ABSTRACT

Title of dissertation: Optimal Variance Swaps Portfolios and Estimating Greeks for Variance-Gamma

Lingyan Cao, Doctor of Philosophy, 2011

Dissertation directed by: Dr. Michael C. Fu
Department of Decision, Operations, and Information Technologies
R. H. Smith School of Business

Dr. Dilip B. Madan
Department of Finance
R. H. Smith School of Business

In this dissertation, we investigate two problems: constructing optimal variance swaps portfolios and estimating Greeks for options with underlying assets following a Variance Gamma process. By modeling the dependent non-Gaussian residual in a linear regression model through a Lévy Mixture (LM) model and a Variance Gamma Correlated (VGC) model, and running some optimizations, we construct an optimal variance swap portfolio. By implementing gradient estimation techniques, we estimate the Greeks for a series of basket options called Mountain Range options.

Constructing an optimal variance swap portfolio consists of two steps: evaluations and optimization. Each variance swap has two legs: a fixed leg (also called the variance strike) and a floating leg (also called the realized variance). The value of a variance swap is the discounted difference between the realized variance and

the variance strike. For the latter, one can use an option surface calibration to evaluate. For the former, the procedure is complicated due to the non-negligible residuals from a linear regression model. Through LM and VGC, we can estimate the realized variance on different sample paths and obtain the payoff of a variance swap numerically. Based on these numerical results, we can apply the optimization method to construct an optimal portfolio.

In the second part of this dissertation, we consider gradient estimation for Mountain Range options including Everest options, Atlas options, Altiplano/Annapurna options and Himalayan options. Assuming the underlying assets follow a Variance-Gamma (VG) process, we derive estimators for sensitivities such as Greeks through Monte Carlo simulation. We implement and compare using numerical experiments several gradient estimation approaches: finite difference methods (forward difference), infinitesimal perturbation analysis (IPA), and likelihood ratio (LR) method using either the density function or the characteristic function.

Optimal Variance Swaps Portfolios and Estimating Greeks for Variance-Gamma

by

Lingyan Cao

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor Michael C. Fu, Chair/Advisor
Professor Dilip B. Madan, Co-Advisor
Professor Mark Loewenstein
Professor Paul J. Smith
Professor Victor M. Yakovenko

Dedication

To My Parents and Yong

Acknowledgments

I want to express my gratitude to all the people who have made this thesis successful.

First and foremost, I would like to thank my dissertation advisors, Professor Michael C. Fu and Professor Dilip Madan for their guidance, encouragement, patience and support on my research. I appreciate for their helping me enter the world of mathematical finance and Monte Carlo simulation in financial engineering.

I would also like to thank Professor Paul Smith, Professor Mark Loewenstein, and Professor Victor M. Yakovenko for agreeing to serve on my defense committee and spending their invaluable time reviewing the manuscript.

I would also like to acknowledge all the help and support I have received from professors in the mathematics department. I want to thank Professor von Petersdorff for the one year seminar course sequence on mathematical finance and many valuable discussions during the preparation of the thesis. I also thank Professor Konstanitina Trivisa for her support and encouragement during my graduate life.

In addition, I owe gratitude to all the participants of the mathematical finance research interaction team (RIT). I have gained a lot of knowledge from weekly discussions in the RIT.

I would also like to express my appreciation to my friends and colleagues. I would like to thank Andrew, Bo, Charles, Hui, Huitian, Hui, Kuan, Li, Linbao, Min, Qiyi, Rongrong, Shu, Will, Xinxian, Ying, Yunhe, and Zhengfeng for their encouragement. I also thank my fellow students Geping, Guoyuan, Huaqiang, Jun,

Tong, and Yun for many helpful discussions.

I thank the staff members Alverda McCoy, Sharon Welton, Fletcher Kinne, Linette Berry, Haydee Hidalgo and Celeste Regalado in the mathematics department. They have offered much support and help.

Moreover, I am also grateful for the financial support from the math department.

Last but not the least, I would express my special appreciation to my parents for their support and understanding. Special thanks to my husband Yong. His everlasting love, persistent support and constant encouragement are what I rely on whenever I encounter difficulties.

Table of Contents

List of Tables viii

List of Figures ix

List of Abbreviations x

1 Introduction 1
- 1.1 Financial Models . . . 1
- 1.2 Simulation and Financial Engineering . . . 3
  - 1.2.1 Gradient Estimation . . . 3
- 1.3 Variance Swaps . . . 5
  - 1.3.1 Multidimensional Non-Gaussian Models . . . 7
- 1.4 Two Essays . . . 9
- 1.5 Outline of the Dissertation . . . 10

2 Preliminaries 12
- 2.1 Introduction to the Lévy Process . . . 12
  - 2.1.1 Lévy Process and Infinitely Divisible Distribution . . . 12
  - 2.1.2 Lévy-Khintchine Representation . . . 15
  - 2.1.3 Lévy-Itô decomposition . . . 16
- 2.2 Variance Gamma Process . . . 17
  - 2.2.1 Definition of VG Process . . . 18
  - 2.2.2 Properties of VG Process . . . 19
  - 2.2.3 Simulation of the VG Process . . . 21
- 2.3 The Fast Fourier Transform Method and Option Pricing . . . 23
  - 2.3.1 The Carr-Madan FFT Method . . . 25

3 Multidimensional Dependent Non-Gaussian Models 28
- 3.1 Overview . . . 28
- 3.2 Full Rank Gaussian Copula (FGC) . . . 29
  - 3.2.1 Introduction to the FGC model . . . 29
  - 3.2.2 Estimation Procedures for FGC . . . 31
- 3.3 Lévy Mixture (LM) . . . 32
  - 3.3.1 Introduction to the LM model . . . 32
  - 3.3.2 Estimation Procedure for LM . . . 34
- 3.4 Variance Gamma Correlated (VGC) . . . 34
  - 3.4.1 Introduction to the VGC model . . . 35
  - 3.4.2 Estimation Procedures for VGC . . . 37

4 Monte Carlo Simulation and Gradient Estimation 38
- 4.1 Overview . . . 38
- 4.2 Gradient Estimation Techniques . . . 38
  - 4.2.1 Indirect Methods . . . 41

4.2.2 Infinitesimal Perturbation Analysis . . . 43
4.2.3 Likelihood Ratio Method . . . 46
4.2.4 GL method . . . 48

5 Constructing Optimal Portfolios of Variance Swaps by Lévy Mixture and Variance Gamma Correlated 52
5.1 Overview . . . 52
5.2 Definition of Variance Swaps . . . 54
5.3 Fixed Leg (Variance Strike) . . . 56
5.3.1 Calculation of the Variance Strike . . . 56
5.4 Floating Leg (Realized Variance) . . . 66
5.4.1 Hardy-Littlewood Transform and Linear Regression Model . . . 67
5.4.2 Lévy Mixture . . . 70
5.4.3 Variance Gamma Correlated Model . . . 71
5.4.4 Simulations . . . 73
5.5 Optimization . . . 74
5.5.1 A New Performance Measure . . . 75
5.6 Numerical Experiments and Conclusions . . . 80

6 Estimating Greeks for Variance-Gamma Processes 84
6.1 Overview . . . 84
6.2 Background . . . 85
6.2.1 Greeks . . . 85
6.2.2 Mountain Range Options . . . 86
6.3 Gradient Estimation . . . 87
6.3.1 Indirect Methods . . . 88
6.3.2 Direct Methods . . . 89
6.3.3 GL Method . . . 91
6.4 Problem Setting . . . 93
6.5 A European Call Example . . . 94
6.5.1 IPA for European Call Option . . . 95
6.5.2 LR for European Call Option . . . 99
6.5.3 GL for European Call Option . . . 101
6.5.4 Numerical Experiments . . . 101
6.6 Independent Mountain Range Options . . . 102
6.6.1 Everest Option . . . 104
6.6.2 Atlas Option . . . 107
6.6.3 Altiplano/Annapurna Option . . . 110
6.6.4 Himalayan Option . . . 112
6.7 Dependent Mountain Range Options . . . 115
6.7.1 Everest Option . . . 116
6.7.2 Atlas Option . . . 120
6.8 Conclusions and Future Work . . . 121

A Derivation of the Variance Strike 130
A.1 Spot Value of M_1 . . . 132
A.2 Spot Value of M_2 . . . 133
A.3 Spot Value of M_3 . . . 133

B Some Calculations 135
B.1 Calculation of $\frac{d\gamma_t^{(\nu)}}{dt}$ and $\frac{dW_{\gamma_t^{(\nu)}}}{dt}$ . . . 135
B.2 Calculation of $\frac{d\gamma_t^{(\nu)}}{d\nu}$ . . . 136
B.3 Calculation of $\frac{d\ln h(z)}{d\xi}$ . . . 138
B.4 Calculation of of $\frac{d\ln f(X_1^t, X_2^t;\xi)}{d\xi}$ . . . 141
B.5 Calculation of Derivatives of $\frac{dX_i}{d\xi}$ . . . 145
B.6 Calculation of Derivatives of $\phi_{X(t)}(u)$ . . . 147

C C Code 149
C.1 Implementation of Variance-Gamma process . . . 149
C.2 Implementation of Stock Price under VG . . . 150
C.3 Implementation of A European Call Option Price . . . 150
C.4 Implementation of Uniform Distributed Random Variables . . . 151
C.5 Implementation of Normal Distributed Random Variables . . . 151
C.6 Implementation of IPA Estimates . . . 152
C.7 Implementation of LR Estimates . . . 153

Bibliography 155

List of Tables

5.1 Tickers and Companies . . . 57
5.2 VGSSD Parameters on 20071019 . . . 60
5.3 One Month Variance Strike on 20071019 . . . 60
5.4 Linear Regression Results . . . 69
5.5 VG estimates for independent variates Y for LM . . . 72
5.6 The Mixing Matrix A for LM . . . 73
5.7 VG estimates for variates Y for VGC . . . 74
5.8 The Covariance Matrix for Standard Normal Variates Z . . . 75
5.9 Portfolio by Maximizing the MinMaxVar Expected Distortion by LM 82
5.10 Portfolio by Maximizing the MaxMinVar Expected Distortion by LM 82
5.11 Portfolio by Maximizing the MinMaxVar Expected Distortion by VGC 82
5.12 Portfolio by Maximizing the MaxMinVar Expected Distortion by VGC 82
5.13 Realized Results of One-Month Variance Swaps by MinMaxVar . . . 82
5.14 Realized Results of One-Month Variance Swaps by MaxMinVar . . . 83
5.15 Realized Results of One-Month Variance Swaps by MinMaxVar in 2007 83

6.1 Simulated European Option Price . . . 101
6.2 European option simulation results . . . 123
6.3 Simulated Everest Option Price . . . 123
6.4 Independent Everest option simulation results . . . 124
6.5 Simulated Atlas Option Price . . . 124
6.6 Independent Atlas option simulation results . . . 125
6.7 Simulated Independent Altiplano Option Price . . . 125
6.8 Independent Altiplano options simulation results . . . 126
6.9 Simulated Himalayan Option Price . . . 126
6.10 Independent Himalayan option simulation results . . . 127
6.11 Independent Everest option simulation results for $t > \nu_i$ . . . 127
6.12 Independent Atlas option simulation results for $t > \nu_i$ . . . 128
6.13 Independent Altiplano options simulation results for $t > \nu_i$ . . . 128
6.14 Independent Himalayan option simulation results for $t > \nu_i$ . . . 128
6.15 Dependent Everest option simulation results . . . 129
6.16 Dependent Atlas option simulation results . . . 129

List of Figures

5.1 Graph of Fitted Option Surface of XOM on 20071019 61
5.2 Graph of Fitted Option Surface of AAPL on 20071019 61
5.3 Graph of Fitted Option Surface of MMM on 20071019 62
5.4 Graph of Fitted Option Surface of C on 20071019 62
5.5 Graph of Fitted Option Surface of ADBE on 20071019 63
5.6 Graph of Fitted Option Surface of AMZN on 20071019 63
5.7 Graph of Fitted Option Surface of GS on 20071019 64
5.8 Graph of Fitted Option Surface of COH on 20071019 64
5.9 Graph of Fitted Option Surface of GOOG on 20071019 65
5.10 Graph of Fitted Option Surface of BAC on 20071019 65

List of Abbreviations

σ	sigma
ν	nu
θ	theta
Δ	delta
vega	vega
rho	rho
Θ	Theta

a.s.	almost surely
i.i.d.	independent identically distributed
c.d.f.	cumulative distribution function
FD	forward difference
FGC	full rank Gaussian copula
LM	Lévy Mixture
VGC	Variance Gamma Correlated
MVO	market valus of options bought
MLE	maximum likelihood estimation
VG	Variance Gamma
IPA	infinitesimal perturbation analysis method
LR	likelihood ratio method
GL	method of Glasserman and Liu
IC	independent component
ICA	independent component analysis
NIG	normal inverse Gaussian
CGMY	Carr, Geman, Madan and Yor
SP	simultaneous perturbation
MVO	market value of options bought
BSM	Blach-Scholes-Merton model

Chapter 1

Introduction

1.1 Financial Models

The most well-known continuous stochastic model for option pricing is the classical Black-Scholes-Merton (BSM) model (Merton [69], Black and Scholes [10]), in which the underlying stock price is assumed to follow a geometric Brownian motion process. This widely applied model has some drawbacks. Firstly, the BSM model is based on some crucial assumptions, for example, no taxes or transaction costs, etc. Secondly, empirical evidence suggests that the classical Black-Scholes model does not describe the statistical properties of financial time series well. It is observed that log returns of market prices do not behave as a normal distribution. They are skewed and have excess kurtosis. Thirdly, volatilities or parameters of estimated uncertainty change stochastically over time and are clustered, but they are assumed to be some constant in the BSM model.

In order to define a stochastic process with independent and stationary increments, a good model with an infinitely divisible probability distribution is necessary to price and hedge derivative securities. One candidate with such desirable properties is a Lévy process (Bertoin [9], Sato [75], Applebaum [2]), named in honor of Paul Lévy the pioneer of the theory. A variety of Lévy models whose distributions are infinitely divisible and can represent skewness and excess kurtosis were proposed

and studied in a vast literature during the 1980s and 1990s, for example, the Variance Gamma (VG) distribution, the normal inverse Gaussian (NIG) distribution, the CGMY (named after Carr, Geman, Madan and Yor) (Carr, Geman, Madan and Yor [18]) distribution, and the generalized hyperbolic distribution. Corresponding to these distributions, a wide variety of univariate financial models to price financial data have been proposed and studied. The VG model was introduced to the finance community as a model for log-price returns and option pricing in Madan and Seneta [60], and was developed in Madan and Milne [59] and Madan, Carr, and Chang [57]; the normal inverse Gaussian (NIG) model was studied in Barndorff-Nielsen [7]; the CGMY model was proposed in Carr, Geman, Madan and Yor [18]; the hyperbolic model was introduced in Eberlein and Keller [28]. The successes with these univariate models indicate that they are capable of explaining the unconditional return densities on asset prices.

However, multivariate models are required in many financial applications, such as basket option pricing, portfolio optimization and simulation of risk scenarios for portfolios. Thus, multidimensional models with dependence between components are more applicable in the finance community, and there has been an increasing interest in the multivariate Lévy process modeling. Similar to univariate models, jumps in the market price must be taken into account. However, multidimensional models with jumps are much more difficult to construct. Many tools have been provided to build multivariate dependent models with jumps. The copula method is one of the most popular techniques for extending univariate Lévy processes to multivariate processes; many copulas including Gaussian, Student-t, Clayton copula have

been successfully introduced and applied in the mathematical finance world. For example, the Lévy copula model was introduced in Tankov [81], in which the joint law of the multivariate Lévy process is characterized by the copula method on the Lévy measure. The multivariate time-changed Brownian motion subordinated by a common subordinator is studied and tested in Cont and Tankov [24], Luciano and Schoutens [51]; for a similar model with multivariate subordinators, see Semeraro [78] and Luciano and Semeraro [52]. The copula methods have proved useful and applied widely, for example in pricing credit derivatives, and structural models (see Burtschell, Gregory and Laurent [15], Laurent and Gregory [50], Madan, Konikov and Marinescu [58], Berd, Engle and Voronov [8], etc).

1.2 Simulation and Financial Engineering

Monte Carlo simulation is a method broadly used in the financial community for derivatives pricing and hedging. Gradient estimation, which is required for hedging, is a technique to estimate gradients of financial derivatives based on Monte Carlo simulation. It has proved useful in sensitivity analysis, as input to optimization, and has been employed in financial engineering.

1.2.1 Gradient Estimation

Gradient estimates play an important role in measuring and managing risks. Sensitivities estimated from gradient estimation measure the effect of change of parameters on the price of derivatives, so that the investors or the holders of the

derivative can adjust their hedging strategy corresponding to market changes.

Gradient estimation techniques were first applied to option pricing focusing on infinitesimal perturbation analysis (IPA) estimation for European and American options by Fu and Hu [31]. Both IPA and the likelihood ratio (LR) method were applied to European and Asian options by Broadie and Glasserman [14]; see also Glasserman [38]. Fu [34] provided a general survey of gradient estimation in stochastic simulation. Fu [33] explained more details on Monte Carlo simulation for financial engineering and various methods for estimating the Greeks using simulation. Glasserman and Liu [39] provided a new estimation method similar to LR which relies only on the characteristic function and does not require the explicit probability density function of the transition. This method is especially relevant for the simulation of Lévy processes, where the characteristic function is readily available and the density function is complicated.

Fu [35] gave a general introduction to the VG process in the context of stochastic (Monte Carlo) simulation and showed how to simulate the stock price and price derivative securities. Hall [41] considered gradient estimation for a class of financial derivatives on a basket of stocks called Mountain Range options under an asset price model of geometric Brownian motion.

Most of the literature above assumes that the underlying asset price follows a geometric Brownian motion process, also called the BSM model. It turns out that the BSM model has some shortcomings in describing the statistical properties of empirical results of market stock prices, such as: (1) the stock price is continuous in the BSM model but discontinuous in the real market; (2) the volatility of stock

price is assumed to be a constant in the BSM model but stochastically changing in real market. If we use the market price to estimate the volatility, a volatility smile will appear, in which the volatility obtained from market is not a constant, contradicting a key BSM model assumption. Based on the imperfections of the BSM model, the Greeks estimated from Monte Carlo-based gradient estimation would lead to errors in estimating the sensitivities of the financial derivatives, possibly resulting in inefficient hedging. Empirical results indicate that log-returns of stock prices are not normally distributed and have fat tails; thus a model whose distribution is non-Gaussian and has fat tails seems more reasonable. Therefore, we estimate Greeks of mountain range options by assuming the stock prices follow a VG process in the first essay of this dissertation.

1.3 Variance Swaps

Variance swaps have been used to trade equity-index volatility and have demonstrated some advantages over other volatility-based assets. A variance swap is an instrument that allows investors to trade future realized (or historical) volatility against current implied volatility. Through a variance swap, investors can achieve long or short exposure to market volatility. For example, when a stock investor wants to speculate on the possible change direction of the stock market, or a bond investor thinks he can foresee the probable change direction of interest rates, he can buy or sell the stocks or bonds. Similarly, investors may also have some thoughts about the change direction of the volatility.

We briefly review some literature on variance swaps. Demeterfi, Derman, Kamal and Zou [26] showed how to replicate and value variance swaps from the perspective of the Black-Scholes-based account of the fundamental strategy, as well as how to dynamically hedge a variance swap and analyze the imperfections of replications. Bossu, Strasser, and Guichard [11] explained hedging strategies of variance swaps both from an intuitive view and from theoretical insights; see also Gairat [36]. Carr and Lee [19] showed how to develop trading strategies for volatility derivatives in a nonparametric setting. They found nonparametric formulas to price variance swaps and other volatility derivatives and claimed that their results were valid if the volatility satisfies an independent condition. Madan [54] provided an introduction to the variance swap contract and proved how to calculate the price of variance strike. Although variance swaps have been traded in the market for a period of time and studied a lot, there was no work in the literature showing how to construct an optimized portfolio of variance swaps until Madan [55] first proposed the variance swap portfolio theory, which shows how to construct optimal portfolios of variance swaps considering both autocorrelation and cross asset dependencies, for which a regression model and a full rank Gaussian copula model is employed. The portfolio is obtained by maximizing the index of acceptability given some distortion function.

1.3.1 Multidimensional Non-Gaussian Models

The description of the joint law of asset returns has drawn a lot of interest in the area of financial modeling. However, a vast literature (such as Boyarchenko and Levendorskii [12], Menn, Fabozzi and Rachev [68], McNeil, Frey and Embrechts [66], and Jondeau, Poon, and Rockinger [48]) provided some examples in which the marginal distribution of each asset return taken separately is not Gaussian. A number of applications call for the study of non-Gaussian multivariate return distributions, and a lot of studies have been done on this topic. These could be applied in many areas, such as pricing options on a basket of stocks where the return distribution is risk neutral, or designing optimal portfolios where the interest is in the physical multivariate return distribution. There are many kinds of multivariate elliptical distributions, such as the multivariate $t-$distribution (Kotz and Nadarajah [49]) or the multivariate VG (Madan and Seneta [60], Schoutens and Cariboni [77]), any of which can be applied to model multidimensional non-Gaussian distributed models. We introduce three multidimensional dependent non-Gaussian models: full rank Gaussian copula, a Lévy mixture and a correlated Variance Gamma. The particular advantage of these three models is that one can estimate high-dimensional models by reducing the problem to a suitable sequence of univariate estimation problems. Thus, modeling the high-dimensional non-Gaussian returns can be simplified to a univariate non-Gaussian modeling problem.

The full rank Gaussian copula (FGC) model was proposed and studied in Malevergne and Sornette [63], where they correlated the set of standard normal

variates generated from a linear transform of each asset return. They provided a simple representation of multivariate distributions of returns which can describe the non-Gaussian fat-tailed properties of the empirical distribution with a nonlinear dependence between each single asset. Moreover, they provided analytical results on the moments and cumulants of the distributions of returns for the portfolio based on a class of multivariate Weibull distributions to parameterize the non-Gaussian properties. Madan [55] used the FGC model to take cross asset dependencies into account. In this dissertation, we present the first essay "Constructing Portfolios of Variance Swaps" based on the other two multidimensional non-Gaussian models: Variance-Gamma Correlated (VGC) and Lévy Mixture (LM).

The VGC model was proposed in Eberlein and Madan [30], where the Lévy process is written as a time-changed Brownian motion and the correlation is put in the Brownian motions part. This is a relatively simple approach to correlate unit period returns resulting from a Lévy process. They first evaluated the model statistically at the level of explaining pairwise joint daily returns. They showed that when the marginal laws have been estimated by the univariate densities, the pairwise joint law only requires an estimation of the correlation between the two Brownian motions that were marginally subjected to a time change. They also showed that sample correlations understate the correlation between the Brownian motions; the correlation between the two Brownian motions is in absolute value always greater than the sample correlation between the asset returns. A chi-squared test of model performance in terms of p-values for the time-changed model was conducted to verify that the model makes a significant improvement in explaining the pairwise

joint structure of daily asset returns. Finally, a exact sample correlation is provided through laws of marginal time changes. The VGC model was employed in Madan [55] to study option pricing on a basket of stocks.

The Lévy Mixture (LM) model is a linear mixture of independent but non-Gaussian variates, which are infinitely divisible and associated with the unit time distribution of a Lévy process. This model was proposed in Madan and Yen [61] to design portfolios for asset allocation. They employed a signal processing technique called independent component analysis (ICA) (Hyvärinen [45], and Hyvärinen, Karhuen and Oja [46]) for multivariate financial time series. They first decomposed the observed time series into statistically independent components (ICs), and assume the ICs follow the VG process. A portfolio is then constructed through application of this model.

1.4 Two Essays

In this dissertation, we present two essays: "Optimal Variance Swaps Portfolios by Lévy mixture and Variance Gamma Correlated" and "Estimating Greeks for Variance-Gamma".

In the first essay, we focus on introducing the variance swaps and constructing optimal portfolios following the theory proposed in Madan [56]. The variance swap contract contains two legs; one is the fixed leg, also called the variance strike; the other is the floating leg, also called the realized variance. The variance strike is calculated from the option surface calibration. The realized variance is calculated

through the Hardy-Littlewood-Gauss transform, and a linear regression model considering the highly correlated autocorrelation and dependencies of cross assets. Two non-Gaussian models: Lévy Mixture and Varanance Gamma Correlated (VGC) are employed to describe the residuals in the regression model. Optimal portfolios of variance swaps are constructed in a new performance measure called acceptability indices.

In the second essay, assuming the underlying assets follow a VG process, we consider the problem of estimating sensitivities, also called the Greeks, of options on a basket of assets by Monte Carlo simulation. We focus on a class of derivatives called Mountain Range options, including the Atlas option, the Everest option, the Altiplano option, the Annapurna option and the Himalayan option. Estimators of gradients are calculated through indirect methods (finite difference techniques such as forward differences) and two direct methods: infinitesimal perturbation analysis (IPA) and the likelihood ratio (LR) method, where the LR is also implemented via a recently proposed numerical technique developed in Glasserman and Liu [39] (the GL method). We carry out numerical simulation experiments to evaluate the efficiency of the different estimators and discuss the strengths and weakness of each method.

1.5 Outline of the Dissertation

The remainder of the dissertation is organized as follows.

In Chapter 2, we give an introduction to the Lévy process and the VG process.

A Fast Fourier Transform (FFT) method of pricing options through characteristic functions is also introduced.

In Chapter 3, we provide an introduction to multidimensional dependent non-Gaussian models including FGC, LM and VGC.

In Chapter 4, we introduce gradient estimation techniques, including FD, IPA, LR and GL.

In Chapter 5, we present the first essay on constructing optimal portfolios of variance swaps by FGC, LM and VGC.

In Chapter 6, we present the second essay on estimating Greeks of options with underlying assets following a VG process.

Some derivations, mathematical calculations and computer codes are included in the Appendices for completeness.

Chapter 2

Preliminaries

2.1 Introduction to the Lévy Process

A Lévy process is a superposition of a Wiener process and a number of independent Poisson processes. In this section, we give a brief introduction to the Lévy process and its properties.

2.1.1 Lévy Process and Infinitely Divisible Distribution

Let $D(f)$ be the domain of $f(x)$. A function $f(x)$ that has a limit from left and is continuous from right is called a cádlág function, i.e., for $x \in D(f)$,

- the left limit $f(x_-) := \lim_{t \to x^-} f(t)$ exists.
- the right limit $f(x_+) := \lim_{t \to x^+} f(t)$ exists and equals $f(x)$.

In other words, f is cádlág if it is right-continuous with left limits.

Definition 1 (**Lévy Process**). *A cádlág stochastic process $(X_t)_{t \geq 0}$ on probability space (Ω, F, P) starting with $X_0 = 0$ is a Lévy process if it satisfies the following properties:*

- *X_t has independent increments, i.e., for any $n \geq 1$ and $0 \leq t_0 \leq t_1 \leq ... \leq t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \cdots, X_{t_n} - X_{t_{n-1}}$ are independent.*

- *X_t has stationary increments, i.e., the law of $X_{t+h} - X_t$ does not depend on t.*
- *X_t has stochastic continuity, i.e., for any $\epsilon > 0$, $\lim_{h \to 0} P(|X_{t+h} - X_t| \leq \epsilon) = 0$.*

The sample paths of a Lévy process are almost surely continuous from the right and have left limits, which is easily verified for the Poisson process. Moreover, if we sample a Lévy process at any fixed time intervals with equal increments, that is at $0, \Delta t, 2\Delta t, ...$, we obtain a random walk.

Example 1. *A Gamma process $\gamma_t^{(\mu,\nu)}$ is a Lévy process with independent Gamma increments. It is a pure-jump increasing process which make it possible to be a subordinator. The marginal distribution of a Gamma process is a Gamma distribution with mean μt and variance νt.*

Lévy processes can also be defined through the properties of infinitely divisible distributions. Let's first define the infinite divisibility.

Definition 2 (**Infinite divisibility**). *A probability distribution function F defined on R^d is infinitely divisible if for any integer $n \geq 2$, there exists n i.i.d random variables $Z_1, Z_2, \ldots, Z_n$ such that*

$$Z_1 + Z_2 + \cdots + Z_n$$

has distribution F.

Example 2. *Let $X \sim N(\mu, \sigma^2)$ and let Z_k, $k = 1, \ldots, n$, be i.i.d random variables with distribution $N(\mu/n, \sigma^2/n)$. Thus,*

$$X \stackrel{d}{=} \sum_{k=1}^{n} Z_k.$$

In other words, the normal distribution is infinitely divisible.

Example 3. *Let $X \sim$ Unif$(0,1)$, i.e., X is uniformly distributed on $(0,1)$. There do not exist i.i.d random variables Z_k, $k = 1, \ldots, n$, such that*

$$X \overset{d}{=} \sum_{k=1}^{n} Z_k.$$

In other words, the uniform distribution is not infinitely divisible.

Two propositions taken from Cont and Tankov [24] are shown as follows.

Proposition 2.1.1 (**Lévy process and infinite divisibility**). *Let X_t be a Lévy process. Then for every t, X_t has an infinite divisible distribution. Conversely, if a probability distribution F is infinitely divisible, then there exists a Lévy process X_t such that the distribution of X_1 is given by F.*

More precisely, if X_t is a Lévy process and $\phi_X(u)$ is called the characteristic function of the random variable X_1, then for $s, t \geq 0$, the characteristic function of $X_{t+s} - X_s$ is $(\phi_X(u))^t$ and it turns out that there is $\psi_X(u)$, called the characteristic exponent of X, such that

$$\phi_X(u) = \exp[\psi_X(u)].$$

The equation above established at time $t = 1$ can be generalized to any time point $t > 0$ of the process.

Proposition 2.1.2. *Let $(X_t)_{t\geq 0}$ be a Lévy process on R^d. Then there exists a continuous function $\psi : R^d \to R$ called the characteristic exponent of X such that*

$$\phi_{X_t}(u) = E\left[e^{iuX_t}\right] = e^{t\psi_{X_1}(u)}, \quad z \in R^d.$$

2.1.2 Lévy-Khintchine Representation

The Lévy-Khintchine Representation connects the property of infinitely divisible distribution and the Lévy processes. From the representation, we can study the Lévy processes simply from these infinitely divisible distributions which can be represented in the form of characteristic functions. Such representation is called the Lévy-Khintchine formula.

Theorem 2.1.3 (**Lévy-Khintchine**). *Let $(X_t)_{t\geq 0}$ be a Lévy process on R^d. There exists $\gamma \in R^d$, a positive definite matrix A and a measure $\hat{\nu}$ such that*

$$E[\exp(iu \cdot X_t)] = \exp(t\psi(u)),$$

for $u \in R^d$. The triplet $(A, \hat{\nu}, \gamma)$ is the characteristic triplet of $(X_t)_{t\geq 0}$, and

$$\psi(u) = -\frac{1}{2}u^T Au + i\gamma \cdot u + \int_{R^d} (e^{iu\cdot X_t} - 1 - iu \cdot x1_{|x|\leq 1})\hat{\nu}(dx).$$

Conversely,

let F be an infinitely divisible distribution on R^d. Its characteristic function can be represented as:

$$\begin{aligned} E[e^{iu\cdot X_t}] &= e^{t\psi(u)}, u \in R^d \\ with \quad \psi(u) &= -\frac{1}{2}u^T Au + i\gamma \cdot u + \int_{R^d} (e^{iu.X_t} - 1 - iu \cdot x1_{|x|\leq 1})\hat{\nu}(dx), \end{aligned}$$

where A is a symmetric positive $d \times d$ matrix, $\gamma \in R^d$ and $\hat{\nu}$ is called the Lévy measure of the distribution F satisfying:

$$\int_{|x|\leq 1} |x|^2 \hat{\nu}(dx) < \infty,$$

$$\int_{|x|\leq 1} \hat{\nu}(dx) < \infty.$$

Recall that the total variation of a function $f : [a, b] \to R^d$ is defined by

$$\sup_{a=t_0\le t_1\ldots\le t_n=b} \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|.$$

Although a compound Poisson process is of finite variation, a general Lévy process may be of infinite variation without further conditions added.

Proposition 2.1.4. *A Lévy process is of finite variation if and only if its characteristic triplet $(A, \hat{\nu}, \gamma)$ satisfies:*

$$A = 0 \quad \text{and} \quad \int_{|x|\le 1} |x|\hat{\nu}(dx) < \infty.$$

2.1.3 Lévy-Itô decomposition

Theorem 2.1.5 (Lévy-Itô decomposition)**.** *Let $X_t, t \ge 0$ be a Lévy process on R^d. Then there exist a vector γ and a Brownian motion $B_t, t \ge 0$ in R^d with covariance matrix A such that the Lévy process X_t can be decomposed into three parts*

$$X_t = X_t^B + X_t^l + \lim_{\epsilon\downarrow 0} X_t^\epsilon. \tag{2.1}$$

Here $X_t^B = \gamma t + B_t$ is a $d-$dimensional continuous Gaussian Lévy process with drift γ and covariance matrix A;

$$X_t^l = \int_{|x|\ge 1, s\in[0,t]} xN(ds, dx)$$

is a compound Poisson process with jump sizes $|x| \ge 1$; $N(ds, dx)$ is a Poisson random measure on $R^+ \times (R^d \backslash \{0\})$;

$$\tilde{X}_t^\epsilon = \int_{\epsilon\le|x|\le 1, s\in[0,t]} x\tilde{N}(ds, dx)$$

is a compensated compound Poisson process, where $\tilde{N}(ds, dx) = N(ds, dx) - \hat{\nu}(dx)ds$, *and* $\hat{\nu}$ *is a Lévy measure on* $R^d \backslash \{0\}$ *satisfying* $\int\limits_{R^d \backslash \{0\}} (1 \wedge |x|^2)\hat{\nu}(dx) < \infty$.

The Lévy-Itô decomposition implies that every Lévy process consists of two parts: a Brownian motion with drift and a possibly infinite sum of independent compound Poisson processes. The component X_t^l is a discontinuous compound Poisson process incorporating with finite number of "large jumps" whose absolute value is greater than 1. Thus X_t^l can be written as a summation of an almost surely finite number of jumps; that is,

$$X_t^l = \sum_{0 \le s \le t}^{\Delta X_s \ge 1} \Delta X_s.$$

The component X_t^ϵ is a discontinuous compound Poisson process with possibly infinitely many "small jumps"; that is,

$$X_t^\epsilon = \sum_{0 \le s \le t}^{\epsilon \le \Delta X_s \le 1} \Delta X_s = \int\limits_{\epsilon \le |x| < 1, s \in [0,t]} xN(ds, dx).$$

The summation may not converge in general. To make the compound Poisson process be a martingale, one needs to compensate it so that the summation will not explode. Therefore, we get the compensated compound Poisson process $\tilde{X}_t^\epsilon$ in the decomposition (2.1).

2.2 Variance Gamma Process

The Variance Gamma (VG) process was introduced to the finance community as a model for pricing assets returns and option pricing in Madan and Seneta [60]. The VG process is a Lévy process of finite variation with infinite but relatively low

activity of small jumps. It is a pure jump process with infinite activity, which was first introduced as a Gamma-time-changed Brownian motion by Madan and Seneta [60] and developed in Madan, Carr and Chang [57] as a difference of two independent Gamma processes.

2.2.1 Definition of VG Process

The VG process is a Lévy process which can be expressed as a Gamma-time-changed Brownian motion or a difference of two independent Gamma processes. Let $\gamma_t^{(\mu,\nu)}$ be the Gamma process with drift parameter μ and variance parameter ν. The VG process can be defined in the following two ways.

The VG process can be defined by setting the subordinator to be a Gamma process.

Definition 3 (Variance Gamma)**.** *The VG process can be defined as* ***Gamma-time-changed Brownian motion*** *with the subordinator being a Gamma process. Let W_t denote the standard Brownian motion, $B_t^{(\mu,\sigma)} = \mu t + \sigma W_t$ denote the Brownian motion with constant drift rate μ and volatility parameter σ, $\gamma_t^{(\nu)} = \gamma_t^{(1,\nu)}$ be the Gamma process with drift parameter $\mu = 1$ and variance parameter ν. The representation of the VG process is:*

$$X_t = B_{\gamma_t^\nu}^{(\theta,\sigma)} = \theta\gamma_t^{(\nu)} + \sigma W_{\gamma_t^{(\nu)}}. \tag{2.2}$$

Alternatively, the VG process can be defined as a difference of two Gamma processes.

Definition 4 (Variance Gamma)**.** *The VG process is the* ***difference of two Gamma processes****. The representation of the VG process could be defined as:*

$$X_t = \gamma_t^{(\mu_+,\nu_+)} - \gamma_t^{(\mu_-,\nu_-)}, \tag{2.3}$$

where $\mu_\pm = (\sqrt{\theta^2 + 2\sigma^2/\nu} \pm \theta)/2$, *and* $\nu_\pm = \mu_\pm^2 \cdot \nu$.

2.2.2 Properties of VG Process

For the VG process with corresponding parameters (σ, ν, θ), the Lévy density is provided by

$$k(x) = \frac{1}{\nu|x|} \exp\left(\frac{\theta}{\sigma^2}x - \frac{1}{\sigma}\sqrt{\frac{2}{\nu} + \frac{\theta^2}{\sigma^2}}|x|\right),$$

where $\nu, \sigma > 0$.

Let $\phi_X(u) = E[e^{iuX}]$ denote the characteristic function of a random variable X. The characteristic function for the VG process VG(σ, ν, θ, t) at time t is

$$\phi_{VG}(u, \sigma, \nu, \theta, t) = (1 - iu\theta\nu + 0.5\sigma^2\nu u^2)^{-t/\nu}, \tag{2.4}$$

which can be expressed in two forms

$$\begin{aligned} \phi_{X_t}(u) &= \Big((1 - iu\nu\mu_+)(1 + iu\nu\mu_-)\Big)^{-t/\nu}, \\ \phi_{X_t}(u) &= \Big(1 - i\nu(u\theta + i\sigma^2u^2/2)\Big)^{-t/\nu}, \end{aligned}$$

reflecting the two representations above.

Under the risk-neutral measure, with no dividends and with constant risk-free interest rate r, the stock price is given by

$$S_t = S_0 \exp((r + \omega)t + X_t),$$

where $\omega = \ln(1 - \theta\nu - \sigma^2\nu/2)/\nu$ is the parameter that makes the discounted asset price a martingale, such that $E[e^{-rt}S_t] = S_0$.

Theorem 2.2.1. *(Madan, Carr and Chang [57]) The density function of the log-price* $z = \ln(S_t/S_0)$ *is*

$$h(z) = \frac{2\exp(\theta x/\sigma^2)}{\upsilon^{t/\upsilon}\sqrt{2\pi}\sigma\Gamma(\frac{t}{\upsilon})}\left(\frac{x^2}{2\sigma^2/\upsilon + \theta^2}\right)^{\frac{t}{2\upsilon}-\frac{1}{4}} K_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{x^2(2\sigma^2/\nu + \theta^2)}\right), \quad (2.5)$$

where $K_{\frac{t}{\nu}-\frac{1}{2}}$ *is the modified Bessel function of the second kind, and* $x = z - rt - t/\nu\ln(1 - \theta\nu - \sigma^2\nu/2)$.

A closed-form expression for pricing the European call option under VG with strike K is derived in Madan, Carr and Chang [57]. The form of the option price formula is similar to the Black-Scholes formula but requires calculation of the modified Bessel function of the second kind.

Theorem 2.2.2. *Assuming the stock price follows a VG process, the European call option price on the stock under the risk-neutral measure (with risk neutral parameters* σ, ν, θ*) is*

$$\begin{aligned} C(S_0; k, t) &= S_0\hat{K}\left(d\sqrt{\frac{1-c_1}{\nu}}, (\alpha+s)\sqrt{\frac{1-c_1}{\nu}}, \frac{t}{\nu}\right) \\ &- K\exp(-rt)\hat{K}\left(d\sqrt{\frac{1-c_2}{\nu}}, \alpha s\sqrt{\frac{\nu}{1-c_2}}, \frac{t}{\nu}\right), \end{aligned} \quad (2.6)$$

where

$$d = \frac{1}{s}\left(\ln\left(\frac{S_0}{k}\right) + rt + \frac{t}{\nu}\ln\left(\frac{1-c_1}{1-c_2}\right)\right),$$

$$\alpha = -\frac{\theta}{\sigma\sqrt{1 + (\theta/\sigma)^2\nu/2}},$$

$$c_1 = \frac{\nu(\alpha+s)^2}{2},$$

$$c_2 = \frac{\nu\alpha^2}{2},$$

and $\hat{K}$ is defined as the modified Bessel function of the second kind.

Though the option price under the VG model can be calculated through this closed form, it would be time-consuming and complicated. A more efficient way to price the option is a numerical method called the Carr-Madan FFT method, which is calculated simply from the characteristic function of the VG process. This method will be introduced in the next section.

2.2.3 Simulation of the VG Process

There are two main techniques to simulate the VG process: **sequential sampling** and **bridge sampling**. The sequential sampling technique is called incremental path construction in Jäckel [47], whereas bridge sampling is described in Avramidis and L'Ecuyer [5]. Fu [33] reviewed several ways to simulate the VG process including sequential sampling and bridge sampling.

Sequential Sampling

In this paper, we only introduce the two main methods of simulating the VG process, which are based on the two representations presented in the previous sections. We present two algorithms for sequential sampling of the VG process. Algorithm 1 is based on the representation of the VG process as a Gamma time-changed Brownian motion; Algorithm 2 is based on the representation of the VG process as a difference of two Gamma processes.

Algorithm 1 Simulating VG as Gamma Time-Changed Brownian Motion

Input: the VG parameters (σ, ν, θ); time increments $\Delta t_1, ..., \Delta t_N$ s.t. $\sum_{i=1}^{N} \Delta t_i = T$.

Initialize: $X_0 = 0$, $t_0 = 0$.

Loop: for $i = 1$ to N:

- Generate $\Delta G_i \sim \Gamma(\Delta t_i/\nu, \nu)$.
- Generate a standard normal random variable $Z_i \sim N(0, 1)$.
- Return $t_i = t_{i-1} + \Delta t_i$ and $X_{t_i} = X_{t_{i-1}} + \sigma\sqrt{\Delta G_i} Z_i$.

Algorithm 2 Simulating VG as Differences of two Gamma processes

Input: the VG parameters (σ, ν, θ); time increments $\Delta t_1, ..., \Delta t_N$ s.t. $\sum_{i=1}^{N} \Delta t_i = T$.

Initialize: $X_0 = 0$, $t_0 = 0$.

Loop: for $i = 1$ to N:

- Generate $\Delta\gamma_i^- \sim \Gamma(\Delta t_i/\nu, \nu\mu_-)$.
- Generate $\Delta\gamma_i^+ \sim \Gamma(\Delta t_i/\nu, \nu\mu_+)$.
- Return $t_i = t_{i-1} + \Delta t_i$ and $X_{t_i} = X_{t_{i-1}} + \Delta\gamma_i^+ - \Delta\gamma_i^-$.

Bridge Sampling

Ribeiro and Webber [72] introduced bridge sampling for the time-changed Brownian motion representation along with stratified sampling and quasi-Monte Carlo to further reduce variance. Avramidis and L'Ecuyer [5] introduced bridge sampling for the difference of Gammas representation combined with randomized quasi-Monte Carlo. We present two algorithms for sampling the VG process. Algorithm 3 is based on the representation of the VG process as a Gamma time-changed Brownian motion; Algorithm 4 is based on the representation of the VG process as a difference of two Gamma processes.

2.3 The Fast Fourier Transform Method and Option Pricing

The Black-Scholes model provides a closed-form solution for the European call option price. However, for the Lévy models, the density functions are complicated and it is not easy to calculate the option price through the density function. Carr and Madan [20] first described a new approach for numerically determining the option price through the characteristic function, which is designed to to value the option price by the fast Fourier transform (FFT) method efficiently. The analytical characteristic function of the risk-neutral density is assumed to be known. Using a simple analytic expression for the Fourier transform of the option value or its time value, the fast Fourier transform (FFT) can be applied to solve numerically the option price. We refer to this method as Carr-Madan FFT method.

Algorithm 3 Simulating VG via Brownian (Gamma-Time-Changed) Bridge

Input: the VG parameters (σ, ν, θ); number of bridges $N = 2^M (T = t_N)$.

Initialize: $X_0 = 0, \gamma_0 = 0, t_0 = 0$.

Generate $\gamma_{t_N} \sim \Gamma(t_N/\nu, \nu)$, $X_{t_N} \sim N(\theta\gamma_{t_N}, \sigma^2\gamma_{t_N})$ independently.

Loop: for $k = 1$ to $M : n \leftarrow 2^{M-k}$;

Loop from $j = 1$ to 2^{k-1} :

- $i \leftarrow (2j-1)n$;
- Generate $Y_i \sim \beta((t_i - t_{i-n})/\nu, (t_{i+n} - t_i)/\nu)$ independent of past random variables;
- $\gamma_{t_i} = \gamma_{t_{i-n}} + [\gamma_{t_{i+n}} - \gamma_{t_{i-n}}]Y_i$;
- Generate $Z_i \sim N(0, [\gamma_{t_{i+n}} - \gamma_{t_i}]\sigma^2 Y_i)$ independent of random variables;
- Return $X_{t_i} = Y_i X_{t_{i+n}} + (1 - Y_i)X_{t_{i-n}} + Z_i$.

Algorithm 4 Simulating VG via Difference-of-Gammas Bridge

Input: the VG parameters (σ, ν, θ); number of bridges $N = 2^M (T = t_N)$.

Initialize: $\gamma_0^+ = \gamma_0^- = 0$

Generate $\gamma_{t_N}^+ \sim \Gamma(t_N/\nu, \nu\mu_+)$, $\gamma_{t_N}^- \sim \Gamma(t_N/\nu, \nu\mu_-)$ independently.

Loop: for $k = 1$ to $M : n \leftarrow 2^{M-k}$;

Loop from $j = 1$ to 2^{k-1} :

- $i \leftarrow (2j-1)n$;
- Generate $Y_i^+, Y_i^- \sim \beta((t_i - t_{i-n})/\nu, (t_{i+n} - t_i)/\nu)$ independently;
- $\gamma_{t_i}^+ = \gamma_{t_{i-n}}^+ + [\gamma_{t_{i+n}}^+ - \gamma_{t_{i-n}}^+]Y_i^+$, $\gamma_{t_i}^- = \gamma_{t_{i-n}}^- + [\gamma_{t_{i+n}}^- - \gamma_{t_{i-n}}^+]Y_i^-$;
- Return $X_{t_i} = \gamma_{t_i}^+ - \gamma_{t_i}^-$.

2.3.1 The Carr-Madan FFT Method

Much literature has successfully applied Fourier analysis to determine options price, for example, Bakshi and Chen [6], Scott [76], Heston [43], etc. The Carr-Madan FFT method evaluates the value of an option by taking an inverse Fourier transform of the characteristic function of the log price. This method is much faster and easier than the analytic formula which contains a modified Bessel function of the second type. It should be pointed out that this method introduces a dampening factor to deal with the singularity of the integrand by multiplying the call pricing function. A sketch of the Carr-Madan FFT method is presented in the following.

Let k be the log of the strike value K, let $C_T(k)$ be the value of a call option with maturity T and strike price $\exp(k)$, and let the risk-neutral density of the log

price be $q_T(s)$. The characteristic function of this density is

$$\phi_T(u) = \int_{-\infty}^{\infty} e^{ius} q_T(s) ds.$$

To make the call option price function square integrable, we modify the call price by multiplying a damping factor $\exp(\alpha k)$ where $\alpha > 0$. Let $\psi_T(v)$ be the Fourier transform of the modified call price $c_T(k) = \exp(\alpha k) C_T(k)$:

$$\psi_T(v) = \int_{-\infty}^{\infty} e^{ivk} c_T(k) dk.$$

Integration gives

$$\psi_T(v) = \frac{e^{-rT}\phi_T(v - (\alpha+1)i)}{\alpha^2 + \alpha - v^2 + i(2\alpha+1)v}.$$

Therefore, the call price is

$$C_T(k) = \frac{\exp(-\alpha k)}{\pi} \int_0^{\infty} \psi_T(v) dv, \tag{2.7}$$

where $\alpha > 0$.

Using the trapezoidal rule for the integral in (2.7), the call price can be approximated by

$$C_T(k) \approx \frac{\exp(-\alpha k)}{\pi} \sum_{j=1}^{N} e^{-iv_j k} \psi_T(v_j) \eta. \tag{2.8}$$

The approximation is described as follows. First substitute a for the upper limit for the integral in (2.7) and set $v_j = (j-1)\eta$, then $N = a/\eta$. By choosing the step size of the log strike $k = \lambda$, i.e., $k_u = -b + \lambda(u-1)$, for $u = 1, ..., N$, we can approximate the call price as

$$C_T(k_u) \approx \frac{\exp(-\alpha k)}{\pi} \sum_{j=1}^{N} \exp[-i\lambda\eta(j-1)(u-1)] e^{ibv_j} \psi_T(v_j) \eta. \tag{2.9}$$

Moreover, assuming $\lambda\eta = \frac{2\pi}{N}$ and applying Simpson's rule, we can write the call price as

$$C_T(k_u) \approx \frac{exp(-\alpha k)}{\pi} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(u-1)} e^{ibv_j} \psi_T(v_j)\eta \frac{3 + (-1)^j - \delta_{j-1}}{3}, \qquad (2.10)$$

where δ_n is the Kronecker delta function which equals 1 for $n = 1$ and 0 otherwise.

Recall that the FFT is an efficient algorithm to compute the sum

$$\omega(k) = \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} x(j),$$

where N is a power of 2. Thus we could employ the FFT to calculate the call price in (2.10). By choosing the appropriate α and η, the Carr-Madan FFT method can calculate the call price efficiently. For a single run, this method compute the option prices across all the strikes. One more benefit of this method is that it only requires the characteristic function to calculate the option price.

Chapter 3

Multidimensional Dependent Non-Gaussian Models

3.1 Overview

It is well known that joint laws of asset returns are required in many applications of financial engineering. It is generally recognized in a vast literature (Jondeau, Poon and Rockinger [48], Menn, Fabozzi and Rachev [68], Boyarchenko and Levendorskii [12]) that the marginal distribution of each asset individually is non-Gaussian. As a result, many studies of multidimensional dependent non-Gaussian models have been conducted, and several models and methods have been proposed.

In this chapter, we provide an introduction to three multidimensional dependent non-Gaussian models: (1) the full rank Gaussian copula (FGC) proposed by Malevergne and Sornette [63]; (2) the Lévy Mixture implemented by Madan and Yen [61]; (3) the Variance Gamma correlated model proposed in Eberlein and Madan [30] and employed in Madan [55] in pricing options on a basket of stocks. The particular advantage of these models is their capability of estimating high dimensional non-Gaussian models by reducing the problem to a suitable sequence of univariate estimations.

3.2 Full Rank Gaussian Copula (FGC)

The key point of the FGC model is to transform the data which is assumed to follow a multidimensional non-Gaussian distribution to a series of dependent Gaussian variates with a possibly full rank correlation matrix. In this section, the FGC model is introduced first, and then the procedure of how to estimate through this model is explained.

3.2.1 Introduction to the FGC model

The closed form of the joint probability law of the multidimensional dependent non-Gaussian variates is difficult to obtain, even though the correlation or covariance of the variates can be calculated without any difficulty.

In the full rank Gaussian copula (FGC) model, each asset return is assumed to be a nonlinear transform of a set of standard normal variates which are correlated with each other. Based on this assumption, we transform the data into standard normal variates, and then calculate the covariance. Accordingly, the joint multivariate law of the transformed standard normal variates, which are correlated with each other due to the correlation of the original non-Gaussian variates, can be obtained easily.

Let $X = (X_1, X_2, ..., X_N)$ be an $N-$dimensional vector. Let $F_{X_i}(x)$, $i = 1, 2, ..., N$, be the corresponding marginal distributions for $X_i, i = 1, 2, ..., N$. That is $P(X_i \leq x) = F_{X_i}(x)$. Assume the F_{X_i} are continuous. Let Z_i be a standard normal variate, and let Φ be the cumulative distribution function of a standard

normal random variable.

By setting the probability integral transforms of X_i and Z_i equal to each other, that is,

$$\Phi(Z_i) = F_{X_i}(X_i),$$

we can transform continuous non-Gaussian variates X_i to standard normal variates Z_i. Due to the difficulty of calculating the value of cumulative distribution function (c.d.f.) of VG vairate X_i, we employ the Carr-Madan FFT method to calculate the c.d.f from the characteristic function. Simple integration leads to the proposition below:

Proposition 3.2.1. *Let X be a random variable with characteristic function $\psi(x)$, and $\exp(-\alpha x)$ be the dampening factor. Thus the cumulative distribution function $F_X(x)$ can be calculated from:*

$$F_X(x) = \exp(\alpha x)\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-iux)\frac{\psi(u+i\alpha)}{-iu+\alpha}du.$$

It turns out that the Carr-Madan FFT method can be applied to approximate the c.d.f. using the characteristic function.

Moreover, this transform is nonlinear and can be written as

$$Z_i = \Phi^{-1}[F_{X_i}(x_i)]. \tag{3.1}$$

Since the non-Gaussian variates X_i, $i = 1, 2, ..., N$, are correlated with each other, so are the transformed standard normal variates Z_i. The correlation matrix of Z_i is required to calculate the joint multivariate normal law. On top of that, the joint

density function of X can be calculated through the joint density function of Z by a change of variables. In contrast, the transform from Z_i to X_i takes the inverse of the transform from X_i to Z_i, that is,

$$X_i = F_{X_i}^{-1}(\Phi(Z_i)). \tag{3.2}$$

To calculate the inverse of the c.d.f. of $F_{X_i}(\cdot)$ in equation (3.2), we employ the Carr-Madan FFT method to calculate a series of grids of values of $F_{X_i}(s)$, where $0 \leq s \leq 1$ and then use interpolation to estimate the X_i from $\Phi(Z_i)$ by equation (3.2).

3.2.2 Estimation Procedures for FGC

Assume we have T dependent $N-$dimensional data $X_t = (X_{1t}, X_{2t}, ..., X_{Nt})$, $t = 1, ..., T$. After transforming the data X_t to have a zero mean value (also called centering), we apply maximum likelihood estimation (MLE) to estimate the corresponding parameters $\sigma_i, \nu_i, \theta_i$ from the univariate data on each component of X_t. Accordingly, we obtain a series of estimated marginal Variance Gamma parameters in a matrix:

$$\hat{\sigma}_i, \hat{\nu}_i, \hat{\theta}_i, i = 1, ..., N.$$

Applying the transform in (3.1) to X_{it}, one can obtain the standard normal variates Z_{it}, that is,

$$Z_{it} = \Phi^{-1}(F_{\mathrm{VG}}(X_{it}; \hat{\sigma}_i, \hat{\nu}_i, \hat{\theta}_i)). \tag{3.3}$$

It is easy to see that the covariance of Z_{it} can be estimated directly.

With the covariance of Gaussian variates Z_t, we can easily simulate correlated multidimensional standard normal variates $Z_s = (Z_{1s}, Z_{2s}, ..., Z_{Ns})$, $s = 1, 2, ..., M$ on $M = 10000$ sample paths. Then by plugging newly generated normal distributed data Z_s back into the transform in (3.3), we can get newly generated correlated Variance Gamma distributed data $X_{is}, i = 1, 2, ..., N$, for $s = 1, 2, ..., M$ by

$$X_{js} = F_{\mathrm{VG}}^{-1}(\Phi(Z_{js}), \sigma_j, \nu_j, \theta_j). \tag{3.4}$$

The inverse of F_{VG} can be approximated through interpolation by Carr-Madan FFT method.

3.3 Lévy Mixture (LM)

The main idea of the LM model is derived from the independent component analysis technique to multidimensional dependent non-Gaussian variates to a transform or a linear mixture of independent non-Gaussian variates. In this section, the LM model is introduced first, and then the procedure of how to estimate through this model is explained.

3.3.1 Introduction to the LM model

Recogning the fact that multivariate Gaussian variates with a nonsingular covariance matrix can be viewed as a linear combination of an equal number of independent standard normal random variates, we generalize the Gaussian variates to a linear mixture of independent non-Gaussian variates that are infinitely divisible. In this dissertation, we only consider the case that non-Gaussian variates are

assumed to follow a Variance Gamma distribution. This model has been employed by Madan and Yen [61] to optimize portfolios for asset allocation.

Let $X = (X_1, X_2, ..., X_N)$ be an $N-$dimensional vector. It is postulated that there exist a mixing matrix A and independent variables Y_i such that the non-Gaussian vectors X is a transform (or a linear mixture) of independent variables Y, that is,

$$X = AY. \tag{3.5}$$

As mentioned earlier, we assume the independent variates Y_i follow a VG process.

The main difficulty of this model lies in how to identify the transform (mixing) matrix A and independent components Y. Madan and Yen [61] used a technique called independent components analysis (ICA) developed in the signal processing (Hyvärinen, Karhuen and Oja [46]), in particular the fast ICA algorithm developed in Hyvärinen [45]. The fast ICA algorithm can help to find the mixing matrix A and independent variates Y. The main idea of ICA is to decompose the observed time series of data into stationary independent components (ICs). Given the observed data, generally ICA contains two main steps: centering and whitening. The fast ICA algorithm employ a fixed point iteration to solve for the optimization problem. Details of the ICA can be found in Madan and Yen [61].

Since Y_i follows a VG distribution, it is straightforward to get the characteristic function of Y_i. The characteristic functions of independent variates Y_i are

$$\Psi_{Y_j}(u) = E[\exp(iuY_j)].$$

From equation (3.5), the joint characteristic function of X is

$$\Psi_X(u) = \prod_{j=1}^{N} \Psi_{Y_j}((A'u)_j),$$

for $j = 1, 2, ..., N$.

3.3.2 Estimation Procedure for LM

Assume we have $N-$dimensional dependent data vectors $X_t = (X_{1t}, X_{2t}, ..., X_{Nt})$, $t = 1, ..., T$. We have a $T \times N$ matrix of data X_t. Through the fast independent component analysis (ICA), we can generate new independent non-Gaussian distributed variates Y_t. Here we assume Y_t follow a VG distribution and the corresponding mixing matrix be A. We have

$$X_t = AY_t.$$

After that, we can apply the MLE method to estimate corresponding parameters $(\sigma_j, \nu_j, \theta_j)$ for the VG variates Y_j in each dimension, respectively. Next, we simulate 10000 sample paths new VG variates Y_s. By substituting back into equation (3.5), we can get the newly generated correlated variates X_s.

3.4 Variance Gamma Correlated (VGC)

In the VGC model, the non-Gaussian variates are assumed to follow a VG distribution, and the VG variates can be written as a Gamma-time-changed Brownian motion. Moreover, the correlation of VG variates is set in the Brownian motion part. In this section, the VGC model is introduced first, and then the procedure of how to estimate the parameters of this model is explained.

3.4.1 Introduction to the VGC model

Assume the marginals follow the law of VG. Thus, each asset X_i can be written as a Gamma-time-changed Brownian motion at unit time. We put the correlation of X_i in the Brownian motion part. Let the Gamma process $G_i(t), t \geq 0, i = 1, ..., N$ be the subordinator of the VG process $X_i(t)$. We assume that the $G_i(1)$ are a sequence of independent Gamma variates with unit mean at time $t = 1$, that is, $E[G_i(1)] = 1$ and $\nu_i^2 = \text{Var}[G_i(1)]$. Under these assumptions, the multidimensional dependent non-Gaussian variates $X_i(t)$ could be written as:

$$X_i(t) = \theta_i(G_i(t) - t) + \sigma_i W_i(G_i(t)), \tag{3.6}$$

where $\theta_i, \sigma_i \geq 0$, and the Brownian motion $(W_i(t), t \geq 0)$ is independent of the subordinator $G_i(t)$. Then the density of the $G_i(t)$ at time $t = 1$ is:

$$f(x) = \frac{1}{\nu_i^{\frac{1}{\nu_i}}\Gamma(\frac{1}{\nu_i})} x^{\frac{1}{\nu_i}-1} e^{-\frac{x}{\nu_i}}.$$

If $X_i(t)$ are independent random variables when $t = 1$, then the $X_i(t)$ are the independent Variance Gamma processes.

Now we consider the dependent case, where $X_i(t), i = 1, \ldots, N$, are correlated processes. There are many kinds of multidimensional dependent models. But in this chapter, we only consider the dependence which can be introduced by only correlating the Brownian motion part while keeping the time changing subordinators independent. The correlated Lévy processes $X_i(t)$ at unit time, $X_i(1) = X_i$, can be written as

$$X_i(1) = \theta_i(G_i(1) - 1) + \sigma_i W_i(G_i(1)), \tag{3.7}$$

where $G_i = G_i(1), i = 1, \ldots, N$ are independent, $W_i(G_i(1)), i = 1, \ldots, N$ are correlated, and standard variates $Z_i = W_i(1)$ are correlated. Let Z_i be standard normal variates with zero mean and correlation ρ_{ij} between Z_i and Z_j for $i \neq j$, i.e.,

$$E[Z_i] = 0$$

and

$$\rho_{i,j} = \text{Corr}(Z_i, Z_j) = \text{Cov}(Z_i, Z_j). \tag{3.8}$$

Therefore, equation (3.7) can be written as

$$X_i = \theta_i(G_i - 1) + \sigma_i\sqrt{G_i}Z_i. \tag{3.9}$$

By equation (3.9), $X_i, i = 1, \cdots, N$ have zero means, i.e., $E[X_i] = 0$ for each i. The covariance of X_i and X_j can be simplified as

$$\begin{aligned} \text{Cov}(X_i, X_j) &= E[X_iX_j] \\ &= \sigma_i\sigma_j E[\sqrt{G_i}]E[\sqrt{G_j}] \cdot E[Z_iZ_j], \end{aligned}$$

where $E[Z_iZ_j] = \text{Cov}(Z_i, Z_j) + E[Z_i]E[Z_j] = \text{Cov}(Z_i, Z_j)$. Thus,

$$\text{Cov}(X_i, X_j) = \sigma_i\sigma_j \cdot \rho_{ij} \cdot E[\sqrt{G_i}]E[\sqrt{G_j}]. \tag{3.10}$$

The expectation of the square root of the Gamma process at unit time can be calculated through

$$\begin{aligned} E[\sqrt{G_i}] &= \int_0^\infty \frac{1}{\nu_i^{1/\nu_i}\Gamma(\frac{1}{\nu_i})}\sqrt{x}x^{\frac{1}{\nu_i}-1}e^{-\frac{x}{\nu_i}}dx \\ &= \frac{\sqrt{\nu_i}\Gamma(\frac{1}{\nu_i} + \frac{1}{2})}{\Gamma(\frac{1}{\nu_i})}, \end{aligned}$$

where

$$\Gamma(x) = \int_0^\infty u^{x-1}e^{-u}du.$$

3.4.2 Estimation Procedures for VGC

Assume we have $N-$dimensional data $X_t = (X_{1t}, X_{2t}, ..., X_{Nt})$, $t = 1, ..., T$. After centering the data X_t, we employ the maximum likelihood ratio (MLE) method to estimate corresponding parameters $\sigma_i, \nu_i, \theta_i$ for the univariates in each dimension of X_{it}. Then, the covariance matrix of X_i and X_j can be calculated directly. Furthermore, the correlation ρ_{ij} of Z_i and Z_j can be calculated by equation (3.10). Thus the correlated standard variates Z_s with correlation ρ_{ij} can be simulated on 10000 sample paths. By substituting the simulated Z_s back into equation (3.9), we can get the newly generated multidimensional correlated non-Gaussian variates X_s.

Chapter 4

Monte Carlo Simulation and Gradient Estimation

4.1 Overview

Monte Carlo simulation is a widely-used technique to study the impact of risk and uncertainty in financial engineering or other forecasting models. Monte Carlo methods were introduced to finance community in Hertz [42] in the context of corporate finance and first applied to derivative valuation in Boyle [13]. A review of gradient estimation techniques in financial engineering is given in Fu [33].

Gradient estimates are useful in hedging risks in markets in the finance community. Generally, the Monte-Carlo based gradient estimation technique contains the following several approaches: finite difference (for example: forward difference), infinitesimal perturbation analysis (IPA) and the likelihood ratio (LR) method, as well as the GL method (similar to the LR method) proposed and tested in Glasserman and Liu [39], which numerically approximates density functions from characteristic functions.

4.2 Gradient Estimation Techniques

Gradient estimation techniques can be applied to estimate gradients of different options, such as a European call option, or options based on a basket of options

(Mountain Range options). Gradients play an important role in hedging strategies since they in fact measure sensitivities of options with respect to corresponding parameters. For instance, Delta measures the sensitivity of option prices to spot prices and determined how many shares of stocks required to purchase or sell to offset risks from changes of stock prices. To calculate gradient estimates, we need to calculate the derivative of option prices with respect to these parameters separately.

Before calculating gradient estimators, we first set up the problem. We begin with $V(\xi)$, the objective function which depends on the parameter ξ. The gradients with respect to the parameter ξ can be calculated as

$$\frac{dV(\xi)}{d\xi},$$

which is the derivative of the objective function with respect to the corresponding parameter.

It is assumed that the performance measure must be estimated in this context. Accordingly, we will assume that the performance measure is an expectation of some sample performances. Therefore, the objective function of interest is assumed to be an expectation of the sample performance measure J, that is,

$$V(\xi) = E[J(\xi)] = E[J(X_1, X_2, \cdots, X_n; \xi)], \tag{4.1}$$

where $X = X_1, X_2, \cdots, X_n$ depends on ξ, n is the fixed number of random variables, and J is referred to as the sample performance. From the fact that Monte Carlo simulation is based on the law of large numbers, we can get a good estimate of the performance measure of interest by taking the average of the sample performance J

over many simulations. Therefore, the problem of gradient estimation is to estimate

$$\frac{dE[J(\xi)]}{d\xi}. \tag{4.2}$$

By law of the unconscious statistician, the expectation can be written as:

$$E[J(X)] = \int y dF_J(y) = \int J(x) dF_X(x), \tag{4.3}$$

where F_J is the distribution of J and F_X is the distribution of the input random variables X. Generally, F_J is not known explicitly; for other cases, simulation is not needed. Nevertheless, the distribution F_X is always known in the simulation since it is required to generate input processes to the simulation model.

The parameter of interest ξ can occur in two places: one is in the input random variables X; the other is the density function f_X of X. To be specific, we express the right hand side of equation (4.3) as:

$$E[J(X)] = \int_0^1 J(X(\xi; u)) du, \tag{4.4}$$

$$E[J(X)] = \int_{-\infty}^{\infty} J(x) f_X(x; \xi) dx, \tag{4.5}$$

where f_X is the probability density function of X. In the first case as shown in equation (4.4), the dependence of parameter ξ is path-wise from the input random variables X; while in the second case shown in equation (4.5), the dependence on the parameter ξ is in the distribution F_X. Considering that the dependence of the parameter ξ can be in two different ways, we have two different kinds of methods (called direct and indirect methods) to estimate the objective function.

A "brute-force" finite difference can be obtained by taking additional simulations at parameter value $(\xi + \Delta\xi)$. A forward difference estimate of (4.2) can be

calculated by subtracting simulated estimates of J at $\xi + \Delta\xi$ from simulated estimates of J at ξ, and dividing by the perturbation $\Delta\xi$. However, this would require additional simulations. Moreover, a suitable perturbation $\Delta\xi$ must be selected to trade off between variance and bias, since larger $\Delta\xi$ means lower variance but higher bias.

In this section, we provide an introduction to indirect methods and direct methods for estimating gradients.

4.2.1 Indirect Methods

Indirect methods for estimating a gradient at ξ is simply to use finite differences, i.e., perturbing the value of each component of θ separately while holding all the other components still. As mentioned before, there is a trade-off between bias and variance when selecting the appropriate value of perturbation $\Delta\xi$. If the perturbation is too small, the resulting estimator may be noisy, whereas a large perturbation is likely to lead to a large bias.

Assume for each ξ, after substituting the generated random variable $X(\xi)$ into the deterministic function $J(X, \xi)$, we can calculate the expectation $\hat{J}(x) = \hat{J}(x, \xi) = E[J(X, \xi)]$. The objective is to get the simulation estimate of $d/d\xi \hat{J}(x, \xi)$.

From the definition of finite difference including forward difference, backward

difference, and central difference, we can get the following formulas:

$$\frac{d}{d\xi}\hat{J}(\xi) = \lim_{h\to 0}\frac{\hat{J}(\xi+h)-\hat{J}(\xi)}{h} \tag{4.6}$$

$$= \lim_{h\to 0}\frac{\hat{J}(\xi)-\hat{J}(\xi-h)}{h} \tag{4.7}$$

$$= \lim_{h\to 0}\frac{\hat{J}(\xi+h/2)-\hat{J}(\xi-h/2)}{h} \tag{4.8}$$

From equation (4.6), the one-sided forward difference gradient estimator in the i-th direction is:

$$\frac{J(\xi+c_i e_i)-J(\xi)}{c_i},$$

where c_i is the scalar of perturbation in the $i-$th direction and e_i is the unit vector in the i-th direction.

From equation (4.7), the one-sided backward difference gradient estimator in the ith direction is:

$$\frac{J(\xi)-J(\xi-c_i e_i)}{c_i},$$

where c_i is the scalar of perturbation in the $i-$th direction and e_i is the unit vector in the $i-$th direction.

From equation (4.8), the two-sided symmetric difference estimator, or central difference estimator is given by

$$\frac{J(\xi+c_i e_i)-J(\xi-c_i e_i)}{2c_i}.$$

An additional gradient estimation technique designed for stochastic approximation is called the simultaneous perturbation (SP) estimator. The $i-$th component of the SP gradient estimator is given by:

$$\frac{J(\xi+c\Delta_i)-J(\xi-c\Delta_i)}{2c\Delta_i},$$

where c is the set of differences for each component, $\Delta = (\Delta_1, \cdots, \Delta_d)$ is a d dimensional vector of perturbations, which are generally assumed i.i.d. In fact, c is a diagonal matrix with the differences c_i on the diagonal.

4.2.2 Infinitesimal Perturbation Analysis

Infinitesimal perturbation analysis (IPA) estimates the sample path derivative of the parameters of interest.That is for $J(X(\xi, \omega))$ on the sample path ω,

$$\frac{dJ(X(\xi,\omega))}{d\xi} = \lim_{h\to 0} \frac{J(X(\xi+h,\omega)) - J(X(\xi,\omega))}{h} \quad \text{w.p.1}$$

To derive direct gradient estimators for IPA, we express the expectation as in equation (4.4), i.e.,

$$E[J(X)] = \int_0^1 J(X(\xi; u))du.$$

After taking the derivative of the equation with respect to the parameter ξ, we can derive the estimates for IPA. However, it should be pointed out that IPA estimates require the interchangeability condition which is easily satisfied when the performance function is continuous with respect to the given parameter. Assume that we can interchange the expectation and differentiation, the IPA estimate is:

$$\frac{dE[J(X)]}{d\xi} = \int_0^1 \frac{dJ(X(\xi;u))}{d\xi}du = \int_0^1 \frac{dJ}{dX}\frac{dX(\xi)}{d\xi}du.$$

Hence, the IPA estimator is:

$$\frac{dJ}{dX}\frac{dX(\xi)}{d\xi}. \tag{4.9}$$

From the theorem, the condition of uniform integrability of $\frac{dJ}{dX}\frac{dX(\xi)}{d\xi}$ must be satisfied to make the interchangeability possible. In other words, in order to

make the IPA estimator be an unbiased stochastic gradient estimator, we need the interchangeability of the derivative and expectation, i.e.,

$$\frac{dE[J(X)]}{d\xi} = E\left[\frac{dJ(X)}{d\xi}\right], \tag{4.10}$$

which is the condition for the unbiased IPA estimator exists. Let $X = (X_1, ..., X_N)$. To make it easy to explain, we start by assuming the parameter ξ only appears in X_1, which is independently of other input random variables. Thus, the IPA estimates can be written as:

$$\begin{aligned}\frac{dE[J(X)]}{d\xi} &= \int_0^1 \frac{dJ(X_1(\xi;u), X_2, ..., X_N)}{d\xi} du \\ &= \int_0^1 \frac{dJ}{dX_1}\frac{dX_1(\xi)}{d\xi} du.\end{aligned}$$

Hence, the IPA estimator is:

$$\frac{dJ}{dX_1}\frac{dX_1(\xi)}{d\xi}.$$

From the definition of derivative, the condition for making IPA estimator unbiased can be expressed as:

$$\lim_{\Delta\xi\to 0} E[g_{\Delta\xi}] = E[\lim_{\Delta\xi\to 0} g_{\Delta\xi}],$$

where $g_{\Delta\xi} = \frac{J(\xi+\Delta\xi)-J(\xi)}{\Delta\xi}$.

Then one question appears: how do we check interchangeability of the expectation and differentiation? Before answering this question, we recall the Dominated Convergence Theorem, given here in the form presented in Fu and Hu [31].

Theorem 4.2.1. ***(Dominated Convergence Theorem).*** *If* $\lim_{\Delta\xi\to 0} g_{\Delta\xi} = g$ *with probability 1, and there exits* $\epsilon > 0$, *such that* $|g_{\Delta\xi}| \leq K$ *with probability 1, for any*

$\Delta\xi \in [0, \epsilon]$ *with* $E[K] < \infty$ *and* K *is independent of* ξ*, then*

$$\lim_{\Delta\xi \to 0} E[g_{\Delta\xi}] = E[g].$$

By the dominated convergence theorem, the following lemma (also taken from Asmussen and Glynn [4]) makes the interchange possible.

Lemma 4.2.2. *Assume that* $J(X) = J(X, \xi)$ *is almost surely differentiable at* ξ_0 *and that* $J(X)$ *satisfies the Lipschitz condition*

$$|J(X, \xi_1) - J(X, \xi_2)| \leq |\xi_1 - \xi_2| M$$

for ξ_1, ξ_2 *in a non-random neighborhood of* ξ_0*, where* $E[M] < \infty$*. Then equation (4.10) holds at* $\xi = \xi_0$*.*

Proof. We write $\hat{J}(x) = \hat{J}(x, \xi) = E[J(\xi)]$ for some random variable $J(X, \xi)$ depending on ξ. Note

$$\begin{aligned} \hat{J}'(x, \xi_0) &= \lim_{h \to 0} \frac{\hat{J}(x, \xi_0 + h) - \hat{J}(x, \xi_0)}{h} \\ &= \lim_{h \to 0} E[\frac{J(X, \xi_0 + h) - J(X, \xi_0)}{h}] \\ &\leq \lim_{h \to 0} E[M]. \end{aligned}$$

Since

$$\lim_{h \to 0} \frac{J(X, \xi_0 + h) - J(X, \xi_0)}{h} = J'(\xi_0),$$

almost surely, we can get

$$\frac{d}{d\xi}\Big(E[J(\xi)]\Big)\Big|_{\xi_0} = E\Big[\frac{d}{d\xi} J(\xi)\Big|_{\xi_0}\Big],$$

by the dominated convergence theorem. □

4.2.3 Likelihood Ratio Method

The likelihood ratio (LR) method constructs estimators by taking the derivative of the probability density function with respect to the parameter of interest. The probability density function f_X of X is assumed to be differentiable with respect to the parameter of interest.

To derive direct gradient estimators for LR method, we have to express the expectation as in (4.5), i.e.,

$$E[J(X)] = \int_{-\infty}^{\infty} J(x) f_X(x;\xi) dx.$$

After taking the derivative of the equation with respect to the parameter ξ, we can derive the estimates for LR. However, it should be pointed out that the conditions for interchangeability required for LR differ from those for IPA.

Assuming we can interchange the expectation and differentiation, the LR estimate is:

$$\frac{dE[J(X)]}{d\xi} = \int_{-\infty}^{\infty} J(x) \frac{df(x;\xi)}{d\xi} dx = \int_{-\infty}^{\infty} J(x) \frac{d\ln f(x;\xi)}{d\xi} f(x;\xi) dx$$

and the estimator is

$$J(x) \frac{d\ln f(x;\xi)}{d\xi},$$

where $\frac{d\ln f(x;\xi)}{d\xi}$ is called the score function.

Let $X = (X_1, ..., X_N)$ be an N-dimensional input process. Assume that X_1 has marginal probability density function $f_1(\cdot;\xi)$ and that the joint p.d.f for the remaining input random variables $(X_2, ..., X_N)$ is given by $f_{-1}(\cdot)$ which is independent

of ξ. Accordingly, the corresponding LR estimate is:

$$\begin{aligned}\frac{dE[J(X)]}{d\xi} &= \int_{-\infty}^{\infty} J(x)\frac{df_1(x_1;\xi)}{d\xi}f_{-1}(x_2,...,x_N)dx \\ &= \int_{-\infty}^{\infty} J(x)\frac{d\ln f_1(x_1;\xi)}{d\xi}f(x)dx\end{aligned}$$

and the corresponding estimator is

$$J(X)\frac{d\ln f_1(X_1;\xi)}{d\xi}.$$

By dominated convergence theorem, the following lemma makes the interchange possible. See previous sentence.

Lemma 4.2.3. *Let $(f_\xi(x))_{\theta\in\Theta}$ be a family of densities on R such that, $f_\xi(x)$ is continuously differentiable in ξ with derivative $f'_\xi(x)$ for x, almost everywhere. Then*

$$\frac{d}{d\xi}\int J(x)f_\xi(x)dx = \int J(x)f'_\xi(x)dx,$$

for all ξ in a given open subintegral $\Theta_0 \subset \Theta$ provided that there exist p, q with $1/p + 1/q = 1$ such that

$$x \in L_q, \text{ and } |f'_\xi(x)| \leq M(x),$$

for some $M \in L_p$ and for all $\theta \in \Theta_0$ and for x almost everywhere.

Proof. Assume $(\xi-\epsilon, \xi+\epsilon) \subseteq \Theta_0$. For $|h| < \epsilon$, we have

$$\begin{aligned}\frac{1}{h}&\left[\int J(x)f_{\xi+h}(x)dx - \int J(x)f_\xi dx\right] \\ &= \int J(x)\frac{f_{\xi+h}(x)-f_\xi(x)}{h}dx \\ &= \int J(x)f'_{\xi+h^*(x)}(x)dx,\end{aligned}$$

for some $h^*(x) \in (-h, h)$. We also have

$$\lim_{h \to 0} J(x) f'_{\xi + h^*(x)}(x) dx = J(x) f'_{\xi}(x),$$

and

$$||J(x) f'_{\xi + h^*(x)}(x) dx||_{L_1} \leq ||J(x) M(x)||_{L_1}.$$

Then, we can apply the dominated convergence theorem to complete the proof.

□

4.2.4 GL method

An approach based on the LR method to estimate gradients was proposed in Glasserman and Liu [39], abbreviated here as the GL method. Unlike the LR method which directly uses the density function, the GL method numerically approximates the density function $g_\xi(x)$ and the derivative $(d/d\xi)g_\xi(x)$ through the characteristic function or the Laplace transform.

Let G_ξ be the distribution function associated with the density function g_ξ. The main idea of the GL method is to sample the input X from the distribution function through the inverse transform method, that is,

$$X = G_\xi^{-1}(U),$$

where U is uniformly distributed on $[0, 1]$. Once the density function and its derivative are numerically approximated, we can estimate the gradients similar to LR method.

The general steps of this GL method are follows: Pick a finite grid of x values, and pre-compute values of $G_\xi, g_\xi, \dot{g}_\xi$, through numerical transform inversions.

- Using the Abate-Whitt algorithm [1], each transform inversion is approximated using a finite weighted sum of transform values, given by

$$I_{\xi,x}^{N,h}(L_g) = \frac{he^{\sigma x}}{2\pi}L_g(\sigma) + \frac{he^{\sigma x}}{\pi}\sum_{k=1}^{N}\Big(\mathrm{Re}[L_g(\sigma+ikh)]\cos(khx) - \mathrm{Im}[L_g(\sigma+ikh)]\sin(khx)\Big),$$

where N is the truncation point, and L_g is the characteristic function of g.

- To calculate G_j on the grid, we pick $\sigma_+ \in (0, \sigma_u)$ and $\sigma_- \in (\sigma_l, 0)$. Let

$$G_j = \begin{cases} I_{\sigma_+,x_j}^{N,h}(L_{G_\xi}), \text{ if } x_j \le 0; \\ 1 - I_{\sigma_-,x_j}^{N,h}(L_{\bar{G}_\xi}), \text{if } x_j > 0. \end{cases} \tag{4.11}$$

Let $x_0 = -d/d\xi L_{g_\xi}(0)$ and calculate G_0 by (4.11). Let $x_j = x_0 + j\delta$ and $x_{-j} = x_0 - j\delta$. Compute G_j by (4.11). If $G_j < G_{j-1}$, set $G_j = G_{j-1}$; if $G_{-j} > G_{-(j-1)}$, set $G_{-j} = G_{-(j-1)}$. Continue for $j = 1, 2, ...$ until we find $j_{\max} > 0$ and $j_{\min} < 0$ such that $G_{j_{\max}} \approx 1$ and $G_{j_{\min}} \approx 0$. Then we set $J = \{j_{\min}, j_{\min}+1, ..., j_{\max}-1, j_{\max}\}$. Then $\{x_j, j \in J\}$ are our grids.

- At each simulation step, we approximate g_ξ and $\dot{g}_\xi$ through G_j as:

$$\hat{g}_\xi(x) = \begin{cases} (G_j - G_{j-1})/\delta, if \;\; x \in [x_{j-1}, x_j), j \in J \\ 0, \quad if \;\; x < x_{\min} \; or \; x > x_{\max} \end{cases}$$

$$\hat{\dot{g}}_\xi(x) = \begin{cases} (\dot{G}_j - \dot{G}_{j-1})/\delta, \; if \;\; x \in [x_{j-1}, x_j), j \in J \\ 0, \; if \;\; x < x_{\min} \; or \; x > x_{\max} \end{cases}$$

where $\dot{G}_j \approx \dot{G}_\xi(x_j)$ is calculated through $\dot{G}_\xi = \frac{dG_\xi}{d\xi}$. Then we can estimate the approximated score function $\hat{S}_\xi = \hat{\dot{g}}_\xi / \hat{g}_\xi$ at X.

- We generate $\hat{X}$ from the approximation $\hat{G}_\xi$ by setting $X = \hat{G}_\xi^{-1}(U)$, where $U \sim \text{Unif}(G_{\min}, G_{\max})$ as in

$$\hat{X} = \frac{U\delta + x_{j-1}G_j - x_jG_{j-1}}{G_j - G_{j-1}}.$$

- At the end of each path, the LR estimator of the derivative of $E_\xi[V(X))]$ is $V(\hat{X})\hat{S}_\xi(\hat{X})$.

The two-sided Laplace transform of a function g is given by

$$L_g(u) = \int_{-\infty}^{\infty} e^{-ux} g(x)dx,$$

where u is a complex variable. Assume the region of convergence in the complex plane includes a real interval σ_l, σ_u, where $\sigma_l < 0$ and $\sigma_u > 0$ for the Laplace transform L_{g_ξ} of g_ξ.

We have

$$L_{G_\xi}(u) = \frac{L_{g_\xi}(u)}{u},$$

for $\text{Re}(u) \in (0, \sigma_u)$ and

$$L_{\bar{G}_\xi}(u) = -L_{g_\xi}(u)/u,$$

for $\text{Re}(u) \in (0, \sigma_u)$, where $\bar{G}_\xi = 1 - G_\xi$. For the VG process in this dissertation, the Laplace transform of the density function is

$$L_g(u) = (1 - iu\theta\nu + \frac{1}{2}\sigma^2\nu u^2)^{-t/\nu}.$$

Again, the interchangeability can be guaranteed by the following lemma in Glasserman and Liu [39] or Asmussen and Glynn [4] as follows:

Lemma 4.2.4. *If there is an integrable function H for which $e^{-\sigma x}|g'_\xi(x)| \leq H(x)$ for all real x and for any ξ in a neighborhood of ξ_0, then we have the following result:*

$$\begin{aligned} L_{g'_\xi(u)} &= \int_{-\infty}^{\infty} e^{-ux} \frac{\partial}{\partial \xi} g_\xi(x) dx \\ &= \frac{\partial}{\partial \xi} \int_{-\infty}^{\infty} e^{-ux} g_\xi(x) dx \\ &= \frac{\partial}{\partial \xi} L_{g_\xi}(u) \end{aligned}$$

at $\xi = \xi_0$, where $Re(u) = \sigma \in (\sigma_l, \sigma_u)$.

Chapter 5

Constructing Optimal Portfolios of Variance Swaps by Lévy Mixture and Variance Gamma Correlated

5.1 Overview

A variance swap is a forward contract on annualized variance in which two parties agree to buy or sell the realized variance of an index or a single stock on a fixed future day for a fixed price. This fixed price is called a variance strike or strike price.

This chapter focuses on constructing optimal portfolios of variance swaps. Our approach follows the classical portfolio theory for stock investment proposed in Markowitz [64]. Madan [54] provided an introduction of a variance swap contract and explained the determination of the variance strike price. According to the result, the variance strike equals the spot value of the realised variance at maturity from the fact that the market values of the variance swap contract is zero. Madan [56] proposed the portfolio theory of variance swaps. However, his implementation to handle the residuals of a linear regression model for the realized variance is based on a full rank Gaussian copula (FGC) model, proposed in Malevergne and Sornette [62] and developed in Madan and Khanna [53]. In this chapter, we study the variance swap from a different viewpoint. More precisely, we employ two other

multidimensional dependent non-Gaussian models: the variance Gamma correlated model (VGC), developed in Eberlein and Madan [30], and the Lévy Mixture model (LM), proposed in Madan and Yen [61].

Madan and Khanna [53] analyzed and compared three non-Gaussian dependent models: FGC, VGC and LM, and concluded that (i) all three models are comparable; (ii) generally LM has a superior performance to VGC and FGC except that occasionally VGC and FGC may dominate.

Cherny and Madan [22] introduced an optimization theory of performance evaluation and several distortion functions. Several optimization methods were used in Eberlein and Madan [29] and Madan [55], such as maximizing the index of acceptability and maximizing the expected distortion given a fixed acceptable index. In this chapter, we seek a portfolio of variance swaps using a new performance measure called acceptability indices. We employed some function to distort the distribution to add more weights to losses and discount weights to gains. Then, we maximize the expected value of the cash flow using distorted distribution function given some fixed acceptability indices.

The remainder of this chapter is organized as follows. Section 5.2 presents the definition of variance swaps and gives a brief introduction to calculating the cash flow of the variance swap. Section 5.3 explains how to price the variance strike of a contract. A calibration on the option surface through a Variance Gamma Specific Self Decomposable (VGSSD) model is conducted to estimate the strike price. Section 5.4 investigates the simulation method of the realized variance by a linear regression model. Recognizing the difficulty in keeping the terms of the realized

variance in a regression model always positive, we employ a transform called the Hardy-Littlewood-Gauss transform. Moreover, considering the residuals from the regression model have excess kurtosis and skewness, we employ three multidimensional dependent non-Gaussian models: full rank Gaussian copula (FGC), Lévy Mixture (LM) and Variance Gamma Correlated (VGC) to handle cross-dependencies between residuals of different assets. Section 5.5 explains the construction of portfolios of variance swaps. For this purpose, we optimize by maximizing the expected distortion given some fixed indices of acceptability. Section 5.6 shows the numerical results during the implementation. Several results of portfolios with different acceptable indices and different non-Gaussian models are also provided.

5.2 Definition of Variance Swaps

A variance swap is an over-the-counter financial derivative that allows one to speculate on or hedge risks associated with the magnitude of movement, that is, the volatility of some underlying product, like an exchange rate or interest rate. Through a variance swap, investors could achieve long or short exposure to market volatility. It is not really a swap in the traditional sense; it is in fact a kind of forward contract signed by two parties who agree to exchange cash flows based on the measured variance of a specified underlying asset during a certain time period. More precisely, on the trading day specified by the contract, the two parties trade the variance swap according to the variance strike, the realized variance and the notional amount. This contract allows them to gain exposure to changes in the

variance of the underlying index, so that they can trade the variance swap to hedge off exposure from other areas of their business or to profit from anticipated changes in the variance of an asset.

The features of a variance swap include the variance strike, which is also called the fixed leg, the realized variance, that is, the floating leg and the notional amount. The floating leg of the swap is the amount paid based on the realized variance of the price changes of the underlying product. The fixed leg of the swap is the fixed amount, which is also the strike quoted at the deal's inception paid on the maturity.

Therefore, the payoff of a variance swap at expiration or maturity is

$$\begin{aligned} &\text{Payoff of a Variance Swap} \\ &= \text{Notional Amount} \times (\text{Realized Variance} - \text{Variance Strike}). \end{aligned}$$

Let P be the notional principal, let σ_r^2 be the realized variance, and let σ_k^2 be the variance strike. The payoff of a variance swap can be expressed as:

$$P \times (\sigma_r^2 - \sigma_k^2).$$

Due to the fact that each year has 252 trading days, the realized variance σ_r^2 can be written (Madan [54]) as

$$\sigma_r^2 = \frac{252}{T} \sum_{t=1}^{T} x_t^2,$$

and the strike price can be written as

$$\sigma_k^2 = k^2,$$

where $x_t = \log(S_t/S_{t-1})$, S_t is the stock price of the underlying asset at the end of day t, and k is the annualized volatility quotation. Hence, the payoff of a variance

swap at the end of day T is:

$$P \times \left(\frac{252}{T} \sum_{t=1}^{T} \left(\log(\frac{S_t}{S_{t-1}}) \right)^2 - k^2 \right). \tag{5.1}$$

5.3 Fixed Leg (Variance Strike)

5.3.1 Calculation of the Variance Strike

Let us start from calculating the strike price of a variance swap. We select 10 assets, of which the tickers whose ticker symbols are: xom, aapl, mmm, c, adbe, amzn, gs, coh, goog, bac on the S&P 500 index. The corresponding companies to the tickers are shown in Table 5.1. On the portfolio construction date, we need to calculate the strike price for each asset. From Appendix A, the fixed swap payment is

$$k = \sqrt{\frac{252 e^{rT/365}}{T}} \mathrm{R},$$

where

$$\mathrm{R} = \mathrm{MVO}. \tag{5.2}$$

by setting the interest rate $r = 0$ and the dividend rate $q = 0$ (proposed by Madan [56]).

The market value of options (MVO) can be calculated from an option surface calibration. We use a model called Variance Gamma Scaled Self-Decomposable (VGSSD) to calibrate the option surfaces. Let's define the γ-self-similar process first.

Definition 5 (γ-self-similar process)**.** *A stochastic process* $(Y(t), t \leq 0)$ *is called a*

Table 5.1: Tickers and Companies

ticker	Company
xom	Exxon Mobil Corp.
aapl	Apple Inc.
mmm	3M Co.
c	Citigroup Inc.
adbe	Adobe Systems Inc.
amzn	Amazon Com Inc.
gs	Goldman Sachs Groups Inc.
coh	Coach Inc.
goog	Google Inc.
bac	Bank of America Corp.

γ-self-similar process, if it satisfies that for any $\lambda > 0$, and all time t, such that

$$Y(\lambda t) \stackrel{law}{=} \lambda^{\gamma} Y(t).$$

The additive process is called as an additive process with inhomogeneous and independent increments. In particular, when the increments are time homogeneous, the process is called a Lévy process. A law is self-decomposable if and only if it is the the law at unit time of an additive process, which is also a self-similar process. Consequently, such processes are called as Sato processes in Sato [74]. If $Y(t)$ is the value at time t of a self-similar additive process with paths of bounded variation. If the law of the self-similar additive process at unit time be the self-decomposable

law of the random variable X:

$$Y(1) \stackrel{\text{law}}{=} X.$$

By the scaling property, we let the law of $Y(t)$ equals the law of $t^\gamma X(1)$, i.e.,

$$Y(t) \stackrel{\text{law}}{=} t^\gamma X(1),$$

where $X(1)$ is a VG process at time $t = 1$. Such $Y(t)$ is said to follow a VGSSD process.

If $Y(t)$ follows a (VGSSD) model, Carr, Geman, Madan and Yor [18] showed that the law of $Y(t)$ is equivalent to the law of $t^\gamma X(1)$ by the scaling property. Thus, the characteristic function of $Y(t)$ is:

$$\begin{aligned}
\phi_{Y(t)}(u) &= E[e^{iuY(t)}] \\
&= E[e^{iut^\gamma X(1)}] \\
&= (1 - iu\theta\nu t^\gamma + \frac{1}{2}\sigma^2\nu u^2 t^{2\gamma})^{-\frac{1}{\gamma}}.
\end{aligned}$$

The stock price $S(t)$ under the risk-neutral measure can be expressed as:

$$S(t) = S(0)e^{rt}\frac{e^{Y(t)}}{E[e^{Y(t)}]}. \tag{5.3}$$

The price of a variance swap contract $\sigma_k^2 = k^2$ can be calculated from

$$\sigma_k^2 t = 2i\frac{\partial\phi_M(u,t)}{\partial u}\bigg|_{u=0}, \tag{5.4}$$

where

$$\phi_M(u,t) = E^{\mathrm{Q}}[\exp(iu\ \ln(M(t)))].$$

Let r be the risk-free interest rate and q be the dividend rate. We can express the stock price in the risk-neutral measure by

$$S(t) = S(0)e^{(r-q)t}M(t).$$

This implies that the characteristic function of $M(t)$ is related to that of $S(t)$, which can be calculated from Equation (5.3). Consequently,

$$\begin{aligned}\sigma_k^2 t &= -2\mathrm{E}[\ln \mathrm{M}] \\ &= -2\mathrm{E}[\ln S_t - \ln S_0 - (r-q)t] \\ &= -2qt - 2\theta t^{\gamma} - \frac{2}{\nu}\ln(1 - \theta\nu t^{\gamma} - \frac{1}{2}\sigma^2\nu t^{2\gamma}),\end{aligned}$$

where $r = 0$, $q = 0$ to make Equation (5.5) equivalent to the variance strike as proposed in Madan [54]. Therefore, the variance strike is

$$\sigma_k^2 = (-2\theta t^{\gamma} - \frac{2}{\nu}\ln(1 - \theta\nu t^{\gamma} - \frac{1}{2}\sigma^2\nu t^{2\gamma}))t^{-1}.$$

Using the data of option prices of the ten assets mentioned above on October 19 2007, from Wharton Research Data Services (WRDS), we employ the VGSSD model to calibrate the option surface for each asset. More precisely, by choosing difference parameters, we employ the VGSSD model to estimate the corresponding option prices for different strike prices and maturities, which can be considered the model price. Then we can estimate the corresponding parameters that minimize the root mean square error (RMSE) of the model prices to the market prices. The RMSE is defined as

$$\mathrm{RMSE} = \sqrt{\sum_{\text{options}} \frac{\big(\text{market price} - \text{model price}\big)^2}{\text{number of options}}}.$$

To measure the overall quality fit, we use model price calculated from the estimated parameters to calculate the APE, which is the average absolute error as a percentage of the mean price. The APE is defined as

$$\mathrm{APE} = \frac{1}{\text{mean of option price}} \cdot \sum_{\text{options}} \frac{|\text{market price} - \text{model price}|}{\text{number of options}}.$$

Table 5.2: VGSSD Parameters on 20071019

ticker	σ	ν	θ	γ	RMSE	APE
xom	0.2511	0.3708	-0.2480	0.4961	0.0511	0.0320
aapl	0.4207	0.0648	-0.3021	0.3912	0.2460	0.0413
mmm	0.2013	0.1177	-0.4167	0.4685	0.0518	0.0395
c	0.2978	0.3147	-0.3405	0.4427	0.0348	0.0432
adbe	0.3182	0.2501	-0.2370	0.5628	0.0277	0.0289
amzn	0.4711	0.1265	-0.6917	0.4135	0.1993	0.0492
gs	0.3236	0.4495	-0.3659	0.4406	0.1526	0.0241
coh	0.3566	0.1027	-0.7032	0.4132	0.0658	0.0372
goog	0.3006	0.3513	-0.1737	0.4954	0.3885	0.0341
bac	0.2540	0.5081	-0.2585	0.5405	0.0425	0.0419

Table 5.3: One Month Variance Strike on 20071019

ticker	xom	aapl	mmm	c	adbe	amzn	gs	coh	goog	bac
	0.26305	0.53385	0.22557	0.36647	0.27097	0.59576	0.37237	0.48358	0.33763	0.24381

The calibrated parameters are presented in Table 5.2, and the corresponding calibrated option surfaces are depicted in Figure 5.1−5.10, respectively. The quotations of prices of the variance strike are presented in Table 5.3.

Figure 5.1: Graph of Fitted Option Surface of XOM on 20071019

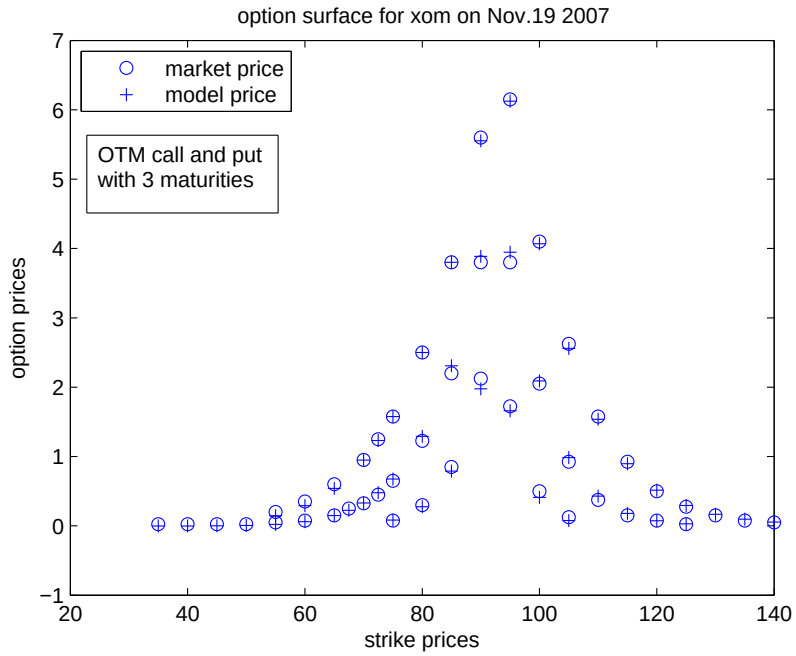


Figure 5.2: Graph of Fitted Option Surface of AAPL on 20071019

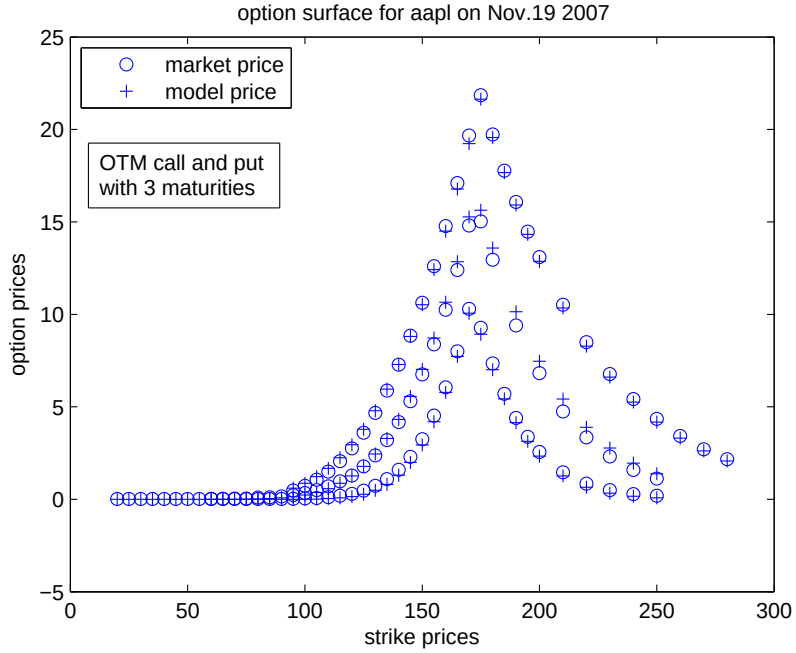


Figure 5.3: Graph of Fitted Option Surface of MMM on 20071019

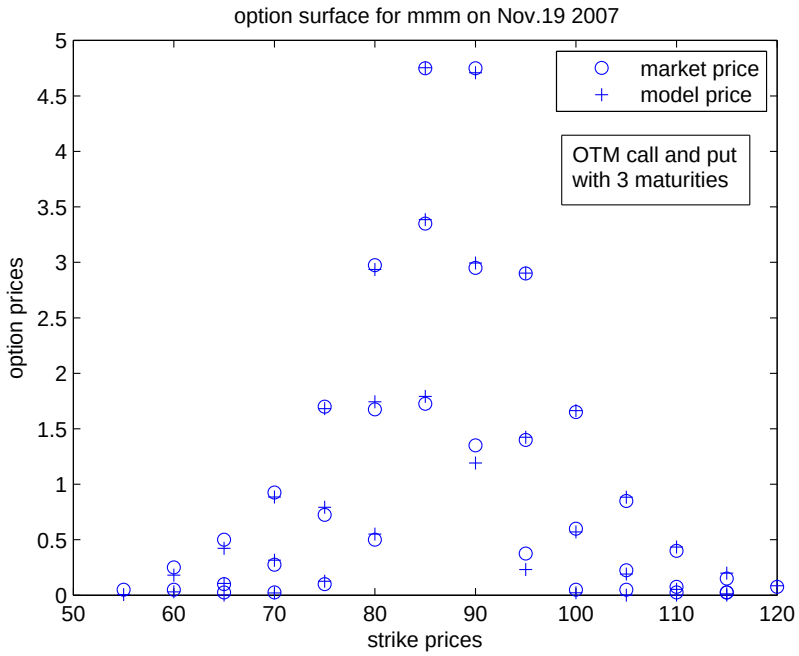


Figure 5.4: Graph of Fitted Option Surface of C on 20071019

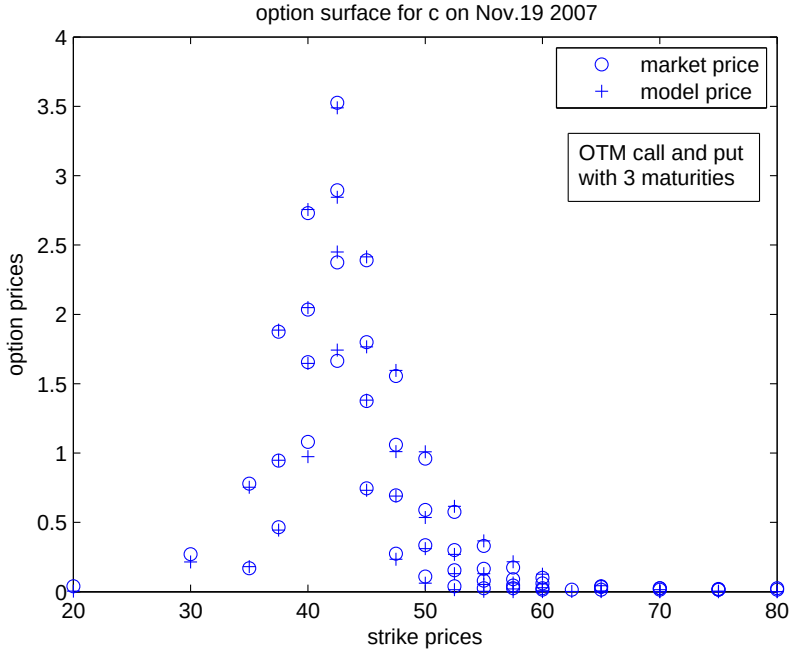


Figure 5.5: Graph of Fitted Option Surface of ADBE on 20071019

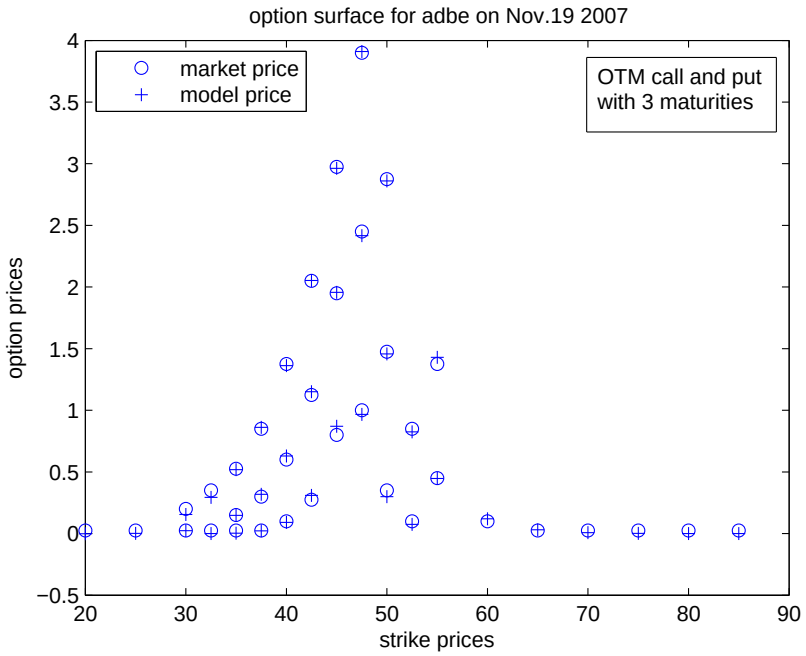


Figure 5.6: Graph of Fitted Option Surface of AMZN on 20071019

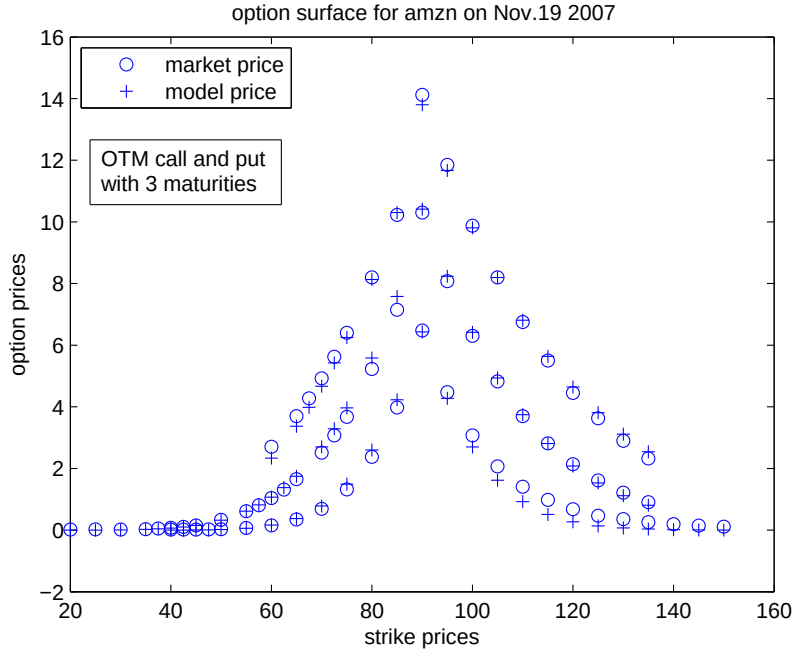


Figure 5.7: Graph of Fitted Option Surface of GS on 20071019

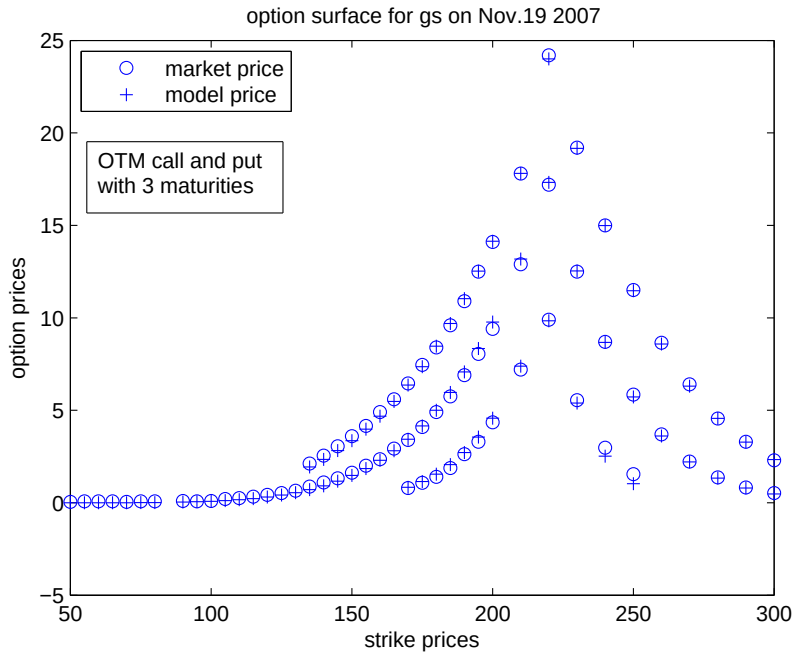


Figure 5.8: Graph of Fitted Option Surface of COH on 20071019

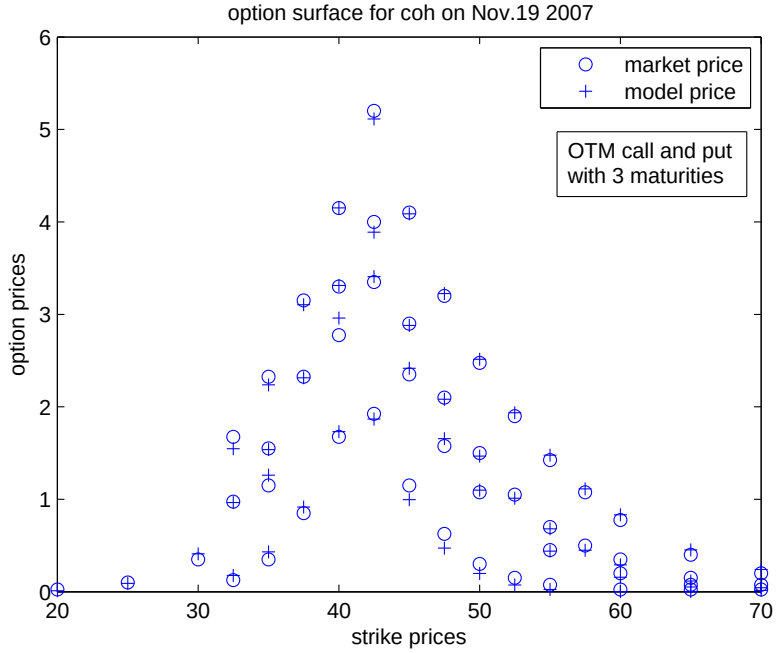


Figure 5.9: Graph of Fitted Option Surface of GOOG on 20071019

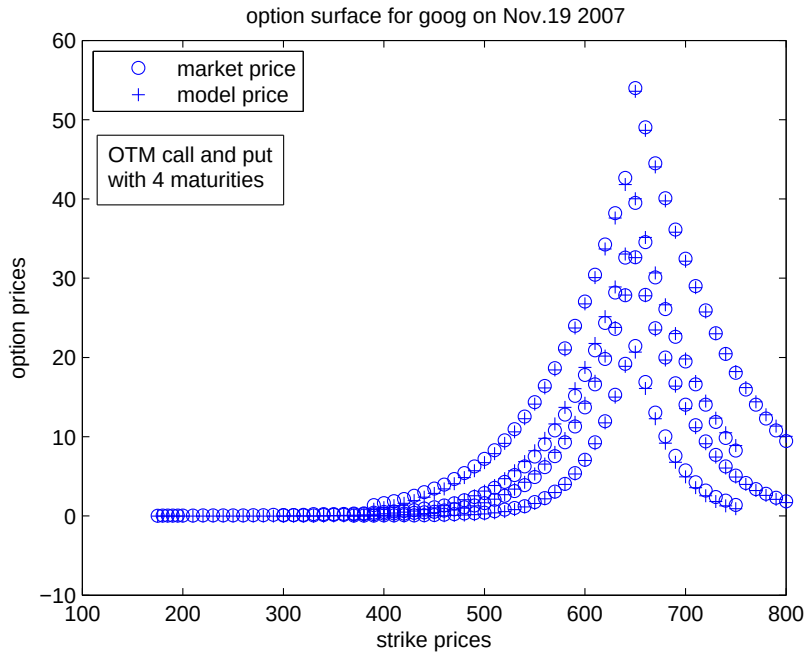
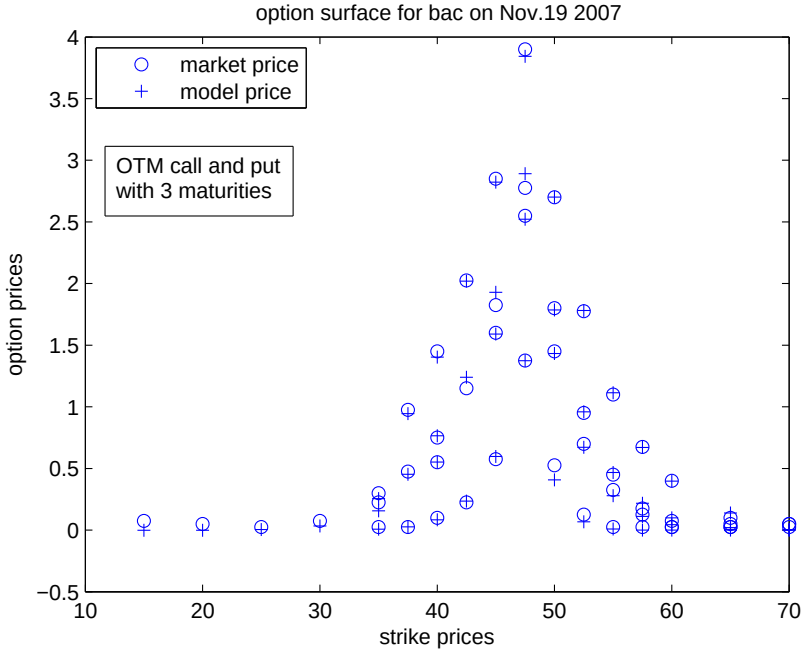


Figure 5.10: Graph of Fitted Option Surface of BAC on 20071019



5.4 Floating Leg (Realized Variance)

We consider the one month variance swap based on the 10 assets whose ticker symbols are: xom, aapl, mmm, c, adbe, amzn, gs, coh, goog, bac on the S&P500 index matured on November 18, 2007.

Let $S_{i,t}$ denote the stock price of asset i at market close on day t for $i = 1, ..., 10$. The daily realized variance for asset i on day t defined as $v_{i,t}$ is:

$$v_{i,t} = \left[\ln(\frac{S_{i,t}}{S_{i,t-1}})\right]^2. \tag{5.5}$$

From equation (5.1), the cash flow of a variance swap on day $t_0 + T$ entered on day t_0 is calculated as:

$$P \times \left(\frac{252}{T} \cdot \sum_{t=t_0+1}^{t_0+T} v_{i,t} - k_{i,t}^2\right). \tag{5.6}$$

In this dissertation, we consider the one month variance swap. Thus $T = 21$ and $k_{i,t}$ is annualized volatility for asset i on day t.

Taking the high autocorrelation of $v_{i,t}$ into account, we fit a linear autoregressive model to $v_{i,t}$, the daily squared log returns calculated from the time series data from stock prices of the 10 assets. It is obvious that the squared log returns $v_{i,t}$ are positive; However, it is difficult to keep the predicted $v_{i,t}$ always positive in the regression model. Hence, we seek a transform to map $v_{i,t}$ from positive numbers to all real numbers such that the regression model can be applied. Due to this, one may consider the log transform, i.e. taking log of $v_{i,t}$. Obviously, $\log(v_{i,t})$ could be positive or negative.

However, this would cause a double exponential on the returns and poor returns on residuals. Madan [56] employed a transform called the Hardy-Littlewood

transform behaving almost like a linear function, which would not cause any unpredictable problems. In the following sections, we introduce this transform and the linear regression model.

5.4.1 Hardy-Littlewood Transform and Linear Regression Model

Recognizing squared log returns are highly autocorrelated and subject to some levels of clustering, we follow Madan [56] and apply the Hardy-Littlewood transform to the squared daily log returns. This transform can deal with highly correlated auto-correlations, as well as transform $v_{i,t}$'s from positive values to all real values required by the linear regression model because of the difficulty of keeping $v_{i,t}$ in the linear regression model positive in the future simulations. We considered the log transform of the $v_{i,t}$, but this would make a double exponent and result in bad effects.

Definition 6 (**Hardy-Littlewood Transform**). *Let $f(x)$ be a symmetric density function on the real line having finite expectation of absolute value of x. The Hardy-Littlewood transform is defined as:*

$$g(x) = \frac{\int_x^\infty u f(u) du}{\int_x^\infty f(u) du}. \tag{5.7}$$

As $x \to -\infty$, $g(x)$ would be close to 0; when x is large enough, $g(x)$ would behave like x, which indicates this transform is close to "linear". Since $g(x)$ is always positive, the inverse of $g(x)$ will transform the positive squared log returns $v_{i,t}$ to all real values. In this chapter, we set the density $f(x)$ to be a standard normal density function, i.e., $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, and get the so-called Hardy-Littlewood-

Gauss transform (also known as the Mills ratio):

$$\begin{aligned} g_{\mathrm{HLG}}(x) &= \frac{\int_x^\infty u/\sqrt{2\pi}\exp(-\frac{u^2}{2})du}{\int_x^\infty 1/\sqrt{2\pi}\exp(-\frac{u^2}{2})du} \\ &= \frac{f_{\mathrm{norm}}(x)}{1-F_{\mathrm{norm}}(x)} \\ &= \frac{1/\sqrt{2\pi}\exp(-\frac{u^2}{2})}{\int_x^\infty 1/\sqrt{2\pi}\exp(-\frac{u^2}{2})du}, \end{aligned} \tag{5.8}$$

where $f_{\mathrm{norm}}(x)$ is the density function and $F_{\mathrm{norm}}(x)$ is the cumulative distribution function of a standard normal distributed variable. Thus, the definition of the Hardy-Littlewood-Gauss transform is shown as follows.

Definition 7 (**Hardy-Littlewood-Gauss Transform**). *Let x be a real value. The transform defined as*

$$g_{HLG}(x) = \frac{1/\sqrt{2\pi}\exp(-\frac{u^2}{2})}{\int_x^\infty 1/\sqrt{2\pi}\exp(-\frac{u^2}{2})du}$$

is called the Hardy-Littlewood-Gauss transform.

By taking the six-month historical data of stock prices from the 10 assets from Mar.18, 2007 to Nov.18, 2007 into equation (5.5), we get the squared log-returns $v_{i,t}$. Then we employ the inverse of Hardy-Littlewood-Gauss transform defined above to map $v_{i,t}$'s to all real values $x_{i,t}$'s, i.e.,

$$x_{i,t} = g_{\mathrm{HLG}}^{-1}(v_{i,t}),$$

where $x_{i,t}$'s are newly generated data which are all real numbers.

To handle the highly correlated auto-correlations of the series of data $x_{i,t}$ into account, we fit a 5-lag linear auto-regressive model to data $x_{i,t}$, that is,

$$x_{i,t} = a_i + \sum_{j=1}^{5} b_{i,j}x_{i,t-j} + u_{i,t}, \tag{5.9}$$

Table 5.4: Linear Regression Results

ticker	Constant	Lag1	Lag2	Lag3	lag4	Lag5
xom	-3.0442	0.1229	-0.0228	0.1332	0.0342	0.0120
aapl	-2.7226	0.0780	-0.0223	0.0903	0.1602	0.0150
mmm	-4.6246	-0.0619	0.1056	-0.0015	-0.0569	-0.0326
c	-2.2779	-0.0688	0.0501	0.1408	0.1890	0.1441
adbe	-4.4866	-0.1239	-0.0198	0.0488	0.0286	-0.0014
amzn	-2.9798	0.1353	-0.0407	0.1019	0.1125	-0.0511
gs	-2.6134	-0.0168	0.0634	0.2381	0.1593	-0.0886
coh	-2.8390	0.0439	0.0903	0.0030	0.2077	-0.0610
goog	-3.7481	0.0254	-0.0263	0.0812	-0.0984	0.1390
bac	-2.8081	0.1144	0.1606	0.0322	0.0054	0.04023

where a_i's are constant terms, $b_{i,1}, b_{i,2}, b_{i,3}, b_{i,4}, b_{i,5}$ are the corresponding five coefficients in the regression model, and $u_{i,t}$ are residuals. The regression results of the five lags and the constant terms by taking $x_{i,t}$ into the linear regression model are provided in Table 5.4.

However, it should be pointed out that time series data of residuals $u_{i,t}$ calculated from the regression model in (5.9) do not follow a Gaussian distribution. Instead, they have skewness and excess kurtosis. Hence, we can not ignore this and seek some multidimensional dependent non-Gaussian models to study them. In the following sections, we introduce the LM model and VGC model and explain the implementation of them to investigate the residuals.

5.4.2 Lévy Mixture

Considering that the residuals $u_{i,t}$ have heavy tail and skewness, we cannot ignore the residuals $u_{i,t}$ in the linear regression. We have to take the residuals into account and apply a multidimensional non-Gaussian model called the Lévy Mixture introduced in Section 3.3.

The Lévy Mixture model is conducted by transforming correlated multidimensional non-Gaussian variates to independent non-Gaussian variates, so that the parameters of each independent non-Gaussian variate can be estimated by maximum likelihood estimation (MLE).

To estimate the realized variance, we suggest the estimation procedures as follows:

- Take historical data of stock prices of the 10 assets from Mar.18, 2007 to Nov.18, 2007 to calculate the series of real data of squared log-return $v_{i,t}$ through equation (5.5).

- Take the inverse of Hardy-Littlewood-Gauss transform to the historical data of $v_{i,t}$

  $$x_{i,t} = g^{-1}(v_{i,t})$$

  to get series of newly generated data $x_{i,t}$.

- Considering highly correlated autocorrelations of the series of data $x_{i,t}$, set a linear regression to the newly generated time series data $x_{i,t}$:

  $$x_{i,t} = a_i + \sum_{j=1}^{5} b_{i,j} x_{i,t-j} + u_{i,t}.$$

- Applying ICA to the time series data $U = \{u_{i,t}, i = 1, ..., N; t = 1, ..., \bar{T}\}$, we get the mixture matrix A and independent variates Y, which are assumed to follow an independent Variance Gamma process.

- Applying MLE to the N-dimensional data Y, of which all assets are independent and follow variance Gamma distribution, we get the corresponding parameters $\sigma_i, \nu_i, \theta_i$.

- Employ the estimated parameters $\sigma_i, \nu_i, \theta_i$ to simulate the independent variables $\hat{Y}_i$ following a variance Gamma distribution.

- Through $\hat{U} = A\hat{Y}$, we can get the simulated dependent residual series of data $\hat{U}$.

The results of the lags of the linear regression model are displayed in Table 5.4. The estimated parameters of the independent VG variates Y are shown in Table 5.5, and the mixing matrix A is presented in Table 5.6.

5.4.3 Variance Gamma Correlated Model

Another model we will employ to deal with the residual data $u_{i,t}$ is Variance Gamma Correlated (VGC) introduced in section (3.4). We write the residual data $u_{i,t}$ as multidimensional correlated non-Gaussian distributed $U_i(t), i = 1, ..., N$.

We use the series of $u_i(t)$ derived from the historical data and estimate the parameters $\sigma_i, \nu_i, \theta_i$ and calculate the covariance from

$$\text{Cov}(U_i, U_j) = \sigma_i \sigma_j E(\sqrt{g_i}) E(\sqrt{g_j}) \rho_{i,j} \tag{5.10}$$

Table 5.5: VG estimates for independent variates Y for LM

ticker	σ	ν	θ
xom	0.9761	0.0745	-0.1299
aapl	0.9456	0.7436	0.0908
mmm	0.9903	0.3091	0.2364
c	0.9033	0.0561	-0.0240
adbe	0.9774	0.5718	-0.2177
amzn	0.9972	0.3658	0.0417
gs	0.8817	0.0077	-0.1546
coh	1.0024	0.5754	-0.0295
goog	0.9871	0.2984	0.3349
bac	1.0282	0.00000029	0.0657

shown in section (3.4). Then we could simulate the multidimensional non-Gaussian process $U_i(t)$ which are correlated with each other. The estimating procedure for applying the VGC model to the residual variates is summarized as following:

- Apply MLE to the time series data $U = u_{i,t}, i = 1, ..., N$, in each dimension separately; each would follow the VG distribution with the corresponding parameters $\sigma_i, \nu_i, \theta_i$.

- Apply the calculated covariance of $u_{i,t}$ to equation (5.10) to get the correlation $\rho_{i,j}$ of the standard normal variable Z_i.

- Simulate the N-dimensional correlated standard normal variable $\hat{Z}$ with the correlation $\rho_{i,j}$ between different assets.

Table 5.6: The Mixing Matrix A for LM

-0.0203	0.0759	-0.3521	-0.2047	0.0121	-0.2031	0.1630	0.0586	-0.0717	0.1436
0.2988	0.1018	-0.1435	0.1439	0.3072	-0.0333	-0.0078	-0.2397	-0.2112	-0.0606
-0.1122	-0.1979	0.0957	-0.0214	0.2095	-0.2476	0.1331	-0.0417	-0.1386	-0.0355
0.2398	0.0222	0.0901	-0.0280	-0.0455	-0.3031	0.0501	-0.04865	0.1470	0.1090
0.2167	-0.0748	0.0530	-0.0004	0.1353	0.0742	0.1883	0.2467	-0.1074	0.1763
0.0445	0.4342	0.0486	0.0201	0.2601	-0.2437	0.0261	0.1074	0.0234	-0.1286
0.1310	0.1026	0.0848	-0.0490	-0.2381	-0.2462	-0.0085	0.0087	-0.3350	0.03909
0.0773	0.0103	-0.0279	0.1338	-0.1566	-0.0899	0.3803	0.0275	0.0222	-0.2656
0.2302	-0.0849	0.0135	-0.2598	0.0498	-0.1008	-0.0264	0.1243	-0.0246	-0.2664
0.1226	0.0968	0.1723	-0.2785	0.0290	-0.0916	0.2261	-0.1695	-0.0344	0.1009

- By the estimated parameters and the newly simulated Z_i, and plugging back into

$$U_i(t) = \theta_i(G_i(t) - t) + \sigma_i W_i(G_i(t)) \tag{5.11}$$

shown in section (3.4). we can get the newly simulated series data $\hat{U}$.

The estimated parameters of the independent VG variates Y are in Table 5.7, and the covariance matrix of the standard normal variable Z are in Table 5.8.

5.4.4 Simulations

Following the procedures mentioned above, we simulate 10000 sample paths. The annualized unit realized variance of the asset i on day t on sample path s is

$$\sigma^2_{i,t,s} = \frac{252}{21} \sum_{j=t+1}^{t+21} v_{i,j,s}. \tag{5.12}$$

It should be pointed out that we are using calendar of trading days, which means that each year has 252 days and each month has 21 days.

Table 5.7: VG estimates for variates Y for VGC

ticker	σ	ν	θ
xom	0.3462	0.1181	-1.0660
aapl	0.4027	0.2765	-0.8307
mmm	0.3991	0.1179	-0.6366
c	0.5070	0.3379	-0.1823
adbe	0.1812	0.1576	-1.0970
amzn	0.5842	0.4243	-0.1553
gs	0.2868	0.1080	-1.3080
coh	0.3933	0.4173	-0.5764
goog	0.4962	0.0540	-0.4227
bac	0.2112	0.0883	-1.4846

The simulated cash flow to asset i on the variance swap on path s is then obtained by

$$c_{i,s} = \sigma^2_{i,t,s} - k^2_{i,t,s}. \tag{5.13}$$

5.5 Optimization

Having the simulated cash flow to asset i on the variance swap, we then construct the optimal portfolios of $\omega = (\omega_i, i = 1, ..., N)$. The magnitude ω_i for asset i is a dollar notional amount that could be positive or negative depending on whether the swap is purchased or sold. More precisely, ω_i is positive, if the variance swap is purchased; ω_i is negative, if the variance swap is sold. With portfolio weights ω, we

Table 5.8: The Covariance Matrix for Standard Normal Variates Z

2.3313	0.1848	0.2617	0.2570	0.5993	0.2871	0.6386	0.1729	0.1490	0.9059
0.1848	2.2776	0.2799	0.1420	0.5811	0.5530	0.2907	0.0289	0.1640	0.4178
0.2617	0.2799	1.3459	0.1469	0.4177	0.1317	0.2606	0.1992	0.1948	0.6956
0.2570	0.1420	0.1469	0.8496	0.3184	0.2681	0.6036	0.2162	0.2105	1.0608
0.5993	0.5811	0.4177	0.3184	6.9588	0.0314	0.3632	0.2397	0.4339	1.3021
0.2871	0.5530	0.1317	0.2681	0.0314	1.1467	0.2536	0.1850	0.1943	0.4705
0.6386	0.2907	0.2606	0.6036	0.3632	0.2536	3.3790	0.3895	0.3428	1.3985
0.1729	0.0289	0.1992	0.2162	0.2397	0.1850	0.3895	1.9780	0.2554	0.3340
0.1490	0.1640	0.1948	0.2105	0.4339	0.1943	0.3428	0.2554	0.9470	0.5116
0.9059	0.4178	0.6956	1.0608	1.3021	0.4705	1.3985	0.3340	0.5116	5.3622

can write the corresponding portfolio of cash flow on sample path s as

$$c_s(\omega) = \sum_{i=1}^{N} \omega_i c_{i,s}.$$

5.5.1 A New Performance Measure

Modern portfolio theory proposed by Markowitz [64, 65] provided selection principles to maximize the portfolio's expected return while fixing the variance, or minimize the variance while fixing the expected return. These two pcriteria define the efficient frontiers. Thus, an investor's preference is a trade-off between risk (represented by the standard deviation of return) and gain (the expected return). Diversification is another important concept. Due to the correlations between assets,

a suitable portfolio's variance is smaller than the sum of the variances of all the assets. Consequently, the risk can be reduced with a diversified portfolio in different assets.

Several optimization strategies can be employed to estimate optimal portfolios. If we employ the optimal portfolio strategy which maximizes the Sharpe ratio (the ratio of mean of return over standard deviation of mean), the selection in the non-Gaussian context here will diverge from the multivariate Gaussian model. This is due to the recognition that investors are not indifferent to other aspects of a return distribution. Moreover, the Sharpe ratio measure does not actually respect risks; for example, for positive cash flows with finite means but infinite variance, the Sharpe ratio is zero while there is an arbitrage. Another popular portfolio strategy called Gain-Loss Ratio has the limitation that small losses and large losses are considered equally with the same weights. Hence, we choose a new performance called acceptability indices which take risks into consideration and are suitable for the non-Gaussian context. More precisely, we maximize the expected distortion which is considered as a utility function given some fixed indices of acceptability.

We employ the arbitrage consistent performance measures developed in Madan [55], generalized from the Sharpe ratio. These measures directly measure the quality of cash flow distributions whose cost is zero. The zero-cost cash flow is a random variable X. First, we define the convex set containing all non-negative cash flows. For an investor who prefers to maximize the expected utility, the zero-cost random variables X's in the convex set satisfy the condition: for any random variable X,

there exists

$$\mathrm{E}[u(W+X)] \geq \mathrm{E}[u(W)],$$

given a random initial position W and the utility function u. We can model the acceptable cash flows by the smallest convex cone containing all the convex sets satisfying the definition above. Such cones of acceptable cash flows are supported by a set of probability measures and these cash flows have the positive expectation under all supporting probability measures. These measures are connected with the acceptability indices. The measure of performance means a map α from L^{∞} to the extended positive half-line $[0, \infty]$. Let $X \in L^{\infty}$ be the terminal cash flow from a trading strategy, and $\alpha(X)$ be the performance or quality of X. Let's first define the set A_x containing all trades acceptable at level x as those with performance above x:

$$A_x = \{X : \alpha(X) \leq x\}, \quad x \in R_+.$$

The following four properties must be satisfied for a performance measure $\alpha(x)$ to define an acceptability index.

1. **Quasi-Concavity:** If $\alpha(X) \geq x$ and $\alpha(Y) \geq x$, then $\alpha(\lambda X + (1-\lambda)Y) \geq x$ for any $\lambda \in [0,1]$.

2. **Monotonicity:** If $X \leq Y$ almost surely, then $\alpha(X) \leq \alpha(Y)$.

3. **Scale Invariance:** The level of acceptability of X does not change under scaling, that is, $\alpha(\lambda X) = \alpha(X)$ for $\lambda > 0$.

4. **Fatou Property:** This property requires that if (X_n) is a sequence of random

variables such that $|X_n| \leq 1$, $\alpha(X_n) \geq x$, and X_n converges to X in probability, then $\alpha(X) \geq x$.

It is recognized that the higher the index is, the smaller the set of acceptable cash flow distributions at this level is. The law of invariant cones of acceptability for the purpose of constructing operational cones of acceptability is developed in Cherny and Madan [22]. These invariant cones are related to a series of concave distortion functions $\Phi^\gamma(y)$, which distort the distribution function by adding more weights to losses and discounting weights to gains. $y = F(x)$ and $F(x)$ is the cumulative distribution function of the cash flow X. The decision of acceptability of the cash flow depends on the the distribution function for each concave distortion function.

The concave distortion function $\Phi^\gamma(y)$ is defined on the unit interval with values in the unit interval which is point wise increasing in the level of the distortion γ. A random variable X with distribution function $F(x)$ is accepted at level γ if

$$\int_{-\infty}^{\infty} x d\Phi^\gamma(F(x)) \geq 0, \tag{5.14}$$

which means that the expected value of the cash flow under the distortion function Φ^γ is nonnegative. The index of acceptability of cash flow X is the highest level of index that the distorted expectation of the cash flow is nonnegative, i.e.,

$$\gamma^*(X) = \sup\{\gamma \Big| \int_{-\infty}^{\infty} x d\Phi^\gamma(F(x)) \geq 0\}.$$

Cherny and Madan [22] considered four kinds of distortion functions: MAXVAR, MINMAXVAR and MAXMINVAR, based on four different types of stressed sampling. They are defined as follows:

- The distortion function Φ^γ of **MINVAR** at level γ is

$$\Phi^\gamma(u) = 1 - (1-u)^{1+\gamma},$$

for which one constructs a stressed sample on forming the expectation of the minimum of $(1+\gamma)$ independent draws from the cash flow distribution.

- The distortion function Φ^γ of **MAXVAR** at level γ is

$$\Phi^\gamma(u) = u^{\frac{1}{1+\gamma}},$$

in which one constructs a distribution from which one draws numerous times and takes the maximum to get the cash flow distribution being evaluated.

- The distortion function Φ^γ of **MINMAXVAR** at level γ is

$$\Phi^\gamma(u) = 1 - (1-u^{\frac{1}{1+\gamma}})^{1+\gamma}, 0 \leq u \leq 1.$$

- The distortion function Φ^γ of **MAXMINVAR** at level γ is

$$\Phi^\gamma(u) = 1 - (1-u^{1+\gamma})^{\frac{1}{1+\gamma}}, 0 \leq u \leq 1.$$

The distortion functions MAXMINVAR and MINMAXVAR combine the properties of MINVAR and MAXVAR. A notable feature for MAXMINVAR and MINVAR is that these two distortions can reweight large losses to infinitely large and reweight large gains to zero.

In this chapter, the portfolios ω_i are constructed so that the distorted expectation in equation (5.14) is maximized given some acceptable index γ. A random variable X with distribution function $F(x)$ is accepted at level γ if

$$\int_{-\infty}^{\infty} x d\phi^\gamma(F_X(x)) \geq 0,$$

which means that the expected value of the cash flow under the distortion Φ^γ is nonnegative. We optimize the portfolios by maximizing the distorted expectation in equation (5.14), given some acceptable index γ. Given the portfolio of cash flows C, the distorted expectation would be

$$\int_{\infty}^{\infty} cd\phi^\gamma(F(c)), \tag{5.15}$$

where F_C is the cumulative distribution function of C. The computation of distorted expectation is facilitated in terms of an ordered sample from the relevant distribution with $c_{(1)} < c_{(2)} < ... < c_{(N)}$ as:

$$\sum_{i=1}^{N} c_{(i)}(\phi(\frac{i}{N}) - \phi(\frac{i-1}{N})).$$

During the optimization, we restrict the portfolios to the unit sphere. That is they satisfy the condition that

$$\sum_{i=1}^{51} \omega_i^2 = 1.$$

Moreover, the aggregated portfolio is zero dollar:

$$\sum_{i=1}^{51} k_{i,t}^2 \omega_i = 0.$$

In addition, we have to impose a zero Vega constraint as

$$\sum_{i=1}^{51} k_{i,t} \omega_i = 0.$$

Therefore, we can apply the restrictions to construct the optimization.

5.6 Numerical Experiments and Conclusions

We conduct several numerical experiments and employ two different kinds of objective functions to construct optimization under the LM and VGC models. We

design the portfolio to maximize the expected distortion for various acceptable index values, namely, $\gamma = 0.3$, $\gamma = 0.6$, and $\gamma = 0.8$. The results of portfolios of variance swaps by maximizing the MinMaxVar expected distortion through LM model are presented in Table 5.9; the results for the MaxMinVar expected distortion through LM model are displayed in Table 5.10. The results of portfolios of variance swaps by maximizing the MinMaxVar expected distortion through VGC model are presented in Table 5.11; the results for the MaxMinVar expected distortion through VGC model are displayed in Table 5.12. The corresponding realized results are shown in Table 5.14 and Table 5.13.

Through the same methodology presented in this chapter, we construct the portfolios of variance swaps by maximizing the expected distortion for $\gamma = 0.6$ each month in 2007, and the realized results are shown in Table 5.15.

From Table 5.14 and Table 5.13, we find that when the acceptable index is smaller, the realized results, i.e. the profits from holding the one-month variance swap would be bigger. We also notice that for the same acceptable index, the profits from the variance swap by MaxMinVar distortion function are always bigger than the profits by MinMaxVar distortion function. From Table 5.15, we find that the profits using the VGC model exceed the profits using the LM model in almost all the cases.

Table 5.9: Portfolio by Maximizing the MinMaxVar Expected Distortion by LM

γ	xom	aapl	mmm	c	adbe	amzn	gs	coh	goog	bac
0.3	0.3913	0.0186	-0.1898	-0.4835	0.1705	-0.0182	0.5947	0.0330	-0.4309	-0.0827
0.6	0.4840	0.0118	-0.3215	-0.4377	0.2258	-0.0131	0.5062	0.0283	-0.3901	-0.1008
0.8	0.5678	0.0043	-0.4225	-0.3844	0.2059	-0.0135	0.4430	0.0163	-0.3084	-0.1309

Table 5.10: Portfolio by Maximizing the MaxMinVar Expected Distortion by LM

γ	xom	aapl	mmm	c	adbe	amzn	gs	coh	goog	bac
0.3	0.2745	0.1481	-0.1890	-0.4364	0.0663	-0.1986	0.6660	0.1394	-0.3907	-0.1307
0.6	0.2812	0.1764	-0.1993	-0.4386	0.0796	-0.2357	0.6239	0.1856	-0.3915	-0.1374
0.8	0.2869	0.1824	-0.2038	-0.4414	0.0894	-0.2444	0.6050	0.2017	-0.3927	-0.1391

Table 5.11: Portfolio by Maximizing the MinMaxVar Expected Distortion by VGC

γ	xom	aapl	mmm	c	adbe	amzn	gs	coh	goog	bac
0.3	-0.1670	-0.1208	-0.2435	0.3016	0.0328	-0.0044	0.7888	-0.1813	-0.3498	-0.1698
06	-0.1584	-0.1187	-0.2532	0.2937	0.0700	0.0024	0.7776	-0.1848	-0.3749	-0.1618
0.8	-0.1514	-0.1170	-0.2604	0.2909	0.0945	0.0067	0.7672	-0.1867	-0.3927	-0.1559

Table 5.12: Portfolio by Maximizing the MaxMinVar Expected Distortion by VGC

γ	xom	aapl	mmm	c	adbe	amzn	gs	coh	goog	bac
0.3	-0.1735	-0.1220	-0.2349	0.3090	-0.0011	-0.0109	0.7965	-0.1773	-0.3277	-0.1762
06	-0.1723	-0.1220	-0.2350	0.3005	0.0034	-0.0100	0.7997	-0.1768	-0.3291	-0.1754
0.8	-0.1717	-0.1220	-0.2353	0.2972	0.0061	-0.0094	0.8005	-0.1767	-0.3303	-0.1750

Table 5.13: Realized Results of One-Month Variance Swaps by MinMaxVar

γ	LM	VGC
0.3	0.0798	0.3000
0.6	0.0757	0.2988
0.8	0.0708	0.2977

Table 5.14: Realized Results of One-Month Variance Swaps by MaxMinVar

γ	LM	VGC
0.3	0.0808	0.3003
0.6	0.0792	0.3001
0.8	0.0786	0.3000

Table 5.15: Realized Results of One-Month Variance Swaps by MinMaxVar in 2007

Dates	LM	VGC
20070319	0.1244	0.4730
20070420	0.1257	0.2314
20070518	0.1322	0.1929
20070618	0.1219	0.2902
20070719	0.1179	0.1984
20070820	0.1279	0.1584
20070917	0.1316	0.1165
20071019	0.0757	0.2988
20071119	0.1153	0.3349
20071219	0.1211	0.0932

Chapter 6

Estimating Greeks for Variance-Gamma Processes

6.1 Overview

Simulation-based derivative estimates are useful in financial engineering, especially in estimating the Greeks, which are critical for hedging financial derivatives.

Gradient estimation techniques were first applied to option pricing using infinitesimal perturbation analysis (IPA) for both European and American options by Fu and Hu [31]. Broadie and Glasserman [14] applied both IPA and likelihood ratio (LR) method to European and Asian options; see also Glasserman [38], which reviews various Monte Carlo methods for financial engineering. Fu [33] reviewed various methods of gradient estimation in stochastic simulation, including both direct and indirect methods; see also Fu [34] for more details on Monte Carlo simulation for financial engineering and various methods for estimating the Greeks through simulation.

In this chapter, we consider gradient estimation for Mountain Range options by assuming the dynamics of aseet prices is a VG process. We first price several mountain range products and then turn to gradient estimation of Greeks. We derive IPA and LR estimators for the various sensitivities where applicable, and compare them in numerical experiments to each other and to finite difference estimates. We also compare these estimates to GL estimates. The GL estimates are especially

relevant for simulation of Lévy processes, where the characteristic function is readily available. We then analyze the strengths and weaknesses of each method.

The remainder of the chapter is organized as follows. We outline some background in section 6.2. In section 6.3, we introduce several gradient estimation methods. In section 6.5, we provide a European call example. In section 6.6, we conduct numerical experiments on independent Mountain Range options. In section 6.7, we introduce a multivariate VG model. In section 6.8, we present conclusions and future work.

6.2 Background

6.2.1 Greeks

Greeks are the quantities representing the sensitivities of derivatives, such as options as in [44]. Each Greek letter measures a different dimension of the risk in an option position and the aim of a trader is to manage the Greeks so that all risks are acceptable. In this section, we provide several examples of the Greeks: Delta, Rho and Theta, defined as follows:

- Delta: Δ is the rate of change of the value of the portfolio of options with respect to the underlying asset price. In general,

$$\Delta = \frac{\partial V}{\partial S}.$$

- Vega is the rate of change of the value of the portfolio of options with respect

to the volatility of the underlying asset price. That is

$$\text{vega} = \frac{\partial V}{\partial \sigma}.$$

- Rho ρ is the rate of change of the value of the portfolio of options with respect to the interest rate. It is defined as:

$$\rho = \frac{\partial V}{\partial r}.$$

- Theta: Θ is the rate of change of the value of the portfolio of options with respect to the passage of time (maturity time) with all else remaining the same. It is defined as:

$$\Theta = \frac{\partial V}{\partial t}.$$

6.2.2 Mountain Range Options

Mountain ranges are exotic options originally marketed by Société Générale in 1998; see also Overhaus [70], Quessette [71], and Meaney [67]. These options combine characteristics of basket options and range options by basing the value of the option on several underlying assets, and by setting a time frame for the option. We first consider the case where the underlying assets are independent, and then turn to the case where the underlying assets are dependent. We subdivide the Mountain Range options into four types, depending on the specific terms of the options.

- *Atlas*: a long-term option in which the best and worst-performing securities are removed from the basket prior to execution of the option.

- *Everest*: a long-term option in which the option holder gets a payoff based on the worst-performing securities in the basket.

- *Altiplano*: a long-term option in which a vanilla option is combined with a compensatory coupon payment if the underlying security never reaches its strike price during a given period.

- *Annapurna*: a long-term option in which the option holder is rewarded if all securities in the basket never fall below a certain price during the relevant time period.

- *Himalayan*: a long-term option which is based on the performance of the best asset in the portfolio.

6.3 Gradient Estimation

Gradient estimates have been broadly applied in the finance community, for example hedging risks which depend on sensitivities of corresponding parameters of derivatives. In order to calculate gradient estimates, we just take derivatives of the price with respect to these parameters separately.

We begin with $V(\xi)$, the objective function which depends on the parameter ξ, and calculate

$$\frac{dV(\xi)}{d\xi}.$$

Suppose the objective function is an expectation of the sample performance measure

L, that is:

$$V(\xi) = E[J(\xi)] = E[J(X_1, X_2, \cdots, X_n; \xi)], \tag{6.1}$$

where $X = X_1, X_2, \cdots, X_n$ are dependent on ξ, and n is the fixed number of random variables. By law of the unconscious statistician, the expectation can be written as:

$$E[J(X)] = \int y dF_J(y) = \int J(x) dF_X(x), \tag{6.2}$$

where F_J is the distribution of J and F_X is the distribution of the input random variables X.

To make sense of the right hand side of (6.2),

$$E[J(X)] = \int_0^1 J(X(\xi; u)) du, \tag{6.3}$$

$$E[J(X)] = \int_{-\infty}^{\infty} J(x) f_X(x; \xi) dx, \tag{6.4}$$

where f_X is the probability density function of X. The parameter ξ dependence can be path-wise from the input random variables X, as shown in (6.3), or in the distribution F_X, as shown in (6.4). Considering that the parameter ξ dependence can be in two different ways, we have two different kind of ways to estimate the objective function. Therefore, we have three different direct methods, i.e. IPA, LR and GL, to estimate the gradient of the objective function sections. We apply both the indirect methods and direct methods to calculate the gradient in the following.

6.3.1 Indirect Methods

The indirect methods for estimating a gradient at ξ is simply to use finite difference, i.e., perturbing the value of each component of θ separately while hold-

ing all the other components still. Assuming for each ξ, we generate a random variable $X(\xi)$, get the value of deterministic function $J(X, \xi)$ with expectation $\hat{J}(x) = \hat{J}(x, \xi) = E[J(X, \xi)]$. The objective is to get the simulation estimate of $\frac{d}{d\xi}\hat{J}(x, \xi)$.

The one-sided forward difference gradient estimator in the ith direction is:

$$\frac{J(\xi + c) - J(\xi)}{c},$$

where c is the scalar perturbation in the $i-$th direction.

The one-sided backward difference gradient estimator in the ith direction is:

$$\frac{J(\xi) - J(\xi - c)}{c},$$

where c is the scalar perturbation in the $i-$th direction.

The two-sided symmetric difference estimator, i.e., central difference estimator is given by

$$\frac{J(\xi + c) - J(\xi - c)}{2c}.$$

6.3.2 Direct Methods

Depending on where the dependence of the parameter ξ is, either in the input random variables as in (6.3), or in the density function as in (6.4), we have two gradient estimation methods IPA and LR. It should be pointed out that when we take the derivative of the objective function $E[J(X)]$, we have to make sure that the interchangeability condition must be satisfied, i.e., we can interchange the derivative and the integral. The dominated convergence theorem can be employed to check the condition of interchangeability.

IPA

The method of infinitesimal perturbation analysis (IPA) is to estimate the sample path derivative of the parameters of interest, i.e., for $J(X(\xi,\omega))$, on the sample path ω,

$$\frac{dJ(X(\xi,\omega))}{d\xi} = \lim_{h\to 0} \frac{J(X(\xi+h,\omega)) - J(X(\xi,\omega))}{h},$$

with probability 1.

IPA estimates require the integrability condition which is easily satisfied when the performance function is continuous with respect to the given parameter. Assume the we can interchange the expectation and differentiation, the IPA estimate is:

$$\begin{aligned} \frac{dE[J(X)]}{d\xi} &= \int_0^1 \frac{dJ(X(\xi;u))}{d\xi} du \\ &= \int_0^1 \frac{dJ}{dX}\frac{dX(\xi)}{d\xi} du, \end{aligned}$$

and the estimator is:

$$\frac{dJ}{dX}\frac{dX(\xi)}{d\xi}. \tag{6.5}$$

According to the dominated convergence theorem, the condition of uniform integrability of $\frac{dJ}{dX}\frac{dX(\xi)}{d\xi}$ must be satisfied to make the interchangeability feasible. In order to make the IPA estimator an unbiased stochastic gradient estimator, we need the interchangeability of the derivative and expectation, i.e.,

$$\frac{dE[J(X)]}{d\xi} = E\left[\frac{dJ(X)}{d\xi}\right], \tag{6.6}$$

which can be considered as the condition for the unbiased IPA estimator exists.

LR

Estimating through the method of likelihood Ratio (LR) is to construct the estimators from the derivatives of the probability density function through simulation. The probability density function f of X is differentiable. The calculation of gradients through likelihood ratio method is:

$$\begin{aligned}\frac{dE[J(X)]}{d\xi} &= \int_{-\infty}^{\infty} J(x)\frac{df(x;\xi)}{d\xi}dx \\ &= \int_{-\infty}^{\infty} J(x)\frac{d\ln f(x;\xi)}{d\xi}f(x)dx\end{aligned}$$

and the estimator is

$$J(x)\frac{d\ln f(x;\xi)}{d\xi}, \tag{6.7}$$

where $\frac{d\ln f(x;\xi)}{d\xi}$ is the score function. The required condition for making interchangeability feasible of LR is placed on the density function.

6.3.3 GL Method

From the LR estimator shown in (6.7), we have to calculate the derivative of the log of the density function. However, the closed form of the density function either does not exist or exists butcomplicated. For example, the density function of the VG process is complicated, so is the derivative of the density function. To overcome this difficulty, Glasserman and Liu [39] propose a similar to "LR" method, referred as GL method, where the density or the derivative of the density are numerically approximated by the characteristic function or Laplace transform.

In the GL method, we use a set of grids to approximate the density function $g_\xi(x)$ and the derivative of the density function $\frac{dg_\xi(x)}{d\xi}$. The main idea of the algorithm

is as follows: Pick a finite grid of x values, we pre-compute values of $G_\xi, g_\xi, \dot{g}_\xi$, through numerical transform inversions. Then we can follow the general idea to approximate the estimators.

- Use the Abate-Whitt algorithm [1], each transform inversion is approximated using a finite weighted sum of transform values, given in:

$$\begin{aligned} I^{N,h}_{\xi,x}(L_f) &= \frac{he^{\sigma x}}{2\pi} L_f(\sigma) + \frac{he^{\sigma x}}{\pi} \sum_{k=1}^{N} \Big(\mathrm{Re}[L_f(\sigma + ikh)] \cos(khx) \\ &\quad - \mathrm{Im}[L_f(\sigma + ikh)] \sin(khx) \Big), \end{aligned}$$

  where N is the truncation point, and L_f is the characteristic function. Then we try to calculate G_j on the grid, pick $\sigma_+ \in (0, \sigma_u)$ and $\sigma_- \in (\sigma_l, 0)$ and let

$$G_j = \begin{cases} I^{N,h}_{\sigma_+,x_j}(L_{G_\xi}), \ \ if \ \ x_j \leq 0; \\ 1 - I^{N,h}_{\sigma_-,x_j}(L_{\bar{G}_\xi}), if \ \ x_j > 0. \end{cases}$$

- We generate $\hat{X}$ from the approximation $\hat{G}_\xi$ by setting $X = \hat{G}^{-1}_\xi(U), U \sim Unif(0,1)$ as in

$$\hat{X} = \frac{U\delta + x_{j-1}G_j - x_j G_{j-1}}{G_j - G_{j-1}}.$$

- At each simulation step, we approximate g_ξ and $\dot{g}_\xi$ through G_j as in:

$$\hat{g}_\xi(x) = \begin{cases} (G_j - G_{j-1})/\delta, if \ \ x \in [x_{j-1}, x_j), j \in J \\ 0, \quad if \ \ x < x_{\min} \ or \ x > x_{\max} \end{cases}$$

$$\dot{\hat{g}}_\xi(x) = \begin{cases} (\dot{G}_j - \dot{G}_{j-1})/\delta, \ if \ \ x \in [x_{j-1}, x_j), j \in J \\ 0, \ if \ \ x < x_{\min} \ or \ x > x_{\max} \end{cases}$$

  where $\dot{G}_j \approx \dot{G}_\xi(x_j)$ is calculated through $\dot{G}_\xi = \frac{dG_\xi}{d\xi}$. Then we can estimate the approximation score function $\hat{S}_\xi = \frac{\dot{\hat{g}}_\xi}{\hat{g}_\xi}$ at X.

- At the end of each path, the LR estimator of the derivative of $E_\xi[V(X))]$ is $V(\hat{X})\hat{S}_\xi(\hat{X})$.

To distinguish this from the general LR estimator, we will henceforth refer to it as the GL estimator. The two-sided Laplace transform of a function g is given by

$$L_g(u) = \int_{-\infty}^{\infty} e^{-ux} g(x) dx,$$

where u is a complex variable.

In this chapter, we apply the GL method to the VG process whose characteristic function is

$$L_g(u) = (1 - iu\theta\nu + \frac{1}{2}\sigma^2\nu u^2)^{-t/\nu}.$$

The required condition for interchangeability provided in the previous chapters has to be satisfied to make GL method feasible.

6.4 Problem Setting

The objective is to estimate the sensitivity

$$\frac{\partial V}{\partial \xi},$$

where V is the value (or price) of the financial derivative and ξ is the parameter of interest. For example, if V is the price of an option written on a single underlying stock and ξ is the current stock price, then this sensitivity would correspond to estimating perhaps the most well-known financial Greek, the **Delta**.

In this chapter, the derivative price will take the following form:

$$V = e^{-rt} E[J_t],$$

where t is the maturity or expiration date, r is the risk-free interest rate (assumed deterministic and constant), and J_t is an option payoff function. Two other sensitivities of interest are the **Rho** and **Theta** given by $\partial V/\partial t$ and $\partial V/\partial r$, respectively. The setting assumes that the expected payoff $E[J_t]$ cannot be easily computed, necessitating Monte Carlo simulation to estimate it. This chapter focuses on options written on a basket of underlying assets following a VG process. However, it is difficult to verify the condition for interchangebility. We use a European call to get some heuristic idea on implementation to Mountain Range options.

6.5 A European Call Example

A call option gives the buyer the right, not the obligation, to buy a certain amount of financial instrument from the seller at a certain time for a certain price. The payoff function of the European Call option with expiration time T, strike price K and risk free interest rate r is

$$J_T = (S_T - K)^+ ,$$

and the price (value) of the European call option is

$$V_T(K) = e^{-rT} (S_T - K)^+ ,$$

where $S_T = S_0 \exp((r + \omega)T + X_T)$, and X_T follows the VG process. We have two different ways to represent the VG process X_T, as a Gamma-time-changed Brownian motion, or as a difference of two Gamma processes. Hence, the gradients can be calculated in these two ways.

6.5.1 IPA for European Call Option

Using Representation 1 of VG process X_T

In the representation 1 of a VG process, the VG process X_T is a Gamma-time-changed Brownian motion, that is, $X_T = \theta\gamma_T^{(\nu)} + \sigma W_{\gamma_T^{(\nu)}}$.

1. Delta (gradient w.r.t S_0):

$$\begin{aligned}\frac{dV_T(K)}{dS_0} &= \frac{d(e^{-rT}(S_T-K)^+)}{dS_0} \\ &= e^{-rT}1_{\{S_T\geq K\}}\frac{dS_T}{dS_0},\end{aligned}$$

where $(d/dS_0)S_T = \exp((r+\omega)T + X_T) = S_T/S_0$.

2. Rho (gradient w.r.t r):

$$\frac{dV_T(K)}{dr} = -Te^{-rT}\cdot(S_T-K)^+ + e^{-rT}1_{\{S_T\geq K\}}\cdot\frac{dS_T}{dr},$$

where $(d/dr)S_T = TS_T$.

3. Gradient with respect to σ:

$$\frac{dV_T(K)}{d\sigma} = e^{-rT}\cdot 1_{\{S_T\geq K\}}\cdot\frac{dS_T}{d\sigma},$$

where

$$\frac{dS_T}{d\sigma} = S_T(\frac{dX_T}{d\sigma} + T\frac{d\omega}{d\sigma}),$$
$$\frac{d\omega}{d\sigma} = -\sigma(1-\theta\nu-\frac{1}{2}\sigma^2\nu)^{(-1)},$$

and

$$\frac{dX_T}{d\sigma} = W_{\gamma_T^{(\nu)}}.$$

4. Theta (gradient w.r.t T):

$$\frac{dV_T(K)}{dT} = -re^{-rT}(S_T - K)^+ + e^{-rT} \cdot 1_{\{S_T \geq K\}} \cdot \frac{dS_T}{dT},$$

where

$$\frac{dS_T}{dT} = S_T \cdot \left((r + \omega) + \frac{dX_T}{dT}\right),$$

and

$$\frac{dX_T}{dT} = \theta \cdot \frac{d\gamma_T^{(\nu)}}{dT} + \sigma \cdot \frac{dW_{\gamma_T^{(\nu)}}}{dT}.$$

5. Gradient with respect to ν:

$$\frac{dV_T(K)}{d\nu} = e^{-rT} I_{\{S_T \geq K\}} \frac{dS_T}{d\nu},$$

where

$$\frac{dS_T}{d\nu} = S_T\left(T\frac{d\omega}{d\nu} + \frac{dX_T}{d\nu}\right),$$
$$\frac{d\omega}{d\nu} = -\left(\frac{\theta + \sigma^2/2}{1 - \theta\nu - \sigma^2\nu/2} + \omega\right)\Big/\nu,$$

and

$$\frac{dX_T}{d\nu} = \theta\frac{d\gamma_T^{(\nu)}}{d\nu} + \sigma\frac{1}{2}(\gamma_t^{(\nu)})^{-1/2} Z \frac{d\gamma_T^{(\nu)}}{d\nu}.$$

6. Gradient with respect to θ,

$$\frac{dV_T(K)}{d\theta} = e^{-rT} 1_{\{S_T \geq K\}} \frac{dS_T}{d\theta},$$

where

$$\frac{dS_T}{d\theta} = S_T(T\frac{d\omega}{d\theta} + \frac{dX_T}{d\theta}),$$
$$\frac{d\omega}{d\theta} = -(1 - \theta\nu - \sigma^2\nu/2)^{-1},$$

and

$$\frac{dX_T}{d\theta} = \gamma_t^{(\nu)} = \nu \cdot Y.$$

In the above, Y is a Gamma process with mean t/ν and variance t/ν, Z is a standard normal distributed random variable with mean 0 and variance 1, and $W_{\gamma_T^{(\nu)}}$ is time-changed Brownian motion with mean 0 and variance $\gamma_T^{(\nu)}$. The detailed calculation of $(d/dT)W_{\gamma_T^{(\nu)}}$ and $(d/dT)\gamma_T^{(\nu)}$ is in the Appendix B.1, and the calculation for $(d/d\nu)\gamma_T^{(\nu)}$ is in the Appendix B.2.

Using Representation 2 of VG process X_T

In the representation 2 of a VG process, the X_T is a difference of two Gamma processes, i.e., $X_T = \gamma_T^{(\mu_+,\nu_+)} - \gamma_T^{(\mu_-,\nu_-)}$.

1. Delta:

$$\frac{dV_T(K)}{dS_0} = e^{-rT} 1_{\{S_T \geq K\}} \frac{S_T}{S_0}.$$

2. Rho:

$$\begin{aligned} \frac{dV_T(K)}{dr} &= -Te^{-rT} \cdot (S_T - K)^+ + e^{-rT} 1_{S_T \geq K} \cdot \frac{dS_T}{dr} \\ &= -Te^{-rT}(S_T - K)^+ + e^{-rT} 1_{\{S_T \geq K\}} \cdot TS_T. \end{aligned}$$

3. Gradient with respect to σ:

$$\frac{dV_T(K)}{d\sigma} = e^{-rT} 1_{\{S_T \geq K\}} \frac{dS_T}{d\sigma},$$

where

$$\frac{dS_T}{d\sigma} = \frac{dX_T}{d\sigma} + T\frac{d\omega}{d\sigma},$$

and

$$\frac{dX_T}{d\sigma} = \frac{d\gamma_T^{(\mu_+,\nu_+)}}{d\sigma} - \frac{d\gamma_T^{(\mu_-,\nu_-)}}{d\sigma}.$$

Since $\gamma_T^{(\mu_+,\nu_+)} = \mu_+ \cdot \nu Y$ and $\gamma_T^{(\mu_-,\nu_-)} = \mu_- \cdot \nu Y$, one has

$$\frac{dX_T}{d\sigma} = \mu Y(\frac{d\mu_+}{d\sigma} - \frac{d\mu_-}{d\sigma}) = 0.$$

It implies

$$\frac{dV_T(K)}{d\sigma} = e^{-rT} 1_{\{S_T \geq K\}} T \frac{d\omega}{d\sigma},$$

where $\frac{d\omega}{d\sigma} = -\sigma \cdot (1 - \theta\nu - \frac{1}{2}\sigma^2\nu)^{(-1)}$.

4. Theta:

$$\frac{dV_T(K)}{dT} = -re^{-rT}(S_T - K)^+ + e^{-rT} 1_{\{S_T \geq K\}} \frac{dS_T}{dT},$$

where

$$\frac{dS_T}{dT} = S_T\bigg((r + \omega) + \frac{dX_T}{dT}\bigg),$$

and

$$\begin{aligned} \frac{dX_T}{dT} &= \frac{d\gamma_T^{(\mu_+,\nu_+)}}{dT} - \frac{d\gamma_T^{(\mu_-,\nu_-)}}{dT} \\ &= (\mu_+ \cdot \nu)\frac{dY}{dT} - (\mu_- \cdot \nu)\frac{dY}{dT}. \end{aligned}$$

The calculation of $(d/d\nu)Y$ is shown in the Appendix (B.1).

5. Gradient with respect to ν:

$$\frac{dV_T(K)}{d\nu} = e^{-rT} 1_{\{S_T \geq K\}} \frac{dS_T}{d\nu}.$$

We can get that

$$\frac{dS_T}{d\nu} = S_T(T\frac{d\omega}{d\nu} + \frac{dX_T}{d\nu}),$$

where

$$\begin{aligned} \frac{dX_T}{d\nu} &= \frac{d\gamma_T^{(\mu_+,\nu_+)}}{d\nu} - \frac{d\gamma_T^{(\mu_-,\nu_-)}}{d\nu} \\ &= \mu_+ Y + \mu_+ \nu \frac{dY}{d\nu} - \mu_- Y - \mu_- \nu \frac{dY}{d\nu}. \end{aligned}$$

The detailed calculation of $(d/d\nu)Y$ is shown in the Appendix B.2.

6. Gradient with respect to θ:

$$\frac{dV_T(K)}{d\theta} = e^{-rT} 1_{\{S_T \geq K\}} \frac{dS_T}{d\theta},$$

where

$$\frac{dS_T}{d\theta} = S_T \cdot (T\frac{d\omega}{d\theta} + \frac{dX_T}{d\theta}),$$

and

$$\frac{dX_T}{d\theta} = \frac{d\gamma_T^{(\mu_+,\nu_+)}}{d\theta} - \frac{d\gamma_T^{(\mu_-,\nu_-)}}{d\theta} = \nu Y(\frac{d\mu_+}{d\theta} - \frac{d\mu_-}{d\theta}).$$

Since

$$\frac{d\mu_+}{d\theta} = \frac{1}{2}\theta \cdot (\theta^2 + 2\sigma^2/\nu)^{(-\frac{1}{2})} + \frac{1}{2}$$

and

$$\frac{d\mu_-}{d\theta} = \frac{1}{2}\theta \cdot (\theta^2 + 2\sigma^2/\nu)^{(-\frac{1}{2})} - \frac{1}{2},$$

we have

$$\frac{d\mu_+}{d\theta} - \frac{d\mu_-}{d\theta} = 1.$$

Hence,

$$\frac{dX_T}{d\theta} = \nu Y.$$

6.5.2 LR for European Call Option

Let $h(z)$ be the density function of $Z = \ln(S_T/S_0)$. Since $h(z)$ does not contain S_0, we use the Jacobian transform to get the density of S_T and calculate the derivative with respect to S_0. Let

$$f_{S_T}(s) \cdot \left|\frac{\partial S_T}{\partial Z}\right| = h(z).$$

Then the density function of S_T is

$$f_{S_T}(s) = h(\ln s - \ln S_0) \times \frac{1}{s}.$$

To calculate Delta, we use $f_{S_T}(s)$ to implement the LR method. The other gradients can be calculated through the density function $h(z)$ as follows:

1. Delta:
$$\frac{dE[V_T]}{dS_0} = \int_0^\infty e^{-rT}(s-K)^+ \cdot \frac{d\ln f_{S_T}(s)}{dS_0} \cdot f_{S_T}(s)ds.$$

2. Rho:
$$\frac{dE[V_T]}{dr} = \int_0^\infty e^{-rT}(S_0e^z - K)^+ \cdot \left(-T + \frac{d\ln h(z)}{dr} \right) \cdot h(z)dz.$$

3. Gradient w.r.t σ:
$$\frac{dE[V_T]}{d\sigma} = \int_0^\infty e^{-rT}(S_0e^z - K)^+ \cdot \frac{d\ln h(z)}{d\sigma} \cdot h(z)dz.$$

4. Gradient w.r.t. θ:
$$\frac{dE[V_T]}{d\theta} = \int_0^\infty e^{-}rT(S_0e^z - K)^+ \cdot \frac{d\ln h(z)}{d\theta} \cdot h(z)dz.$$

5. Gradient w.r.t ν:
$$\frac{dE[V_T]}{d\nu} = \int_0^\infty e^{-rT}(S_0e^z - K)^+ \cdot \frac{d\ln h(z)}{d\nu} \cdot h(z)dz.$$

6. Gradient w.r.t T:
$$\frac{dE[V_T]}{dT} = \int_0^\infty e^{-rT}(S_0e^z - K)^+ \cdot \left(-r + \frac{d\ln h(z)}{dT} \right) \cdot h(z)dz.$$

The calculations of the derivatives of $\ln h(z)$ are in the Appendix B.3.

Table 6.1: Simulated European Option Price

	VG1	VG2	GL
Price	0.2948	0.2921	0.3060
StdErr	0.0033	0.0032	0.0034

6.5.3 GL for European Call Option

Recall the characteristic function of VG process at a fixed time t, VG(σ, ν, θ, t) is given by

$$\phi_{\mathrm{VG}}(u, \sigma, \nu, \theta, t) = (1 - iu\theta\nu + 0.5\sigma^2\nu u^2)^{-t/\nu}. \tag{6.8}$$

Using this equation (6.8), and applying GL method, we can get the estimators above to estimate the gradients.

6.5.4 Numerical Experiments

Using the formulas of the estimators above, we apply Monte Carlo to do the estimation on 10000 sample paths. By selecting spot value $S_0 = 10$, $K = 10$, $r - \delta = 0.057$, $\nu = 0.2686$, $\theta = -0.1436$, $\sigma = 0.1213$ and $T = 0.2$, we get the numerical results of simulated European call option price in Table 6.1, and simulated results of gradients of European call option in Table 6.2.

From the numerical results in Table 6.1, we find that the European call option price from representation 1 of VG (VG1) and representation of VG (VG2) are closer than the one from the GL. representation 1. Numerical results in Table 6.2 indicate that both IPA and LR are applicable to most gradients, and are close to

the results from forward difference (FD), which is considered to be the benchmark of true values, with small standard errors; The FD can get good results but require additional simulations and takes a longer time. The results obtained from IPA in $\frac{dV}{dT}$ and $\frac{dV}{d\nu}$ are far away from the results from forward difference (FD) methods, which may due to the failure of interchangeability. The results obtained from LR in $\frac{dV}{d\theta}$ are far away from the results from FD, which may due to the failure of interchangeability. Moreover, the results of LR in $\frac{dV}{d\nu}$ have large standard errors, thus additional simulations are required. The GL method get similar results to the LR method, and have advantage over $\frac{dV}{d\nu}$ because the standard errors are smaller.

6.6 Independent Mountain Range Options

Mountain range options are exotic options originally marketed by Société Générale in 1998; see also Overhaus [70], Quessette [71] and Meaney [67]. The options combine characteristics of basket options and range options by basing the value of the option on several underlying assets, and by setting a time frame for the option. In this section, we only consider the case where the underlying assets are independent, and treat four types of mountain range options: Everest, Atlas, Altiplano/Annapurna, and Himalayan. To price these options, we take $(X_1(t), X_2(t))^T$ as a two-dimensional independent VG process. Since they are independent, we can deal with them separately. Two different ways of representing X_i are as follows:

- Gamma-time-changed Brownian motion

$$\begin{aligned} X_j(t) &= B^{(\theta_j,\sigma_j)}_{\gamma_j^{(\nu_j)}(t)} \\ &= \theta_j \gamma_j^{(\nu_j)}(t) + \sigma_j W_{\gamma_j^{(\nu_j)}(t)}, \end{aligned}$$

  for $j = 1, 2$.

- Difference of two Gamma processes

$$X_j(t) = \gamma_j^{(\mu_j^+,\nu_j^+)}(t) - \gamma_j^{(\mu_j^-,\nu_j^-)}(t),$$

  where $\mu_j^{\pm} = (\sqrt{\theta_j^2 + 2\sigma_j^2/\nu_j} \pm \theta_j)/2$ and $\nu_j^{\pm} = (\mu_j^{\pm})^2 \cdot \nu_j$, for $j = 1, 2$.

  The characteristic function of VG process $X_j(t)$ is given by

$$\phi_{\mathrm{VG}}(u, \sigma_j, \nu_j, \theta_j, t) = (1 - iu\theta_j\nu_j + \frac{1}{2}\sigma_j^2\nu_j u^2)^{-t/\nu_j}. \tag{6.9}$$

  Under the risk-neutral measure, the stock price would be

$$S_j^t = S_0^t \exp((r + \omega_j)t + X_j(t)),$$

where $\omega_j = \frac{1}{\nu_j}\log(1 - \theta_j\nu_j - \sigma_j^2\nu_j/2)$, for $j = 1, 2$.

  The density function of log-price $z_j = \ln(S_j^t/S_0^t)$ is:

$$h_j(z_j) = \frac{2\exp(\theta_j x_j/\sigma_j^2)}{\nu_j^{t/\nu_j}\sqrt{2\pi}\sigma_j\Gamma(\frac{t}{\nu_j})}\left(\frac{x_j^2}{2\sigma_j^2/\nu_j + \theta_j^2}\right)^{\frac{t}{2\nu_j}-\frac{1}{4}} \kappa_{\frac{t}{\nu_j}-\frac{1}{2}}\left(\frac{1}{\sigma_j^2}\sqrt{x_j^2(2\sigma_j^2/\nu_j + \theta_j^2)}\right), \tag{6.10}$$

where κ is the modified Bessel function of 2nd kind, and

$$x_j = z_j - rt - \frac{t}{\nu_j}\ln(1 - \theta_j\nu_j - \sigma_j^2\nu_j/2),$$

for $j = 1, 2$.

6.6.1 Everest Option

Given n stocks $S_1, S_2, \cdots, S_n$, the payoff for an Everest option is given by

$$J_T = \min_{i=1,\cdots,n} \left(\frac{S_i^T}{S_i^0} \right). \tag{6.11}$$

Notice that the payoff function is a continuous and monotonically non-decreasing piecewise linear function of S_i^T.

IPA for Everest Option

For ξ be $\sigma_i, \nu_i, \theta_i$, the IPA estimator is

$$\exp(-rT)\frac{dJ_T}{d\xi}.$$

For ξ be T, the IPA estimator is

$$\exp(-rT) \cdot \frac{dJ_T}{dT} - rJ_T \cdot \exp(-rT).$$

For ξ be r, the IPA estimator is

$$-TJ_T \cdot \exp(-rT),$$

this is zero.

Taking the derivative with respect to ξ, we get that

$$\begin{aligned} \frac{dJ_T}{d\xi} &= \sum_{i=1}^{n} \frac{dJ_T}{dS_i^T}\frac{dS_i^T}{d\xi} \\ &= \sum_{i=1}^{n} \left(\frac{1}{S_i^0}\frac{dS_i^T}{d\xi} - \frac{S_i^T}{(S_i^0)^2}\frac{dS_i^0}{d\xi} \right) 1_{\{\frac{s_i^T}{s_i^0} \leq \frac{s_j^T}{s_j^0}, \forall j \neq i\}}. \end{aligned}$$

Notice that (as pointed out in Hall [41])

$$\Delta = \frac{dJ_T}{dS_i^0} = \sum_{i=1}^{n} (\frac{1}{S_i 0}\frac{S_i^T}{S_i^0} - \frac{S_i^T}{(S_i^0)^2}) = 0. \tag{6.12}$$

Obviously, $dS_i^0/d\xi = 0$, where ξ could be $\sigma_i, \nu_i, \theta_i, T$. Therefore

$$\frac{dJ_T}{d\xi} = \frac{1}{S_i^0}\frac{dS_i^T}{d\xi} 1_{\{\frac{s_i^T}{s_i^0} \leq \frac{s_j^T}{s_j^0}, \forall j \neq i\}}, \tag{6.13}$$

for $\xi = \sigma_i, \nu_i, \theta_i, T$.

Fu [35] provided the IPA estimators of the derivative of S_i^T with respect to different parameters ξ, which can be written as:

$$\begin{aligned}
\frac{dS_i^t}{dr} &= tS_i^t, \\
\frac{dS_i^t}{dt} &= S_i^t(r + \omega_i + \frac{dX_i^t}{dt}), \\
\frac{dS_i^t}{d\theta_i} &= S_i^t(t\frac{d\omega_i}{d\theta_i} + \frac{dX_i^t}{d\theta_i}), \\
\frac{dS_i^t}{d\sigma_i} &= S_i^t(t\frac{d\omega_i}{d\sigma_i} + \frac{dX_i^t}{d\sigma_i}), \\
\frac{dS_i^t}{d\nu_i} &= S_i^t(t\frac{d\omega_i}{d\nu_i} + \frac{dX_i^t}{d\nu_i}),
\end{aligned}$$

for $i = 1, 2$. By taking derivative of ω, we have

$$\begin{aligned}
\frac{d\omega_i}{d\theta_i} &= -\left(1 - \theta\nu - \sigma^2\nu/2\right)^{-1}, \\
\frac{d\omega_i}{d\sigma_i} &= -\sigma_i \times \left(1 - \theta_i\nu_i - \sigma_i^2\nu_i/2\right)^{-1}, \\
\frac{d\omega_i}{d\sigma_i} &= -\left[(\theta_i + \sigma_i^2/2) \times (1 - \theta_i\nu_i - \sigma_i^2\nu_i/2)^{-1} + \omega_i\right],
\end{aligned}$$

where $i = 1, 2$. Moreover, for the independent mountain range options, we have

$$\begin{aligned}
\frac{dS_i^t}{d\theta_j} &= 0, \\
\frac{dS_i^t}{d\sigma_j} &= 0, \\
\frac{dS_i^t}{d\nu_j} &= 0,
\end{aligned}$$

for $i \neq j$.

LR for Everest Option

The LR estimator is

$$\exp(-rT) \cdot \min_{i=1,\cdots,n} \left(\frac{S_i^T}{S_i^0} \right) \frac{d \ln f(X_1^T, X_2^T, \cdots, X_n^T; \xi)}{d\xi}. \tag{6.14}$$

We consider $n = 2$ in this chapter, the density function for the 2-dimensional case is

$$f(X_1^T, X_2^T; \xi) = h_1(z_1) \cdot h_2(z_2),$$

where $h_1(z_1)$ and $h_2(z_2)$ are the density functions of X_1^T, X_2^T in equation (6.10) for $j = 1, 2$ respectively. The detailed calculation of $\frac{d \ln f(X_1^T, X_2^T; \xi)}{d\xi}$ are presented in Appendix (B.4).

GL for Everest Option

The GL estimators are the same as the LR estimators in equation (6.14). However, for the GL method, we approximate $h_i(z_i)$ and $\frac{dh_i(z_i)}{d\xi}$ numerically through the characteristic functions shown in equation (6.9).

Numerical Results

To compare the performance of the IPA, LR, GL and FD estimators for the Everest option, 10000 independent replications were simulated, by setting parameters as spot values $S_1^0 = 10$, $S_2^0 = 10$, srike price $K = 10$, $T = 0.2$ years, $\nu_1 = 0.2686$, $\nu_2 = 0.2976$, $\theta_1 = -0.1436$, $\theta_2 = -0.1033$, $\sigma_1 = 0.1213$, $\sigma_2 = 0.1532$, and $r = 0.0570$. The numerical results are shown in Table 6.4, where VG1 and VG2 represents to

the two representations the VG processes, respectively. From the numerical results in Table 6.3, we find that the independent Everest option price from VG1, VG2 and GL are similar with very small standard errors. The numerical results in Table 6.4 indicate that both IPA and LR are applicable to most gradients, and are close to the results from FD method with small standard errors; The FD can get good results but require additional sample paths and cost a longer time. The numerical results indicate that the IPA estimator matches the FD estimator with smaller standard errors. The results obtained from IPA in $\frac{dV}{dT}$ and $\frac{dV}{d\nu_i}$ are far away from the results from FD method, which may due to the failure of interchangeability. The results obtained from LR in $\frac{dV}{d\theta_i}$ are far away from the results from FD methods, which may due to the failure of interchangeability. Moreover, the results of LR in VG2 have larger standard errors compared other estimates in other methods. The results of estimates $\frac{dV}{d\nu_i}$ by LR have large difference with ones by other methods with very large standard errors, thus additional simulations are required. The GL method get similar results to the LR method, and have advantage in $\frac{dV}{d\nu_i}$ with small standard errors.

6.6.2 Atlas Option

Given two positive integers n_1, n_2 where $n_1+n_2 < n$, and n stocks $S_1, S_2, \cdots, S_n$, with strike K, the payoff for an Atlas option is

$$J_T = \left(\sum_{j=1+n_1}^{n-n_2} \frac{R_{(j)}^T}{n-(n_1+n_2)} - K \right)^+, \tag{6.15}$$

where $R^t_{(i)} = \frac{S^t_{(i)}}{S^0_{(i)}}$ is the ith smallest return from $\left\{\frac{S^t_1}{S^0_1}, \frac{S^t_2}{S^0_2}, \cdots, \frac{S^t_n}{S^0_n}\right\}$, i.e.,

$$\frac{S^t_{(1)}}{S^0_{(1)}} \leq \frac{S^t_{(2)}}{S^0_{(2)}} \leq \cdots \leq \frac{S^t_{(n)}}{S^0_{(n)}}.$$

IPA for Atlas Option

For ξ be $\sigma_i, \nu_i, \theta_i$, the IPA estimator is

$$\exp(-rT)\frac{dJ_T}{d\xi}.$$

For ξ be T, the IPA estimator is

$$\exp(-rT) \cdot \frac{dJ_T}{dT} - rJ_T \cdot \exp(-rT).$$

For ξ be r, the IPA estimator is

$$-TJ_T \cdot \exp(-rT).$$

Through calculation, we can get that

$$\begin{aligned}
\frac{dJ_T}{d\xi} &= \frac{dJ_T}{dS^T_i}\frac{dS^T_i}{d\xi}1_{\{1+n_1 \leq i \leq n-n_2\}} \\
&= \frac{1}{(n-(n_1+n_2))S^0_{(i)}}1_{\{\sum_{j=1+n_1}^{n-n_2} \frac{R^T_{(j)}}{n-n_1-n_2} > K\}}\frac{dS^T_i}{d\xi}1_{\{1+n_1 \leq i \leq n-n_2\}},
\end{aligned}$$

for $\xi = \sigma_i, \nu_i, \theta_i, T$. Note that $\frac{dS^T_i}{d\sigma_j}$, $\frac{dS^T_i}{d\nu_j}$, $\frac{dS^T_i}{d\theta_j}$, $\frac{dS^T_i}{dr}$, $\frac{dS^T_i}{dT}$ are the same as in the Everest option.

LR for Atlas Option

The LR estimator is

$$\exp(-rT) \cdot \left(\sum_{j=1+n_1}^{n-n_2} \frac{R^T_{(j)}}{n-(n_1+n_2)} - K\right)^+ \frac{d\ln f(X^T_1, X^T_2, \cdots, X^T_n; \xi)}{d\xi}. \qquad (6.16)$$

For $n = 2$, the density function is $f(X_1^T, X_2^T; \xi) = h_1(z_1) \cdot h_2(z_2)$. The calculation of $\frac{d\ln f(X_1^T, X_2^T; \xi)}{d\xi}$ is the same as the one in the Everest option.

GL for Atlas Option

For the GL method, we use the same estimator as the one of the LR method in (6.16), but approximate $h_i(z_i)$ and $\frac{dh_i(z_i)}{d\xi}$ through the characteristic function as in (6.9) instead of the density function directly.

Numerical Results

Again, the performance of the IPA, LR, GL and FD estimators are compared through 10000 independent replications of simulations, with spot values $S_1^0 = 10$, $S_2^0 = 10$, strike price $K = 0.95$ and $n_1 = 0$, $n_2 = 1$ and using the same values as in the Atlas option for the other parameter settings: $T = 0.2$ years, $\nu_1 = 0.2686$, $\nu_2 = 0.2976$, $\theta_1 = -0.1436$, $\theta_2 = -0.1033$, $\sigma_1 = 0.1213$, $\sigma_2 = 0.1532$, and $r = 0.0570$.

From the numerical results in Table 6.5, we find that the Atlas option price from VG1, VG2 and GL are similar with very small standard errors. The numerical results in Table 6.6 indicate that both IPA and LR are applicable to most gradients, and are close to the results from FD methods with small standard errors; The FD can get good results but require additional sample paths and cost a longer time. Again IPA is generally closer to the FD results, with smaller standard error than the LR and GL method. The numerical results indicate that the IPA estimator matches the FD estimator with smaller standard errors. The results obtained from IPA in

$\frac{dV}{dT}$ and $\frac{dV}{d\nu}$ are far away from the results from FD methods, which may due to the failure of interchangeability. The results obtained from LR in $\frac{dV}{d\theta}$ are far away from the results from FD methods, which may due to the failure of interchangeability. Moreover, the results of LR and GL are a little different from the ones of FD in $\frac{dV}{dT}$ but with smaller standard errors. The results of $\frac{dV}{d\sigma_i}$ by LR method are different from the results from the GL and FD methods and with large standard errors. In most cases, the GL method will gain better results than the lR method and with smaller standard errors.

6.6.3 Altiplano/Annapurna Option

Given n stocks $S_1, S_2, \cdots, S_n$, a coupon amount C, a limit L and strike K, the barrier period from t_1 to t_2, the payoff for Altiplano option is

$$J_T = \begin{cases} C & \text{if } \max\left(\frac{S_i^t}{S_i^0}\right) \leq L, \forall i, \forall t \in (t_1, t_2) \\ \left(\sum_{j=1}^{n} \frac{S_j^T}{S_j^0} - K\right)^+ & \text{otherwise} \end{cases}. \tag{6.17}$$

If the limit is a floor rather than a ceiling, the option is Annapurna.

Due to the discontinuities in the payoff functions, IPA is not applicable for Altiplano or Annapurna options.

LR for Altiplano/Annapurna Option

The LR estimator is

$$\exp(-rT) \cdot J_T \cdot \frac{d \ln f(X_1^T, X_2^T, \cdots, X_n^T; \xi)}{d\xi}. \tag{6.18}$$

For $n = 2$, the density function is $f(X_1^T, X_2^T; \xi) = h_1(z_1) \cdot h_2(z_2)$. The calculation of $\frac{d\ln f(X_1^T, X_2^T; \xi)}{d\xi}$ is the same as the one in the Everest option.

GL for Altiplano/Annapurna Option

Again, we use the LR estimator in equation (6.18), but approximate $h_i(z_i)$ and $\frac{dh_i(z_i)}{d\xi}$ through the characteristic function as in (6.9).

Numerical Results for Altiplano Option

Again, 10000 independent replications were simulated to compare the performance of the LR, GL and FD estimators, with spot values $S_1^0 = 10$, $S_2^0 = 10$, strike price $K = 1.8$, boundary levels $L = 0.75$, $C = 0.75$, barrier period $t_1 = 0$, $t_2 = 1/3$, $\nu_1 = 0.2686$, $\nu_2 = 0.2976$, $\theta_1 = -0.1436$, $\theta_2 = -0.1033$, $\sigma_1 = 0.1213$, $\sigma_2 = 0.1532$, and risk-free interest rate $r = 0.0570$. Again, the results in Table 6.7 indicate that the option price calculated from VG1, VG2 and Gl are similar. The results in Table 6.8 indicate similar conclusions as before, with the LR estimates have similar results with FD estimates except in the $\frac{dJ}{d\sigma_i}$. But FD methods require additional simulation work and cost more time. Furthermore, not surprisingly the GL method is computationally far more intensive than the usual LR method, so knowing the density saves a lot of computational burden.

6.6.4 Himalayan Option

Define

$$\mathcal{R}_i = \left\{ \frac{S_1^t}{S_1^0}, \frac{S_2^t}{S_2^0}, \cdots, \frac{S_n^t}{S_n^0} \right\}$$

$$i_1^* = arg\max \mathcal{R}_1$$

$$i_2^* = arg\max \mathcal{R}_2 \backslash \left\{ \frac{S_{i_1^*}^t}{S_{i_1^*}^0} \right\}$$

$$i_3^* = arg\max \mathcal{R}_3 \backslash \left\{ \frac{S_{i_1^*}^{t_1}}{S_{i_1^*}^0}, \frac{S_{i_2^*}^{t_2}}{S_{i_2^*}^0} \right\}$$

Given a basket of n stocks and a number of time points $\{t_0, t_1, \cdots, T\}$, first construct

$$\left\{ R_{i_1^*}^{t_1}, R_{i_2^*}^{t_2}, \cdots, R_{i_n^*}^{T} \right\},$$

the payoff of a Himalayan option is given by

$$J_T = \begin{cases} \left(\sum_{j=1}^{n} (R_{i_j^*}^{t_j} - 1) \right)^+ & \text{if globally floored,} \\ \sum_{j=1}^{n} (R_{i_j^*}^{t_j} - 1)^+ & \text{if locally floored.} \end{cases} \tag{6.19}$$

Again, since the Himalayan option has a discontinuous payoff, IPA is not applicable.

LR for Himalayan Option

The LR estimator is

$$\exp(-rT) \cdot J_T \cdot \frac{d \ln f(X_1^T, X_2^T, \cdots, X_n^T; \xi)}{d\xi}. \tag{6.20}$$

For $n = 2$, the density is $f(X_1^T, X_2^T; \xi) = h_1(z_1) \cdot h_2(z_2)$. The calculation of $\frac{d \ln f(X_1^T, X_2^T; \xi)}{d\xi}$ is the same as the one in the Everest option.

GL for Himalayan Option

For the GL method, we use the same estimator as the one of the LR method in equation (6.20), but approximate $h_i(z_i)$ and $\frac{dh_i(z_i)}{d\xi}$ through the characteristic function as in equation (6.9) instead of the density function directly.

Numerical Results for the Himalayan Option

Again, 10000 independent replications were simulated to compare the performance of the LR, GL and FD estimators, using a local floor over 0.2 year with strike price $K = 1.8$, with spot values $S_1^0 = 10$, $S_2^0 = 10$, strike price $K = 1.8$, boundary levels $\nu_1 = 0.2686$, $\nu_2 = 0.2976$, $\theta_1 = -0.1436$, $\theta_2 = -0.1033$, $\sigma_1 = 0.1213$, $\sigma_2 = 0.1532$, and risk-free interest rate $r = 0.0570$.

Again, the results in Table 6.9 indicate that the option price calculated from VG1, VG2 and Gl are similar. The results in Table 6.10 indicate similar conclusions as before, with the LR estimates all having much larger standard error than the FD estimates except in the $\frac{dJ}{d\nu_i}$. The results of GL method are closer than ones from LR method. But GL costs much more time. The FD method has much smaller standard error but requires additional number of simulations. The LR method here is very effective.

Numerical Results for $t > \nu_i$

Dr. Madan requested that for options on the VG process to be well behaved, it is generally necessary to have $T > \nu_i$. Thus, we provide numerical results with

parameters satisfying $t > \nu_i$.

For the Everest option, 10000 independent replications were simulated, by setting parameters as spot values $S_1^0 = 10$, $S_2^0 = 10$, srike price $K = 10$, $T = 0.2$ years, $\nu_1 = 0.1686$, $\nu_2 = 0.1576$, $\theta_1 = -0.1436$, $\theta_2 = -0.1033$, $\sigma_1 = 0.1213$, $\sigma_2 = 0.1532$, and $r = 0.0570$. The numerical results are shown in Table 6.11.

For independent Atlas options, we conduct 10000 independent replications of simulations, with spot values $S_1^0 = 10$, $S_2^0 = 10$, strike price $K = 0.95$ and $n_1 = 0$, $n_2 = 1$ and using the values as follows for the other parameter settings: $T = 0.2$ years, $\nu_1 = 0.1686$, $\nu_2 = 0.1576$, $\theta_1 = -0.1436$, $\theta_2 = -0.1033$, $\sigma_1 = 0.1213$, $\sigma_2 = 0.1532$, and $r = 0.0570$. The numerical results are shown in Table 6.12.

Again, 10000 independent replications were simulated for independent Altiplano options, with spot values $S_1^0 = 10$, $S_2^0 = 10$, strike price $K = 1.8$, boundary levels $L = 0.75$, $C = 0.75$, barrier period $t_1 = 0$, $t_2 = 1/3$, $\nu_1 = 0.1686$, $\nu_2 = 0.1576$, $\theta_1 = -0.1436$, $\theta_2 = -0.1033$, $\sigma_1 = 0.1213$, $\sigma_2 = 0.1532$, and risk-free interest rate $r = 0.0570$. The results are shown in Table 6.13.

Again, 10000 independent replications were simulated to compare the performance of the LR, GL and FD estimators, using a local floor over 0.2 year with strike price $K = 1.8$, with spot values $S_1^0 = 10$, $S_2^0 = 10$, strike price $K = 1.8$, boundary levels $\nu_1 = 0.2686$, $\nu_2 = 0.2976$, $\theta_1 = -0.1436$, $\theta_2 = -0.1033$, $\sigma_1 = 0.1213$, $\sigma_2 = 0.1532$, and risk-free interest rate $r = 0.0570$. The results are shown in Table 6.14.

6.7 Dependent Mountain Range Options

The payoff of dependent mountain range options is based on a basket of stocks which are correlated with each other, therefore the main difficulty in pricing such products lies in the correlation of the assets within the basket. To price these dependent options, we define the multi-dimensional VG model shown in Wang[W1]. Assume X_1 and X_2 are two correlated marginal VG processes,

$$X_1 \sim \mathrm{VG}(\theta_1, \sigma_1, \nu_1), \quad \text{and} \quad X_2 \sim \mathrm{VG}(\theta_2, \sigma_2, \nu_2).$$

We build the dependence with two additional parameters ρ and ν_0 as follows:

$$X_i = A_i + Y_i,$$

$$A_i \sim \mathrm{VG}(\theta_i \frac{\nu_i}{\nu_0}, \sigma_i \sqrt{\frac{\nu_i}{\nu_0}}, \nu_0),$$
$$Y_i \sim \mathrm{VG}\left(\theta_i(1 - \frac{\nu_i}{\nu_0}), \sigma_i \sqrt{1 - \frac{\nu_i}{\nu_0}}, \frac{1}{\frac{1}{\nu_i} - \frac{1}{\nu_0}}\right),$$

where (A_1, A_2), Y_1 and Y_2 are independent. (A_1, A_2) is a 2-dimensional $\rho-$correlated Brownian motion with associated mean and covariance matrix subordinated by a common gamma process $\Gamma(t; 1, \nu_0)$, where $\nu_0 \geqslant \max(\nu_1, \nu_2)$. Moreover, the pairwise correlation between X_i and X_j is:

$$\mathrm{Corr}(X_i, X_j) = \frac{\theta_i\theta_j\nu_i\nu_j/\nu_0 + \sigma_i\sigma_j\rho\sqrt{\nu_i\nu_j}/\nu_0}{\sqrt{\theta_j^2\nu_j + \sigma_j^2} \cdot \sqrt{\theta_i^2\nu_i + \sigma_i^2}}.$$

The joint characteristic function of the 2-dimensional correlated VG process

is:

$$\begin{aligned}\phi_{X_1(t),X_2(t)}(u_1,u_2) &= (\frac{1}{1-iu_1\theta_1\nu_1-iu_2\theta_2\nu_2+u^T\Sigma u/2})^{\frac{t}{\nu_0}} \\ &\cdot(\frac{1}{1-i\theta_1u_1\nu_1+(\sigma_1^2\nu_1/2)u_1^2})^{\frac{t}{\nu_1}-\frac{t}{\nu_0}} \\ &\cdot(\frac{1}{1-i\theta_2u_2\nu_2+(\sigma_2^2\nu_2/2)u_2^2})^{\frac{t}{\nu_2}-\frac{t}{\nu_0}},\end{aligned} \tag{6.21}$$

where $u=(u_1,u_2)^T$, and the covariance matrix

$$\Sigma=\begin{bmatrix}\sigma_1^2\nu_1 & \sigma_1\sigma_2\rho\sqrt{\nu_1\nu_2} \\ \sigma_1\sigma_2\rho\sqrt{\nu_1\nu_2} & \sigma_2^2\nu_2\end{bmatrix}. \tag{6.22}$$

Wang [82] designed an algorithm to simulate the dependent multi-dimensional Variance Gamma process as in Algorithm 5.

Under the risk-neutral measure, the stock price would be

$$S_i^t=S_i^0\exp(rt+X_i(t)+\omega_it),$$

where $X_i=A_i+Y_i$ with $\omega_i=\frac{1}{\nu_i}\log(1-\theta_i\nu_i-\sigma_i^2\nu_i/2)$, for $i=1,2$.

6.7.1 Everest Option

The payoff function J_T of an Everest Option is given by equation (6.11).

IPA for Everest Option

The IPA gradient estimators $\frac{dJ_T}{d\xi}$ is similar to the estimators in section (6.6.1). We also can see that $\Delta=0$ as in equation (6.12). The other Greeks are calculated as follows:

Algorithm 5 Algorithm to Simulate Multivariate VG process

Input: the VG parameters $(\sigma_j, \nu_j, \theta_j), j = 1, 2$; time spacing $\Delta t_1, ..., \Delta t_N$ s.t. $\sum_{i=1}^{N} \Delta t_i = T$.

Initialize: $X_1(0) = X_2(0) = 0$.

Loop: for $i = 1$ to N:

- Generate $\Delta\Delta G_i^0 \sim \Gamma(\Delta t_i/\nu_0, \nu_0)$.
- Generate $\Delta\Delta G_i^j \sim \Gamma(\Delta t_i/(1/\nu_j - 1/\nu_0), 1/1/\nu_j - 1/\nu_0)$, for $j = 1, 2$.
- Generate a multi-dimensional normal distributed vector, $\Delta W_i \sim N(0, \Sigma)$, where Σ is defined in equation (6.22).
- Generate normal distributed variables $Y_{j,i} \sim N(0, \sigma_j\sqrt{1 - \nu_j/\nu_0})$, for $j = 1, 2$.
- Return $X_j(t_{i+1}) = X_j(t_i) + \frac{\theta_j\nu_j}{\nu_0}G_i^0 + \sqrt{G_i^0}\Delta W_{i,j} + \theta_j(1 - \nu_j/\nu_0)G_i^j + \sqrt{G_i^j}Y_{j,i}$.

1. Rho:

$$\frac{dJ_T}{dr} = T\frac{S_i^T}{S_i^0}1_{\{\frac{s_i^T}{s_i^0} \leq \frac{s_j^T}{s_j^0}, \forall j \neq i\}}.$$

2. Vega:

$$\frac{dJ_T}{d\sigma_i} = \frac{S_i^T}{S_i^0} \cdot (\frac{dX_i}{d\sigma_i} + \frac{d\omega_i}{d\sigma_i}T)1_{\{\frac{s_i^T}{s_i^0} \leq \frac{s_j^T}{s_j^0}, \forall j \neq i\}},$$

where

$$\frac{d\omega_i}{d\sigma_i} = -\sigma_i(1 - \theta_i\nu_i - \sigma_i^2\nu_i/2)^{-1},$$

and

$$\frac{dX_i}{d\sigma_i} = \frac{dA_i}{d\sigma_i} + \frac{dY_i}{d\sigma_i}.$$

3. Gradient w.r.t ν_i:

$$\frac{dJ_T}{d\sigma_i} = \frac{S_i^T}{S_i^0} \cdot (\frac{dX_i}{d\nu_i} + \frac{d\omega_i}{d\nu_i}T)1_{\{\frac{s_i^T}{s_i^0} \leq \frac{s_j^T}{s_j^0}, \forall j \neq i\}},$$

where

$$\frac{d\omega_i}{d\nu_i} = -\frac{1}{\nu_i^2}\log(1 - \theta_i\nu_i - \sigma_i^2\nu_i/2) + \frac{1}{\nu_i}\frac{-\theta_i - \frac{1}{2}\sigma_i^2}{1 - \theta_i\nu_i - \sigma_i^2\nu_i/2},$$

and

$$\frac{dX_i}{d\nu_i} = \frac{dA_i}{d\nu_i} + \frac{dY_i}{d\nu_i}$$

.

4. Gradient w.r.t θ_i:

$$\frac{dJ_T}{d\theta_i} = \frac{S_i^T}{S_i^0} \cdot (\frac{dX_i}{d\theta_i} + \frac{d\omega_i}{d\theta_i}T)1_{\{\frac{s_i^T}{s_i^0} \leq \frac{s_j^T}{s_j^0}, \forall j \neq i\}},$$

where

$$\frac{d\omega_i}{d\theta_i} = \frac{-1}{1 - \theta_i\nu_i - \frac{1}{2}\sigma_i^2\nu_i},$$

and

$$\frac{dX_i}{d\theta_i} = \frac{dA_i}{d\theta_i} + \frac{dY_i}{d\theta_i}$$

.

5. Theta:

$$\frac{dJ_T}{dT} = \frac{S_i^T}{S_i^0} \cdot (r + \frac{dX_i}{dT} + \omega_i),$$

where $\frac{dX_i}{dT} = \frac{dA_i}{dT} + \frac{dY_i}{dT}$.

The calculations of $\frac{dX_i^T}{d\xi} = \frac{dA_i}{d\xi} + \frac{dY_i}{d\xi}$ for $\xi = \sigma_i, \nu_i, \theta_i, T$ refer to Appendix (B.5).

GL Method

The LR/SF estimator is

$$\min_{i=1,\cdots,n} \left(\frac{S_i^T}{S_i^0} \right) \frac{d \ln f(S_1^T, S_2^T, \cdots, S_n^T; \theta)}{d\theta}$$

Applying the 2-dimensional characteristic function in equation (6.21) and using the fact that

$$\frac{d\mu_i}{d\sigma_j} = -\sigma_i T \delta_{ij}, \quad \frac{d\mu_i}{dr} = T, \quad \frac{d\mu_i}{dT} = r - \frac{1}{2}\sigma_i^2,$$

we could apply the GL method to a 2-dimensional dependent case. The detailed calculations are in Appendix (B.6).

Numerical Results

We apply the IPA method and FD method to the dependent Everest option. We simulate 1000 paths for a period of one year. We set the strike price $K = 1.0$,

$\rho = 0.002$ and the other parameters are the same as in the independent Everest option. The numerical results are given in Table 6.15.

The numerical results indicate that for the $\frac{dV}{d\theta_i}$, both IPA and FD have similar results with relatively small standard errors; for the $\frac{dV}{d\nu_i}$, the results from IPA have larger difference with smaller standard errors than the ones from the FD method. Thus, FD methods require additional sample paths.

6.7.2 Atlas Option

The payoff function J_T of the Atlas Option is given by (6.15).

IPA Estimates

The IPA gradient estimator is

$$\begin{aligned}\frac{dJ_T}{d\xi} &= \frac{dJ_T}{dS_i^T}\frac{dS_i^T}{d\xi}1_{\{1+n_1\le i\le n-n_2\}} \\ &= \frac{1}{(n-(n_1+n_2))S_{(i)}^0}1_{\{\sum_{j=1+n_1}^{n-n_2}\frac{R_{(j)}^T}{n-n_1-n_2}>K\}}\frac{dS_i^T}{d\xi}1_{\{1+n_1\le i\le n-n_2\}},\end{aligned}$$

where $\frac{dS_i^T}{d\sigma_j}$, $\frac{dS_i^T}{dr}$, $\frac{dS_i^T}{dT}$ are calculated in the dependent Everest option.

GL Method

The LR estimator is

$$\left(\sum_{j=1+n_1}^{n-n_2}\frac{R_{(j)}^T}{n-(n_1+n_2)}-K\right)^+\frac{d\ln f(X_1^T,X_2^T,\cdots,X_n^T;\theta)}{d\theta}.$$

The numerical approximation of the score function is the same as the one in the Everest option.

Numerical Results

We apply the IPA , LR method and FD method to the dependent Atlas option. We simulate for 1000 paths for a period of one year. We set the strike price $K = 4.0$, $n_1 = 0$, and $n_2 = 1$. The numerical results in the Table 6.16.

The numerical results indicate that for both IPA and FD have similar results and small errors. But for generally, standard errors from FD are larger than the ones from IPA, and FD estimators require additional sample paths.

For the dependent Altiplano/Annapurna and Himalaya options, since IPA can't be implemented and the joint law is difficult to obtain. Future work includes applying the GL method to this setting.

6.8 Conclusions and Future Work

The IPA method performs well where applicable, but it is not applicable in all cases. When the density is available, the direct LR method is preferred to the numerical approximation, as they have essentially the same statistical properties in most cases, but the numerical approximation is computationally intensive.

From the numerical results, we offer the following conclusions:

- When applicable, the IPA estimator should be the choice for Mountain Range Options.
- Indirect methods are easy to implement, but need additional sample paths, which increases simulation costs.

- The LR and GL method has almost the same result; but GL method costs longer time and smaller standard errors; when the density function is hard to get, then the GL method can be used to approximate the density and derivative of density funtion.

- Once the pricing algorithm is implemented, the LR/SF estimator can be completed by re-using of the score function.

In the future, we will do the following work:

- Try to resolve and/or explain the apparent discrepancies in the reported numerical experiments, such as Theta by IPA, $\frac{dV}{d\nu_i}$ by IPA, and $\frac{dV}{d\theta_i}$ by LR.

- Provide a theoretical proof or check of the interchangeability for the IPA and LR method. More precisely, prove and disprove unbiasedness of $\frac{dV}{dT}$ and $\frac{dV}{d\nu_i}$ for IPA, as well as $\frac{dV}{d\theta_i}$ in LR.

- Generalize the GL method to the dependent Mountain Range Options.

- Estimate the Greeks for the Mountain Range Options using other Lévy process models, such as the normal inverse Gaussian.

Table 6.2: European option simulation results

VG1	Delta	Rho	Vega	Theta	$\frac{dV}{d\theta}$	$\frac{dV}{d\nu}$
FD	0.6967	1.3481	1.1086	0.9908	-0.548	0.2715
StdErr	0.0045	0.0045	0.046	0.4584	0.044	0.4513
IPA	0.7053	1.3578	1.1432	-1.539*	-0.4273	0.5417*
StdErr	0.0048	0.0092	0.031	0.1632*	0.0147	0.1207*
LR	0.6997	1.3589	1.4035	1.0623	1.4482*	0.24
StdErr	0.0048	0.0099	0.0889	0.0344	0.0594*	0.6801
VG2						
FD	0.7027	1.3506	1.2254	1.1208	-0.4304	0.2269
StdErr	0.0048	0.0091	0.0297	0.4545	0.0099	0.4494
IPA	0.6995	1.3404	1.2644	-2.922*	-0.4693	2.835*
StdErr	0.0048	0.0092	0.0335	0.0476*	0.0085	0.0356*
LR	0.6995	1.3579	1.5357	1.0741	1.8176*	0.259
StdErr	0.0048	0.0066	0.1252	0.0346	0.0341*	0.121
GL	0.6994	1.3568	1.4035	1.0167	1.245*	0.2754
StdErr	0.0047	0.0011	0.097	0.0003	0.0283*	0.0083

Table 6.3: Simulated Everest Option Price

	VG1	VG2	GL
Price	0.9665	0.9655	0.9681
StdErr	0.0006	0.0006	0.0006

Table 6.4: Independent Everest option simulation results

VG1	Theta	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\nu_1}$	$\frac{dV}{d\nu_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	-0.054	-0.099	-0.112	-0.0139	-0.020	0.0233	0.0111
StdErr	0.0027	0.0027	0.0025	0.0666	0.0694	0.0017	0.0017
IPA	-1.339*	-0.096	-0.116	0.4401*	0.3221*	0.0243	0.0121
StdErr	0.1164*	0.0026	0.0026	0.0603*	0.0547*	0.0016	0.0017
LR	-0.048	-0.092	-0.111	-0.0165	-0.0187	0.1944*	-0.1187*
StdErr	0.0743	0.0027	0.0025	0.0024	0.0043	0.027*	0.017*
VG2							
FD	-0.063	-0.107	-0.118	-0.014	-0.0201	0.0230	0.0229
StdErr	0.0605	0.0025	0.0025	0.088	0.0301	0.0013	0.0012
IPA	-0.1594*	-0.103	-0.115	-0.137*	-0.101*	0.0251	0.0219
StdErr	0.0229*	0.0001	0.0002	0.0156*	0.0126*	0.0017	0.0018
LR	-0.058	-0.112	-0.146	-0.053	-0.075	-0.1889*	0.1164*
StdErr	0.087	0.0273	0.014	0.098	0.097	0.0268*	0.0171*
GL	-0.057	-0.105	-0.116	-0.017	-0.021	-0.3418*	-0.2316*
StdErr	0.0232	0.0008	0.0028	0.0004	0.0004	0.0269*	0.0253*

Table 6.5: Simulated Atlas Option Price

	VG1	VG2	GL
Price	0.0426	0.0425	0.0440
StdErr	0.00035	0.00035	0.00036

Table 6.6: Independent Atlas option simulation results

VG1	Theta	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\nu_1}$	$\frac{dV}{d\nu_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	-0.016	-0.0208	-0.0282	0.0149	0.0583	-0.0086	-0.0189
StdErr	0.025	0.0015	0.0012	0.0322	0.0334	0.0011	0.0012
IPA	0.3483*	-0.0206	-0.0285	0.0848*	0.0575*	-0.0084	-0.0185
StdErr	0.0312*	0.0016	0.0012	0.0129*	0.0045*	0.0011	0.0012
LR	-0.033	-0.6823	-0.3185	0.0159	0.0575	0.1154*	0.1106*
StdErr	0.0025	0.1032	0.0591	0.0013	0.0045	0.0025*	0.0017*
VG2							
FD	-0.0151	-0.0245	-0.0286	0.0176	0.0541	-0.0107	-0.0180
StdErr	0.0262	0.0013	0.0018	0.0493	0.0494	0.0701	0.0008
IPA	-0.208*	-0.0285	-0.0211	0.0361*	0.0268*	-0.0066	-0.0121
StdErr	0.0023*	0.0001	0.0002	0.0009*	0.0007*	0.0007	0.0082
LR	-0.0478	-0.7414	-0.3755	0.0183	0.0542	0.1906*	0.1152*
StdErr	0.0019	0.1230	0.0763	0.0015	0.0012	0.0027*	0.0019*
GL	-0.046	-0.0279	-0.0272	0.0191	0.0571	-0.007*	-0.005*
StdErr	0.0027	0.0014	0.0011	0.0037	0.0038	0.0002*	0.0001*

Table 6.7: Simulated Independent Altiplano Option Price

	VG1	VG2	GL
Price	1.00083	0.9833	0.9982
StdErr	0.00016	0.00056	0.00035

Table 6.8: Independent Altiplano options simulation results

VG1	Theta	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\nu_1}$	$\frac{dV}{d\nu_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	-0.0067	0.00031	0.00037	-0.0102	-0.016	-0.0013	-0.0014
StdErr	0.0217	0.00079	0.00079	0.0214	0.0210	0.00037	0.00004
LR	-0.0047	0.04467	0.00053	-0.0238	-0.0178	0.0032*	0.0101*
StdErr	0.0085	0.013	0.0087	0.042	0.034	0.0024*	0.0065*
VG2							
FD	-0.0061	0.0071	0.0012	-0.0341	-0.032	-0.0011	-0.0010
StdErr	0.0792	0.0024	0.026	0.0784	0.0394	0.0033	0.0054
LR	-0.0043	0.0099	0.00503	-0.0094	-0.064	2.3141*	1.0956*
StdErr	0.0353	0.0024	0.0671	0.0109	0.011	0.0416*	0.0313*
GL	-0.0063	0.0081	0.00034	-0.0246	-0.0268	0.4542*	0.387*
StdErr	0.0106	0.0233	0.0216	0.0092	0.0095	0.0105*	0.0116*

Table 6.9: Simulated Himalayan Option Price

	VG1	VG2	GL
Price	0.8967	0.8967	0.8654
StdErr	0.0031	0.0031	0.0037

Table 6.10: Independent Himalayan option simulation results

VG1	Theta	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\nu_1}$	$\frac{dV}{d\nu_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	0.0239	0.0081	0.0107	-0.0118	-0.0216	-0.0025	-0.0082
StdErr	0.0409	0.0023	0.0023	0.0392	0.0403	0.0018	0.0105
LR	0.0278	0.0084	0.0175	-0.0339	-0.0228	0.0759*	0.0395*
StdErr	0.0047	0.0062	0.035	0.0031	0.0021	0.0021*	0.0015*
VG2							
FD	0.0343	0.0084	0.0109	-0.0096	-0.012	-0.003	-0.0083
StdErr	0.0063	0.0020	0.0023	0.0558	0.0533	0.0015	0.0014
LR	0.0282	0.0074	0.0179	-0.0341	-0.0229	0.0816*	0.0403*
StdErr	0.0047	0.0061	0.0037	0.0031	0.0021	0.0015*	0.0016*
GL	0.0261	0.0081	0.0169	-0.0161	-0.0157	0.1817*	0.1847*
StdErr	0.0134	0.0150	0.0119	0.0068	0.0069	0.0080*	0.0076*

Table 6.11: Independent Everest option simulation results for $t > \nu_i$

VG1	Theta	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\nu_1}$	$\frac{dV}{d\nu_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	-0.5570	-0.0834	-0.1483	-0.8667	-0.6209	0.4161	0.3523
StdErr	0.0102	0.0054	0.0043	0.0164	0.0143	0.0072	0.0078
IPA	-1.232*	-0.0793	-0.1466	0.5401*	0.4122*	0.4136	0.3506
StdErr	0.1221*	0.0055	0.0044	0.1102*	0.0952*	0.0071	0.0077
LR	-0.582	-0.0906	-0.1428	-0.8769	-0.6801	-2.961*	-1.8446*
StdErr	0.0784	0.0247	0.1111	0.0247	0.0470	0.0552*	0.0353*

Table 6.12: Independent Atlas option simulation results for $t > \nu_i$

VG1	Theta	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\nu_1}$	$\frac{dV}{d\nu_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	0.0384	0.0409	0.0154	-0.0408	-0.0121	0.0770	0.0694
StdErr	0.0014	0.0024	0.0015	0.0017	0.0023	0.0024	0.0033
IPA	0.1732*	0.0413	0.0162	-0.1227*	-0.1977*	0.0461	0.0403
StdErr	0.1262*	0.0024	0.0015	0.1421*	0.2125*	0.0019	0.0022
LR	0.0092	0.0307	0.0210	-0.0328	-0.0097	0.0733	0.0551
StdErr	0.0036	0.0109	0.0071	0.0018	0.0038	0.0026	0.0019

Table 6.13: Independent Altiplano options simulation results for $t > \nu_i$

VG1	Theta	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\nu_1}$	$\frac{dV}{d\nu_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	0.0024	0.00011	0.00019	0.0135	0.0109	-0.00014	-0.0001
StdErr	0.0199	0.0007	0.0007	0.0197	0.0196	0.0003	0.00031
LR	0.0849	0.0015	0.00020	0.0019	0.0399	0.2532*	0.1338*
StdErr	0.0199	0.0098	0.0113	0.0111	0.0125	0.0021*	0.0022*

Table 6.14: Independent Himalayan option simulation results for $t > \nu_i$

VG1	Theta	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\nu_1}$	$\frac{dV}{d\nu_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	0.0281	0.0096	0.0126	-0.0112	-0.0233	-0.0021	-0.00016
StdErr	0.0333	0.0022	0.0022	0.0317	0.0341	0.0012	0.0014
LR	0.0234	0.1012	0.0156	-0.0195	-0.0198	-0.1287	-0.0832
StdErr	0.0212	0.0293	0.2927	0.0586	0.0818	0.0087	0.0052

Table 6.15: Dependent Everest option simulation results

Multi VG	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	-0.0931	-0.0635	-0.0080	-0.0267
StdErr	0.0211	0.0253	0.0028	0.0037
IPA	-0.0122	-0.0049	-0.0081	-0.0267
StdErr	0.0037	0.0062	0.0029	0.0038

Table 6.16: Dependent Atlas option simulation results

Multi VG	$\frac{dV}{d\sigma_1}$	$\frac{dV}{d\sigma_2}$	$\frac{dV}{d\theta_1}$	$\frac{dV}{d\theta_2}$
FD	-0.191	-0.0152	-0.0544	-0.0138
StdErr	0.0229	0.0511	0.0039	0.0099
IPA	-0.172	-0.0121	-0.0594	-0.0152
StdErr	0.0163	0.0082	0.0098	0.0018

Appendix A

Derivation of the Variance Strike

The variance strike is determined such that the variance swap has a market value of 0 initially, which indicates that the initial market value of

$$P(\sum_{t=1}^{T} \frac{x_t^2}{T} 252 - \sigma_k^2). \tag{A.1}$$

is zero. In other words, the variance strike σ_k^2 must equal the spot value of the annualized realized variance, i.e.,

$$\sigma_k^2 = \text{spot value of } \sum_{t=1}^{T} \frac{x_t^2}{T} 252. \tag{A.2}$$

252 is due to the fact that one year has 252 trading days. For simplicity, let R be the spot value of $\sum_{t=1}^{T} x_t^2$. Following Madan [54], we start with the Taylor expansion of e^x up to the second order, i.e.,

$$e^x = 1 + x + \frac{x^2}{2} + \mathrm{O}(x) \approx 1 + x + \frac{x^2}{2}.$$

Then, we have

$$x^2 = 2 \times e^x - 2 - 2 \times x.$$

By substituting x by x_t, we have

$$x_t^2 = 2 \times e_t^x - 2 - 2 \times x_t, \tag{A.3}$$

and,

$$\sum_{t=1}^{T} x_t^2 \approx 2 \times \sum_{t=1}^{T} (e^{x_t} - 1 - x_t).$$

Hence, equation (A.2) is simplified as

$$\begin{aligned}\sigma_k^2 &= \text{R} \\ &= \text{spot value of } \frac{252}{T} \times 2 \times \sum_{t=1}^{T}(e^{x_t} - 1 - x_t). \end{aligned} \tag{A.4}$$

In order to obtain R, we need to figure out the spot value of $2 \times \sum_{t=1}^{T}(e^{x_t} - 1 - x_t)$.

By substituting x_t by $\log \frac{S_t}{S_{t-1}}$, we have

$$\begin{aligned}2 \times \sum_{t=1}^{T}(e^{x_t} - 1 - x_t) &= 2 \times \sum_{t=1}^{T}(\frac{S_t}{S_{t-1}} - 1 - \log\frac{S_t}{S_{t-1}}) \\ &= 2 \times \sum_{t=1}^{T}(\frac{S_t}{S_{t-1}} - 1) - 2 \times \sum_{t=1}^{T}(\log\frac{S_t}{S_{t-1}}).\end{aligned}$$

Since

$$\begin{aligned}2 \times \sum_{t=1}^{T}\log\frac{S_t}{S_{t-1}} &= 2 \times \sum_{t=1}^{T}(\log(S_t) - \log(S_{t-1})) \\ &= 2 \times (\log(S_T) - \log(S_0)),\end{aligned}$$

we have

$$\begin{aligned}\sum_{t=1}^{T} x_t^2 &\approx 2 \sum_{t=1}^{T}(\frac{S_t}{S_{t-1}} - 1) - 2\log(S_T) + 2\log(S_0) \\ &= \sum_{t=1}^{T}\left(\frac{2}{S_{t-1}}(S_t - S_{t-1})\right) - 2\log(S_T) + 2\log(S_0). \end{aligned} \tag{A.5}$$

Consequently,

$$\begin{aligned}\sigma_k^2 &= \text{Spot Value of}\Big\{\frac{252}{T} \\ &\times \left[\sum_{t=1}^{T}\left(\frac{2}{S_{t-1}}(S_t - S_{t-1})\right) - 2\log(S_T) + 2\log(S_0)\right]\Big\}.\end{aligned}$$

For simplicity, we split the equation (A.5) into three parts:

$$\text{M}_1 = \sum_{t=1}^{T}(\frac{2}{S_{t-1}}(S_t - S_{t-1})),$$

$$\mathrm{M}_2 = -2\log(S_T),$$

$$\mathrm{M}_3 = 2\log(S_0).$$

Therefore the variance strike σ_k^2 can be expressed as

$$\sigma_k^2 = \frac{252}{T} \times \text{Spot Value of } \Big\{\mathrm{M}_1 + \mathrm{M}_2 + \mathrm{M}_3\Big\}. \tag{A.6}$$

In the following sections, we focus on deriving the spot values.

A.1 Spot Value of M_1

In order to get the spot value of M_1, we construct the portfolio as follows: from $t = 1, 2, ..., T-1$,

- On day $t-1$, we buy a bond with face value a and maturity T;
- On day $t-1$, we borrow certain amount of cash which is required to be repayed on day T to buy b shares of stock at the stock price S_{t-1};
- On day t, we sell the stocks.
- On day T, we pay back all the cash borrowed at a risk-free interest rate r.

Consequently, the cash flow on day T from the portfolios above, i.e., from bonds, stocks and paying back the borrowed money would be

$$a + bS_t \exp(r\frac{(T-t)}{365}) - bS_{t-1}\exp(r\frac{(T-t+1)}{365}). \tag{A.7}$$

We construct these portfolios daily from day $t = 0$ through day $t = T-1$. Since only the bond has a current cost, the spot value of these series of trades would equal

the spot value of the summation of the series of all the bonds we bought during these days. Hence, we buy a total of T shares of bonds, and each has the face value a.

By setting $a = 2(\exp(\frac{r}{365}) - 1)$ and $b = \frac{2\exp(\frac{-r(T-t)}{365})}{S_{t-1}}$, we find that the cash flow in equation (A.7) is equal to M_1, i.e.,

$$a + bS_t \exp(\frac{r(T-t)}{365}) - bS_{t-1} \exp(\frac{r(T-t+1)}{365}) = 2\frac{(S_t - S_{t-1})}{S_{t-1}}.$$

Therefore, the spot value of M_1 at T would equal the spot value of total of T shares of bonds whose face value is $a = 2(\exp(\frac{r}{365}) - 1)$ with maturity T. Since the spot value of equation (A.7) is

$$2(\exp(\frac{r}{365}) - 1) \times e^{-\frac{rT}{365}},$$

it implies that

$$\text{Spot Value of } M_1 = 2(e^{\frac{r}{365}} - 1) \times e^{-\frac{rT}{365}}.$$

A.2 Spot Value of M_2

The spot value of the cash flow $2\log S_0$ at the T is $2\log S_0 e^{\frac{-rT}{365}}$.

A.3 Spot Value of M_3

Madan [54] shows the value of $-2\log(S_T)$ is

$$\begin{aligned} -2 \quad & \log(S_0)e^{-rT/365} - 2(1 - e^{-rT/365}) \\ + \quad & \text{market value of options bought (MVO).} \end{aligned} \tag{A.8}$$

In conclusion, we can get the spot value R of $\sum_{t=1}^{T} x_t^2$ as

$$\begin{aligned} \mathrm{R} &= 2(e^{\frac{r}{365}} - 1) \times e^{-\frac{rT}{365}} + 2\log S_0 e^{\frac{-rT}{365}} \\ &\quad -2\log(S_0)e^{-rT/365} - 2(1 - e^{-rT/365}) + \mathrm{MVO} \\ &= \mathrm{MVO} + 2T(e^{r/365} - 1)e^{-rT/365} - 2(1 - e^{-rT/365}). \end{aligned}$$

It should be noted that if the risk-free interest rate is 0, the spot value R can be simplified as

$$\mathrm{R} = \mathrm{MVO}.$$

Appendix B

Some Calculations

B.1 Calculation of $\frac{d\gamma_t^{(\nu)}}{dt}$ and $\frac{dW_{\gamma_t^{(\nu)}}}{dt}$

$\gamma_t^{(\nu)}$ is a Gamma process with unit drift, mean t and variance νt. $W_{\gamma_t^{(\nu)}}$ is a time-changed Brownian motion, which is normally distributed with mean 0 and variance $\gamma_t^{(\nu)}$. Assuming Z is a standard normal distributed random variable with mean 0 and variance 1, we can write $W_{\gamma_t^{(\nu)}}$ as:

$$W_{\gamma_t^{(\nu)}} = \sqrt{\gamma_t^{(\nu)}} \cdot Z.$$

Therefore, the derivative of $W_{\gamma_t^{(\nu)}}$ with respect to t is:

$$\frac{dW_{\gamma_t^{(\nu)}}}{dt} = \frac{1}{2} \cdot (\gamma_t^{(\nu)})^{(-\frac{1}{2})} Z \cdot \frac{d\gamma_t^{(\nu)}}{dt}.$$

Everything reduces to calculating $\frac{d\gamma_t^{(\nu)}}{dt}$.

In our dissertation, we apply the inverse transform method to generate $\gamma_t^{(\nu)}$, with the shape parameter $\frac{t}{\nu}$ and the scale parameter ν. Let the cumulative distribution function (c.d.f.) of $\gamma_t^{(\nu)}$ be $F(x, \frac{t}{\nu}, \nu)$, then

$$F(x, \frac{t}{\nu}, \nu) = \frac{\int_0^{\frac{x}{\nu}} s^{\frac{t}{\nu}-1} e^{-s} ds}{\int_0^{\infty} s^{\frac{t}{\nu}-1} e^{-s} ds}.$$

Let $\gamma_t^{(\nu)} = \nu Y$, where Y is a Gamma process with mean $\frac{t}{\nu}$ and variance $\frac{t}{\nu}$, then

$$\frac{d\gamma_t^{(\nu)}}{dt} = \nu \cdot \frac{dY}{dT}$$

and the c.d.f. of Y is

$$F_Y(y) = \frac{\int_0^y s^{\frac{t}{\nu}-1}e^{-s}ds}{\int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}ds}.$$

It implies that the derivative of X_t with respect to t is

$$\frac{dX_t}{dt} = \theta \cdot \frac{d\gamma_t^{(\nu)}}{dt} + \sigma\frac{1}{2} \cdot (\gamma_t^{(\nu)})^{(-\frac{1}{2})} \cdot Z \cdot \frac{d\gamma_t^{(\nu)}}{dt},$$

where

$$\frac{d\gamma_t^{(\nu)}}{dt} = \nu\frac{dY}{dt}.$$

According to the inverse transform method, we generate a random variable u which is uniformly distributed on $[0, 1]$, and set u equals the c. d. f. of Y.

Thus, by setting

$$u = F_Y(Y) = \frac{\int_0^Y s^{\frac{t}{\nu}-1}e^{-s}ds}{\int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}ds},$$

we have

$$\int_0^Y s^{\frac{t}{\nu}-1}e^{-s}ds = u \cdot \int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}ds.$$

Taking the derivative with respect to t on both sides, we get:

$$\int_0^Y s^{\frac{t}{\nu}-1}e^{-s}\ln s\frac{1}{\nu}ds + Y^{\frac{t}{\nu}-1}e^{-Y}\frac{dY}{dt} = u \cdot \int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\frac{1}{\nu}ds.$$

Finally, we have

$$\frac{dY}{dt} = e^Y Y^{1-\frac{t}{\nu}}\frac{1}{\nu} \cdot \left(u\int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds - \int_0^Y s^{\frac{t}{\nu}-1}e^{-s}\ln s ds\right). \tag{B.1}$$

B.2 Calculation of $\frac{d\gamma_t^{(\nu)}}{d\nu}$

For the calculation of $\frac{d\gamma_t^{(\nu)}}{d\nu}$, we have

$$\frac{d\gamma_t^{(\nu)}}{d\nu} = Y + \nu\frac{dY}{d\nu}.$$

Apply the same method as in the section (B.1). Generating a uniformly distributed random variable u between 0 and 1, and setting

$$u = F_Y(Y) = \frac{\int_0^Y s^{\frac{t}{\nu}-1}e^{-s}ds}{\int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}ds},$$

we have

$$\int_0^Y s^{\frac{t}{\nu}-1}e^{-s}ds = u \cdot \int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}ds.$$

Taking the derivative with respect to ν on both sides, we get:

$$\int_0^Y s^{\frac{t}{\nu}-1}e^{-s}\ln s(-\frac{t}{\nu^2})ds + Y^{\frac{t}{\nu}-1}e^{-Y}\frac{dY}{d\nu} = u \cdot \int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s(-\frac{t}{\nu^2})ds.$$

Finally, we get

$$\frac{dY}{d\nu} = Y^{1-\frac{t}{\nu}}e^Y\frac{t}{\nu^2}(-u\int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds + \int_0^Y s^{\frac{t}{\nu}-1}e^{-s}\ln s ds). \qquad \text{(B.2)}$$

Denote $I_1 = \int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds$ and $I_2 = \int_0^Y s^{\frac{t}{\nu}-1}e^{-s}\ln s ds$. As in Section (B.1), we need to test whether the integrals I_1 and I_2 are integrable or not, if not, the derivative $\frac{dY}{dt}$ can not be calculated; if exist, then the derivative above could be approximated numerically. Therefore, we provide the two lemmas to show that I_1 and I_2 are integrable, as well as the proof.

Lemma B.2.1. *For $t > 0, \nu > 0$, the integral $I_1 = \int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds$ are integrable.*

Proof. Let's separate the integral I_1 into two parts,

$$\begin{aligned} I_1 &= \int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds \\ &= \int_0^1 s^{\frac{t}{\nu}-1}e^{-s}\ln s ds + \int_1^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds \end{aligned}$$

For the 2nd term, it is easy to see that

$$\int_1^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds < \int_1^\infty s^{\frac{t}{\nu}}e^{-s}ds.$$

For $t > 0$ and $\nu > 0$, there exists $M > 1$ such that, for all $s > M$, we have

$$s^{\frac{t}{\nu}} < e^{\frac{s}{2}}.$$

Then,

$$\begin{aligned}\int_1^{+\infty} s^{\frac{t}{\nu}} e^{-s} ds &= \int_1^M s^{\frac{t}{\nu}} e^{-s} ds + \int_M^{+\infty} s^{\frac{t}{\nu}} e^{-s} ds \\ &\leq \int_1^M s^{\frac{t}{\nu}} e^{-s} ds + \int_M^{+\infty} e^{-\frac{s}{2}} ds \\ &< +\infty,\end{aligned}$$

Therefore,

$$0 < \int_1^{+\infty} s^{\frac{t}{\nu}} e^{-s} \ln s ds < +\infty.$$

For the 1st term, it is easy to see that, there exists a constant l,

$$0 > \int_0^1 s^{\frac{t}{\nu}-1} e^{-s} \ln s ds > -l \int_0^1 \frac{1}{s^{1-\frac{t}{2\nu}}} ds > -\infty.$$

Therefore, the summation of 1st term and 2nd term I_1 is convergent. □

Lemma B.2.2. *For $t > 0, \nu > 0$, the integral $I_2 = \int_0^Y s^{\frac{t}{\nu}-1} e^{-s} \ln s ds$ converges.*

Proof. The proof is similar to lemma (B.2.1). □

B.3 Calculation of $\frac{d \ln h(z)}{d\xi}$

The density function of $Z = \ln(\frac{S_t}{S_0})$ is

$$h(z) = \frac{2\exp(\frac{\theta x}{\sigma^2})}{\nu^{t/\nu}\sqrt{2\pi}\sigma\Gamma(\frac{t}{\nu})} \left(\frac{x^2}{2\sigma^2/\nu + \theta^2}\right)^{\frac{t}{2\nu}-\frac{1}{4}} \kappa_{\frac{t}{\nu}-\frac{1}{2}} \left(\frac{\sqrt{x^2(\frac{2\sigma^2}{\nu} + \theta)}}{\sigma^2}\right),$$

where $x = z - rt - t\omega$.

Taking the natural logarithm of the density function, we get

$$\begin{aligned}\ln h(z) &= \ln 2 + \frac{\theta x}{\sigma^2} - \frac{t}{\nu}\ln\nu - \ln\sqrt{2\pi} - \ln\sigma - \ln\Gamma(\frac{t}{\nu}) \\ &+ \left(\frac{t}{2\nu} - \frac{1}{4}\right)\cdot\left(2\ln(x) - \ln(2\sigma^2/\nu + \theta^2)\right) \\ &+ \ln(\kappa_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{\sqrt{(x^2(2\sigma^2/\nu+\theta^2))}}{\sigma^2}\right).\end{aligned}$$

Let $\tau = \frac{\sqrt{\frac{2x^2\sigma^2}{\nu} + x^2\theta^2}}{\sigma^2}$, and let

$$\begin{aligned}KB = KB(\tau) &= \kappa_{\frac{t}{\nu}-\frac{1}{2}}(\frac{\sqrt{(x^2(2\sigma^2/\nu+\theta^2))}}{\sigma^2}) \\ &= \frac{1}{2}(\frac{\tau}{2})^{\frac{t}{\nu}-\frac{1}{2}}\int_0^\infty \frac{e^{-s-\frac{\tau^2}{4s}}}{s^{\frac{t}{\nu}+\frac{1}{2}}}ds,\end{aligned}$$

where κ is the modified Bessel function of 2nd kind. The calculation of the derivatives of the log-density function $\ln h(z)$ w.r.t different parameters is as follows.

- The derivative of log-density with respect to σ is:

$$\begin{aligned}\frac{d\ln h(z)}{d\sigma} &= -\frac{2\theta x}{\sigma^3} + \frac{\theta}{\sigma^2}\cdot(-t\frac{d\omega}{d\sigma}) - \frac{1}{\sigma} + (\frac{t}{2\nu} - \frac{1}{4}) \\ &\times \left(\frac{2}{x}\frac{dx}{d\sigma} - \frac{4\sigma/\nu}{2\sigma^2/\nu+\theta^2}\right) + \frac{1}{KB}\cdot\frac{dKB}{d\sigma}.\end{aligned}$$

- The derivative of log-density with respect to θ is:

$$\begin{aligned}\frac{d\ln h(z)}{d\theta} &= \frac{x}{\theta^2} + \frac{\theta}{\sigma^2}(-t\frac{d\omega}{d\theta}) + (\frac{t}{2\nu} - \frac{1}{4}) \\ &\times \left(\frac{2}{x}(-t\frac{d\omega}{d\theta}) - \frac{2\theta}{2\sigma^2/\nu+\theta^2}\right) + \frac{1}{KB}\frac{dKB}{d\theta}.\end{aligned}$$

- The derivative of log-density with respect to ν is:

$$\begin{aligned}\frac{d\ln h(z)}{d\nu} &= \frac{\theta}{\sigma^2}(-t\frac{d\omega}{d\nu}) - \frac{t}{\nu} + \frac{t\ln\nu}{\nu^2} - \frac{1}{\Gamma(\frac{t}{\nu})}\frac{d\Gamma(\frac{t}{\nu})}{d\nu} \\ &- \frac{t}{2\nu^2}(2\ln x - \ln(2\sigma^2/\nu+\theta^2)) \\ &+ (\frac{t}{2\nu} - \frac{1}{4})(\frac{2}{x}(-t\frac{d\omega}{d\nu}) + \frac{\frac{2\sigma^2}{\nu^2}}{\frac{2\sigma^2}{\nu}+\theta^2}) + \frac{1}{KB}\frac{dKB}{d\nu}.\end{aligned}$$

- The derivative of log-density with respect to maturity time t is:

$$\begin{aligned}\frac{\ln h(z)}{dt} &= \frac{\theta}{\sigma^2}(-r-\omega) - \frac{1}{\nu}\ln\nu - \frac{1}{\Gamma(\frac{t}{\nu})}\frac{d\Gamma(\frac{t}{\nu})}{dt} \\ &+ \frac{1}{2\nu}(2\ln x - \ln(2\sigma^2/\nu + \theta^2)) \\ &+ (\frac{t}{2\nu} - \frac{1}{4})(\frac{2}{x}(-r-\omega)) + \frac{1}{KB}\frac{dKB}{dt}.\end{aligned}$$

Since the Gamma function $\Gamma(\frac{t}{\nu}) = \int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}ds$, we have:

- the derivative of Gamma function with respect to ν:

$$\frac{d\Gamma(\frac{t}{\nu})}{d\nu} = -\frac{t}{\nu^2}\int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds = -\frac{t}{\nu^2}\times I_1,$$

- the derivative of Gamma function with respect to t:

$$\frac{d\Gamma(\frac{t}{\nu})}{dt} = \frac{1}{\nu}\int_0^\infty s^{\frac{t}{\nu}-1}e^{-s}\ln s ds = \frac{1}{\nu}\times I_1.$$

Then we begin to calculate the derivative of $KB(\tau)$ w.r.t different parameters.

- The derivative of $KB(\tau)$ with respect to σ is:

$$\frac{dKB(\tau)}{d\sigma} = \frac{dKB(\tau)}{d\tau}\cdot\frac{d\tau}{d\sigma},$$

where

$$\begin{aligned}\frac{d\tau}{d\sigma} &= -\frac{2}{\sigma^3}\left(\frac{2x^2\sigma^2}{\nu} + x^2\theta^2\right)^{\frac{1}{2}} + \frac{1}{2\sigma^2}\left(\frac{2x^2\sigma^2}{\nu} + x^2\theta^2\right)^{-\frac{1}{2}} \\ &\cdot \left(-2\theta^2 xt\frac{d\omega}{d\sigma} + \frac{4x^2\sigma}{\nu} - \frac{4x\sigma^2 t}{\nu}\cdot\frac{d\omega}{d\sigma}\right).\end{aligned}$$

- The derivative of $KB(\tau)$ with respect to θ is:

$$\frac{dKB(\tau)}{d\theta} = \frac{dKB(\tau)}{d\tau}\cdot\frac{d\tau}{d\theta},$$

where $\frac{d\tau}{d\theta} = \frac{1}{2\sigma^2}(\frac{2x^2\sigma^2}{\nu} + x^2\theta^2)^{-\frac{1}{2}}\cdot(-\frac{4x\sigma^2}{\nu}t\frac{d\omega}{d\theta} + 2x^2\theta - 2x\theta^2 t\frac{d\omega}{d\theta})$.

- The derivative of $KB(\tau)$ with respect to ν is:

$$\frac{dKB(\tau)}{d\nu} = \frac{1}{2}\frac{dM_t}{d\nu}\cdot P_t + \frac{1}{2}M_t\cdot\frac{dP_t}{d\nu},$$

where $P_t = \int_0^\infty \frac{e^{-s-\frac{\tau^2}{4s}}}{s^{\frac{t}{\nu}+\frac{1}{2}}}ds$, and $M_t = (\frac{\tau}{2})^{\frac{t}{\nu}-\frac{1}{2}}$.

The derivatives are

$$\begin{aligned}\frac{dM_t}{d\nu} &= M_t\Bigg[-\frac{t}{\nu^2}\ln\left(\frac{1}{2\sigma^2}\sqrt{\frac{2x^2\sigma^2}{\nu}+x^2\theta^2}\right)\\ &- (\frac{t}{\nu}-\frac{1}{2})(\frac{2x^2\sigma^2}{\nu}+x^2\theta^2)^{-1}\cdot\frac{2x^2\sigma^2}{\nu^2}\Bigg],\end{aligned}$$

and

$$\begin{aligned}\frac{dP_t}{d\nu} &= \int_0^\infty \exp(-s-\frac{x^2}{2s\sigma^2\nu}-\frac{x^2\theta^2}{4s\sigma^4})\cdot\left(\frac{x^2}{2s\sigma^2\nu^2}\right)s^{-\frac{t}{\nu}-\frac{1}{2}}ds\\ &+ \int_0^\infty \exp(-s-\frac{x^2}{2s\sigma^2\nu}-\frac{x^2\theta^2}{4s\sigma^4})\cdot\left(\frac{x^2}{2s\sigma^2\nu^2}\right)s^{-\frac{t}{\nu}-\frac{1}{2}}\ln s\cdot(\frac{t}{\nu^2})ds.\end{aligned}$$

- The derivative of $KB(\tau)$ with respect to t is:

$$\begin{aligned}\frac{dKB(\tau)}{dt} &= \frac{1}{2}(\frac{\tau}{2})^{\frac{t}{\nu}-\frac{1}{2}}\ln(\frac{\tau}{2})\cdot\frac{1}{\nu}\int_0^\infty\frac{e^{-s-\frac{\tau^2}{4s}}}{s^{\frac{t}{\nu}+\frac{1}{2}}}ds\\ &+ \frac{1}{2}(\frac{\tau}{2})^{\frac{t}{\nu}-\frac{1}{2}}\int_0^\infty e^{-s-\frac{\tau^2}{4s}}s^{-\frac{t}{\nu}-\frac{1}{2}}\ln s(-\frac{1}{\nu})ds.\end{aligned}$$

- The derivative of $KB(\tau)$ with respect to τ is:

$$\begin{aligned}\frac{dKB(\tau)}{d\tau} &= \frac{1}{4}(\frac{\tau}{2})^{\frac{t}{\nu}-\frac{3}{2}}(\frac{t}{\nu}-\frac{1}{2})\int_0^\infty\frac{e^{-s-\frac{\tau^2}{4s}}}{s^{\frac{t}{\nu}+\frac{1}{2}}}ds\\ &+ \frac{1}{2}(\frac{\tau}{2})^{\frac{t}{\nu}-\frac{1}{2}}(-\frac{\tau}{2})\int_0^\infty\frac{e^{-s-\frac{\tau^2}{4s}}}{s^{\frac{t}{\nu}+\frac{1}{2}}}\frac{1}{s}ds.\end{aligned}$$

B.4 Calculation of of $\frac{d\ln f(X_1^t,X_2^t;\xi)}{d\xi}$

Since $f(X_1^t, X_2^t;\xi) = h_1(z_1)\cdot h_2(z_2)$, we have

$$\frac{d\ln f(X_1^t,X_2^t;\xi)}{d\xi} = \frac{d\ln h_1(z_1)}{d\xi} + \frac{d\ln h_2(z_2)}{d\xi}.$$

If $i \neq j$, we have

$$\frac{d \ln h_i(z_i)}{d\sigma_j} = 0,$$
$$\frac{d \ln h_i(z_i)}{d\theta_j} = 0,$$
$$\frac{d \ln h_i(z_i)}{d\nu_j} = 0.$$

If $i = j$, the calculation of $\frac{d \ln h_i(z_i)}{d\xi}$ is the same as $\frac{d \ln h_i(z_i)}{d\xi}$ calculated in the last section for each i.

The density function $h_i(z)$ is

$$h_i(z) = \frac{2\exp(\frac{\theta_i x}{\sigma_i^2})}{\nu_i^{t/\nu_i}\sqrt{2\pi}\sigma_i\Gamma(\frac{t}{\nu_i})}\left(\frac{x_i^2}{2\sigma_i^2/\nu_i + \theta_i^2}\right)^{\frac{t}{2\nu_i}-\frac{1}{4}} \kappa_{\frac{t}{\nu_i}-\frac{1}{2}}\left(\frac{\sqrt{x^2(\frac{2\sigma_i^2}{\nu_i}+\theta_i)}}{\sigma_i^2}\right),$$

where $x_i = z - rt - t\omega_i$.

Taking the natural logarithm of the density function, we get

$$\begin{aligned}
\ln h_i(z) &= \ln 2 + \frac{\theta_i x_i}{\sigma_i^2} - \frac{t}{\nu_i}\ln\nu_i - \ln\sqrt{2\pi} - \ln\sigma_i - \ln\Gamma(\frac{t}{\nu_i}) \\
&+ \left(\frac{t}{2\nu_i} - \frac{1}{4}\right)\cdot\left(2\ln(x_i) - \ln(2\sigma_i^2/\nu_i + \theta_i^2)\right) \\
&+ \ln\left(\kappa_{\frac{t}{\nu_i}-\frac{1}{2}}(\frac{\sqrt{(x_i^2(2\sigma_i^2/\nu_i+\theta_i^2))}}{\sigma_i^2})\right).
\end{aligned}$$

Let $\tau_i = \frac{\sqrt{\frac{2x_i^2\sigma_i^2}{\nu_i}+x_i^2\theta_i^2}}{\sigma_i^2}$, and let

$$\begin{aligned}
KB_i = KB_i(\tau_i) &= \kappa_{\frac{t}{\nu_i}-\frac{1}{2}}\left(\frac{\sqrt{(x_i^2(2\sigma_i^2/\nu_i+\theta_i^2))}}{\sigma_i^2}\right) \\
&= \frac{1}{2}(\frac{\tau_i}{2})^{\frac{t}{\nu_i}-\frac{1}{2}}\int_0^\infty \frac{e^{-s-\frac{\tau_i^2}{4s}}}{s^{\frac{t}{\nu_i}+\frac{1}{2}}}ds,
\end{aligned}$$

where κ is the modified Bessel function of 2nd kind. The calculation of the derivatives of the log-density function $\ln h_i(z)$ w.r.t different parameters is as follows.

- The derivative of log-density with respect to σ_i is:

$$\begin{aligned}\frac{d\ln h_i(z)}{d\sigma_i} &= -\frac{2\theta_i x_i}{\sigma_i^3}+\frac{\theta_i}{\sigma_i^2}\cdot(-t\frac{d\omega_i}{d\sigma_i})-\frac{1}{\sigma_i}+(\frac{t}{2\nu_i}-\frac{1}{4})\\ &\times \left(\frac{2}{x_i}\frac{dx_i}{d\sigma_i}-\frac{4\sigma_i/\nu_i}{2\sigma_i^2/\nu_i+\theta_i^2}\right)+\frac{1}{KB_i}\cdot\frac{dKB_i}{d\sigma_i}.\end{aligned}$$

- The derivative of log-density with respect to θ_i is:

$$\begin{aligned}\frac{d\ln h_i(z)}{d\theta_i} &= \frac{x_i}{\theta_i^2}+\frac{\theta_i}{\sigma_i^2}(-t\frac{d\omega_i}{d\theta_i})+(\frac{t}{2\nu_i}-\frac{1}{4})\\ &\times \left(\frac{2}{x_i}(-t\frac{d\omega_i}{d\theta_i})-\frac{2\theta_i}{2\sigma_i^2/\nu_i+\theta_i^2}\right)+\frac{1}{KB_i}\frac{dKB_i}{d\theta_i}.\end{aligned}$$

- The derivative of log-density with respect to ν_i is:

$$\begin{aligned}\frac{d\ln h_i(z)}{d\nu_i} &= \frac{\theta_i}{\sigma_i^2}(-t\frac{d\omega_i}{d\nu_i})-\frac{t}{\nu_i}+\frac{t\ln\nu_i}{\nu_i^2}-\frac{1}{\Gamma(\frac{t}{\nu_i})}\frac{d\Gamma(\frac{t}{\nu_i})}{d\nu_i}\\ &- \frac{t}{2\nu_i^2}(2\ln x_i-\ln(2\sigma_i^2/\nu_i+\theta_i^2))\\ &+ (\frac{t}{2\nu_i}-\frac{1}{4})(\frac{2}{x_i}(-t\frac{d\omega_i}{d\nu_i})+\frac{\frac{2\sigma_i^2}{\nu_i^2}}{\frac{2\sigma_i^2}{\nu_i}+\theta_i^2})+\frac{1}{KB_i}\frac{dKB_i}{d\nu_i}.\end{aligned}$$

- The derivative of log-density with respect to maturity time t is:

$$\begin{aligned}\frac{\ln h_i(z)}{dt} &= \frac{\theta_i}{\sigma_i^2}(-r-\omega_i)-\frac{1}{\nu_i}\ln\nu_i-\frac{1}{\Gamma(\frac{t}{\nu_i})}\frac{d\Gamma(\frac{t}{\nu_i})}{dt}\\ &+ \frac{1}{2\nu_i}(2\ln x_i-\ln(2\sigma_i^2/\nu_i+\theta_i^2))\\ &+ (\frac{t}{2\nu_i}-\frac{1}{4})(\frac{2}{x_i}(-r-\omega_i))+\frac{1}{KB_i}\frac{dKB_i}{dt}.\end{aligned}$$

Since the Gamma function $\Gamma(\frac{t}{\nu_i})=\int_0^\infty s^{\frac{t}{\nu_i}-1}e^{-s}ds$, we have:

- the derivative of Gamma function with respect to ν_i:

$$\frac{d\Gamma(\frac{t}{\nu_i})}{d\nu_i}=-\frac{t}{\nu_i^2}\int_0^\infty s^{\frac{t}{\nu_i}-1}e^{-s}\ln s ds=-\frac{t}{\nu_i^2}\times I_1^{(i)},$$

- the derivative of Gamma function with respect to t:

$$\frac{d\Gamma(\frac{t}{\nu_i})}{dt} = \frac{1}{\nu_i}\int_0^\infty s^{\frac{t}{\nu_i}-1}e^{-s}\ln s ds = \frac{1}{\nu_i}\times I_1^{(i)}.$$

Then we begin to calculate the derivative of $KB_i(\tau_i)$ w.r.t different parameters.

- The derivative of $KB_i(\tau_i)$ with respect to σ_i is:

$$\frac{dKB_i(\tau_i)}{d\sigma_i} = \frac{dKB_i(\tau_i)}{d\tau_i}\cdot\frac{d\tau_i}{d\sigma_i},$$

where

$$\begin{aligned}\frac{d\tau_i}{d\sigma_i} &= -\frac{2}{\sigma_i^3}(\frac{2x_i^2\sigma_i^2}{\nu_i}+x^2\theta_i^2)^{\frac{1}{2}} \\ &+ \frac{1}{2\sigma_i^2}(\frac{2x_i^2\sigma_i^2}{\nu_i}+x^2\theta_i^2)^{-\frac{1}{2}}\cdot(-2\theta_i^2xt\frac{d\omega_i}{d\sigma_i}+\frac{4x_i^2\sigma_i}{\nu_i}-\frac{4x_i\sigma_i^2t}{\nu_i}\cdot\frac{d\omega_i}{d\sigma_i}).\end{aligned}$$

- The derivative of $KB_i(\tau_i)$ with respect to θ_i is:

$$\frac{dKB_i(\tau_i)}{d\theta_i} = \frac{dKB_i(\tau_i)}{d\tau_i}\cdot\frac{d\tau_i}{d\theta_i},$$

where

$$\frac{d\tau_i}{d\theta_i} = \frac{1}{2\sigma_i^2}(\frac{2x_i^2\sigma_i^2}{\nu_i}+x^2\theta_i^2)^{-\frac{1}{2}}\cdot(-\frac{4x_i\sigma_i^2}{\nu_i}t\frac{d\omega_i}{d\theta_i}+2x_i^2\theta_i-2x_i\theta_i^2t\frac{d\omega_i}{d\theta_i}).$$

- The derivative of $KB_i(\tau_i)$ with respect to ν_i is:

$$\frac{dKB_i(\tau_i)}{d\nu_i} = \frac{1}{2}\frac{dM_t^{(i)}}{d\nu_i}\cdot P_t^{(i)}+\frac{1}{2}M_t^{(i)}\cdot\frac{dP_t^{(i)}}{d\nu_i},$$

where $P_t^{(i)} = \int_0^\infty \frac{e^{-s-\frac{\tau_i^2}{4s}}}{s^{\frac{t}{\nu_i}+\frac{1}{2}}}ds$, and $M_t^{(i)} = (\frac{\tau_i}{2})^{\frac{t}{\nu_i}-\frac{1}{2}}$.

The derivatives are

$$\begin{aligned}\frac{dM_t^{(i)}}{d\nu_i} &= M_t^{(i)}\Big[-\frac{t}{\nu_i^2}\ln\Big(\frac{1}{2\sigma_i^2}\sqrt{\frac{2x_i^2\sigma_i^2}{\nu_i}+x_i^2\theta_i^2}\Big) \\ &- (\frac{t}{\nu_i}-\frac{1}{2})(\frac{2x_i^2\sigma_i^2}{\nu_i}+x^2\theta_i^2)^{-1}\cdot\frac{2x_i^2\sigma_i^2}{\nu_i^2}\Big],\end{aligned}$$

and

$$\begin{aligned}\frac{dP_t^{(i)}}{d\nu_i} &= \int_0^\infty \exp(-s-\frac{x_i^2}{2s\sigma_i^2\nu_i}-\frac{x_i^2\theta_i^2}{4s\sigma_i^4})\cdot\left(\frac{x_i^2}{2s\sigma_i^2\nu_i^2}\right)s^{-\frac{t}{\nu_i}-\frac{1}{2}}ds \\ &+ \int_0^\infty \exp(-s-\frac{x_i^2}{2s\sigma_i^2\nu_i}-\frac{x_i^2\theta_i^2}{4s\sigma_i^4})\cdot\left(\frac{x_i^2}{2s\sigma_i^2\nu_i^2}\right)s^{-\frac{t}{\nu_i}-\frac{1}{2}}\ln s\cdot(\frac{t}{\nu_i^2})ds.\end{aligned}$$

- The derivative of $KB_i(\tau_i)$ with respect to t is:

$$\begin{aligned}\frac{dKB_i(\tau_i)}{dt} &= \frac{1}{2}(\frac{\tau_i}{2})^{\frac{t}{\nu_i}-\frac{1}{2}}\ln(\frac{\tau_i}{2})\cdot\frac{1}{\nu_i}\int_0^\infty\frac{e^{-s-\frac{\tau_i^2}{4s}}}{s^{\frac{t}{\nu_i}+\frac{1}{2}}}ds \\ &+ \frac{1}{2}(\frac{\tau_i}{2})^{\frac{t}{\nu_i}-\frac{1}{2}}\int_0^\infty e^{-s-\frac{\tau_i^2}{4s}}s^{-\frac{t}{\nu_i}-\frac{1}{2}}\ln s(-\frac{1}{\nu_i})ds.\end{aligned}$$

- The derivative of $KB_i(\tau_i)$ with respect to τ_i is:

$$\begin{aligned}\frac{dKB_i(\tau_i)}{d\tau_i} &= \frac{1}{4}(\frac{\tau_i}{2})^{\frac{t}{\nu_i}-\frac{3}{2}}(\frac{t}{\nu_i}-\frac{1}{2})\int_0^\infty\frac{e^{-s-\frac{\tau_i^2}{4s}}}{s^{\frac{t}{\nu_i}+\frac{1}{2}}}ds \\ &+ \frac{1}{2}(\frac{\tau_i}{2})^{\frac{t}{\nu_i}-\frac{1}{2}}(-\frac{\tau_i}{2})\int_0^\infty\frac{e^{-s-\frac{\tau_i^2}{4s}}}{s^{\frac{t}{\nu_i}+\frac{1}{2}}}\frac{1}{s}ds.\end{aligned}$$

B.5 Calculation of Derivatives of $\frac{dX_i}{d\xi}$

According to the algorithm for sampling the multi-dimensional VG process $X_i(t)$, we have to generate two-dimensional independent normal distributed random variables $(\check{Z}_1, \check{Z}_2)'$. In each simulation step, we can calculate the derivatives as follows.

- The derivative of $X_i(t)$ with respect to σ_i is:

$$\frac{dX_i(t)}{d\sigma_i} = \sqrt{G_t^i}\sqrt{1-\frac{\nu_i}{\nu_0}}\tilde{Z} + \sqrt{G_t^0}\frac{dW_{t,i}}{d\sigma_i}$$

where $\frac{dW_{t,1}}{d\sigma_1} = \check{Z}_1\sqrt{\nu_1}$, and $\frac{dW_{t,2}}{d\sigma_2} = \rho\sqrt{\nu_2}\check{Z}_2 + \sqrt{1-\rho^2}\sqrt{\nu_2}\check{Z}_2$.

For $i \neq j$, we have

$$\frac{dX_i(t)}{d\sigma_j} = 0.$$

- The derivative of $X_i(t)$ with respect to ν_i is:

$$\begin{aligned}\frac{dX_i(t)}{d\nu_i} &= \frac{\theta_i}{\nu_0}G_t^0 + \theta_i(1-\frac{\nu_i}{\nu_0})\frac{dG_t^i}{d\nu_i} - \frac{\theta_i}{\nu_0}G_t^i + \frac{1}{2}(G_t^i)^{-\frac{1}{2}}\sigma_i\sqrt{1-\frac{\nu_i}{\nu_0}}\tilde{Z}\frac{dG_t^i}{d\nu_i} \\ &+ \sqrt{G_t^i}\sigma_i\frac{1}{2}(1-\frac{\nu_i}{\nu_0})^{-\frac{1}{2}}(-\frac{1}{\nu_0})\tilde{Z} + \sqrt{G_t^0}\frac{dW_{t,i}}{d\nu_i},\end{aligned}$$

where $\frac{dW_{t,1}}{d\nu_1} = \check{Z}_1\sigma_1\frac{1}{2}\nu_1^{-\frac{1}{2}}$ and $\frac{dW_{t,2}}{d\nu_2} = \sigma_2\rho\check{Z}_1\frac{1}{2}\nu_2^{-\frac{1}{2}} + \sigma_2\sqrt{1-\rho^2}\check{Z}_2 + \frac{1}{2}\nu_2^{-\frac{1}{2}}$.

For $i \neq j$, we have

$$\frac{dX_i(t)}{d\nu_j} = 0.$$

- The derivative of $X_i(t)$ with respect to θ_i is:

$$\frac{dX_i(t)}{d\theta_i} = \frac{\nu_i}{\nu_0}G_t^0 + (1-\frac{\nu_i}{\nu_0})G_t^i.$$

For $i \neq j$, we have

$$\frac{dX_i(t)}{d\theta_j} = 0.$$

- The derivative of $X_i(t)$ with respect to t is:

$$\begin{aligned}\frac{dX_i(t)}{dt} &= \frac{\theta_i\nu_i}{\nu_0}\frac{dG_t^0}{dt} + \frac{1}{2}(G_t^0)^{-\frac{1}{2}}\frac{dG_t^0}{dt}W_{t,i} \\ &+ \theta_i(1-\frac{\nu_i}{\nu_0})\frac{dG_t^i}{dt} + \frac{1}{2}(G_t^i)^{-\frac{1}{2}}\frac{dG_t^i}{dt}\sigma_i\sqrt{1-\frac{\nu_i}{\nu_0}}\hat{Z}.\end{aligned}$$

Now we calculate $\frac{dG_t^i}{d\nu_i}$, $\frac{dG_t^0}{dt}$, and $\frac{dG_t^i}{dt}$.

Assume $G_t^0 \sim \Gamma(\frac{t}{\nu_0}, \nu_0)$ is Gamma distributed with mean t and variance $t\nu_0$. G_t^i is distributed with mean t and variance $\frac{t}{\frac{1}{\nu_i}-\frac{1}{\nu_0}}$, i.e. $G_t^i \sim \Gamma(\frac{t}{\frac{1}{\nu_i}-\frac{1}{\nu_0}}, \frac{1}{\frac{1}{\nu_i}-\frac{1}{\nu_0}})$. For simplicity, we let $Y \sim \Gamma(\frac{t}{\hat{\nu}}, 1)$. $\frac{dY}{dt}$ and $\frac{dY}{d\nu}$ can be calculated as in equation(B.1) and equation (B.2). The steps of calculation are:

- Since $G_t^i = \nu \cdot Y = (\frac{1}{\nu_i} - \frac{1}{\nu_0}) \cdot Y$, and $\nu = \frac{1}{\nu_i} - \frac{1}{\nu_0}$, we get

$$\frac{dG_t^i}{d\nu_i} = (-\frac{1}{\nu_i^2}) \cdot (Y + \nu \frac{dY}{d\tau}).$$

- Let $\nu = (\frac{1}{\nu_i} - \frac{1}{\nu_0})$, then we get

$$\frac{dG_t^i}{dt} = (\frac{1}{\nu_i} - \frac{1}{\nu_0}) \frac{dY}{dt}.$$

- Let $\nu = (\frac{1}{\nu_i} - \frac{1}{\nu_0})$, then $G_t^0 = \nu_0 \cdot Y$. We get

$$\frac{dG_t^0}{dt} = \nu_0 \frac{dY}{dt}.$$

B.6 Calculation of Derivatives of $\phi_{X(t)}(u)$

Let X_t follow a one-dimensional Variance Gamma process, the characteristic function is

$$\begin{aligned} \phi_{X(t)}(u) &= \int_0^\infty e^{iux} f_X(x) dx \\ &= E[\exp(iuX)] \\ &= \left(1 - i\theta\nu u + \sigma^2 \nu u^2/2\right)^{-t/\nu} \end{aligned}$$

The Laplace transform of the function g is

$$L_g(u) = \int_0^\infty e^{-ux} g(x) dx.$$

Hence, the Laplace transform of the density function of $X(t)$ is

$$\begin{aligned} L_{f_{\mathrm{VG}}} &= \int_0^\infty e^{-ux} f_{\mathrm{VG}}(x) dx \\ &= E[e^{-uX}] \\ &= \left(1 + \theta\nu u - \sigma^2 \nu u^2/2\right)^{-t/\nu}. \end{aligned}$$

If the Laplace transform of a density function g is $L_g(u)$, the Laplace transform of the corresponding distribution function G is

$$L_G(u) = L_g(u)/u. \tag{B.3}$$

It is easy to see that if $\bar{G} = 1 - G$, then the Laplace transform $L_{\bar{G}}$ of $\bar{G}$ can also be calculated from $L_{\bar{G}}(u) = -L_g(u)/u$. Therefore, the Laplace transform L_{VG} of the distribution function of $X(t)$ can be calculated from equation (B.3) by

$$L_{\mathrm{VG}}(u) = \left(1 + \theta\nu u - \sigma^2\nu u^2/2\right)^{-t/\nu} \Big/ u.$$

Consequently, the derivative of the Laplace transform $L_{\mathrm{VG}}(u)$ of the distribution function of $X(t)$ with respect to σ can be expressed as

$$\frac{dL_{\mathrm{VG}}}{d\sigma} = \sigma u t \cdot \left(1 + \theta\nu u - \sigma^2\nu u^2/2\right)^{-t/\nu-1}. \tag{B.4}$$

The derivative of $L_{\mathrm{VG}}(u)$ with respect to θ is

$$\frac{dL_{\mathrm{VG}}}{d\theta} = -t \cdot \left(1 + \theta\nu u - \sigma^2\nu u^2/2\right)^{-t/\nu-1}. \tag{B.5}$$

The derivative of $L_{\mathrm{VG}}(u)$ with respect to ν is

$$\frac{dL_{\mathrm{VG}}}{d\nu} = L_{\mathrm{VG}} \cdot \left(\frac{t}{\nu^2}\ln(1 + \theta\nu u - \sigma^2\nu u^2/2) - \frac{t}{\nu}\frac{\theta u - \sigma^2 u^2/2}{1 + \theta\nu u - \sigma^2\nu u^2/2}\right) \Big/ u. \tag{B.6}$$

The derivative of $L_{\mathrm{VG}}(u)$ with respect to t is

$$\frac{dL_{\mathrm{VG}}}{dt} = L_{\mathrm{VG}} \cdot \ln(1 + \theta\nu u - \sigma^2\nu u^2/2) \cdot (-\frac{1}{\nu})/u. \tag{B.7}$$

Appendix C

C Code

C.1 Implementation of Variance-Gamma process

```
/* the VG process X */
double VarianceGammaprocess(double tt, double X_first)
{
        double next, delta_G;
        if (tt/nu ≥ 1)
           delta_G = nu*gamma_process2(tt/nu);
        else
           delta_G = nu * gamma_process1(tt/nu);
        W_gamma_t = sqrt(delta_G) * Z;
        gamma_t = delta_G;
        gamma_temp = gamma_t / nu;
        next = X_first + theta * delta_G + sigma * sqrt(delta_G) * Z;
        gamma_temp_ips = gamma_temp * theta_ips;
        fprintf(outfile3, %"lf\n",next);
        return next;
}
```

C.2 Implementation of Stock Price under VG

```
/* the stock price S(t) */
double stock_process(double sspot, double t, double XX)
{
        double next;
        next = sspot * exp( (rate + w) * t + XX);
        fprintf(outfile2, "%lf\n",next);
        return next;
}
```

C.3 Implementation of A European Call Option Price

```
for (sim_num = 0; sim_num < sim_limit; ++sim_num)
{
        uniform1 = mcg();
        uniform2 = mcg();
        Z = boxmuller(uniform1,uniform2);
        X_t = Variance_Gamma_process(h, X_0);
        stock = stock_process(spot, h, X_t);
        z_t = (rate + w) * T + X_t;
        price_temp = ((stock > K) ? exp(- rate * h) * (stock - K) : 0);
        price += price_temp/sim_limit;
        fprintf(outfile1,"%lf\n", price_temp);
```

```
}
```

C.4 Implementation of Uniform Distributed Random Variables

```
/* Generate uniform random variates */
double mcg()
{
        hi = (int)(seed1/q);
        lo = seed1 - q * hi;
        test = a * lo - r * hi;
        if (test > 0.0)
           seed1 = test;
        else
           seed1 = test + m;
        rand_u = seed1/m;
        return rand_u;
}
```

C.5 Implementation of Normal Distributed Random Variables

```
/* The Box Muller Algorithm */
double boxmuller(double random_u1, double random_u2)
{
        double R; double V;
```

```
        R = - 2 * log(random_u1);
        V = 2 * pi * random_u2;
        gauss1 = sqrt(R) * cos(V);
        gauss2 = sqrt(R) * sin(V);
        return gauss1;
}
```

C.6 Implementation of IPA Estimates

```
/* (1) w.r.t: S_0 = spot; —Delta*/
delta_IPA_temp = exp(- rate * h) * stock/spot * ((stock ≥ K)? 1.0:0.0);
delta_IPA += delta_IPA_temp/sim_limit;
/* (2) w.r.t: rate; */
rho_IPA_temp = - h * exp( - rate * h) * ((stock ≥ K) ? (stock - K):0.0)
        + exp( - rate * h) * ((stock ≥ K)? 1.0:0) * h * stock;
rho_IPA += rho_IPA_temp /sim_limit;
/* (3) w.r.t: vol = sigma; —-vega */
vega_IPA_temp = exp(- rate * h) * ((stock ≥ K) ? 1.0:0.0) * stock
        * ( h * de_w_de_sigma + de_X_t_de_sigma);
vega_IPA += vega_IPA_temp /sim_limit;
/* (4) w.r.t: T; */
grad_T_IPA_temp = - rate * exp(- rate * h)* ((stock ≥ K)?(stock - K):0.0)
        + exp( - rate * h) * ((stock ≥ K) ? 1.0:0.0) *
```

```
stock * ((rate-w) + de_X_t_de_T);

grad_T_IPA += grad_T_IPA_temp /sim_limit;

/* (5) w.r.t: nu; */

grad_nu_IPA_temp = exp(- rate * h) * ((stock ≥ K)? 1.0:0.0) * stock
        * (h * de_w_de_nu + de_X_t_de_nu);

grad_nu_IPA += grad_nu_IPA_temp /sim_limit;

/* (6) w.r.t: theta; */

grad_theta_IPA_temp = exp(-rate * h) * ((stock ≥ K)? 1.0:0.0) * stock
        * (h * de_w_de_theta + de_X_t_de_theta);

grad_theta_IPA += grad_theta_IPA_temp /sim_limit;
```

C.7 Implementation of LR Estimates

```
/* (1) w.r.t: S_0 = spot; —Delta*/

delta_LR_temp = exp(- rate * T) * exp(z_t) * ( ((spot * exp(z_t)) ≥ K)? 1.0:0.0 );

delta_LR += delta_LR_temp/sim_limit;

/* (2) w.r.t: rate; */

rho_LR_temp = - T * exp( - rate * T) * ( ((spot * exp(z_t)) ≥ K)? (spot
        * exp(z_t) - K) : 0.0);

rho_LR += rho_LR_temp/sim_limit;

/* (3) w.r.t: vol = sigma; —-vega */

vega_LR_temp = exp(- rate * T) * ( ((spot * exp(z_t)) ≥ K) ? (spot
```

```
        * exp(z_t) - K) : 0.0) * de_logh_de_sigma;
vega_LR += vega_LR_temp /sim_limit;
/* (4) w.r.t: theta; */
grad_theta_LR_temp = exp(- rate * T) * ( ((spot * exp(z_t)) ≥ K)? (spot
        * exp(z_t) - K) : 0.0) * de_logh_de_theta;
grad_theta_LR += grad_theta_LR_temp /sim_limit;
/* (5) w.r.t: nu; */
grad_nu_LR_temp = exp(-rate * T) * ( ((spot * exp(z_t)) ≥ K)? (spot
        * exp(z_t) - K) : 0.0) * de_logh_de_nu;
grad_nu_LR += grad_nu_LR_temp /sim_limit;
/* (6) w.r.t: T; */
grad_T_LR_temp = exp(- rate * T) * ( ((spot * exp(z_t)) ≥ K)? (spot
        * exp(z_t) - K) : 0.0) * (-rate + de_logh_de_T);
grad_T_LR +=grad_T_LR_temp /sim_limit;
```

Bibliography

[1] Abate, J., and W. Whitt. (1992). *The Fourier-Series Method for Inverting Transforms of Probability Distributions*, Queueing Systems: Theory and Applications 10:5-88.

[2] Applebaum, D. (2003). *Lévy processes and stochastic calculus*. Cambridge University Press.

[3] Artzner, P., Delbaen, F., Eber, J. M., and D. Heath. (1999). *Coherent Measures of Risk*, Mathematical Finance, 9, 203-228.

[4] Asmussen, S., and P. W. Glynn. (2007). *Stochastic Simulation: Algorithm and Analysis*, Springer.

[5] Avramidis, A. N., and P. L'Ecuyer. (2006). *Efficient Monte Carlo and quasi-Monte Carlo option pricing under the Variance-Gamma model*. Management Science, 52: 1930-1944, 2006.

[6] Bakshi, G. and Chen, Z. (1997). *An alternative valuation model for contingent claims*. Journal of Financial economics, 44, 1, 123-165.

[7] Barndorff-Nielsen, O. E. (1997). *Normal inverse Gaussian Distributions and Stochastic Volatility modelling*. Scandinavian Journal of Statistics, 24, 1-13.

[8] Berd, A., Engle, R., and A. Voronov. (2005). *The underlying dynamics of credit correlations*, preprint.

[9] Bertoin, J. (1996). *Lévy Processes*. Cambridge University Press: Cambridge.

[10] Black, F., and M. Scholes. (1973). *The Pricing of Options and Corporate Liabilities*. Journal of Political Economy 81 (3): 637-654.

[11] Bossu, S., Strasser, E., and R. Guichard. (2005). *Just What You Need To Know About Variance Swaps*. JPMorgan Equity Derivatives report. http://math.uchicago.edu˜sbossu/VarSwaps.pdf.

[12] Boyarchenko, S. I., and S. Z. Levendorskii. (2002). *Non-Gaussian Merton-Black-Scholes Theory*. World Scientific, Singapore.

[13] Boyle, P. (1977). *Options: A Monte Carlo Approach*. Journal of Financial Economics, May 1977, pp.323-338.

[14] Broadie, M., and P. Glasserman. (1996). *Estimating security price derivatives using simulation*, Management Science 42(2) 269-285, 1996.

[15] Burtschell, X., Gregory, J., and J.-P. Laurent. (2005). *A comparative analysis of CDO pricing models*, working paper.

[16] Cao, L., and M. C. Fu. (2011). *Estimating Greeks for Variance-Gamma*, Proceedings of the 2010 Winter Simulation Conference, WSC 2010, Baltimore, Maryland. pp. 2620-2628.

[17] Carr, P., Geman, H., Madan, D. B., and M. Yor. (2007). *Self-Decomposability and Option Pricing*, Mathematical Finance, vol.17, No.1, 31-57.

[18] Carr, P., Geman, H., Madan, D. B., and M. Yor. (2002). *The fine structure of asset returns: an empirical investigation*, Journal of Business, April 2002, 75 2, 305-32.

[19] Carr, P., and R. Lee. (2009). *Robust replication of volatility derivatives*.

[20] Carr, P., and D. Madan. (1999). *Option valuation using the fast Fourier transform*. Journal of Computational Finance, 2:61-73, 1999.

[21] Carr, P., Madan, D. B., and J. J. V. Alvarez. (2010). *Markets, Profits, Capital, Leverage and Return*

[22] Cherny, A., and D. B. Madan. (2009). *New Measures for Performance Evaluation*, Review of Financial Studies, 22, 2571-2606.

[23] Choi, J. Y., Salandro, D., and K. Shastri. (1988). *On the Estimation of Bid-Ask Spreads: Theory and Evidence*, Journal of Financial and Quantitative Analysis, 23, 219-230.

[24] Cont, R., and P. Tankov. (2004). *Financial Modelling with jump processes*. Chapman and Hall-CRC financial mathematical series.

[25] Copeland, T. C., and D. Galai. (1983). *Information effects on the bid ask spread*, Journal of Finance, 1457-1469.

[26] Demeterfi, K., Derman, E., Kamal, M., and J. Zou (1999). *More Than You Ever Wanted To Know About Volatility Swaps*. Goldman Sachs Quantitative Strategies Research Notes. http://www.ederman.com/new/docs/gs−volatility_swaps.pdf.

[27] Easley, D., and M. O'Hara. (1987), *Price, trade size and information in securities markets*, Journal of Financial Economics, 19, 69-90.

[28] Eberlein, E., and U. Keller. (1995) *Hyperbolic distributions in finance*. Bernoulli, Volume 1, Number 3 (1995), 281-299.

[29] Eberlein E., and D. B. Madan. (2009). *Maximally Acceptable Portfolios*. Preprint, University of Maryland.

[30] Eberlein E., and D. B. Madan. (2008). *On Correlating Lévy Processes*. Preprint, University of Maryland.

[31] Fu, M. C., and J.-Q. Hu. (1995). *Sensitivity analysis for Monte Carlo simulation of option pricing*, Probability in the Engineering and Informational Sciences 9(3) 417-446, 1995.

[32] Fu, M. C., and J.-Q. Hu. (1997) *Conditional Monte Carlo: Gradient Estimation and Optimization Applications*, Kluwer Academic Publishers, Chapter 1, Chapter 2.

[33] Fu, M. C. (2008). *Stochastic Gradient Estimation*, S. G. Henderson, B. L. Nelson, eds., Simulation. Handbooks in Operations Research and Management Science, Elsevier, Amsterdam, The Netherlands, 575-616. Chapter 19.

[34] Fu, M. C. (2006). *What you should know about simulation and derivatives*, Naval Research Logistics 55(8), 723-736, 2006.

[35] Fu, M. C. (2007). *Variance-Gamma and Monte Carlo*, Advances in Mathematical Finance (eds. Fu, Jarrow, Yen, Elliott), 21-35, Birkhäuser.

[36] Gairat, A. (2004). *Variance Swaps*. In collaboration with IVolatility.com. www.ivolatility.com/doc/VarianceSwaps.pdf.

[37] George, T. J., Kaul, G., and M. Nimalendran. (1991). *Estimation of the Bid-Ask Spreads and its Components: A New Approach*, Review of Financial Studies, 4, 623-656.

[38] Glasserman, P. (2004). *Monte Carlo methods in financial engineering*, Applications of mathematics 53, Springer, New York, 2004.

[39] Glasserman, P., and Z. Liu. (2007). *Sensitivity Estimates From Characteristic Functions*, pp.932-940 in Proceedings of the Winter Simulation Conference, IEEE Press, Piscataway, New Jersey, 2007.

[40] Glosten, L. R. and P. R. Milgrom. (1985). *Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders*, Journal of Financial Economics, 14, 71-100.

[41] Hall, A. (2009). *Gradient Estimation and Mountain Range options*, PhD. dissertation, 2009.

[42] Hertz, D. (1964). *Risk Analysis in Capital Investment*. Harvard Business Review, Jan-Feb 1964, vol. 42.

[43] Heston, S. (1993). *A closed-form solution for options with stochastic volatility with applications to bond and currency options*. Review of financial studies, 6, 327-343.

[44] Hull, J. (2003). *Options, Futures,and Other Derivatives 5th Edition*, Pearson, Prentice Hall, New Jersey.

[45] Hyvärinen, A. (1999). *Fast and Robust Fixed-point Algorithms for Independent Component Analysis*, IEEE Trans. on Neural Networks, 10, 626-634.

[46] Hyvärinen, A., Karhuen, J., and E. Oja. (2001). *Independent Component Analysis*. John Wiley & Sons inc., New York.

[47] Jäckel, P. (2002). *Monte Carlo Methods in Finance*. Wiley.

[48] Jondeau, E., Poon, S., and M. Rockinger. (2007). *Financial Modeling under Non Gaussian Distributions*, Springer Verlag, London.

[49] Kotz, S., and S. Nadarajah. (2004). *Multivariate t Distributions and Their Applications*, Cambridge University Press, Cambridge.

[50] Laurent, J-P., and J. Gregory. (2005). *Basket Default Swaps, CDO and Factor Copulas*. Journal of Risk, Vol. 7, No. 4, (Summer 2005).

[51] Luciano, E., and W. Schoutens. (2006). *A multivariate Jump-Driven Financial asset Model*. Quantitative Finance 6(5), 385-402.

[52] Luciano, E., and P. Semeraro. (2007). *Extending time-changed Levy asset models through multivariate Subordinators*. Collegio Carlo Alberto Working paper series.

[53] Madan, D. B., and A. Khanna. (2009), *Non Gaussian Models of Dependence in Returns*. Preprints, University of Maryland.

[54] Madan, D. B. (2007). *The Variance Swap Contract*, Notes.

[55] Madan, D. B. (2009). *Pricing and Hedging Basket Options to prespecified levels of Acceptability*. Preprints, University of Maryland.

[56] Madan, D. B. (2010). *Variance Swap Portfolio Theory*. Preprints, University of Maryland.

[57] Madan, D. B., Carr, P., and E. Chang. (1998). *The Variance Gamma Processes and Option Pricing*, European Finance Review 2:79-105, 1998.

[58] Madan, D. B., Konikov, M., and M. Marinescu. (2004). *Credit and Basket default swaps*, working paper.

[59] Madan, D. B., and F. Milne. (1991). *Option pricing with V.G. martingale components*, Mathematical Finance, 1:39-55, 1991.

[60] Madan, D. B., and E. Seneta. (1990). *The Variance Gamma(VG) Model for Share Market Returns*, Journal of Business, 1990, vol. 63, No. 4. pp. 511 - 524.

[61] Madan, D. B., and J-Y J. Yen. (2008). *Asset Allocation with Multivariate Non-Gaussian Returns*, Handbooks in Operations Research and Management Science: Financial Engineering, Vol 15, 949-969, Eds. John R. Birge and Vadim Linetsky, North-Holland, The Netherlands.

[62] Malevergne, Y., and D. Sornette. (2003). *Testing the Gaussian Copula Hypothesis for Financial Asset Dependencies,* Quantitative Finance, 3, 231-250.

[63] Malevergne, Y., and D. Sornette. (2005). *High-Order Moments and Cumulants of Multivariate Weibull Asset Returns Distributions: Analytical Theory and Empirical Tests: II.* Finance Letters, 2005, 3(1), 54-63.

[64] Markowitz, H. M. (1952). *Portfolio Selection*, Journal of Finance, 7, 77-91.

[65] Markowitz, H. M. (1959). *Portfolio selection: efficient diversification of investments*, John Wiley & Sons., New York.

[66] McNeil, A. J., Frey, R., and P. Embrechts. (2005). *Quantitative Risk Management*, Princeton University Press, Princeton.

[67] Meaney, J. *Deling with the volatility smile of Himalayan Options*, 5th Actuarial and Financial Mathematics Day,Universa Press, Flemish Royal Academy of Arts and Sciences for Belgium.

[68] Menn, C., Fabozzi, F. J., and S. T. Rachev. (2005). *Fat Tailed and Skewed Asset Return Distributions: Implications for Risk Management, Portfolio Selection and Option Pricing*, John Wiley and Sons, New York.

[69] Merton, R. C. (1973). *Theory of Rational Option Pricing*. Bell Journal of Economics and Management Science (The RAND Corporation) 4 (1): 141-183.

[70] Overhaus, M. (2002). *Himalayan Options*,Masterclass with Deutsche Bank, 2002.

[71] Quessette, R. (2002). *New products, New risks*, Masterclass with Deutsche Bank, 2002.

[72] Ribeiro C., and N. Webber. (2004). *Valuing path-dependent options in the Variance-Gamma model by Monte Carlo with a gamma bridge*. Journal of Computational Finance, 7:81-100, 2004.

[73] Roll, R. (1984). *A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market*, Journal of Finance, 39, 1127-1139.

[74] Sato, K. (1991). *Self similar processes with independent increments*. Probability Theory and Related Fields, 89, 285-300.

[75] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press: Cambridge, UL, 1999.

[76] Scott, L. (1997). *Pricing Stock options in a jump-diffusion model with stochastic volatility and interest rates: applications of Fourier inversion methods*, Mathematical finance, 7, 413-426.

[77] Schoutens, W., and J. Cariboni. (2009). *Lévy Processes in credit risk*, John Wiley and Sons, New York.

[78] Semeraro, P. (2008). *A multivariate variance gamma model for financial application*. International Journal of Theoretical and Applied Finance, vol 11, 1, 1-18.

[79] Stoll, H. R. (1989). *Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests*, Journal of Finance, 44, 115-134.

[80] Tankov, P. (2003). *Dependence Structures of Specially positive Multidimensional Lévy processes*, working paper.

[81] Tankov, P. (2003). *Simulation and Option Pricing in Levy Copula Models*, preprints.

[82] Wang, J. (2010). *The New Multivariate Variance Gamma Model*, PhD. dissertation, 2010. University of Maryland.