satisfies the IPDE*

$$\frac{\partial \wp}{\partial t} + \mathcal{L}_{\mathbf{s},v}\wp = 0, \tag{A.28}$$

for $(t, s_1, s_2, v) \in [0, T] \times \mathbb{R}^3_+$, subject to the terminal condition

$$\wp(T, u_1, u_2, b; T, s_1, s_2, v) = \delta(s_1 - u_1)\delta(s_2 - u_2)\delta(v - b),$$

where $\delta(\cdot)$ is the Dirac-delta function.

Proof. Let $h : \mathbb{R}^3_+ \to \mathbb{R}$ be an arbitrary bounded Borel-measurable function and define the function $g : [0, T] \times \mathbb{R}^3_+ \to \mathbb{R}$ by

$$\begin{aligned} g(t, s_1, s_2, v) &= \mathbb{E}_{\mathbb{Q}}^{(t,s_1,s_2,v)}[h(S_1(T), S_2(T), v(T))] \\ &= \int_{\mathbb{R}^3_+} h(u_1, u_2, b)\wp(T, u_1, u_2, b; t, s_1, s_2, v)\,\mathrm{d}u_1\,\mathrm{d}u_2\,\mathrm{d}b, \end{aligned} \tag{A.29}$$

where the second equality follows from (A.27).

Consider the process $\mathfrak{G} = \{\mathfrak{G}(t)\}_{t\in[0,T]}$ defined by $\mathfrak{G}(t) = g(t, S_1(t), S_2(t), v(t))$. From (A.26), it follows that $\mathfrak{G}$ is a $\mathbb{Q}$-martingale. Indeed, if $s \le t$, then

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[\mathfrak{G}(t)|\mathcal{F}_s] &= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}[h(S_1(T), S_2(T), v(T))|\mathcal{F}_t]|\mathcal{F}_s\right] \\ &= \mathbb{E}_{\mathbb{Q}}[h(S_1(T), S_2(T), v(T))|\mathcal{F}_s] \\ &= \mathfrak{G}(s). \end{aligned}$$

Using Itô's formula, we obtain the following stochastic differential equation for $\mathfrak{G}$,

$$\begin{aligned} \mathrm{d}\mathfrak{G}(t) = &\left[\frac{\partial g}{\partial t} + \sum_{j=1}^{2}(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j)S_j(t)\frac{\partial g}{\partial s_j} + (\xi\eta - (\xi + \Lambda)v(t))\frac{\partial g}{\partial v}\right. \\ &+ \frac{1}{2}\sum_{j=1}^{2}\sigma_j^2 v(t)S_j^2(t)\frac{\partial^2 g}{\partial s_j^2} + \frac{1}{2}\omega^2 v(t)\frac{\partial^2 g}{\partial v^2} \\ &\left.+ \rho_w\sigma_1\sigma_2 v(t)S_1(t)S_2(t)\frac{\partial^2 g}{\partial s_1\partial s_2} + \sum_{j=1}^{2}\rho_{wz_j}\sigma_j\omega v(t)S_j(t)\frac{\partial^2 g}{\partial s_j\partial v}\right]\mathrm{d}t \\ &+ \sum_{j=1}^{2}\sigma_j\sqrt{v(t)}S_j(t)\frac{\partial g}{\partial s_j}\,\mathrm{d}\tilde{W}_j(t) + \omega\sqrt{v(t)}\frac{\partial g}{\partial v}\,\mathrm{d}\tilde{Z}(t) \\ &+ \int_{\mathbb{R}}\left[g(t, S_1(t)e^y, S_2(t), v(t)) - g(t, S_1(t), S_2(t), v(t))\right]p_1(\mathrm{d}y, \mathrm{d}t) \\ &+ \int_{\mathbb{R}}\left[g(t, S_1(t), S_2(t)e^y, v(t)) - g(t, S_1(t), S_2(t), v(t))\right]p_2(\mathrm{d}y, \mathrm{d}t) \\ = &\left[\frac{\partial g}{\partial t} + \mathcal{L}_{\mathbf{s},v}g\right]\mathrm{d}t + \sum_{j=1}^{2}\sigma_j\sqrt{v(t)}S_j(t)\frac{\partial g}{\partial s_j}\,\mathrm{d}\tilde{W}_j(t) + \omega\sqrt{v(t)}\frac{\partial g}{\partial v}\,\mathrm{d}\tilde{Z}(t) \end{aligned}$$

$$+\int_{\mathbb{R}}\left[g(t,S_1(t)e^y,S_2(t),v(t))-g(t,S_1(t),S_2(t),v(t))\right]q_1(\mathrm{d}y,\mathrm{d}t)$$
$$+\int_{\mathbb{R}}\left[g(t,S_1(t),S_2(t)e^y,v(t))-g(t,S_1(t),S_2(t),v(t))\right]q_2(\mathrm{d}y,\mathrm{d}t),$$

From the martingale property of $\mathfrak{G}$, we deduce that the function g satisfies the IPDE

$$\begin{aligned}&\frac{\partial g}{\partial t}+\mathcal{L}_{\mathbf{s},v}g=0,\\ &g(T,s_1,s_2,v)=h(s_1,s_2,v),\qquad \forall(s_1,s_2,v)\in\mathbb{R}^3_+.\end{aligned}\tag{A.30}$$

This is the Feynman-Kac representation result corresponding to the proportional SVJD model (3.8) and (3.9). The terminal condition of the IPDE above follows from (A.29) and (A.25) since

$$\begin{aligned}g(T,s_1,s_2,v)&=\mathbb{E}_{\mathbb{Q}}^{(T,s_1,s_2,v)}[h(S_1(T),S_2(T),v(T)]\\ &=\mathbb{E}_{\mathbb{Q}}[h(S_1^{(T,s_1,v)}(T),S_2^{(T,s_2,v)}(T),v^{(T,v)}(T))]\\ &=\mathbb{E}_{\mathbb{Q}}[h(s_1,s_2,v)]\\ &=h(s_1,s_2,v).\end{aligned}$$

Differentiating g in (A.29) and using (A.30), we obtain

$$0=\int_{\mathbb{R}^3_+}h(u_1,u_2,b)\left[\frac{\partial\wp}{\partial t}+\mathcal{L}_{\mathbf{s},v}\wp\right]\mathrm{d}u_1\,\mathrm{d}u_2\,\mathrm{d}b.$$

Since this equation holds for any bounded, Borel-measurable function h, $\wp$ must satisfy (A.28). Furthermore, from (A.29) and the terminal condition in (A.30), we know that

$$h(s_1,s_2,v)=\int_{\mathbb{R}^3_+}h(u_1,u_2,b)\wp(T,u_1,u_2,b;T,s_1,s_2,v)\,\mathrm{d}u_1\,\mathrm{d}u_2\,\mathrm{d}b.$$

For this equation to hold, we impose the terminal condition

$$\wp(T,u_1,u_2,b;T,s_1,s_2,v)=\delta(u_1-s_1)\delta(u_2-s_2)\delta(b-v),$$

on IPDE (A.28). ■

APPENDIX B

Proofs of Selected Results in Chapter 3

B.1 Proof of Proposition 3.7

The proof of the proposition involves several steps, which are outlined as follows. First, we solve the PDE in terms of the unknown function $f(\tau)$ using the method of characteristics. Next, condition (3.28) will used to determine the function $f(\tau)$. Lastly, the resulting expression will be simplified to obtain the form stated in the proposition.

The characteristic equations: Suppose the solution surface $(\tau, \vartheta, \bar{U})$ is parameterized by s. The characteristic equations corresponding to equation (3.27) are given by

$$\begin{aligned}
\frac{d\tau}{ds} &= 1, \\
\frac{d\vartheta}{ds} &= \frac{1}{2}\omega^2\vartheta^2 - \Theta\vartheta + \frac{1}{2}\varepsilon, \\
\frac{d\bar{U}}{ds} &= (\alpha\vartheta + \Theta + \Psi)\,\bar{U} + f(\tau),
\end{aligned}$$

where, for brevity, we introduce the notation $\alpha = \xi\eta - \omega^2$. The above system of ODEs may also be written compactly as

$$\mathrm{d}\tau = \frac{\mathrm{d}\vartheta}{\frac{1}{2}\omega^2\vartheta^2 - \Theta\vartheta + \frac{1}{2}\varepsilon} = \frac{\mathrm{d}\bar{U}}{(\alpha\vartheta + \Theta + \Psi)\,\bar{U} + f(\tau)}. \tag{B.1}$$

The first characteristic pair: Integrating both sides of the first characteristic pair, we obtain

$$\tau + c_1 = \frac{2}{\omega^2}\int \frac{1}{\vartheta^2 - (\frac{2\Theta}{\omega^2})\vartheta + \frac{\varepsilon}{\omega^2}}\,\mathrm{d}\vartheta,$$

where c_1 is an arbitrary constant. Note that the quadratic equation $\vartheta^2 - (\frac{2\Theta}{\omega^2})\vartheta + \frac{\varepsilon}{\omega^2} = 0$ has roots

$$\vartheta = \frac{(2\Theta/\omega^2) \pm \sqrt{(4\Theta^2/\omega^4) - (\varepsilon/\omega^2)}}{2} = \frac{\Theta \pm \sqrt{\Theta^2 - \varepsilon\omega^2}}{\omega^2} = \frac{\Theta \pm \delta}{\omega^2},$$

where $\delta = \sqrt{\Theta^2 - \varepsilon\omega^2}$. With this, the integrand may be written as

$$\frac{1}{\vartheta^2 - (\frac{2\Theta}{\omega^2})\vartheta + \frac{\varepsilon}{\omega^2}} = \frac{\omega^2}{2\delta}\left[\frac{1}{\vartheta - \left(\frac{\Theta+\delta}{\omega^2}\right)} - \frac{1}{\vartheta - \left(\frac{\Theta-\delta}{\omega^2}\right)}\right],$$

and so we find that

$$\delta\tau + c_2 = \int \left[\frac{1}{\vartheta - \left(\frac{\Theta+\delta}{\omega^2}\right)} - \frac{1}{\vartheta - \left(\frac{\Theta-\delta}{\omega^2}\right)} \right] \mathrm{d}\vartheta = \ln\left[\frac{\omega^2\vartheta - \Theta - \delta}{\omega^2\vartheta - \Theta + \delta}\right].$$

This implies that

$$c_3 e^{\delta\tau} = \frac{\omega^2\vartheta - \Theta - \delta}{\omega^2\vartheta - \Theta + \delta},$$

where c_3 is a constant. This constant may be written as a function of τ and ϑ as

$$c_3 = c_3(\tau, \vartheta) \equiv \frac{\left(\omega^2\vartheta - \Theta - \delta\right) e^{-\delta\tau}}{\omega^2\vartheta - \Theta + \delta}. \tag{B.2}$$

We can also find a relation between τ and ϑ in terms of c_3 by rearranging the above equation, giving us

$$\vartheta = \frac{\Theta - \delta}{\omega^2} - \frac{2\delta e^{-\delta\tau}}{\omega^2(c_3 - e^{-\delta\tau})}. \tag{B.3}$$

The second characteristic pair: By rearranging the second characteristic pair, we obtain the differential equation

$$\frac{d\bar{U}}{d\tau} + (-\alpha\vartheta - \Theta - \Psi)\bar{U} = f(\tau).$$

To solve this equation, we introduce an integrating factor $R(\tau)$ that satisfies the equation

$$\frac{dR}{d\tau} = (-\alpha\vartheta - \Theta - \Psi)R.$$

Using the expression for ϑ in equation (B.3) and solving for R, we have

$$\begin{aligned}\ln R &= \int \left(-\Theta - \Psi - \alpha \left[\frac{\Theta - \delta}{\omega^2} - \frac{2\delta e^{-\delta\tau}}{\omega^2(c_3 - e^{-\delta\tau})} \right] \right) \mathrm{d}\tau \\ &= \left[\frac{\alpha(\delta - \Theta)}{\omega^2} - \Theta - \Psi \right] \tau + \alpha \int \frac{2\delta e^{-\delta\tau}}{\omega^2(c_3 - e^{-\delta\tau})}\, \mathrm{d}\tau.\end{aligned}$$

If we let $u = c_3 - e^{-\delta\tau}$, then

$$\ln R = \left[\frac{\alpha(\delta - \Theta)}{\omega^2} - \Theta - \Psi \right] \tau + \frac{2\alpha}{\omega^2} \ln\left| c_3 - e^{-\delta\tau} \right|.$$

This implies that

$$R(\tau) = \exp\left\{ \left[\frac{\alpha(\delta - \Theta)}{\omega^2} - \Theta - \Psi \right] \tau \right\} \left| \frac{1}{c_3 - e^{-\delta\tau}} \right|^{-\frac{2\alpha}{\omega^2}}.$$

Using the method of variation of parameters (see Duffy, 2011), the solution $\bar{U}$ of the inhomogenous ODE is given by

$$R(\tau)\bar{U} = \int_0^\tau R(z) f(z)\, \mathrm{d}z + c_4,$$

where c_4 is a constant. Using the expression above for $R(\tau)$, we find that

$$\begin{aligned}\bar{U}(\tau, \phi_1, \phi_2, \vartheta) &= \exp\left\{ \left[\frac{\alpha(\Theta - \delta)}{\omega^2} + \Theta + \Psi \right] \tau \right\} \left| c_3 - e^{-\delta\tau} \right|^{-2\alpha/\omega^2} \\ &\times \left[c_4 + \int_0^\tau f(z) \exp\left\{ \left[\frac{\alpha(\delta - \Theta)}{\omega^2} - \Theta - \Psi \right] z \right\} \left| \frac{1}{c_3 - e^{-\delta z}} \right|^{-\frac{2\alpha}{\omega^2}} \mathrm{d}z \right].\end{aligned} \tag{B.4}$$

Determining c_3 and c_4: Assume that $c_4 = F(c_3)$ for some function F of c_3. Note that if $\tau = 0$, then equation (B.4) gives us

$$\bar{U}(0, \phi_1, \phi_2, \vartheta) = \left| c_3 - e^{-\delta\tau} \right|^{-2\alpha/\omega^2} F(c_3).$$

When $\tau = 0$, equation (B.3) yields for ϑ the expression

$$\vartheta = \frac{\Theta - \delta}{\omega^2} - \frac{2\delta}{\omega^2(c_3 - 1)}.$$

Solving for $F(c_3)$, we find that

$$c_4 = F(c_3) = \left| c_3 - e^{-\delta\tau} \right|^{2\alpha/\omega^2} \bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta - \delta}{\omega^2} - \frac{2\delta}{\omega^2(c_3 - 1)}\right). \tag{B.5}$$

The above equation, in conjunction with equation (B.2), allows us to write $\bar{U}$ free of c_3 and c_4. We first note that

$$\begin{aligned}
\left| \frac{c_3 - e^{-\delta\tau}}{c_3 - 1} \right| &= \left| \frac{\frac{\left(\omega^2\vartheta - \Theta - \delta\right) e^{-\delta\tau}}{\omega^2\vartheta - \Theta + \delta} - e^{-\delta\tau}}{\frac{\left(\omega^2\vartheta - \Theta - \delta\right) e^{-\delta\tau}}{\omega^2\vartheta - \Theta + \delta} - 1} \right| \\
&= \left| \frac{(\omega^2\vartheta - \Theta - \delta)e^{-\delta\tau} - e^{-\delta\tau}(\omega^2\vartheta - \Theta + \delta)}{(\omega^2\vartheta - \Theta - \delta)e^{-\delta\tau} - (\omega^2\vartheta - \Theta + \delta)} \right| \\
&= \left| \frac{-2\delta e^{-\delta\tau}}{(\omega^2\vartheta - \Theta + \delta)e^{-\delta\tau} - (\omega^2\vartheta - \Theta + \delta) - 2\delta e^{-\delta\tau}} \right| \\
&= \left| \frac{2\delta e^{-\delta\tau}}{(\omega^2\vartheta - \Theta + \delta)(1 - e^{-\delta\tau}) + 2\delta e^{-\delta\tau}} \right|.
\end{aligned}$$

Also, for $0 \leq z \leq \tau$, we have

$$\left| \frac{c_3 - e^{-\delta\tau}}{c_3 - e^{-\delta z}} \right| = \left| \frac{2\delta e^{-\delta\tau}}{(\omega^2\vartheta + \delta)(e^{-\delta z} - e^{-\delta\tau}) + 2\delta e^{-\delta\tau}} \right|.$$

Lastly, we note that

$$\begin{aligned}
\frac{2\delta}{\omega^2(c_3 - 1)} &= \frac{2\delta}{\omega^2}\left[\frac{\omega^2\vartheta - \Theta + \delta}{(\omega^2\vartheta - \Theta - \delta)e^{-\delta\tau} - (\omega^2\vartheta - \Theta + \delta)} \right] \\
&= \frac{2\delta}{\omega^2}\left[\frac{\omega^2\vartheta - \Theta + \delta}{(\omega^2\vartheta - \Theta + \delta)(e^{-\delta\tau} - 1) - 2\delta e^{-\delta\tau}} \right].
\end{aligned}$$

We note that all real arguments in the absolute values are positive, and so the absolute value symbols may be omitted. In light of the preceding calculations, equation (B.4) may now be written

as

$$\begin{aligned}
&\bar{U}(\tau, \phi_1, \phi_2, \vartheta) \\
&\quad = \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right]\tau\right\} \\
&\qquad \times \left(\frac{2\delta e^{-\delta\tau}}{(\omega^2\vartheta - \Theta + \delta)(1-e^{-\delta\tau}) + 2\delta e^{-\delta\tau}}\right)^{-\frac{2\alpha}{\omega^2}} \\
&\qquad \times \bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta-\delta}{\omega^2} - \frac{2\delta(\omega^2\vartheta - \Theta + \delta)}{\omega^2\left[(\omega^2\vartheta - \Theta + \delta)(e^{-\delta\tau}-1) - 2\delta e^{-\delta\tau}\right]}\right) \\
&\qquad + \int_0^\tau f(z)\exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right](\tau - z)\right\} \\
&\qquad \times \left(\frac{2\delta e^{-\delta\tau}}{(\omega^2\vartheta + \delta)(e^{-\delta z} - e^{-\delta\tau}) + 2\delta e^{-\delta\tau}}\right)^{-\frac{2\alpha}{\omega^2}} \mathrm{d}z
\end{aligned} \tag{B.6}$$

Determining $f(\tau)$: The function $f(\tau)$ must be determined to satisfy condition (3.28). Taking the limit of equation (B.6) as $\vartheta \to \infty$, we have

$$\begin{aligned}
0 &= \lim_{\vartheta\to\infty} \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right]\tau\right\} \\
&\quad \times \left(\frac{2\delta e^{-\delta\tau}}{(\omega^2\vartheta - \Theta + \delta)(1-e^{-\delta\tau}) + 2\delta e^{-\delta\tau}}\right)^{-\frac{2\alpha}{\omega^2}} \\
&\quad \times \bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta-\delta}{\omega^2} - \frac{2\delta(\omega^2\vartheta - \Theta + \delta)}{\omega^2\left[(\omega^2\vartheta - \Theta + \delta)(e^{-\delta\tau}-1) - 2\delta e^{-\delta\tau}\right]}\right) \\
&\quad + \lim_{\vartheta\to\infty} \int_0^\tau f(z)\exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right](\tau - z)\right\} \\
&\quad \times \left(\frac{2\delta e^{-\delta\tau}}{(\omega^2\vartheta + \delta)(e^{-\delta z} - e^{-\delta\tau}) + 2\delta e^{-\delta\tau}}\right)^{-\frac{2\alpha}{\omega^2}} \mathrm{d}z.
\end{aligned}$$

Rearranging the terms and simplifying, we have

$$\begin{aligned}
&- \lim_{\vartheta\to\infty} \bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta-\delta}{\omega^2} - \frac{2\delta(\omega^2\vartheta - \Theta + \delta)}{\omega^2\left[(\omega^2\vartheta - \Theta + \delta)(e^{-\delta\tau}-1) - 2\delta e^{-\delta\tau}\right]}\right) \\
&\quad = \lim_{\vartheta\to\infty} \int_0^\tau f(z)\left[\frac{(\omega^2\vartheta - \Theta + \delta)(1-e^{-\delta\tau}) + 2\delta e^{-\delta\tau}}{(\omega^2\vartheta - \Theta + \delta)(e^{-\delta z} - e^{-\delta\tau}) + 2\delta e^{-\delta\tau}}\right]^{-\frac{2\alpha}{\omega^2}} \\
&\qquad \times \exp\left\{-\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right]z\right\}\mathrm{d}z.
\end{aligned}$$

Note that Assumption 3.5 ensures the existence of $\hat{U}$ and $\partial\hat{U}/\partial\vartheta$, therefore ensuring that $\hat{U}$ is continuous in ϑ. This hence allows us to evaluate the limit of the left-hand side of the equation, giving us

$$\begin{aligned}
&- \hat{U}\left(0, \phi_1, \phi_2, \frac{\Theta-\delta}{\omega^2} - \frac{2\delta}{\omega^2(e^{-\delta\tau}-1)}\right) \\
&\quad = \int_0^\tau f(z)\exp\left\{-\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right]z\right\}\left(\frac{1-e^{-\delta t}}{e^{-\delta z} - e^{-\delta\tau}}\right)\mathrm{d}z.
\end{aligned} \tag{B.7}$$

To proceed, we introduce the substitutions

$$r = \frac{1}{1 - e^{-\delta z}} \quad \text{and} \quad s = \frac{1}{1 - e^{-\delta \tau}}. \tag{B.8}$$

Equation (B.7) can therefore be written in r and s as

$$\begin{aligned} &- \bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta - \delta}{\omega^2} + \frac{2\delta s}{\omega^2}\right) \\ &\quad = \int_s^\infty f(z) \exp\left\{ - \left[\frac{\alpha(\Theta - \delta)}{\omega^2} + \Theta + \Psi \right] z \right\} \left(\frac{r}{r - s} \right)^{-\frac{2\alpha}{\omega^2}} \left(\frac{1}{\delta(r-1)r} \right) \mathrm{d}r. \end{aligned}$$

Rearranging the terms, we have

$$\begin{aligned} &- \delta \bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta - \delta}{\omega^2} + \frac{2\delta s}{\omega^2}\right) \\ &\quad = \int_s^\infty f(z) \exp\left\{ - \left[\frac{\alpha(\Theta - \delta)}{\omega^2} + \Theta + \Psi \right] z \right\} \frac{r^{-2\alpha/\omega^2}}{r(r-1)} (r - s)^{2\alpha/\omega^2} \, \mathrm{d}r. \end{aligned}$$

Define the function $g(r)$ by

$$g(r) = f(z) \exp\left\{ - \left[\frac{\alpha(\Theta - \delta)}{\omega^2} + \Theta + \Psi \right] z \right\} \frac{r^{-2\alpha/\omega^2}}{r(r-1)}. \tag{B.9}$$

Then the last equation can be written as

$$- \delta \bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta - \delta}{\omega^2} + \frac{2\delta s}{\omega^2}\right) = \int_s^\infty g(r)(r - s)^{2\alpha/\omega^2} \, \mathrm{d}r. \tag{B.10}$$

This last equation will be used to determine $g(r)$ by applying the definition of the Laplace transform (equation (3.26)) to expand the left-hand side of the equation. After obtaining g, the function f follows from equation (B.9).

Using equation (3.26), we have

$$\bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta - \delta}{\omega^2} + \frac{2\delta s}{\omega^2}\right) = \int_0^\infty \hat{U}(0, \phi_1, \phi_2, b) \exp\left\{ - \left(\frac{\Theta - \delta}{\omega^2} + \frac{2\delta s}{\omega^2} \right) b \right\} \mathrm{d}b.$$

Introducing a gamma function $\Gamma(\frac{2\alpha}{\omega^2} + 1)$ and using equation (3.34), we find that

$$\begin{aligned} &\bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta - \delta}{\omega^2} + \frac{2\delta s}{\omega^2}\right) \\ &\quad = \frac{\Gamma(\frac{2\alpha}{\omega^2} + 1)}{\Gamma(\frac{2\alpha}{\omega^2} + 1)} \int_0^\infty \hat{U}(0, \phi_1, \phi_2, b) \exp\left\{ - \left(\frac{\Theta - \delta}{\omega^2} + \frac{2\delta s}{\omega^2} \right) b \right\} \mathrm{d}b \\ &\quad = \frac{1}{\Gamma(\frac{2\alpha}{\omega^2} + 1)} \int_0^\infty \int_0^\infty e^{-x} x^{2\alpha/\omega^2} \hat{U}(0, \phi_1, \phi_2, b) \\ &\qquad \times \exp\left\{ - \left(\frac{\Theta - \delta}{\omega^2} \right) b \right\} \exp\left\{ - \frac{2\delta s b}{\omega^2} \right\} \mathrm{d}x \, \mathrm{d}b. \end{aligned}$$

With the substitution $x = (2\delta b/\omega^2)y$, we obtain

$$\bar{U}\left(0, \phi_1, \phi_2, \frac{\Theta - \delta}{\omega^2} + \frac{2\delta s}{\omega^2}\right)$$

$$= \frac{1}{\Gamma(\frac{2\alpha}{\omega^2}+1)} \int_0^\infty \hat{U}(0,\phi_1,\phi_2,b) \exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\}\left(\frac{2\delta b}{\omega^2}\right)$$
$$\times \left[\int_0^\infty \exp\left\{-\frac{2\delta b y}{\omega^2}\right\}\left(\frac{2\delta b y}{\omega^2}\right)^{\frac{2\alpha}{\omega^2}} \exp\left\{-\frac{2\delta s b}{\omega^2}\right\} \mathrm{d}y\right] \mathrm{d}b$$
$$= \frac{1}{\Gamma(\frac{2\alpha}{\omega^2}+1)} \int_0^\infty \hat{U}(0,\phi_1,\phi_2,b) \exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\}\left(\frac{2\delta b}{\omega^2}\right)^{\frac{2\alpha}{\omega^2}+1}$$
$$\times \left[\int_0^\infty \exp\left\{-\left(\frac{2\delta b}{\omega^2}\right)(s+y)\right\} y^{2\alpha/\omega^2} \mathrm{d}y\right] \mathrm{d}b.$$

If we let $r = s + y$, then

$$\bar{U}\left(0,\phi_1,\phi_2,\frac{\Theta-\delta}{\omega^2}+\frac{2\delta s}{\omega^2}\right)$$
$$= \frac{1}{\Gamma(\frac{2\alpha}{\omega^2}+1)} \int_0^\infty \hat{U}(0,\phi_1,\phi_2,b) \exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\}\left(\frac{2\delta b}{\omega^2}\right)^{\frac{2\alpha}{\omega^2}+1}$$
$$\times \left[\int_s^\infty \exp\left\{-\frac{2\delta b}{\omega^2} r\right\} (r-s)^{2\alpha/\omega^2} \mathrm{d}r\right] \mathrm{d}b.$$

Rearranging the order of integration, we have

$$\bar{U}\left(0,\phi_1,\phi_2,\frac{\Theta-\delta}{\omega^2}+\frac{2\delta s}{\omega^2}\right)$$
$$= \int_s^\infty (r-s)^{2\alpha/\omega^2} \left[\int_0^\infty \frac{\hat{U}(0,\phi_1,\phi_2,b)}{\Gamma(\frac{2\alpha}{\omega^2}+1)}\left(\frac{2\delta b}{\omega^2}\right)^{\frac{2\alpha}{\omega^2}+1} \exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}+\frac{2\delta b}{\omega^2}\right)b\right\} \mathrm{d}b\right] \mathrm{d}r \tag{B.11}$$

Comparing equations (B.10) and (B.11), we see that $g(r)$ is given by

$$g(r) = -\frac{\delta}{\Gamma(\frac{2\alpha}{\omega^2}+1)} \int_0^\infty \hat{U}(0,\phi_1,\phi_2,b)\left(\frac{2\delta b}{\omega^2}\right)^{\frac{2\alpha}{\omega^2}+1} \exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}+\frac{2\delta b}{\omega^2}\right)b\right\} \mathrm{d}b. \tag{B.12}$$

The function $f(z)$ thus follows from equation (B.9), having obtained $g(r)$.

Simplifying $\bar{U}(\tau,\phi_1,\phi_2,\vartheta)$: Denote by J_1 and J_2 the first and second terms, respectively, of the right-hand side of equation (B.6). In terms of $s = (1-e^{-\delta\tau})^{-1}$, J_1 can be written as

$$J_1 = \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2}+\Theta+\Psi\right]\tau\right\}\left[\frac{2\delta(\frac{s-1}{s})}{(\omega^2\vartheta-\Theta+\delta)(\frac{1}{s})+2\delta(\frac{s-1}{s})}\right]^{-\frac{2\alpha}{\omega^2}}$$
$$\times \bar{U}\left(0,\phi_1,\phi_2,\frac{\Theta-\delta}{\omega^2}-\frac{2\delta(\omega^2\vartheta-\Theta+\delta)}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)(-\frac{1}{s})-2\delta(\frac{s-1}{s})\right]}\right)$$
$$= \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2}+\Theta+\Psi\right]\tau\right\}\left[\frac{2\delta(s-1)}{(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)}\right]^{-\frac{2\alpha}{\omega^2}}$$
$$\times \bar{U}\left(0,\phi_1,\phi_2,\frac{\Theta-\delta}{\omega^2}+\frac{2\delta(\omega^2\vartheta-\Theta+\delta)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right).$$

Likewise, we apply the substitutions in equation (B.8) to rewrite J_2 as

$$\begin{aligned}
J_2 &= \int_s^\infty f(z) \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right](\tau - z)\right\} \\
&\quad \times \left[\frac{2\delta(\frac{s-1}{s})}{(\omega^2\vartheta - \Theta + \delta)(\frac{r-s}{rs}) + 2\delta(\frac{s-1}{s})}\right]^{-\frac{2\alpha}{\omega^2}} \frac{1}{\delta(r-1)r}\,\mathrm{d}r \\
&= \int_s^\infty f(z) \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right](\tau - z)\right\} \frac{r^{-2\alpha/\omega^2}}{\delta(r-1)r} \\
&\quad \times \left[\frac{2\delta(s-1)r}{(\omega^2\vartheta - \Theta + \delta)(r-s) + 2\delta(s-1)r}\right]^{-\frac{2\alpha}{\omega^2}} \mathrm{d}r.
\end{aligned}$$

Using equations (B.9) and (B.12), we have

$$\begin{aligned}
J_2 &= \frac{1}{\delta} \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right]\tau\right\} \\
&\quad \times \int_s^\infty g(r) \left[\frac{2\delta(s-1)r}{(\omega^2\vartheta - \Theta + \delta)(r-s) + 2\delta(s-1)r}\right]^{-\frac{2\alpha}{\omega^2}} \mathrm{d}r \\
&= \frac{1}{\delta} \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right]\tau\right\} \int_s^\infty \frac{-\delta}{\Gamma(\frac{2\alpha}{\omega^2}+1)} \\
&\quad \times \left[\frac{2\delta(s-1)r}{(\omega^2\vartheta - \Theta + \delta)(r-s) + 2\delta(s-1)r}\right]^{-\frac{2\alpha}{\omega^2}} \\
&\quad \times \left[\int_0^\infty \hat{U}(0,\phi_1,\phi_2,b) \exp\left\{-\left(\frac{\Theta-\delta}{\omega^2} + \frac{2\delta r}{\omega^2}\right)b\right\}\left(\frac{2\delta b}{\omega^2}\right)^{\frac{2\alpha}{\omega^2}+1} \mathrm{d}b\right] \mathrm{d}r \\
&= -\frac{[2\delta(s-1)]^{-2\alpha/\omega^2}}{\Gamma(\frac{2\alpha}{\omega^2}+1)} \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2} + \Theta + \Psi\right]\tau\right\} \\
&\quad \times \int_0^\infty \hat{U}(0,\phi_1,\phi_2,b)\left(\frac{2\delta b}{\omega^2}\right)^{\frac{2\alpha}{\omega^2}+1} \exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\} G(b)\,\mathrm{d}b,
\end{aligned}$$

where $G(b)$ is given by

$$G(b) = \int_s^\infty \exp\left\{-\frac{2\delta b}{\omega^2}r\right\}\left[(\omega^2\vartheta - \Theta + \delta)(r-s) + 2\delta(s-1)r\right]^{2\alpha/\omega^2} \mathrm{d}r.$$

To simplify $G(b)$, we consider the substitution

$$u = (\omega^2\vartheta - \Theta + \delta)(r-s) + 2\delta(s-1)r.$$

The resulting integral is

$$\begin{aligned}
G(b) &= \int_{2\delta(s-1)s}^\infty \exp\left\{-\frac{2\delta b}{\omega^2} \times \frac{u + (\omega^2\vartheta - \Theta + \delta)s}{(\omega^2\vartheta - \Theta + \delta) + 2\delta(s-1)}\right\} u^{2\alpha/\omega^2} \\
&\quad \times \frac{1}{(\omega^2\vartheta - \Theta + \delta) + 2\delta(s-1)}\,\mathrm{d}u \\
&= \frac{1}{(\omega^2\vartheta - \Theta + \delta) + 2\delta(s-1)} \exp\left\{-\frac{2\delta b(\omega^2\vartheta - \Theta + \delta)s}{\omega^2\left[(\omega^2\vartheta - \Theta + \delta) + 2\delta(s-1)\right]}\right\}
\end{aligned}$$

$$\times \int_{2\delta(s-1)s}^{\infty} \exp\left\{-\frac{2\delta bu}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right\} u^{2\alpha/\omega^2}\,\mathrm{d}u.$$

If we let

$$x = \frac{2\delta bu}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}$$

and define $\beta(b)$ by

$$\beta(b) = \frac{4\delta^2 b(s-1)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]},$$

then $G(b)$ can be written as

$$\begin{aligned} G(b) = {} & \frac{1}{(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)} \exp\left\{-\frac{2\delta b(\omega^2\vartheta-\Theta+\delta)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right\} \\ & \times \frac{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}{2b\delta}\left(\frac{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}{2b\delta}\right)^{\frac{2\alpha}{\omega^2}} \\ & \times \int_{\beta(b)}^{\infty} e^{-x}x^{2\alpha/\omega^2}\,\mathrm{d}x \end{aligned}$$

Note that

$$\begin{aligned} \int_{\beta(b)}^{\infty} e^{-x}x^{2\alpha/\omega^2}\,\mathrm{d}x &= \int_0^{\infty} e^{-x}x^{2\alpha/\omega^2}\,\mathrm{d}x - \int_0^{\beta(b)} e^{-x}x^{2\alpha/\omega^2}\,\mathrm{d}x \\ &= \Gamma\left(\frac{2\alpha}{\omega^2}+1\right) - \int_0^{\beta(b)} e^{-x}x^{2\alpha/\omega^2}\,\mathrm{d}x. \end{aligned}$$

In terms of the incomplete gamma function (equation (3.33)), we may write

$$\begin{aligned} \int_{\beta(b)}^{\infty} e^{-x}x^{2\alpha/\omega^2}\,\mathrm{d}x &= \Gamma\left(\frac{2\alpha}{\omega^2}+1\right) - \Gamma\left(\frac{2\alpha}{\omega^2}+1\right)\Gamma\left(\frac{2\alpha}{\omega^2}+1;\beta(b)\right) \\ &= \Gamma\left(\frac{2\alpha}{\omega^2}+1\right)\left[1-\Gamma\left(\frac{2\alpha}{\omega^2}+1;\beta(b)\right)\right]. \end{aligned}$$

Therefore, $G(b)$ simplifies to

$$\begin{aligned} G(b) = {} & \frac{\omega^2}{2b\delta}\left(\frac{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}{2b\delta}\right)^{\frac{2\alpha}{\omega^2}} \\ & \times \exp\left\{-\frac{2b\delta(\omega^2\vartheta-\Theta+\delta)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right\} \\ & \times \Gamma\left(\frac{2\alpha}{\omega^2}+1\right)\left[1-\Gamma\left(\frac{2\alpha}{\omega^2}+1;\beta(b)\right)\right] \\ = {} & \left(\frac{\omega^2}{2b\delta}\right)^{\frac{2\alpha}{\omega^2}+1}\left(\frac{1}{(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)}\right)^{-\frac{2\alpha}{\omega^2}} \\ & \times \exp\left\{-\frac{2b\delta(\omega^2\vartheta-\Theta+\delta)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right\} \\ & \times \Gamma\left(\frac{2\alpha}{\omega^2}+1\right)\left[1-\Gamma\left(\frac{2\alpha}{\omega^2}+1;\beta(b)\right)\right]. \end{aligned}$$

Consequently, J_2 can now be expressed as

$$\begin{aligned}J_2 = \exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2}+\Theta+\Psi\right]\tau\right\}&\left[\frac{2\delta(s-1)}{(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)}\right]^{-\frac{2\alpha}{\omega^2}}\\ &\times\int_0^\infty \hat{U}(0,\phi_1,\phi_2,b)\exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\}\exp\left\{-\frac{2b\delta(\omega^2\vartheta-\Theta+\delta)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right\}\\ &\times\left[\Gamma\left(\frac{2\alpha}{\omega^2}+1;\beta(b)\right)-1\right]\mathrm{d}b.\end{aligned}$$

Since $\bar{U}=J_1+J_2$, it follows that

$$\begin{aligned}&\bar{U}(\tau,\phi_1,\phi_2,\vartheta)\\ &\quad=\exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2}+\Theta+\Psi\right]\tau\right\}\left[\frac{2\delta(s-1)}{(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)}\right]^{-\frac{2\alpha}{\omega^2}}\\ &\qquad\times\left\{\int_0^\infty \hat{U}(0,\phi_1,\phi_2,b)\exp\left\{-\left[\frac{\Theta-\delta}{\omega^2}+\frac{2\delta(\omega^2\vartheta-\Theta+\delta)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right]b\right\}\mathrm{d}b\right.\\ &\qquad\quad\left.+\int_0^\infty \hat{U}(0,\phi_1,\phi_2,b)\exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\}\right.\\ &\qquad\quad\left.\times\exp\left\{-\frac{2b\delta(\omega^2\vartheta-\Theta+\delta)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right\}\left[\Gamma\left(\frac{2\alpha}{\omega^2}+1;\beta(b)\right)-1\right]\mathrm{d}b\right\}\\ &\quad=\exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2}+\Theta+\Psi\right]\tau\right\}\left[\frac{2\delta(s-1)}{(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)}\right]^{-\frac{2\alpha}{\omega^2}}\\ &\qquad\times\int_0^\infty \hat{U}(0,\phi_1,\phi_2,b)\exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\}\\ &\qquad\times\exp\left\{-\frac{2b\delta(\omega^2\vartheta-\Theta+\delta)s}{\omega^2\left[(\omega^2\vartheta-\Theta+\delta)+2\delta(s-1)\right]}\right\}\Gamma\left(\frac{2\alpha}{\omega^2}+1;\beta(b)\right).\end{aligned}$$

From equation (3.25), we note that $\hat{U}(0,\phi_1,\phi_2,b)=\hat{u}(\phi_1,\phi_2,b)$. Reversing the substitution $s=1/(1-e^{-\delta\tau})$ produces the expression for $\bar{U}(\tau,\phi_1,\phi_2,\vartheta)$ and $\beta^*(b)$ as shown in the proposition.

B.2 Proof of Proposition 3.8

Similar to the proof of the previous proposition, this proof will be broken down into several steps. First, we consider some substitutions, which we denote by $A\equiv A(b)$ and $z\equiv z(\vartheta)$, that will facilitate the inversion of the Laplace transform. Next, since we introduced a substitution z involving the original Laplace transform parameter ϑ, we need to establish a connection between the inverse Laplace transform with respect to ϑ and the inverse Laplace transform with respect to z. The last step consists in reversing the substitutions to obtain the expression stated in the proposition.

Some simplifying substitutions: Define A and z by

$$\begin{aligned}A\equiv A(b)&=\frac{2\delta b}{\omega^2(1-e^{-\delta\tau})},\\ z\equiv z(\vartheta)&=\frac{(\omega^2\vartheta-\Theta+\delta)(e^{\delta\tau}-1)+2\delta}{2\delta}.\end{aligned}\tag{B.13}$$

Note that these substitutions are defined such that $\beta^*(b)=A/z$. Furthermore, let H be the

function defined as

$$H(\tau,\phi_1,\phi_2,b)=\exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2}+\Theta+\Psi\right]\tau\right\}\hat{u}(\phi_1,\phi_2,b)\exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\}. \tag{B.14}$$

With these substitutions, we may rewrite $\bar{U}(\tau,\phi_1,\phi_2,\vartheta)$ (as given in equation (3.31)) as

$$\begin{aligned}\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z)) &= \int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)z^{2\alpha/\omega^2}\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\\ &\quad\times\exp\left\{-\frac{A(e^{\delta\tau}-1)(\omega^2\vartheta-\Theta+\delta)}{2\delta z}\right\}\Gamma\left(\frac{2\alpha}{\omega^2}+1;\frac{A}{z}\right)\mathrm{d}A\\ &= \int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}z^{2\alpha/\omega^2}\\ &\quad\times\exp\left\{-\frac{A(e^{\delta\tau}-1)(\omega^2\vartheta-\Theta+\delta)}{2\delta z}\right\}\frac{1}{\Gamma(\frac{2\alpha}{\omega^2}+1)}\\ &\quad\times\int_0^{A/z}e^{-u}u^{2\alpha/\omega^2}\,\mathrm{d}u\,\mathrm{d}A.\end{aligned}$$

If we let $x=1-(z/A)u$, then we have

$$\begin{aligned}&\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))\\ &\quad= \int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{z^{2\alpha/\omega^2}}{\Gamma(\frac{2\alpha}{\omega^2}+1)}\\ &\qquad\times\exp\left\{-\frac{A(e^{\delta\tau}-1)(\omega^2\vartheta-\Theta+\delta)}{2\delta z}\right\}\int_0^1 e^{-A/z}e^{Ax/z}\left[\frac{A}{z}(1-x)\right]^{\frac{2\alpha}{\omega^2}}\frac{A}{z}\,\mathrm{d}x\,\mathrm{d}A\\ &\quad= \int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{A^{2\alpha/\omega^2+1}}{\Gamma(\frac{2\alpha}{\omega^2}+1)}\\ &\qquad\times\exp\left\{-\frac{A(e^{\delta\tau}-1)(\omega^2\vartheta-\Theta+\delta)}{2\delta z}-\frac{A}{z}\right\}\\ &\qquad\times\int_0^1 e^{Ax/z}(1-x)^{2\alpha/\omega^2}z^{-1}\,\mathrm{d}z\,\mathrm{d}A\end{aligned}$$

Noting that

$$\frac{A(e^{\delta\tau}-1)(\omega^2\vartheta-\Theta+\delta)}{2\delta z}+\frac{A}{z}=\frac{A\left[(\omega^2\vartheta-\Theta+\delta)(e^{\delta\tau}-1)+2\delta\right]}{2\delta}\cdot\frac{1}{z}=A,$$

we find that

$$\begin{aligned}&\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))\\ &\quad= \int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{A^{2\alpha/\omega^2+1}}{\Gamma(\frac{2\alpha}{\omega^2}+1)}e^{-A}\\ &\qquad\times\int_0^1 e^{Ax/z}(1-x)^{2\alpha/\omega^2}z^{-1}\,\mathrm{d}z\,\mathrm{d}A.\end{aligned} \tag{B.15}$$

An inverse Laplace transform in z: Note that $\bar{U}(\tau,\phi_1,\phi_2,\vartheta)$ as defined in equation (3.26) is the Laplace transform parameterized by ϑ. In order to invert the Laplace transform $\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))$ as given by (B.15) (which is expressed as a function of z), we need to establish a relationship between

the Laplace transform with parameter ϑ to one that is parameterized by z.

From the expression for z in equation (B.13), we obtain the following expression for ϑ,

$$\vartheta \equiv \vartheta(z) = \frac{2\delta(z-1)}{\omega^2(e^{\delta\tau}-1)} + \frac{\Theta-\delta}{\omega^2}. \tag{B.16}$$

With this choice of ϑ, the Laplace transform of $\hat{U}(\tau,\phi_1,\phi_2,v)$ in the v-variable is given by

$$\mathscr{L}_v\left\{\hat{U}(\tau,\phi_1,\phi_2,v)\right\}(\vartheta(z)) = \int_0^\infty \exp\left\{-\left[\frac{2\delta(z-1)}{\omega^2(e^{\delta\tau}-1)} + \frac{\Theta-\delta}{\omega^2}\right]v\right\}\hat{U}(\tau,\phi_1,\phi_2,v)\,\mathrm{d}v.$$

Consider the substitution

$$y = \frac{2\delta v}{\omega^2(e^{\delta\tau}-1)}. \tag{B.17}$$

Then $\mathscr{L}_v\left\{\hat{U}(\tau,\phi_1,\phi_2,v)\right\}$ can be written as

$$\begin{aligned}
&\mathscr{L}_v\left\{\hat{U}(\tau,\phi_1,\phi_2,v(y))\right\} \\
&\quad= \int_0^\infty e^{-yz}\exp\left\{y - \left(\frac{\Theta-\delta}{\omega^2}\right)\frac{\omega^2(e^{\delta\tau}-1)}{2\delta}y\right\}\frac{\omega^2(e^{\delta\tau}-1)}{2\delta} \\
&\qquad\times \hat{U}(\tau,\phi_1,\phi_2,v(y))\,\mathrm{d}y \\
&\quad= \frac{\omega^2(e^{\delta\tau}-1)}{2\delta}\int_0^\infty e^{-yz}\exp\left\{-\left[\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}-1\right]y\right\}\hat{U}(\tau,\phi_1,\phi_2,v(y))\,\mathrm{d}y.
\end{aligned}$$

Define the Laplace transform of a function $f(y)$ with respect to y (parameterized by z) by

$$\mathcal{L}_y\{f(y)\}(z) = \int_0^\infty e^{-yz}f(y)\,\mathrm{d}y. \tag{B.18}$$

It follows that

$$\begin{aligned}
&\mathscr{L}_v\left\{\hat{U}(\tau,\phi_1,\phi_2,v(y))\right\} \\
&\quad= \frac{\omega^2(e^{\delta\tau}-1)}{2\delta}\mathscr{L}_y\left\{\exp\left\{-\left[\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}-1\right]y\right\}\hat{U}(\tau,\phi_1,\phi_2,v(y))\right\}.
\end{aligned}$$

By definition (see equation (3.26)), we note that

$$\mathscr{L}_v\left\{\hat{U}(\tau,\phi_1,\phi_2,v(y))\right\} = \bar{U}(\tau,\phi_1,\phi_2,\vartheta(z)),$$

and so the last equation can be rearranged to yield

$$\begin{aligned}
&\mathscr{L}_y\left\{\exp\left\{-\left[\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}-1\right]y\right\}\hat{U}(\tau,\phi_1,\phi_2,v(y))\right\} \\
&\qquad= \frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z)).
\end{aligned}$$

This is equivalent to

$$\hat{U}(\tau,\phi_1,\phi_2,v(y)) = \frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\exp\left\{\left[\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}-1\right]y\right\}\mathscr{L}_y^{-1}\left\{\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))\right\}.$$

Again, by definition, we note that $\hat{U}(\tau,\phi_1,\phi_2,v(y)) = \mathscr{L}_v^{-1}\left\{\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))\right\}$, and so we find that

$$\begin{aligned}&\mathscr{L}_v^{-1}\left\{\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))\right\}\\&\quad= \frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\exp\left\{\left[\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}-1\right]y\right\}\mathscr{L}_y^{-1}\left\{\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))\right\}\end{aligned} \tag{B.19}$$

This last equation establishes the relationship between the inverse Laplace transform with parameter ϑ and that with parameter z.

Evaluating the inverse Laplace transform: Taking the inverse Laplace transform (with respect to v) of both sides of equation (B.15) (and noting the linearity property of Laplace transforms), we find that

$$\begin{aligned}&\mathscr{L}_v^{-1}\left\{\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))\right\}\\&\quad= \mathscr{L}_v^{-1}\left\{\int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{A^{2\alpha/\omega^2+1}}{\Gamma(\frac{2\alpha}{\omega^2}+1)}e^{-A}\right.\\&\qquad\left.\times\int_0^1 e^{Ax/z}(1-x)^{2\alpha/\omega^2}z^{-1}\,\mathrm{d}z\,\mathrm{d}A\right\}\\&\quad= \int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{A^{2\alpha/\omega^2+1}}{\Gamma(\frac{2\alpha}{\omega^2}+1)}e^{-A}\\&\qquad\times\int_0^1(1-x)^{2\alpha/\omega^2}\mathscr{L}_v^{-1}\left\{z^{-1}e^{Ax/z}\right\}\mathrm{d}z\,\mathrm{d}A.\end{aligned}$$

In view of equation (B.19), the inverse Laplace transform in v may be written as an inverse Laplace transform in y as

$$\begin{aligned}&\mathscr{L}_v^{-1}\left\{\bar{U}(\tau,\phi_1,\phi_2,\vartheta(z))\right\}\\&\quad= \int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{A^{2\alpha/\omega^2+1}}{\Gamma(\frac{2\alpha}{\omega^2}+1)}e^{-A}\cdot\frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\\&\qquad\times\exp\left\{\left[\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}-1\right]y\right\}\int_0^1(1-x)^{2\alpha/\omega^2}\mathscr{L}_y^{-1}\left\{z^{-1}e^{Ax/z}\right\}\mathrm{d}z\,\mathrm{d}A.\end{aligned}$$

The inverse Laplace transform which appears above is given by

$$\mathscr{L}_y^{-1}\left\{z^{-1}e^{Ax/z}\right\} = I_0\left(2\sqrt{Axy}\right),$$

where $I_k(u)$ is defined in equation (3.36) (refer to equations 9.6.10 and 29.3.81 of Abramowitz and Stegun, 1964). It therefore follows that

$$\begin{aligned}&\hat{U}(\tau,\phi_1,\phi_2,v(y))\\&\quad= \int_0^\infty H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\frac{A^{2\alpha/\omega^2+1}}{\Gamma(\frac{2\alpha}{\omega^2}+1)}e^{-A}\end{aligned}$$

$$\times \exp\left\{\left[\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}-1\right]y\right\}\int_0^1(1-x)^{2\alpha/\omega^2}I_0\left(2\sqrt{Axy}\right)\mathrm{d}x\,\mathrm{d}A.$$

The calculations that follow show the evaluation of the $\mathrm{d}x$-integral in the last expression above. Expressing the Bessel function as a power series (using equation (3.36)), the integral may be written as

$$\begin{aligned}\int_0^1(1-x)^{2\alpha/\omega^2}I_0\left(2\sqrt{Axy}\right)\mathrm{d}x &= \int_0^1(1-x)^{2\alpha/\omega^2}\sum_{n=0}^{\infty}\frac{(Ay)^n}{\Gamma(n+1)n!}x^n\,\mathrm{d}x\\ &= \sum_{n=0}^{\infty}\frac{(Ay)^n}{\Gamma(n+1)n!}\int_0^1(1-x)^{2\alpha/\omega^2}x^n\,\mathrm{d}x.\end{aligned}$$

It is known that for $m,n>0$ (see p. 343, equation 586 of Zwillinger, 2018)

$$\int_0^1 x^{n-1}(1-x)^{m-1}\,\mathrm{d}x=\frac{\Gamma(n)\Gamma(m)}{\Gamma(n+m)}.$$

This implies that

$$\begin{aligned}&\int_0^1(1-x)^{2\alpha/\omega^2}I_0\left(2\sqrt{Axy}\right)\mathrm{d}x\\ &\quad=\sum_{n=0}^{\infty}\frac{(Ay)^n}{\Gamma(n+1)n!}\cdot\frac{\Gamma(n+1)\Gamma(\frac{2\alpha}{\omega^2}+1)}{\Gamma(n+\frac{2\alpha}{\omega^2}+2)}\\ &\quad=\Gamma\left(\frac{2\alpha}{\omega^2}+1\right)\sum_{n=0}^{\infty}\frac{\left(\sqrt{Ay}\right)^{2n}}{\Gamma(n+\frac{2\alpha}{\omega^2}+1+1)}\\ &\quad=\Gamma\left(\frac{2\alpha}{\omega^2}+1\right)(Ay)^{-\frac{\alpha}{\omega^2}-\frac{1}{2}}\sum_{n=0}^{\infty}\frac{\left(\sqrt{Ay}\right)^{2n+\frac{2\alpha}{\omega^2}+1}}{\Gamma(n+\frac{2\alpha}{\omega^2}+1+1)}\\ &\quad=\Gamma\left(\frac{2\alpha}{\omega^2}+1\right)(Ay)^{-\frac{\alpha}{\omega^2}-\frac{1}{2}}I_{\left(\frac{2\alpha}{\omega^2}+1\right)}\left(2\sqrt{Ay}\right).\end{aligned}$$

Having evaluated the innermost integral, $\hat{U}$ may therefore be written as

$$\begin{aligned}\hat{U}(\tau,\phi_1,\phi_2,v(y)) &= \int_0^{\infty}H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\\ &\quad\times e^{-A-y}\frac{A^{\frac{2\alpha}{\omega^2}+1}}{\Gamma(\frac{2\alpha}{\omega^2}+1)}\exp\left\{\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}y\right\}\\ &\quad\times\Gamma\left(\frac{2\alpha}{\omega^2}+1\right)\frac{A^{-\frac{\alpha}{\omega^2}-\frac{1}{2}}}{y^{\frac{\alpha}{\omega^2}+\frac{1}{2}}}I_{\left(\frac{2\alpha}{\omega^2}+1\right)}\left(2\sqrt{Ay}\right)\mathrm{d}A\\ &= \int_0^{\infty}H\left(\tau,\phi_1,\phi_2,\frac{\omega^2(1-e^{-\delta\tau})A}{2\delta}\right)\frac{\omega^2(1-e^{-\delta\tau})}{2\delta}\frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\\ &\quad\times e^{-A-y}\left(\frac{A}{y}\right)^{\frac{\alpha}{\omega^2}+\frac{1}{2}}\exp\left\{\frac{(\Theta-\delta)(e^{\delta\tau}-1)}{2\delta}y\right\}I_{\left(\frac{2\alpha}{\omega^2}+1\right)}\left(2\sqrt{Ay}\right)\mathrm{d}A.\end{aligned}$$

Using equations (B.13) and (B.17), we may recover the original variables b and v from A and y. We also replace the function H with the right-hand side of equation (B.14). Doing so and simplifying, we have

$$\hat{U}(\tau,\phi_1,\phi_2,v)$$

$$
\begin{aligned}
&= \int_0^\infty H(\tau,\phi_1,\phi_2,b)\frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\exp\left\{-\frac{2\delta(be^{\delta\tau}+v)}{\omega^2(e^{\delta\tau}-1)}\right\} \\
&\quad\times\left(\frac{be^{\delta\tau}}{v}\right)^{\frac{\alpha}{\omega^2}+\frac{1}{2}}\exp\left\{\frac{(\Theta-\delta)}{\omega^2}v\right\}I_{\left(\frac{2\alpha}{\omega^2}+1\right)}\left(\frac{4\delta\sqrt{bve^{\delta\tau}}}{\omega^2(e^{\delta\tau}-1)}\right)\mathrm{d}b \\
&= \int_0^\infty \hat{u}(\phi_1,\phi_2,b)\exp\left\{\left[\frac{\alpha(\Theta-\delta)}{\omega^2}+\Theta+\Psi\right]\tau\right\}\exp\left\{-\left(\frac{\Theta-\delta}{\omega^2}\right)b\right\} \\
&\quad\times\frac{2\delta}{\omega^2(e^{\delta\tau}-1)}\exp\left\{-\frac{2\delta(be^{\delta\tau}+v)}{\omega^2(e^{\delta\tau}-1)}\right\}\left(\frac{be^{\delta\tau}}{v}\right)^{\frac{\alpha}{\omega^2}+\frac{1}{2}}\exp\left\{\frac{(\Theta-\delta)}{\omega^2}v\right\} \\
&\quad\times I_{\left(\frac{2\alpha}{\omega^2}+1\right)}\left(\frac{4\delta\sqrt{bve^{\delta\tau}}}{\omega^2(e^{\delta\tau}-1)}\right)\mathrm{d}b \\
&= \int_0^\infty \hat{u}(\phi_1,\phi_2,b)\exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)(\alpha\tau+v-b)\right\}e^{(\Theta+\Psi)\tau}\frac{2\delta}{\omega^2(e^{\delta\tau}-1)} \\
&\quad\times\left(\frac{be^{\delta\tau}}{v}\right)^{\frac{\alpha}{\omega^2}+\frac{1}{2}}\exp\left\{-\frac{2\delta(be^{\delta\tau}+v)}{\omega^2(e^{\delta\tau}-1)}\right\}I_{\left(\frac{2\alpha}{\omega^2}+1\right)}\left(\frac{4\delta\sqrt{bve^{\delta\tau}}}{\omega^2(e^{\delta\tau}-1)}\right)\mathrm{d}b.
\end{aligned}
$$

Recalling that $\alpha = \xi\eta - \omega^2$, we note that

$$
\exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)(\alpha\tau+v-b)\right\} = \exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)(\xi\eta\tau+v-b)\right\}e^{-\Theta\tau}e^{\delta\tau}.
$$

Replacing all other instances of α with $\xi\eta-\omega^2$ in the last expression for $\hat{U}$ produces the equation stated in the proposition.

B.3 Proof of Proposition 3.9

Using the inversion formula (3.37) with $\hat{U}$ as given in equation (3.35), we find that

$$
\begin{aligned}
&\tilde{U}(\tau,X_1,X_2,v) \\
&\quad= \frac{1}{4\pi^2}\iint_{\mathbb{R}^2} e^{-i\phi_1X_1-i\phi_2X_2}\left[\int_0^\infty \hat{u}(\phi_1,\phi_2,b)\exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)(\xi\eta\tau+v-b)\right\}\right. \\
&\qquad\times e^{\Psi\tau}\frac{2\delta e^{\delta\tau}}{\omega^2(e^{\delta\tau}-1)}\left(\frac{be^{\delta\tau}}{v}\right)^{\frac{\xi\eta}{\omega^2}-\frac{1}{2}}\exp\left\{-\frac{2\delta(be^{\delta\tau}+v)}{\omega^2(e^{\delta\tau}-1)}\right\} \\
&\qquad\left.\times I_{\left(\frac{2\xi\eta}{\omega^2}-1\right)}\left(\frac{4\delta\sqrt{bve^{\delta\tau}}}{\omega^2(e^{\delta\tau}-1)}\right)\mathrm{d}b\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2.
\end{aligned}
$$

We first expand the exponential term $e^{\Psi\tau}$ using the definition of Ψ in equation (3.24). We thus have

$$
\begin{aligned}
e^{\Psi\tau} &= \exp\left\{\sum_{j=1}^{2} i\phi_j\left[\frac{\tilde{\lambda}_j}{i\phi_j}(\varphi_j(\phi_j)-1)-r+q_j+\tilde{\lambda}_j\tilde{\kappa}_j\right]\tau\right\} \\
&= \exp\left\{-\sum_{j=1}^{2} i\phi_j(r-q_j-\tilde{\lambda}_j\tilde{\kappa}_j)\tau\right\}\exp\left\{\tilde{\lambda}_1\tau(\varphi_1(\phi_1)-1)\right\} \\
&\quad\times\exp\left\{\tilde{\lambda}_2\tau(\varphi_2(\phi_2)-1)\right\}.
\end{aligned}
$$

The Maclaurin series expansion of $\exp\left\{\tilde{\lambda}_1\tau\varphi_1(\phi_1)\right\}$ is

$$\exp\left\{\tilde{\lambda}_1\tau\varphi_1(\phi_1)\right\} = \sum_{m=0}^{\infty}\frac{(\tilde{\lambda}_1\tau)^m}{m!}[\varphi_1(\phi_1)]^m.$$

Recall that $\varphi_1(\phi)$ is the $\mathbb{Q}$-characteristic function of the jump random variable Y_1 with $\mathbb{Q}$-density function $g_1(y)$. Let $Y_{1,1},\ldots,Y_{1,m}$ be a random sample from this density function and define $\Upsilon_{1,m} = \sum_{k=1}^{m} Y_{1,k}$. Following the independence of the random variables, $[\varphi_1(\phi)]^m$ gives the $\mathbb{Q}$-characteristic function of $\Upsilon_{1,m}$.[1] We can therefore write

$$\exp\left\{\tilde{\lambda}_1\tau(\varphi_1(\phi_1)-1)\right\} = e^{-\tilde{\lambda}_1\tau}\sum_{m=0}^{\infty}\frac{(\tilde{\lambda}_1\tau)^m}{m!}\mathbb{E}_{\mathbb{Q}}\left[e^{-i\phi_1\Upsilon_{1,m}}\right].$$

A similar argument yields

$$\exp\left\{\tilde{\lambda}_2\tau(\varphi_2(\phi_2)-1)\right\} = e^{-\tilde{\lambda}_2\tau}\sum_{n=0}^{\infty}\frac{(\tilde{\lambda}_2\tau)^n}{n!}\mathbb{E}_{\mathbb{Q}}\left[e^{-i\phi_2\Upsilon_{2,n}}\right],$$

where $\Upsilon_{2,n} = \sum_{l=1}^{n} Y_{2,l}$ and $Y_{2,1},\ldots,Y_{2,n}$ are i.i.d. random variables with common density function $g_2(y)$ and characteristic function $\varphi_2(\phi)$ under $\mathbb{Q}$. It follows that

$$\begin{aligned} e^{\Psi\tau} = \exp&\left\{-\sum_{j=1}^{2} i\phi_j(r-q_j-\tilde{\lambda}_j\tilde{\kappa}_j)\tau\right\}e^{-\left(\tilde{\lambda}_1+\tilde{\lambda}_2\right)\tau} \\ &\times\left[\sum_{m=0}^{\infty}\frac{(\tilde{\lambda}_1\tau)^m}{m!}\mathbb{E}_{\mathbb{Q}}\left[e^{-i\phi_1\Upsilon_{1,m}}\right]\right]\left[\sum_{n=0}^{\infty}\frac{(\tilde{\lambda}_2\tau)^n}{n!}\mathbb{E}_{\mathbb{Q}}\left[e^{-i\phi_2\Upsilon_{2,n}}\right]\right]. \end{aligned}$$

Since the jumps in the price of one asset are independent from those in the other asset, we may write

$$\begin{aligned} e^{\Psi\tau} = \exp&\left\{-\sum_{j=1}^{2} i\phi_j(r-q_j-\tilde{\lambda}_j\tilde{\kappa}_j)\tau\right\}e^{-\left(\tilde{\lambda}_1+\tilde{\lambda}_2\right)\tau} \\ &\times\sum_{m,n=0}^{\infty}\frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!}\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{-i\phi_1\Upsilon_{1,m}-i\phi_2\Upsilon_{2,n}\right\}\right]. \end{aligned}$$

Next, we note that $\hat{u}(\phi_1,\phi_2,b)$ (as defined in equation (3.25)) denotes the Fourier transform of initial condition (3.17) of IPDE (3.16). As we are already using the variables X_1 and X_2, we rewrite $\hat{u}(\phi_1,\phi_2,b)$ as

$$\hat{u}(\phi_1,\phi_2,b) = \iint_{\mathbb{R}^2} e^{i\phi_1y_1+i\phi_2y_2}(e^{y_1}-e^{y_2})^+\,\mathrm{d}y_1\,\mathrm{d}y_2.$$

In light of the preceding calculations and with $F(b)$ as defined in equation (3.40), we can therefore write $\tilde{U}(\tau,X_1,X_2,v)$ as

$$\tilde{U}(\tau,X_1,X_2,v)$$

[1] Given that the $Y_{1,k}$'s are i.i.d. with common characteristic function $\varphi_1(\phi)$, we have

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-i\phi\Upsilon_{1,m}}\right] = \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{-i\phi\sum_{k=1}^{m}Y_{1,k}\right\}\right] = \mathbb{E}_{\mathbb{Q}}\left[\prod_{k=1}^{m}e^{-i\phi Y_{1,k}}\right] = \prod_{k=1}^{m}\mathbb{E}_{\mathbb{Q}}\left[e^{-i\phi Y_{1,k}}\right] = [\varphi_1(\phi)]^m.$$

$$
\begin{aligned}
&= \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} e^{-i\phi_1 X_1 - i\phi_2 X_2} \int_0^\infty \left[\iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} (e^{y_1} - e^{y_2})^+ \, \mathrm{d}y_1 \, \mathrm{d}y_2 \right] \\
&\quad \times \exp\left\{ -\sum_{j=1}^{2} i\phi_j (r - q_j - \tilde{\lambda}_j \tilde{\kappa}_j)\tau \right\} e^{-\left(\tilde{\lambda}_1 + \tilde{\lambda}_2\right)\tau} \\
&\quad \times \left(\sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1 \tau)^m (\tilde{\lambda}_2 \tau)^n}{m! n!} \mathbb{E}_{\mathbb{Q}} \left[\exp\left\{ -i\phi_1 \Upsilon_{1,m} - i\phi_2 \Upsilon_{2,n} \right\} \right] \right) \\
&\quad \times F(b) \, \mathrm{d}b \, \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.
\end{aligned}
$$

We recall that $X_j = \ln S_j$, where S_j is the price of asset j observed at time to maturity τ (see equation (3.12)). Rearranging the terms and the order of integration, we can thus write

$$
\tilde{U}(\tau, X_1, X_2, v) = \int_0^\infty \iint_{\mathbb{R}^2} \left[(e^{y_1} - e^{y_2})^+ \right] G(\tau, y_1, y_2, b; S_1, S_2, v) \, \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}b,
$$

where $G(\tau, y_1, y_2, b; S_1, S_2, v)$ is given by equation (3.39).

Finally, we note the relation

$$
C^E(\tau, S_1, S_2, v) = e^{-r\tau} \tilde{U}(\tau, X_1, X_2, v)
$$

as a result of the representations made under equations (3.12) and (3.15). From here, the expression for $C^E(\tau, S_1, S_2, v)$ shown in the proposition follows.

B.4 Proof of Lemma 3.10

We consider the substitutions A and y from equations (B.13) and (B.17). It follows that

$$
\begin{aligned}
\int_0^\infty F(b) \, \mathrm{d}b &= \int_0^\infty \left(\frac{A}{y} \right)^{\xi\eta/\omega^2 - 1/2} e^{-A-y} \exp\left\{ \left(\frac{\Theta - \delta}{\omega^2} \right) \xi\eta\tau \right\} \\
&\quad \times \exp\left\{ \left(\frac{\Theta - \delta}{\omega^2} \right) \frac{\omega^2 (e^{\delta\tau} - 1)}{2\delta} y \right\} \exp\left\{ -\left(\frac{\Theta - \delta}{\omega^2} \right) \frac{\omega^2 (1 - e^{-\delta\tau})}{2\delta} A \right\} \\
&\quad \times I_{\left(\frac{2\xi\eta}{\omega^2} - 1\right)} \left(2\sqrt{Ay} \right) \mathrm{d}A \\
&= e^{-y} \exp\left\{ \left(\frac{\Theta - \delta}{\omega^2} \right) \xi\eta\tau \right\} \exp\left\{ \left(\frac{\Theta - \delta}{\omega^2} \right) \frac{\omega^2 (e^{\delta\tau} - 1)}{2\delta} y \right\} \\
&\quad \times \int_0^\infty \left(\frac{A}{y} \right)^{\xi\eta/\omega^2 - 1/2} e^{-A} \exp\left\{ -\left(\frac{\Theta - \delta}{\omega^2} \right) \frac{\omega^2 (1 - e^{-\delta\tau})}{2\delta} A \right\} \\
&\quad \times \sum_{n=0}^{\infty} \frac{\left(\sqrt{Ay}\right)^{2n + 2\xi\eta/\omega^2 - 1}}{n! \Gamma(n + \frac{2\xi\eta}{\omega^2})} \mathrm{d}A
\end{aligned}
$$

Let $F_1(y)$ denote the integral with respect to A. Simplifying, we have

$$
\begin{aligned}
F_1(y) &= \int_0^\infty A^{\xi\eta/\omega^2 - 1/2} y^{1/2 - \xi\eta/\omega^2} \exp\left\{ -\left[\frac{(\Theta - \delta)(1 - e^{-\delta\tau})}{2\delta} + 1 \right] A \right\} \\
&\quad \times \sum_{n=0}^{\infty} \frac{y^{n + \xi\eta/\omega^2 - 1/2} A^{n + \xi\eta/\omega^2 - 1/2}}{n! \Gamma(n + \frac{2\xi\eta}{\omega^2})} \mathrm{d}A \\
&= \sum_{n=0}^{\infty} y^n \int_0^\infty \exp\left\{ -\left[\frac{(\Theta - \delta)(1 - e^{-\delta\tau})}{2\delta} + 1 \right] A \right\} \frac{A^{n + 2\xi\eta/\omega^2 - 1}}{n! \Gamma(n + \frac{2\xi\eta}{\omega^2})} \mathrm{d}A.
\end{aligned}
$$

Consider the substitution

$$u = \left[\frac{(\Theta-\delta)(1-e^{-\delta\tau})}{2\delta} + 1\right] A.$$

We then have

$$\begin{aligned} F_1(y) &= \sum_{n=0}^{\infty} \frac{y^n}{n!\Gamma(n+\frac{2\xi\eta}{\omega^2})} \int_0^{\infty} e^{-u} u^{n+2\xi\eta/\omega^2-1} \left[\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right]^{n+\frac{2\xi\eta}{\omega^2}-1} \\ &\quad \times \frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\,\mathrm{d}u \\ &= \sum_{n=0}^{\infty} \frac{y^n}{n!\Gamma(n+\frac{2\xi\eta}{\omega^2})} \left[\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right]^{n+\frac{2\xi\eta}{\omega^2}} \\ &\quad \times \int_0^{\infty} e^{-u} u^{n+2\xi\eta/\omega^2-1}\,\mathrm{d}u. \end{aligned}$$

We recognize the last integral as a gamma function, and so we have

$$F_1(y) = \left[\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right]^{\frac{2\xi\eta}{\omega^2}} \sum_{n=0}^{\infty} \frac{1}{n!} \left[\frac{2\delta y}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right]^n \frac{\Gamma(n+\frac{2\xi\eta}{\omega^2})}{\Gamma(n+\frac{2\xi\eta}{\omega^2})}.$$

Finally, noting that the sum is the power series expansion of an exponential function, we find that

$$F_1(y) = \left[\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right]^{\frac{2\xi\eta}{\omega^2}} \exp\left\{\frac{2\delta y}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right\}.$$

Going back to $\int_0^{\infty} F(b)\,\mathrm{d}b$, we then see that

$$\begin{aligned} \int_0^{\infty} F(b)\,\mathrm{d}b &= e^{-y} \exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)\xi\eta\tau\right\} \exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)\frac{\omega^2(e^{\delta\tau}-1)}{2\delta}y\right\} \\ &\quad \times \left[\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right]^{\frac{2\xi\eta}{\omega^2}} \exp\left\{\frac{2\delta y}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right\} \end{aligned}$$

Reversing the substitution of y for v in equation (B.17), we obtain the desired result.

B.5 Proof of Lemma 3.11

From equation (3.39), we have

$$\begin{aligned} &\int_0^{\infty} G(\tau, y_1, y_2, b; S_1, S_2, v)\,\mathrm{d}b \\ &\quad = \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau} \exp\{-i\phi_1 \ln S_1 - i\phi_2 \ln S_2\} \\ &\qquad \times \exp\left\{-\sum_{j=1}^{2} i\phi_j(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j)\tau\right\} \\ &\qquad \times \left[\sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!} \mathbb{E}_{\mathbb{Q}}\left[\exp\{-i\phi_1\Upsilon_{1,m} - i\phi_2\Upsilon_{2,n}\}\right]\right] \\ &\qquad \times \int_0^{\infty} F(b)\,\mathrm{d}b\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2 \end{aligned}$$

$$
\begin{aligned}
&= \sum_{m,n=0}^{\infty} \frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau} \mathbb{E}_{\mathbb{Q}}\Bigg[\frac{1}{4\pi^2}\iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} \\
&\quad \times \exp\left\{-i\phi_1\left(\ln S_1 + \Upsilon_{1,m}\right) - i\phi_2\left(\ln S_2 + \Upsilon_{2,n}\right)\right\} \\
&\quad \times \exp\left\{-\sum_{j=1}^{2} i\phi_j(r - q_j - \tilde{\lambda}_j\tilde{\kappa}_j)\tau\right\} \exp\left\{\left(\frac{\Theta-\delta}{\omega^2}\right)(\xi\eta\tau + v)\right\} \\
&\quad \times \exp\left\{-\frac{2\delta v}{\omega^2(e^{\delta\tau}-1)}\right\} \exp\left\{\frac{2\xi\eta}{\omega^2}\ln\left[\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\right]\right\} \\
&\quad \times \exp\left\{\frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\cdot\frac{2\delta v}{\omega^2(e^{\delta\tau}-1)}\right\} \mathrm{d}\phi_1\,\mathrm{d}\phi_2\Bigg].
\end{aligned}
$$

Note that this integral may be written as in (3.43), where $B_1(\tau, -\phi_1, -\phi_2)$ and $B_2(\tau, -\phi_1, -\phi_2)$ are given by

$$
\begin{aligned}
B_1(\tau, -\phi_1, -\phi_2) &= -i\phi_1(r-q_1)\tau - i\phi_2(r-q_2)\tau \\
&\quad + \frac{\xi\eta}{\omega^2}\left\{(\Theta-\delta)\tau - 2\ln\left[\frac{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}{2\delta}\right]\right\}
\end{aligned} \tag{B.20}
$$

and

$$
B_2(\tau, -\phi_1, -\phi_2) = \frac{\Theta-\delta}{\omega^2} - \frac{2\delta}{\omega^2(e^{\delta\tau}-1)} + \frac{2\delta}{(\Theta-\delta)(1-e^{-\delta\tau})+2\delta}\cdot\frac{2\delta}{\omega^2(e^{\delta\tau}-1)}. \tag{B.21}
$$

Using the representations made in (3.46), the expressions for B_1 and B_2 can be further simplified.[2] First, we have from equation (B.20)

$$
\begin{aligned}
B_1(\tau, \phi_1, \phi_2) &= i\phi_1(r-q_1) + i\phi_2(r-q-2) \\
&\quad + \frac{\xi\eta}{\omega^2}\left\{(\Theta_2-\delta_2)\tau - 2\ln\left[\frac{(\Theta_2-\delta_2)(1-e^{-\delta_2\tau})+2\delta_2}{2\delta_2}\right]\right\}.
\end{aligned}
$$

We note that

$$
\begin{aligned}
\frac{(\Theta_2-\delta_2)(1-e^{-\delta_2\tau})+2\delta_2}{2\delta_2} &= \frac{(\Theta_2-\delta_2)(e^{\delta_2\tau}-1)+2\delta_2 e^{\delta_2\tau}}{2\delta_2 e^{\delta_2\tau}} \\
&= \frac{(\Theta_2-\delta_2)e^{\delta_2\tau} - (\Theta_2-\delta_2) + 2\delta_2 e^{\delta_2\tau}}{2\delta_2 e^{\delta_2\tau}}
\end{aligned}
$$

[2] From equation (3.24), we note that

$$
\Theta_2 = \Theta(-\phi_1, -\phi_2) = \xi + \Lambda - i\sum_{j=1}^{2}\rho_{wz_j}\sigma_j\omega\phi_j.
$$

Likewise, from the definition of $\delta = \delta(\phi_1, \phi_2)$ in equation (3.32), we note that

$$
\delta_2 = \delta(-\phi_1, -\phi_2) = \sqrt{\Theta^2(-\phi_1, -\phi_2) - \varepsilon(-\phi_1, -\phi_2)\omega^2} = \sqrt{\Theta_2^2 - \varepsilon_2\omega^2},
$$

where we define

$$
\begin{aligned}
\varepsilon_2 &= \varepsilon_2(\phi_1, \phi_2) \equiv \varepsilon(-\phi_1, -\phi_2) \\
&= -\sum_{j=1}^{2}\sigma_j^2(i\phi_j + \phi_j^2) - 2\rho_w\sigma_1\sigma_2\phi_1\phi_2,
\end{aligned}
$$

following the definition of $\varepsilon(\phi_1, \phi_2)$ in (3.24).

$$
\begin{aligned}
&= \frac{(\Theta_2+\delta_2)e^{\delta_2\tau}-(\Theta_2-\delta_2)}{2\delta_2}\cdot e^{-\delta_2\tau} \\
&= \frac{[(\Theta_2+\delta_2)/(\Theta_2-\delta_2)]e^{\delta_2\tau}-1}{2\delta_2/(\Theta_2-\delta_2)}\cdot e^{-\delta_2\tau} \\
&= \frac{\chi_2 e^{\delta_2\tau}-1}{[(\Theta_2+\delta_2)-(\Theta_2-\delta_2)]/(\Theta_2-\delta_2)}\cdot e^{-\delta_2\tau} \\
&= \frac{\chi_2 e^{\delta_2\tau}-1}{\chi_2-1}\cdot e^{-\delta_2\tau}.
\end{aligned}
$$

Using this result in the above expression for $B_1(\tau,\phi_1,\phi_2)$ and rearranging the terms produces equation (3.44). Similarly, if we use the substitutions (3.46) in equation (B.21), then we obtain

$$
\begin{aligned}
B_2(\tau,\phi_1,\phi_2) &= \frac{\Theta_2-\delta_2}{\omega^2} - \frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)} + \frac{2\delta_2}{(\Theta_2-\delta_2)(1-e^{-\delta_2\tau})+2\delta_2}\cdot\frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)} \\
&= \frac{\Theta_2-\delta_2}{\omega^2} - \frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)} + \frac{e^{\delta_2\tau}(\chi_2-1)}{\chi_2 e^{\delta_2\tau}-1}\cdot\frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)},
\end{aligned}
$$

where the last equation follows from the calculations made immediately prior. Performing additional manipulations on B_2, we have

$$
\begin{aligned}
B_2(\tau,\phi_1,\phi_2) &= \frac{\Theta_2-\delta_2}{\omega^2} - \frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)}\left[1-\frac{e^{\delta_2\tau}(\chi_2-1)}{\chi_2 e^{\delta_2\tau}-1}\right] \\
&= \frac{\Theta_2-\delta_2}{\omega^2} - \frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)}\left[\frac{\chi_2 e^{\delta_2\tau}-1-e^{\delta_2\tau}(\chi_2-1)}{\chi_2 e^{\delta_2\tau}-1}\right] \\
&= \frac{\Theta_2-\delta_2}{\omega^2} - \frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)}\left[\frac{e^{\delta_2\tau}-1}{\chi_2 e^{\delta_2\tau}-1}\right] \\
&= \left[\frac{\Theta_2-\delta_2}{\omega^2}\cdot\frac{\chi_2 e^{\delta_2\tau}-1}{e^{\delta_2\tau}-1} - \frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)}\right]\left[\frac{e^{\delta_2\tau}-1}{\chi_2 e^{\delta_2\tau}-1}\right] \\
&= \left[\frac{\Theta_2-\delta_2}{\omega^2}\cdot\frac{\left(\frac{\Theta_2+\delta_2}{\Theta_2-\delta_2}\right)e^{\delta_2\tau}-1}{e^{\delta_2\tau}-1} - \frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)}\right]\left[\frac{e^{\delta_2\tau}-1}{\chi_2 e^{\delta_2\tau}-1}\right] \\
&= \left[\frac{\Theta_2-\delta_2}{\omega^2}\cdot\frac{(\Theta_2+\delta_2)e^{\delta_2\tau}-(\Theta_2-\delta_2)}{(\Theta_2-\delta_2)(e^{\delta_2\tau}-1)} - \frac{2\delta_2}{\omega^2(e^{\delta_2\tau}-1)}\right]\left[\frac{e^{\delta_2\tau}-1}{\chi_2 e^{\delta_2\tau}-1}\right] \\
&= \left[\frac{(\Theta_2+\delta_2)e^{\delta_2\tau}-(\Theta_2-\delta_2)+2\delta_2}{\omega^2(e^{\delta_2\tau}-1)}\right]\left[\frac{e^{\delta_2\tau}-1}{\chi_2 e^{\delta_2\tau}-1}\right]
\end{aligned}
$$

Equation (3.45) quickly follows from the last statement.

B.6 Proof of Proposition 3.14

The proof starts with equation (3.48). For brevity, we set

$$
f_2(-\phi_1,-\phi_2) \equiv f_2(\tau, S_1 e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau}, S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau}, v; -\phi_1,-\phi_2).
$$

Let $\mathbb{I}$ be the four-dimensional integral inside the expectation operator in equation (3.48). We note that

$$\begin{aligned}\mathbb{I} &= \iint_{\mathbb{R}^2} f_2(-\phi_1,-\phi_2)\int_{-\infty}^{\infty}\int_{y_2}^{\infty}\left[e^{(i\phi_1+1)y_1+i\phi_2 y_2}-e^{i\phi_1 y_1+(i\phi_2+1)y_2}\right]\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2\\ &= \iint_{\mathbb{R}^2} f_2(-\phi_1,-\phi_2)\int_{-\infty}^{\infty} e^{i\phi_1 y_2}\int_{y_2}^{\infty} e^{i(\phi_1-i)y_1}\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\ &\quad - \iint_{\mathbb{R}^2} f_2(-\phi_1,-\phi_2)\int_{-\infty}^{\infty} e^{i(\phi_2-i)y_2}\int_{y_2}^{\infty} e^{i\phi_1 y_2}\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2.\end{aligned} \tag{B.22}$$

We denote by $\mathbb{I}_1$ and $\mathbb{I}_2$ the first and second terms, respectively, of the right-hand side.

Evaluating $\mathbb{I}_1$: We first note that

$$\begin{aligned}&\int_{-\infty}^{\infty} e^{i\phi_1 y_2}\int_{y_2}^{\infty} e^{i(\phi_1-i)y_1}\,\mathrm{d}y_1\,\mathrm{d}y_2\\ &\quad= \int_{-\infty}^{\infty} e^{i\phi_2 y_2}\lim_{b_1\to\infty}\int_{y_2}^{b_1} e^{i(\phi_1-i)y_1}\,\mathrm{d}y_1\,\mathrm{d}y_2\\ &\quad= \int_{-\infty}^{\infty} e^{i\phi_2 y_2}\lim_{b_1\to\infty}\left[\frac{e^{i(\phi_1-i)b_1}-e^{i(\phi_1-i)y_2}}{i(\phi_1-i)}\right]\mathrm{d}y_2\\ &\quad= \lim_{b_1\to\infty}\frac{1}{i(\phi_1-i)}\left\{\int_0^{\infty}\left[e^{i\phi_2 y_2}e^{i(\phi_1-i)b_1}-e^{(i(\phi_1-i)+i\phi_2)y_2}\right]\mathrm{d}y_2\right.\\ &\qquad\left.+\int_0^{\infty}\left[e^{-i\phi_2 y_2}e^{i(\phi_1-i)b_1}-e^{-(i(\phi_1-i)+i\phi_2)y_2}\right]\mathrm{d}y_2\right\}\\ &\quad= \lim_{b_1\to\infty}\frac{1}{i(\phi_1-i)}\left\{\lim_{c_1\to\infty}\int_0^{c_1}\left[e^{i\phi_2 y_2}e^{i(\phi_1-i)b_1}-e^{(i(\phi_1-i)+i\phi_2)y_2}\right]\mathrm{d}y_2\right.\\ &\qquad\left.+\lim_{c_2\to\infty}\int_0^{c_2}\left[e^{-i\phi_2 y_2}e^{i(\phi_1-i)b_1}-e^{-(i(\phi_1-i)+i\phi_2)y_2}\right]\mathrm{d}y_2\right\}\\ &\quad= \lim_{b_1\to\infty}\frac{1}{i(\phi_1-i)}\left\{\lim_{c_1\to\infty}\left[\frac{e^{i\phi_2 y_2}e^{i(\phi_1-i)b_1}}{i\phi_2}-\frac{e^{(i(\phi_1-i)+i\phi_2)y_2)}}{i(\phi_1-i)+i\phi_2}\right|_0^{c_1}\right]\right.\\ &\qquad\left.+\lim_{c_2\to\infty}\left[\frac{e^{-i\phi_2 y_2}e^{i(\phi_1-i)b_1}}{-i\phi_2}-\frac{e^{-(i(\phi_1-i)+i\phi_2)y_2}}{-(i(\phi_1-i)+i\phi_2)}\right|_0^{c_2}\right]\right\}\\ &\quad= \lim_{b_1,c_1,c_2\to\infty}\frac{1}{i(\phi_1-i)}\left[\frac{e^{i(\phi_1-i)b_1+i\phi_2 c_1}-e^{i(\phi_1-i)b_1-i\phi_2 c_2}}{i\phi_2}\right.\\ &\qquad\left.-\frac{e^{i(\phi_1-i)c_1+i\phi_2 c_1}-e^{-i(\phi_1-i)c_2+i\phi_2 c_2}}{i(\phi_1-i)+i\phi_2}\right]\end{aligned}$$

Consider the substitution $\xi_1=\phi_1-i$. Then $\mathbb{I}_1$ becomes

$$\begin{aligned}\mathbb{I}_1 &= \iint_{\mathbb{R}^2} f_2(-\phi_1,-\phi_2)\int_{-\infty}^{\infty} e^{i\phi_1 y_2}\int_{y_2}^{\infty} e^{i(\phi_1-i)y_1}\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(-\xi_1-i,-\phi_2)\lim_{b_1,c_1,c_2\to\infty}\left[\frac{e^{i\xi_1 b_1+i\phi_2 c_1}-e^{i\xi_1 b_1-i\phi_2 c_2}}{i\xi_1 i\phi_2}\right.\\ &\quad\left.-\frac{e^{i\xi_1 c_1+i\phi_2 c_1}-e^{-i\xi_1 c_2-i\phi_2 c_2}}{i\xi_1(i\xi_1+i\phi_2)}\right]\mathrm{d}\xi_1\,\mathrm{d}\phi_2\\ &= \lim_{b_1,c_1,c_2\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(-\xi_1-i,-\phi_2)\left[\frac{e^{i\xi_1 b_1+i\phi_2 c_1}-e^{i\xi_1 b_1-i\phi_2 c_2}}{i\xi_1 i\phi_2}\right]\mathrm{d}\xi_1\,\mathrm{d}\phi_2\end{aligned}$$

$$-\lim_{c_1,c_2\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(-\xi_1-i,-\phi_2)\left[\frac{e^{i\xi_1c_1+i\phi_2c_1}-e^{-i\xi_1c_2-i\phi_2c_2}}{i\xi_1(i\xi_1+i\phi_2)}\right]\mathrm{d}\xi_1\,\mathrm{d}\phi_2$$

Starting with the last expression above, we replace ξ_1 with ϕ_1 as the first variable of integration. Furthermore, we note that we can write

$$\mathbb{I}_1=\lim_{b_1,c_2\to\infty}\mathbb{I}_{11}(b_1,c_1)-\lim_{b_1,c_2\to\infty}\mathbb{I}_{12}(b_1,c_2)-\lim_{c_1,c_2\to\infty}\mathbb{I}_{13}(c_1,c_2),$$

where

$$\begin{aligned}
\mathbb{I}_{11}(b_1,c_1)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f_2(-\phi_1-i,-\phi_2)e^{i\phi_1b_1+i\phi_2c_1}}{i\phi_1i\phi_2}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2,\\
\mathbb{I}_{12}(b_1,c_2)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f_2(-\phi_1-i,-\phi_2)e^{i\phi_1b_1-i\phi_2c_2}}{i\phi_1i\phi_2}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2,\\
\mathbb{I}_{13}(c_1,c_2)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_2(-\phi_1-i,-\phi_2)\left[\frac{e^{i\phi_1c_1+i\phi_2c_1}-e^{-i\phi_1c_2-i\phi_2c_2}}{i\phi_1(i\phi_1+i\phi_2)}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2.
\end{aligned}$$

With some modifications to the respective integrands, we can use Lemma 3.12 to evaluate $\mathbb{I}_{11}$ and $\mathbb{I}_{12}$.

We start with $\mathbb{I}_{11}$. First, by splitting the region of integration into quadrants, we observe that the integral can be written as

$$\begin{aligned}
\mathbb{I}_{11}(b_1,c_1)&=\int_0^{\infty}\int_0^{\infty}\left[\frac{f_2(-\phi_1-i,-\phi_2)e^{i\phi_1b_1+i\phi_2c_1}}{i\phi_1i\phi_2}-\frac{f_2(\phi_1-i,-\phi_2)e^{-i\phi_1b_1+i\phi_2c_1}}{i\phi_1i\phi_2}\right.\\
&\qquad\left.-\frac{f_2(-\phi_1-i,\phi_2)e^{i\phi_1b_1-i\phi_2c_1}}{i\phi_1i\phi_2}+\frac{f_2(\phi_1-i,\phi_2)e^{-i\phi_1b_1-i\phi_2c_1}}{i\phi_1i\phi_2}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2\\
&=\int_0^{\infty}\int_0^{\infty}\Delta_{\phi_1}\Delta_{\phi_2}\left[\frac{f_2(\phi_1-i,\phi_2)e^{-i\phi_1b_1-i\phi_2c_1}}{i\phi_1i\phi_2}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2,
\end{aligned}$$

where we used the Δ-notation in Lemma 3.12.

Before we proceed, we first look at $f_2(\phi_1-i,\phi_2)$. Letting $z_1=S_1e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau}$ and $z_2=S_2e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau}$, we note that

$$\begin{aligned}
f_2(\phi_1-i,\phi_2)&=f_2(\tau,z_1,z_2,v;\phi_1-i,\phi_2)\\
&=\exp\left\{i(\phi_1-i)\ln z_1+i\phi_2\ln z_2+B_1(\tau,\phi_1-i,\phi_2)+B_2(\tau,\phi_1-i,\phi_2)v\right\}\\
&=z_1\exp\left\{i\phi_1\ln z_1+i\phi_2\ln z_2+B_1(\tau,\phi_1-i,\phi_2)+B_2(\tau,\phi_1-i,\phi_2)v\right\}.
\end{aligned}$$

We note further that

$$\begin{aligned}
B_1(\tau,\phi_1-i,\phi_2)&=i(\phi_1-i)(r-q_1)\tau+i\phi_2(r-q_2)\tau\\
&\qquad+\frac{\xi\eta}{\omega^2}\left\{\left(\Theta_2(\phi_1-i,\phi_2)+\delta_2(\phi_1-i,\phi_2)\right)\tau\right.\\
&\qquad\left.-2\ln\left[\frac{1-\chi_2(\phi_1-i,\phi_2)e^{\delta_2(\phi_1-i,\phi_2)\tau}}{1-\chi_2(\phi_1-i,\phi_2)}\right]\right\}\\
&=(r-q_1)\tau+B_{11}(\tau,\phi_1,\phi_2),
\end{aligned}$$

where $B_{11}(\tau,\phi_1,\phi_2)$ is defined in Definition 3.13. Also, we note that $B_2(\tau,\phi_1-i,\phi_2)=B_{21}(\tau,\phi_1,\phi_2)$, which is also defined in Definition 3.13. With the aforementioned notation, $f_2(\phi_1-i,\phi_2)$ can therefore be written as

$$f_2(\phi_1-i,\phi_2)=z_1e^{(r-q_1)\tau}f_{21}(\phi_1,\phi_2),$$

where f_{21} is given in Definition 3.13.

We can now rewrite $\mathbb{I}_{11}$ as

$$\mathbb{I}_{11}(b_1, c_1) = z_1 e^{(r-q_1)\tau} \int_0^\infty \int_0^\infty \Delta_{\phi_1}\Delta_{\phi_2} \frac{f_{21}(\phi_1, \phi_2) e^{-i\phi_1 b_1 - i\phi_2 c_1}}{i\phi_1 i\phi_2} \, \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.$$

Let $F_1(x_1, x_2)$ be a bivariate cumulative distribution function whose corresponding characteristic function is $f_{21}(\phi_1, \phi_2)$, and let $F_{11}(x_1)$ and $F_{12}(x_2)$ be its associated marginal cdfs (in x_1 and x_2, respectively). Then by Lemma 3.12, we have

$$\mathbb{I}_{11}(b_1, c_1) = \pi^2 z_1 e^{(r-q_1)\tau} \left\{ 4F_1(b_1, c_1) - 2[F_{11}(b_1) + F_{12}(c_1)] + 1 \right\}.$$

Taking the limit as $b_1, c_1 \to \infty$, we obtain

$$\lim_{b_1, c_1 \to \infty} \mathbb{I}_{11}(b_1, c_1) = \pi^2 z_1 e^{(r-q_1)\tau}.$$

The same strategy can be used to evaluate $\mathbb{I}_{12}$. By splitting the region of integration and using the Δ-notation, we find that

$$\begin{aligned}
\mathbb{I}_{12}(b_1, c_2) = \int_0^\infty \int_0^\infty \Bigg[& \frac{f_2(-\phi_1 - i, -\phi_2) e^{i\phi_1 b_1 - i\phi_2 c_2}}{i\phi_2 i\phi_2} - \frac{f_2(\phi_1 - i, -\phi_2) e^{-i\phi_1 b_1 - i\phi_2 c_2}}{i\phi_1 i\phi_2} \\
& - \frac{f_2(-\phi_1 - i, \phi_2) e^{i\phi_1 b_1 + i\phi_2 c_2}}{i\phi_1 i\phi_2} + \frac{f_2(\phi_1 - i, \phi_2) e^{-i\phi_1 b_1 + i\phi_2 c_2}}{i\phi_1 i\phi_2} \Bigg] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2 \\
= - \int_0^\infty \int_0^\infty & \Delta_{\phi_1}\Delta_{\phi_2} \left[\frac{f_2(\phi_1 - i, -\phi_2) e^{-i\phi_1 b_1 - i\phi_2 c_2}}{i\phi_1 i\phi_2} \right] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.
\end{aligned}$$

In terms of $f_{21}(\phi_1, \phi_2)$, we can write

$$\mathbb{I}_{12}(b_1, c_2) = -z_1 e^{(r-q_1)\tau} \int_0^\infty \int_0^\infty \Delta_{\phi_1}\Delta_{\phi_2} \left[\frac{f_{21}(\phi_1, -\phi_2) e^{-i\phi_1 b_1 - i\phi_2 c_2}}{i\phi_1 i\phi_2} \right] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.$$

Let $h_1(\phi_1, \phi_2) = f_{21}(\phi_1, -\phi_2)$. Furthermore, let $F_2(x_1, x_2)$ be a bivariate cumulative distribution function whose corresponding characteristic function is $h_1(\phi_1, \phi_2)$, and let $F_{21}(x_1)$ and $F_{22}(x_2)$ be its associated marginal cdfs (in x_1 and x_2, respectively). Then by Lemma 3.12, $\mathbb{I}_{12}$ can be simplified as

$$\mathbb{I}_{12} = -\pi^2 z_1 e^{(r-q_1)\tau} \left\{ 4F_2(b_1, c_2) - 2[F_{21}(b_1) + F_{22}(c_2)] + 1 \right\}.$$

Taking the limit as $b_1, c_2 \to \infty$, we find that

$$\lim_{b_1, c_2 \to \infty} \mathbb{I}_{12}(b_1, c_2) = -\pi^2 z_1 e^{(r-q_1)\tau}.$$

Finally, we note that $\mathbb{I}_{13}$ can be written in terms of f_{21} as

$$\mathbb{I}_{13}(c_1, c_2) = z_1 e^{(r-q_1)\tau} \iint_{\mathbb{R}^2} \frac{f_{21}(-\phi_1, -\phi_2) \left[e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2} \right]}{i\phi_1(i\phi_1 + i\phi_2)} \, \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.$$

For now, we retain this form for $\mathbb{I}_{13}$.

Having simplified $\mathbb{I}_{11}$ and $\mathbb{I}_{12}$, we therefore have

$$\begin{aligned}\mathbb{I}_1 &= 2\pi^2 z_1 e^{(r-q_1)\tau} \\ &\quad - z_1 e^{(r-q_1)\tau} \lim_{c_1,c_2\to\infty} \iint_{\mathbb{R}^2} \frac{f_{21}(-\phi_1,-\phi_2)\left[e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2}\right]}{i\phi_1(i\phi_1+i\phi_2)}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\ &= S_1 e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau} e^{(r-q_1)\tau} \\ &\quad \times \left[2\pi^2 - \lim_{c_1,c_2\to\infty} \iint_{\mathbb{R}^2} \frac{f_{21}(-\phi_1,-\phi_2)\left[e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2}\right]}{i\phi_1(i\phi_1+i\phi_2)}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2\right].\end{aligned} \tag{B.23}$$

Evaluating $\mathbb{I}_2$: The calculations above can adapted to simplify $\mathbb{I}_2$. Evaluating the double integral in y_1 and y_2, we have

$$\begin{aligned}&\int_{-\infty}^{\infty} e^{i(\phi_2-i)y_2} \int_{y_2}^{\infty} e^{i\phi_1 y_1}\,\mathrm{d}y_1\,\mathrm{d}y_2 \\ &= \lim_{b_1,d_1,d_2\to\infty} \frac{1}{i\phi_1}\left\{\frac{e^{i\phi_1 b_1 + i(\phi_2-i)d_1} - e^{i\phi_1 b_1 - i(\phi_2-i)d_2}}{i(\phi_2-i)} - \frac{e^{i\phi_1 d_1 + i(\phi_2-i)d_1} - e^{i\phi_1 d_2 - i(\phi_2-i)d_2}}{i\phi_1 + i(\phi_2-i)}\right\}.\end{aligned}$$

If we let $\xi_2 = \phi_2 - i$, then

$$\begin{aligned}\mathbb{I}_2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(-\phi_1,-\xi_2-i) \lim_{b_1,d_1,d_2\to\infty}\left\{\frac{e^{i\phi_1 b_1 + i\xi_2 d_1} - e^{i\phi_1 b_1 - i\xi_2 d_2}}{i\phi_1 i\xi_2}\right. \\ &\quad \left. - \frac{e^{i\phi_1 d_1 + i\xi_2 d_1} - e^{-i\phi_1 d_2 - i\xi_2 d_2}}{i\phi_1(i\phi_1+i\xi_2)}\right\}\mathrm{d}\phi_1\,\mathrm{d}\xi_2 \\ &= \lim_{b_1,d_1,d_2\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(-\phi_1,-\xi_2-i)\left[\frac{e^{i\phi_1 b_1 + i\xi_2 d_1} - e^{i\phi_1 b_1 - i\xi_2 d_2}}{i\phi_2 i\xi_2}\right]\mathrm{d}\phi_1\,\mathrm{d}\xi_2 \\ &\quad - \lim_{b_1,d_1,d_2\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(-\phi_1,-\xi_2-i)\left[\frac{e^{i\phi_1 d_1 + i\xi_2 d_1} - e^{-i\phi_1 d_2 - i\xi_2 d_2}}{i\phi_1(i\phi_1+i\xi_2)}\right]\mathrm{d}\phi_1\,\mathrm{d}\xi_2.\end{aligned}$$

Starting here, we return to the original variable ϕ_2 in place of ξ_2. From the last expression, we can write $\mathbb{I}_2$ as

$$\mathbb{I}_2 = \lim_{b_1,d_1\to\infty} \mathbb{I}_{21}(b_1,d_1) - \lim_{b_1,d_2\to\infty} \mathbb{I}_{22}(b_1,d_2) - \lim_{d_1,d_2\to\infty} \mathbb{I}_{23}(d_1,d_2),$$

where

$$\begin{aligned}\mathbb{I}_{21}(b_1,d_1) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{f_2(-\phi_1,-\phi_2-i)e^{i\phi_1 b_1 + i\phi_2 d_1}}{i\phi_1 i\phi_2}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2, \\ \mathbb{I}_{22}(b_1,d_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{f_2(-\phi_1,-\phi_2-i)e^{i\phi_1 b_1 - i\phi_2 d_2}}{i\phi_1 i\phi_2}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2, \\ \mathbb{I}_{23}(d_1,d_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(-\phi_1,-\phi_2-i)\left[\frac{e^{i\phi_1 d_1 + i\phi_2 d_1} - e^{-i\phi_1 d_2 - i\phi_2 d_2}}{i\phi_1(i\phi_1+i\phi_2)}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2.\end{aligned}$$

The integrals in $\mathbb{I}_{21}$ and $\mathbb{I}_{22}$ will also be evaluated using Lemma 3.12.

First, we consider the function $f_2(\phi_1,\phi_2-i)$. As before, if we set $z_1 = S_1 e^{\Upsilon_{1,m}-\tilde{\lambda}_1\tilde{\kappa}_1\tau}$ and $z_2 = S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2\tau}$, we find that

$$\begin{aligned}f_2(\phi_1,\phi_2-i) &= f_2(\tau,z_1,z_2,v;\phi_1,\phi_2-i) \\ &= \exp\left\{i\phi_1\ln z_1 + i(\phi_2-i)\ln z_2 + B_1(\tau,\phi_1,\phi_2-i) + B_2(\tau,\phi_1,\phi_2-i)v\right\} \\ &= z_2\exp\left\{i\phi_1\ln z_1 + i\phi_2\ln z_2 + B_1(\tau,\phi_1,\phi_2-i) + B_2(\tau,\phi_1,\phi_2-i)v\right\}.\end{aligned}$$

Note that

$$B_1(\tau, \phi_1, \phi_2 - i) = (r - q_2)\tau + B_{12}(\tau, \phi_1, \phi_2),$$

where B_{12} is given in Definition 3.13. We also observe that $B_2(\tau, \phi_1, \phi_2 - i) = B_{22}(\tau, \phi_1, \phi_2)$, which is defined in Definition 3.13. With these representations, we can thus write

$$f_2(\phi_1, \phi_2 - i) = z_2 e^{(r-q_2)\tau} f_{22}(\phi_1, \phi_2),$$

where f_{22} has also been defined in Definition 3.13.

By splitting the region of integration and using the Δ-notation, we find that

$$\begin{aligned}\mathbb{I}_{21} &= \int_0^\infty \int_0^\infty \Delta_{\phi_1}\Delta_{\phi_2} \left[\frac{f_2(\phi_1, \phi_2 - i)e^{-i\phi_1 b_1 - i\phi_2 d_1}}{i\phi_1 i\phi_2}\right] \mathrm{d}\phi_1\, \mathrm{d}\phi_2 \\ &= z_2 e^{(r-q_2)\tau} \int_0^\infty \int_0^\infty \Delta_{\phi_1}\Delta_{\phi_2} \left[\frac{f_{22}(\phi_1, \phi_2)e^{-i\phi_1 b_1 - i\phi_2 d_1}}{i\phi_1 i\phi_2}\right] \mathrm{d}\phi_1\, \mathrm{d}\phi_2.\end{aligned}$$

Using Lemma 3.12, we find that

$$\mathbb{I}_{21} = \pi^2 z_2 e^{(r-q_2)\tau} \left\{4F_3(b_1, d_1) - 2[F_{31}(b_1) + F_{32}(d_1)] + 1\right\},$$

for some joint cumulative distribution function $F_3(x_1, x_2)$, with marginal distribution functions $F_{31}(x_1)$ and $F_{32}(x_2)$, whose corresponding characteristic function is $f_{22}(\phi_1, \phi_2)$. Taking the limit as $b_1, d_1 \to \infty$, it follows that

$$\lim_{b_1, d_1 \to \infty} \mathbb{I}_{21}(b_1, d_1) = \pi^2 z_2 e^{(r-q_2)\tau}.$$

The same calculations can be followed to evaluate $\mathbb{I}_{22}$. First note that

$$f_2(\phi_1, -\phi_2 - i) = z_2 e^{(r-q_2)\tau} f_{22}(\phi_1, -\phi_2) = z_2 e^{(r-q_2)\tau} h_2(\phi_1, \phi_2),$$

where we set $h_2(\phi_1, \phi_2) = f_{22}(\phi_1, -\phi_2)$. With the Δ-notation, we can write

$$\begin{aligned}\mathbb{I}_{22} &= -\int_0^\infty \int_0^\infty \Delta_{\phi_1}\Delta_{\phi_2} \left[\frac{f_2(\phi_1, -\phi_2 - i)e^{-i\phi_1 b_1 - i\phi_2 d_2}}{i\phi_1 i\phi_2}\right] \mathrm{d}\phi_1\, \mathrm{d}\phi_2 \\ &= -z_2 e^{(r-q_2)\tau} \int_0^\infty \int_0^\infty \Delta_{\phi_1}\Delta_{\phi_2} \left[\frac{h_2(\phi_1, \phi_2)e^{-i\phi_1 b_1 - i\phi_2 d_2}}{i\phi_1 i\phi_2}\right] \mathrm{d}\phi_1\, \mathrm{d}\phi_2.\end{aligned}$$

Lemma 3.12 allows us to write

$$\mathbb{I}_{22} = -\pi^2 z_2 e^{(r-q_2)\tau} \left\{4F_4(b_1, d_2) - 2[F_{41}(b_1) + F_{42}(d_2)] + 1\right\},$$

for some joint cumulative distribution function $F_4(x_1, x_2)$, with marginal distributions $F_{41}(x_1)$ and $F_{42}(x_2)$, whose characteristic function is $h_2(\phi_1, \phi_2)$. Taking the limit as $b_1, d_2 \to \infty$, we find that

$$\lim_{b_1, d_2 \to \infty} \mathbb{I}_{22} = -\pi^2 z_2 e^{(r-q_2)\tau}.$$

Lastly, we can write $\mathbb{I}_{23}$ in terms of f_{22} as

$$\mathbb{I}_{23}(d_1, d_2) = z_2 e^{(r-q_2)\tau} \iint_{\mathbb{R}^2} \frac{f_{22}(-\phi_1, -\phi_2)\left[e^{i\phi_1 d_1 + i\phi_2 d_1} - e^{-i\phi_1 d_2 - i\phi_2 d_2}\right]}{i\phi_1(i\phi_1 + i\phi_2)} \mathrm{d}\phi_1\, \mathrm{d}\phi_2.$$

As before, we will leave $\mathbb{I}_{23}$ as an integral in the Fourier transform variables.

With the preceding calculations on $\mathbb{I}_{21}$, $\mathbb{I}_{22}$, and $\mathbb{I}_{23}$, it follows therefore that

$$\begin{aligned}\mathbb{I}_2 &= 2\pi^2 z_2 e^{(r-q_2)\tau} \\ &\quad - z_2 e^{(r-q_2)\tau} \lim_{d_1,d_2\to\infty} \iint_{\mathbb{R}^2} \frac{f_{22}(-\phi_1,-\phi_2)\left[e^{i\phi_1 d_1 + i\phi_2 d_1} - e^{-i\phi_1 d_2 - i\phi_2 d_2}\right]}{i\phi_1(i\phi_1 + i\phi_2)}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\ &= S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2\tau} e^{(r-q_2)\tau} \\ &\quad \times \left[2\pi^2 - \lim_{d_1,d_2\to\infty} \iint_{\mathbb{R}^2} \frac{f_{22}(-\phi_1,-\phi_2)\left[e^{i\phi_1 d_1 + i\phi_2 d_1} - e^{-i\phi_1 d_2 - i\phi_2 d_2}\right]}{i\phi_1(i\phi_1 + i\phi_2)}\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2\right].\end{aligned} \tag{B.24}$$

Rewriting the European exchange option price: Given $\mathbb{I}$ in equation (B.22), we can now write equation (3.48) as

$$\begin{aligned}C^E(\tau,S_1,S_2,v) &= \frac{e^{-r\tau}}{4\pi^2}\sum_{m,n=0}^{\infty}\frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau}\mathbb{E}_{\mathbb{Q}}\left(\mathbb{I}\right) \\ &= \sum_{m,n=0}^{\infty}\frac{(\tilde{\lambda}_1\tau)^m(\tilde{\lambda}_2\tau)^n}{m!n!}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)\tau}\mathbb{E}_{\mathbb{Q}}\left[\frac{e^{-r\tau}}{4\pi^2}\mathbb{I}_1 - \frac{e^{-r\tau}}{4\pi^2}\mathbb{I}_2\right].\end{aligned}$$

The functions P_1 and P_2 in equation (3.53) arise as the expressions enclosed in square brackets when the terms $e^{-r\tau}\mathbb{I}_1/4\pi^2$ and $e^{-r\tau}\mathbb{I}_2/4\pi^2$ are multiplied out given the expressions for $\mathbb{I}_1$ and $\mathbb{I}_2$ in equations (B.23) and (B.24), respectively. These final operations result to the expression for $C^E(\tau,S_1,S_2,v)$ seen in the proposition.

B.7 Proof of Proposition 3.19

Observe that the indicator function $\mathbf{1}_{\mathcal{A}_t}$ in the inhomogeneous term Ξ (see equation (3.62)) is nonzero when $y_1 \geq \ln B(u,b) + y_2$, where $y_2 \in \mathbb{R}$. Thus, equation (3.66) can be written as

$$\begin{aligned}C^P(\tau,S_1,S_2,v) = \int_0^\tau e^{-r(\tau-u)}\int_0^\infty\int_{-\infty}^\infty\int_{\ln B(u,b)+y_2}^\infty \Bigg\{ & q_1 e^{y_1} - q_2 e^{y_2} \\ & - \tilde{\lambda}_1\int_{-\infty}^{\ln B(u,b)+y_2-y_1}\left[C^A\left(u,e^{y_1+y},e^{y_2},b\right) - \left(e^{y_1+y}-e^{y_2}\right)\right]g_1(y)\,\mathrm{d}y \\ & - \tilde{\lambda}_2\int_{-\ln B(u,b)-y_2+y_1}^{\infty}\left[C^A\left(u,e^{y_1},e^{y_2+y},b\right) - \left(e^{y_1}-e^{y_2+y}\right)\right]g_2(y)\,\mathrm{d}y\Bigg\} \\ & \times G(\tau-u,y_1,y_2,b;S_1,S_2,v)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}b\,\mathrm{d}u.\end{aligned}$$

From here, we deduce that C^P can be decomposed in terms of C_D^P, $C_{J_1}^P$, and $C_{J_2}^P$ as given in equation (3.73), where we identify

$$\begin{aligned}
C_D^P(\tau, S_1, S_2, v) &= \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \int_{-\infty}^\infty \int_{\ln B(u,b)+y_2}^\infty (q_1 e^{y_1} - q_2 e^{y_2}) \\
&\quad \times G(\tau-u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u, \\
C_{J_1}^P(\tau, S_1, S_2, v) &= \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \int_{-\infty}^\infty \int_{\ln B(u,b)+y_2}^\infty \int_{-\infty}^{\ln B(u,b)+y_2-y_1} \\
&\quad \left[C^A\left(u, e^{y_1}, e^{y_2+y}, b\right) - \left(e^{y_1} - e^{y_2+y}\right)\right] g_2(y) \\
&\quad \times G(\tau-u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u, \\
C_{J_2}^P(\tau, S_1, S_2, v) &= \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \int_{-\infty}^\infty \int_{\ln B(u,b)+y_2}^\infty \int_{-\ln B(u,b)-y_2+y_1}^\infty \\
&\quad \left[C^A\left(u, e^{y_1}, e^{y_2+y}, b\right) - \left(e^{y_1} - e^{y_2+y}\right)\right] g_2(y) \\
&\quad \times G(\tau-u, y_1, y_2, b; S_1, S_2, v)\, \mathrm{d}y\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u.
\end{aligned} \tag{B.25}$$

The proof that follows will be presented in three parts, each of which will focus on evaluating the three components enumerated above.

Evaluating C_D^P: Substituting the expression for the function G in equation (3.72) to the expression for C_D^P given above, we have

$$\begin{aligned}
C_D^P(\tau, S_1, S_2, v) = \sum_{m,n=0}^\infty \frac{\tilde{\lambda}_1^m \tilde{\lambda}_2^n}{m!n!} \mathbb{E}_{\mathbb{Q}} \Bigg[\frac{1}{4\pi^2} \int_0^\tau \int_0^\infty (\tau-u)^{m+n} e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(\tau-u)} e^{-r(\tau-u)} \\
\times \int_{-\infty}^\infty \int_{\ln B(u,b)+y_2}^\infty (q_1 e^{y_1} - q_2 e^{y_2}) \iint_{\mathbb{R}^2} e^{i\phi_1 y_1 + i\phi_2 y_2} \\
\times h(-\phi_1, -\phi_2)\, \mathrm{d}\phi_1\, \mathrm{d}\phi_2\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}b\, \mathrm{d}u \Bigg],
\end{aligned}$$

where, for notational brevity, we suppress the other arguments in the function h and write

$$h(-\phi_1, -\phi_2) \equiv h\left(\tau-u, S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1\tilde{\kappa}_1(\tau-u)}, S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}, v; -\phi_1, -\phi_2, b\right).$$

Observe that the innermost four-dimensional integral, which we denote by $\mathcal{I}$, may be written as

$$\begin{aligned}
\mathcal{I} = q_1 \iint_{\mathbb{R}^2} h(-\phi_1, -\phi_2) \int_{-\infty}^\infty e^{i\phi_2 y_2} \int_{\ln B(u,b)+y_2}^\infty e^{i(\phi_1 - i)y_1}\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}\phi_1\, \mathrm{d}\phi_2 \\
- q_2 \iint_{\mathbb{R}^2} h(-\phi_1, -\phi_2) \int_{-\infty}^\infty e^{i(\phi_2 - i)y_2} \int_{\ln B(u,b)+y_2}^\infty e^{i\phi_1 y_1}\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}\phi_1\, \mathrm{d}\phi_2.
\end{aligned}$$

Denote by $\mathcal{I}_1$ and $\mathcal{I}_2$ the first and second terms, respectively, of $\mathcal{I}$.

Evaluating $\mathcal{I}_1$: Evaluating the innermost two-dimensional integrals with respect to y_1 and y_2, we find that

$$\begin{aligned}
\mathcal{I}_1 &= q_1 \iint_{\mathbb{R}^2} h(-\phi_1, -\phi_2) \lim_{b_1, c_1, c_2 \to \infty} \Bigg[\frac{e^{i(\phi_1 - i)b_1 + i\phi_2 c_1} - e^{i(\phi_1 - i)b_1 - i\phi_2 c_2}}{i(\phi_1 - i)i\phi_2} \\
&\quad - e^{i(\phi_1 - i)\ln B(u,b)} \frac{e^{i(\phi_1 - i)c_1 + i\phi_2 c_1} - e^{-i(\phi_1 - i)c_2 - i\phi_2 c_2}}{i(\phi_1 - i)[i(\phi_1 - i) + i\phi_2]} \Bigg] \mathrm{d}\phi_1\, \mathrm{d}\phi_2 \\
&= q_1 \iint_{\mathbb{R}^2} h(-\phi_1 - i, -\phi_2) \lim_{b_1, c_1, c_2 \to \infty} \Bigg[\frac{e^{i\phi_1 b_1 + i\phi_2 c_1} - e^{i\phi_1 b_1 - i\phi_2 c_2}}{i\phi_1 i\phi_2}
\end{aligned}$$

$$- e^{i\phi_1 \ln B(u,b)} \frac{e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2}}{i\phi_1(i\phi_1 + i\phi_2)} \Bigg] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.$$

Additionally, define $\mathcal{I}_{11}$, $\mathcal{I}_{12}$, and $\mathcal{I}_{13}$ as

$$\begin{aligned}
\mathcal{I}_{11}(b_1, c_1) &= \iint_{\mathbb{R}^2} \frac{h(-\phi_1 - i, -\phi_2) e^{i\phi_1 b_1 + i\phi_2 c_1}}{i\phi_1 i\phi_2} \, \mathrm{d}\phi_1 \, \mathrm{d}\phi_2, \\
\mathcal{I}_{12}(b_1, c_2) &= \iint_{\mathbb{R}^2} \frac{h(-\phi_1 - i, -\phi_2) e^{i\phi_1 b_1 - i\phi_2 c_2}}{i\phi_1 i\phi_2} \, \mathrm{d}\phi_1 \, \mathrm{d}\phi_2, \\
\mathcal{I}_{13}(c_1, c_2) &= \iint_{\mathbb{R}^2} e^{i\phi_1 \ln B(u,b)} h(-\phi_1 - i, -\phi_2) \left[\frac{e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2}}{i\phi_1(i\phi_1 + i\phi_2)} \right] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.
\end{aligned}$$

Then $\mathcal{I}$ may be written as

$$\mathcal{I} = q_1 \left[\lim_{b_1, c_1 \to \infty} \mathcal{I}_{11}(b_1, c_1) - \lim_{b_1, c_2 \to \infty} \mathcal{I}_{12}(b_1, c_2) - \lim_{c_1, c_2 \to \infty} \mathcal{I}_{13}(c_1, c_2) \right].$$

Using the Δ-notation in Lemma 3.12, we may write $\mathcal{I}_{11}$ as

$$\mathcal{I}_{11}(b_1, c_1) = \iint_{\mathbb{R}_+^2} \Delta_{\phi_1} \Delta_{\phi_2} \left[\frac{h(\phi_1 - i, \phi_2) e^{-i\phi_1 b_1 - i\phi_2 c_1}}{i\phi_1 i\phi_2} \right] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.$$

We note however that

$$h(\tau, z_1, z_2, v; \phi_1 - i, \phi_2, b) = z_1 e^{(r - q_1)\tau} h_1(\tau, z_1, z_2, v; \phi_1, \phi_2, b),$$

where h_1 is given by equation (3.70), so $\mathcal{I}_{11}$ may be expressed as

$$\begin{aligned}
\mathcal{I}_{11}(b_1, c_1) = S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1 \tilde{\kappa}_1 (\tau - u)} e^{(r - q_1)(\tau - u)} \\
\times \iint_{\mathbb{R}_+^2} \Delta_{\phi_1} \Delta_{\phi_2} \left[\frac{h_1(\phi_1, \phi_2) e^{-i\phi_1 b_1 - i\phi_2 c_1}}{i\phi_1 i\phi_2} \right] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.
\end{aligned}$$

Here, we argue in the same manner as we did in evaluating $\mathbb{I}_{11}$ in the proof of Proposition 3.14. That is, using Lemma 3.12, we find that

$$\lim_{b_1, c_1 \to \infty} \mathcal{I}_{11}(b_1, c_1) = \pi^2 S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1 \tilde{\kappa}_1 (\tau - u)} e^{(r - q_1)(\tau - u)}.$$

Similarly, the quantity $\mathcal{I}_{12}$ can also be written using the Δ-notation as

$$\begin{aligned}
\mathcal{I}_{12}(b_1, c_2) &= - \iint_{\mathbb{R}_+^2} \Delta_{\phi_1} \Delta_{\phi_2} \left[\frac{h(\phi_1 - i, \phi_2) e^{-i\phi_1 b_1 - i\phi_2 c_2}}{i\phi_1 i\phi_2} \right] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2 \\
&= -S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1 \tilde{\kappa}_1 (\tau - u)} e^{(r - q_1)(\tau - u)} \\
&\quad \times \iint_{\mathbb{R}_+^2} \Delta_{\phi_1} \Delta_{\phi_2} \left[\frac{h_1(\phi_1, \phi_2) e^{-i\phi_1 b_1 - i\phi_2 c_2}}{i\phi_1 i\phi_2} \right] \mathrm{d}\phi_1 \, \mathrm{d}\phi_2.
\end{aligned}$$

Using Lemma 3.12 again, we find that

$$\lim_{b_1, c_2 \to \infty} \mathcal{I}_{12}(b_1, c_2) = -\pi^2 S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1 \tilde{\kappa}_1 (\tau - u)} e^{(r - q_1)(\tau - u)}.$$

Finally, we note that $\mathcal{I}_{13}$ can be written in terms of h_1 as

$$\mathcal{I}_{13}(c_1, c_2) = S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1 \tilde{\kappa}_1 (\tau - u)} e^{(r - q_1)(\tau - u)}$$

$$\times \iint_{\mathbb{R}^2} e^{i\phi_1 \ln B(u,b)} h_1(-\phi_1,-\phi_2) \left[\frac{e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2}}{i\phi_1(i\phi_1 + i\phi_2)}\right] \mathrm{d}\phi_1\,\mathrm{d}\phi_2.$$

Similar to $\mathbb{I}_{13}$ in the proof of Proposition 3.14, we cannot use Lemma 3.12 to simplify $\mathcal{I}_{13}$, so we leave it as it is.

Given these calculations, we can thus write $\mathcal{I}_1$ as

$$\mathcal{I}_1 = q_1 S_1 e^{\Upsilon_{1,m} - \tilde{\lambda}_1 \tilde{\kappa}_1 (\tau - u)} e^{(r-q_1)(\tau-u)} \left[2\pi^2 - \lim_{c_1,c_2\to\infty} \iint_{\mathbb{R}^2} e^{i\phi_1 \ln B(u,b)} \times h_1(-\phi_1,-\phi_2) \left(\frac{e^{i\phi_1 c_1 + i\phi_2 c_1} - e^{-i\phi_1 c_2 - i\phi_2 c_2}}{i\phi_1(i\phi_1 + i\phi_2)}\right) \mathrm{d}\phi_1\,\mathrm{d}\phi_2\right].$$

Evaluating $\mathcal{I}_2$: We first note that the innermost integrals with respect to y_1 and y_2 can be evaluated, giving us

$$\begin{aligned}\mathcal{I}_2 &= q_2 \iint_{\mathbb{R}^2} h(-\phi_1,-\phi_2) \lim_{b_1,d_1,d_2\to\infty} \left[\frac{e^{i\phi_1 b_2 + i(\phi_2 - i)d_1} - e^{i\phi_1 b_2 - i(\phi_2 - i)d_2}}{i\phi_1 + i(\phi_2 - i)}\right. \\ &\quad \left. - e^{i\phi_1 \ln B(u,b)} \frac{e^{i\phi_1 d_1 + i(\phi_2 - i)d_1} - e^{-i\phi_1 d_2 - i(\phi_2 - i)d_2}}{i\phi_1[i\phi_1 + i(\phi_2 - i)]}\right] \mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\ &= q_2 \iint_{\mathbb{R}^2} h(-\phi_1,-\phi_2 - i) \lim_{b_1,d_1,d_2\to\infty} \left[\frac{e^{i\phi_1 b_2 + i\phi_2 d_1} - e^{i\phi_1 b_2 - i\phi_2 d_2}}{i\phi_1 + i\phi_2}\right. \\ &\quad \left. - e^{i\phi_1 \ln B(u,b)} \frac{e^{i\phi_1 d_1 + i\phi_2 d_1} - e^{-i\phi_1 d_2 - i\phi_2 d_2}}{i\phi_1(i\phi_1 + i\phi_2)}\right] \mathrm{d}\phi_1\,\mathrm{d}\phi_2.\end{aligned}$$

If we consider $\mathcal{I}_{21}$, $\mathcal{I}_{22}$, and $\mathcal{I}_{23}$ given by

$$\begin{aligned}\mathcal{I}_{21}(b_2,d_1) &= \iint_{\mathbb{R}^2} \frac{h(-\phi_1,-\phi_2 - i) e^{i\phi_1 b_2 + i\phi_2 d_1}}{i\phi_1 i\phi_2} \mathrm{d}\phi_1\,\mathrm{d}\phi_2, \\ \mathcal{I}_{22}(b_2,d_2) &= \iint_{\mathbb{R}^2} \frac{h(-\phi_1,-\phi_2 - i) e^{i\phi_1 b_2 - i\phi_2 d_2}}{i\phi_1 i\phi_2} \mathrm{d}\phi_1\,\mathrm{d}\phi_2, \\ \mathcal{I}_{23}(d_1,d_2) &= \iint_{\mathbb{R}^2} e^{i\phi_1 \ln B(u,b)} h(-\phi_1,-\phi_2 - i) \left[\frac{e^{i\phi_1 d_1 + i\phi_2 d_1} - e^{-i\phi_1 d_2 - i\phi_2 d_2}}{i\phi_1(i\phi_2 + i\phi_2)}\right] \mathrm{d}\phi_1\,\mathrm{d}\phi_2,\end{aligned}$$

then we can write $\mathcal{I}_2$ as

$$\mathcal{I}_2 = q_2 \left[\lim_{b_2,d_1\to\infty} \mathcal{I}_{21}(b_2,d_1) - \lim_{b_2,d_2\to\infty} \mathcal{I}_{22}(b_2,d_2) - \lim_{d_1,d_2\to\infty} \mathcal{I}_{23}(d_1,d_2)\right].$$

Using the Δ-notation, we may write $\mathcal{I}_{21}$ as

$$\mathcal{I}_{21}(b_2,d_1) = \iint_{\mathbb{R}^2_+} \Delta_{\phi_1}\Delta_{\phi_2} \left[\frac{h(\phi_1,\phi_2 - i) e^{-i\phi_1 b_2 - i\phi_2 d_1}}{i\phi_1 i\phi_2}\right] \mathrm{d}\phi_1\,\mathrm{d}\phi_2.$$

But since we have

$$h_1(\tau,z_1,z_2,v;\phi_1,\phi_2 - i,b) = z_2 e^{(r-q_2)\tau} h_2(\tau,z_1,z_2,v;\phi_1,\phi_2,b),$$

where h_2 is defined in equation 3.70, it follows that

$$\mathcal{I}_{21}(b_2,d_1) = S_2 e^{\Upsilon_{2,n} - \tilde{\lambda}_2 \tilde{\kappa}_2 (\tau - u)} e^{(r-q_2)(\tau-u)}$$

$$\times \iint_{\mathbb{R}_+^2} \Delta_{\phi_1}\Delta_{\phi_2}\left[\frac{h_2(\phi_1,\phi_2)e^{-i\phi_1 b_2 - i\phi_2 d_1}}{i\phi_1 i\phi_2}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2.$$

Invoking Lemma 3.12, we find that

$$\lim_{b_2,d_1\to\infty}\mathcal{I}_{21}(b_2,d_1) = \pi^2 S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}e^{(r-q_2)(\tau-u)}.$$

A similar calculation can be done for $\mathcal{I}_{22}$. With the Δ-notation and the function h_2, we write

$$\begin{aligned}\mathcal{I}_{22}(b_2,d_2) &= -\iint_{\mathbb{R}_+^2}\Delta_{\phi_1}\Delta_{\phi_2}\left[\frac{h(\phi_1,\phi_2-i)e^{-i\phi_1 b_2 - i\phi_2 d_2}}{i\phi_1 i\phi_2}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2\\ &= -S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}e^{(r-q_2)(\tau-u)}\\ &\quad\times\iint_{\mathbb{R}_+^2}\Delta_{\phi_1}\Delta_{\phi_2}\left[\frac{h_2(\phi_1,\phi_2)e^{-i\phi_1 b_2 - i\phi_2 d_2}}{i\phi_1 i\phi_2}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2,\end{aligned}$$

and in light of Lemma 3.12, we have

$$\lim_{b_2,d_2\to\infty}\mathcal{I}_{22}(b_2,d_2) = -\pi^2 S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}e^{(r-q_2)(\tau-u)}.$$

Lastly, we note that $\mathcal{I}_{23}(d_1,d_2)$ can be expressed as

$$\begin{aligned}\mathcal{I}_{23}(d_1,d_2) &= S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}e^{(r-q_2)(\tau-u)}\\ &\quad\times\iint_{\mathbb{R}^2}e^{i\phi_1\ln B(u,b)}h_2(-\phi_1,-\phi_2)\left[\frac{e^{i\phi_1 d_1+i\phi_2 d_1}-e^{-i\phi_1 d_2-i\phi_2 d_2}}{i\phi_1(i\phi_1+i\phi_2)}\right]\mathrm{d}\phi_1\,\mathrm{d}\phi_2.\end{aligned}$$

It therefore follows that

$$\begin{aligned}\mathcal{I}_2 = q_2 S_2 e^{\Upsilon_{2,n}-\tilde{\lambda}_2\tilde{\kappa}_2(\tau-u)}e^{(r-q_2)(\tau-u)}&\left[2\pi^2 - \lim_{d_1,d_2\to\infty}\iint_{\mathbb{R}^2}e^{i\phi_1\ln B(u,b)}\right.\\ &\left.\times\, h_2(-\phi_1,-\phi_2)\left(\frac{e^{i\phi_1 d_1+i\phi_2 d_1}-e^{-i\phi_1 d_2-i\phi_2 d_2}}{i\phi_1(i\phi_1+i\phi_2)}\right)\mathrm{d}\phi_1\,\mathrm{d}\phi_2\right].\end{aligned}$$

At this point, we note that C_D^P can be written in terms of $\mathcal{I}_1$ and $\mathcal{I}_2$ as

$$C_D^P(\tau,S_1,S_2,v) = \sum_{m,n=0}^{\infty}\frac{\tilde{\lambda}_1^m\tilde{\lambda}_2^n}{m!n!}\mathbb{E}_{\mathbb{Q}}\left[\int_0^\tau\int_0^\infty(\tau-u)^{m+n}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(\tau-u)}\frac{1}{4\pi^2}e^{-r(\tau-u)}\left(\mathcal{I}_1-\mathcal{I}_2\right)\right].$$

Substituting into the above equation our derived expressions for $\mathcal{I}_1$ and $\mathcal{I}_2$ and writing the resulting quantity in terms of the functions P_1^D and P_2^D, we obtain the representation for C_D^P as stated in the proposition.

Evaluating $C_{J_1}^P$: In the expression for $C_{J_1}^P$ in equation (B.25), we change the order of integration of the innermost three-dimensional integral from $\mathrm{d}y\,\mathrm{d}y_1\,\mathrm{d}y_2$ to $\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}y$. The original region of integration, which we denote by $\mathcal{R}_{J_1}$, is given by

$$\mathcal{R}_{J_1} = \left\{(y,y_1,y_2) : y \le \ln B(u,b) + y_2 - y_1, \ln B(u,b) + y_2 \le y_1, y_2 \in \mathbb{R}\right\}.$$

This can be rewritten as

$$\mathcal{R}_{J_1} = \left\{(y,y_1,y_2) : \ln B(u,b) + y_2 \le y_1 \le \ln B(u,b) + y_2 - y, y_2 \in \mathbb{R}, y < 0\right\}.$$

(Note that we restrict y to negative values so that the upper bound on y_1 remains greater than

corresponding the lower bound.) Therefore, $C^P_{J_1}$ may be written as

$$C^P_{J_1}(\tau, S_1, S_2, v) = \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \int_{-\infty}^0 g_1(y) \int_{-\infty}^\infty \int_{\ln B(u,b)+y_2}^{\ln B(u,b)+y_2-y} \left[C^A\left(u, e^{y_1+y}, e^{y_2}, b\right) - \left(e^{y_1+y} - e^{y_2}\right)\right] \times G\left(\tau - u, y_1, y_2, b; S_1, S_2, v\right) \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}y \, \mathrm{d}b \, \mathrm{d}u.$$

Substituting in the expression for G in equation (3.72) and writing the result in terms of the function h gives us the representation for $C^P_{J_1}$ stated in the proposition.

Evaluating $C^P_{J_2}$: Similar to what was done above, we also rearrange the order of integration in the innermost three-dimensional integral. The original region of integration, denoted by $\mathcal{R}_{J_2}$, is given by

$$\mathcal{R}_{J_2} = \left\{(y, y_1, y_2) : -\ln B(u,b) - y_2 + y_1 \leq y, \ln B(u,b) + y_2 \leq y_1, y_2 \in \mathbb{R}\right\},$$

and can be reorganized as

$$\mathcal{R}_{J_2} = \left\{(y, y_1, y_2) : \ln B(u,b) + y_2 \leq y_1 \leq \ln B(u,b) + y_2 + y, y_2 \in \mathbb{R}, y > 0\right\}.$$

With this, we may write $C^P_{J_2}$ as

$$C^P_{J_2}(\tau, S_1, S_2, v) = \int_0^\tau e^{-r(\tau-u)} \int_0^\infty \int_0^\infty g_2(y) \int_{-\infty}^\infty \int_{\ln B(u,b)+y_2}^{\ln B(u,b)+y_2+y} \left[C^A\left(u, e^{y_1}, e^{y_2+y}, b\right) - \left(e^{y_1} - e^{y_2+y}\right)\right] \times G\left(\tau - u, y_1, y_2, b; S_1, S_2, v\right) \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}y \, \mathrm{d}b \, \mathrm{d}u.$$

Again, substituting in the expression for G in equation (3.72) and writing the result in terms of the function h gives us the representation for $C^P_{J_2}$ provided in the proposition.

APPENDIX C

Proofs of Selected Results in Chapter 4

C.1 Proof of Proposition 4.15

Using the inversion formula (4.64) and the expression for $\hat{H}(t,\phi,v)$ in equation (4.62), $H(t,x,v)$ is given by

$$H(t,x,v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\phi x}\exp\left\{i\phi x_T - i\phi\Psi(T-t)\right\}h(T-t,\phi,v;v_T)\,\mathrm{d}\phi,$$

where h is the function given by equation (4.66). Using the expression for Ψ in equation (4.54), we have

$$\begin{aligned}H(t,x,v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}&\exp\left\{-i\phi\left[x - x_T - \tilde{\lambda}_1\tilde{\kappa}_1(T-t) - \tilde{\lambda}_2\tilde{\kappa}_2^-(T-t)\right]\right\}\\ &\times\exp\left\{\tilde{\lambda}_1(T-t)[\varphi_1(\phi)-1]\right\}\exp\left\{\tilde{\lambda}_2(T-t)[\varphi_2(-\phi)-1]\right\}\\ &\times h(T-t,\phi,v;v_T)\,\mathrm{d}\phi.\end{aligned}$$

Next we take the power series expansion of $e^{\tilde{\lambda}_1(T-t)[\varphi_1(\phi)-1]}$ and $e^{\tilde{\lambda}_2(T-t)[\varphi_2(-\phi)-1]}$. Note that

$$e^{\tilde{\lambda}_1(T-t)\varphi_1(\phi)} = \sum_{m=0}^{\infty}\frac{(\tilde{\lambda}_1(T-t))^m}{m!}[\varphi_1(\phi)]^m.$$

If $Y_{1,1},\ldots,Y_{1,m}$ are i.i.d. with density function $G_1(y)$, then $[\varphi_1(\phi)]^m = \mathbb{E}_{\hat{\mathbb{Q}}}^{(m)}[e^{-i\phi\Upsilon_{1,m}}]$, where $\Upsilon_{1,m} = \sum_{k=1}^m Y_{1,k}$, since $\varphi_1(\cdot)$ is the characteristic function of each of the $Y_{1,k}$'s. Here, we note that the expectation operator $\mathbb{E}_{\hat{\mathbb{Q}}}^{(m)}$ acts only on $\Upsilon_{1,m}$. It follows that

$$\exp\left\{\tilde{\lambda}_1(T-t)[\varphi_1(\phi)-1]\right\} = e^{-\tilde{\lambda}_1(T-t)}\sum_{m=0}^{\infty}\frac{(\tilde{\lambda}_1(T-t))^m}{m!}\mathbb{E}_{\hat{\mathbb{Q}}}^{(m)}\left[e^{-i\phi\Upsilon_{1,m}}\right].$$

A similar analysis finds that

$$\exp\left\{\tilde{\lambda}_2(T-t)[\varphi_2(-\phi)-1]\right\} = e^{-\tilde{\lambda}_2(T-t)}\sum_{n=0}^{\infty}\frac{(\tilde{\lambda}_2(T-t))^n}{n!}\mathbb{E}_{\hat{\mathbb{Q}}}^{(n)}\left[e^{i\phi\Upsilon_{2,n}}\right],$$

where $\Upsilon_{2,n} = \sum_{l=1}^n Y_{2,l}$, $\{Y_{2,1},\ldots,Y_{2,n}\}$ is a sample of i.i.d. random variables from a population with density function $G_2(y)$, and $\mathbb{E}_{\hat{\mathbb{Q}}}^{(n)}$ is the expectation operator acting only on $\Upsilon_{2,n}$. With these

expressions and from the independence of $\Upsilon_{1,n}$ and $\Upsilon_{2,n}$, we then obtain

$$H(t,x,v) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!} \times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left\{-i\phi\left[x - x_T - \tilde{\lambda}_1\tilde{\kappa}_1(T-t) - \tilde{\lambda}_2\tilde{\kappa}_2^{-}(T-t)\right]\right\}\right. \times \exp\left\{-i\phi\left(\Upsilon_{1,m}-\Upsilon_{2,n}\right)\right\} h(T-t,\phi,v;v_T)\,\mathrm{d}\phi\Bigg]. \tag{C.1}$$

Here, $\mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}$ is the expectation operator acting only on $\Upsilon_{1,m}$ and $\Upsilon_{2,n}$.

Finally, we recall that $x = \ln\tilde{s}$ and $x_T = \ln\tilde{s}_T$. Using these substitutions will give us the original transition density function $Q(t,\tilde{s},v)$,

$$Q(t,\tilde{s},v) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!} \times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left\{-i\phi\left[\ln(\tilde{s}/\tilde{s}_T) - (\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^{-})(T-t)\right]\right\}\right. \times \exp\left\{-i\phi\left(\Upsilon_{1,m}-\Upsilon_{2,n}\right)\right\} h(T-t,\phi,v;v_T)\,\mathrm{d}\phi\Bigg].$$

Combining the exponent in the above expression under one natural logarithm yields the desired result.

C.2 Proof of Lemma 4.16

From $H(T,x_T,v_T;t,x,v)$ given in equation (C.1), we can write

$$\int_0^{\infty} H(T,x_T,v_T;t,\ln\tilde{s},v)\,\mathrm{d}v_T = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!} \times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\phi x_T}\exp\left\{-i\phi\ln\left(\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^{-})(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\right)\right\}\right. \times \int_0^{\infty} h(T-t,\phi,v;v_T)\,\mathrm{d}v_T\,\mathrm{d}\phi\Bigg].$$

As seen here, we first need to evaluate $\int_0^{\infty} h(\tau,\phi,v;v_T)\,\mathrm{d}v_T$. For this calculation, we will integrate equation (4.66) over $(0,\infty)$ in v_T. To this end, we consider the substitutions

$$A = \frac{2\digamma v_T e^{\digamma\tau}}{\omega^2(e^{\digamma\tau}-1)} \qquad \text{and} \qquad y = \frac{2\digamma v}{\omega^2(e^{\digamma\tau}-1)}. \tag{C.2}$$

It follows that

$$\int_0^{\infty} h(\tau,\phi,v;v_T)\,\mathrm{d}v_T = e^{-y}\exp\left\{\frac{(\Theta-\digamma)}{\omega^2}\alpha\tau\right\}\exp\left\{\frac{(\Theta-\digamma)}{\omega^2}\frac{\omega^2(e^{\digamma\tau}-1)}{2\digamma}y\right\}$$

$$\times \int_0^\infty \left(\frac{A}{y}\right)^{\frac{\alpha}{\omega^2}-\frac{1}{2}} \exp\left\{-\left[\frac{(\Theta-\digamma)}{\omega^2}\frac{\omega^2(e^{\digamma\tau}-1)}{2\digamma e^{\digamma\tau}}+1\right]A\right\} I_{\frac{2\alpha}{\omega^2}-1}\left(2\sqrt{Ay}\right)\mathrm{d}A.$$

$$= e^{-y}\exp\left\{\frac{(\Theta-\digamma)}{\omega^2}\alpha\tau\right\}\exp\left\{\frac{(\Theta-\digamma)}{\omega^2}\frac{\omega^2(e^{\digamma\tau}-1)}{2\digamma}y\right\}$$

$$\times \int_0^\infty \left(\frac{A}{y}\right)^{\frac{\alpha}{\omega^2}-\frac{1}{2}} \exp\left\{-\left[\frac{(\Theta-\digamma)}{\omega^2}\frac{\omega^2(e^{\digamma\tau}-1)}{2\digamma e^{\digamma\tau}}+1\right]A\right\}$$

$$\times \sum_{n=0}^\infty \frac{\left(\sqrt{Ay}\right)^{2n+2\alpha/\omega^2-1}}{n!\Gamma\left(n+\frac{2\alpha}{\omega^2}\right)}\mathrm{d}A.$$

Let $h_1(y)$ denote the integral with respect to A. After some algebra, we find that

$$h_1(y) = \sum_{n=0}^\infty y^n \int_0^\infty \exp\left\{-\left[\frac{(\Theta-\digamma)}{\omega^2}\frac{\omega^2(e^{\digamma\tau}-1)}{2\digamma e^{\digamma\tau}}+1\right]A\right\}\frac{A^{n+2\alpha/\omega^2-1}}{n!\Gamma\left(n+\frac{2\alpha}{\omega^2}\right)}\mathrm{d}A.$$

Let $u = \left[\frac{(\Theta-\digamma)}{\omega^2}\frac{\omega^2(e^{\digamma\tau}-1)}{2\digamma e^{\digamma\tau}}+1\right]A$. Implementing this substitution and using the definition of the gamma function, we find that

$$h_1(y) = \sum_{n=0}^\infty \frac{y^n}{n!}\left[\frac{2\digamma e^{\digamma\tau}}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}}\right]^{n+\frac{2\alpha}{\omega^2}},$$

which, upon the use of the Maclaurin series expansion for e^x, simplifies to

$$h_1(y) = \left[\frac{2\digamma e^{\digamma\tau}}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}}\right]^{\frac{2\alpha}{\omega^2}}\exp\left\{\frac{2\digamma e^{\digamma\tau}y}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}}\right\}.$$

Using this last expression and reversing the substitution with y for v (see equation (C.2)), the required integral thus becomes

$$\int_0^\infty h(\tau,\phi,v;v_T)\,\mathrm{d}v_T$$

$$= \exp\left\{\frac{(\Theta-\digamma)}{\omega^2}(v+\alpha\tau)\right\}\exp\left\{-\frac{2\digamma v}{\omega^2(e^{\digamma\tau}-1)}\right\}\left[\frac{2\digamma e^{\digamma\tau}}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}}\right]^{\frac{2\alpha}{\omega^2}}$$

$$\times \exp\left\{\frac{2\digamma e^{\digamma\tau}}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}}\times\frac{2\digamma}{\omega^2(e^{\digamma\tau}-1)}v\right\}$$

$$= \exp\left\{\frac{(\Theta-\digamma)}{\omega^2}v+\frac{(\Theta-\digamma)}{\omega^2}\alpha\tau-\frac{2\digamma}{\omega^2(e^{\digamma\tau}-1)}v\right.$$

$$-\frac{2\alpha}{\omega^2}\ln\left(\frac{2\digamma e^{\digamma\tau}}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}}\right)$$

$$\left.+\frac{2\digamma e^{\digamma\tau}}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}}\times\frac{2\digamma}{\omega^2(e^{\digamma\tau}-1)}v\right\}$$

The integral may be written as

$$\int_0^\infty h(\tau,\phi,v;v_T)\,\mathrm{d}v_T = e^{B(\tau,\phi)+D(\tau,\phi)v},$$

where, by collecting the coefficients of v, $B(\tau,\phi)$ and $D(\tau,\phi)$ are given by

$$B(\tau,\phi) = \frac{\alpha}{\omega^2}\left[(\Theta-\digamma)\tau - 2\ln\left(\frac{2\digamma e^{\digamma\tau}}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}}\right)\right],$$
$$D(\tau,\phi) = \frac{\Theta-\digamma}{\omega^2} - \frac{2\digamma}{\omega^2(e^{\digamma\tau}-1)} + \frac{2\digamma e^{\digamma\tau}}{(\Theta-\digamma)(e^{\digamma\tau}-1)+2\digamma e^{\digamma\tau}} \times \frac{2\digamma}{\omega^2(e^{\digamma\tau}-1)}.$$

With some algebraic manipulations, $B(\tau,\phi)$ and $D(\tau,\phi)$ may be written in terms of χ as shown in equation (4.70) in the statement of the proposition. See Appendix B.5 for details.

Therefore, the integral of the transition density function is now given by

$$\begin{aligned}
&\int_0^\infty H(T,x_T,v_T;t,\ln\tilde{s},v)\,\mathrm{d}v_T \\
&\quad= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!} \\
&\qquad\times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\phi x_T}\exp\left\{-i\phi\ln\left(\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\right)\right\}\right. \\
&\qquad\qquad\left.\times\exp\left\{B(T-t,\phi)+D(T-t,\phi)v\right\}\mathrm{d}\phi\right].
\end{aligned}$$

Equation (4.69) follows by writing the above result in terms of the function $f(\tau,z,v;\phi)$ defined in equation (4.70). Details of calculations similar to the ones presented here can be found in Section 3.7 of this thesis.

C.3 Proof of Proposition 4.17

From equation (4.68) and from the results in Lemma 4.16, $\tilde{V}(t,\tilde{s},v)$ may be written as

$$\begin{aligned}
\tilde{V}(t,\tilde{s},v) = e^{-q_1 T}&\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!} \\
&\times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{(q_1-q_2)T}^{\infty}\left(e^{x_T}-e^{(q_1-q_2)T}\right)e^{i\phi x_T}\right. \\
&\qquad\left.\times f\left(T-t,\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}},v;-\phi\right)\mathrm{d}x_T\,\mathrm{d}\phi\right].
\end{aligned}$$

To compress notation, temporarily set $f(-\phi)\equiv f(T-t,\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}},v;-\phi)$ and $K\equiv(q_1-q_2)T$. Furthermore, let $\mathbb{I}$ denote the double integral inside the above expectation. It follows that

$$\mathbb{I} = \int_{-\infty}^{\infty}\int_{K}^{\infty} e^{i(\phi-i)x_T}f(-\phi)\,\mathrm{d}x_T\,\mathrm{d}\phi - e^{K}\int_{-\infty}^{\infty}\int_{K}^{\infty} e^{i\phi x_T}f(-\phi)\,\mathrm{d}x_T\,\mathrm{d}\phi.$$

Denote by $\mathbb{I}_1$ and $\mathbb{I}_2$ the first and second terms, respectively, of $\mathbb{I}$. Integrating with respect to x_T, we have

$$\mathbb{I}_1 = \int_{-\infty}^{\infty} f(-\phi)\int_{K}^{\infty} e^{i(\phi-i)x_T}\,\mathrm{d}x_T\,\mathrm{d}\phi = \lim_{b\to\infty}\int_{-\infty}^{\infty} f(-\phi)\left[\frac{e^{i(\phi-i)b}-e^{i(\phi-i)K}}{i(\phi-i)}\right]\mathrm{d}\phi.$$

Using the substitution $u = \phi - i$, we obtain

$$\mathbb{I}_1 = \lim_{b\to\infty} \int_{-\infty}^{\infty} f(-u-i) \left[\frac{e^{iub} - e^{iuK}}{iu}\right] \mathrm{d}u,$$

and with $\phi = -u$, we have

$$\begin{aligned}
\mathbb{I}_1 &= \lim_{b\to\infty} \int_{-\infty}^{\infty} f(\phi - i) \left[\frac{e^{-i\phi K} - e^{-i\phi b}}{i\phi}\right] \mathrm{d}\phi \\
&= \int_0^{\infty} \frac{f(\phi - i)e^{-i\phi K} - f(-\phi - i)e^{i\phi K}}{i\phi} \mathrm{d}\phi \\
&\quad - \lim_{b\to\infty} \int_0^{\infty} \frac{f(\phi - i)e^{-i\phi b} - f(-\phi - i)e^{i\phi b}}{i\phi} \mathrm{d}\phi,
\end{aligned}$$

where the last equality is a result of splitting the domain of integration on ϕ. Observe that

$$\begin{aligned}
f(\phi - i) &= f(T-t, \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}}, v; \phi - i) \\
&= \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}} f_1\left(T-t, \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}}, v; \phi\right),
\end{aligned}$$

where f_1 and its components are defined in equation (4.73). Thus, we may write

$$\begin{aligned}
\mathbb{I}_1 &= \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}} \\
&\quad \times \left[\int_0^{\infty} \frac{f_1(\phi)e^{-i\phi K} - f_1(-\phi)e^{i\phi K}}{i\phi} \mathrm{d}\phi - \lim_{b\to\infty} \int_0^{\infty} \frac{f_1(\phi)e^{-i\phi b} - f_1(-\phi)e^{i\phi b}}{i\phi} \mathrm{d}\phi\right],
\end{aligned}$$

where $f_1(\phi) \equiv f_1(T-t, \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}}, v; \phi)$.

The limit appearing above may be evaluated using the relationship of cumulative distribution functions and characteristic functions discussed by Shephard (1991b) and Shephard (1991a, Theorem 3).[1] Let $F_1(x)$ be the cumulative density function of a random variable whose mean exists and whose characteristic function is $f_1(\phi)$. Then by Shephard (1991b, Theorem 3), we have

$$F_1(b) = \frac{1}{2} - \frac{1}{2\pi} \int_0^{\infty} \frac{f_1(\phi)e^{-i\phi b} - f_1(-\phi)e^{i\phi b}}{i\phi} \mathrm{d}\phi.$$

Since $F_1(b) \to 1$ as $b \to \infty$ (from the property of cdfs), it follows therefore that

$$\mathbb{I}_1 = \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)} e^{\Upsilon_{1,m} - \Upsilon_{2,n}} \left[\pi + \int_0^{\infty} \frac{f_1(\phi)e^{-i\phi K} - f_1(-\phi)e^{i\phi K}}{i\phi} \mathrm{d}\phi\right].$$

The same calculations apply for evaluating $\mathbb{I}_2$. Evaluating the integral with respect to x_T and splitting the domain of integration in ϕ, we have

$$\begin{aligned}
\mathbb{I}_2 &= \lim_{b\to\infty} \int_{-\infty}^{\infty} f(-\phi) \left[\frac{e^{i\phi b} - e^{i\phi K}}{i\phi}\right] \mathrm{d}\phi \\
&= \lim_{b\to\infty} \int_{-\infty}^{\infty} f(\phi) \left[\frac{e^{-i\phi K} - e^{-i\phi b}}{i\phi}\right] \mathrm{d}\phi \\
&= \int_0^{\infty} \frac{f(\phi)e^{-i\phi K} - f(-\phi)e^{i\phi K}}{i\phi} \mathrm{d}\phi - \lim_{b\to\infty} \int_0^{\infty} \frac{f(\phi)e^{-i\phi b} - f(-\phi)e^{i\phi b}}{i\phi} \mathrm{d}\phi.
\end{aligned}$$

[1] See also Appendix 8 of Chiarella et al. (2010).

Using Shephard (1991b, Theorem 3) to evaluate the limit, we find that

$$\mathbb{I}_2 = \pi + \int_0^\infty \frac{f(\phi)e^{-i\phi K} - f(-\phi)e^{i\phi K}}{i\phi}\,\mathrm{d}\phi.$$

From the above calculations, we may now write $\tilde{V}(t,\tilde{s},v)$ as

$$\begin{aligned}\tilde{V}(t,\tilde{s},v) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{(\tilde{\lambda}_1(T-t))^m(\tilde{\lambda}_2(T-t))^n e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(T-t)}}{m!n!} \\ \times \mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\Bigg[e^{-q_1 T}\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(T-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}} \\ \times\left(\frac{1}{2}+\frac{1}{2\pi}\int_0^\infty \frac{f_1(\phi)e^{-i\phi K}-f_1(-\phi)e^{i\phi K}}{i\phi}\,\mathrm{d}\phi\right) \\ - e^{-q_2 T}\left(\frac{1}{2}+\frac{1}{2\pi}\int_0^\infty \frac{f(\phi)e^{-i\phi K}-f(-\phi)e^{i\phi K}}{i\phi}\,\mathrm{d}\phi\right)\Bigg].\end{aligned}$$

Putting back $(q_1-q_2)T$ in place of K and writing the resulting expression in terms of P_1^E and P_2^E yields the expression presented in the proposition.

C.4 Proof of Proposition 4.18

Some introductory remarks are in order. The representation of the early exercise premium via equation (4.77) is selected such that $\tilde{V}_D^P$, $\tilde{V}_{J_1}^P$, and $\tilde{V}_{J_2}^P$ denote the first, second, and third expectations, respectively, that appear in equation (4.38). These terms may be interpreted as the diffusion and jump components, respectively, of the early exercise premium. In each term, the indicator function $\mathbf{1}(\mathcal{A}(u))$ appears, where $u \in [t,T]$. In terms of x_u and v_u, the event $\mathcal{A}(u)$ may be written as $\mathcal{A}(u) = \{x_u \geq \ln A(u,v_u)\}$. Lastly, our calculations will require the transition density function of (X,v), which is $H(u,x_u,v_u;t,x,v) = H(t,x,v)$ now given by (C.1), replacing T, x_T, and v_T with u, x_u, and v_u, respectively, where $(x_u,v_u) \in \mathbb{R}\times\mathbb{R}_+$ and $t < u \leq T$ (see also equation (4.51) for its original definition).

Evaluating V_D^P: From equation (4.38), V_D^P is given by

$$V_D^P(t,\tilde{s},v) = \mathbb{E}_{\hat{Q}}\left[\int_t^T e^{-q_1 u}\left(q_1\tilde{s}_u - q_2 e^{(q_1-q_2)u}\right)\mathbf{1}(\mathcal{A}(u))\,\middle|\,\tilde{S}(t)=\tilde{s}, v(t)=v\right],$$

and in terms of the transition density function H, we have

$$V_D^P(t,\tilde{s},v) = \int_t^T\int_0^\infty\int_{\ln A(u,v_u)}^\infty e^{-q_1 u}\left(q_1 e^{x_u} - q_2 e^{K(u)}\right)H(u,x_u,v_u;t,\ln\tilde{s},v)\,\mathrm{d}x_u\,\mathrm{d}v_u\,\mathrm{d}t,$$

where $K(u) \equiv (q_1-q_2)u$. Using equation (C.1), we find that

$$\begin{aligned}V_D^P(t,\tilde{s},v) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\tilde{\lambda}_1^m\tilde{\lambda}_2^n}{m!n!}\mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\Bigg[\int_t^T\int_0^\infty (u-t)^{m+n}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(u-t)}e^{-q_1 u} \\ \times\frac{1}{2\pi}\int_{\ln A(u,v_u)}^\infty\left(q_1e^{x_u}-q_2e^{K(u)}\right)\int_{-\infty}^\infty e^{i\phi x_u} \\ \times\exp\left\{-i\phi\ln\left(\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\right)\right\} \\ \times h(u-t,\phi,v;v_u)\,\mathrm{d}\phi\,\mathrm{d}x_u\,\mathrm{d}v_u\,\mathrm{d}u\Bigg].\end{aligned}$$

In terms of the function f_2 introduced in the proposition, we can write

$$V_D^P(t,\tilde{s},v) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\tilde{\lambda}_1^m\tilde{\lambda}_2^n}{m!n!}\mathbb{E}_{\hat{\mathbb{Q}}}^{(m,n)}\Bigg[\int_t^T\int_0^{\infty}(u-t)^{m+n}e^{-(\tilde{\lambda}_1+\tilde{\lambda}_2)(u-t)}e^{-q_1u} \\ \times\frac{1}{2\pi}\int_{-\infty}^{\infty}f_2(-\phi)\int_{\ln A(u,v_u)}^{\infty}\left(q_1e^{x_u}-q_2e^{K(u)}\right)e^{i\phi x_u}\,\mathrm{d}x_u\,\mathrm{d}\phi\,\mathrm{d}v_u\,\mathrm{d}u\Bigg], \tag{C.3}$$

where, for brevity, we set $f_2(-\phi)\equiv f_2(u-t,\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}},y;-\phi,v_u)$.

We next integrate the innermost double integral, which we denote by $\mathbb{I}^*$, using the techniques employed in the proof of Proposition 4.17. To this end, we make a secondary representation $\mathbb{I}^*=q_1\mathbb{I}_1^*-q_2e^{K(u)}\mathbb{I}_2^*$, where

$$\mathbb{I}_1^*=\int_{-\infty}^{\infty}f_2(-\phi)\int_{\ln A(u,v_u)}^{\infty}e^{i(\phi-i)x_u}\,\mathrm{d}x_u\,\mathrm{d}\phi,$$
$$\mathbb{I}_2^*=\int_{-\infty}^{\infty}f_2(-\phi)\int_{\ln A(u,v_u)}^{\infty}e^{i\phi x_u}\,\mathrm{d}x_u\,\mathrm{d}\phi.$$

Evaluating the integral with respect to x_u and splitting the domain of integration on ϕ, $\mathbb{I}_1^*$ can be written as

$$\begin{aligned}
\mathbb{I}_1^* &= \lim_{b\to\infty}\int_{-\infty}^{\infty}f_2(-\phi)\left[\frac{e^{i(\phi-i)b}-e^{i(\phi-i)\ln A(u,v_u)}}{i(\phi-i)}\right]\mathrm{d}\phi\\
&= \lim_{b\to\infty}\int_{-\infty}^{\infty}f_2(\phi-i)\left[\frac{e^{-i\phi\ln A(u,v_u)}-e^{-i\phi b}}{i\phi}\right]\mathrm{d}\phi\\
&= \lim_{b\to\infty}\Bigg\{\int_0^{\infty}f_2(\phi-i)\left[\frac{e^{-i\phi\ln A(u,v_u)}-e^{-i\phi b}}{i\phi}\right]\mathrm{d}\phi\\
&\qquad+\int_0^{\infty}f_2(-\phi-i)\left[\frac{e^{i\phi\ln A(u,v_u)}-e^{i\phi b}}{-i\phi}\right]\mathrm{d}\phi\Bigg\}\\
&= \int_0^{\infty}\frac{f_2(\phi-i)e^{-i\phi\ln A(u,v_u)}-f_2(-\phi-i)e^{i\phi\ln A(u,v_u)}}{i\phi}\,\mathrm{d}\phi\\
&\qquad-\lim_{b\to\infty}\int_0^{\infty}\frac{f_2(\phi-i)e^{-i\phi b}-f_2(-\phi-i)e^{i\phi b}}{i\phi}\,\mathrm{d}\phi.
\end{aligned}$$

Note that

$$\begin{aligned}
f_2(\phi-i) &\equiv f_2\left(u-t,\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}},y;\phi-i,v_u\right)\\
&= \exp\left\{i(\phi-i)\ln\left(\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\right)\right\}\\
&\qquad\times h(u-t,-(\phi-i),v;v_u)\\
&= \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\\
&\qquad\times\exp\left\{i\phi\ln\left(\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}\right)\right\}h(u-t,-(\phi-i),v;v_u)\\
&= \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}}f_{21}\left(u-t,\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}},v;\phi,v_u\right),
\end{aligned}$$

where f_{21} is defined in the proposition. For ease of notation, let $f_{21}(\phi)$ be shorthand for

$$f_{21}(u-t,\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}},v;\phi,v_u).$$

With this representation, $\mathbb{I}_1^*$ can therefore be written as

$$\begin{aligned}\mathbb{I}_1^* &= \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}} \\ &\quad\times\left[\int_0^\infty \frac{f_{21}(\phi)e^{-i\phi\ln A(u,v_u)}-f_{21}(-\phi)e^{i\phi\ln A(u,v_u)}}{i\phi}\,\mathrm{d}\phi\right. \\ &\quad\left.-\lim_{b\to\infty}\int_0^\infty \frac{f_{21}(\phi)e^{-i\phi b}-f_{21}(-\phi)e^{i\phi b}}{i\phi}\,\mathrm{d}\phi\right] \\ &= \tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}} \\ &\quad\times\left[\pi+\int_0^\infty \frac{f_{21}(\phi)e^{-i\phi\ln A(u,v_u)}-f_{21}(-\phi)e^{i\phi\ln A(u,v_u)}}{i\phi}\,\mathrm{d}\phi\right],\end{aligned}$$

where the limit was evaluated using Theorem 3 of Shephard (1991b) (for additional details, refer to the proof of Proposition 4.17).

Similar calculations apply in simplifying $\mathbb{I}_2^*$. Integration with respect to x_u yields

$$\begin{aligned}\mathbb{I}_2^* &= \lim_{b\to\infty}\int_{-\infty}^\infty f_2(-\phi)\left[\frac{e^{i\phi b}-e^{i\phi\ln A(u,v_u)}}{i\phi}\right]\mathrm{d}\phi \\ &= \lim_{b\to\infty}\int_{-\infty}^\infty f_2(\phi)\left[\frac{e^{-i\phi\ln A(u,v_u)}-e^{-i\phi b}}{i\phi}\right]\mathrm{d}\phi \\ &= \lim_{b\to\infty}\left\{\int_0^\infty f_2(\phi)\left[\frac{e^{-i\phi\ln A(u,v_u)}-e^{-i\phi b}}{i\phi}\right]\mathrm{d}\phi+\int_0^\infty f_2(-\phi)\left[\frac{e^{i\phi\ln A(u,v_u)}-e^{i\phi b}}{-i\phi}\right]\mathrm{d}\phi\right\} \\ &= \int_0^\infty \frac{f_2(\phi)e^{-i\phi\ln A(u,v_u)}-f_2(-\phi)e^{i\phi\ln A(u,v_u)}}{i\phi}\,\mathrm{d}\phi \\ &\quad-\lim_{b\to\infty}\int_0^\infty \frac{f_2(\phi)e^{-i\phi b}-f_2(-\phi)e^{i\phi b}}{i\phi}\,\mathrm{d}\phi.\end{aligned}$$

By Theorem 3 of Shephard (1991b), the limit reduces to $-\pi$, giving us

$$\mathbb{I}_2^* = \pi+\int_0^\infty \frac{f_2(\phi)e^{-i\phi\ln A(u,v_u)}-f_2(-\phi)e^{i\phi\ln A(u,v_u)}}{i\phi}\,\mathrm{d}\phi.$$

From our calculations, we now find that $\mathbb{I}^*$ is given by

$$\begin{aligned}\mathbb{I}^* &= q_1\tilde{s}e^{-(\tilde{\lambda}_1\tilde{\kappa}_1+\tilde{\lambda}_2\tilde{\kappa}_2^-)(u-t)}e^{\Upsilon_{1,m}-\Upsilon_{2,n}} \\ &\quad\times\left(\pi+\int_0^\infty \frac{f_{21}(\phi)e^{-i\phi\ln A(u,v_u)}-f_{21}(-\phi)e^{i\phi\ln A(u,v_u)}}{i\phi}\,\mathrm{d}\phi\right) \\ &\quad-q_2e^{(q_1-q_2)u}\left(\pi+\int_0^\infty \frac{f_2(\phi)e^{-i\phi\ln A(u,v_u)}-f_2(-\phi)e^{i\phi\ln A(u,v_u)}}{i\phi}\,\mathrm{d}\phi\right).\end{aligned}$$

Substituting this expression in place of the innermost double integral in equation (C.3), multiplying by $e^{-q_1u}/(2\pi)$, and expressing the resulting expression in terms of P_1^A and P_2^A (as defined in the proposition) yields equation (4.78) for the diffusion component of the early exercise premium.

Evaluating the jump components: From equation (4.38) and with a change of variable from $\tilde{s}_u$ to x_u, $\tilde{V}_{J_1}^P$ can be expressed as

$$\tilde{V}_{J_1}^P(t,\tilde{s},v) = \mathbb{E}_{\hat{\mathbb{Q}}}\left[\int_t^T\int_{-\infty}^{\ln A(u,v_u)-x_u}\left(\tilde{V}^A(u,e^{x_u+y},v_u)-\left(e^{-q_1u}e^{x_u+y}-e^{-q_2u}\right)\right)\right.$$

$$\times G_1(y)\mathbf{1}(\mathcal{A}_u))\,\mathrm{d}y\,\mathrm{d}u\Big|X(t)=\ln\tilde{s},v(t)=v\Big],$$

and in terms of the transition density function H, we have

$$\begin{aligned}\tilde{V}^P_{J_1}(t,\tilde{s},v)&\\ =\int_t^T\int_0^\infty\int_{\ln A(u,v_u)}^\infty\int_{-\infty}^{\ln A(u,v_u)-x_u}&\left(\tilde{V}^A(u,e^{x_u+y},v_u)-\left(e^{-q_1u}e^{x_u+y}-e^{-q_2u}\right)\right)\\ &\times G_1(y)H(u,x_u,v_u;t,\ln\tilde{s},v)\,\mathrm{d}y\,\mathrm{d}x_u\,\mathrm{d}v_u\,\mathrm{d}u.\end{aligned}$$

We switch the order of integration from $\mathrm{d}y\,\mathrm{d}x_u$ to $\mathrm{d}x_u\,\mathrm{d}y$ by noting that the region $\mathcal{R}_{J_1}$ of integration of the innermost double integral can be written equivalently as

$$\begin{aligned}\mathcal{R}_{J_1}&=\left\{(x_u,y):y\le\ln A(u,v_u)-x_u,x_u\ge\ln A(u,v_u)\right\}\\ &=\left\{(x_u,y):\ln A(u,v_u)\le x_u\le\ln A(u,v_u)-y,y\le 0\right\}.\end{aligned}$$

The addition of the condition $y\le 0$ is needed to ensure that the upper bound on x_u remains above the lower bound. It therefore follows that

$$\begin{aligned}\tilde{V}^P_{J_1}(t,\tilde{s},v)=\int_t^T\int_0^\infty\int_{-\infty}^0\int_{\ln A(u,v_u)}^{\ln A(u,v_u)-y}&\left(\tilde{V}^A(u,e^{x_u+y},v_u)-\left(e^{-q_1u}e^{x_u+y}-e^{-q_2u}\right)\right)\\ &\times G_1(y)H(u,x_u,v_u;t,\ln\tilde{s},v)\,\mathrm{d}x_u\,\mathrm{d}y\,\mathrm{d}v_u\,\mathrm{d}u.\end{aligned}$$

Substitution of equation (C.1) for H in the equation above yields equation (4.80) in the proposition.

We follow similar steps in evaluating the second jump component. From equation (4.38) and in terms of x_u, $\tilde{V}^P_{J_1}$ can be written as

$$\begin{aligned}\tilde{V}^P_{J_2}(t,\tilde{s},v)=\mathbb{E}_{\hat{\mathbb{Q}}}\Bigg[\int_t^T\int_{x_u-\ln A(u,v_u)}^\infty&\left(\tilde{V}^A(u,e^{x_u+y},v_u)-\left(e^{-q_1u}e^{x_u+y}-e^{-q_2u}\right)\right)\\ &\times G_2(y)\mathbf{1}(\mathcal{A}_u))\,\mathrm{d}y\,\mathrm{d}u\Big|X(t)=\ln\tilde{s},v(t)=v\Bigg].\end{aligned}$$

We may equivalently write this in terms of the transition density function as

$$\begin{aligned}\tilde{V}^P_{J_2}(t,\tilde{s},v)&\\ =\int_t^T\int_0^\infty\int_{\ln A(u,v_u)}^\infty\int_{x_u-\ln A(u,v_u)}^\infty&\left(\tilde{V}^A(u,e^{x_u+y},v_u)-\left(e^{-q_1u}e^{x_u+y}-e^{-q_2u}\right)\right)\\ &\times G_2(y)H(u,x_u,v_u;t,\ln\tilde{s},v)\,\mathrm{d}y\,\mathrm{d}x_u\,\mathrm{d}v_u\,\mathrm{d}u.\end{aligned}$$

Noting that the region of integration $\mathcal{R}_{J_2}$ for the innermost double integral may be written as

$$\begin{aligned}\mathcal{R}_{J_2}&=\left\{(x_u,y):y\ge x_u-\ln A(u,v_u),x_u\ge\ln A(u,v_u)\right\}\\ &=\left\{(x_u,y):\ln A(u,v_u)\le x_u\le\ln A(u,v_u)+y,y\ge 0\right\},\end{aligned}$$

we may switch the order of integration to obtain

$$\begin{aligned}\tilde{V}^P_{J_2}(t,\tilde{s},v)=\int_t^T\int_0^\infty\int_0^\infty\int_{\ln A(u,v_u)}^{\ln A(u,v_u)+y}&\left(\tilde{V}^A(u,e^{x_u+y},v_u)-\left(e^{-q_1u}e^{x_u+y}-e^{-q_2u}\right)\right)\\ &\times G_2(y)H(u,x_u,v_u;t,\ln\tilde{s},v)\,\mathrm{d}x_u\,\mathrm{d}y\,\mathrm{d}v_u\,\mathrm{d}u.\end{aligned}$$

Replacing H with its expression in equation (C.1) gives us equation (4.81) for the second jump component.

APPENDIX D

The Trapezoidal Rule and Least Squares Monte Carlo

This appendix contains supplementary notes on the trapezoidal rule for the Riccati equation (for the MOL solution in Chapter 5) and on the implementation of the Longstaff and Schwartz (2001) least squares Monte Carlo (LSMC) algorithm.

D.1 Trapezoidal Rule for the Sweep Equations

In this appendix, we provide details on the trapezoidal rule used to solve the sweep equations associated to the Riccati transform. Following Meyer (2015), we shall consider a general two-point boundary problem that resembles system (5.12) of equations for the option price and delta and provide a numerical solution of the forward and reverse sweep equations based on this general notation.

D.1.1 The Riccati Transform and the Sweep Equations

Consider the two-point boundary value problem

$$\begin{aligned} u'(x) &= A(x)u(x) + B(x)v(x) + f(x), \\ v'(x) &= C(x)u(x) + D(x)v(x) + g(x), \end{aligned} \tag{D.1}$$

with boundary conditions

$$u(s) = \Gamma v(s) + \alpha, \qquad G(u(S), v(S)) = 0, \tag{D.2}$$

where Γ and α are constants and G is a known function of two variables, possibly nonlinear. In this discussion, we assume that $s < S$, although this is not necessary. It is necessary however that at least one of the boundary conditions at $x = s$ or $x = S$ is linear (or at most affine).

The Riccati transformation is given by

$$u(x) = R(x)v(x) + w(x), \tag{D.3}$$

where $R(x)$ and $w(x)$ are solutions of well-defined initial value problems

$$R'(x) = B(x) + \big[A(x) - D(x)\big] R(x) - C(x)R^2(x) \tag{D.4}$$

$$w'(x) = \big[A(x) - C(x)R(x)\big] w(x) + f(x) - R(x)g(x). \tag{D.5}$$

Evaluating equation (D.3) at $x = s$ gives

$$u(s) = R(s)v(s) + w(s)$$

and replacing the left-hand side with the first boundary condition (D.2) yields

$$\Gamma v(s) + \alpha = R(s)v(s) + w(s).$$

Comparing coefficients again gives us

$$R(s) = \Gamma, \qquad w(s) = \alpha \tag{D.6}$$

which will be used as initial conditions to complement the system (D.4) and (D.5).

Equation (D.4) with initial condition $R(s) = \Gamma$ is known as the (scalar) Riccati equation. From the theory of initial value problems, this equation has a unique local solution (Meyer, 1973, Theorem 1.3.1). Furthermore, a solution $w(x)$ of equation (D.5) exists as long as $R(x)$ exists. At this point, we now solve both of these equations by integrating them from $x = s$ to $x = S$. Once $R(x)$ has been determined from integrating (D.4), equation (D.5) can be integrated to solve for $w(x)$. Values for $R(S)$ and $w(S)$ consequently follow.

From the boundary condition (D.2) at $s = S$ and the Riccati transformation (D.3), we can then solve the scalar equation

$$G\left(R(S)v + w(S), v\right) = 0$$

for $v(S) = v$, possibly using some root-finding algorithm. Any solution v^* of this equation can be set equal to $v(S)$, which will then serve as a boundary condition for the equation

$$v'(x) = \left[C(x)R(x) + D(x)\right] v(x) + C(x)w(x) + g(x). \tag{D.7}$$

Note that this equation was obtained by using equation (D.3) in the second equation of (D.1). This equation can then be integrated from $x = S$ to $x = s$, giving us $v(x)$. With knowledge of $v(x)$, $R(x)$, and $w(x)$, we can then solve for $u(x)$ using equation (D.3).

D.1.2 The Trapezoidal Rule

Given the nature of the source terms in the MOL approximation of the pricing equation, the integration of the forward sweep equation depends on the numerical solution from the preceding time steps, which are only available at discrete points $\{x_j\}$. Thus, it is not recommended to use an adaptive ODE solver to perform the forward sweep since these integrators require all functions to be defined for all x in the domain. Hence, we will use the trapezoidal rule to avoid having to resort to interpolating values of the solution in between the discrete points. Using the trapezoidal rule also allows us to use a fixed (i.e. time-invariant) but possibly non-uniform grid on x and improve the accuracy of the numerical solution by simply reducing the mesh size.

Consider an N-dimensional system of first-order equations

$$u' = F(x, u), \qquad u(x_0) = u_0 \tag{D.8}$$

to be solved for $x \in [x_0, x_J]$. Define the partition

$$x_0 < x_1 < x_2 < \cdots < x_J,$$

which may not be evenly spaced, but must remain fixed for all time steps. Given a numerical solution $u_{j_0} \equiv u(x_{j_0})$ at the point x_{j_0}, we can find the numerical solution $u_j \equiv u(x_j)$ at a neighboring point x_j via the equation

$$\frac{u_j - u_{j_0}}{x_j - x_{j_0}} = \frac{1}{2}\left[F(x_j, u_j) + F(x_{j_0}, u_{j_0})\right]. \tag{D.9}$$

Note that x_j may be to the left or right of x_{j_0}.

Let $x_0 = s$ and $x_J = S$ be known boundaries. Applying the trapezoidal rule (D.9) to equation

(D.4) gives us

$$\frac{R_j - R_{j_0}}{x_j - x_{j_0}} = \frac{1}{2}\left\{\left[B_j + (A_j - D_j)R_j - C_j R_j^2\right] + \left[B_{j_0} + (A_{j_0} - D_{j_0})R_{j_0} - C_{j_0} R_{j_0}^2\right]\right\},$$

where $R_j \equiv R(x_j)$, $A_j \equiv A(x_j)$, $B_j \equiv B(x_j)$, $C_j \equiv C(x_j)$, and $D_j \equiv D(x_j)$. We may rearrange this expression to obtain a quadratic equation in R_j,

$$\begin{aligned} 0 = \frac{1}{2}C_j R_j^2 + \left[\frac{1}{x_j - x_{j_0}} - \frac{1}{2}(A_j - D_j)\right] R_j + \frac{1}{2}C_{j_0} R_{j_0}^2 \\ - \left[\frac{1}{x_j - x_{j_0}} + \frac{1}{2}(A_{j_0} - D_{j_0})\right] R_{j_0} - \frac{1}{2}(B_j + B_{j_0}). \end{aligned}$$

If we let

$$\begin{aligned} \alpha_j &= \frac{1}{2}C_j, \\ \beta_j &= \frac{1}{x_j - x_{j_0}} - \frac{1}{2}(A_j - D_j), \\ \gamma_j &= \frac{1}{2}C_{j_0} R_{j_0}^2 - \left[\frac{1}{x_j - x_{j_0}} + \frac{1}{2}(A_{j_0} - D_{j_0})\right] R_{j_0} - \frac{1}{2}(B_j + B_{j_0}), \end{aligned}$$

then R_j is given by

$$R_j = \frac{-\beta_j \pm \sqrt{\beta_j^2 - 4\alpha_j\gamma_j}}{2\alpha_j}. \tag{D.10}$$

The choice of which root to take in equation (D.10) is dependent on the sign of $x_j - x_{j_0}$ (i.e. the direction of integration). To see this, we first note that in most applications, $B(x) > 0$ for all x. Evaluating the forward sweep equation (D.4) at $x = x_{j_0}$ and assuming that $R(x_{j_0}) = 0$, we find that $R'(x_{j_0}) = B(x_{j_0}) > 0$. From the defintion of the derivative, we note that

$$0 < R'(x_{j_0}) = \lim_{x \to x_{j_0}} \frac{R(x) - R(x_{j_0})}{x - x_{j_0}} = \lim_{x \to x_{j_0}} \frac{R(x)}{x - x_{j_0}}.$$

Thus, $R(x) > 0$ if $x > x_{j_0}$ and $R(x) < 0$ if $x < x_{j_0}$. Since $B(x)$ is always positive and $R_{j_0} = 0$, we find that $\gamma_j < 0$. If $x_j > x_{j_0}$, then $\beta_j \to +\infty$ as $x_j - x_{j_0} \to 0$, and so $R_j > 0$ only if we take the positive branch of (D.10). Likewise, if $x_j < x_{j_0}$, then $\beta_j \to -\infty$ as $x_j - x_{j_0} \to 0$, and so $R_j < 0$ if the negative branch is chosen. Thus, we have

$$\begin{aligned} R_j &= \frac{-\beta_j + \sqrt{\beta_j^2 - 4\alpha_j\gamma_j}}{2\alpha_j} = \frac{-2\gamma_j}{\beta_j + \sqrt{\beta_j^2 - 4\alpha_j\gamma_j}} \qquad \text{if } x_j > x_{j_0}, \\ R_j &= \frac{-\beta_j - \sqrt{\beta_j^2 - 4\alpha_j\gamma_j}}{2\alpha_j} = \frac{-2\gamma_j}{\beta_j - \sqrt{\beta_j^2 - 4\alpha_j\gamma_j}} \qquad \text{if } x_j < x_{j_0}. \end{aligned} \tag{D.11}$$

Once $\{R_j\}$ is known, $\{w_j\}$ can be determined from the linear equation

$$\frac{w_j - w_{j_0}}{x_j - x_{j_0}} = \frac{1}{2}\left\{\left[(A_j - C_j R_j)w_j + f_j - R_j g_j\right] + \left[(A_{j_0} - C_{j_0} R_{j_0})w_{j_0} + f_{j_0} - R_{j_0} g_{j_0}\right]\right\}, \tag{D.12}$$

which is obtained by applying the trapezoidal rule (D.9) to equation (D.5). Note that we can solve this equation for w_j each time a value of R_j is obtained from equation (D.11). Solving equation

(D.12) for w_j, we find that

$$w_j = \frac{w_{j_0} + \frac{1}{2}(x_j - x_{j_0})\left[(A_{j_0} - C_{j_0}R_{j_0})w_{j_0} - (R_j g_j + R_{j_0} g_{j_0}) + f_j + f_{j_0}\right]}{1 - \frac{1}{2}(x_j - x_{j_0})(A_j - C_j R_j)}. \tag{D.13}$$

Using the trapezoidal rule (D.9) on the reverse sweep equation (D.7), we obtain[1]

$$\frac{v_j - v_{j_0}}{x_j - x_{j_0}} = \frac{1}{2}\left\{\left[(C_j R_j + D_j)v_j + C_j w_j + g_j\right] + \left[(C_{j_0}R_{j_0} + D_{j_0})v_{j_0} + C_{j_0}w_{j_0} + g_{j_0}\right]\right\}. \tag{D.14}$$

Having determined $\{R_j, w_j\}$, we can solve the above linear equation for v_j, which we find is given by

$$v_j = \frac{v_{j_0} + \frac{1}{2}(x_j - x_{j_0})\left[(C_{j_0}R_{j_0} + D_{j_0})v_{j_0} + C_j w_j + C_{j_0}w_{j_0} + g_j + g_{j_0}\right]}{1 - \frac{1}{2}(x_j - x_{j_0})(C_j R_j + D_j)}. \tag{D.15}$$

As before, we can then use the Riccati transform (D.3) to find $\{u_j\}$.

D.2 Least-Squares Monte Carlo Simulation

In this appendix, we provide a brief account of the least-squares Monte Carlo (LSMC) algorithm developed by Longstaff and Schwartz (2001) and details on how the algorithm was adapted to the exchange option problem.

D.2.1 Simulating the Underlying State Variables

The LSMC algorithm first requires simulated values of the underlying state variables, namely the asset yield ratio and the instantaneous variance. To avoid the recursive relationship that arises when discretizing the stochastic differential equation for $\tilde{S}$, we instead work with the *log asset yield ratio* $X = \ln \tilde{S}$. An application of Itô's formula yields the following dynamics for X,

$$\begin{aligned} \mathrm{d}X(t) = &-\left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \frac{1}{2}\sigma^2 v(t)\right)\mathrm{d}t + \sigma_1\sqrt{v(t)}\,\mathrm{d}\bar{W}_1(t) + \sigma_2\sqrt{v(t)}\,\mathrm{d}\bar{W}_2(t) \\ &+ \int_{\mathbb{R}} y_1 p(\mathrm{d}y_1, \mathrm{d}t) - \int_{\mathbb{R}} y_2 p(\mathrm{d}y_2, \mathrm{d}t), \end{aligned} \tag{D.16}$$

where where $\bar{W}_1$, $\bar{W}_2$, and $\bar{Z}$ are standard $\hat{\mathbb{Q}}$-Wiener processes with correlations

$$\mathbb{E}_{\hat{\mathbb{Q}}}[\mathrm{d}\bar{W}_1(t)\,\mathrm{d}\bar{W}_2(t)] = \rho_w\,\mathrm{d}t \qquad \mathbb{E}_{\hat{\mathbb{Q}}}[\mathrm{d}\bar{W}_j(t)\,\mathrm{d}\bar{Z}(t)] = \rho_j\,\mathrm{d}t, \qquad j = 1, 2,$$

with the rest of the quantities defined in the first section of this chapter.

A Cholesky decomposition of the Wiener processes allows us to write the SDEs of the log asset yield ratio and the instantaneous variance using a standard 3-dimensional Brownian motion $\mathbf{B} = (B_1, B_2, B_3)^\top$, as

$$\begin{bmatrix} \mathrm{d}\bar{W}_1(t) \\ \mathrm{d}\bar{W}_2(t) \\ \mathrm{d}\bar{Z}(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \rho_w & \Sigma_{22} & 0 \\ \rho_1 & \Sigma_{32} & \Sigma_{33} \end{bmatrix} \begin{bmatrix} \mathrm{d}B_1(t) \\ \mathrm{d}B_2(t) \\ \mathrm{d}B_3(t) \end{bmatrix},$$

where $\Sigma_{22} = \sqrt{1 - \rho_w^2}$, $\Sigma_{32} = (\rho_2 - \rho_1\rho_2)/\Sigma_{22}$, and $\Sigma_{33} = \sqrt{1 - \rho_1^2 - \Sigma_{32}^2}$. Denote by $\mathbf{\Sigma}$ the coefficient matrix of the matrix equation above.

Under $\hat{\mathbb{Q}}$, the second asset yield process $\mathsf{DF}_{0,t} = (1/(S_2(t)e^{q_2 t}))_{t\geq 0}$ serves as a stochastic discount factor. In the absence of arbitrage opportunities, the discount factor satisfies $\mathsf{DF}_{0,t} = \mathsf{DF}_{0,s} \cdot \mathsf{DF}_{s,t}$

[1] If the forward sweep proceeds in the direction of increasing x (i.e. $j_0 = j - 1$), then the reverse sweep proceeds such that $j_0 = j + 1$.

for any $0 \le s \le t$, so the general discounting factor for any period $[s,t] \subseteq [0,T]$ is given by

$$\mathsf{DF}_{s,t} = \frac{S_2(s)e^{q_2 s}}{S_2(t)e^{q_2 t}}.$$

We thus require the stochastic differential equation characterizing the second asset price process under $\hat{\mathbb{Q}}$. Under the market measure $\mathbb{P}$, S_2 satisfies the equation

$$\frac{\mathrm{d}S_2(t)}{S_2(t)} = (\mu_2 - \lambda_2\kappa_2)\,\mathrm{d}t + \sigma_2\sqrt{v(t)}\,\mathrm{d}t + \int_{\mathbb{R}}(e^{y_2}-1)p(\mathrm{d}y_2,\mathrm{d}t).$$

Using Proposition 3.2, we can write $\mathrm{d}W_2(t) = -\psi_2(t)\,\mathrm{d}t + \mathrm{d}\bar{W}_2(t)$, where

$$\psi_2(t) = \frac{\mu_2 + q_2 - r - \sigma_2^2 v(t) - \lambda_2\kappa_2 - \tilde{\lambda}_2\tilde{\kappa}_2^-}{\sigma_2\sqrt{v(t)}},$$

and thus we have

$$\frac{\mathrm{d}S_2(t)}{S_2(t)} = \left(r - q_2 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \sigma_2^2 v(t)\right)\mathrm{d}t + \sigma_2\sqrt{v(t)}\,\mathrm{d}\bar{W}_2(t) + \int_{\mathbb{R}}(e^{y_2}-1)p(\mathrm{d}y_2,\mathrm{d}t).$$

Using Itô's Lemma, the log-price process $X_2 = \ln S_2$ satisfies the equation

$$\mathrm{d}X_2(t) = \left(r - q_2 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \frac{1}{2}\sigma_2^2 v(t)\right)\mathrm{d}t + \sigma_2\sqrt{v(t)}\,\mathrm{d}\bar{W}_2(t) + \int_{\mathbb{R}} y_2 p(\mathrm{d}y_2,\mathrm{d}t). \tag{D.17}$$

For simplicity, we shall employ an Euler-Maruyama discretization scheme to simulate X, X_2, and v. We partition the time interval $[0,T]$ using a uniform mesh $0 = t_0 < t_1 < \cdots < t_N = T$ with width Δt. Then at each time step, we can simulate time paths of X, X_2, and v using the recursions

$$\begin{aligned}
X_2(t_{i+1}) &= X_2(t_i) + \left(r - q_2 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \frac{1}{2}\sigma_2^2 v(t_i)\right)\Delta t + \sigma_2\sqrt{v(t_i)\Delta t}\,\boldsymbol{\Sigma}_2\boldsymbol{\epsilon} + \sum_{n_2=1}^{\mathcal{N}_2} Y_{2,n_2},\\
X(t_{i+1}) &= X(t_i) - \left(\tilde{\lambda}_1\tilde{\kappa}_1 + \tilde{\lambda}_2\tilde{\kappa}_2^- + \frac{1}{2}\sigma^2 v(t_i)\right)\Delta t + \sigma_1\sqrt{v(t_i)\Delta t}\,\boldsymbol{\Sigma}_1\boldsymbol{\epsilon}\\
&\qquad + \sigma_2\sqrt{v(t_i)\Delta t}\,\boldsymbol{\Sigma}_2\boldsymbol{\epsilon} + \sum_{n_1=1}^{\mathcal{N}_1} Y_{1,n} - \sum_{n_2=1}^{\mathcal{N}_2} Y_{2,n_2},\\
v(t_{i+1}) &= v(t_i) + \big(\xi\eta - (\xi+\Lambda)v(t_i)\big)\,\Delta t + \omega\sqrt{v(t_i)\Delta t}\,\boldsymbol{\Sigma}_3\boldsymbol{\epsilon},
\end{aligned} \tag{D.18}$$

where $\boldsymbol{\Sigma}_i$ is the ith row of $\boldsymbol{\Sigma}$, $\boldsymbol{\epsilon} = (\epsilon_1,\epsilon_2,\epsilon_3)^\top$ is a 3-dimensional standard normal random vector, $\mathcal{N}_j \sim \text{Poisson}(\lambda_j\Delta t)$ counts the number of jumps occurring over the time interval $(t_i, t_{i+1}]$ in the process X_j, and $(Y_{1,n_1})_{n_1=1,\ldots,\mathcal{N}_1}$ and $(Y_{2,n_2})_{n_2=1,\ldots,\mathcal{N}_2}$ are random samples from the distribution with density $m_{\hat{\mathbb{Q}}}(\mathrm{d}y_1)$ and $m_{\hat{\mathbb{Q}}}(\mathrm{d}y_2)$, respectively.

We now approximate the sample space Ω with L scenarios, each denoted by ω_l, $l = 1,\ldots,L$. We write $\mathbf{X}(\omega_l, t_i)$ to denote the value of $\mathbf{X} = (X, X_2, v)$ at time t_i in scenario l. The steps to simulating $(\mathbf{X}(\omega_l, t_i))$ are as follows: for each t_i,

1. Generate L observations each for ϵ_1, ϵ_2, and ϵ_3 from the standard normal distribution $N(0,1)$. If antithetic variates are to be used, then simulate only $L/2$ observations (assuming L is even) and append the additive inverse to form a sample of size L.

2. Generate L random numbers each for $\mathcal{N}_1$ and $\mathcal{N}_2$ from a Poisson distribution with means $\lambda_1\Delta t$ and $\lambda_2\Delta t$, respectively.

3. Generate a sample $Y_{1,1},\ldots,Y_{1,\mathcal{N}_1}$ from the common distribution of Y_1 and compute $M_1 = \sum_{n_1=1}^{\mathcal{N}_1} Y_{1,n_1}$. If $\mathcal{N}_1 = 0$, then set $M_1 = 0$. Do the same for $M_2 = \sum_{n_2=1}^{\mathcal{N}_2} Y_{2,n_2}$. This step produces L values each for M_1 and M_2.

4. For each $l = 1, \ldots, L$, compute $\mathbf{X}(\omega_l, t_i)$ using equation (D.18). Note that $\mathbf{X}(\omega_l, t_0) = \mathbf{X}(0)$ for all l. To avoid negative values for $v(\omega_l, t_i)$, we apply a reflecting condition at the origin: $v(\omega_l, t_i) \leftarrow |v(\omega_l, t_i)|$.

This process is then repeated for the next time step t_{i+1}.

D.2.2 The Least-Squares Monte Carlo Algorithm

Assuming that the time mesh introduced earlier also represents the possible dates of exercising the option, at each t_i, the option holder compares the immediate exercise value, $(S_1(t) - S_2(t))^+$, and the expected (discounted) payoff from continuing the option. The option is exercised if the exercise value is higher. The issue however is estimating the conditional expectation of the discounted payoff as viewed from time t_i. To this end, Longstaff and Schwartz (2001) proposed the *least squares Monte Carlo* (LSM) approach where at each exercise time, this conditional expectation is estimated via a least squares regression against suitably chosen basis functions of the stated variables (in our case, the asset prices and the instantaneous variance). The LSM algorithm starts at the expiry date of the option, where the optimal exercise strategy (i.e. exercise the option if it is in the money) is known, and proceeds backward in time. The algorithm then produces a pathwise approximation of the optimal exercise strategy/stopping rule that maximizes the value of the American option.

The LSM algorithm consists primarily of calculating the cash flows generated by the option conditional on no exercise prior to the current time step. Let $C(\omega, t_j; t_i, T)$ denote the path of cash flows from the option conditional on the option not being exercised at or prior to time t_i and the holder adhering to the optimal exercise strategy for all $s \in (t_i, T]$. Here, t_i denotes the current time step whereas t_j indexes all possible exercise times after t_i. Note that for a fixed ω, the sequence $C(\omega, t_j; t_i, T)$ for $j = i+1, \ldots, N$ has zero value except at the optimal exercise time for this particular sample path. Since for some scenarios it is not optimal to exercise the option at all, there may be sample paths for C which consist of zeros. Most importantly, it is possible that $(C(\omega, t_j; t_i, T))_j$ is different from $(C(\omega, t_j; t_{i+1}, T))_j$ since it may be optimal to exercise the option at time t_{i+1}, effectively changing the subsequent cash flows along the path. Thus, $C(\omega, t_j; t_i, T)$ is determined recursively, working backwards from $i = N$ to $i = 1$.

At each time step t_i, the option holder compares the immediate exercise cash flow

$$I(\omega, t_i) = \big(S_1(\omega, t_i) - S_2(\omega, t_i)\big)^+ = S_2(\omega, t_i)\left(\tilde{S}(\omega, t_i)e^{-(q_1-q_2)t_i} - 1\right)^+. \tag{D.19}$$

and the continuation value, the value of the option given that it cannot be exercised until after t_i. The continuation value, which we denote by $F(\omega, t_i)$, is given by the expected value under $\mathbb{Q}$ of any future cash flows $C(\omega, s; t_i, T)$, $t_i < s \leq T$, discounted to time t_i. That is,

$$\begin{aligned} F(\omega, t_i) &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\sum_{j=i+1}^{N} \mathsf{DF}_{t_i,t_j}(\omega) C(\omega, t_j; t_i, T)\middle|\mathcal{F}_{t_i}\right] \\ &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\sum_{j=i+1}^{N} \frac{S_2(\omega, t_i)e^{q_2 t_i}}{S_2(\omega, t_j)e^{q_2 t_j}} C(\omega, t_j; t_i, T)\middle|\mathcal{F}_{t_i}\right]. \end{aligned} \tag{D.20}$$

Since the dynamics of the state variables are Markov, the conditional expectation above is equivalent to an expectation conditional on the value of $\mathbf{X}(t_i)$. The option is then exercised at time t_i if $I(\omega, t_i) > 0$ and $I(\omega, t_i) > F(\omega, t_i)$.

Since the exact (functional) relationship between the continuation value $F(\omega, t_i)$ and the underlying state variables is unknown, we assume that the conditional expectation D.20 can be represented as a linear combination of a countable set of $\mathcal{F}_{t_k}$-measurable *Hilbert basis functions*. Specifically, if $F(\omega, t_i) \in L^2(\Omega, \mathcal{F}, \mathbb{Q})$ which is a Hilbert space, then we can write

$$F(\omega, t_i) = F(\omega, t_i, \mathbf{X}(t_i)) = \sum_{\ell=0}^{\infty} \alpha_\ell E_\ell(\mathbf{X}(t_i)), \tag{D.21}$$

where $\{E_\ell(\mathbf{X}(t_i))\}_{\ell\geq 0}$ is an orthonormal basis of $L^2(\Omega, \mathcal{F}, \mathbb{Q})$. Examples of basis functions include, Chebyshev, Hermite, Laguerre, Legendre, and Jacobi polynomials (see Abramowitz and Stegun, 1964, Chapter 22 for more examples).

In the LSM algorithm, $F(\omega, t_i)$ is approximated by the first $N_b < \infty$ elements of the orthonormal basis. Denote this approximation by $\hat{F}(\omega, t_i)$. The coefficients $\alpha_1, \ldots, \alpha_{N_b}$ of (D.21) are estimated using a least squares regression of the discounted cash flows $\sum_{j=i+1}^{N} e^{-r(t_j - t_i)} C(\omega, t_j; t_i, T)$ against $\{E_\ell(\mathbf{X}(\omega, t_i))\}_{\ell=0,\ldots,N_b}$ *for sample paths in which the option is in the money.* Note that when an option is out-of-the-money, it is better for the option holder to postpone exercise, so we only need to consider in-the-money paths when estimating the continuation value. Considering only in-the-money paths also reduces the number of basis functions needed to obtain an accurate approximation for $F(\omega, t_i)$. When there is more than one state variable (as in our case), the set of basis functions should also include terms in each state variable as well as their cross products. Glasserman (2003) also suggests including the exercise value $I(\omega, t_i)$ as one of the regressors. It is then optimal to exercise the option at time t_i if $I(\omega, t_i) > \hat{F}(\omega, t_i)$.

Once the exercise decision has been determined at time t_i, the algorithm then proceeds backward to the next time step t_{i-1}, repeating the process outlined above starting from the calculation of the cash flow matrix $C(\omega, t_j; t_{i-1}, T)$.

After the process has been performed for $t_N, t_{N-1}, \ldots, t_1$, we can also generate an $L \times N$ optimal stopping rule matrix $\mathcal{T}$, where $\mathcal{T}(\omega_l, t_i) = 1$ if and only if it is optimal to exercise the option at time t_i in path ω_l, and is zero elsewhere. Equivalently, this means that $\mathcal{T}(\omega_l, t_i) = 1$ if and only if $\tau^*(\omega_l) = t_i$. It is possible for $\mathcal{T}(\omega_l, t_i) = 0$ for all i if it never becomes optimal to exercise the option in path ω_l, in which case $\tau^*(\omega_l) = +\infty$.

Once the optimal stopping rule has been established, the American option is then valued at $t = 0$ by computing the cash flow at the optimal stopping time for each path ω_l, discounting it to $t = 0$ by multiplying the cash flow by $\mathsf{DF}_{0,\tau^*(\omega_l)}$, and taking the average of the discounted cash flows over ω. That is,

$$V_0^{LSM} = \frac{1}{L} \sum_{l=1}^{L} \sum_{j=1}^{N} \mathsf{DF}_{0,t_j} C(\omega_l, t_j; 0, T) = \frac{1}{L} \sum_{l=1}^{L} \mathsf{DF}_{0,\tau^*(\omega_l)} I(\omega_l, \tau^*(\omega_l)) \tag{D.22}$$

APPENDIX E

Methods for Parameter Estimation and Calibration of the SVJD Model

In this appendix, we give an overview of some existing methods for parameter estimation and model calibration that are relevant to our stochastic volatility and jump-diffusion models and discuss potential challenges in implementing these estimation and calibration methods in our specific problem of exchange option pricing. We emphasize that the list of methods discussed here is not exhaustive and that we do not provide an actual implementation of the parameter estimation or calibration methods.

We first distinguish parameter estimation from model calibration. In parameter estimation, we are interested in determining model parameters such that the asset price model (usually in the form of a stochastic differential equation) best replicates a given historical time series of asset price data. In contrast, model calibration, specifically the calibration of models to option price data, is the estimation of a *risk-neutral* model of the underlying assets which best approximates the prices of options written on these underlying assets. In the sections that follow, we shall make more concise the notion by which an asset price model "best replicates" historical asset prices and that by which a risk-neutral model "best approximates" option prices. As such, it is clear that parameter estimation determines the historical price process under the market probability space $(\Omega, \mathcal{F}, \mathbb{P})$, whereas model calibration seeks to estimate the asset price dynamics within the risk-neutral space $(\Omega, \mathcal{F}, \mathbb{Q})$ where option pricing occurs. This distinction is important since in incomplete markets (such as those produced by assets which exhibit both stochastic volatility and jumps) there is no unique risk-neutral measure that gives a single unique price for an option. As such, a model of historical asset prices is not sufficient to accurately price options. Thus, as noted by Cont and Tankov (2004b), model calibration can be seen as a way to augment or correct prior knowledge of asset price dynamics (as reflected by the parameter estimates in the market measure) so that observed option prices are correctly approximated by the model.

In Section E.1, we set our notation for a general stochastic differential equation with jumps that is typically used to model the price of underlying assets. We then describe a number of methods to estimate the parameters of these asset pricing models. In addition to the typical parameter estimation problem, we also briefly discuss the specialized problems of jump activity analysis and Lévy density estimation. In Section E.2, we then discuss the basic ideas behind model calibration and how single-asset and multi-asset option pricing models can be calibrated to market data.

E.1 Parameter Estimation

In this section, we shall discuss the estimation of parameters for stochastic differential equations with jumps driven by Lévy processes. We refer to Kloeden and Platen (1992), Nielsen et al. (2000), Phillips and Yu (2009), Platen and Bruti-Liberati (2010, Chapter 9), Prakasa Rao (1999), and Sørensen (2004) for a discussion of parameter estimation for pure diffusion processes.

On the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ acting as the historical or objective probability space, the relevant market state variables, for example, asset price returns and instantaneous variances, are modelled as an n-dimensional process $\mathbf{Y}$ which takes values in a state space $\mathscr{D} \subset \mathbb{R}^n$ and whose ith component satisfies the stochastic differential equation

$$\begin{aligned} \mathrm{d}Y_i(t) &= \mu_i(\mathbf{Y}(t-);\theta)\,\mathrm{d}t + \sigma_i^\top(\mathbf{Y}(t-);\theta)\,\mathrm{d}\mathbf{B}(t) \\ &\quad + \sum_{k=1}^{n} \int_{\mathbb{R}_0} \gamma_{ik}(u;\theta)(p_k(\mathrm{d}u,\mathrm{d}t) - \lambda_k(\mathbf{Y}(t);\theta) m_{\mathbb{P}}^{(k)}(\mathrm{d}u;\theta)\,\mathrm{d}t), \end{aligned} \tag{E.1}$$

where $\mathbf{B}$ is a standard d-dimensional Brownian motion under $\mathbb{P}$, $\mu_i : \mathscr{D} \times \Theta \to \mathbb{R}$, $\sigma_i : \mathscr{D} \times \Theta \to \mathbb{R}^d$, $\gamma_{ik} : \mathbb{R}_0 \times \Theta \to \mathbb{R}$, $p_k(\mathrm{d}u, \mathrm{d}t)$ for $k = 1, \ldots, n$ are independent Poisson random measures on $\mathbb{R}_0 = \mathbb{R}\backslash\{0\}$ and $[0, \infty)$ with intensity $\lambda_k(\mathbf{Y}(t);\theta)$ (for some $\lambda_k : \mathscr{D} \times \Theta \to [0, \infty)$) and jump size density $m_{\mathbb{P}}^{(k)}(\mathrm{d}u;\theta)$. Here, $\theta \in \Theta$ is the vector of parameters of (E.1) with parameter space Θ that is assumed to be a compact convex subset of $\mathbb{R}^{\mathcal{P}}$ (where $\mathcal{P}$ is the total number of parameters). The initial value $\mathbf{Y}(0)$ of the process is assumed to be known or to follow a known distribution. Whenever required, we write $\mu = (\mu_1, \ldots, \mu_n)^\top$ and $\sigma = [\sigma_{ij}] \in \mathbb{R}^{n \times d}$.

We assume that the coefficients μ, σ, and γ satisfy linear growth and Lipschitz continuity conditions to ensure that there exists a unique adapted solution $\{\mathbf{Y}(t)\}_{0 \le t \le T}$ that is right-continuous with left limits and satisfies $\mathbb{E}_{\mathbb{P}}[\|\mathbf{Y}(t)\|^2] < \infty$ for all t (see for example, Øksendal and Sulem, 2019, Theorem 1.19).

An important class of stochastic processes within the general model (E.1) is the class of *affine jump-diffusions* (Duffie et al., 2000). Affine jump-diffusion models are models for which we can write

$$\begin{aligned} \mu(\mathbf{Y}(t-);\theta) &= K_0 + K_1 \mathbf{Y}(t-), \\ [\sigma(\mathbf{Y}(t-);\theta)\sigma^\top(\mathbf{Y}(t-);\theta)]_{ij} &= [H_0]_{ij} + [H_1]_{ij}\mathbf{Y}(t-), \\ \lambda_k(\mathbf{Y}(t-);\theta) &= \ell_{0,k} + \ell_{1,k}^\top \mathbf{Y}(t-), \qquad k = 1, \ldots, n \end{aligned} \tag{E.2}$$

for some $K_0 \in \mathbb{R}^n$, $K_1 \in \mathbb{R}^{n \times n}$, $H_0 \in \mathbb{R}^{n \times n}$, $H_1 \in \mathbb{R}^{n \times n \times n}$, $\ell_{0,k} \in \mathbb{R}$, and $\ell_{1,k} \in \mathbb{R}^n$ (for $k = 1, \ldots, n$). This is to say that the drift coefficients, the entries of $\sigma\sigma^\top$, and the jump intensities are affine functions of the state variable $\mathbf{Y}(t-)$. In this case, the parameters of the model are the coefficients K_0, K_1, H_0, H_1, ℓ_0, and ℓ_1, in addition to the parameters of the jump size distribution. The availability of a closed analytical form for the conditional characteristic function

$$\psi(\phi; \mathbf{Y}(t+\tau)|\mathbf{Y}(t)) = \mathbb{E}_{\mathbb{P}}[\exp\{i\phi^\top \mathbf{Y}(t+\tau)\}|\mathbf{Y}(t)],$$

known up to the solution of a pair of complex-valued Ricatti differential equations (see for example, Duffie et al. (2000) or Section 2.5.2), is useful for parameter estimation.

The goal of parameter estimation is to find an estimator $\hat{\theta} \in \Theta$ based on observations of $\mathbf{Y}$ that will allow us to make statistical inferences about the true parameters θ. Typically, estimators for $\hat{\theta}$ are obtained from discrete observations $\{\mathbf{y}_{t_j}\}_{j=0}^N = \{\mathbf{y}_j\}_{j=0}^N$ over a period $[0, T]$, with $\mathbf{y}_0 = \mathbf{Y}(0)$ and $\mathbf{y}_N = \mathbf{Y}(T)$, of the continuous-time process $\mathbf{Y}$. Within this framework, there are at least four estimation techniques that are commonly used: (1) maximum likelihood estimation (MLE); (2) generalized method of moments (GMM); (3) indirect inference or efficient method of moments (EMM); and (4) Bayesian Monte Carlo Markov Chain (MCMC) methods. Below, each method shall be discussed briefly. However, we immediately point out that the vector process $\mathbf{Y}$ may not be completely observable as is the case, for example, when there is stochastic volatility, thus complicating the application of these estimation methods. For specific models of $\mathbf{Y}$, like the affine jump-diffusion, this problem can be addressed.

E.1.1 Maximum Likelihood Estimation and Other Likelihood-Based Approaches

In maximum likelihood estimation, the estimator $\hat{\theta}$ is chosen such that it maximizes the *likelihood* of the parameters or the probability (parameterized by θ) that the random variables $\{\mathbf{Y}(0), \ldots, \mathbf{Y}(T)\}$

take on the observed values $\{\mathbf{y}_{t_0}, \ldots, \mathbf{y}_{t_N}\} = \{\mathbf{y}_0, \ldots, \mathbf{y}_N\}$. The likelihood function is derived from the transition probability density function of $\mathbf{Y}(0), \ldots, \mathbf{Y}(T)$. The transition probability density function of $\mathbf{Y}$ is denoted by $p(\tau, \mathbf{y}; t, \mathbf{y}', \theta)$. The likelihood function of θ is then given by

$$L(\theta) = p(\mathbf{y}_0; \theta) \prod_{i=1}^{N} p(t_i, \mathbf{y}_i; 0, \mathbf{y}_0, \theta),$$

where $p(\mathbf{y}_0; \theta)$ is the probability density function of the initial state $\mathbf{Y}(0)$. If $\mathbf{Y}$ is Markov (as is the case for jump-diffusions, see Øksendal and Sulem (2019, Theorem 1.20)), then the likelihood function can be written as

$$L(\theta) = p(\mathbf{y}_0; \theta) \prod_{i=1}^{N} p(t_i, \mathbf{y}_i; t_{i-1}, \mathbf{y}_{i-1}, \theta).$$

Assuming that all the states are observed discretely without measurement error, the maximum likelihood (ML) estimator $\hat{\theta}_{ML}$ is then given by

$$\hat{\theta}_{ML} = \arg\max_{\theta \in \Theta} L(\theta).$$

For computational purposes, it is the log-likelihood function $\ell(\theta) = \log L(\theta)$ that is often maximized.

Several issues arise for the maximum likelihood approach:

1. The transition density function for $\mathbf{Y}$ as modelled by equation (E.1) is almost never available in closed analytical form. It is known that p satisfies a Kolmogorov backward equation corresponding to $\mathbf{Y}$, but even the Kolmogorov backward equation requires a numerical solution (see for example, Section 4.6 and Appendix A.5). One such model for which the transition density function is available in closed form as an infinite series expansion is the one-dimensional Merton (1976) jump-diffusion model (see for example, Ball and Torous, 1983; Aït-Sahalia, 2004). The Ornstein-Ulhenbeck process with exponentially decaying jumps also has a transition density function that is available as an infinite series (see for example, Jiang, 1998).

2. A numerical optimization algorithm (see for example, Kochenderfer and Wheeler, 2019) is inevitably required to maximize the likelihood function. The likelihood function is also not guaranteed to be concave in Θ, therefore allowing the possibility of several local maxima. In addition, convergence to a local maximum is heavily contingent on the choice of a starting point and algorithm.

3. Not all components of $\mathbf{Y}$ are observable.

Given the general unavailability of the transition density function, several advances have been made towards a likelihood-based parameter estimation. A number of these approaches involve likelihood methods based on approximations of the transition density function. There are at least three approaches in this direction: (1) maximum likelihood estimation using the transition density obtained from a discretization of the stochastic differential equation; (2) closed-form approximations of the transition density; and (3) simulated maximum likelihood (Phillips and Yu, 2009).

The first approach involves discretizing the stochastic differential equation (see for example, Platen and Bruti-Liberati, 2010, Chapter 13 for discretization schemes for SDEs with jumps) and taking advantage of the distributional properties of the increments of the Brownian and Poisson components to obtain the transition density function corresponding to the increments of the process. We note however that this introduces a discretization bias—the quality of approximation largely depends on the time scale used. For the Euler-Maruyama scheme, in particular, the resulting estimator is inconsistent if the width of the time grid is fixed (see for example, Lo, 1988). As such, Elerian (1998) argued that the Milstein discretization scheme should be used to reduce discretization biases.

To avoid having to discretize the SDE, closed form approximations via series expansions for the transition density function have been proposed. We briefly describe here the approach of Yu (2007) who extended the Hermite series expansion employed by Aït-Sahalia (2002) and Aït-Sahalia (2008) for pure diffusion processes to the case of jump-diffusion processes. Yu (2007) proposed that the transition density $p(t+\Delta, \mathbf{y}; t, \mathbf{y}', \theta)$ be expanded as

$$p(t+\Delta, \mathbf{y}; t, \mathbf{y}', \theta) = \Delta^{-n/2} \exp\left\{ -\frac{C^{(-1)}(\mathbf{y}', \mathbf{y})}{\Delta} \right\} \sum_{k=0}^{\infty} C^{(k)}(\mathbf{y}', \mathbf{y}) \Delta^k + \sum_{k=0}^{\infty} D^{(k)}(\mathbf{y}', \mathbf{y}) \Delta^k$$

for some functions $C^{(k)}$ and $D^{(k)}$ which are determined using the forward and backward Kolmogorov equations for the transition density. In addition to the restrictions induced by the Kolmogorov equations, it is also required that $C^{(-1)}(\mathbf{y}', \mathbf{y}) = 0$ if and only if $\mathbf{y}' = \mathbf{y}$ and $C^{(0)}(\mathbf{y}, \mathbf{y}) = (2\pi)^{-n/2}[\det(\sigma(\mathbf{y})\sigma^\top(\mathbf{y}))]^{-1/2}$. Yu (2007) argues that the first term captures the behavior of the transition density when $\mathbf{y}'$ is near $\mathbf{y}$, which is when the diffusion terms dominate, and the second term captures the effect of jumps on the tails of the distribution. An mth order approximation (with $m > 0$) can be obtained by truncating the infinite series after m terms,

$$p^{(m)}(t+\Delta, \mathbf{y}; t, \mathbf{y}', \theta) = \Delta^{-n/2} \exp\left\{ -\frac{C^{(-1)}(\mathbf{y}', \mathbf{y})}{\Delta} \right\} \sum_{k=0}^{m} C^{(k)}(\mathbf{y}', \mathbf{y}) \Delta^k + \sum_{k=0}^{m} D^{(k)}(\mathbf{y}', \mathbf{y}) \Delta^k.$$

With observations $\{\mathbf{y}_j\}_{j=1}^N$ for the process $\mathbf{Y}$, the log-likelihood function has mth order approximation

$$\ell^{(m)}(\theta) = \sum_{i=1}^{N} \log p^{(m)}(t_i, \mathbf{y}_i; t_{i-1}, \mathbf{y}_{i-1}, \theta)$$

and the approximate ML estimator of θ is

$$\hat{\theta}^{(m)} = \arg\max_{\theta \in \Theta} \ell^{(m)}(\theta).$$

With some technical assumptions, Yu (2007) showed that $\hat{\theta}^{(m)}$ inherits the asymptotic properties of the true ML estimator of θ. Here, asymptotic properties hold by fixing the order of approximation and letting $\Delta \to 0$.

Another closed-form approximation technique involves the use of saddlepoint approximations. Here, one approximates the conditional cumulant generating function of the transition density. See Aït-Sahalia and Yu (2007) for an application of this method to multi-dimensional Markov processes, including jump-diffusion processes, or Preston and Wood (2012) for an application to diffusion processes in conjunction with Itô-Taylor expansions.

The third likelihood-based approach is the simulated infill ML approach, also known as simulated maximum likelihood (SML). SML further reduces the discretization bias induced when the SDE is discretized by further splitting each time interval and constructing a high-frequency time series. That is, between two observations $\mathbf{Y}(t_i) = \mathbf{y}_i$ and $\mathbf{Y}(t_{i-1}) = \mathbf{y}_{i-1}$, one introduces $M-1$ auxiliary points at times $\tau_1, \ldots, \tau_{M-1}$, taking $\tau_0 = t_{i-1}$ and $\tau_M = t_i$. Assuming that $\mathbf{Y}$ is Markov, the transition density $p(t_i, \mathbf{y}_i; t_{i-1}, \mathbf{y}_{i-1}, \theta)$ can be calculated by carrying out an $(M-1)$-dimensional integration

$$\begin{aligned} p(t_i, \mathbf{y}_i; t_{i-1}, \mathbf{y}_{i-1}, \theta) &= \int \cdots \int p(\mathbf{y}_{\tau_M}, \mathbf{y}_{\tau_{M-1}} \ldots, \mathbf{y}_{\tau_1}; y_{\tau_0}, \theta) \, \mathrm{d}\mathbf{y}_{\tau_{M-1}} \ldots \mathrm{d}\mathbf{y}_{\tau_1} \\ &= \int \cdots \int \prod_{m=1}^{M} p(\tau_m, \mathbf{y}_{\tau_m}; \tau_{m-1}, \mathbf{y}_{\tau_{m-1}}, \theta) \, \mathrm{d}\mathbf{y}_{\tau_{M-1}} \ldots \mathrm{d}\mathbf{y}_{\tau_1}. \end{aligned}$$

One way to approximate the auxiliary transition densities $p(\tau_m, \mathbf{y}_{\tau_m}; \tau_{m-1}, \mathbf{y}_{\tau_{m-1}}, \theta)$ is to use an Euler discretization over $[\tau_{m-1}, \tau_m]$. The integration can then be approximated by Monte Carlo integration by drawing samples of $\{\mathbf{Y}(\tau_1), \ldots, \mathbf{Y}(\tau_{M-1})\}$ (Pedersen, 1995).

The key issues with SML are the introduction of further discretization biases by introducing auxiliary points and the computational cost of performing the multi-dimensional integration. Refinements to the basic SML methodology for pure diffusion processes have been proposed by Elerian et al. (2001), Brandt and Santa-Clara (2002) and Durham and Gallant (2002). These refinements include the use of alternative discretization methods and making more efficient use of the knowledge of $\mathbf{y}_{\tau_M} = \mathbf{y}_i$ and $\mathbf{y}_{\tau_0} = \mathbf{y}_{i-1}$ in simulating the auxiliary points. While these refinements have been proposed in the context of diffusion processes, SML is intimately linked to MCMC methods for parameter estimation which have been applied to estimating jump-diffusions, including stochastic volatility models (see for example, Eraker et al., 2003).

Outside the aforementioned transition density-based approximate ML methods, Shimizu and Yoshida (2006) and Ogihara and Yoshida (2011) proposed a quasi-likelihood estimation method where the likelihood function is estimated by a function with two parts: (1) the log-likelihood of a local Gaussian process and (2) an estimator based on the log-likelihood of Poisson random measures. Increments of the observed process are tested for the presence or absence of a jump to determine which part of the estimating function they shall be used. They show that the consistency and asymptotic normality of the estimator as the sampling frequency rapidly increases. In addition Ogihara and Yoshida (2011) provide adaptive Bayes-type estimators for the same class of jump-diffusion processes.

For the class of affine diffusion processes and affine jump diffusion processes, Singleton (2001) uses an inversion formula on the conditional characteristic function to obtain a conditional log-likelihood function. Bates (2006) extended the strategy of inverting the characteristic function to the case where some components of the affine process are latent or unobservable variables, thereby integrating the parameter estimation and the filtration stages into an approximate maximum likelihood estimation method. In addition, Bates (2006) proposed a Bayesian updating algorithm for the characteristic function of the observable and latent components of the process in terms of the characteristic function of the latent component conditional on data on the observable component.

Bégin and Boudreault (2020) proposed a likelihood-based method using discrete nonlinear filters for a stochastic volatility jump-diffusion model with jumps in the volatility and stochastic jump arrival intensity. Similar to Bates (2006), stochastic volatility and jump intensity are considered latent variables. At each time t, the discrete nonlinear filter involves estimating the contribution of the observed state at time t conditional on past observed states to the likelihood function and calculating the posterior distribution of the latent variable at time t conditional on current and past observations. The likelihood function is then calculated by multiplying all of the likelihood contributions. The main challenge in the algorithm is the multidimensional integration problem involved in calculating the likelihood contribution and the posterior distribution. Thus, numerical integration methods need to be considered. Therefore, the estimation problem becomes increasingly complex as the dimensionality increases, as is also the case for the AML approach of Bates (2006).

Statistical packages are also able to perform maximum likelihood estimation of specific classes of SDEs with jumps. YUIMA package in `R` (Brouste et al., 2014; Iacus and Yoshida, 2018) provides a quasi-maximum likelihood estimation of scalar jump-diffusions with Poisson jumps based on the QMLE of Shimizu and Yoshida (2006) and Ogihara and Yoshida (2011). The DiffusionRjgqd package (Pienaar and Varughese, 2015), also for `R`, provides utilities for inference of generalized quadratic jump-diffusion processes, including a likelihood-based parameter estimation. The DiffusionRjgqd package accommodates multi-dimensional jump-diffusion processes where the jump intensity can be both state-dependent and stochastic (Pienaar and Varughese, 2016).

E.1.2 Generalized Method of Moments (GMM) Estimation

Unlike maximum likelihood estimation, the generalized method of moments (GMM) approach requires only the specification of some moment conditions rather than the full transition density function. The GMM is hence robust with respect to distributional misspecification, but at the cost of some information loss relative to the complete information provided by the transition density.

The required moment conditions are expressed as

$$\mathbb{E}_{\mathbb{P}}[\mathbf{f}(\mathbf{y}_j;\theta)] = \mathbf{0},$$

where $\mathbf{f}(\mathbf{y}_j;\theta)$ is an r-dimensional vector-valued function (with $r \geq \mathcal{P}$) and $\mathbf{0}$ is the zero vector in $\mathbf{R}^r$. Specficially, the equation above provides r moment conditions on the discretized version of stochastic differential equation (E.1). The expectation can then be estimated by

$$\mathbf{g}(\theta;\mathbf{y}) = \frac{1}{N}\sum_{j=1}^{N}\mathbf{f}(\mathbf{y}_j;\theta)$$

where $g : \mathbb{R}^{\mathcal{P}} \to \mathbb{R}^r$. The GMM estimator $\hat{\theta}_{GMM}$ of θ is then the solution of the equation

$$\mathbf{g}(\hat{\theta}_{GMM};\mathbf{y}) = 0.$$

This equation is solved by means of a quadratic minimization problem

$$\hat{\theta}_{GMM} = \arg\min_{\theta\in\Theta} \mathbf{g}^{\top}(\theta;\mathbf{y})\mathcal{W}_N\mathbf{g}(\theta,\mathbf{y}),$$

where $\mathcal{W}_N$ is an $r\times r$ symmetric positive semi-definite weight matrix. The weight matrix is typically chosen to minimize the asymptotic variance of the GMM estimator (Hansen, 1982). Under some regularity conditions, Hansen (1982) showed that the GMM estimator is consistent and asymptotically normal, i.e. $\sqrt{N}(\hat{\theta}_{GMM} - \theta) \to \mathcal{N}(\mathbf{0}, \mathbf{V}_{GMM})$, where the variance matrix $\mathbf{V}_{GMM}$ can be estimated consistently by

$$\hat{\mathbf{V}}_{GMM} = \frac{1}{N}\hat{\mathbf{D}}^{\top}\hat{\mathbf{S}}^{-1}(\hat{\mathbf{D}}^{\top})^{-1}.$$

Here, $\hat{\mathbf{D}}$ is the Jacobian matrix of $\mathbf{g}(\theta;\mathbf{y})$ with respect to θ evaluated at $\hat{\theta}_{GMM}$ and $\hat{\mathbf{S}}$ is a consistent estimator of the covariance matrix of moment restrictions $\mathbf{f}(\mathbf{y}_j;\theta)$.

A crucial consideration with the GMM is that the estimator depends largely on the chosen moment conditions for which there exists no hard and fast rule. In addition, when moments for $\mathbf{Y}$ are not known, the SDE has to be discretized in order to obtain suitable moment conditions. As discussed prior, this introduces discretization biases depending on the discretization scheme used. As such, errors would then arise and are difficult to correct for in the estimation. However, for models whose joint characteristic function is available in closed analytical form (as is the case for affine models), moment conditions can be derived explicitly (Singleton, 2001; Jiang and Knight, 2002). Furthermore, the asymptotic properties of the GMM estimator hold provided that the chosen moments are finite. This may not be guaranteed when the transition density of $\mathbf{Y}$ exhibits heavy tails, as is the case for models with jumps.

In the case that $\mathbf{Y}$ is an affine jump-diffusion with observable $\mathbf{Y}^{(O)}$ and latent $\mathbf{Y}^{(L)}$ components, Jiang and Knight (2002) note that although $\mathbf{Y}$ is Markov, the observable component $\mathbf{Y}^{(O)}$ by itself is not. As such, the transition density of $\mathbf{Y}^{(O)}$ does not enjoy the iterative property of the transition density of the whole process $\mathbf{Y}$. However, Jiang and Knight (2002) show that if the characteristic function of the latent component $\mathbf{Y}^{(L)}$ is available in closed form, then the unconditional characteristic function of $\mathbf{Y}^{(O)}$ has a closed analytical form. Otherwise, one can still obtain in closed form the conditional characteristic function of $\mathbf{Y}^{(O)}$ conditional on $\mathbf{Y}^{(L)}$. Using these properties, Jiang and Knight (2002) presented a GMM approach to estimating the parameters of $\mathbf{Y}$. Pan (2002) augments the historical series of asset prices with a time series of observed option prices, resulting to an "implied state" GMM approach. Aït-Sahalia et al. (2015) use a GMM estimation approach to estimate the parameters of mutually exciting jump-diffusions driven by Hawkes processes.

E.1.3 Indirect Inference or Efficient Method of Moments (EMM) Estimation

When the original model or stochastic differential equation (E.1) results to a complicated and intractable likelihood function, the indirect inference method developed by Gourieroux et al. (1993) instead considers an auxiliary, oftentimes simpler, model to approximate and indirectly estimate the original model. Indirect inference is intimately related to the efficient method of moments (EMM), developed by Gallant and Tauchen (1996), in which the auxiliary model is used to generate moment conditions to be used in GMM estimation. In this section, we give an overview of the indirect inference method following the discussion of Jiang (1998). A detailed discussion of the efficient method of moments is given by Nielsen et al. (2000) and Gallant and Tauchen (2010). Bakshi et al. (1997), Andersen et al. (2002) and Chernov et al. (2003), among others, have employed the EMM to estimate the parameters of stochastic volatility and jump-diffusion models.

The first step in indirect inference is to choose and fit an auxiliary or instrumental model to the observed data $\{\mathbf{y}_j\}_{j=1}^N$. The auxiliary model is characterized by the conditional density function $f_I(\mathbf{y}; \mathcal{Y}^N, \theta_I)$, where $\theta_I \in \mathbb{R}^q$ is the parameter vector of the auxiliary model and $\mathcal{Y}^N$ contains lagged values of $\{\mathbf{y}_j\}_{j=1}^N$. Here, the dimension of θ_I need not be the same as that of the original parameter vector θ and not all components of the process $\mathbf{Y}$ need to be observable. For a one-dimensional time series, the auxiliary model may be an ARMA-GARCH model of any order. Jiang (1998) provides a simple auxiliary model for a one-dimensional jump-diffusion process based on a first-order discretization of the SDE. The auxiliary model must be chosen so that its parameters θ_I can easily be estimated, say with maximum likelihood estimation.

For simplicity, we assume that the auxiliary model is also Markov, as in Jiang (1998), so we can write the conditional density function as $f_I(\mathbf{y}_i; \mathbf{y}_{i-1}, \theta_I)$. The MLE for θ_I is therefore

$$\hat{\theta}_I = \arg\max_{\theta_I} \sum_{j=1}^{N} \log f_I(\mathbf{y}_i; \mathbf{y}_{i-1}, \theta).$$

The next step is to simulate paths from the original model (E.1) and estimate the auxiliary model based on the simulated sample paths. The auxiliary model f_I is typically misspecified with respect to the true model, so the estimator $\hat{\theta}_I$ is biased and inconsistent for θ. Simulations from the original model are then used to correct for the asymptotic biases of $\hat{\theta}_I$.

Given a value θ_0 of the parameter vector θ and the initial value $\mathbf{Y}(0)$ of $\mathbf{Y}$, we can simulate values $\{\tilde{\mathbf{Y}}(t_j; \theta_0)\}_{j=1}^N$ from (E.1). See Appendix D.2 for a simple simulation method or Glasserman (2003) and Platen and Bruti-Liberati (2010) for a comprehensive analysis of available simulation methods for SDEs with jumps. We simplify our notation to $\tilde{\mathbf{y}}_j(\theta_0) = \tilde{\mathbf{Y}}(t_j; \theta_0)$. A total of H sample paths are generated, of which the hth simulation is denoted by $\tilde{\mathbf{y}}^{(h)}(\theta_0) = \{\tilde{\mathbf{y}}_j^{(h)}(\theta_0)\}_{j=1}^N$. We then again estimate θ_I using the ML method, this time with f_I evaluated at the simulated sample paths. The estimator $\hat{\theta}_I^{HN}$ is given by

$$\hat{\theta}_I^{HN}(\theta_0) = \arg\max_{\theta_I} \sum_{h=1}^{H} \sum_{j=1}^{N} \log f_I(\tilde{\mathbf{y}}_j^{(h)}(\theta_0); \tilde{\mathbf{y}}_j^{(h)}(\theta_0), \theta_I).$$

The last step is to estimate the original parameter θ using $\hat{\theta}_I$ and $\hat{\theta}_I^{HN}$. The original parameters θ are estimated using a GMM criterion

$$\hat{\theta}^{HN}(\mathcal{W}) = \arg\min_{\theta_0 \in \Theta} (\hat{\theta}_I - \hat{\theta}_I^{HN}(\theta_0))^\top \mathcal{W} (\hat{\theta}_I - \hat{\theta}_I^{HN}(\theta_0)),$$

where, as in the GMM approach, $\mathcal{W}$ is a $q \times q$ symmetric positive semi-definite weighting matrix. This estimator is motivated by the idea that if the original model (E.1) is correctly specified and the true value of θ is θ_0, then there should be little difference between the data $\{\mathbf{y}_j\}$, as fitted by the auxiliary model $f_I(\mathbf{y}_j; \mathbf{y}_{j-1}, \theta_I)$, and the simulated values $\{\tilde{\mathbf{y}}_j(\theta_0)\}$, as fitted by the auxiliary

model $f_I(\tilde{\mathbf{y}}_j^{(h)}(\theta_0); \tilde{\mathbf{y}}_j^{(h)}(\theta_0), \theta_I)$. One possible choice for the weight matrix is

$$\mathcal{W} = \frac{1}{N}\sum_{j=1}^{N}\left[\nabla_{\theta_I}\log f_I(\mathbf{y}_j;\mathbf{y}_{j-1},\hat{\theta}_I)\right]\left[\nabla_{\theta_I}\log f_I(\mathbf{y}_j;\mathbf{y}_{j-1},\hat{\theta}_I)\right]^\top,$$

where $\nabla_{\theta_I}\log f_I(\mathbf{y}_j;\mathbf{y}_{j-1},\hat{\theta}_I)$ is the gradient of $\log f_I$ with respect to θ_I evaluated at $\hat{\theta}_I$.

In literature on the efficient method of moments, the estimator $\hat{\theta}^{HN}(\mathcal{W})$ is formulated as

$$\hat{\theta}^{HN}(\mathcal{W}) = \arg\min_{\theta\in\Theta}(m(\theta,\hat{\theta}_I))^\top \mathcal{W} m(\theta,\hat{\theta}_I),$$

where $m(\theta,\hat{\theta}_I)$ is given by

$$m(\theta,\hat{\theta}_I) = \frac{1}{NH}\sum_{h=1}^{H}\sum_{j=1}^{N}\nabla_{\theta_I}\log f_I(\tilde{\mathbf{y}}_j^{(h)}(\theta);\tilde{\mathbf{y}}_{j-1}^{(h)}(\theta),\hat{\theta}_I)$$

(Gallant and Tauchen, 1996; Nielsen et al., 2000). The quantity $m(\theta,\hat{\theta}_I)$ estimates the expected value of the score function, which is the gradient of the log of the conditional density of the auxiliary model evaluated at simulations of the original model (E.1) and at the ML estimate $\hat{\theta}_I$. In view of the GMM approach, $m(\theta,\hat{\theta}_I)$ gives the moment conditions for the (simulated) method of moments estimation of θ. The EMM formulation has computational advantage if the gradient is available in closed form, since only one optimization in θ is needed, but this is typically not the case.

Similar to the GMM estimator, the indirect inference or EMM estimator $\hat{\theta}^{HN}$ is also consistent and asymptotically normal as $N\to\infty$ for fixed H under some regularity conditions (Gourieroux et al., 1993).

E.1.4 Markov Chain Monte Carlo (MCMC) Approach

Markov Chain Monte Carlo (MCMC) algorithms offer a Bayesian solution to the parameter estimation problem by characterizing the conditional distribution of the parameters and the unobservable components of $\mathbf{Y}$, if there are any, conditional on the observable components $\mathbf{Y}^{(O)}$. As such, MCMC algorithms naturally incorporate a state-space interpretation of the asset price dynamics (see for example, Duffie et al., 2000; Jiang and Knight, 2002) where the components of $\mathbf{Y}$ can be partitioned into observable $\mathbf{Y}^{(O)}$ and latent $\mathbf{Y}^{(L)}$ components. Eraker et al. (2003) employed MCMC algorithms in estimating the parameters of an affine stochastic volatility jump-diffusion where the instantaneous variance process also exhibits jumps, although as discussed by Johannes and Polson (2010), MCMC algorithms may also be applied to solve the joint problem of parameter estimation and model calibration if option price data is available. This connection to model calibration shall be discussed in the next section. Our discussion below follows that of Johannes and Polson (2010), but with more focus on parameter estimation as in Eraker et al. (2003).

Before we discuss the basic ideas behind MCMC algorithms, we first state additional assumptions on the model (E.1) and set our notation. As above, suppose the n-dimensional process $\mathbf{Y}$ can be partitioned into $\mathbf{Y}^{(O)}\in\mathbb{R}^{n_O}$ and $\mathbf{Y}^{(L)}\in\mathbb{R}^{n_L}$, where $n_O+n_L=n$. Following Eraker et al. (2003) and Johannes and Polson (2010), we assume that each component of $\mathbf{Y}$ has only one jump term. Furthermore, we assume that the coefficients of the SDE for the latent part $\mathbf{Y}^{(L)}$ do not depend on the observable components, but the coefficients of the SDE for $\mathbf{Y}^{(O)}$ may depend on both $\mathbf{Y}^{(O)}$ and $\mathbf{Y}^{(L)}$. We note that this assumption applies to our two-asset SVJD model in Chapter 2. We thus assume that the components of $\mathbf{Y}^{(O)}$ and $\mathbf{Y}^{(L)}$ satisfy the stochastic differential equations

$$\begin{aligned}\mathrm{d}Y_i^{(O)}(t) &= \mu_i^{(O)}(\mathbf{Y}^{(O)}(t-),\mathbf{Y}^{(L)}(t-);\theta)\,\mathrm{d}t + \left(\sigma_i^{(O)}(\mathbf{Y}^{(O)}(t-),\mathbf{Y}^{(L)}(t-);\theta)\right)^\top \mathrm{d}\mathbf{B}^{(O)}(t)\\ &\quad + \mathrm{d}\left(\sum_{j=1}^{N_i^{(O)}(t)} Z_{i,j}^{(O)}\right), \qquad i=1,\ldots,n_O\end{aligned} \tag{E.3}$$

$$\begin{aligned} \mathrm{d}Y_i^{(L)}(t) &= \mu_i^{(L)}(\mathbf{Y}^{(L)}(t-);\theta)\,\mathrm{d}t + \left(\sigma_i^{(L)}(\mathbf{Y}^{(L)}(t-);\theta)\right)^\top \mathrm{d}\mathbf{B}^{(L)}(t) \\ &\quad + \mathrm{d}\left(\sum_{j=1}^{N_i^{(L)}(t)} Z_{i,j}^{(L)}\right), \qquad i = 1,\ldots,n_L \end{aligned} \tag{E.4}$$

where $\mathbf{B}^{(O)}(t) \in \mathbb{R}^{n_O}$ and $\mathbf{B}^{(L)}(t) \in \mathbb{R}^{n_L}$ are standard Brownian motions under $\mathbb{P}$, $N_i^{(O)}(t)$ and $N_i^{(L)}(t)$ are counting processes with (possibly stochastic) intensity $\lambda_i^{(O)}(\mathbf{Y}^{(O)}(t-), \mathbf{Y}^{(L)}(t-);\theta)$ and $\lambda_i^{(L)}(\mathbf{Y}^{(L)}(t-);\theta)$, respectively, and $Z_{i,j}^{(O)} \sim \Pi_i^{(O)}(\theta)$ and $Z_{i,j}^{(L)} \sim \Pi_i^{(L)}(\theta)$ are i.i.d. draws from a jump size distributions $\Pi_i^{(O)}(\theta)$ and $\Pi_i^{(L)}(\theta)$ parameterized by θ. We also assume that $\mu_i^{(O)}$, $\mu_i^{(L)}$, $\sigma_i^{(O)}$, and $\sigma_i^{(L)}$ are known parametric functions.

As mentioned above, the Bayesian solution involves characterizing the conditional posterior distribution of θ and the collection of latent or hidden factors $\mathcal{H}$ (whose elements will be clarified later) given the observable component $\mathbf{Y}^{(O)}$. We denote this posterior distribution by $\mathsf{p}(\theta, \mathcal{H}|\mathbf{Y}^{(O)})$. The posterior distribution summarizes information about the model (E.3), the evolution of the latent components (E.4), and any prior information about the parameters. To see this, we note that by Bayes' Rule the posterior distribution factors, up to proportionality, as

$$\mathsf{p}(\theta, \mathcal{H}|\mathbf{Y}^{(O)}) \propto \mathsf{p}(\mathbf{Y}^{(O)}|\theta, \mathcal{H})\mathsf{p}(\mathcal{H}|\theta)\mathsf{p}(\theta),$$

where $\mathsf{p}(\mathbf{Y}^{(O)}|\theta, \mathcal{H})$ is the likelihood function of the observable component, $\mathsf{p}(\mathcal{H}|\theta)$ is the conditional distribution of latent variables, and $\mathsf{p}(\theta)$ is the prior of the parameters. The prior, often introduced by the modeller, allows one to include non-sample information about the parameters, like the positivity of certain parameters or known jump size characteristics.

However, the main challenge of implementing MCMC methods is characterizing the conditional posterior $\mathsf{p}(\theta, \mathcal{H}|\mathbf{Y}^{(O)})$, which is often very complicated and high-dimensional. To address this issue, MCMC algorithms rely on the Clifford-Hammersley Theorem (see for example, Robert and Casella, 2004, Theorem 9.3 for a statement), which states that $\mathsf{p}(\theta, \mathcal{H}|\mathbf{Y}^{(O)})$ is completely determined by the lower-dimensional distributions $\mathsf{p}(\theta|\mathcal{H}, \mathbf{Y}^{(O)})$, the posterior distribution of the parameters, and $\mathsf{p}(\mathcal{H}|\theta, \mathbf{Y}^{(O)})$, the smoothing distribution of the latent factors.

MCMC algorithms then allow us to form samples $\{\theta^{(g)}, \mathcal{H}^{(g)}\}_{g=1}^G$ from the posterior distributions $\mathsf{p}(\theta, \mathcal{H}|\mathbf{Y}^{(O)})$ and $\mathsf{p}(\theta|\mathcal{H}, \mathbf{Y}^{(O)})$ as follows. Given initial values $\theta^{(0)}$ and $\mathcal{H}^{(0)}$, we draw $\mathcal{H}^{(1)} \sim \mathsf{p}(\mathcal{H}|\theta^{(0)}, \mathbf{Y}^{(O)})$, then we draw $\theta^{(1)} \sim \mathsf{p}(\theta|\mathcal{H}^{(1)}, \mathbf{Y}^{(O)})$. Note that this is the Gibbs sampling approach, which is applicable when the conditional posteriors are known distributions or when direct sampling from the conditional posteriors is possible. This procedure produces a Markov chain which is known to have a stationary distribution $\mathsf{p}(\theta, \mathcal{H}|\mathbf{Y}^{(O)})$. The constructed Markov chain can then be used to obtain consistent estimates of expectations

$$\mathbb{E}_{\mathbb{P}}[f(\theta, \mathcal{H})|\mathbf{Y}^{(O)}] = \int\int f(\theta, \mathcal{H})\mathsf{p}(\theta, \mathcal{H}|\mathbf{Y}^{(O)})\,\mathrm{d}\theta\,\mathrm{d}\mathcal{H}$$

via the estimator

$$\hat{\mathbb{E}}_{\mathbb{P}}[f(\theta, \mathcal{H})|\mathbf{Y}^{(O)}] = \frac{1}{G}\sum_{g=1}^{G} f(\theta^{(g)}, \mathcal{H}^{(g)}).$$

See Johannes and Polson (2010) for an overview of convergence results related to MCMC algorithms.

To use MCMC algorithms to estimate the parameters of the model (E.3) and (E.4), we start with an Euler-Maruyama discretization, assuming the sampling interval Δt is equal to the observed frequency, which we assume to be equal to 1. The discretized equations are

$$\begin{aligned} Y_i^{(O)}(t+1) &= Y_i^{(O)}(t) + \mu_i^{(O)}(\mathbf{Y}^{(O)}(t), \mathbf{Y}^{(L)}(t);\theta) + \left(\sigma_i^{(O)}(\mathbf{Y}^{(O)}(t), \mathbf{Y}^{(L)}(t);\theta)\right)^\top \epsilon^{(O)}(t+1) \\ &\quad + Z_i^{(O)}(t+1)J_i^{(O)}(t+1), \qquad i = 1,\ldots,n_O \end{aligned} \tag{E.5}$$

$$Y_i^{(L)}(t+1) = Y_i^{(L)}(t) + \mu_i^{(L)}(\mathbf{Y}^{(L)}(t);\theta) + \left(\sigma_i^{(L)}(\mathbf{Y}^{(L)}(t);\theta)\right)^\top \epsilon^{(L)}(t+1) + Z_i^{(L)}(t+1)J_i^{(L)}(t+1), \qquad i = 1,\dots,n_L, \tag{E.6}$$

where $\epsilon^{(O)}(t+1) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{n_O})$, $\epsilon^{(L)}(t+1) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{n_L})$ ($\mathbf{I}_{n_O}$ and $\mathbf{I}_{n_L}$ are the $n_O \times n_O$ and $n_L \times n_L$ identity matrices), $Z_i^{(O)}(t+1) \sim \Pi_i^{(O)}(\theta)$, $Z_i^{(L)}(t+1) \sim \Pi_i^{(L)}(\theta)$, $J_i^{(O)}(t+1) \sim \text{Bernoulli}(\lambda_i^{(O)}(\mathbf{Y}^{(L)}(t);\theta))$, and $J_i^{(L)}(t+1) \sim \text{Bernoulli}(\lambda_i^{(L)}(\mathbf{Y}^{(L)}(t);\theta))$. To simplify the illustration of the main ideas, we assume that the observable and latent components of $\mathbf{Y}$ each have dimension 2, as is the case for the two-asset SVJD model in Chapter 2.

Given the discretized equations (E.5) and (E.6), we now properly identify our latent space $\mathcal{H}$ as

$$\mathcal{H} = \left\{ \left(\mathbf{Y}^{(L)}(t-1), \{Z_i^{(O)}(t)\}_{i=1,2}, \{Z_i^{(L)}(t)\}_{i=1,2}, \{J_i^{(O)}(t)\}_{i=1,2}, \{J_i^{(L)}(t)\}_{i=1,2} \right) \right\}_{t=1,\dots,N},$$

(assuming there are N observation points available). That is, in addition to the latent components of $\mathbf{Y}$, we also consider part of the latent space the jump sizes and the random variables which determine the presence of jumps during the period $(t, t+1]$.

In this model, the conditional posteriors $\mathsf{p}(\theta|\mathcal{H}, \mathbf{Y}^{(O)})$ and $\mathsf{p}(\mathcal{H}|\theta, \mathbf{Y}^{(O)})$ are still high-dimensional and difficult to characterize. In this case, the Clifford-Hammersley Theorem can be used to assert that these conditional posteriors can be determined by other conditional posteriors that have an even lower dimensionality. For example, supposing we can partition θ into $K \leq \mathcal{P}$ blocks $\theta = (\theta_1, \dots, \theta_K)$, then the conditional posteriors

$$\mathsf{p}(\theta_k|\theta_{(-k)}, \mathcal{H}, \mathbf{Y}^{(O)}), \qquad k = 1,\dots,K$$

where $\theta_{(-k)}$ is θ excluding the kth block, completely determines $\mathsf{p}(\theta|\mathcal{H}, \mathbf{Y}^{(O)})$. Similarly, the conditional posterior $\mathsf{p}(\mathcal{H}|\theta, \mathbf{Y}^{(O)})$ may be fully determined by the simpler conditional posteriors

(Jump time posteriors)

$$\mathsf{p}\left(J_i^{(O)}(t) \middle| \theta, \{J_i^{(L)}(t)\}_{i=1,2}, \{Z_i^{(O)}(t)\}_{i=1,2}, \{Z_i^{(L)}(t)\}_{i=1,2}, \mathbf{Y}^{(L)}(t), \mathbf{Y}^{(O)}(t)\right), \quad i = 1,2$$

$$\mathsf{p}\left(J_i^{(L)}(t) \middle| \theta, \{J_i^{(O)}(t)\}_{i=1,2}, \{Z_i^{(O)}(t)\}_{i=1,2}, \{Z_i^{(L)}(t)\}_{i=1,2}, \mathbf{Y}^{(L)}(t), \mathbf{Y}^{(O)}(t)\right), \quad i = 1,2$$

(Jump size posteriors)

$$\mathsf{p}\left(Z_i^{(O)}(t) \middle| \theta, \{J_i^{(L)}(t)\}_{i=1,2}, \{J_i^{(O)}(t)\}_{i=1,2}, \{Z_i^{(L)}(t)\}_{i=1,2}, \mathbf{Y}^{(L)}(t), \mathbf{Y}^{(O)}(t)\right), \quad i = 1,2$$

$$\mathsf{p}\left(Z_i^{(L)}(t) \middle| \theta, \{J_i^{(L)}(t)\}_{i=1,2}, \{J_i^{(O)}(t)\}_{i=1,2}, \{Z_i^{(O)}(t)\}_{i=1,2}, \mathbf{Y}^{(L)}(t), \mathbf{Y}^{(O)}(t)\right), \quad i = 1,2$$

(Latent $\mathbf{Y}$ posteriors)

$$\mathsf{p}\Big(\mathbf{Y}^{(L)}(t) \Big| \theta, \mathbf{Y}^{(L)}(t+1), \mathbf{Y}^{(L)}(t+1), \{J_i^{(L)}(t)\}_{i=1,2}, \{J_i^{(O)}(t)\}_{i=1,2}, \{Z_i^{(O)}(t)\}_{i=1,2}, \{Z_i^{(L)}(t)\}_{i=1,2}, \mathbf{Y}^{(O)}(t)\Big).$$

By Bayes' Rule, the calculation of each of these conditional posteriors involve the prior of the parameters $\mathsf{p}(\theta)$, likelihood function $\mathsf{p}(\mathbf{Y}^{(O)}|\theta, \mathcal{H})$ and the conditional density $\mathsf{p}(\mathcal{H}|\theta)$, noting that the latter two can be derived from the discretized equations (E.5) and (E.6).

If these conditional posteriors are a known distribution, depending on how the priors are chosen, then the MCMC algorithm may be implemented via Gibbs sampling, the direct sampling method described above. Otherwise, indirect sampling methods, known collectively as Metropolis-Hastings algorithms are required. Metropolis-Hastings algorithms are based on acceptance-rejection schemes on known candidate distributions which aim to approximate the unknown conditional posterior of interest. In asset price models with stochastic volatility, both methods are employed, as was done

by Eraker et al. (2003). Amengual and Xiu (2018) also used MCMC methods to estimate the parameters of non-affine jump-diffusion and stochastic volatility models. See Johannes and Polson (2009) and Johannes and Polson (2010) for further examples of designing MCMC algorithms for commonly used asset pricing and term structure models. Gudkov and Ignatieva (2021) apply the MCMC algorithm to estimate models for electricity prices with stochastic volatility and jumps. For a textbook discussion of MCMC algorithms and their theoretical underpinnings, see Robert and Casella (2004).

As evident from this discussion, MCMC algorithms are unified estimation procedures in the sense that both parameter estimation and the estimation of the latent variables are both carried out. In fact, augmenting the observable data on $\mathbf{Y}^{(O)}$ with option price data, Johannes and Polson (2010) show that even model calibration can be done with one MCMC infrastructure. In addition, we bypass the need to maximize the likelihood, which itself has a number of issues, although we still need access to the likelihood function. Via the prior distribution of parameters, the users are also able to incorporate non-sample information about parameters. However, the benefits of the MCMC algorithms, included increased computational efficiency, are significantly dependent, among other factors, on the construction of the posterior distributions, the choice of priors (which is not a trivial task), and how sampling from these posteriors can be done.

E.1.5 Jump Activity Analysis

The statistical problem of estimating specific components of the dynamics of $\mathbf{Y}$ has also gained considerable traction. In this section, we look at techniques for estimating the jump component of $\mathbf{Y}$, particularly the number of jumps that occur within a specified period of time, the size of the jumps, and the jump intensity.

For the Merton (1976) jump-diffusion model (where jumps are modelled using a compound Poisson process) and the jump-diffusion model driven by Cauchy jumps (see for example, Cont and Tankov, 2004a), both of which can be expressed using the scalar jump SDE

$$\mathrm{d}Y(t) = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W(t) + \mathrm{d}Z(t)$$

for the log-return $Y(t)$, Aït-Sahalia (2004) uses the maximum likelihood approach to solve a more specific problem: how to isolate the diffusive component from the jump component solely using the time series of historical asset prices. This problem is understandably difficult as multiple (if not infinite) combinations of jump parameters and σ^2 may result to the same observable variance or excess kurtosis of observed log-returns. Nonetheless, Aït-Sahalia (2004) shows that it is possible for maximum likelihood estimation to "disentagle" diffusion from jumps in the sense that the asymptotic variance of the MLE estimator for σ^2 is the same whether or not jumps are present. The same result also holds when jumps are smaller and more frequent, as captured by the Cauchy distribution. Furthermore, as the sampling of returns becomes for frequent, the ability to distinguish σ^2 from jumps also improves. In other words, the separation of diffusive and jump components is possible (and made easier) using high-frequency data.

Non-parametric approaches have also been proposed for separating diffusion and jump effects using high-frequency data. Before we provide an overview of these non-parametric approaches, we first establish some notation. We consider a scalar vector process Y satisfying the jump SDE

$$\mathrm{d}Y(t) = \mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t) + \int_{\mathbb{R}_0} \gamma(u)(p(\mathrm{d}u, \mathrm{d}t) - \lambda(t)m_{\mathbb{P}}(\mathrm{d}u)\,\mathrm{d}t). \tag{E.7}$$

Here, we suppress the dependence of the coefficients and the jump intensity and density on the parameter θ and current state $Y(t-)$ and assume that μ and σ are predictable processes. We also assume that high-frequency data for Y is available and is denoted by $\{y_{i,t}\}$, where $y_{i,t}$ is the ith intraday observation on trading day t. Assume that for each trading day, there are M available intraday observations.

One non-parametric approach is to consider the squared increments of Y as was done by Mancini (2004). Specifically, Mancini (2004) estimates that a jump occurred between the ith and $(i-1)$th

observations $y_{i-1,t}$ and $y_{i,t}$ if the squared increment $(\Delta y_{i,t})^2 = (y_{i,t} - y_{i-1,t})^2$, which is also the squared return if Y represents the log-price of an asset, exceeds a threshold $g(\delta)$. Here, δ is the sampling frequency or time between the ith and $(i-1)$th observations (i.e. $\delta = 1/M$) and g is a positive function such that $g(\delta) \to 0$ as $\delta \to 0$. Mancini (2004) chose $g(\delta) = 8M^2\beta\delta\log(1/\delta)$, where $\beta \in (1, 2]$ and $M > 0$ is an upper bound of μ and σ, i.e. $|\mu(t)| \leq M$ and $|\sigma(t)| \leq M$ a.s., although any other function of δ that goes to 0 slower than g may be used. One can then formulate a consistent estimator of the number of jumps that occur on the trading day t based on g.

Another non-parametric approach for estimating the jump components is that which uses the bipower variation (BV) introduced by Barndorff-Nielsen and Shephard (2004) as an extension of the realized variance (RV). Letting $r_{i,t} = y_{i,t} - y_{i-1,t}$, $i = 1, \ldots, M$ denote the increments of y, the realized variance and bipower variance observed on trading day t are given by

$$\mathrm{RV}_t = \sum_{i=1}^{M} r_{i,t}^2, \qquad \mathrm{BV}_t = \frac{\pi}{2}\left(\frac{M}{M-1}\right)\sum_{i=1}^{M} |r_{i-1,t}||r_{i,t}|.$$

As $\delta \to 0$, it is known that RV_t converges in probability to $[Y](t) - [Y](t-1)$, where $[Y]$ is the quadratic variation of Y which, given that Y satisfies (E.7), is

$$[Y](t) = \int_0^t \sigma^2(s)\,\mathrm{d}s + \int_0^t \int_{\mathbb{R}_0} \gamma^2(u)p(\mathrm{d}u, \mathrm{d}s).$$

Thus, RV_t is a consistent estimator of the integrated variance observed between trading days t and $t-1$ plus any jump contributions. However, Barndorff-Nielsen and Shephard (2004) show that as $\delta \to 0$, BV_t converges in probability to $\int_{t-1}^t \sigma^2(s)\,\mathrm{d}s$, meaning that the bipower variation is a jump-robust consistent estimator the integrated variance of Y. It follows therefore that $\mathrm{RV}_t - \mathrm{BV}_t$ is a consistent estimator of the pure jump part of the quadratic variation of Y Barndorff-Nielsen and Shephard (2004). Based on this, Huang and Tauchen (2005) developed a jump detection test using the test statistic

$$\mathrm{ZJ}_t = \frac{(\mathrm{RV}_t - \mathrm{BV}_t)/\mathrm{RV}_t}{\sqrt{((\frac{\pi}{2})^2 + \pi - 5)\delta\max\{1, \mathrm{TP}_t/\mathrm{BV}_t^2\}}},$$

where TP_t is the tripower quarticity defined as

$$\mathrm{TP}_t = M\Upsilon_{4/3}^{-3}\frac{M}{M-1}\sum_{i=3}^{M} |r_{i-2,t}|^{4/3}|r_{i-1,t}|^{4/3}|r_{i,t}|^{4/3},$$

where $\Upsilon_k = 2^{k/2}\frac{\Gamma((k+1)/2)}{\Gamma(1/2)}$, $k > 0$, and $\Gamma(\cdot)$ is the gamma function. As $\delta \to 0$, ZJ_t converges in law to the standard normal distribution.

The non-parametric jump detection test of Huang and Tauchen (2005) is one of many that are based on high-frequency data. Another is that of Aït-Sahalia and Jacod (2009) which applies to any Itô semimartingale. For a survey of related non-parametric jump detection tests, see Bjursell and Gentle (2012) and Dumitru and Urga (2012).

For the multi-dimensional case, we are also interested in whether jumps in two processes that occur at the same time are joint or disjoint. The presence of joint jumps in two or more stochastic processes may point to a common risk factor that causes these processes to jump at the same time, possibly also with correlated jump magnitudes. To this end, tests for joint jumps have also been developed (see for example, Bollerslev et al., 2008; Jacod and Todorov, 2009).

E.2 Calibration

In this section, we give an overview of the basic ideas of model calibration, following the discussion of Cont and Tankov (2004b). As Cont and Tankov (2004b) focuses on the calibration of single-asset pricing models driven by Lévy processes, we also tackle issues concerning the calibration of multi-asset pricing models. To more concretely illustrate the idea of calibration, we refer to the

two-asset Bates (1996) model used in Chapter 2.

E.2.1 Main Ideas

As evident in the previous section, parameter estimation is performed to determine the parameters θ of the asset pricing model (E.1) so that it best replicates the time series of historical observations $\{\mathbf{y}_j\}_{j=1}^N$. However, since markets under which prices jump and exhibit stochastic volatility are incomplete markets, historical asset prices are insufficient to uniquely determine option prices. Arbitrage-free option prices are defined as expectations under an equivalent martingale measure (EMM) $\mathbb{Q}$, but in incomplete markets the EMM is not unique. Thus, in addition to correctly characterizing the historical price process (E.1), we must also "choose" the EMM $\mathbb{Q}$ so that the option prices generated by the model are consistent with option prices observed in the market. This is the goal of *model calibration*.

The primary objective of model calibration is to determine the EMM $\mathbb{Q}$ under which the price of *benchmark options*, as determined by the option pricing model or equation, match their market prices. Benchmark options are typically chosen to be vanilla call options on the individual assets or actively traded options on baskets of these assets over a collection of strike prices and maturities. Once the model has been calibrated, it can then be used to price other options or derivatives, such as exotic options and non-traded or illiquid derivatives, or to compute hedge ratios.

The asset pricing model (2.1)-(2.2) is parameterized by

$$\theta = \left\{\mu_1, \mu_2, \lambda_1, \lambda_2, m_{\mathbb{P}}^{(1)}(\mathrm{d}y), m_{\mathbb{P}}^{(2)}(\mathrm{d}y), \xi_1, \xi_2, \eta_1, \eta_2, \sigma_1, \sigma_2, \boldsymbol{\Sigma}\right\}.$$

We leave the jump size densities $m_{\mathbb{P}}^{(1)}(\mathrm{d}y)$ and $m_{\mathbb{P}}^{(2)}(\mathrm{d}y)$ unspecified, but for a fully parametric model we require that these densities are characterized as a parametric probability density function. The historical parameters θ can be estimated with a fully parametric or semi-parametric approach using any of the methods able to deal with observable (stock prices) and latent (instantaneous variances) components as discussed in the previous section.

Proposition 2.3 states that we can construct a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ with the Radon-Nikodým derivative $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = L(T)$, where L is the process defined by equation (2.8). From the definition of L, we say that the measure $\mathbb{Q}$ is parameterized by $\vartheta = \{\{\boldsymbol{\theta}(t)\}_{t\in[0,T]}, \gamma_1, \gamma_2, \nu_1, \nu_2\}$, where $\{\boldsymbol{\theta}(t)\}_{t\in[0,T]}$ is a four-dimensional predictable process and $\gamma_1, \gamma_2, \nu_1, \nu_2$ are constants. For $\mathbb{Q}$ to be an EMM, i.e. the measure under which the discounted yield processes $\{e^{-(r-q_i)t}S_i(t)\}_{t\in[0,T]}$ for $i = 1, 2$ are martingales, $\boldsymbol{\theta}$ must be given by equation (2.17), which is jointly parameterized by θ, $\{\gamma_1, \gamma_2, \nu_1, \nu_2\}$, and two additional parameters Λ_1 and Λ_2 corresponding to the market price of volatility risk. Therefore, the EMM $\mathbb{Q}$ is jointly parameterized by θ and $\tilde{\vartheta} = \{\gamma_1, \gamma_2, \nu_1, \nu_2, \Lambda_1, \Lambda_2\}$. To emphasize this parameterization, we shall write $\mathbb{Q}(\theta, \tilde{\vartheta})$ and denote by $L(T; \theta, \tilde{\vartheta})$ the corresponding Radon-Nikodým derivative $\frac{\mathrm{d}\mathbb{Q}(\theta,\tilde{\vartheta})}{\mathrm{d}\mathbb{P}} = L(T; \theta, \tilde{\vartheta})$ such that $\mathbb{Q}(\theta, \tilde{\vartheta})$ is an EMM.[1] For the remainder of this section, we shall only consider EMMs $\mathbb{Q}(\theta, \tilde{\vartheta})$.

Since the parameter estimation process is only able to determine θ so that the model (2.1)-(2.2) best captures the time series of historical prices for S_1 and S_2, model calibration must be performed to determine $\tilde{\vartheta}$ so that a collection $\mathcal{I}$ of benchmark options with payoffs $\mathsf{C}_i^{\text{payoff}}(S_1(\mathsf{T}_i), S_2(\mathsf{T}_i))$, maturities T_i and current market prices $\mathsf{C}_i^{\text{mkt}}$ are properly priced by the model; that is, we must solve for $\tilde{\vartheta}$ such that

$$\mathsf{C}_i^{\text{mkt}} = \mathsf{C}_i^{\text{mod}}(\theta, \tilde{\vartheta}) \triangleq \mathbb{E}_{\mathbb{Q}(\theta,\tilde{\vartheta})}\left[e^{-r\mathsf{T}_i}\mathsf{C}_i^{\text{payoff}}(S_1(\mathsf{T}_i), S_2(\mathsf{T}_i))\big|S_1(0), S_2(0), v_1(0), v_2(0)\right], \qquad \forall i \in \mathcal{I} \tag{E.8}$$

where $\mathsf{C}_i^{\text{mod}}(\theta, \tilde{\vartheta})$ is the model price of the option i defined as an expectation in the EMM $\mathbb{Q}(\theta, \tilde{\vartheta})$.

[1] As emphasized in Chapter 2, any probability measure equivalent to the historical probability measure $\mathbb{P}$ can be defined via the Radon-Nikodým derivative $\frac{\mathrm{d}\mathbb{Q}(\vartheta)}{\mathrm{d}\mathbb{P}} = L(T; \vartheta)$, where $L(t, \vartheta)$ is the process L in equation (2.8) parameterized by $\vartheta = \{\{\boldsymbol{\theta}(t)\}_{t\in[0,T]}, \gamma_1, \gamma_2, \nu_1, \nu_2\}$. The specific parameterization $\mathbb{Q}(\theta, \tilde{\vartheta})$ indicates that $\mathbb{Q}$ is chosen to be an equivalent martingale measure, where the choice of ϑ is guided by the constraint on the parameter process $\{\boldsymbol{\theta}(t)\}_{t\in[0,T]}$ introduced by (2.17).

In the formulation above, it is understood that the model option prices may also depend on a number of inputs, such as the current asset prices.

However, the calibration problem (E.8) is often ill-posed: there may be multiple EMMs which generate the same price of benchmark options, especially since the set of benchmark options is finite and often sparse (Cont and Tankov, 2004b). In addition, there may not exist a pricing model $\mathsf{C}^{\text{mod}}(\theta, \tilde{\vartheta})$ that exactly replicates observed option prices. Thus for practical reasons, the model calibration problem is often formulated as an optimization problem

$$\tilde{\vartheta} = \arg\min_{\vartheta_0} \sum_{i \in \mathcal{I}} \omega_i \left(\mathsf{C}_i^{\text{mkt}} - \mathsf{C}_i^{\text{mod}}(\theta, \vartheta_0) \right)^2, \tag{E.9}$$

where ω_i is the weight of the ith benchmark option and $\mathsf{C}_i^{\text{mod}}(\theta, \vartheta_0)$ is the model option price defined as an expectation under *some other* EMM $\mathbb{Q}_0 = \mathbb{Q}_0(\theta, \vartheta_0)$. Formulation (E.9) means that we choose $\tilde{\vartheta}$ that minimizes the squared in-sample option pricing error. Owing to the complex nature of the option pricing formula $\mathsf{C}_i^{\text{mod}}(\theta, \vartheta_0)$, the optimization problem (E.9) is generally a non-convex problem. Thus, solving (E.9) is typically computationally intensive since $\mathsf{C}_i^{\text{mod}}(\theta, \vartheta_0)$ must be calculated via Monte Carlo simulations, numerical integration, or other such methods, and because numerical optimizers may converge to a local, non-global minimum of the objective function.

Nonetheless, (E.9) is the version of the calibration problem that is often used in calibrating stochastic volatility and jump-diffusion asset pricing models, as was done by Bates (1996) and Duffie et al. (2000), among others. Working with jump-diffusions with state-dependent local volatility, Andersen and Andreasen (2000) separated the fitting of jump parameters and the construction of the volatility surface. Jump parameters were fitted using a least-squares criterion with respect to Black and Scholes (1973) implied volatility then constructed the volatility surface with respect to Merton (1976) implied volatilities. He et al. (2006) also used a least-squares criterion for calibration jump-diffusions with local volatility, but included American options in the collection of benchmark options, so the calibration becomes a variational least-squares calibration problem. In addition to the squared in-sample option pricing error, Andersen et al. (2015) also included in the objective function a measure of fit on the volatility, with particular focus on the model's ability to also fit the Black and Scholes (1973) implied volatility. Cont and Tankov (2004b) and Christoffersen et al. (2009) suggest using the option vega as weights ω_i so that the pricing error is related to the option bid-ask spread.

Cont and Tankov (2004b) contends that the calibration problem (E.9) is still ill-posed in the sense that uniqueness and stability issues prevail. As such, they propose adding a regularization term equal to the relative entropy or the Kullback-Leibler distance $\mathcal{D}(\mathbb{Q}_0(\theta, \vartheta_0)|\mathbb{P})$ between the candidate EMM $\mathbb{Q}_0(\theta, \vartheta_0)$ and the historical measure $\mathbb{P}$,

$$\mathcal{D}(\mathbb{Q}_0(\theta, \vartheta_0)|\mathbb{P}) = \mathbb{E}_{\mathbb{Q}_0(\theta,\vartheta_0)} \left[\ln \frac{d\mathbb{Q}_0(\theta, \vartheta_0)}{d\mathbb{P}} \right] = \mathbb{E}_{\mathbb{Q}}[\ln L(T; \theta, \vartheta_0)] \triangleq K(\vartheta_0).$$

Thus, the regularized calibration problem

$$\tilde{\vartheta} = \arg\min_{\vartheta_0} \left[\alpha K(\vartheta_0) + \sum_{i \in \mathcal{I}} \omega_i \left(\mathsf{C}_i^{\text{mkt}} - \mathsf{C}_i^{\text{mod}}(\theta, \vartheta_0) \right)^2 \right], \tag{E.10}$$

seeks not only to minimize the squared in-sample option pricing error but also the amount of information that needs to be added to the historical measure $\mathbb{P}$ to correctly reproduce observed option prices. Viewed in this perspective, model calibration thus augments parameter estimation in that calibration corrects the initial knowledge of the asset price dynamics, as reflected by the historical parameters θ, so that available option prices can be accurately reproduced (Cont and Tankov, 2004b). A suitable prior measure $\mathbb{Q}_0$ that reflects reasonable parameter values may be chosen instead of the historical measure $\mathbb{P}$ if it is unavailable or unreliable. The weight α is ideally selected based on the available option price data, as explained by Cont and Tankov (2004b).

The parameter estimation and calibration problem may also be addressed via the stochastic filtering method. The stochastic filtering method assumes that the market price of risk associated to the equivalent martingale measure $\mathbb{Q}$ (see for example the process $\boldsymbol{\theta}$ defined in Remark 2.17) evolves dynamically according to some stochastic differential equation. Using historical information on stock prices and derivative prices, one then aims to compute the conditional filter distribution of the market price of risk. This is done by constructing the filter model consisting of the dynamics of the market price of risk, the dynamics of the stock prices, and the dynamics of the derivative prices under the historical probability measure. The filter model is then estimated using maximum likelihood approaches. For an overview of the stochastic filtering method and its application to derivatives pricing, see for example Runggaldier (2004).

E.2.2 Calibrating Multi-Asset Models

When dealing with multi-asset pricing models, the calibration procedure must also be able to incorporate market information about the dependency structure among the individual underlying assets, which can be done by including multi-asset options in the set of benchmark options. However, there seems to be very little to no available data for multi-asset options since these options are typically considered exotic or illiquid. In addition, it is also possible that the underlying assets themselves may not be actively traded, thus limiting the amount of market information available.

Given the general unavailability of multi-asset option price data, Leoni and Schoutens (2007), who sought to calibrate a multivariate asset pricing model based on variance gamma (VG) processes (Madan et al., 1998), proposed two alternative calibration methods. The first approach assumes that market information is not available for all underlying assets but only for a basket of these assets or an index written on these assets. In this case, calibration is numerically performed using a double bootstrapping method. The second approach assumes implied market data (for example, implied volatilities, single-asset options) are available for each underlying. In this case, each component of the multi-asset pricing model is calibrated with respect to the corresponding collection of benchmark options via the Carr and Madan (1999) FFT option pricing formula. Historical information on the joint behavior of assets in the form of correlation matrices estimated from historical asset price time series are then incorporated after the fact. Luciano and Schoutens (2006) also considered a multidimensional asset pricing model based on VG processes but assumed that the Brownian components are uncorrelated. In their case, calibration can then be done individually per asset.

For exchange options written on two assets S_1 and S_2 with GBM or lognormal dynamics (i.e. the Margrabe (1978) model specification), Alexander and Venkatramanan (2009) suggested that the volatility σ_1 of S_1 be set equal to the implied volatility of a call option on S_1 with strike S_2 and the volatility σ_2 of S_2 be set equal to the implied volatility of a call option on S_2 with strike S_1. Cross-asset correlations can then be calibrated from any liquid spread option or basket option on S_1 and S_2, if available, using volatilities calibrated as stated.

For a multi-asset Heston (1993) model (composed of correlated single-asset Heston (1993) models), Dimitroff et al. (2011) proposed a calibration procedure similar to the second approach of Leoni and Schoutens (2007). That is, each component single-asset model is first calibrated to vanilla benchmark options using any available calibration approach and the correlation between the Brownian components are then derived from the empirical cross-asset correlations calculated from historical time series of asset prices. This is justified since the correlation of the Brownian components is invariant under an equivalent measure change (see for example, Section 2.3). The method of calibrating the correlation of the Brownian components proposed by Dimitroff et al. (2011) involves the minimization of the difference between the empirical cross-asset correlation matrix and an estimator of the true cross-correlation matrix, which is a function of the unknown correlation matrix of Brownian components.

Bibliography

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, ninth Dover printing, tenth GPO printing edition, 1964.

Y. Achdou and N. Tchou. Variational analysis for the Black and Scholes equation with stochastic volatility. *ESAIM Mathematical Modelling and Numerical Analysis*, 36(3):373–395, 2002.

T. Adolfsson, C. Chiarella, A. Ziogas, and J. Ziveyi. Representation and numerical approximation of American option prices under Heston stochastic volatility dynamics. *Quantitative Finance Research Center Research Paper 327, University of Technology Sydney*, 2013.

Y. Aït-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, 70(1):223–262, 2002.

Y. Aït-Sahalia. Disentangling diffusion from jumps. *Journal of Financial Economics*, 74:487–528, 2004.

Y. Aït-Sahalia. Closed-form likelihood expansions for multivariate diffusions. *The Annals of Statistics*, 36(2):906–937, 2008.

Y. Aït-Sahalia and J. Jacod. Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1):184–222, 2009.

Y. Aït-Sahalia and J. Yu. Saddlepoint approximations for continuous-time Markov processes. *Journal of Econometrics*, 134:507–551, 2007.

Y. Aït-Sahalia, J. Cacho-Diaz, and R. J. A. Laeven. Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117:585–606, 2015.

Y. Aït-Sahalia, M. Karaman, and L. Mancini. The term structure of equity and variance risk premia. *Swiss Finance Institute Research Paper Series No. 18-37*, 2018.

C. Alexander and A. Venkatramanan. Analytic approximations for multi-asset option pricing. *ICMA Centre Discussion Papers in Finance DP2009-05*, 2009. URL https://ssrn.com/abstract=1424985.

M. Alfeus and E. Schlögl. On spread option pricing using two-dimensional Fourier transform. *International Journal of Theoretical and Applied Finance*, 22(5):1–20, 2019.

W. Allegretto, G. Barone-Adesi, and R. J. Elliott. Numerical evaluation of the critical price and American options. *Journal of Computational Finance*, 3:33–52, 1995.

A. Almendral and C. W. Oosterlee. Numerical valuation of options with jumps in the underlying. *Applied Numerical Mathematics*, 53:1–18, 2005.

E. Alòs and T. Rheinlander. Pricing and hedging Margrabe options with stochastic volatilities. *Economic Working Papers 1475, Department of Economics and Business, Universitat Pompeu Fabra*, 2017.

D. Amengual and D. Xiu. Resolution of policy uncertainty and sudden declines in volatility. *Journal of Econometrics*, 203:297–315, 2018.

L. Andersen and J. Andreasen. Jump diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4:231–262, 2000.

L. B. G. Andersen and V. V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11:29–50, 2007.

T. G. Andersen, L. Benzoni, and J. Lund. An empirical investigation of continuous-time equity return models. *The Journal of Finance*, 57(3):1239–1284, 2002.

T. G. Andersen, N. Fusari, and V. Todorov. The risk premia embedded in index options. *Journal of Financial Economics*, 117:558–584, 2015.

F. Antonelli and S. Scarlatti. Exchange option pricing under stochastic volatility: a correlation expansion. *Finance and Stochastics*, 13(2):269–303, 2010.

F. Antonelli, A. Ramponi, and S. Scarlatti. Pricing options under stochastic volatility: a power series approach. *Review of Derivatives Research*, 13(1):45–73, 2009.

M. R. Armada, L. Kryzanowski, and P. J. Pereira. A modified finite-lived American exchange option methodology applied to real options valuation. *Global Finance Journal*, 17(3):419–438, 2007.

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

L. Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure*, 3e série, 17:21–86, 1900. doi: 10.24033/asens.476. URL http://www.numdam.org/item/ASENS_1900_3_17__21_0.

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

G. Bakshi, C. Cao, and Z. Zhong. Assessing models of individual equity option prices. *Review of Quantitative Finance and Accounting*, 2021. URL https://doi.org/10.1007/s11156-020-00951-4.

C. A. Ball and A. Roma. Stochastic volatility option pricing. *The Journal of Financial and Quantitative Analysis*, 29(4):589–607, 1994.

C. A. Ball and W. N. Torous. A simplified jump process for common stock returns. *The Journal of Financial and Quantitative Analysis*, 18(1):53–65, 1983.

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

W. Bao, J. Yue, and Y. Rao. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE*, 12(7), 2017.

C. Bardgett, E. Gourier, and M. Leippold. Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *Journal of Financial Economics*, 131:593–618, 2019.

O. E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2: 41–68, 1998.

O. E. Barndorff-Nielsen and N. Shephard. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37, 2004.

G. Barone-Adesi. The saga of the American put. *Journal of Banking and Finance*, 29:2909–2918, 2005.

G. Barone-Adesi and R. E. Whaley. Efficient analytic approximation of American option values. *The Journal of Finance*, 17(2):301–320, 1987.

D. S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *The Review of Financial Studies*, 9(1):69–107, 1996.

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

D. S. Bates. Maximum likelihood estimation of latent affine processes. *The Review of Financial Studies*, 19(3):909–965, 2006.

C. Bayer, P. Friz, and J. Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6): 887–904, 2016.

C. Bayer, R. Tempone, and S. Wolfers. Pricing American options by exercise rate optimization. *Quantitative Finance*, 2020. URL https://doi.org/10.1080/14697688.2020.1750678.

D. Becherer. Utility-indifference hedging and valuation via reaction-diffusion systems. *Proceedings of the Royal Society A*, 460:27–51, 2004.

C. Beck, W. E, and A. Jentzen. Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations. *Journal of Nonlinear Science*, 29:1563–1619, 2019.

J.-F. Bégin and M. Boudreault. Likelihood evaluation of jump-diffusion models using deterministic nonlinear filters. *Journal of Computational and Graphical Statistics*, 2020. URL https://doi.org/10.1080/10618600.2020.1840995.

N. Bellamy and M. Jeanblanc. Incompleteness of markets driven by a mixed diffusion. *Finance and Stochastics*, 4:209–222, 1999.

S. Benninga, T. Björk, and Z. Wiener. On the use of numéraires in option pricing. *The Journal of Derivatives*, 10(2):43–58, 2002.

A. Bensoussan. On the theory of option pricing. *Acta Applicandae Mathematica*, 2(2):139–158, 1984.

A. Bensoussan and J. L. Lions. *Applications of Variational Inequalities in Stochastic Control*. North Holland, 1982.

F. E. Benth. Option pricing: General principles. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Inc., 2010.

F. E. Benth and H. Zdanowicz. Pricing energy spread options. In A. Roncoroni, G. Fusai, and M. Cummins, editors, *Handbook of Multi-Commodity Markets and Products: Structuring, Trading and Risk Management*, chapter 17, pages 801–825. John Wiley & Sons, Ltd., West Sussex, UK, 2015.

F. E. Benth, G. Di Nunno, A. Løkka, B. Øksendal, and F. Proske. Explicit representation of the minimal variance portfolio in markets driven by Lévy processes. *Mathematical Finance*, 13(1): 55–72, 2003.

C. Bernard and Z. Cui. A note on exchange options under stochastic interest rates, 2010. URL https://ssrn.com/abstract=1626020.

L. Bertini and L. Passalacqua. Modelling interest rates by correlated multi-factor CIR-like processes, 2008. URL https://arxiv.org/abs/0807.3898v1.

S. Biagini, M. Frittelli, and M. Grasselli. Indifference price with general semimartingales. *Mathematical Finance*, 21(3):423–446, 2010.

P. Bjerskund and G. Stensland. American exchange options and a put-call transformation: a note. *Journal of Business Finance and Accounting*, 20(5):761–764, 1993.

P. Bjerskund and G. Stensland. Closed form spread option valuation. *Quantitative Finance*, 14 (10):1785–1794, 2011.

T. Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, New York, 2009.

T. Björk, G. Di Masi, Y. Kabanov, and W. Runggaldier. Towards a general theory of bond markets. *Finance and Stochastics*, 1:141–174, 1997a.

T. Björk, Y. Kabanov, and W. Runggaldier. Bond market structure in the presence of marked point processes. *Mathematical Finance*, 7(2):211–223, 1997b.

J. Bjursell and J. E. Gentle. Identifying jumps in asset prices. In J.-C. Duan, W. K. Hardle, and J. E. Gentle, editors, *Handbook of Computational Finance*, chapter 14, pages 371–399. Springer-Verlag, 2012.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

L. P. Blenman and S. P. Clark. Power exchange options. *Finance Research Letters*, 2:97–106, 2005.

T. Bollerslev, T. H. Law, and G. Tauchen. Risk, jumps, and diversification. *Journal of Econometrics*, 144:234–256, 2008.

S. I. Boyarchenko and S. Z. Levendorskiĭ. *Non-Gaussian Merton-Black-Scholes Theory*. World Scientific, Singapore, 2002.

P. Brandimarte. *Numerical Methods in Finance and Economics: A MATLAB-Based Introduction*. John Wiley & Sons, Inc., 2nd edition, 2006.

M. W. Brandt and P. Santa-Clara. Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. *Journal of Financial Economics*, 63:161–210, 2002.

P. Brémaud. *Point Processes and Queues: Martingale Dynamics*. Springer-Verlag, 1981.

M. Briani, C. La Chioma, and R. Natalini. Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory. *Numerische Mathematik*, 98:607–646, 2004.

M. Broadie and J. Detemple. American option valuation: new bounds, approximations, and a comparison of existing methods. *The Review of Financial Studies*, 9(4):1211–1250, 1996.

M. Broadie and J. Detemple. The valuation of American options on multiple assets. *Mathematical Finance*, 7(3):241–286, 1997.

A. Brouste, M. Fukasawa, H. Hino, S. M. Iacus, K. Kamatani, Y. Koike, H. Masuda, R. Nomura, T. Ogihara, Y. Shimuzu, M. Uchida, and N. Yoshida. The yuima project: A computational framework for simulation and inference of stochastic differential equations. *Journal of Statistical Software*, 57(4):1–51, 2014. doi: 10.18637/jss.v057.i04. URL http://www.jstatsoft.org/v57/i04/.

R. Caldana and G. Fusai. A general closed-form spread option pricing formula. *Journal of Banking and Finance*, 37:4893–4906, 2013.

R. Caldana, G. H. L. Cheang, C. Chiarella, and G. Fusai. Correction: Exchange options under jump-diffusion dynamics. *Applied Mathematical Finance*, 22(1):99–103, 2015.

M. Cane and P. Olivares. Pricing spread options under stochastic correlation and jump-diffusion models, 2014. URL https://arxiv.org/abs/1409.1175.

R. Carmona and M. Coulon. A survey of commodity market and structural models for electricity prices. In F. E. Benth, V. A. Kholodnyi, and P. Laurence, editors, *Quantitative Energy Finance: Modeling, Pricing, and Hedging in Energy and Commodity Markets*, chapter 2, pages 41–83. Springer Science + Business Media, New York, 2012.

R. Carmona and V. Durrleman. Pricing and hedging spread options. *SIAM Review*, 45(4):627–685, 2003a.

R. Carmona and V. Durrleman. Pricing and hedging basket options in a log-normal model. *Technical Report, Department of Operations Research and Financial Engineering, Princeton University*, 2003b.

P. Carr. The valuation of American exchange options with application to real options. In L. Trigeorgis, editor, *Real Options in Capital Investment: Models, Strategies, and Applications*, pages 109–120. Praeger, Westport, CT, 1995.

P. Carr and A. Hirsa. Why be backward? Forward equations for American options. *Risk*, pages 103–107, 2003.

P. Carr and D. B. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999.

P. Carr and L. Wu. A tale of two indices. *The Journal of Derivatives*, 13(3):13–29, 2006.

P. Carr, R. Jarrow, and R. Myneni. Alternative characterizations of American put options. *Mathematical Finance*, 2(2):87–106, 1992.

P. Carr, H. Geman, D. B. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75(2):305–332, 2002.

P. Carr, H. Geman, D. B. Madan, and M. Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13(3):345–382, 2003.

R. C. Cavalcante, R. C. Brasileiro, V. L. F. Souza, J. P. Nobrega, and A. L. I. Oliveira. Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, 55:194–211, 2016.

Cboe Global Markets Inc. White paper: Cboe volatility index. http://www.cboe.com/micro/vix/vixwhite.pdf, 2019.

A. Cerny. *Mathematical Techniques in Finance: Tools for Incomplete Markets*. Princeton University Press, 2nd edition, 2009.

G. H. L. Cheang and C. Chiarella. Exchange options under jump-diffusion dynamics. *Applied Mathematical Finance*, 18(3):245–276, 2011.

G. H. L. Cheang and C. Chiarella. A modern view on merton's jump-diffusion model. In S. N. Cohen, D. Madan, T. K. Siu, and H. Yang, editors, *Stochastic Processes, Finance and Control: A Festschrift in Honor of Robert J. Elliot*, chapter 9, pages 217–234. World Scientific, 2012.

G. H. L. Cheang and G. Lian. Perpetual exchange options under jump-diffusion dynamics. *Applied Mathematical Finance*, 22(5):450–462, 2015.

G. H. L. Cheang and G.-A. Teh. Change of numéraire and a jump-diffusion option pricing formula. In R. Dieci, X.-Z. He, and C. Hommes, editors, *Nonlinear Economic Dynamics and Financial Modelling: Essays in Honour of Carl Chiarella*, pages 371–389. Springer International Publishing, Cham, 2014.

G. H. L. Cheang, C. Chiarella, and A. Ziogas. On exchange options with jumps. In M. Holder, editor, *Proceedings of the Third IASTED International Conference on Financial Engineering and Applications (FEA 2006)*, pages 104–109, Calgary, Canada, Oct. 2006. ACTA Press.

G. H. L. Cheang, C. Chiarella, and A. Ziogas. An analysis of American options under Heston stochastic volatility and jump-diffusion dynamics. *Quantitative Finance Research Center Research Paper 256*, 2011.

G. H. L. Cheang, C. Chiarella, and A. Ziogas. The representation of American option prices under stochastic volatility and jump-diffusion dynamics. *Quantitative Finance*, 13(2):241–253, 2013.

A. Chen and M. Suchanecki. Parisian exchange options. *Quantitative Finance*, 11(8):1207–1220, 2011.

Y. Cheng, M. Escobar-Anel, and Z. Gong. Generalized mean-reverting 4/2 factor model. *Journal of Risk and Financial Management*, 12(4):1–21, 2019.

M. Chernov, A. R. Gallant, E. Ghysels, and G. Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics*, 116:225–257, 2003.

C. Chiarella and A. Ziogas. McKean's method applied to American call options on jump-diffusion processes. *Quantitative Finance Research Center Research Paper 117, University of Technology Sydney*, 2004.

C. Chiarella and A. Ziogas. Evaluation of American strangles. *Journal of Economic Dynamics and Control*, 29:31–62, 2005.

C. Chiarella and A. Ziogas. American call options under jump-diffusion processes: A Fourier transform approach. *Applied Mathematical Finance*, 16(1):37–79, 2009.

C. Chiarella and J. Ziveyi. Method of lines approach for pricing American spread options. https://dx.doi.org/10.2139/ssrn.2019353, 2011.

C. Chiarella and J. Ziveyi. American option pricing under two stochastic volatility processes. *Applied Mathematics and Computation*, 224:283–310, 2013.

C. Chiarella and J. Ziveyi. Pricing American options written on two underlying assets. *Quantitative Finance*, 14(3):409–426, 2014.

C. Chiarella, B. Kang, G. H. Meyer, and A. Ziogas. 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*, 13(3):393–425, 2009.

C. Chiarella, A. Ziogas, and J. Ziveyi. Representation of American option prices under Heston stochastic volatility dynamics using integral transforms. In C. Chiarella and A. Novikov, editors, *Contemporary Quantitative Finance*, pages 281–315. Springer-Verlag, 2010.

C. Chiarella, B. Kang, and G. H. Meyer. *The Numerical Solution of the American Option Pricing Problem: Finite Difference and Transform Approaches*. World Scientific, Singapore, 2015.

C. Chiarella, C. Sklibosios Nikotopoulos, E. Schlögl, and H. Yang. Pricing American options under regime switching using the method of lines. *Quantitative Finance Research Center Research Paper 368, University of Technology Sydney*, 2016.

P. Christoffersen, S. Heston, and K. Jacobs. The shape and term structure of the index option smirk: why multifactor stochastic volatility models work so well. *Management Science*, 55(12): 1914–1932, 2009.

P. Christoffersen, K. Jacobs, and K. Mimouni. Volatility dynamics for the S& P500: Evidence from realized volatility, daily returns, and option prices. *The Review of Financial Studies*, 23(8): 3141–3189, 2010.

N. Clarke and K. Parrott. Multigrid for American option pricing with stochastic volatility. *Applied Mathematical Finance*, 6(3):177–195, 1999.

S. N. Cohen and R. J. Elliott. *Stochastic Calculus and Applications*. Springer Science + Business Media, New York, USA, 2nd edition, 2015.

R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, Florida, 2004a.

R. Cont and P. Tankov. Non-parametric calibration of jump-diffusion option pricing models. *Journal of Computational Finance*, 7(3):1–49, 2004b.

R. Cont, P. Tankov, and E. Voltchkova. Hedging with options in models with jumps. In F. E. Benth, G. Di Nunno, T. Lindstrom, B. Oksendal, and T. Zhang, editors, *Stochastic Analysis and Applications: The Abel Symposium 2005*, pages 197–217. Springer-Verlag, 2007.

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

J. C. Cox, S. A. Ross, and M. Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

J. C. Cox, J. Jonathan E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

C. Cuchiero, M. Keller-Ressel, and J. Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16:711–740, 2012.

N. Cufaro-Petroni and P. Sabino. Pricing exchange options with correlated jump diffusion processes. *Quantitative Finance*, 20(11):1811–1823, 2020.

J. Da Fonseca and K. Ignatieva. Jump activity analysis for affine jump-diffusion models: Evidence from the commodity market. *Journal of Banking and Finance*, 99:45–62, 2019.

B. Davies. *Integral Transforms and their Applications*. Springer Science + Business Media, LLC, New York, 2nd edition, 1985.

M. H. A. Davis. Options: Basic definitions. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Inc., 2010a.

M. H. A. Davis. Black-scholes formula. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Inc., 2010b.

D. De Vallière, Y. Kabanov, and E. Lépinette. Consumption-investment problem with transaction costs for Lévy-driven price processes. *Finance and Stochastics*, 20:705–740, 2016.

L. Debnath. The double Laplace transforms and their properties with applications to functional, integral, and partial differential equations. *International Journal of Applied and Computational Mathematics*, 2:223–241, 2016.

F. Delbaen and W. Schachermayer. The variance-optimal martingale measure for continuous processes. *Bernoulli*, 2(1):81–105, 1996.

F. Delbaen, P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer, and C. Stricker. Exponential hedging and entropic penalties. *Mathematical Finance*, 12(2):99–123, 2002.

M. A. H. Dempster and S. S. G. Hong. Spread option valuation and the fast Fourier transform. In H. Geman, D. Madan, S. R. Pliska, and T. Vorst, editors, *Mathematical Finance: Bachelier Congress 2000*, pages 203–220. Springer-Verlag, New York, 2002.

M. A. H. Dempster and K. Tang, editors. *Commodities*, Boca Raton, FL, 2016. Taylor & Francis Group, LLC.

S. Deng, M. Li, and J. Zhou. Closed-form approximations for spread option prices and greeks. *Journal of Derivatives*, 15(3):58–80, 2008.

E. Derman and I. Kani. Riding on a smile. *Risk Magazine*, 7:32–39, 1994.

Y. d'Halluin, P. A. Forsyth, and G. Labahn. A penalty method for American options with jump diffusion processes. *Numerische Mathematik*, 97:321–352, 2004a.

Y. d'Halluin, P. A. Forsyth, and K. R. Vetzal. Robust numerical methods for contingent claims under jump diffusion processes. *Cornell Theory Center (CTC) Research Report*, 2004b.

G. Dimitroff, S. Lorenz, and A. Szimayer. A parsimonious multi-asset Heston model: calibration and derivative pricing, 2011. URL https://ssrn.com/abstract=1435199.

J.-C. Duan and C.-Y. Yeh. Jump and volatility risk premiums implied by VIX. *Journal of Economic Dynamics & Control*, 34:2232–2244, 2010.

D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

D. G. Duffy. *Advanced Engineering Mathematics with MATLAB*. CRC Press, Florida, 3rd edition, 2011.

D. J. Duffy. *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. John Wiley & Sons, Inc., England, 2006.

A.-M. Dumitru and G. Urga. Identifying jumps in financial assets: a comparison between nonparametric jump tests. *Journal of Business & Economic Statistics*, 30(2):242–255, 2012.

B. Dupire. Pricing with a smile. *Risk Magazine*, 7:18–20, 1994.

G. B. Durham and A. R. Gallant. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business & Economic Statistics*, 20(3):297–338, 2002.

R. Durrett. *Stochastic Calculus: A Practical Introduction*. CRC Press, Inc., Florida, 1996.

W. E, J. Han, and A. Jentzen. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 5:349–380, 2017.

E. Eberlein. Jump-type Lévy processes. In T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch, editors, *Handbook of Financial Time Series*, pages 439–455. Springer-Verlag, 2009.

E. Eberlein and J. Jacod. On the range of option prices. *Finance and Stochastics*, 1:131–140, 1997.

E. Eberlein and A. Papapantoleon. Symmetries and pricing of exotic options in Lévy models. In A. Kyprianou, W. Schoutens, and P. Wilmott, editors, *Exotic Option Pricing and Advanced Lévy Models*, pages 99–128. John Wiley & Sons, Ltd., 2005.

E. Eberlein, U. Keller, and K. Prause. New insights into smile, mispricing, and value at risk: the hyperbolic model. *Journal of Business*, 71(3):371–405, 1998.

E. Ekström and J. Tysk. Properties of option prices in models with jumps. *Mathematical Finance*, 17(3):381–397, 2007a.

E. Ekström and J. Tysk. Convexity preserving jump-diffusion models for option pricing. *Journal of Mathematical Analysis and Applications*, 330:715–728, 2007b.

E. Ekström, P. Lötstedt, and J. Tysk. Boundary values and finite difference methods for the single factor term structure equation. *Applied Mathematical Finance*, 16(3):253–259, 2009.

N. El Karoui and M.-C. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal of Control and Optimization*, 33(1):29–66, 1995.

O. Elerian. A note on the existence of a closed form conditional transition density function for the Milstein scheme. *Economics Discussion Paper 1998-W18, Nuffield College, Oxford*, 1998.

O. Elerian, S. Chib, and N. Shephard. Likelihood inference for discretely observed nonlinear diffusions. *Econometrica*, 69(4):959–993, 2001.

C. M. Elliott and J. R. Ockendon. *Weak and Variational Methods for Moving Boundary Problems.* Pitman, London, 1982.

B. Eraker. Do stock prices and volatility jump? reconciling evidence from spot and option prices. *The Journal of Finance*, 59(3):1367–1403, 2004.

B. Eraker, M. Johannes, and N. Polson. The impact of jumps in volatility and returns. *The Journal of Finance*, 58(3):1269–1300, 2003.

J. Fajardo and E. Mordecki. Pricing derivatives on two-dimensional Lévy processes. *International Journal of Theoretical and Applied Finance*, 9(2):185–197, 2006.

W. Feller. Two singular diffusion problems. *Annals of Mathematics (Second Series)*, 54(1):173–182, 1951.

S. Fischer. Call option pricing when the exercise price is uncertain, and the valuation of index bonds. *The Journal of Finance*, 33(1):169–176, 1978.

G. B. Folland. *Real Analysis: Modern Techniques and Their Applications.* John Wiley & Sons, Inc., Canada, 1999.

J.-P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna. *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives.* Cambridge University Press, Cambridge, UK, 2011.

M. Frittelli. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance*, 10(1):39–52, 2000.

T. Fujiwara. From the minimal entropy martingale measures to the optimal strategies for the exponential utility maximization: The case of geometric Lévy processes. *Asia-Pacific Financial Markets*, 11:367–391, 2006.

T. Fujiwara and Y. Miyahara. The minimal entropy martingale measures for geometric Lévy processes. *Finance and Stochastics*, 7:509–531, 2003.

M. Fukasawa. Volatility has to be rough. *Quantitative Finance*, 21(1):1–8, 2020.

A. R. Gallant and G. Tauchen. Which moments to match? *Econometric Theory*, 12(4):657–681, 1996.

A. R. Gallant and G. Tauchen. Simulated score methods and indirect inference for continuous-time models. In Y. Aït-Sahalia and L. P. Hansen, editors, *Handbook of Financial Econometrics*, chapter 8, pages 427–477. North Holland, 2010.

M. J. Gander and F. Kwok. *Numerical Analysis of Partial Differential Equations using Maple and MATLAB*. Society for Industrial and Applied Mathematics, United States of America, 2018.

J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.

J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6): 933–949, 2018.

H. Geman. *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy*. John Wiley & Sons, Inc., 2005.

H. Geman, N. El Karoui, and J.-C. Rochet. Changes of numéraire, changes of probability measure and option pricing. *Journal of Applied Probability*, 32(2):443–458, 1995.

H. Geman, D. B. Madan, and M. Yor. Time changes for lévy processes. *Mathematical Finance*, 11 (1):79–96, 2001.

H. U. Gerber and E. S. W. Shiu. Option pricing by Esscher transforms. *Transactions of Society of Actuaries*, 46:99–140, 1994.

H. U. Gerber and E. S. W. Shiu. Martingale approach to pricing perpetual American options on two stocks. *Mathematical Finance*, 6(3):303–322, 1996.

R. Geske and H. E. Johnson. The American put valued analytically. *Journal of Finance*, 39: 1511–1524, 1984.

P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2003.

T. Goll and L. Rüschendorf. Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance and Stochastics*, 5:557–581, 2001.

C. Gourieroux, A. Monfort, and E. Renault. Indirect inference. *Journal of Applied Econometrics*, 8:S85–S118, 1993.

M. Grasselli. The 4/2 stochastic volatility model: a unified approach for the Heston and the 3/2 model. *Mathematical Finance*, 27(4):1013–1034, 2017.

N. Gudkov and K. Ignatieva. Electricity price modelling with stochastic volatility and jumps: an empirical investigation. *Energy Economics*, 2021. URL https://doi.org/10.1016/j.eneco.2021.105260.

C. R. Gukhal. Analytical valuation of American options on jump-diffusion processes. *Mathematical Finance*, 11(1):97–115, 2001.

J. Han, A. Jentzen, and W. E. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences of the United States of America*, 115 (34):8505–8510, 2018.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054, 1982.

J. M. Harrison and D. M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3):215 – 260, 1981.

E. G. Haug. *The Complete Guide to Option Pricing Formulas*. McGraw-Hill, 2nd edition, 2007.

E. G. Haug and J. Haug. Knock-in/out Margrabe. *Wilmott Magazine*, 1:183–204, 2002.

C. He, J. S. Kennedy, T. F. Coleman, P. A. Forsyth, Y. Li, and K. R. Vetzal. Calibration and hedging under jump diffusion. *Review of Derivatives Research*, 9:1–35, 2006.

V. Henderson and D. Hobson. Utility indifference pricing: An overview. In R. Carmona, editor, *Indifference Pricing: Theory and Applications*, chapter 2, pages 44–73. Princeton University Press, Princeton, NJ, 2009.

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

S. L. Heston. A simple new formula for options with stochastic volatility. *Technical Report, Washington University*, 1997.

N. Hilber, O. Reichmann, C. Schwap, and C. Winter. *Computational Methods for Quantitative Finance: Finite Element Methods for Derivative Pricing*. Springer, 2013.

S. D. Hodges and A. Neuberger. Optimal replication of contingent claims under transaction costs. *Review of Futures Markets*, 8(2):223–242, 1989.

K. Hornik, M. Stinchcombe, and H. White. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3:551–560, 1990.

B. Horvath, A. Muguruza, and M. Tomas. Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models. *Quantitative Finance*, 21(1):11–27, 2021.

J. Huang and J.-S. Pang. Option pricing and linear complementarity. *Journal of Computational Finance*, 2(1):31–60, 1998.

X. Huang and G. Tauchen. The relative contribution of jumps to total price variance. *Journal of Financial econometrics*, 3(4):456–499, 2005.

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

J. C. Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, USA, 8th edition, 2012.

T. R. Hurd and Z. Zhou. A Fourier transform method for spread option pricing. *SIAM Journal of Financial Mathematics*, 1:142–157, 2010.

S. M. Iacus and N. Yoshida. *Simulation and Inference for Stochastic Processes with YUIMA: A Comprehensive `R` Framework for SDEs and Other Stochastic Processes*. Springer, 2018.

A. Ibáñez and F. Zapatero. Monte Carlo valuation of American options through computation of the optimal exercise frontier. *Journal of Financial and Quantitative Analysis*, 39(2):253–275, 2004.

S. Ikonen and J. Toivanen. Operator splitting methods for American option pricing. *Applied Mathematics Letters*, 17:809–814, 2004.

S. Ikonen and J. Toivanen. Efficient numerical methods for pricing American options under stochastic volatility. *Numerical Methods for Partial Differential Equations*, 24:104–126, 2007.

S. Ikonen and J. Toivanen. Operator splitting methods for pricing American options under stochastic volatility. *Numerische Mathematik*, 113:299–324, 2009.

A. Itkin. Deep learning calibration of option pricing models: some pitfalls and solutions, 2019. URL https://arxiv.org/abs/1906.03507.

S. D. Jacka. Optimal stopping and the American put. *Mathematical Finance*, 1(2):1–14, 1991.

J. Jacod and V. Todorov. Testing for common arrivals of jumps for discretely observed multidimensional processes. *The Annals of Statistics*, 37(4):1792–1838, 2009.

P. Jalliet, D. Lamberton, and B. Lapeyre. Variational inequalities and the pricing of American options. *Acta Applicandae Mathematica*, 21:263–289, 1990.

F. Jamshidian. An analysis of American options. *Review of Futures Markets*, 11(1):72–80, 1992.

F. Jamshidian. Exchange options. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Inc., 2010.

R. A. Jarrow. Option pricing theory: Historical perspectives. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Inc., 2010.

M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets*. Springer-Verlag London Ltd., 2009.

G. J. Jiang. Jump-diffusion model of exchange rate dynamics–estimation via indirect inference. *Research Report 98A40, University of Groningen, Research Institute SOM (Systems, Organisations and Management)*, 1998.

G. J. Jiang and J. L. Knight. Estimation of continuous-time processes via the empirical characteristic function. *Journal of Business & Economic Statistics*, 20(2):198–212, 2002.

M. Johannes and N. Polson. Markov chain monte carlo. In T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch, editors, *Handbook of Financial Time Series*, pages 1001–1014. Springer-Verlag, 2009.

M. Johannes and N. Polson. Mcmc methods for continuous-time financial econometrics. In Y. Aït-Sahalia and L. P. Hansen, editors, *Handbook of Financial Econometrics*, chapter 13. North Holland, 2010.

B. Kang and G. H. Meyer. Pricing an American call under stochastic volatility and interest rates. In R. Dieci, X.-Z. He, and C. Hommes, editors, *Nonlinear Economic Dynamics and Financial Modelling: Essays in Honour of Carl Chiarella*, pages 291–314. Springer International Publishing, Switzerland, 2014.

I. Karatzas. On the pricing of American options. *Applied Mathematics and Optimization*, 17:37–60, 1988.

I. Karatzas and J. Ruf. Distribution of the time to explosion for one-dimensional diffusions. *Probability Theory and Related Fields*, 164:1027–1069, 2015.

I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag New York, Inc., New York, USA, 1988.

J. W. Kensinger. The capital investment project as a set of exchange options. *Managerial Finance*, 14(2/3):16–27, 1988.

I. J. Kim. The analytic valuation of American options. *Review of Financial Studies*, 3:547–572, 1990.

J.-H. Kim and C.-R. Park. A multiscale extension of the Margrabe formula under stochastic volatility. *Chaos, Solitons and Fractals: Nonlinear Science, and Nonequilibrium and Complex Phenomena*, 97:59–65, 2017.

E. Kirk. Correlation in energy markets. In V. Kaminski, editor, *Managing Energy Price Risk*. Risk Publications and Enron, 1995.

P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Berlin, 1992.

S. Klöppel and M. Schweizer. Dynamic indifference valuation via convex risk measures. *Mathematical Finance*, 17(4):599–627, 2007.

M. J. Kochenderfer and T. A. Wheeler. *Algorithms for Optimization*. The MIT Press, Cambridge, MA, 2019.

S. G. Kou. A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101, 2002.

S. G. Kou. Jump-diffusion models for asset pricing in financial engineering. In J. R. Birge and V. Linetski, editors, *Handbooks in Operations Research and Management Science*, chapter 2, pages 73–116. Elsevier, Amsterdam, The Netherlands, 2008.

P. Kovalov, V. Linetsky, and M. Marcozzi. Pricing multi-asset American options: A finite element method-of-lines with smooth penalty. *Journal of Scientific Computing*, 33:209–237, 2007.

D. O. Kramkov. Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. *Probability Theory and Related Fields*, 105:459–479, 1996.

H.-H. Kuo. *Introduction to Stochastic Integration*. Springer Science+Business Media, Inc., USA, 2006.

Y. K. Kwok, K. S. Leung, and H. Y. Wong. Efficient options pricing using the fast Fourier transform. In J.-C. Duan, W. K. Hardle, and J. E. Gentle, editors, *Handbook of Computational Finance*, chapter 21, pages 579–604. Springer-Verlag, 2012.

I. E. Lagaris, A. Likas, and D. I. Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks*, 9(8):987–1000, 1998.

R. W. Lee. Option pricing by transform methods: Extensions, unification, and error control. *Journal of Computational Finance*, 7(3):51–86, 2004.

P. Leoni and W. Schoutens. Multivariate smiling, 2007. URL https://perswww.kuleuven.be/~u0009713/smiling.pdf.

A. L. Lewis. *Option Valuation under Stochastic Volatility with Mathematica Code*. Finance Press, California, USA, 2000.

A. L. Lewis. A simple option formula for general jump-diffusion and other exponential Lévy processes. *Working Paper, Envision Financial Systems and OptionCity.net*, 2001.

M. Li, J. Zhou, and S.-J. Deng. Multi-asset spread option pricing and hedging. *Quantitative Finance*, 10(3):305–324, 2010.

W. Li, L. Liu, G. Lv, and C. Li. Exchange option pricing in jump-diffusion models based on Esscher transform. *Communications in Statistics - Theory and Models*, 47(19):4661–4672, 2018.

Z. Li. Equity option pricing with systemtatic and idiosyncratic volatility and jump risks. *Journal of Risk and Financial Management*, 13(1):1–18, 2020.

Y.-N. Lin. Pricing VIX futures: Evidence from integrated physical and risk-neutral probability measures. *The Journal of Futures Markets*, 27(12):1175–1217, 2007.

S. Lindset and S.-A. Persson. A note on a barrier exchange option: The world's simplest option formula? *Finance Research Letters*, 3:207–211, 2006.

S. Liu, A. Borovykh, L. A. Grzelak, and C. W. Oosterlee. A neural network-based framework for financial model calibration. *Journal of Mathematics in Industry*, 9(9):1–28, 2019a.

S. Liu, C. W. Oosterlee, and S. M. Bohte. Pricing options and computing implied volatilites using neural networks. *Risks*, 7(16), 2019b.

A. W. Lo. Maximum likeligood estimation of generalized Itô processes with discretely sampled data. *Econometric Theory*, 4(2):231–247, 1988.

J. D. Logan. *Applied Partial Differential Equations*. Springer International Publishing Switzerland, 3rd edition, 2015.

A. Løkka. Martingale representation of functionals of Lévy processes. *Stochastic Analysis and Applications*, 22(4):867–892, 2004.

F. A. Longstaff and E. S. Schwartz. Valuing American options by simulation: A simple least-squares approach. *Review of Financial Studies*, 14:113–147, 2001.

E. Luciano and W. Schoutens. A multivariate jump-driven financial asset model. *Quantitative Finance*, 6(5):385–402, 2006.

Y. Ma, D. Pan, and T. Wang. Exchange options under clustered jump dynamics. *Quantitative Finance*, 2020. URL https://doi.org/10.1080/14697688.2019.1704045.

D. B. Madan, P. R. Carr, and E. C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

C. Mancini. Estimation of the characteristics of the jumps of a general Poisson-diffusion model. *Scandinavian Actuarial Journal*, 1:42–52, 2004.

W. Margrabe. The value of an option to exchange one asset for another. *The Journal of Finance*, 33(1):177–186, 1978.

A.-M. Matache, C. Schwab, and T. P. Wihler. Fast numerical solution of parabolic integrodifferential equations with applications in finance. *SIAM Journal of Scientific Computing*, 27(2): 369–393, 2005.

R. L. McDonald and D. R. Siegel. Investment and the valuation of firms when there is an option to shut down. *International Economic Review*, 26(2):331–349, 1985.

H. P. McKean. Appendix: A free boundary value problem for the heat equation arising from a problem in mathematical economics. *Industrial Management Review*, 6(2):32–39, 1965.

R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):229–271, 1973.

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

G. H. Meyer. *Initial Value Methods for Boundary-Value Problems*. Academic Press, 1973.

G. H. Meyer. The numerical valuation of options with underlying jumps. *Acta Mathematica Universitatis Comenianae*, 67(1):69–82, 1998.

G. H. Meyer. *The Time-Discrete Method of Lines for Options and Bonds: A PDE Approach*. World Scientific, 2015.

G. H. Meyer and J. van der Hoek. The valuation of American options with the method of lines. In P. Boyle, G. Pennacchi, and P. Ritchken, editors, *Advances in Futures and Options Research*, pages 265–286. JAI Press, Inc., Greenwich, CT, 1997.

S. M. Miller. Booms and busts as exchange options. *Multinational Finance Journal*, 16(3/4): 189–223, 2012.

K. F. Mina, G. H. L. Cheang, and C. Chiarella. Approximate hedging of options under jump-diffusion processes. *International Journal of Theoretical and Applied Finance*, 18(4), 2015.

Y. Mishura and G. Shevchenko. The optimal time to exchange one asset for another on finite interval. In F. Delbaen, M. Rásonyi, and C. Stricker, editors, *Optimality and Risk–Modern Trends in Mathematical Finance: The Kabanov Festschrift*, pages 197–210. Springer-Verlag Berlin Heidelberg, Heidelberg, Germany, 2009.

Y. Miyahara. Canonical martingale measures of incomplete asset markets. In S. Watanabe, M. Fukushima, Y. V. Prohorov, and A. N. Shiryaev, editors, *Probability Theory and Mathematical Statistics: Proceedings of the Seventh Japan-Russia Symposium*, pages 343–352. World Scientific, 1996.

Y. Miyahara. *Option Pricing in Incomplete Markets: Modeling Based on Geometric Lévy Processes and Minimal Entropy Martingale Measures*. World Scientific, Singapore, 2012.

V. V. Morozov and K. V. Khizhnyak. An upper bound on the value of an infinite American call option on difference and sum of two assets. *Computational Mathematics and Modeling*, 24(1), 2013.

V. V. Morozov and K. V. Khizhnyak. A bound on the value of a two-sided Margrabe American option with finite expiration. *Computational Mathematics and Modeling*, 27(4):460–471, 2016.

M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling*. Springer-Verlag, Germany, 2nd edition, 2005.

M. Musiela and T. Zariphopoulou. An example of indifference prices under exponential preferences. *Finance and Stochastics*, 8:229–239, 2004.

R. Myneni. The pricing of the American option. *The Annals of Applied Probability*, 2(1):1–23, 1992.

V. Naik. Option valuation and hedging strategies with jumps in the volatility of asset returns. *The Journal of Finance*, 48(5):1969–1984, 1993.

V. Naik and M. Lee. General equilibrium pricing of options on the market portfolio with discontinuous returns. *The Review of Financial Studies*, 5(4):493–521, 1990.

J. N. Nielsen, H. Madsen, and P. C. Young. Parameter estimation in stochastic differential equations: An overview. *Annual Reviews in Control*, 24:83–94, 2000.

K. Nyström. On the behavior near expiry for multi-dimensional American options. *Journal of Mathematical Analysis and Applications*, 339:644–654, 2008.

T. Ogihara and N. Yoshida. Quasi-likelihood analysis for the stochastic differential equation with jumps. *Statistical Inference for Stochastic Processes*, 14:189–229, 2011.

B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, Heidelberg, (6th corrected printing) 6th edition, 2013.

B. Øksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions*. Springer, Switzerland, 3rd edition, 2019.

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

P. Parischa and A. Goel. Pricing power exchange options with Hawkes jump diffusion processes. *Journal of Industrial and Management Optimization*, 17(1):133–149, 2019.

P. Parischa and A. Goel. A closed-form pricing formula for European exchange options with stochastic volatility. *Probability in the Engineering and Informational Sciences*, 2021. URL https://doi.org/10.1017/S0269964820000698.

A. R. Pedersen. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics*, 22(1):55–71, 1995.

B. Peng and F. Peng. Pricing Bermudan exchange option under jump-diffusion process. *Mathematics in Practice and Theory*, 46(8):35–42, 2016.

H. Pham. Optimal stopping, free boundary, and American option in a jump-diffusion model. *Applied Mathematics and Optimization*, 35(2):145–164, 1997.

H. Pham. Optimal stopping of controlled jump diffusion processes: a viscosity solution approach. *Journal of Mathematical Systems, Estimation, and Control*, 8(1):1–27, 1998.

P. C. B. Phillips and J. Yu. Maximum likelihood and Gaussian estimation of continuous time models in finance. In T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch, editors, *Handbook of Financial Time Series*, pages 497–530. Springer-Verlag, 2009.

E. A. D. Pienaar and M. M. Varughese. *DiffusionRjgqd: An R Package for Performing Inference and Analysis on Time-Inhomogeneous Quadratic Jump Diffusions with State-Dependent Intensity*, 2015. URL https://CRAN.R-project.org/package=DiffusionRjgqd. R package version 0.1.1.

E. A. D. Pienaar and M. M. Varughese. *Likelihood inference for non-linear, multivariate jump diffusions with state-dependent intensity*, 2016. URL https://CRAN.R-project.org/package=DiffusionRjgqd. R Vignette.

E. Platen. A non-linear stochastic volatility model. *Financial Mathematics Research Report No. FMRR 005-97, Center for Financial Mathematics, Australian National University*, 1997.

E. Platen and N. Bruti-Liberati. *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*. Springer-Verlag, Berlin, 2010.

R. Poulsen. Margrabe formula. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Inc., 2010.

B. L. S. Prakasa Rao. *Statistical Inference for Diffusion Type Processes*. Wiley, 1999.

S. P. Preston and A. T. A. Wood. Approximation of transition densities of stochastic differential equations by saddlepoint methods applied to small-time Itô-Taylor sample-path expansions. *Statistics and Computing*, 22:205–217, 2012.

P. E. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin, 2nd edition, 2004.

F. Quittard-Pinon and R. Randrianarivony. Exchange options when one underlying price can jump. *Finance*, 31(1):33–53, 2010.

B. Radi and A. El Hami. *Advanced Numerical Methods with MATLAB*. John Wiley & Sons, Inc., 2nd edition, 2018.

R. Rannacher. Finite element solution of diffusion problems with irregular data. *Numerische Mathematik*, 43:309–327, 1984.

G. Rapuch. American options and the free boundary exercise region: a PDE approach. *Interfaces and Free Boundaries*, 7:79–98, 2005.

I. Richards and H. Youn. *Theory of Distributions: A Non-Technical Introduction*. Cambridge University Press, Cambridge, UK, 1990.

C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Science + Business Media, Inc., New York, USA, 2nd edition, 2004.

W. Rudin. *Functional Analysis*. McGraw-Hill, Inc., New York, 2nd edition, 1991.

W. J. Runggaldier. Jump-diffusion models. In S. T. Rachev, editor, *Handbook of Heavy Tailed Distributions in Finance*, volume 1, chapter 5, pages 169 – 209. North-Holland, Amsterdam, The Netherlands, 2003.

W. J. Runggaldier. Estimation via stochastic filtering in financial market models. In G. Yin and Q. Zhang, editors, *Mathematics of Finance: Proceedings of an AMS-IMS-SIAM Joint Summer Research Conference on Mathematics of Finance*, pages 309–318. American Mathematical Society, Providence, RI, 2004.

S. Salmi, J. Toivanen, and L. von Sydow. Iterative methods for pricing American options under the bates model. *Procedia Computer Science*, 18:1136–1144, 2013.

D. M. Salopek. *American Put Options*. Addison Welsey Longman Limited, England, 1997.

Y. F. Saporito, X. Yang, and J. P. Zubelli. The calibration of stochastic-local volatility models: an inverse problem perspective. *Computers & Mathematics with Applications*, 77(12):3054–3067, 2019.

W. E. Schiesser and G. W. Griffiths. *A Compendium of Partial Differential Equation Models: Method of Lines Analysis with MATLAB*. Cambridge University Press, 2009.

W. Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives*. John Wiley & Sons, Ltd., England, 2003.

M. Schweizer. Approximation pricing and the variance-optimal martingale measure. *The Annals of Probability*, 24(1):206–236, 1996.

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

L. O. Scott. Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Applications of Fourier inversion methods. *Mathematical Finance*, 7(4):413–424, 1997.

R. U. Seydel. *Tools for Computational Finance*. Springer-Verlag, London, 2017.

Y. Shen, M. Sherris, and J. Ziveyi. Valuation of guaranteed minimum maturity benefits in variable annuities with surrender options. *Insurance: Mathematics and Economics*, 69:127–137, 2016.

N. Shephard. General introduction. In N. Shephard, editor, *Stochastic Volatility: Selected Readings*, chapter 1, pages 1–33. Oxford University Press, 2005a.

N. Shephard, editor. *Stochastic Volatility: Selected Readings*. Oxford University Press, 2005b.

N. G. Shephard. Numerical integration rules for multivariate inversions. *Journal of Statistical Computation and Simulation*, 39:37–46, 1991a.

N. G. Shephard. From characterstic function to distribution function: A simple framework for the theory. *Econometric Theory*, 7:519–528, 1991b.

Y. Shimizu and N. Yoshida. Estimation of parameters for diffusion processes with jumps from discrete observations. *Statistical Inference for Stochastic Processes*, 9:227–277, 2006.

S. E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer-Verlag, New York, 2004.

A. F. Siegel. Measuring systematic risk using implicit beta. *Management Science*, 41(1):124–128, 1995.

C. A. Sin. Complications with stochastic volatility models. *Advances in Applied Probability*, 30: 256–268, 1998.

K. J. Singleton. Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102:111–141, 2001.

J. Sirignano and K. Spiliopoulos. DGM: a deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, 375:1339–1364, 2018.

J. L. Solé, F. Utzet, and J. Vives. Canonical Lévy process and Malliavin calculus. *Stochastic Processes and their Applications*, 117:165–187, 2007.

H. Sørensen. Parametric inference for diffusion processes observed at discrete points in time: a survey. *International Statistical Review*, 72(3):337–354, 2004.

E. M. Stein and J. C. Stein. Stock price distributions with stochastic volatility: An analytic approach. *The Review of Financial Studies*, 4(4):727–752, 1991.

R. S. Strichartz. *A Guide to Distribution Theory and Fourier Transforms*. World Scientific Publishing Co. Pte. Ltd., Singapore, 2003.

R. Suzuki. A Clark-Ocone type formula under change of measure for Lévy processes with L^2-Lévy measure. *Communications on Stochastic Analysis*, 7(3):383–407, 2013.

R. Suzuki. A Clark-Ocone type formula under change of measure for multidimensional Lévy processes. *Communications on Stochastic Analysis*, 11(1):21–42, 2017.

B. Taruvinga. *Solving Selected Problems on American Option Pricing with the Method of Lines*. PhD thesis, University of Technology Sydney, 2019.

Y. Tian, Z. Zhu, G. Lee, F. Klebaner, and K. Hamza. Calibrating and pricing with a stochastic-local volatility model. *Journal of Derivatives*, 22(3):21–39, 2015.

J. Toivanen. A componentwise splitting method for pricing American options under the Bates model. In W. Fitzgibbon, Y. Kuznetsov, P. Neittaanmaki, J. Periaux, and O. Pironneau, editors, *Applied and Numerical Partial Differential Equations: Scientific Computing in Simulation, Optimization and Control in a Multidisciplinary Context*, pages 213–227. Springer, Heidelberg, 2010a.

J. Toivanen. Finite difference methods for early exercise options. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Inc., 2010b.

A. Torchinsky. The Fourier transform and the wave equation. *The American Mathemaical Monthly*, 118(7):599–609, 2011.

N. Touzi. American options exercise boundary when the volatility changes randomly. *Applied Mathematics and Optimization*, 39:411–422, 1999.

G. Van Damme. Legendre, Laguerre, and Hermite-Gauss quadrature, 2010. https://au.mathworks.com/matlabcentral/fileexchange/26737-legendre-laguerre-and-hermite-gauss-quadrature, MATLAB Central File Exchange, Retrieved: 7 May 2020.

J. van der Hoek and M. W. Korolkiewicz. New analytic approximations for pricing spread options. In S. N. Cohen, D. Madan, T. K. Siu, and H. Yang, editors, *Stochastic Processes, Finance and Control: A Festschrift in Honor of Robert J. Elliot*, chapter 11, pages 259–284. World Scientific, 2012.

G. van Dijk. *Distribution Theory: Convolution, Fourier Transform, and Laplace Transform*. Walter de Gruyter GmbH, Berlin, 2013.

P. Van Moerbeke. On optimal stopping and free boundary problems. *Archive for Rational Mechanics and Analysis*, 60(2):101–148, 1976.

A. Venkatramanan and C. Alexander. Closed form approximations for spread options. *Applied Mathematical Finance*, 18(5):447–472, 2011.

S. Villeneuve. Exercise regions of American options on several assets. *Finance and Stochastics*, 3: 295–322, 1999.

L. Von Sydow, L. J. Höök, E. Larsson, E. Linström, S. Milovanović, J. Persson, V. Shcherbakov, Y. Shpolyanskiy, S. Sirén, J. Toivanen, J. Waldén, M. Wiktorsson, J. Levesley, J. Li, C. W. Oosterlee, M. J. Ruijter, A. Toropov, and Y. Zhao. BENCHOP - The BENCHmarking project in option pricing. *International Journal of Computer Mathematics*, 92(12):2361–2379, 2015.

C. Wang and R. Chen. Option pricing for Merton jump diffusion model with regime-switching using FFT. In *2009 International Conference on Management and Service Science*, pages 1–4, 2009. URL https://doi.org/10.1109/ICMSS.2009.5303110.

W. Wang and Y. Chen. Fast numerical valuation of options with jump under merton's model. *Journal of Computational and Applied Mathematics*, 318:79–92, 1993.

X. Wang. Pricing power exchange options with correlated jump risk. *Finance Research Letters*, 19: 90–97, 2016.

X. Wang, S. Song, and Y. Wang. The valuation of power exchange options with counterparty risk and jump risk. *The Journal of Futures Markets*, 37(5):499–521, 2017.

D. Williams. *Probability with Martingales*. Cambridge University Press, Great Britain, 1991.

P. Wilmott, J. Dewynne, and S. Howison. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, London, 1993.

B. Wong and C. C. Heyde. On the martingale property of stochastic exponentials. *Journal of Applied Probability*, 41:654–664, 2004.

M. Xu. Risk measure pricing and hedging in incomplete markets. *Annals of Finance*, pages 51–71, 2006.

Z. Xu and X. Jia. The calibration of volatility for option pricing models with jump diffusion processes. *Applicable Analysis*, pages 1–18, 2017.

J.-A. Yan. *Introduction to Stochastic Finance*. Springer Nature Singapore Ptd Ltd., Singapore, 2018.

J. Yu. Closed-form likelihood approximation and estimation of jump-diffusions with an application to the realignment risk of the Chinese Yuan. *Journal of Econometrics*, 141:1245–1280, 2007.

E. Zauderer. *Partial Differential Equations of Applied Mathematics*. John Wiley & Sons, Inc., New Jersey, 3rd edition, 2006.

S.-P. Zhu and G.-H. Lian. An analytical formula for VIX futures and its applications. *The Journal of Futures Markets*, 32(2):166–190, 2011.

O. Zhylyevskyy. A fast Fourier transform technique for pricing American options under stochastic volatility. *Review of Derivatives Research*, 13:1–24, 2010.

J. Ziveyi. *The Evaluation of Early Exercise Exotic Options*. PhD thesis, University of Technology Sydney, 2011.

J. Ziveyi, C. Blackburn, and M. Sherris. Pricing European options on deferred annuities. *Insurance: Mathematics and Economics*, 52:300–311, 2013.

D. Zwillinger. *CRC Standard Mathematical Tables and Formulas*. CRC Press, Florida, 33rd edition, 2018.