2015

Estimating Sensitivities of Exotic Options Using Monte Carlo Methods

Wei Yuan

FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

ESTIMATING SENSITIVITIES OF EXOTIC OPTIONS USING MONTE CARLO METHODS

By

WEI YUAN

A Dissertation submitted to the
Department of Mathematics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2015

Wei Yuan defended this dissertation on June 9th, 2015.
The members of the supervisory committee were:

Giray Ökten
Professor Co-Directing Dissertation

Kyounghee Kim
Professor Co-Directing Dissertation

Fred Huffer
University Representative

Alec N. Kercheval
Committee Member

Warren D. Nichols
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.

To my parents and husband
for their endless love and encouragement

ACKNOWLEDGMENTS

I would like to thank my advisor, Professor Giray Ökten, for his invaluable guidance throughout my journey as a graduate student in the mathematics department at the Florida State University. His many years of Monte Carlo expertise have allowed him to focus on the financial mathematics of any problem. It has been a unique privilege to be his student.

I would like to thank the members of my dissertation committee for their valuable feedbacks. Specially, thank to my co-advisor Professor Kyounghee Kim, and Professor Alec N. Kercheval for their insightful comments to the contents.

I would like to thank Dr. Dervis Bayazit for his help in my fixed income security study. I regard him as my off-campus PhD advisor.

I would like to thank my most reliable friend Wen Huang for his help with my research as well as five years companionship in Florida. I would like thank Professor Yijun Sun, Dr. Mary Taylor and Will Benish for their encouragement that makes me believe I could excel in my future career.

Finally, I would like to thank my father Yuping Yuan and mother Weijuan Ji, for fostering me and providing me with an excellent, comfortable, and free educational environment. Thanks to my husband Billi Zhu for his continuous love, support, and encouragement. I also want to thank my cousin Xiaolin Zhao and my aunt Weijing Ji, who made me feel that I had a home at Florida.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In this dissertation, methods of estimating the sensitivity of complex exotic options, including options written on multiple assets, and have discontinuous payoffs, are investigated. The calculation of the sensitivities (Greeks) is based on the finite difference method, pathwise method, likelihood ratio method and kernel method, via Monte Carlo or quasi-Monte Carlo simulation. Direct Monte Carlo estimators for various sensitivities of weather derivatives and mountain range options are given. The numerical results show that the pathwise method outperforms other methods when the payoff function is Lipschitz continuous. The kernel method and the central finite difference methods are competitive when the payoff function is discontinuous.

CHAPTER 1

INTRODUCTION

1.1 Option Pricing under the Black-Scholes Model

In 1973, Black and Scholes published their ground-breaking paper "The pricing of options and corporate liabilities" [71]. The article developed a pricing formula for European options, known as the *Black-Scholes formula*. The assumptions behind the formula were: the risk-free interest rate is constant; the underlying asset is a geometric Brownian motion; trading can take place continuously; the market is frictionless (meaning that there are no transaction costs or taxes); short sell is allowed, and the assets are divisible. The Black-Scholes framework is based on the concept of perfect replication of contingent claims. In other words, a European option can be replicated by continuously rebalancing a self-financing portfolio which contains underlying assets and risk-free bonds.

There are two alternative ways of deriving the Black-Scholes formula. This first one is to design a replication strategy. It gives the valuation formulation by solving the Black-Scholes partial differential equation, and the explicit formula for the replication strategy as well. The second method makes direct use of the risk-neutral valuation formula. It focuses on the explicit computation of the arbitrage price of the option, rather than on the derivation of the hedging strategy.

The European call option is a contract whose holder has the right but no obligation to buy one unit of the underlying asset $X(t)$ on the maturity date T at a strike price K. The payoff function is

$$\varphi(X(T)) = \max(X(T) - K, 0). \tag{1.1}$$

Following the Black-Scholes model, the evolution of the underlying asset $X(t)$ assumes the *geometric Brownian motion*, and it is described by the stochastic differential equation (SDE)

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t), \tag{1.2}$$

where $\mu \in \mathbb{R}$ is a constant drift term, $\sigma > 0$ is a constant volatility. Let W_t, $t \in [0, T]$ be a one-dimensional standard Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}(t)$, $t \in [0, T]$ be a filtration for this Brownian motion. The risk-free bond, also known as *cash bond*, or *money market account*, is continuously compounded in value at the rate r; i.e. $B(t) = e^{rt}$.

1.1.1 The Black-Scholes Partial Differential Equation

We consider a portfolio by holding $\phi(t)$ units of the underlying asset and $\psi(t)$ units of cash bond. The value of this portfolio is

$$P(t) = \phi(t)X(t) + \psi(t)B(t). \tag{1.3}$$

It is assumed this portfolio is self-financing in the sense that the changes in its value only come from the variations of the stock price $X(t)$, and the bond price $B(t)$. Alongwith Eqn.(1.2), we have

$$dP(t) = \phi(t)dX(t) + \psi(t)dB(t) = (\phi(t)\mu X(t) + \psi(t)rB(t))dt + \phi(t)\sigma X(t)dW(t). \tag{1.4}$$

The idea is to choose the strategy $(\phi(t), \psi(t))$ in such a way that it replicates the price of a European call option. Let $V(t, x)$ denote the value of this option at time t. An application of Itô's formula to $V(t, x)$ yields

$$dV(x,t) = \left(\frac{\partial V(t,x)}{\partial t} + \mu x\frac{\partial V(t,x)}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V(t,x)}{\partial x^2}\right) dt + \sigma x \frac{\partial V(t,x)}{\partial x} dW(t). \tag{1.5}$$

If we set the initial values of $P(0)$ and the option value $V(0)$ equal, then it must be the case $P(t) = V(t)$ for all t, since V and P have the same dynamics. Then we set the terms on the right-hand side of Eqn.(1.5) equal to the corresponding terms in Eqn.(1.4). This gives:

$$\begin{aligned} \phi(t) &= \frac{\partial V(t,x)}{\partial x}, & (1.6)\\ \psi(t) &= \frac{1}{rB(t)}\left(\frac{\partial V(t,x)}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V(t,x)}{\partial x^2}\right). & (1.7) \end{aligned}$$

Substitution of ϕ and φ Eqn.(1.3) yields the *Black-Scholes partial differential equation*

$$\mathcal{L}_{BS}(\sigma) = \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2} + rx\frac{\partial V}{\partial x} - rV = 0. \tag{1.8}$$

The solution of this partial differential equation with the terminal condition (1.1) is the celebrated *Black-Scholes formula*

$$V(t,x) = x\mathcal{N}(d_1(t,\ x)) - Ke^{-r(T-t)}\mathcal{N}(d_2(t,\ x)), \tag{1.9}$$

where $\mathcal{N}(\cdot)$ denotes the standard cumulative normal distribution function and

$$d_{1,2}(t,x) = \frac{\ln(x/K) + (r \pm \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}.$$

1.1.2 Pricing under the Risk-Neutral Measure

Risk-neutral valuation arises from one key property of the Black-Scholes differential equation. We observe that Eqn.(1.8) does not contain the drift term μ, which depends on risk preferences. Therefore, it will be simple if we assume all the investors are risk neutral while pricing derivatives. The reason is that risk-free rates are known, while the expected return of an asset in the real world should be the risk-free rate plus some risk premium, which is not observable. We construct the risk-neutral measure Q by Girsanov theorem (see Appendix A 31), which is equivalent to the measure P in the real world; i.e. the physical measure. We apply Itô's formula to the discounted asset price $\tilde{X}(t) = e^{-rt}X(t)$,

$$d\tilde{X}(t) = (\mu - r)\tilde{X}(t)dt + \sigma\tilde{X}(t)dW(t) = \sigma\tilde{X}(t)\left(dW(t) + \left(\frac{\mu - r}{\sigma}\right)dt\right). \tag{1.10}$$

The drift term under the P measure can be removed by introducing

$$\tilde{W}(t) = W(t) + \int_0^t \frac{\mu - r}{\sigma} du, \tag{1.11}$$

and the *Radon-Nikodym derivative* process (see Appendix A, Definition 33)

$$Z(t) = \exp\left(-\Theta W(t) - \frac{1}{2}\Theta^2 t\right), \tag{1.12}$$

where $\Theta = \frac{\mu - r}{\sigma}$, which we call the market price of risk. Alongwith the Brownian motion (1.11), we rewrite Eqn.(1.10) as $d\tilde{X}(t) = \sigma\tilde{X}(t)\tilde{W}(t)$. The risk-neutral measure Q is defined via Girsanov's theorem as below:

$$\frac{dQ}{dP} = \exp\left(-\Theta W(T) - \frac{1}{2}\Theta^2 T\right).$$

Hence, the discounted underlying asset $\tilde{X}(t)$ is a martingale under the measure Q. We can also show that the discounted value of portfolio $\tilde{P}(t) = e^{-rt}P(t)$ is a martingale by applying Ito's formula to Eqn.(1.3) and the self-financing property Eqn.(1.4). The martingale property of $\tilde{P}(t)$ implies that

$$e^{-rt}(t)P(t) = E_Q[e^{-rT}P(T)|\mathcal{F}(t)].$$

The martingale representation theorem tells us that a short hedge exists and $P(t) = V(t)$ for all t. Therefore the discounted option value $e^{-rt}V(t)$ is a martingale as well. We have

$$e^{-rt}V(t) = E_Q[e^{-rT}V(T)|\mathcal{F}(t)].$$

The *risk-neutral pricing formula* is

$$V(t) = e^{-r(T-t)} E_Q[(X(T) - K)^+ | \mathcal{F}(t)]. \tag{1.13}$$

Hull [47] observed that, in a world where investors are risk-neutral, the expected return on all investment assets is the risk-free rate, because investors in this world do not require a risk premium to compensate them to take risks. The solutions obtained via risk-neutral formula are valid in all worlds of risk-preference, not just those where investors are risk-neutral. When we move from a risk-neutral world to a risk-averse world, two things happen. The expected growth rate in the stock price changes and the discount rate that must be used for any payoff from the derivative changes. These two changes always exactly offset each other.

1.2 Monte Carlo and Quasi-Monte Carlo Methods

Hertz [44] first introduced the Monte Carlo methods to finance in 1964 in the application of corporate finance. Later in 1977, Monte Carlo simulation was applied to asset pricing by Boyle in his seminal paper [9]. Since then remarkable progress has been achieved in a wide range of financial problems, making the Monte Carlo approach an essential tool in the pricing of derivative securities and in risk management. There is an extensive literature on the use of Monte Carlo Methods in financial problems. Boyle et al. [10] give an overview of developments in the use of simulation for security pricing. Different approaches to value American type options by Monte Carlo simulation are proposed by Tilley [79], Broadie and Glasserman [14], and Longstaff and Schwartz [53]. A survey of current state of the art is given by the monograph by Glasserman [31].

Monte Carlo simulation is an attractive method for asset pricing because of the generality in the types of assets it can handle, and the ease with which it handles multiple state variables or path dependent problems. The idea is that the arbitrage-free price of a derivative security can be formulated in terms of the expectation of some random variables; then valuing the derivative reduces to computing an expectation. In many cases, when expressing the expectation as an integral, we find the dimension of the integral to be large. This feature makes the Monte Carlo method very attractive, because the error convergence rate of the Monte Carlo method is independent of the dimension of the problem. This is the dominant advantage of the method over other classical deterministic numerical integration methods. Furthermore, the Monte Carlo method is flexible and easy to implement and modify. It is not always possible to derive analytical formulas for option prices and their Greeks, either because the payout or the underlying model, or both, are too complicated to be analytically tractable.

The major disadvantage of the Monte Carlo method is its slow convergence rate $O(1/\sqrt{n})$, where n is the number of random samples. When dealing with complex problems, huge computational effort in sample replications is required to get sufficient accuracy. In order to improve the computational efficiency of Monte Carlo simulations, Boyle et al. [10] suggest using classical variance reduction techniques, such as control variate, antithetic variate, moment matching, important sampling, etc. Another technique to improve the convergence rate replaces random sequences with some deterministic sequences, which are known as low-discrepancy sequences or quasi-random sequences. Niederreiter [57] and Tezuka [78] provide in-depth analysis of low-discrepancy sequences. Moskowitz and Caflisch [55] discuss and suggest some modifications in improving the convergence of quasi-Monte Carlo methods. The application of low-discrepancy sequences in finance appear in Birge [7], Joy et al. [50] and Boyle et.al. [10]. We will explain the quasi-Monte Carlo technique in Section 1.2.2 briefly.

1.2.1 Monte Carlo Method

We consider the problem of estimating the multi-dimensional integral of a Riemann integrable function. Let X be a multivariate random variable uniformly distributed over the d-dimensional half-open unit cube $C_d = [0,1)^d$. For a function $f \in L^2(0,1)$, consider the problem of estimating

$$E[f(X)] = \int_{C_d} f(x)dx. \tag{1.14}$$

By drawing n random samples X_i, for $i = 1, ..., n$, uniformly distributed on C_d, we obtain the unbiased Monte Carlo estimator of the integral (1.14), i.e. the sample mean of function f, and we denote it as $\hat{E}\ [f(X)]$:

$$\hat{E}\ [f(X)] = \frac{1}{n}\sum_{i=1}^{n} f(X_i). \tag{1.15}$$

From the strong law of large numbers

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(X_i) = \int_{C_d} f(x)dx$$

almost surely. According to the central limit theorem, the error $e_n = \hat{E}[f(X)] - E[f(X)]$ is approximately normally distributed, with mean 0 and standard deviation $\sigma_f/\sqrt{n}$, where

$$\sigma_f^2 = \int_{C_d} (f(x) - E[f(x)])^2 dx.$$

In practice, σ_f is unknown and we use its sample standard deviation s_f for proxy. It is calculated through

$$s_f = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \Big(f(X_i) - \hat{E}[f(X)] \Big)^2}. \tag{1.16}$$

From Eqn.(1.16), we observe that the convergence rate of Monte Carlo method is $O(n^{-1/2})$ which is independent of the dimension d.

1.2.2 Quasi-Monte Carlo Method

Quasi-Monte Carlo methods (QMC) are described as the deterministic version of the Monte Carlo methods. This approach uses the low-discrepancy sequences, which are designed to be more evenly dispersed throughout the region of integration than pseudorandom sequences. It provides deterministic convergence, as opposed to the stochastic convergence of the Monte Carlo simulation.

The idea of QMC is straightforward. Suppose we wish to integrate function $f(x)$ over $C_d = [0,1)^d$ using a sequence of n points. Instead of picking pseudorandom numbers, we choose a deterministic low-discrepancy sequence $u_1, u_2, ..., u_n$. The integral is approximated by

$$\int_{C_d} f(x)dx \approx \frac{1}{n} \sum_{i=1}^{n} f(u_i). \tag{1.17}$$

We next define the discrepancy, which is a measure of goodness for low-discrepancy sequences.

Definition 1 *Consider a sequence of* n *points* $\{x_1, x_2, ..., x_n\}$ *in the* d*-dimensional unit cube* $C_d = [0,1)^d$ *and let* J *be a subset of the form* $\prod_{i=1}^{d}[0, u_i)$*, and* $u_i \in C_d$*. The star-discrepancy of the sequence is defined as*

$$D_n^* = \sup_{J \in C_d} \left| \frac{A(J;n)}{n} - V(J) \right|,$$

where $A(J;n)$ *is the number of elements of the sequence in the subset* J *and* $V(J)$ *is the volume of* J*.*

There are many examples of low-discrepancy sequences, for example, Halton sequence [40], Sobol sequence [73], Faure sequence [25], and Niederreiter sequence [56]. The star discrepancy of all these sequences has the order $O((\log n)^d/n)$. It is customary to reserve the informal term "low-discrepancy" for the methods that achieve a star discrepancy of $O((\log n)^d/n)$.

The approximation error for the evaluation of an integral using the quasi-Monte Carlo method is established via Koksma-Hlawka inequality.

Theorem 2 *(Koksma-Hlawka) If f has bounded variation $V(f)$ on $[0,1]^d$ in the sense of Hardy and Krause, then for any $x_1, \dots,\ x_n \in C_d$, we have*

$$\left| \frac{1}{n} \sum_{i=1}^{n} f(x_i) - \int_{C_d} f(x)dx \right| \leq V(f) D_n^*, \tag{1.18}$$

where D_n^* is the star discrepancy of the points $x_1, \dots,\ x_n$.

Generally, the error bound in the Koksma-Hlawka inequality is of little practical value, since it is very hard to evaluate the variation of f in the sense of Hardy and Krause (see [57] for a complete discussion of the Hardy-Krause definition of variation). However this bound allows a separation between the regularity properties of the integrand and the degree of uniformity of the sequence. We can guarantee a reasonable approximation for any function f with bounded total variation $V(f)$ by ensuring the star discrepancy of the sequence D_n^* is small.

1.3 Sensitivities (Greeks)

When taking an option position or setting up an option strategy, there will be risk from many dimensions, such as a change in the underlying's price, a change in volatility or a change in time value decay. Each dimensional risk can be measured or represented by a sensitivity of the option price. These sensitivities are commonly referred to as the Greeks, since most of them are labelled by the Greek letters. The Greeks play a critical role in trading strategy and the traders use them to manage all risks within acceptable levels.

When a strategy is constructed, there are associated delta, vega and theta positions, as well as other position Greeks. By calculating position Greeks, we can know how much risk and potential reward reside in the strategy, whether it is long put or call, or a complex strategy like a strangle, butterfly spread, or ratio spread, among many others. If you can match your correct outlook on a market to the position Greeks in a strategy, you capitalize on favorable changes in the strategy at every level of the Greeks. That is why Greeks are important in finance. Greeks can be incorporated into strategy design at a precise level using mathematical modeling skills and sophisticated softwares. But at a more basic level, the Greeks can be used as guideposts for where the risks and rewards can generally be found. In the following, we will introduce the most common Greeks as well as their use in the pricing and hedging of derivatives.

Let V denote the value of an option written on a single underlying asset S. The volatility of the underlying asset is σ. The maturity is T and the risk-free interest rate is r. The correlation coefficient between two different asset i and asset j is $\rho_{i,j}$. We introduce six essential Greeks as follows:

$$\begin{aligned}
delta\ \Delta &:= \frac{\partial V}{\partial S} \\
gamma\ \Gamma &:= \frac{\partial^2 V}{\partial S^2} \\
vega\ \mathcal{V} &:= \frac{\partial V}{\partial \sigma} \\
theta\ \Theta &:= -\frac{\partial V}{\partial T} \\
rho\ \rho &:= \frac{\partial V}{\partial r} \\
corr.delta\ \Delta_{i,j} &:= \frac{\partial V}{\partial \rho_{i,j}}
\end{aligned}$$

1.3.1 Delta

Delta measures the sensitivity of the value of an option to changes in the price of an underlying asset. If the movement of the underlying asset is small, the price of the option will move by delta times the movement of the underlying asset. As convention, delta can also be shown as a value between -100 to 100 to show the total dollar sensitivity on the value of one option, which comprises of 100 shares of the underlying. For example, if you own a call with a delta 0.3, you should gain \$30 if the stock price goes up by \$1. Under the Black-Scholes model, the interpretation of delta is the probability that the option will expire in the money. We could see this simply by taking the derivative of the Black-Scholes formula (1.9) with respect to x and then obtain the delta $\Delta = \mathcal{N}(d_1)$.

Delta is the most important Greek for hedging. At the outright position, it is positive for a call option and negative for a put option. This reflects the fact that call options increase and put options decrease in value, when the underlying asset price increases. The position delta can be understood through the idea hedge ratio. As we have seen in the derivation of the Black-Scholes formula, delta is a hedge ratio because it tells us how many options contracts are needed to hedge a long or short position in the underlying. By changing the ratio of calls to number of positions in the underlying, we can turn this position delta either positive or negative.

1.3.2 Gamma

Since delta is an important factor, option traders are interested in how delta moves as stock prices go up or down. Gamma measures the rate of change in delta for each one point move of the underlying asset. Therefore gamma can help to forecast changes in delta. For example, a 105 at-the-money put option with -10 delta and 4 gamma, you will find the delta of 110 put will be around -30, i.e $-10 - 4 \times 5 = 30$. For the delta-hedged portfolio, gamma also measures how much or how often a position must be re-balanced to maintain a delta-neutral position. If gamma is small, which means delta changes slowly, then we do not need to adjust frequently to keep delta neutral.

For two options with the same value for delta, the one with higher gamma will have a higher risk because of adverse move of the underlying asset price. Gamma always reaches its maximum for at the money options, and get progressively lower for both in-the-money and out-of-money options. Unlike delta, gamma is always positive for both call and puts, for the underlying prices change in the same direction as delta. All net selling strategies will have negative position of gamma and net buying strategy will have net positive gamma.

Cross gamma is the sensitivity of a multi-asset option to the changes in two of the underlying assets. Let us assume that an option is written on more than one underlying asset, then the cross gamma with respect to the assets S_i and S_j is given by $\Gamma_{i,j} = \dfrac{\partial V^2}{\partial S_i \partial S_j}$. It can be thought of as sensitivity of delta in S_i to the price of asset S_j. Greek gamma of a multi-asset option could be expressed in a matrix form, with elements on the diagonal being gamma and elements off the diagonal being cross gamma. In multi-asset options, it is possible that the delta with respect to one asset can be affected by a movement in another asset, even if the first asset stays as a constant.

1.3.3 Vega

The Greek vega measures the risk of gain or loss resulting from changes in the volatility. For all options, call or put, vega is always positive because options increase in value when volatility increases and decreases when volatility declines. Usually, the higher time premium values are associated with higher vega value. Also, the options with more distant expiration date have higher vega, for example long-term equity anticipation securities (LEAPS) options.

Some strategies are long volatility and others are short volatility, while some can be constructed to be neutral volatility. For example, a put that is purchased is long volatility, which means the value increases when volatility increases and falls when volatility drops. Conversely, a put that is sold is short volatility. If the volatility risk has been neutralized, vega is around zero.

1.3.4 Theta

Theta is a measure of the rate of time premium decay. Theta values are always negative for options because options are always losing time value with each tick of the clock until expiration is reached. In fact, all options, regardless of strikes or markets, will always have zero time value at expiration. For example, if theta of a call option equals to -10, then the option is losing $10 in the time value every day. Theta will have wiped out all time value leaving the option with the intrinsic value. The intrinsic value will represent to what extent the option expires in the money.

1.3.5 Rho

Rho is the Greek letter used to measure the risk related to changes in interest rates. This Greek will be larger for long-dated options and negligible for short-dated ones. Strategists who use LEAPS should take into account rho since over longer time frames the interest rate share of an option's value is more significant. However, when interest rates in an economy are relatively stable, the chance that the value of an option position will change dramatically because of a drop or rise in interest rates is pretty low. Therefore, the interest rate generally doesn't play a role in typical strategy designs and option pricing. A simply rule about the rho is that when interest rates rise, call prices will rise and put prices will fall. The reverse occurs when interest rates fall.

1.3.6 Correlation Delta

Correlation delta is the first-order derivative of the price of a multi-asset option to changes in the correlation coefficient between two underlying assets. The correlation delta with respect to the asset S_i and the asset S_j is given by the derivative $\frac{\partial V}{\partial \rho_{i,j}}$. In the contest of pricing derivatives with multi-assets under the Black-Scholes setting, we measure the degree of the dependence or dispersion of multi-assets through the correlation matrix. Embrechts et al.[23] pointed out that the dependence structure of an elliptical distribution, such as the multivariate normal distribution or the multivariate log-normal distribution, is fully described by its correlation coefficients and covariance.

1.4 Estimating Greeks using Monte Carlo Methods

In general, Monte Carlo methods for estimating Greeks can be categorized into two classes. The first class refers to the *finite difference approximation*. It is easy to implement, but requires simulating at multiple values and produces a biased estimator with large mean squared error. (See Glasserman [31] for general background on this method). Methods in the second category produce unbiased estimators and they are sometimes called "direct Monte Carlo method". Two well-known methods in the second category are the *pathwise method* and the *likelihood ratio method*. In brief, the pathwise method differentiates the payoff function with respect to the parameter of interest. The likelihood ratio method differentiates the density function of the underlying asset with respect to the parameter rather than the payoff. Broadie and Glasserman [13] discuss these two methods in depth and derive estimators for the sensitivities of financial derivatives.

The pathwise method stems from the infinitesimal perturbation analysis in the traditional simulation literature (see Ho and Cao [18], Suri and Zazanis [76]). There is an extensive literature on the pathwise derivative estimation. Fu and Chen [20] estimate the sensitivities in application of the mortgage backed securities using the pathwise method. Glasserman and Zhao [33] investigate using the pathwise method to estimate deltas of caplets.

The pathwise method requires the discounted payoff function to be Lipschitz continuous. This condition is restrictive and not satisfied for most of the exotic options. Even for vanilla options, applying the pathwise method directly to gamma estimation is not possible. Nevertheless, with tailored modification, the pathwise method could circumvent the difficulty of discontinuity and be applicable widely. Fu and Hu [29] smooth the payoff function by conditioning on certain random variables. This method requires finding the appropriate random variables on which to condition, and it may be difficult to implement in practice. The pathwise method can also be applied to American type options. Chen and Liu [22] introduce a generalized pathwise method to resolve the difficulty caused by discontinuity of the optimal decision with respect to the underlying parameter.

The likelihood ratio method (LRM) differentiates the density function of the stochastic model for the underlying asset rather than the payoff function. The method was first proposed to study discrete-event systems and was called score function method (see Glynn [34], Reiman and Weiss [64], Rubinstein [68], Rubinstein and Shapiro [69]). Glasserman and Liu [32] apply saddle approximation to the score function of increments for estimating Greeks in Lévy driven models. In contrast to the pathwise method, the likelihood ratio does not require any smoothness in the discounted payoff because it is based on differentiation of probability densities. This property makes the LRM potentially attractive in exactly the settings in which the pathwise method fails. The application of the LRM requires explicit knowledge of the relevant probability densities, and as reported by Glasserman [31] its estimates often have a larger variance than the pathwise method.

Recently, other methods for estimating Greeks have been proposed. Wang and Caflisch [80] propose an approach called "*modified least-squares Monte Carlo method*" based on the celebrated Longstaff-Schwartz regression method, which can be viewed as a variation of the finite-difference approximation to option sensitivities. This method works well for estimating delta and gamma for the American type of options with one underlying asset. But it fails for high dimensional problems with multiple underlyings, and also fails in estimating Greeks expect for delta and gamma. Liu and Hong [52] propose the kernel estimator method for Greeks. They convert Greeks into the sum of an ordinary expectation and a derivative with respect to an auxiliary parameter and then estimate the derivative with kernel estimator. Fournié et al. [26] propose to use Malliavin calculus to estimate Greeks. This method first derives closed-form expressions of Greeks as integrations that can be estimated through simulations under appropriate discretization schemes. Gobet and Kohatsu-Higa [36] apply it to compute the Greeks for barrier and lookback options. Chen and Glasserman [21] show that the Malliavin calculus method can be viewed as a combination of the pathwise method and likelihood ratio method. Bayazit [3] presents sensitivities when the underlying follows an exponential Lévy process based on a finite-dimensional Malliavin calculus and finite difference methods via Monte-Carlo simulations.

1.4.1 Finite Difference Approximation

Let $V(\theta)$ be the value of an option with the discounted payoff function $\varphi(\theta)$, where θ is a parameter of interest. The parameter of interest could be the initial stock price $S(0)$, the volatility σ, the interest rate r, etc. The finite difference approximation is a way to estimate $V'(\theta)$, the derivative of $V(\theta)$ with respect to θ. The two most common finite difference estimators are the forward difference estimator and the central difference estimator:

- **Forward difference estimator:** We simulate n independent replications $\varphi_1(\theta), \ldots, \varphi_n(\theta)$ at the parameter θ and n additional independent replications $\varphi_1(\theta+h), \ldots, \varphi_n(\theta+h)$ for some small positive value of h. The two sets of replications might be independent or not. Let $\overline{\varphi}_n(\theta)$ and $\overline{\varphi}_n(\theta+h)$ be the arithmetic averages of each set of replications. The forward difference estimator is given by
$$\frac{\overline{\varphi}_n(\theta+h) - \overline{\varphi}_n(\theta)}{h}. \tag{1.19}$$

- **Central difference estimator:** The central difference estimator is given by
$$\frac{\overline{\varphi}_n(\theta+h) - \overline{\varphi}_n(\theta-h)}{2h}. \tag{1.20}$$

- **Second order Greeks:** The cental difference estimator for the second order Greeks is
$$\frac{\overline{\varphi}_n(\theta+h) - \overline{\varphi}_n(\theta) + \overline{\varphi}_n(\theta-h)}{h^2}. \tag{1.21}$$

Finite difference approximations are easy to implement. However, these estimators suffer from some disadvantages:

- First, the finite difference approximation requires twice as many simulations as we conduct at a single point.

- Secondly, the finite difference approximation should be expected to converge to the true value with h small enough. The finite difference estimator is heavily biased if h is large unless the function $V(\theta)$ is close to linear. For the case of the central difference estimator, if the function $V(\theta)$ is smooth enough, we can show that the bias has order $O(h^2)$ by Taylor series expansion:
$$\begin{aligned}
& E\left[\frac{\overline{\varphi}_n(\theta+h) - \overline{\varphi}_n(\theta-h)}{2h}\right] \\
= \;& \frac{V(\theta+h) - V(\theta-h)}{2h} \\
= \;& \frac{1}{2h}(V(\theta) + V'(\theta)h + \frac{1}{2}V''(\theta)h^2 - (V(\theta) - V'(\theta)h + \frac{1}{2}m''(\theta)h^2) + O(h^3)) \\
= \;& V'(\theta) + O(h^2).
\end{aligned}$$

- On the other hand, a small value of h gives a large variance. For example, when $h \to 0$, the variance of the central difference estimator goes to infinity.
$$var\left(\frac{\overline{\varphi}_n(\theta+h) - \overline{\varphi}_n(\theta-h)}{2h}\right) = \frac{1}{4h^2}(var(\overline{\varphi}_n V(\theta+h)) + var(\overline{\varphi}_n(\theta-h))) \to \infty.$$

Although using the same set of random numbers in re-simulation can reduce the error, the variance cannot be wiped out completely, even if we apply variance reduction techniques. In order to balance these two errors, we need to find the optimal relation between them to minimize the mean square error. This tradeoff is analyzed by Glynn [35], Fox and Glynn [27], Zazanis [82], and in early work by Frolov and Chentsov [28].

1.4.2 Pathwise Method

The pathwise method differentiates each simulated outcome with respect to the parameter of interest. We start by defining a probability space $(\Omega, \mathcal{F}, Q)$, where $\mathcal{F}$ is a σ-algebra of events and Q is the risk-neutral probability measure defined on this probability space. We write the underlying asset price as $S(\theta, \omega) : \Theta \times \Omega \to R_+^n$ representing the n-dimensional underlying asset process with the parameter $\theta \in \Theta$, which affects the dynamics of S. For each fixed $\omega \in \Omega$, we take the derivative of $S(\theta)$ with respect to θ. If this derivative exists with probability 1, we call $S'(\theta)$ the *pathwise derivative* of S at θ. In practice, we take each ω as a random sample, and each $S(\theta, \omega)$ could be viewed as a realization or simulated outcome with the random sample ω. In application to option pricing, we consider the discounted payoff function $\varphi : \mathrm{R}_+^n \to \mathrm{R}$, and we are interested in obtaining an unbiased estimator for

$$\frac{d}{d\theta} E_Q[\varphi(S(\theta, \omega))]. \tag{1.22}$$

If the order of expectation and differentiation can be changed, we have

$$\begin{aligned} \frac{d}{d\theta} E[\varphi(S(\theta, \omega))] &= E_Q\left[\lim_{h\to 0} \frac{\varphi(S(\theta+h, \omega)) - \varphi(S(\theta, \omega))}{h}\right] \\ &= E_Q\left[\varphi'(S(\theta, \omega)) \frac{d}{d\theta} S(\theta, \omega)\right]. \end{aligned} \tag{1.23}$$

Then the pathwise estimator is

$$\varphi'(S(\theta,\ \omega)) \frac{d}{d\theta} S(\theta,\ \omega). \tag{1.24}$$

Conditions of unbiasedness for the pathwise method. Broadie and Glasserman [13] establish the conditions to interchange the order of the differentiation and expectation. Suppose that the discounted payoff is a function of a random vector

$$X(\theta) = (X_1(\theta),\ X_2(\theta),\ \ldots,\ X_m(\theta)), \tag{1.25}$$

which is a function of the parameter θ. Thus

$$Y(\theta) = f(X_1(\theta),\ X_2(\theta),\ \ldots,\ X_m(\theta))$$

for some function $f : R^m \to R$ depending on the specific derivative security. We require:

(A1) At each $\theta \in \Theta,\ X_i'(\theta)$ exists with probability 1, for all $i = 1,\ \ldots,\ m$.

(A2) Let $D_f \subseteq R^m$ denote the set of points at which f is differentiable and require $P(X(\theta) \in D_f) = 1$ for all $\theta \in \Theta$. Then we have:

$$Y'(\theta) = \sum_{i=1}^{m} \frac{\partial f(X(\theta))}{\partial X_i} X_i'(\theta) \tag{1.26}$$

exists with probability 1.

(A3) We require f to be Lipschitz continuous, i.e., there exists a constant k_f such that for all $x,\ y \in R^m$

$$|f(x) - f(y)| \leq k_f \|x - y\|. \tag{1.27}$$

(A4) There exist random variables $\kappa_i, i = 1 \ldots, m$, such that for all $\theta_1, \theta_2 \in \Theta$,

$$|X_i(\theta_2) - X_i(\theta_1)| \leq \kappa_i |\theta_2 - \theta_1| \tag{1.28}$$

with $E[\kappa_i] < \infty,\ i = 1,\ \ldots m$.

Proof. The key result used in the proof of the unbiasedness is the Lebesque dominated convergence theorem that allows for the exchange of limit and expectation operator required in Eqn.(1.23):

Theorem 3 ***Lebesgue's Dominated Convergence Theorem*** *[6] : If $|f_n| \leq g$ almost everywhere, where g is integrable, and if $f_n \to f$ almost everywhere, then f and f_n are integrable and $\int f_n d\mu \to \int f d\mu$.*

From (A1) and (A2), we know that the function $Y(\theta)$ is continuous and differentiable at all $\theta \in \Theta$ with the probability 1. Conditions (A3) together with (A4) imply that $Y(\theta)$ is almost surely Lipschitz in θ because the Lipschitz property is preserved by composition (See [24] for details). This continuity of $Y(\theta)$ satisfies the boundness condition required for dominated convergence theorem.

$$\begin{aligned} |Y(\theta_2) - Y(\theta_1)| &= |f(X_1(\theta_2), ..., X_m(\theta_2)) - f(X_1(\theta_1), ..., X_m(\theta_1))| \\ &\leq \kappa_f \|X(\theta_2) - X(\theta_1)\| \\ &\leq \kappa_f \sum_{i=1}^{m} \kappa_i |\theta_2 - \theta_1| \end{aligned}$$

and for κ_Y we may take

$$\kappa_Y = \kappa_f \sum_{i=1}^{m} \kappa_i. \tag{1.29}$$

It follows that $E[\kappa_Y] < \infty$. Observing that

$$\left| \frac{Y(\theta + \Delta\theta) - Y(\theta)}{\Delta\theta} \right| \leq \kappa_Y, \quad \textit{almost surely}. \tag{1.30}$$

Therefore, we can interchange expectation and the limit according to the dominated convergence theorem. In summary, the four conditions (A1-A4) suffice to ensure the pathwise derivative is an unbiased estimator.

■

1.4.3 Likelihood Ratio Method

The likelihood ratio method differentiates the density function of the stochastic model of the underlying asset rather than the payoff function. This method does not require smoothness in the discounted payoff and thus complement the pathwise method. We first express the discounted payoff $\varphi(X)$ as a function of a random vector $X = (X_1, \ \ldots, \ X_n)$ with a known joint density function $f_\theta(X)$, where θ is a parameter of this density. Then the price of the option is written as

$$V(\theta, X) = E[\varphi(X)] = \int_{R^n} \varphi(X) f_\theta(x) dx. \tag{1.31}$$

We take the derivative of the option price with respect to θ, and switch the integral and derivative, if possible, and obtain

$$\frac{\partial}{\partial\theta} E\left[\varphi(X)\right] = \int_{R^n} \varphi(x) \frac{\partial}{\partial\theta} f_\theta(x) dx = \int_{R^n} \varphi(x) \frac{\dot{f}_\theta(x)}{f_\theta(x)} f_\theta(x) dx = E\left[\varphi(X) \frac{\dot{f}_\theta(X)}{f_\theta(X)}\right]. \tag{1.32}$$

From the last expression of Eqn.(1.32), we obtain the unbiased likelihood ratio estimator

$$\varphi(X) \frac{\dot{f}_\theta(X)}{f_\theta(X)}, \tag{1.33}$$

where $\dot{f}_\theta(X)/f_\theta(X)$ is called the score function in the statistical literature.

As with the pathwise method, the validity of the likelihood ratio method relies on an interchange of differentiation and integration. However, the interchange is relatively benign because the probability densities are typically smooth functions of their parameters. The main limitation in the application of this method is that it requires the knowledge of the density, which may not be available when the stochastic model of the underlying asset is complex.

1.4.4 Kernel Method

Liu and Hong [52] extend the pathwise method to options with a certain type of discontinuous payoff. They convert the Greek into the sum of an ordinary expectation and a derivative with respect to an auxiliary parameter and then estimate the expectation and derivative by a sample mean and a kernel estimator, respectively. This method requires little analytical effort and it is very easy to implement. The kernel method has been studied extensively in the literature of nonparametric regression (see Bosq [8] for a comprehensive overview). This method is applicable when the payoff function could be written as

$$\varphi(S) = \mathrm{g(S)} \cdot 1_{\{h(S)\geq 0\}}, \tag{1.34}$$

where $g(S)$ and $h(S)$ are functions that satisfy the following two assumptions (see Theorem 2 in [52]).

Assumption 1 *For any $\theta \in \Theta$, where Θ is an open set, $g(S)$ and $h(S)$ are differentiable with respect to θ with probability 1 and there exist random variables K_g and K_h with finite second moment that may depend on θ, such that $|g(S(\theta+\triangle\theta))-g(S(\theta))| \leq K_g|\triangle\theta|$, and $|h(S(\theta+\triangle\theta))-h(S(\theta))| \leq K_h|\triangle\theta|$ when $|\triangle\theta|$ is small enough.*

Assumption 2 *For any $\theta \in \Theta$, $\partial_\theta\phi(\theta,\ y)$ exists and is continuous at $(\theta, 0)$, where $\phi(\theta, y) = E_Q[g(S)1_{\{h(S)\geq y\}}]$.*

Theorem 4 *Suppose that $E[|g(s)|]^2 < \infty$ and $E[|h(s)|]^2 < \infty$. If the two assumptions above are satisfied, then we have*

$$V'(\theta) = E[\partial_\theta g(S) \cdot 1_{\{h(S)\geq 0\}}] - \partial_y E\left[g(S)\partial_\theta h(S) \cdot 1_{\{h(S)\geq y\}}\right]\Big|_{y=0}. \tag{1.35}$$

Assumption 1 is a typical assumption in the pathwise sensitivity estimation and it is satisfied in practice for most cases. For example, when $S_i(t)$ follows a geometric Brownian motion, $S_i(t)$ is continuously differentiable and $\|S_{\theta+\Delta\theta} - S_\theta\| \leq K_s|\theta|$ for some random variable K_s. (For a proof, see Broadie and Glasserman [13]). Therefore, to satisfy Assumption 1, we need that $g(S)$ and $h(S)$ are Lipschitz continuous and differentiable almost everywhere. Assumption 2 is generally difficult to prove. Liu and Hong [52] propose the sufficient condition (Lemma 5) of Assumption 2, which is easy to verify.

Lemma 5 *Let $f_{S(\theta)}(s)$ denote the density of $S(\theta)$. Suppose for any $\theta \in \Theta$, $f_{S(\theta)}(s)$ is continuously differentiable in θ for almost all $s \in \mathbb{R}^k$, and there exists a function $\beta(s)$ such that $|f_{S(\theta+\Delta\theta)}(s) - f_{S(\theta)}(s)| \leq \beta(s)|\Delta\theta|$, when $|\Delta\theta|$ is small enough, and $\int_{\mathbb{R}^k} |g(s)|\beta(s)ds < +\infty$. Then, Assumption 2 is satisfied.*

Note that the condition of Lemma 5 are typical conditions for the likelihood ratio method to be applicable to estimating $dE[g(S)]/d\theta$ (see Asmussen and Glynn [2]). As pointed out by Broadie and Glasserman [13], these conditions are typically satisfied in the context of estimating the Greeks. Therefore, Lemma 5 shows that Assumption 2 is typically satisfied in practice.

We observe that Eqn.(1.35) converts $V'(\theta)$, the derivative with respect to θ, to a sum of ordinary expectation and a derivative with respect to y, where y is an auxiliary parameter and not used in the sample path simulation. Therefore, it is unnecessary to run simulations to estimate this derivative. In a simple way, the second term on the right-hand side of Eqn.(1.35) can be estimated by the finite difference method

$$\begin{aligned}
& \partial_y E[g(S)\partial_\theta h(S) \cdot 1_{\{h(S)\geq y\}}]|_{y=0} \\
= & \lim_{\delta\to 0} \frac{1}{\delta}\left(E\left[g(S)\partial_\theta h(S) \cdot 1_{\{h(S)\geq \frac{\delta}{2}\}}\right] - E\left[g(S)\partial_\theta h(S) 1_{\{h(S)\geq \frac{\delta}{2}\}}\right]\right) \\
= & -\lim_{\delta\to 0} \frac{1}{\delta} E\left[g(S)\partial_\theta h(S) \cdot 1_{\{-\frac{\delta}{2}\leq h(S)\leq \frac{\delta}{2}\}}\right].
\end{aligned}$$

This motivates the use of the kernel method, because it is a generalization of the finite difference method. A kernel Z is a symmetric density such that $\int_{-\infty}^{+\infty} u^2 Z(u)du < \infty$ and $|u| \to \infty$ as $uZ(u) \to 0$. For example, we can set $Z(u) = 1_{\{-0.5\leq u\leq 0.5\}}$. It is the density of the uniform distribution and is known as the uniform kernel. As shown in [52], we have the following identity

$$\lim_{\delta\to 0} \frac{1}{\delta} E\left[g(S)\partial_\theta h(S) \cdot 1_{\{-\frac{\delta}{2}\leq h(S)\leq \frac{\delta}{2}\}}\right] = \lim_{\delta\to 0} \frac{1}{\delta} E\left[g(S)\partial_\theta h(S) \cdot Z\left(\frac{h(S)}{\delta}\right)\right]. \tag{1.36}$$

The smooth kernels, such as the standard normal density, are widely used as well, since they give more robust estimators. We use the Monte Carlo method to estimate the expectation in the right-hand side of Eqn.(1.35) and obtain the estimator

$$\overline{G}_n = \frac{1}{n}\sum_{l=1}^{n} g_l' \cdot 1_{\{h_l\geq 0\}} + \frac{1}{n\delta_n}\sum_{l=1}^{n} g_l \cdot h_l' \cdot Z\left(\frac{h_l}{\delta_n}\right), \tag{1.37}$$

where δ_n is called the bandwidth parameter in the kernel estimation and (g_l, h_l, g_l', h_l'), with $l = 1, \ldots, n$, denotes the lth independent observation of $(g(S), h(S), \partial_\theta g(S), \partial_\theta h(S))$.

The asymptotic properties of $\overline{G}_n$ are well established in the literature. In [52], it is proved that $\overline{G}_n$ is a consistent estimator and follows a central limit theorem if $\delta_n \to 0$ and $n\delta_n \to \infty$ as $n \to \infty$ and the optimal rate of convergence of the estimator is $n^{-2/5}$ and it is achieved when $\delta_n = cn^{1/5}$ for some constant $c > 0$. The bandwidth δ_n is chosen from a pilot simulation run.

1.4.5 Modified Least Squares Monte Carlo Algorithm

The least squares Monte Carlo (LSM) algorithm is a well-known algorithm for pricing American options introduced by Longstaff and Schwartz [53]. Wang and Caflisch [80] propose that by a simple modification of the algorithm, (called modified least squares monte carlo method(MLSM)), one can not only price the option, but also two derivatives of the option: delta and gamma. These are the first and second derivatives of the payoff function with respect to the initial asset price.

At any exercise time, the option holder compares the payoff from immediate exercise with the expected payoff from not exercising it. If the underlying asset is a Markov process, the continuation expressed as a conditional expectation could be approximated by polynomials of underlying asset variable. Using standard regression techniques, whereby *ex-post* realized payoffs are regressed against the specified *ex-ante* observed values, these functions can be estimated from cross-sectional information obtained in the simulation.

The MLSM is quite powerful in estimating delta and gamma for American type options with single asset. We give a numerical example of MLSM in the application of American-type barrier options. For higher dimensional cases, the choice of the initial distribution and basis functions for regression is rather complicated, even for 2-dimensional problems. This algorithm is not applicable for other Greeks other than delta and gamma, which is another limitation of MLSM. Wang and Caflisch [80] proposed a method to estimate vega. They simulate different values of volatility for each stock path from a carefully chosen distribution, and compute a regression function of volatility values by regressing pathwise payouts against pathwise volatilities. However, the simulation results are unsatisfactory. Based on our numerical simulations, different distributions or different parameters associated with the same distribution give vega values with large variances.

Example: Delta and Gamma for an American Barrier Option. We consider an American-style "up-and-out" put option (at the money) on an asset with strike price K, barrier $H = 50$, and maturity $T = 1/3$. The underlying asset follows the geometric Brownian motion, with the volatility $\sigma = 0.2$, and the risk free interest rate $r = 0.0488$. In contrast to the availability of closed-form solutions for European style barrier options, American barrier option's closed form is unknown in the literature. So we have to resort to numerical methods.

If the asset price hits the barrier, the option becomes worthless. The payoff function is given by

$$\varphi(t) = e^{-rT}(K - S(T))^{+} 1 \left\{ \max_{0 \le t \le T} \mathrm{S(t)} \ < H \right\}. \tag{1.38}$$

We use the MLSM approach with Monte Carlo simulation and quasi-Monte Carlo simulation to evaluate the Greeks delta and gamma. Both simulations use 10,000 sample paths for the asset price and 50 time discretization points per year. The root mean square error (RMSE) is obtained from 10 independent randomizations. We compare the results with the benchmark value in Ritchken [67]. The numerical results are within one root-mean square error of the benchmark values.

Table 1.1: Greeks and RMSE for an American Barrier Put Options (ATM) using MLSM

S_0	MC Delta	QMC Delta	Benchmark Binomial	MC Gamma	QMC Gamma	Benchmark Binomial
40	-0.7682	-0.7660	-0.7696	0.0504	0.0515	0.0512
RMSE	0.0103	0.0053		0.0035	0.0024	
42.5	-0.6306	-0.6263	-0.6285	0.0496	0.0491	0.0483
RMSE	0.0104	0.0081		0.0026	0.0011	
45	-0.5072	-0.5049	-0.5197	0.0534	0.0536	0.0539
RMSE	0.0093	0.0082		0.0024	0.0012	

CHAPTER 2

ESTIMATING GREEKS FOR MOUNTAIN RANGE OPTIONS UNDER THE BLACK-SCHOLES MODEL

In this chapter we consider a class of complex exotic options with dispersion called *Mountain Range Options*, originally introduced by Société Générale in 1998. Mountain range options are essentially a combination of two types of exotic derivatives: basket options and range options. In a basket option, the price is dependent on several underlying assets rather than a single one. In a range option, the option is active for a particular time. In general, the payoffs of mountain range options are not directly related to the overall performance of the basket. Instead, they are related to the best performer or the worst performer of the basket during a particular time period. Variations have been given the names, *Everest, Atlas, Altiplano* and *Himalaya.* In the literature, Quessette [63] gives a general introduction to all the variations; Overhaus [60] discusses Himalaya in depth.

2.1 Everest Option

The *Everest* option gives the holder the right to receive a payoff at maturity based on the worst-performing security of a large basket of stocks or indexes. It has long maturity of 10-15 years and consists of 10 to 25 underlying assets in the basket (see Quessette [63]). We consider an Everest option written on n assets. Let $S_i(t)$ denote the price of the ith asset at time t, and the terminal payoff $\varphi : \mathbb{R}_+^n \mapsto \mathbb{R}$, is

$$\varphi(S_1(T), S_2(T), ..., S_n(T)) = \min_{i=1,\dots,n} \left\{ \frac{S_i(T)}{S_i(0)} \right\}. \tag{2.1}$$

The payoff structure of the Everest option is a special case of the call on minimum with zero strike. Stulz [75] derives the pricing formula for options on the maximum and the minimum of two assets under the Black-Scholes setting. Johnson [49] extends this result to handle European options in the case of multiple assets by using the asset price as numéraire. He measures the call option in units of the asset's price S, and the pricing formula obtained looks like a European put on an asset with current price Ke^{-rT}/S and unit exercise price. Although Johnson's results are in closed form, the resulting pricing formula involves calculating multivariate normal distribution. Boyle and Tse [11] derive an approximation method to compute the European call options on the maximum and minimum of n assets in terms of order statistics.

2.1.1 Everest Option Pricing under the Black-Scholes Model

In this section, we apply Johnson's method [49] to the Everest option pricing under the Black-Scholes model. We first restate his results as follows: Consider an European call on minimum of n assets $S_1, S_2, ..., S_n$ with strike K, time to maturity T, and risk-free interest rate r. The price of a call option on the minimum of n assets is

$$\begin{aligned}
C_{\min} &= S_1\mathcal{N}_n(d_1(S_1,K,\sigma_1^2), -d_1'(S_1,S_2,\sigma_{1,2}^2), .., -d_1'(S_1,S_n,\sigma_{1,n}^2), -\rho_{1,2,1}, -\rho_{1,3,1}, ..., \rho_{2,3,1}, ...) \\
&+ S_2\mathcal{N}_n(d_1(S_2,K,\sigma_2^2), -d_1'(S_2,S_1,\sigma_{1,2}^2), .., -d_1'(S_2,S_n,\sigma_{2,n}^2), -\rho_{1,2,2}, -\rho_{2,3,2}, ..., \rho_{1,3,2}, ...) \\
&+ ... \\
&+ S_n\mathcal{N}_n(d_1(S_n,K,\sigma_n^2), -d_1'(S_n,S_1,\sigma_{1,n}^2), .., -d_1'(S_n,S_{n-1},\sigma_{n-1,n}^2), -\rho_{1,n,n}, \rho_{2,n,n}, ..., \rho_{1,2,n}, ...) \\
&- Ke^{-rT}\mathcal{N}_n(d_2(S_1,K,\sigma_1^2), d_2(S_2,K,\sigma_2^2), ..., d_2(S_n,K,\sigma_n^2), \rho_{12}, \rho_{13}, ...), \qquad (2.2)
\end{aligned}$$

where $\mathcal{N}_n(\cdot)$ denotes the n-variate normal distribution function, and $\sigma_{i,j}^2$, $\rho_{j,k,i}$, $d_{1,2}$, d_1' are given by

$$\begin{aligned}
\sigma_{i,j}^2 &= \sigma_i^2 - 2\rho_{i,j}\sigma_i\sigma_j + \sigma_j^2, \\
\rho_{j,k,i} &= \frac{\sigma_i^2 - \rho_{i,j}\sigma_i\sigma_j - \rho_{i,k}\sigma_i\sigma_k + \rho_{j,k}\sigma_j\sigma_k}{\sigma_{i,j}\sigma_{i,k}}, \\
d_{1,2}(S_i,S_j,\sigma_{i,j}^2) &= \frac{\log(S_i/S_j) + (r \pm \sigma_{i,j}^2/2)T}{\sigma_{i,j}\sqrt{T}}, \\
d_{1,2}(S_i,K,\sigma_i^2) &= \frac{\log(S_i/K) + (r - \sigma_i^2/2)T}{\sigma_i\sqrt{T}}, \\
d_1'(S_i,S_j,\sigma_{i,j}^2) &= \frac{\log(S_i/S_j) + \sigma_{i,j}^2 T/2}{\sigma_{i,j}\sqrt{T}}.
\end{aligned}$$

The pricing formula of an Everest option is similar to the call on minimum, but it cannot be obtained by simply setting the $K = 0$ in Eqn.(2.2). We derive the price of an Everest option in detail as follows.

Let $W(t) = (W_1(t), ..., W_n(t))'$ be an n-dimensional standard Brownian motion on the probability space $(\Omega, \mathcal{F}, Q)$, where Q is the risk-neutral measure and the filtration $\{\mathcal{F}(t), 0 \leq t \leq T\}$ is generated by $W(t)$. Suppose each underlying asset $S_i(t)$ in the basket follows the geometric

Brownian motion

$$dS_i(t) = rS_i(t)dt + S_i(t)\sum_{k=1}^{n} \sigma_{i,k} dW_k(t), \quad i = 1, ..., n \tag{2.3}$$

where the term $\sigma_{i,k}$ determines the correlation between the assets i and k in the basket. An alternative way to express the dynamics of asset $S_i(t)$ is given through

$$dS_i(t) = rS_i(t)dt + S_i(t)\sigma_i d\tilde{W}_i(t), \quad i = 1, ..., n, \tag{2.4}$$

where σ_i is the volatility of asset $S_i(t)$, and $\tilde{W}_i(t) = \sum_{k=1}^{n} \frac{\sigma_{i,k}}{\sigma_i} W_k(t)$. We can show that $\tilde{W}_i(t)$, for $i = 1, 2, ..., n$, are Brownian motions by Lévy's theorem (see Appendix A Theorem 32), since it is a continuous martingale, and $d\tilde{W}_i(t)d\tilde{W}_i(t) = dt$. Moreover, they are jointly correlated with

$$Cov(\tilde{W}_i(t), \tilde{W}_j(t)) = \rho_{i,j}t.$$

The vector $\tilde{W}(t) = (\tilde{W}_1(t), \tilde{W}_2(t), ..., \tilde{W}_n(t))$ follows a multivariate normal distribution with zero mean and covariance matrix $\Sigma_t = t\Sigma$, where Σ is the correlation coefficient matrix with the entries $\rho_{i,j}$. The correlated Brownian motion $\tilde{W}(t)$ can be expressed as $LW(t)$ with $W(t)$ representing a standard n-dimensional Brownian motion and L representing any matrix such that $\Sigma = LL'$. As usual, choosing L to be the Cholesky factor of Σ is the easiest way to decompose Σ. The solution of stochastic differential equation (2.4) is

$$S_i(t) = S_i(0) \exp\left(\left(r - \frac{\sigma_i^2}{2}\right) t + \sigma_i \tilde{W}_i(t)\right). \tag{2.5}$$

According to the risk-neutral pricing formula (1.13), the price of an Everest option is given by

$$V(t, S) = e^{-r(T-t)} E\left[\min_{i=1,...,n} \left(\frac{S_i(T)}{S_i(0)}\right) \middle| \mathcal{F}_t\right].$$

We introduce a group of new processes $X_i(t) = S_i(t)/S_i(0)$, for $i = 1, ..., n$, to represent the performance of each underlying asset with the initial condition $X_i(0) = 1$. In terms of the performance processes, the terminal payoff function (2.1) is written as

$$\varphi(X_1(T), X_2(T), ..., X_n(T)) = \min_{i=1,...,n} \{X_i(T)\}.$$

The option value at time t is

$$V(t, X) = e^{-r(T-t)} E\left[\min_{i=1,...,n} \{X_i(T)\} \middle| \mathcal{F}_t\right]. \tag{2.6}$$

Proposition 6 *The price of an Everest option written on a basket consisting of* n *underlying assets is*

$$V(t,S) = \sum_{k=1}^{n} \left(\frac{S_k(t)}{S_k(0)} \mathcal{N}_{n-1}\left(d_{1,k}(t), ..., d_{k-1,k}(t), d_{k+1,k}(t), ..., d_{n,k}(t), \Sigma_k\right) \right), \tag{2.7}$$

where

$$d_{i,k}(t) = \frac{\ln\left(\frac{S_i(t)}{S_i(0)}\right) - \ln\left(\frac{S_k(t)}{S_k(0)}\right) - \frac{\sigma_{i,k}^2}{2}(T-t)}{\sigma_{i,k}\sqrt{T-t}}$$

with $\sigma_{i,k}^2 = \sigma_i^2 + \sigma_k^2 - 2\rho_{i,k}\sigma_i\sigma_k$. *Here,* Σ_k *denotes the correlation coefficient matrix of all assets except for asset* k *and it has the form*

$$\Sigma_k = \begin{pmatrix} 1 & \rho_{1,2,k} & \cdots & \rho_{1,k-1,k} & \rho_{1,k+1,k} & \cdots & \rho_{1,n,k} \\ \rho_{2,1,k} & 1 & & & & & \\ \vdots & & \ddots & & & & \vdots \\ \rho_{k-1,1,k} & & & 1 & & & \\ \rho_{k+1,1,k} & & & & 1 & & \\ \vdots & & & & & \ddots & \\ \rho_{n,1,k} & \rho_{n,2,k} & \cdots & \rho_{n,k-1,k} & \rho_{n,k+1,k} & \cdots & \rho_{n,n,k} \end{pmatrix}_{(n-1)\times(n-1)}$$

with the entries

$$\rho_{i,j,k} = \frac{\rho_{i,j}\sigma_i\sigma_j - \rho_{i,k}\sigma_i\sigma_k - \rho_{j,k}\sigma_j\sigma_k + \sigma_k^2}{\sigma_{i,k}\sigma_{j,k}}, \text{ with } i, j \neq k, \tag{2.8}$$

where i *and* j *are not equal to* k.

Proof. Suppose the kth asset performs worst at maturity, that is $X_k(T) < X_i(T)$ for all $i \neq k$. Then $X_k(t)$ is set as the numéraire. The process $X_{i,k}(t) = X_i(t)/X_k(t)$ is regarded as the performance of $X_i(t)$ discounted by the performance of $X_k(t)$, or the performance of $X_i(t)$ measured in units of $X_k(t)$. By Ito's formula, the dynamics of $X_{i,k}(t)$ satisfies the stochastic differential equation

$$\frac{dX_{i,k}(t)}{X_{i,k}(t)} = \left(\sigma_k^2 - \sigma_i\sigma_k\rho_{i,k}\right)dt + \sigma_i d\tilde{W}_i(t) - \sigma_k d\tilde{W}_k(t). \tag{2.9}$$

There exists a probability measure Q^k such that $X_{i,k}(t)$ is a Q^k-local martingale, which is called the equivalent martingale measure (EMM) associated with the numéraire $X_k(t)$. The measure Q^k is defined by its Radon-Nikodym derivative with respect to the measure Q. Under Q^k, the dynamics of $X_{i,k}(t)$ satisfies

$$\frac{dX_{i,k}(t)}{X_{i,k}(t)} = \sigma_i d\hat{W}_i(t) - \sigma_k d\hat{W}_k(t), \tag{2.10}$$

where $\hat{W}_i(t)$ and $\hat{W}_k(t)$ are Q^k-Brownian motions. The solution of the SDE (2.10) follows the log-normal distribution

$$X_{i,k}(T) \sim LN\left(\ln(X_{i,k}(t)) - \frac{\sigma_{i,k}^2}{2}(T-t),\ \ \sigma_{i,k}^2(T-t)\right).$$

The price of an Everest option under the measure Q^k is the same as the one under the measure Q. Therefore, when the kth asset performs worst at maturity, the price of the Everest option price is given by

$$\begin{aligned} V_k(t,X) &= X_k(t)E^k\left[1_{\{X_1(T)>X_k(T),X_2(T)>X_k(T),\dots,X_n(T)>X_k(T)\}}\middle|\,\mathcal{F}_t\right] \\ &= X_k(t)E^k\left[1_{\{\min_i\{X_{i,k}(T)\}>1\}}\middle|\,\mathcal{F}_t\right]. \end{aligned} \tag{2.11}$$

We can write the expectation in Eqn.(2.11) in terms of probability as follows:

$$\begin{aligned} & E^k[1_{\{\min_i\{X_{i,k}(T)\}>1\}}|\mathcal{F}_t] \qquad (2.12) \\ = \ & P_t^k(\min_{i\neq k}\{X_{i,k}(T)\}>1) \\ = \ & P_t^k\left(\ln(X_{1,k}(T))>0,\dots,\ln(X_{k-1,k}(T))>0,\ln(X_{k+1,k}(T))>0,\dots,\ln(X_{n,k}(T))>0\right). \end{aligned}$$

For each i with $i \neq k$, we have

$$\begin{aligned} & P_t^k(\ln(X_{i,k}(T))>0) \\ = \ & \int_0^{+\infty}\frac{1}{\sigma_{i,k}\sqrt{2\pi(T-t)}}\exp\left(-\frac{(x-(\ln(X_{i,k}(t))-\sigma_{i,k}^2(T-t)/2))^2}{2(T-t)\sigma_{i,k}^2}\right)dx \\ = \ & \mathcal{N}\left(\frac{\ln(X_{i,k}(t))-\frac{\sigma_{i,k}^2}{2}(T-t)}{\sigma_{i,k}\sqrt{T-t}}\right), \end{aligned}$$

with $\mathcal{N}(\cdot)$ representing the standard normal distribution function. We express the joint density function in Eqn.(2.12) in terms of an integral:

$$\begin{aligned} & \underbrace{\int_{-\infty}^{d_{1,k}}\cdots\int_{-\infty}^{d_{n,k}}}_{n-1}\frac{1}{(\sqrt{2\pi})^{\frac{n-1}{2}}|\Sigma_k|^{1/2}}\exp\left(-\frac{\mathbf{x}^T\Sigma_k^{-1}\mathbf{x}}{2}\right)d\mathbf{x} \qquad (2.13) \\ = \ & \mathcal{N}_{n-1}\left(d_{1,k}(t),\dots,d_{k-1,k}(t),d_{k+1,k}(t),\dots,d_{n,k}(t),\Sigma_k\right), \end{aligned}$$

where

$$d_{i,k}(t) = \frac{\ln(X_{i,k}(t))-\sigma_{i,k}^2(T-t)/2}{\sigma_{i,k}\sqrt{T-t}}$$

and $\mathcal{N}_{n-1}(\cdot)$ denotes the $n-1$ dimensional normal distribution function. Σ_k is the $(n-1)$ by $(n-1)$ correlation coefficient matrix of the random vector $(X_{1,k}, ..., X_{k-1,k}, X_{k+1,k}, ..., X_{n,k})'$ with elements

$$\rho_{i,j,k} = \frac{\rho_{i,j}\sigma_i\sigma_j - \rho_{i,k}\sigma_i\sigma_k - \rho_{j,k}\sigma_j\sigma_k + \sigma_k^2}{\sigma_{i,k}\sigma_{j,k}}, \ for\ i, j \neq k.$$

Besides the kth asset, any other asset in the basket also has a possibility of being the worst performing one. Consequently, the price of the Everest option Eqn.(2.6) would be the sum of the pricing formula (2.11) for all k starting from 1 to n. It is given by

$$\begin{aligned} V(t, X) &= \sum_{k=1}^{n} \left(X_k(t) E^k \left[1_{\{\min_{i,i\neq k} X_{i,k}(T) > 1\}} | \mathcal{F}_t \right] \right) \\ &= \sum_{k=1}^{n} \left(X_k(t) \mathcal{N}_{n-1} \left(d_{1,k}(t), ..., d_{k-1,k}(t), d_{k+1,k}(t), ..., d_{n,k}(t), \Sigma_k \right) \right). \end{aligned}$$

We switch back to asset process $S_i(t)$, from the performance process $X_i(t)$, and obtain the pricing formula

$$V(t, S) = \sum_{k=1}^{n} \left(\frac{S_k(t)}{S_k(0)} \mathcal{N}_{n-1} \left(d_{1,k}(t), ..., d_{k-1,k}(t), d_{k+1,k}(t), ..., d_{n,k}(t), \Sigma_k \right) \right).$$

Specifically, the price of the Everest option at $t = 0$ is

$$V(0) = \sum_{k=1}^{n} \left(\mathcal{N}_{n-1} \left(\frac{-\sigma_{1,k}\sqrt{T}}{2}, ..., \frac{-\sigma_{k-1,k}\sqrt{T}}{2}, \frac{-\sigma_{k+1,k}\sqrt{T}}{2}, ..., \frac{-\sigma_{n,k}\sqrt{T}}{2}, \Sigma_k \right) \right). \tag{2.14}$$

■

2.1.2 Estimating Greeks of Everest Options

Greeks are estimated in order to measure the sensitivity of an Everest option value to the changes in different parameters. The delta and gamma of the Everest option written on multiple assets are vectors of which the elements are the first or second derivatives with respect to every single asset in the basket. Cross gamma measures the sensitivity of delta in one underlying asset to changes in each of another underlying asset. It takes the form of a matrix with entries the second derivatives with respect to two different assets.

Proposition 7 *The delta of the Everest option is the zero vector. For each asset i in the basket, we have*

$$\frac{\partial V(0)}{\partial S_i(0)} = 0.$$

Proof. The pricing formula Eqn.(2.14) does not depend on the initial values of assets. Thus the derivative of $V(0)$ with respect to the initial value $S(0)$ equals zero. ■

Proposition 8 *The gamma matrix of the Everest option is the zero matrix. That is, for all* $i, j = 1, ..., n$,

$$\frac{\partial^2 V(0)}{\partial S_i(0) \partial S_j(0)} = 0.$$

Proof. This result follows trivially, since the first derivative, delta, is zero from Proposition 7. ■

Proposition 9 *The rho of the Everest option is zero.*

Proof. The interest rate r in the drift term of Eqn.(2.9) is cancelled out when we apply the Ito's formula to $X_{i,k}$. The right hand side of pricing formula (2.14) doesn't depend on the interest rate r.

■

There are also other important Greeks: vega measures the sensitivity of the option price with respect to the asset's volatility; theta measures the sensitivity with respect to the maturity; correlation delta measures the sensitivity of the option price to the fixed correlation parameters between any two different assets. However, it is difficult to take the derivative of the option price Eqn.(2.14) with respect to the asset volatility σ, maturity T, or correlation coefficient $\rho_{i,j}$, because they are involved in the integrand and upper bounds of a multi-integral. For example, through Eqn.(2.13), we observe that the volatility σ_i exists in all the upper bounds, in the determinant $|\Sigma|$, and in the inverse matrix Σ^{-1}. Therefore, deriving an analytical formula for $\frac{\partial V}{\partial \sigma_i}$ is not simple. In the following section, we derive the pathwise estimators and likelihood ratio estimators for estimating Greeks.

2.1.3 Pathwise Method

Before applying the pathwise method, we first examine the conditions (A1)-(A4) that were introduced in Chapter 1 to ensure that the pathwise estimators of the Everest option are unbiased.

(A1) The asset price S is differentiable with probability 1 for the parameter θ.

(A2) The payoff function φ is differentiable with probability 1 for the parameter θ.

(A3) The payoff function φ is Lipschitz continuous with respect to the asset price S.

(A4) The asset price S is Lipschitz continuous with respect to the parameter θ.

Proposition 10 *The discounted payoff function of the Everest option (2.1) is Lipschitz continuous with respect to underlying assets $S(0)$, interest rate r, asset volatility σ, maturity T, and correlation coefficients $\rho_{i,j}$ between any two assets in the basket.*

Proof. We will prove the proposition for $S(0)$; the other cases are proved similarly. Condition (A1) is true because $S_i(t)$ follows a geometric Brownian motion. Condition (A2) is also satisfied since the discounted payoff (2.1) is continuous and differentiable everywhere. For condition (A3), we prove that the function $f_n : R^n \to R$,

$$f_n(x_1, x_2, ..., x_n) = \min(x_1, x_2, ..., x_n)$$

is Lipschitz continuous by mathematical induction in three steps.

Step 1: When $n = 1$, $|f_1(x_1) - f_1(y_1)| \leq |x_1 - y_1|$, i.e., f_1 is Lipschitz continuous.

Step 2: Assume $n = m$ and the function $f_m(x) = \min(x_1, x_2, ..., x_m)$ is Lipschitz continuous. Thus, we have

$$|\min(x_1, x_2, ..., x_m) - \min(y_1, y_2, ..., y_m)| \leq |x_1 - y_1| + |x_2 - y_2| + ... + |x_m - y_m|.$$

Step 3: When $n = m + 1$, we have

$$\begin{aligned}
& |\min(x_1, x_2, ..., x_{m+1}) - \min(y_1, y_2, ..., y_{m+1})| \\
= \ & |\min(f_m(x), x_{m+1}) - \min(f_m(y), y_{m+1})| \\
= \ & \left| \frac{f_m(x) + x_{m+1} - |f_m(x) - x_{m+1}|}{2} - \frac{f_m(y) + y_{m+1} - |f_m(y) - y_{m+1}|}{2} \right| \\
= \ & \left| \frac{(f_m(x) - f_m(y)) + (x_{m+1} - y_{m+1}) - (|f_m(x) - x_{m+1}| - |f_m(y) - y_{m+1}|)}{2} \right| \\
\leq \ & \frac{|f_m(x) - f_m(y)| + |x_{m+1} - y_{m+1}| + \big||f_m(x) - x_{m+1}| - |f_m(y) - y_{m+1}|\big|}{2} \\
\leq \ & \frac{|f_m(x) - f_m(y)| + |x_{m+1} - y_{m+1}| + |f_m(x) - x_{m+1} - f_m(y) + y_{m+1}|}{2} \\
\leq \ & |f_m(x) - f_m(y)| + |x_{m+1} - y_{m+1}| \\
= \ & |x_1 - y_1| + |x_2 - y_2| + ... + |x_m - y_m| + |x_{m+1} - y_{m+1}|,
\end{aligned}$$

i.e. $f_{n+1}(x)$ is Lipschitz continuous with respect to x.

Therefore the Everest payoff function (2.1) is Lipschitz continuous with respect to all the underlying assets. Condition (A4) is satisfied since S_i is log-normally distributed. Condition (A3) together with (A4) implies that the discounted payoff is almost surely Lipschitz in θ because the Lipschitz property is preserved by composition. In conclusion, the pathwise estimators of the Everest option are unbiased.

A similar proof applies when θ is taken as the interest rate r, the volatility σ, the maturity T, or the correlation coefficients $\rho_{i,j}$, and we omit the details here. The range of these parameters Θ is set to be a bounded interval. This restriction is harmless because for the purposes of estimating a derivative at θ, we may take an arbitrarily small neighbourhood of θ. ■

Delta and Gamma. The pathwise estimator of delta is the zero vector. The gamma matrix is the zero matrix. The jth entry of the pathwise estimator of delta is given as follows:

$$\begin{aligned}\frac{\partial(e^{-rT}\varphi(\cdot))}{\partial S_j(0)} &= e^{-rT}\left(\frac{\partial}{\partial S_j(0)}\left(\frac{S_j(T)}{S_j(0)}\right)\cdot 1\left\{\min_{i=1,\ldots,n}\left(\frac{S_i(T)}{S_i(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\}\right)\\ &= e^{-rT}\left(\frac{1}{S_j(0)}\frac{S_j(T)}{S_j(0)}-\frac{S_j(T)}{S_j^2(0)}\right)\cdot 1\left\{\min_{i=1,\ldots,n}\left(\frac{S_i(T)}{S_i(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\}\\ &= 0.\end{aligned}$$

Both Hall [39] and Cao [19] derive this result, and our Proposition 7 confirms it.

Rho. From Proposition 9, rho equals zero. The pathwise estimator for rho is also zero:

$$\begin{aligned}\frac{d\left(e^{-rT}\varphi(\cdot)\right)}{dr} &= \frac{d}{dr}\left(e^{-rT}\frac{S_j(T)}{S_j(0)}\cdot 1\left\{\min_{i=1,\ldots,n}\left(\frac{S_i(T)}{S_i(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\}\right)\\ &= \left(e^{-rT}\frac{S_j(T)}{S_j(0)}(-T)+e^{-rT}\frac{S_j(T)}{S_j(0)}T\right)\cdot 1\left\{\min_{i=1,\ldots n}\left(\frac{S_i(T)}{S_i(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\}\\ &= 0.\end{aligned}$$

Vega. The derivative of the discounted payoff with respect to the jth asset is:

$$\begin{aligned}\frac{\partial\left(e^{-rT}\varphi(\cdot)\right)}{\partial\sigma_j} &= e^{-rT}\frac{dS_j(T)}{d\sigma_j}\left(\frac{1}{S_j(0)}\right)\cdot 1\left\{\min_{i=1,\ldots n}\left(\frac{S_i(T)}{S_i(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\}\\ &= e^{-rT}\frac{S_j(T)}{S_j(0)}\left(-\sigma_jT+\sqrt{T}L_jZ\right)\cdot 1\left\{\min_{i=1,\ldots n}\left(\frac{S_i(T)}{S_i(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\},\end{aligned}$$

where

$$\frac{dS_j(T)}{d\sigma_j} = S_j(T)(-\sigma_jT+\sqrt{T}L_jZ),$$

L_j is the jth row of the Cholesky matrix of Σ, and Z is the standard n-dimensional normally distributed vector that is used to generate the sample paths.

Theta. According to Eqn.(2.5), the derivative of each asset price S_j with respect to maturity is given by

$$\frac{dS_j(T)}{dT} = \frac{S_j(T)}{S_j(0)}\left(r - \frac{\sigma_j^2}{2} + \frac{\sigma_j L_j Z}{2\sqrt{T}}\right),$$

where L_j and Z are defined as before. The pathwise estimator for theta is:

$$\begin{aligned}-\frac{d\left(e^{-rT}\varphi(\cdot)\right)}{dT} &= -\frac{d}{dT}\left(e^{-rT}\frac{S_j(T)}{S_j(0)}\right)\cdot 1\left\{\min_{i=1,\dots n}\left(\frac{S_i(T)}{S_i(0)}\right) = \frac{S_j(T)}{S_j(0)}\right\}\\ &= -e^{-rT}\frac{S_j(T)}{S_j(0)}\left(-\frac{\sigma_j^2}{2} + \frac{\sigma_j L_j Z}{2\sqrt{T}}\right)\cdot 1\left\{\min_{i=1,\dots n}\left(\frac{S_i(T)}{S_i(0)}\right) = \frac{S_j(T)}{S_j(0)}\right\}.\end{aligned}$$

Correlation Delta. To calculate the correlation delta, we use a different approach to factor the correlation matrix Σ through the Cholesky decomposition while generating the asset prices. We first explain the access pattern of the Cholesky decomposition. The correlation matrix Σ could be decomposed in the form of $\Sigma = LL'$ with entries of $l_{i,j}$,

$$l_{j,j} = \sqrt{1 - \sum_{k=1}^{j-1} l_{j,k}^2}, \tag{2.15}$$

$$l_{i,j} = \frac{1}{l_{j,j}}\left(\rho_{i,j} - \sum_{k=1}^{j-1} l_{i,k}l_{j,k}\right), \quad for\ i > j. \tag{2.16}$$

According to Eqn.(2.15) and (2.16), the access pattern for the algorithm starts from the leftmost entry to the entry on the diagonal for each row. For example, considering a 4 by 4 matrix, we calculate the entries for L in the order of $l_{1,1}$, $l_{2,1}$, $l_{2,2}$, $l_{3,1}$, $l_{3,2}$, $l_{3,3}$, $l_{4,1}$, $l_{4,2}$, $l_{4,3}$, $l_{4,4}$:

$$\Sigma = \begin{pmatrix} 1 & & & \\ \rho_{2,1} & 1 & & \\ \rho_{3,1} & \rho_{3,2} & 1 & \\ \rho_{4,1} & \rho_{4,2} & \rho_{4,3} & 1 \end{pmatrix} \Rightarrow L = \begin{pmatrix} l_{1,1} & & & \\ l_{2,1} & l_{2,2} & & \\ l_{3,1} & l_{3,2} & & \\ l_{4,1} & l_{4,2} & l_{4,3} & l_{4,4} \end{pmatrix}.$$

When calculating the ij-th element of L, we use all the elements of Σ which are above the ith row, as well as the elements on the ith row but at or before the jth column. For example, to calculate $l_{3,2}$, we use all the entries on the first two rows and $\rho_{3,1}$, $\rho_{3,2}$ of matrix Σ. This decomposition pattern will bring difficulties when we calculate the derivative of the entries on the ith row of L with respect to the coefficient $\rho_{i,j}$ with small row number i. For example, if we want to calculate $\frac{\partial V}{\partial \rho_{2,1}}$, we first have to identify all the elements in L which are expressed via $\rho_{2,1}$ through the Cholesky decomposition. According to Eqn.(2.16) and its access pattern, we know that $l_{2,1}$ is the first element using $\rho_{2,1}$. Then calculating all elements on the second column will involve $l_{2,1}$. Later these elements will be used to generate other elements with larger row numbers. Therefore, it is hard to apply the chain rule under this decomposition. On the other hand, if we want to calculate $\frac{\partial V}{\partial \rho_{4,3}}$, we observe that $\rho_{4,3}$ is used only for calculating $l_{4,3}$ and $l_{4,4}$. The correlation delta is obtained by taking the derivative with respect to these two element $l_{4,3}$, $l_{4,4}$ and then $\rho_{4,3}$ by the chain rule.

The purpose of explaining the access pattern of Cholesky decomposition is that when we calculate $\frac{\partial V}{\partial \rho_{i,j}}$, we hope to localize $\rho_{i,j}$ within the rightmost two entries in the last row of the lower triangular Cholesky decomposition matrix. Therefore, we could easily identify the only two entries that are expressed in terms of $\rho_{i,j}$ in the matrix L, and then take the derivative with respect to it. Now we are ready to introduce the method to derive the pathwise estimator for correlation delta. We point out that this method works for all options with continuous payoffs and written on multiple underlying assets.

Proposition 11 *The pathwise estimator of the correlation delta for the Everest option between any two assets i and j is given by*

$$\frac{\partial \left(e^{-rT}\varphi(\cdot)\right)}{\partial \rho_{i,j}} = e^{-rT}\sigma_j\sqrt{T}\frac{S_j(T)}{S_j(0)}\left(\frac{Z_{n-1}}{l_{n-1,n-1}} - \frac{l_{n,n-1}Z_n}{l_{n-1,n-1}l_{n,n}}\right)\cdot 1\left\{\min_{i=1,\ldots n}\left(\frac{S_i(T)}{S_i(0)}\right) = \frac{S_j(T)}{S_j(0)}\right\},$$

where $l_{i,j}$ is the ij-th entry of the Cholesky matrix; Z_{n-1} and Z_n are the last two components of the normally distributed random vector that is used to generate $S_j(T)$.

Proof. The proof is based on a re-ordering of the assets in the basket and their correlation matrix. For example, when estimating the correlation delta between assets i and j, we write all assets in the form of a vector with the asset i and j in the last two positions:

$$S = (S_1, S_2, ..., S_{i-1}, S_{i+1}, ..., S_{j-1}, S_{j+1}, , ..., S_i, S_j)'.$$

This reordering will help localize the correlation coefficient $\rho_{i,j}$ within the rightmost two entries in the last row of the lower triangular Cholesky decomposition matrix. According to the access pattern, $\rho_{i,j}$ is accessed only for calculating the right-most two entries in the last row. For example, suppose $i = 3$ and $j = 4$, $\rho_{4,3}$ is used only to calculate $l_{4,3}$ and $l_{4,4}$. Moreover, $l_{4,3}$ and $l_{4,4}$ are used only for sampling the 4th assets. Accordingly, the jth asset is generated by

$$S_j(T) = S_j(0)\exp\left(\left(r - \frac{\sigma_j^2}{2}\right)T + \sigma_j\sqrt{T}\sum_{k=1}^{n} l_{j,k}Z_k\right), \tag{2.17}$$

where Z_i is the ith entry of the normal random vector $Z = (Z_1, Z_2, ..., Z_n)$. The parameter $\rho_{i,j}$ is involved in $l_{n,n-1}$ and $l_{n,n}$ of Eqn.(2.17). The derivative of $S_j(T)$ with respect to this parameter is obtained by

$$\begin{aligned}
\frac{dS_j(T)}{d\rho_{i,j}} &= S_j(T)\sigma_j\sqrt{T}\cdot\frac{d}{d\rho_{i,j}}\left(\sum_{k=1}^{n} l_{j,k}Z_k\right)\\
&= S_j(T)\sigma_j\sqrt{T}\left(\frac{dl_{n,n-1}}{d\rho_{i,j}}Z_{n-1}+\frac{dl_{n,n}}{d\rho_{i,j}}Z_n\right)\\
&= S_j(T)\sigma_j\sqrt{T}\left(\frac{1}{l_{n-1,n-1}}Z_{n-1}+\frac{1}{2l_{n,n}}(-2)l_{n,n-1}\frac{dl_{n,n-1}}{d\rho_{i,j}}Z_n\right)\\
&= \frac{S_j(T)\sigma_j\sqrt{T}}{l_{n-1,n-1}}\left(Z_{n-1}-\frac{l_{n,n-1}}{l_{n,n}}Z_n\right).
\end{aligned}$$

The pathwise estimator of the correlation delta between asset i and j is

$$\begin{aligned}
e^{-rT}\frac{\partial\varphi(0,S)}{\partial\rho_{i,j}} &= e^{-rT}\left(\frac{dS_j(T)}{d\rho_{i,j}}\left(\frac{1}{S_j(0)}\right)\cdot 1\left\{\min_{i=1,\ldots,n}\left(\frac{S_k(T)}{S_k(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\}\right)\\
&= e^{-rT}\frac{S_j(T)\sigma_j\sqrt{T}}{S_j(0)}\left(\frac{dl_{n,n-1}}{d\rho_{i,j}}Z_{n-1}+\frac{dl_{n,n}}{d\rho_{i,j}}Z_n\right)\cdot 1\left\{\min_{i=1,\ldots,n}\left(\frac{S_i(T)}{S_i(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\}\\
&= e^{-rT}\frac{S_j(T)\sigma_j\sqrt{T}}{S_j(0)}\left(\frac{Z_{n-1}}{l_{n-1,n-1}}-\frac{l_{n,n-1}Z_n}{l_{n-1,n-1}l_{n,n}}\right)\cdot 1\left\{\min_{i=1,\ldots,n}\left(\frac{S_i(T)}{S_i(0)}\right)=\frac{S_j(T)}{S_j(0)}\right\}.
\end{aligned}$$

■

2.1.4 Likelihood Ratio Method

The likelihood ratio estimators for the Everest option are derived in this section. Eqn.(2.5) suggests that $S_i(T)$ for $i = 1, ..., n$, has the distribution of $\exp(Y_i)$. The vector $Y = (Y_1, Y_2, ..., Y_n)'$ follows the normal distribution $\mathcal{N}(\mu(\theta), \tilde{\Sigma}(\theta))$, where $\mu(\theta; \cdot)$ is a vector with elements $\mu_i = \ln(S_i(0)) + (r - \sigma_i^2/2)T$. Covariance matrix has the form $\tilde{\Sigma}(\theta) = TA\Sigma A'$, where $A = diag(\sigma_1, ..., \sigma_k)$. The probability density function of Y could be written as

$$g(\theta; Y) = \frac{1}{\sqrt{(2\pi)^k|\tilde{\Sigma}(\theta)|}}\exp\left(-\frac{1}{2}(y-\mu(\theta))'\tilde{\Sigma}^{-1}(\theta)(y-\mu(\theta))\right). \tag{2.18}$$

We generate the sample paths Y by $\mu+\sqrt{T}ALZ$, where L is the Cholesky matrix such that $LL' = \Sigma$, and $Z \sim N(0, I)$.

Delta. We write the option price as the expectation of a discounted payoff under the risk-neutral measure,

$$E_{S(0)}[e^{-rT}\varphi(Y)] = \int_{R^n} e^{-rT}\varphi(y)g_{S(0)}(y)dy. \tag{2.19}$$

Since the integrand in Eqn.(2.19) is Lipschitz continuous with respect to $S_i(0)$, we switch the order of the differentiation and the integration to obtain

$$
\begin{aligned}
\frac{d}{dS_i(0)} E_{S_i(0)}[e^{-rT}\varphi(Y)] &= \int_{R^n} e^{-rT}\varphi(y)\frac{d}{dS_i(0)}\left(g_{S_i(0)}(y)\right) dy \\
&= \int_{R^n} e^{-rT}\varphi(y)\frac{\dot{g}_{S_i(0)}(y)}{g_{S_i(0)}(y)} g_{S_i(0)}(y) dy \\
&= E_{S_i(0)}\left[e^{-rT}\varphi(y)\frac{\dot{g}_{S_i(0)}(y)}{g_{S_i(0)}(y)}\right].
\end{aligned}
$$

The score function simplifies as

$$
\frac{\dot{g}_{S_i(0)}(y)}{g_{S_i(0)}(y)} = \frac{d}{dS_i(0)} \log(g_{S_i(0)}(Y)) = (Y-\mu)'\tilde{\Sigma}^{-1}\frac{d\mu}{dS_i(0)} = \frac{(Z'L^{-1}A^{-1})_i}{\sqrt{T}S_i(0)},
$$

and $(Z'L^{-1}A^{-1})_i$ is the ith component of the row vector $Z'L^{-1}A^{-1}$. The likelihood ratio estimator of delta with respect to the ith asset is

$$
e^{-rT}\varphi_{Everest}(Y)\frac{(Z'L^{-1}A^{-1})_i}{\sqrt{T}S_i(0)}.
$$

Gamma. The score function for the second derivative is

$$
\begin{aligned}
\frac{\ddot{g}_{S_i(0)}(y)}{g_{S_i(0)}(y)} &= \left((Y-\mu)'\tilde{\Sigma}^{-1}\frac{d\mu}{dS_i(0)}\right)^2 - \left(\frac{d\mu}{dS_i(0)}\right)'\tilde{\Sigma}^{-1}\frac{d\mu}{dS_i(0)} + (Y-\mu)'\tilde{\Sigma}^{-1}\frac{d^2\mu}{dS_i^2(0)} \\
&= \left(\frac{(Z'L^{-1}A^{-1})_i}{\sqrt{T}S_i(0)}\right)^2 - \frac{(\tilde{\Sigma}^{-1})_{ii}}{S_i^2(0)} - \frac{(Z'L^{-1}A^{-1})_i}{\sqrt{T}S_i^2(0)},
\end{aligned}
$$

where $(\tilde{\Sigma}^{-1})_{ii}$ is the ith element on the diagonal of inverse matrix $\tilde{\Sigma}$. Thus, the likelihood ratio estimator of gamma is

$$
e^{-rT}\varphi_{Everest}(Y)\left(\frac{((Z'L^{-1}A^{-1})_i)^2}{TS_i^2(0)} - \frac{(\tilde{\Sigma}^{-1})_{ii}}{S_i^2(0)} - \frac{(Z'L^{-1}A^{-1})_i}{\sqrt{T}S_i^2(0)}\right).
$$

Rho. The risk-neutral price of the option is

$$
E_r[e^{-rT}\varphi_{Everest}(Y)] = \int_{R^n} e^{-rT}\varphi(y)g_r(y)dy,
$$

where the interest rate r is a parameter of the discount factor and a parameter of the density function $g_r(Y)$ as well. Before taking the derivative, we switch the order of the differentiation and integration due to the continuity of the integrand with respect to r. That is,

$$\begin{aligned}\frac{d}{dr}E_r[e^{-rT}\varphi(Y)] &= \int_{R^n}\varphi(y)\frac{d}{dr}\left(e^{-rT}g_r(y)\right)dy\\ &= \int_{R^n}e^{-rT}\varphi(y)\left(-Tg_r(y)+\dot{g}_r(y)\right)dy\\ &= \int_{R^n}e^{-rT}\varphi(y)\left(-T+\frac{\dot{g}_r(y)}{g_r(y)}\right)g_r(y)dy\\ &= E_r\left[e^{-rT}\varphi(y)\left(-T+\frac{\dot{g}_r(y)}{g_r(y)}\right)\right],\end{aligned}$$

where the score function is

$$\frac{\dot{g}_r(Y)}{g_r(Y)} = Z'L^{-1}A^{-1}\sqrt{T}I_{n\times 1},$$

and $I_{n\times 1}$ is an n-dimensional vector with all elements equal to 1. The likelihood ratio estimator for rho is

$$e^{-rT}\varphi_{Everest}(Y)\cdot(-T+Z'L^{-1}A^{-1}\sqrt{T}I_{n\times 1}).$$

Theta. The parameter of interest for theta is the maturity T. We write the density function (2.18) as a function of T as follows: We take the derivative of the risk-neutral price of the option with respect to the maturity T.

$$\begin{aligned}\frac{d}{dT}E_T[e^{-rT}\varphi_{Everest}(Y)] &= \int_{R^n}\varphi(y)\frac{d}{dT}\left(e^{-rT}g_T(y)\right)dy\\ &= \int_{R^n}e^{-rT}\varphi(y)\left(-r+\frac{\dot{g}_T(y)}{g_T(y)}\right)g_T(y)dy\\ &= E_T\left[e^{-rT}\varphi(y)\left(-r+\frac{\dot{g}_T(y)}{g_T(y)}\right)\right],\end{aligned}$$

where the score function is

$$\begin{aligned}\frac{\dot{g}_T(Y)}{g_T(Y)} &= \frac{d}{dT}\log(g_T(Y))\\ &= \frac{1}{g}\frac{1}{\sqrt{(2\pi)^k|A\Sigma A'|}}\left(-\frac{k}{2T}\right)T^{-\frac{k}{2}}\exp\left(-\frac{1}{2T}(y-\mu)'(A\Sigma A')^{-1}(y-\mu)\right)\\ &+ \frac{1}{2T^2}(y-\mu)'(A\Sigma A')^{-1}(y-\mu)+\frac{1}{T}(Y-\mu)'(A\Sigma A')^{-1}\left(r-\frac{\sigma^2}{2}\right)\\ &= -\frac{k}{2T}+\frac{1}{\sqrt{T}}Z'L^{-1}A^{-1}\left(rI_{n\times 1}-\frac{\sigma^2}{2}\right)+\frac{1}{2T}Z'Z,\end{aligned}$$

and $I_{n\times 1}$ and identity vector. The likelihood ratio estimator for theta is

$$e^{-rT}\varphi_{Everest}(Y)\left(-r-\frac{k}{2T}+\frac{1}{\sqrt{T}}Z'L^{-1}A^{-1}\left(rI_{n\times 1}-\frac{\sigma^2}{2}\right)+\frac{1}{2T}Z'Z\right).$$

Vega. In order to explicitly express the parameter σ_i in the density function, we write the density function (2.18) as

$$g_\sigma(Y) = \frac{1}{\sigma_1 ... \sigma_n \sqrt{(2\pi)^k T^k |\Sigma|}} \exp\left(-\frac{1}{2T}(A^{-1}(Y-\mu(\sigma)))'\Sigma^{-1}A^{-1}(Y-\mu(\sigma))\right),$$

where

$$A^{-1}(Y-\mu(\sigma)) = \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_n} \end{pmatrix} \begin{pmatrix} Y_1-\mu_1 \\ Y_2-\mu_2 \\ \vdots \\ Y_n-\mu_n \end{pmatrix} = \begin{pmatrix} \frac{Y_1-\mu_1}{\sigma_1} \\ \frac{Y_2-\mu_2}{\sigma_2} \\ \vdots \\ \frac{Y_n-\mu_n}{\sigma_n} \end{pmatrix}.$$

The derivative of $A^{-1}(Y-\mu(\theta))$ with respect to σ_i is

$$\frac{d}{d\sigma_i}\left(A^{-1}(Y-\mu(\theta))\right) = \left(0, ..., T - \frac{1}{\sigma_i^2}(Y_i-\mu_i), ..., 0\right)'.$$

Thus the score function is given by

$$\begin{aligned}
\frac{\dot{g}_{\sigma_i}(Y)}{g_{\sigma_i}(Y)} &= \frac{d}{d\sigma_i}\log(g_{\sigma_i}(Y)) \\
&= -\frac{1}{\sigma_i} - \frac{1}{T}(A^{-1}(Y-\mu(\sigma)))'\Sigma^{-1}\frac{d}{d\sigma_i}\left(A^{-1}(Y-\mu(\sigma))\right) \\
&= -\frac{1}{\sigma_i} - \frac{1}{T}(A^{-1}(Y-\mu(\sigma)))'\Sigma^{-1}\left(0, ..., T - \frac{1}{\sigma_i^2}(Y_i-\mu_i), ..., 0\right)' \\
&= -\frac{1}{\sigma_i} - \frac{1}{\sqrt{T}}Z'L^{-1}\left(0, ..., T - \frac{\sqrt{T}}{\sigma_i^2}(ALZ)_i, ..., 0\right)'.
\end{aligned}$$

The likelihood ratio estimator for vega with respect to ith entry is

$$e^{-rT}\varphi_{Everest}(Y)\left(-\frac{1}{\sigma_i} - \frac{1}{\sqrt{T}}Z'L^{-1}\left(0, ..., T - \frac{\sqrt{T}}{\sigma_i^2}(ALZ)_i, ..., 0\right)'\right).$$

Correlation delta. The derivation of the correlation delta estimator involves calculating the derivative of the correlation matrix Σ with respect to its entries $\rho_{i,j}$, for $i, j = 1, ..., n$ with $i \neq j$. The score function for the correlation delta of assets i and j is given as follows:

$$\begin{aligned}
\frac{\dot{g}_{\rho_{i,j}}(Y)}{g_{\rho_{i,j}}(Y)} &= \frac{d}{d\rho_{i,j}}\log(g_{\rho_{i,j}}(Y)) \\
&= -\frac{1}{2}Trace\left(\tilde{\Sigma}^{-1}\frac{\partial\tilde{\Sigma}}{\partial\rho_{i,j}}\right) + \frac{1}{2}(Y-\mu)'\tilde{\Sigma}^{-1}\frac{\partial\tilde{\Sigma}}{\partial\rho_{i,j}}\tilde{\Sigma}^{-1}(Y-\mu) \\
&= -\frac{1}{2}Trace\left(L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}\right) + \frac{1}{2}Z'L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}Z,
\end{aligned}$$

where Z is a standard n-dimensional normal random vector that is used to generate the sample paths. The matrix $\partial\Sigma/\partial\rho_{i,j}$ contains entries at (i,j) and (j,i) equal to 1; all other entries are equal to 0. In summary, the likelihood ratio estimator of the correlation delta between asset i and asset j is

$$e^{-rT}\varphi_{Everest}(Y)\left(-\frac{1}{2}Trace\left(L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}\right)+\frac{1}{2}Z'L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}Z\right).$$

2.1.5 Numerical Results

An Everest option on four correlated stocks is considered in the numerical example. The covariance matrix was estimated from data, using 256 days of data from June 1st, 2006 through June 8th, 2007 collected by Wharton Research Data Services. The selected stocks are from the following financial institutions: Citigroup (C), Freddie Mac (old ticker: FRE), J.P. Morgan Chase (JPM) and Lehman Brothers (LEH). Each stock has a zero dividend yield. The maturity of the Everest option is $T = 15$, and the risk-free rate is $r = 0.04$. The spot prices and the volatilities are shown in Table 2.1, and the correlation coefficient matrix is shown in Table 2.2.

We first compute the Everest option price using the analytical formula given by Eqn.(2.7). The formula requires the computation of the multivariate normal distribution function

$$\mathcal{N}_3(x_1,x_2,x_3,\Sigma)=\frac{1}{(2\pi)^{3/2}\sqrt{|\Sigma|}}\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\int_{-\infty}^{x_3}\exp(-\frac{1}{2}X'\Sigma^{-1}X)dX_3dX_2dX_1. \tag{2.20}$$

The dimension n is set at 3 to account for the fact that one of the four assets in the basket will function as the numéraire. We use an algorithm proposed by Genz [30] to compute the multivariate normal distribution function in Eqn.(2.20). This algorithm uses a quasi-Monte Carlo sequence with a sample size M to approximate the multivariate integral for normal distribution. Table 2.3 presents the price of the Everest option using this algorithm for various values for M. We also compare the price and the standard deviation of the Everest option using a pseudorandom sequence and the scrambled Faure sequence in Table 2.4 and Figure 2.1. For more information about the scrambled Faure sequence, see Appendix C.

The results for estimating the Greeks for the Everest option using Monte Carlo and quasi-Monte Carlo methods are shown in Table 2.5 and Table 2.6. We compare the estimators obtained from the pathwise method, the likelihood ratio method and the finite difference method. The root mean square errors (RMSE) are shown in parenthesis. We set the sample size at 10,000 with 40 independent runs. With the finite difference method, the perturbation size $h = 0.0001$ when $\theta = r$, $h = 1/12$ when $\theta = T$, and $h = 0.1$ when $\theta = T$ or ρ. The standard errors with the likelihood ratio method are typically 1.5 to 3 times greater than the pathwise and finite difference standard error. The scrambled Faure sequence is used in the quasi-Monte Carlo simulation. We observe that using quasi-Monte Carlo reduces the standard error of the likelihood ratio estimator by a factor of 3, and the standard error of pathwise and finite difference estimators by a factor of 2.

Table 2.1: The spot prices and volatilities of four stocks in the basket on June 1st, 2006.

	Spot	Volatility (σ^2)
C	54	0.096313
FRE	66	0.092968
JPM	53	0.113978
LEH	61	0.316459

Table 2.2: Correlation coefficient matrix of the stocks in the basket.

Σ	C	FRE	JPM	LEM
C	1.0000	0.4549	0.6759	0.5559
FRE	0.4549	1.0000	0.5491	0.4337
JPM	0.6759	0.5491	1.0000	0.6430
LEH	0.5559	0.4337	0.6430	1.0000

Table 2.3: Comparison of prices and standard deviations for an Everest option by computing Eqn.(2.7) using Genz's algorithm [30] with three different M values.

M	Price	Std.Dev
1000	0.2374	7.31×10^{-4}
5000	0.2374	1.51×10^{-4}
10000	0.2374	5.04×10^{-5}

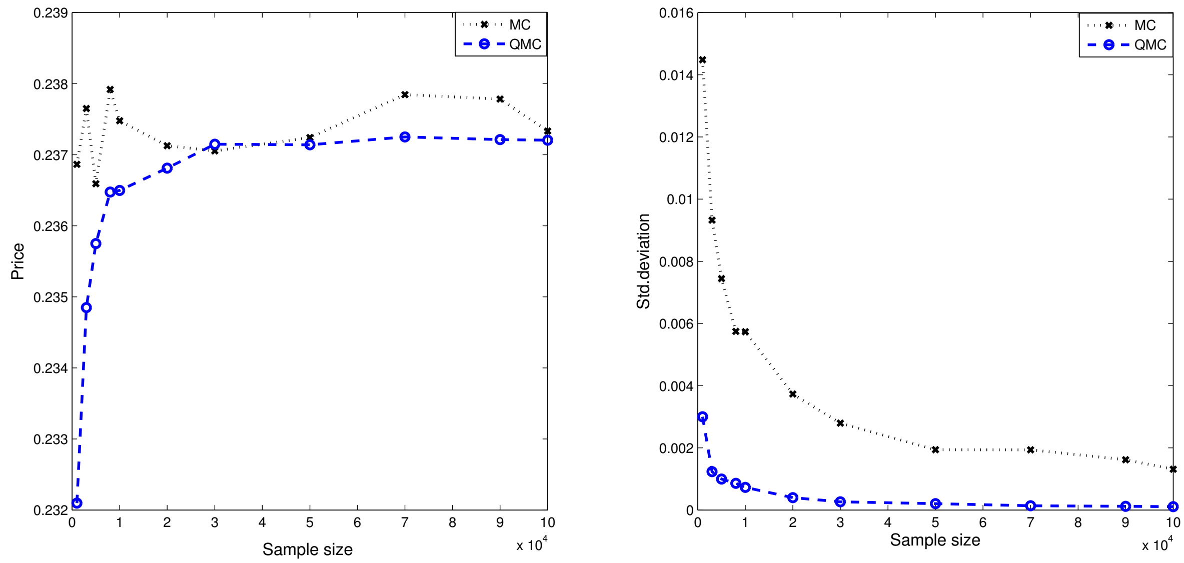

Figure 2.1: Comparison of the Everest option prices and their standard deviations between the Monte Carlo method and quasi-Monte Carlo method.

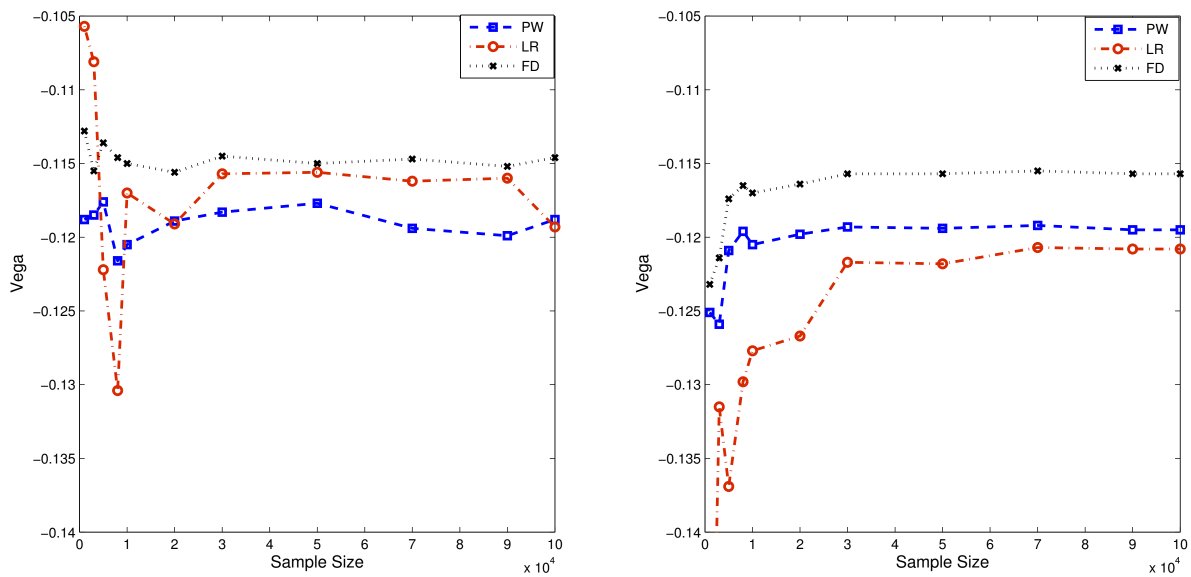

Figure 2.2: Comparison of Greek vega for the Everest option between the Monte Carlo method (left) and quasi-Monte Carlo method (right).

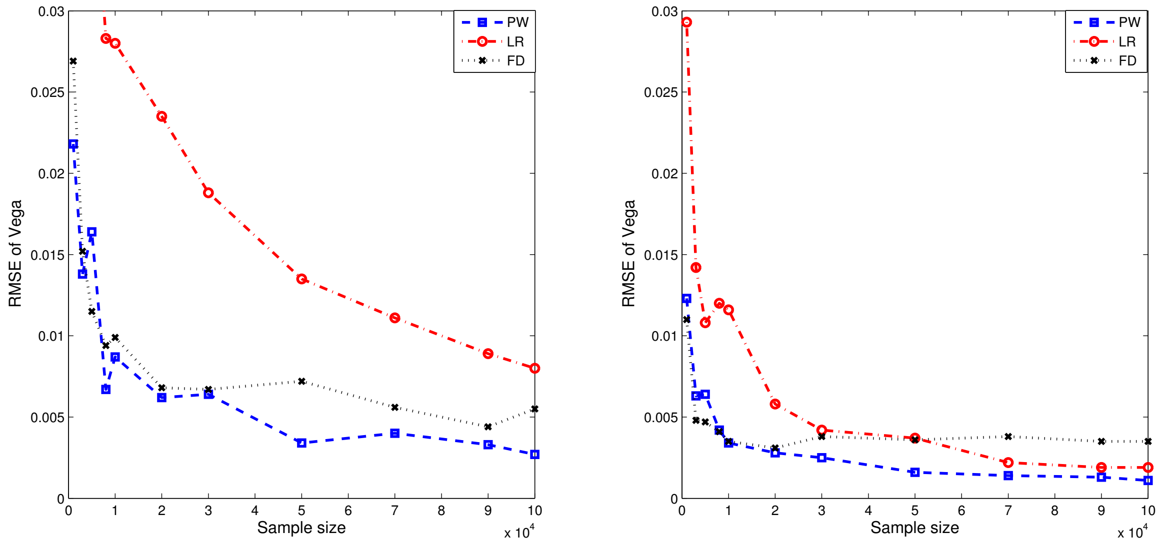

Figure 2.3: Comparison of the standard deviations of Greek vega between the Monte Carlo method and quasi-Monte Carlo method.

Table 2.4: Comparison of the price and standard deviation (in parenthesis) of an Everest option using pseudorandom and quasi-Monte Carlo sequences.

Sample size	Monte Carlo (Std.Dev)	Quasi-Monte Carlo (Std.Dev)
10^3	0.2389 (1.69×10^{-2})	0.2321 (3.41×10^{-3})
10^4	0.2377 (4.81×10^{-3})	0.2366 (7.45×10^{-4})
10^5	0.2373 (1.42×10^{-3})	0.2373 (1.29×10^{-4})
10^6	0.2374 (5.70×10^{-4})	0.2374 (2.42×10^{-5})

Table 2.5: Comparison of Greeks and root mean square errors (RMSE) for an Everest option using the pathwise method, the likelihood ratio method, and the finite difference method under Monte Carlo simulations.

Greeks	Pathwise	Likelihood Ratio	Finite Difference
delta	0	0	0
gamma	0	NE	0
rho	0	0	0
theta	-0.0143 (0.0002)	-0.0141(0.0009)	-0.0143(0.0002)
vega			
$\partial V/\partial\sigma_1$	-0.1192 (0.0104)	-0.1198(0.0243)	-0.1133(0.0191)
$\partial V/\partial\sigma_2$	-0.1660 (0.0100)	-0.1663(0.0381)	-0.1704(0.0293)
$\partial V/\partial\sigma_3$	-0.1102 (0.0067)	-0.1169(0.0248)	-0.1174(0.0181)
$\partial V/\partial\sigma_4$	-0.5419 (0.0088)	-0.5437(0.0122)	-0.5399(0.0174)
corr.delta			
$\partial V/\partial\rho_{1,2}$	0.0471 (0.0052)	0.0481(0.0083)	0.0474(0.0067)
$\partial V/\partial\rho_{1,3}$	0.0412 (0.0044)	0.0394(0.0116)	0.0414(0.0124)
$\partial V/\partial\rho_{1,4}$	0.0785 (0.0078)	0.0804(0.0088)	0.0795(0.0089)
$\partial V/\partial\rho_{2,3}$	0.0426 (0.0053)	0.0415(0.0103)	0.0423(0.0078)
$\partial V/\partial\rho_{2,4}$	0.0923 (0.0069)	0.0909(0.0092)	0.0926(0.0079)
$\partial V/\partial\rho_{3,4}$	0.0732 (0.0068)	0.0727(0.0091)	0.0761(0.0089)

Table 2.6: Comparison of Greeks and RMSE for an Everest option using the pathwise method, the likelihood ratio method and the finite difference method under quasi-Monte Carlo simulations (scrambled Faure sequences).

Greeks	Pathwise	Likelihood Ratio	Finite Difference
delta	0	0	0
gamma	0	NE	0
rho	0	0	0
theta	-0.0144(0.00007)	-0.0147(0.0002)	-0.0144(0.00008)
vega			
$\partial V/\partial\sigma_1$	-0.1198 (0.0043)	-0.1285(0.0093)	-0.1239 (0.0083)
$\partial V/\partial\sigma_2$	-0.1677 (0.0045)	-0.1707(0.0065)	-0.1730 (0.0054)
$\partial V/\partial\sigma_3$	-0.1109 (0.0030)	-0.1263(0.0083)	-0.1042 (0.0079)
$\partial V/\partial\sigma_4$	-0.5417 (0.0027)	-0.5435(0.0033)	-0.5389 (0.0034)
corr.delta			
$\partial V/\partial\rho_{1,2}$	0.0460 (0.0025)	0.0473(0.0021)	0.0483(0.0032)
$\partial V/\partial\rho_{1,3}$	0.0410 (0.0020)	0.0408(0.0025)	0.0392(0.0028)
$\partial V/\partial\rho_{1,4}$	0.0787 (0.0031)	0.0783(0.0023)	0.0772(0.0034)
$\partial V/\partial\rho_{2,3}$	0.0416 (0.0024)	0.0401(0.0027)	0.0389(0.0031)
$\partial V/\partial\rho_{2,4}$	0.0913 (0.0033)	0.0906(0.0023)	0.0921(0.0019)
$\partial V/\partial\rho_{3,4}$	0.0726 (0.0042)	0.0720(0.0021)	0.0714(0.0041)

2.2 Atlas Option

The Atlas option is a European basket option. At the maturity, the m_1 worst performing assets as well as the m_2 best performing assets are removed from the basket, which contains a total of n assets. The terminal payoff $\psi : \mathbb{R}_+^n \to \mathbb{R}$ is based on the average performance of the remaining assets in the basket:

$$\psi(S_1(T), S_2(T), ..., S_n(T)) = \left(\frac{\sum_{n(i)=1+m_1}^{n-m_2} \frac{S_i(T)}{S_i(0)}}{n - m_1 - m_2} - K \right)^+ , \tag{2.21}$$

where K is the strike price. The stochastic model of these assets is the same as in the case of the Everest option, and $X_i(t)$ denotes the performance of the ith asset at time t. At maturity T, we order the asset performances increasingly and label them by $n(i)$. For example, $n(i) = 4$ means that the performance of the asset S_i ranks 4. We rewrite the terminal payoff (2.21) in terms of performance variables.

$$\psi(X_1(T), X_2(T), ..., X_n(T)) = \left(\frac{\sum_{n(i)=1+m_1}^{n-m_2} X_i(T)}{n - m_1 - m_2} - K \right)^+ . \tag{2.22}$$

There is no known analytical formula for the price of an Atlas option under the general Black-Scholes setting. When an Atlas option is written on a basket containing only two assets, it degenerates to call on maximum or call on minimum, and Stulz [75] gives the closed form pricing formula. Pricing an Atlas option is difficult because of the multi-asset nature of this type of option and uncertainty about which assets will be removed at the maturity. For that reason, no closed formula has yet been developed that can be applied to a basket containing more than two assets. For such cases, we estimate the value of the Atlas option in using Monte Carlo and quasi-Monte Carlo methods.

2.2.1 Pathwise Method

In this section, we derive the pathwise and likelihood ratio estimators for Greeks of the Atlas option. The pathwise estimators for the Greeks are unbiased, because the payoff function (2.21) is Lipschitz continuous with respect to the initial asset values $S(0)$, interest rate r, maturity T, volatility σ, and all the correlation coefficients $\rho_{i,j}$ with $i, j = 1, ..., n$. The proof of Lipschitz continuity is quite similar to that of Proposition 10, given that Eqn.(2.21) is a composition of a minimum function and a maximum function. We omit the proof.

Delta and Gamma. The pathwise estimator for delta is the zero vector, and gamma is the zero matrix. It is based on two facts:

1. The payoff function of the Atlas option (2.21) is that of a European call option on the average performance of some assets.
2. When the dynamics of underlying assets is geometric Brownian motion, the derivative of any performance process with respect to its initial asset price is zero, i.e. $\frac{\partial X_i(t)}{\partial S_i(0)} = 0$. This is shown by Eqn.(2.15).

From the chain rule and the above facts, the delta and the gamma of the Atlas option are zero.

Rho. The pathwise estimator of rho for the Atlas option is

$$\begin{aligned}
\frac{d(e^{-rT}\psi(\cdot))}{dr} &= e^{-rT}\left(\frac{\sum_{n(i)=1+m_1}^{n-m_2}\frac{\partial X_i(T)}{\partial r}}{n-m_1-m_2} - \left(\frac{\sum_{n(i)=1+m_1}^{n-m_2} X_i(T)}{n-m_1-m_2} - K\right)T\right)\cdot 1_{\{\psi(\cdot)>0\}} \\
&= e^{-rT}\left(\frac{\sum_{n(i)=1+m_1}^{n-m_2} X_i(T)T}{n-m_1-m_2} - \left(\frac{\sum_{n(i)=1+m_1}^{n-m_2} X_i(T)}{n-m_1-m_2} - K\right)T\right)\cdot 1_{\{\psi(\cdot)>0\}} \\
&= e^{-rT}TK\cdot 1_{\{\psi(\cdot)>0\}}.
\end{aligned}$$

Vega. The derivative of the discounted payoff with respect to the ith asset is:

$$\frac{\partial(e^{-rT}\psi(\cdot))}{\partial\sigma_i} = e^{-rT}X_i(T)\left(\frac{-\sigma_i T + \sqrt{T}L_i Z}{n-m_1-m_2}\right)\cdot 1_{\{1+m_1\le n(i)\le n-m_2\}}\cdot 1_{\{\psi(\cdot)>0\}}$$

where L_i is the ith row of the Cholesky matrix of Σ, and Z is the standard n-dimensional normally distributed vector that is used to generate the sample paths.

Theta. The pathwise estimator for theta is

$$-\frac{d(e^{-rT}\psi(\cdot))}{dT} = -e^{-rT}\left(\frac{\sum_{n(i)=1+m_1}^{n-m_2} X_i(T)\left(-\frac{\sigma_i^2}{2} + \frac{\sigma_i L_i Z}{2\sqrt{T}}\right)}{n-m_1-m_2} + rK\right)\cdot 1_{\{\psi(\cdot)>0\}},$$

where L_i and Z are defined as before.

Correlation delta. To estimate the Greek correlation delta for the Atlas option, we use the same method that is used in the proof for Proposition 9. We take the derivative of the discounted payoff function with respect to $\rho_{i,j}$, the pathwise correlation delta between the asset i and the asset

j which is given by

$$\begin{aligned}
&\frac{\partial(e^{-rT}\psi(\cdot))}{\partial\rho_{i,j}} \\
= & \left(\frac{e^{-rT}X_j(T)\sigma_j\sqrt{T}}{n-m_1-m_2}\right)\frac{d}{d\rho_{i,j}}\left(\sum_{k=1}^{n} l_{j,k}Z_k\right)\cdot 1_{\{1+m_1\leq n(j)\leq n-m_2\}}\cdot 1_{\{\psi(\cdot)>0\}} \\
= & \left(\frac{e^{-rT}X_j(T)\sigma_j\sqrt{T}}{n-m_1-m_2}\right)\cdot\left(\frac{Z_{n-1}}{l_{n-1,n-1}}-\frac{l_{n,n-1}Z_n}{l_{n-1,n-1}l_{n,n}}\right)\cdot 1_{\{1+m_1\leq n(j)\leq n-m_2\}}\cdot 1_{\{\psi(\cdot)>0\}}
\end{aligned}$$

where Z_i is the ith entry of vector $Z=(Z_1,Z_2,...,Z_n)$, and $l_{n-1,n-1}$, $l_{n,n-1}$, $l_{n,n}$ are entries of the Cholesky matrix.

2.2.2 Likelihood Ratio Method

As shown in Section 2.1.4, the asset price S_i at the maturity has the distribution of $\exp(Y_i)$, for $i=1,2,...,n$. The joint density function of $Y=(Y_1,Y_2,...,Y_n)$ is given by Eqn.(2.18). We think of the payoff function $\psi(\cdot)$ in Eqn.(2.21) as a function of the random variable Y. The likelihood ratio estimators for Greeks of an Atlas option are given as follows. We omit the derivations, since the score functions for the parameters of interest are the same as before. The Greeks are obtained by replacing the Everest payoff in their likelihood ratio estimators with the Atlas payoff.

Delta. The likelihood ratio estimator of delta with respect to the ith asset is

$$e^{-rT}\psi_{Atlas}(Y)\frac{(Z'L^{-1}A^{-1})_i}{\sqrt{T}S_i(0)}, \tag{2.23}$$

where L is the Cholesky matrix such that $LL'=\Sigma$, $A=diag(\sigma_1,...,\sigma_k)$ and Z is a standard n-dimensional normal random vector that is used to generate the sample Y.

Gamma. The likelihood ratio estimator of gamma with respect to the ith asset is

$$e^{-rT}\psi_{Atlas}(Y)\left(\frac{((Z'L^{-1}A^{-1})_i)^2}{TS_i^2(0)}-\frac{(\tilde{\Sigma}^{-1})_{ii}}{S_i^2(0)}-\frac{(Z'L^{-1}A^{-1})_i}{\sqrt{T}S_i^2(0)}\right),$$

where L, A, Z are the same as before, and $(\tilde{\Sigma}^{-1})_{ii}$ is the ith element on the diagonal of the inverse matrix of $\tilde{\Sigma}$.

Rho. The likelihood ratio estimator of rho is

$$e^{-rT}\psi_{Atlas}(Y)\left(-T+Z'L^{-1}A^{-1}\sqrt{T}I_{n\times 1}\right).$$

Theta. The likelihood ratio estimator of theta is

$$e^{-rT}\psi_{Atlas}(Y)\left(-r-\frac{k}{2T}+\frac{1}{\sqrt{T}}Z'L^{-1}A^{-1}\left(rI_{n\times 1}-\frac{\sigma^2}{2}\right)+\frac{1}{2T}Z'Z\right).$$

Vega. The likelihood ratio estimator of vega with respect to the ith asset is

$$e^{-rT}\psi_{Atlas}(Y)\left(-\frac{1}{\sigma_i}-\frac{1}{\sqrt{T}}Z'L^{-1}\ \left(0,...,T-\frac{\sqrt{T}}{\sigma_i^2}(ALZ)_i,...,0\right)'\right).$$

Correlation delta. The likelihood ratio estimator of correlation delta between asset i and asset j is

$$e^{-rT}\psi_{Atlas}(Y)\left(-\frac{1}{2}Trace\left(L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}\right)+\frac{1}{2}Z'L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}Z\right),$$

where the matrix $(\partial\Sigma/\partial\rho_{i,j})$ contains entries at (i,j) and (j,i) equal to 1. All other entries are equal to 0.

2.2.3 Numerical Results

We use the same assets for the Atlas option: Citigroup (C), Freddie Mac (old ticker: FRE), J.P. Morgan Chase (JPM) and Lehman Brothers (LEH). Each stock has a zero dividend yield. The maturity of the Atlas option is $T = 6$, and the risk-free rate is $r = 0.04$. The spot prices and volatilities are given in Table 2.1. The correlation coefficient matrix is given in Table 2.2.

Under quasi-Monte Carlo simulation, the Atlas option price is 0.2959 with a standard deviation of 9.08×10^{-5}; under the Monte Carlo simulation, the Atlas option price is 0.2963 with a standard deviation of 2.10×10^{-3}. We take them as our benchmark value for the price of the Atlas option (see Table 2.7). Fig 2.4 shows that the quasi-Monte Carlo simulation (the dotted line) results in a lower standard deviation, about one tenth that of Monte Carlo (the solid line).

We compare the pathwise, likelihood ratio, and finite difference estimators for Greeks of the Atlas option in Table 2.8 (Monte Carlo simulation) and Table 2.9 (quasi-Monte Carlo simulation). We note that the pathwise estimators have consistently a smaller root-mean square error that the likelihood ratio estimators.

The comparison between the pathwise estimators and the likelihood estimators for the correlation delta is shown in Fig. 2.5 (MC vs. QMC). We observe that the standard deviation of the likelihood ratio estimator is twice as large as that of the pathwise estimators (see Fig. 2.6).

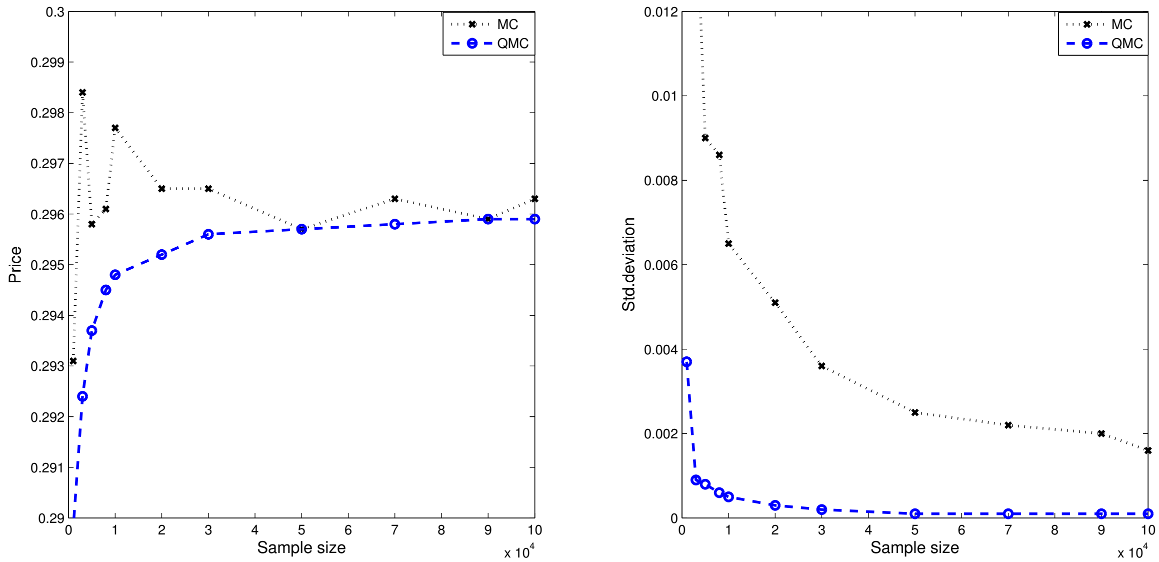

Figure 2.4: Comparison of the Atlas option prices and their RMSE between the Monte Carlo and quasi-Monte Carlo method.

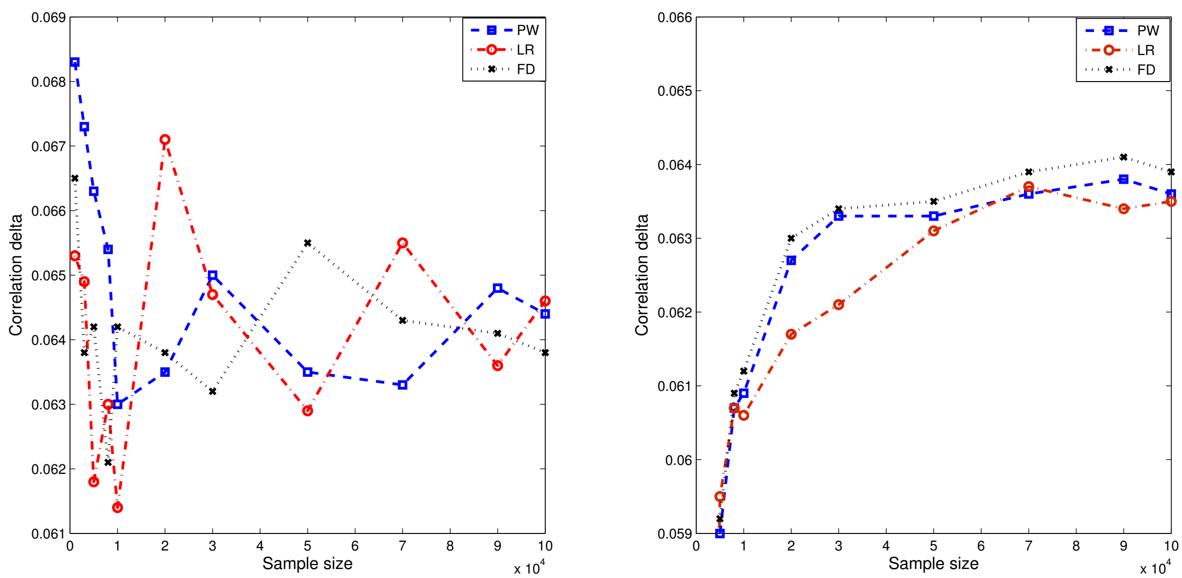

Figure 2.5: Comparison of the Greek correlation delta for Atlas Option between the Monte Carlo method and quasi-Monte Carlo method.

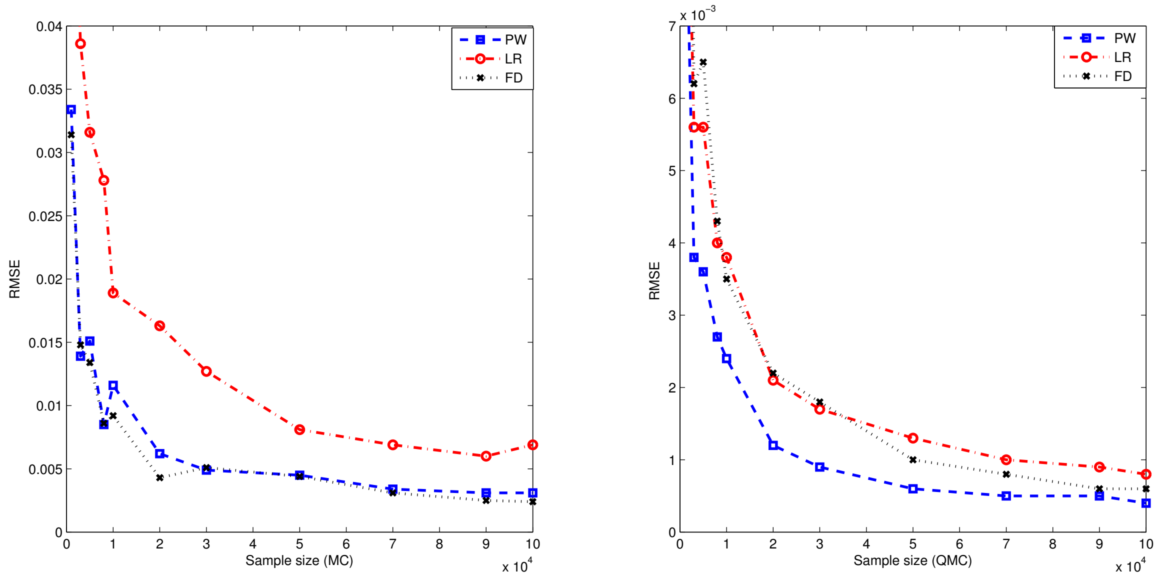

Figure 2.6: Comparison of the RMSE of the correlation delta for Atlas option between the Monte Carlo and quasi-Monte Carlo method.

Table 2.7: Comparison of the price and standard deviation for an Atlas option using pseudorandom sequences (MC) and quasi-Monte Carlo (QMC) sequences.

	MC	QMC
Price	0.2963	0.2959
Std.Dev	0.0021	9.1×10^{-5}

Table 2.8: Comparison of Greeks and their root of mean square errors (RMSE) for an Atlas option using the pathwise method, the likelihood ratio method, and the finite difference method under Monte Carlo simulations.

Greeks	Pathwise	Likelihood Ratio	Finite Difference
delta	0	0	0
gamma	0	0	0
rho	2.0017 (0.0195)	1.9998(0.0984)	2.0022(0.0223)
theta	0.0203 (0.0007)	0.0200(0.0025)	0.0202(0.0008)
vega			
$\partial V/\partial\sigma_1$	0.1383(0.0137)	0.1352(0.0689)	0.1377(0.0103)
$\partial V/\partial\sigma_2$	0.0936(0.0089)	0.0904(0.0350)	0.0909(0.0083)
$\partial V/\partial\sigma_3$	0.1468(0.0142)	0.1496(0.0602)	0.1443(0.0135)
$\partial V/\partial\sigma_4$	-0.0675(0.0036)	-0.0684(0.0235)	-0.0647(0.0033)
corr.delta			
$\partial V/\partial\rho_{1,2}$	0.0271(0.0043)	0.0268(0.0141)	0.0276(0.0051)
$\partial V/\partial\rho_{1,3}$	0.0589(0.0054)	0.0532(0.0210)	0.0597(0.0065)
$\partial V/\partial\rho_{1,4}$	0.0372(0.0071)	0.0387(0.0150)	0.0362(0.0059)
$\partial V/\partial\rho_{2,3}$	0.0426(0.0053)	0.0415(0.0103)	0.0423(0.0078)
$\partial V/\partial\rho_{2,4}$	0.0221(0.0069)	0.0239(0.0135)	0.0224(0.0057)
$\partial V/\partial\rho_{3,4}$	0.0641(0.0089)	0.0670(0.0187)	0.0646(0.0085)

Table 2.9: Comparison of Greeks and RMSE for an Atlas option using the pathwise method, the likelihood ratio method and the finite difference method under quasi-Monte Carlo simulations (scrambled Faure sequences).

Greeks	Pathwise	Likelihood Ratio	Finite Difference
delta	0	0	0
gamma	0	0	0
rho	2.0022(0.0071)	1.9659(0.0127)	2.0031(0.0061)
theta	0.0200 (0.0001)	0.0189 (0.0005)	0.0200(0.0001)
vega			
$\partial V/\partial\sigma_1$	0.1369(0.0046)	0.1262(0.0138)	0.1341(0.0024)
$\partial V/\partial\sigma_2$	0.0929(0.0042)	0.0984(0.0140)	0.0917(0.0025)
$\partial V/\partial\sigma_3$	0.1418(0.0034)	0.1339(0.0164)	0.1412(0.0020)
$\partial V/\partial\sigma_4$	-0.0690(0.0020)	-0.0732(0.0059)	-0.0662 (0.0016)
corr.delta			
$\partial V/\partial\rho_{1,2}$	0.0264 (0.0011)	0.0259(0.0044)	0.0261(0.0014)
$\partial V/\partial\rho_{1,3}$	0.0565 (0.0017)	0.0600(0.0058)	0.0586(0.0011)
$\partial V/\partial\rho_{1,4}$	0.0361 (0.0021)	0.0368(0.0048)	0.0381(0.0019)
$\partial V/\partial\rho_{2,3}$	0.0416 (0.0016)	0.0442(0.0053)	0.0436(0.0011)
$\partial V/\partial\rho_{2,4}$	0.0208 (0.0025)	0.0230(0.0042)	0.0236(0.0015)
$\partial V/\partial\rho_{3,4}$	0.0636 (0.0035)	0.0613(0.0044)	0.0628(0.0022)

2.3 Altiplano Option

The Altiplano option is another basket option with a payoff based on the performance of assets in the basket. It combines the characteristics of a barrier option and an Asian option. The underlying assets are observed at discrete time steps $0 = t_0 < t_1 < ... < t_M = T$. The option entitles the holder to receive a large fixed coupon C at maturity T, if none of the assets in the basket falls below the barrier L before expiry. Otherwise the payoff is the vanilla call on the basket. Let $X(t_m) = (X_1(t_m), X_2(t_m), ..., X_n(t_m))'$, denote the performance of the n assets in the basket at time t_m, for $0 < m \leq M$, and we write the terminal payoff $\phi : \mathbb{R}_+^n \to \mathbb{R}$ in terms of performance processes:

$$\phi(X_1(T), X_2(T), ..., X_n(T)) = \chi \max\left(\Sigma_{i=1}^n X_i(T) - K, 0\right) + (1-\chi)C, \tag{2.24}$$

where K is the strike price and χ is an indicator function given as:

$$\chi = \begin{cases} 1 & \text{if } \min_{1\leq i\leq n, 0<m\leq M}\{X_i(t_m)\} \leq L, \\ 0 & \text{otherwise.} \end{cases}$$

There is no known analytical formula for the Altiplano option price under the Black-Scholes model. We estimate the value of the Altiplano option by the Monte Carlo method and the quasi-Monte Carlo method, and compare them in Table 2.10.

While estimating Greeks for the Altiplano option, we observe that the payoff function (2.24) is discontinuous. The pathwise method is not applicable but the likelihood ratio method still applies. As mentioned in Chapter 1, several techniques have been proposed to circumvent the difficulty of discontinuous payoffs, such as the smoothing payoff technique, the kernel method, the conditional Monte Carlo method, etc. In this section, we estimate Greeks of Altiplano option by the kernel method and the likelihood ratio method, and compare them numerically.

2.3.1 Kernel Method

We recall that the kernel method is applicable when the payoff function can be written in the form

$$\varphi(S) = g(S) \cdot 1_{\{h(S)\geq 0\}},$$

where functions $g(S)$ and $h(S)$ satisfy Assumptions 1 and 2 in Theorem 4 in [52]. Therefore, we first write the Altiplano payoff function (2.24) in this form. Let $\breve{X}_i = \min_{t_0<t_m\leq t_M}\{X_i(t_m)\}$, denote the

worst performance of the ith asset during its lifespan. Then the payoff function can be rewritten as

$$\begin{aligned}\phi(X) &= e^{-rT}C \cdot 1_{\left\{\min_{1\le i\le n}\{\check{X}_i\}-L>0\right\}} \\ &\quad + e^{-rT}(\Sigma_{i=1}^n X_i(T) - K) \cdot 1_{\left\{\min\left\{\sum_{i=1}^n X_i(T)-K, L-\min_{1\le i\le n}\{\check{X}_i\}\right\}>0\right\}}.\end{aligned} \tag{2.25}$$

We break Eqn.(2.25) into two parts. The first part consists of $g_1(X) = e^{-rT}C$ and $h_1(X) = \min_{1\le i\le n}\{\check{X}_i\} - L$. The second part consists of $g_2(X) = e^{-rT}\left(\sum_{i=1}^n X_i(T) - K\right)^+$ and $h_2(X) = L - \min_{1\le i\le n}\{\check{X}_i\}$. Let $V_1 = E[g_1(X) \cdot 1_{\{h_1(X)\ge 0\}}]$ and $V_2 = E[g_2(X) \cdot 1_{\{h_2(X)\ge 0\}}]$. The Altiplano option price is the sum of V_1 and V_2.

Assumption 1 is a typical assumption in the pathwise sensitivity estimation. As $X_i(t)$ follows a geometric Brownian motion, $X_i(t)$ is continuously differentiable and $\|X_{\theta+\Delta\theta} - X_\theta\| \le K_s|\theta|$ for some random variables K_s as Broadie and Glasserman prove in [13]. Therefore, to satisfy Assumption 1, we need g_1, h_1, g_2, h_2 to be Lipschitz continuous and differentiable almost everywhere. Since g_1 and g_2 are simple functions, they are continuously differentiable almost everywhere. The functions h_1, h_2 are composition of minimum functions, and we have proved that they are Lipschitz continuous in Proposition 10. Therefore, Assumption 1 is satisfied for the Altiplano option. Assumption 2 is satisfied since the normal density function is infinitely differentiable with respect to the parameter of volatility σ, interest rate r, and its random variables. In conclusion, the kernel method is applicable to the Altiplano option. We derive the estimators for Greeks as follows.

Delta and Gamma. The delta of the Altiplano option is the zero vector. The gamma is the zero matrix. The reason is the same as we stated in Proposition 7 in Section 2.1.2, that is, the derivative of the performance $X_i(t)$ with respect to its initial asset price $S_i(0)$ is zero under the Black-Scholes setting, and the Altiplano payoff (2.24) can be expressed in terms of performance, without initial asset variables.

In the following proposition, we derive the kernel estimator for Greek vega of the Altiplano option, i.e., the sensitivity of the option price to the change in the volatility.

Proposition 12 *The kernel estimator of the **vega** with respect to the jth asset is*

$$\begin{aligned}\bar{G}_N &= e^{-rT}\frac{1}{N}\sum_{l=1}^N X'_{j,l}(T) 1_{\{\sum_{i=1}^n X_{i,l}(T)-K\ge 0\}} 1_{\{\min_{1\le i\le n}\{\check{X}_{i,l}\}\le L\}} \\ &\quad - e^{-rT}\frac{1}{N\delta_N}\sum_{l=1}^N \left(\left((\Sigma_{i=1}^n X_{i,l}(T) - K)^+ - C\right)\frac{d\check{X}_{j,l}}{d\sigma_j} 1_{\{\min_i\{\check{X}_{i,l}\}=\check{X}_{j,l}\}} Z\left(\frac{L-\check{X}_{j,l}}{\delta_N}\right)\right),\end{aligned}$$

where $Z(\cdot)$ is the normal kernel density function with the bandwidth δ_N, and $(X_{j,l}(T)$, $X'_{j,l}(T)$, $\check{X}_{j,l}$, $\check{X}'_{j,l})$ denotes the lth sample of $(X_j(T), \frac{dX_j(T)}{d\sigma_j}, \check{X}_j, \frac{d\check{X}_j}{d\sigma_j})$ from the Monte Carlo simulation.

Proof. By Theorem 4 of Liu and Hong [52],

$$\frac{\partial V_1}{\partial \sigma_j} = E\left[\frac{\partial g_1(X)}{\partial \sigma_j} \cdot 1_{\{h_1(X)\geq 0\}}\right] - \partial_y E\left[g_1(X)\frac{\partial h_1(X)}{\partial \sigma_j} \cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0}. \tag{2.26}$$

We put $g_1(X) = e^{-rT}C$, $\frac{\partial g_1(X)}{\partial \sigma_j} = 0$ and $h_1(X) = \min_{1\leq i\leq n}\{\check{X}_i\} - L$ into Eqn.(2.26) and get

$$\frac{\partial V_1}{\partial \sigma_j} = -\partial_y E\left[e^{-rT}C\frac{\partial h_1(X)}{\partial \sigma_j} \cdot 1_{\{\min_{1\leq i\leq n}\{\check{X}_i\}-L\geq y\}}\right]\Bigg|_{y=0}. \tag{2.27}$$

If $\check{X}_j$ reaches the minimum value at t_m,

$$\frac{\partial h_1(X)}{\partial \sigma_j} = \frac{d\check{X}_j}{d\sigma_j} \cdot 1_{\{\min_i \check{X}_i=\check{X}_j\}},$$

where

$$\frac{d\check{X}_j}{d\sigma_j} = X_j(t_m)\left(-\sigma_j t_m + L_j\sum_{k=1}^{m}\sqrt{t_k - t_{k-1}}Z_k\right).$$

Let Z_k be the normal distributed n-dimensional vector, for $k = 1, ..., M$, and L_j the jth row of the Cholesky decomposition matrix. We have

$$\frac{\partial V_2}{\partial \sigma_j} = E\left[\frac{\partial g_2(X)}{\partial \sigma_j} \cdot 1_{\{h_2(X)\geq 0\}}\right] - \partial_y E\left[g_2(X)\frac{\partial h_2(X)}{\partial \sigma_j} \cdot 1_{\{h_2(X)\geq y\}}\right]\Bigg|_{y=0}, \tag{2.28}$$

where $g_2(X) = e^{-rT}\left(\sum_{i=1}^{n} X_i(T) - K\right)^+$ and $h_2(X) = L - \min_{1\leq i\leq n}\{\check{X}_i\}$. The derivative of $g_2(X)$ and $h_2(X)$ with respect σ_j are

$$\frac{\partial g_2(X)}{\partial \sigma_j} = e^{-rT} \cdot 1_{\{\sum_{i=1}^{n} X_i(T)-K\geq 0\}}\frac{dX_j(T)}{d\sigma_j} \tag{2.29}$$

and

$$\frac{\partial h_2(X)}{\partial \sigma_j} = -\frac{\partial h_1(X)}{\partial \sigma_j}. \tag{2.30}$$

Eqn.(2.28) can be rewritten as

$$\begin{aligned}\frac{\partial V_2}{\partial \sigma_j} &= E\left[e^{-rT} \cdot 1_{\{\sum_{i=1}^{n} X_i(T)-K\geq 0\}}\frac{dX_j(T)}{d\sigma_j} \cdot 1_{\{h_2(X)\geq 0\}}\right] \\ &\quad + \partial_y\ E\left[e^{-rT}\left(\Sigma_{i=1}^{n} X_i(T) - K\right)^+\frac{\partial h_1(X)}{\partial \sigma_j} \cdot 1_{\{L-\min_{1\leq i\leq n}\{\check{X}_i\}\geq y\}}\right]\Bigg|_{y=0}.\end{aligned} \tag{2.31}$$

The vega of the Altiplano option is the sum of the derivatives:

$$
\begin{aligned}
\frac{\partial V}{\partial \sigma_j} &= \frac{\partial V_1}{\partial \sigma_j} + \frac{\partial V_2}{\partial \sigma_j} \\
&= E\left[e^{-rT}\frac{dX_j(T)}{d\sigma_j}\cdot 1_{\{\sum_{i=1}^n X_i(T)-K\geq 0\}}\cdot 1_{\{\min_{1\leq i\leq n}\check{X}_i\leq L\}}\right] \\
&+ \partial_y\, E\left[e^{-rT}((\Sigma_{i=1}^n X_i(T)-K)^+ - C)\frac{\partial h_1(X)}{\partial \sigma_j}\cdot 1_{\{L-\min_{1\leq i\leq n}\{\check{X}_i\}\geq y\}}\right]\Bigg|_{y=0}.
\end{aligned}
\tag{2.32}
$$

By the definition of derivative, the second part of R.H.S of Eqn.(2.32)

$$
\begin{aligned}
&\partial_y\, E\left[e^{-rT}((\Sigma_{i=1}^n X_i(T)-K)^+ - C)\frac{\partial h_1(X)}{\partial \sigma_j}\cdot 1_{\{L-\min_{1\leq i\leq n}\{\check{X}_i\}\geq y\}}\right]\Bigg|_{y=0} \\
&= -\lim_{\delta\to 0}\frac{1}{\delta}E\left[e^{-rT}((\Sigma_{i=1}^n X_i(T)-K)^+ - C)\frac{d\check{X}_j}{d\sigma_j}\cdot 1_{\{\min_i \check{X}_i=\check{X}_j\}}U\left(\frac{\check{X}_j-L}{\delta}\right)\right]
\end{aligned}
$$

where $U(u) = 1_{\{-0.5\leq u\leq 0.5\}}$ is the density function of uniform $(-1/2, 1/2)$ distribution. We use the Monte Carlo simulation to estimate the expectations in Eqn.(2.32). Let $(X_{j,l}(T)$, $X'_{j,l}(T)$, $\check{X}_{j,l}$, $\check{X}'_{j,l})$ denote the lth observation of $(X_j(T), \frac{dX_j(T)}{d\sigma_j}, \check{X}_j, \frac{d\check{X}_j}{d\sigma_j})$, for $l = 1, 2, ..., N$. The kernel estimator of $\frac{dV}{d\sigma_j}$ is given by

$$
\begin{aligned}
\bar{G}_N &= e^{-rT}\frac{1}{N}\sum_{l=1}^{N}\frac{dX_{j,l}(T)}{d\sigma_j}\cdot 1_{\{\sum_{i=1}^n X_{i,l}(T)-K\geq 0\}}\cdot 1_{\{\min_{1\leq i\leq n}\check{X}_{i,l}\leq L\}} \\
&- e^{-rT}\frac{1}{N\delta_N}\sum_{l=1}^{N}\left[\left((\Sigma_{i=1}^n X_{i,l}(T)-K)^+ - C\right)\frac{d\check{X}_{j,l}}{d\sigma_j}\cdot 1_{\{\min_i \check{X}_{i,l}=\check{X}_{j,l}\}}U\left(\frac{L-\check{X}_{j,l}}{\delta_N}\right)\right],
\end{aligned}
\tag{2.33}
$$

where δ_N is called a bandwidth parameter in the kernel estimation. The choice of kernel is not limited to uniform density function. A range of kernel function are commonly used, such as triangular, biweight, normal, as long as it satisfies that $uU(u)\to 0$ as $|u|\to\infty$ and $\int_{-\infty}^{\infty} uU(u)du < \infty$. In our numerical examples, we use the standard normal density function $Z(u)$ as our kernel instead of the uniform one, because the estimator is more robust with a smooth kernel [8]. We write the second part of the Eqn.(2.33) as a function $\bar{\Gamma}(u)$:

$$
\bar{\Gamma}(u) = e^{-rT}\frac{1}{N\delta_N}\sum_{l=1}^{N}\left[\left((\Sigma_{i=1}^n X_{i,l}(T)-K)^+ - C\right)\frac{d\check{X}_{j,l}}{d\sigma_j}1_{\{\min_i \check{X}_{i,l}=\check{X}_{j,l}\}}Z\left(\frac{L-\check{X}_{j,l}+u}{\delta_N}\right)\right].
$$

In order to the minimize the mean square error of $\bar{\Gamma}$, the asymptotically optimal bandwidth is chosen as $\delta_N^* = c\cdot N^{-1/5}$, where c could be obtained by a pilot simulation. We start with $\delta_N = N^{-1/5}$ and approximate c by

$$\hat{c} = \left(\frac{Var(\bar{\Gamma}(0))}{4\left(\bar{\Gamma}''(0)\int_{-\infty}^{\infty} u^2 Z(u)du\right)^2} \right)^{\frac{1}{5}}, \tag{2.34}$$

and $\int_{-\infty}^{\infty} u^2 Z(u)du = 1$, when $Z(u)$ is chosen as the standard normal density function; $Var(\bar{\Gamma}(0))$ is the sample variance of $\bar{\Gamma}(0)$, and $\bar{\Gamma}''(0)$ is the central difference of the second derivative of $\bar{\Gamma}(0)$:

$$\bar{\Gamma}''(0) = \frac{\bar{\Gamma}(u) + \bar{\Gamma}(-u) - 2\bar{\Gamma}(0)}{s^2}$$

where s is the step size of the finite difference approach. Then we have the new $\delta_N = \hat{c} \cdot N^{-1/5}$ and use it to estimate $\bar{V}_N$ and $\bar{\Gamma}''(0)$. Repeating this procedure for a number of times, we remove the dependence on the initial choice of δ_N when estimating $\hat{c}$ by (2.34). We could achieve this by setting a desired error tolerance to get a stable value of $\hat{c}$ and multiply it by $N^{-1/5}$ as the optimal choice of bandwidth. See details about selection of bandwidth δ_N in the electronic companion of Liu and Hong [52]. ■

Proposition 13 *The Greek **rho** of the Altiplano option estimated via kernel method is given as:*

$$\begin{aligned} \bar{G}_N = & \frac{1}{N}\sum_{l=1}^{N}\left(-CTe^{-rT}1_{\{h_{1,l}(X)\geq 0\}} + \frac{dg_{2,l}(X)}{dr}1_{\{h_{2,l}(X)\geq 0\}}\right) \\ & - \frac{1}{N\delta_N}\sum_{l=1}^{N}\left(e^{-rT}\left((\Sigma_{i=1}^{n}X_{i,l}(T)-K)^+ - C\right)\frac{dh_{1,l}(X)}{dr}Z\left(\frac{h_{1,l}(X)}{\delta_N}\right)\right), \end{aligned} \tag{2.35}$$

where $(h_{1,l}(X), h_{2,l}(X), \frac{dg_{2,l}(X)}{dr}, \frac{dh_{1,l}(X)}{dr})$ are the lth sample of $(h_1(X),\ h_2(X),\ \frac{dg_2(X)}{dr},\ \frac{dh_1(X)}{dr})$, with $l = 1, 2, ..., N$. Let Z denote the normal kernel function with the bandwidth $\delta_N > 0$. The optimal value of δ_N is obtained by a pilot simulation.

Proof. The derivation of Eqn.(2.35) is the same as in Proposition 12. We denote $V_1 = E[g_1(X) \cdot 1_{\{h_1(X)\geq 0\}}]$ and $V_2 = E[g_2(X) \cdot 1_{\{h_2(X)\geq 0\}}]$, where $g_1(X) = e^{-rT}C$, $h_1(X) = \min_{1\leq i\leq n}\{\check{X}_i\} - L$, $g_2(X) = e^{-rT}\left(\sum_{i=1}^{n} X_i(T) - K\right)^+$ and $h_2(X) = L - \min_{1\leq i\leq n}\{\check{X}_i\}$. When X_i reaches its minimum value $\check{X}_i$ at t_m,

$$\frac{dh_1(X)}{\partial r} = \sum_{j=1}^{n} \check{X}_j t_m \cdot 1_{\{\min_{1\leq i\leq n}\{\check{X}_i\}=\check{X}_j\}}, \tag{2.36}$$

and then

$$\frac{dV_1}{dr} = -CTe^{-rT}E\left[1_{\{h_1(X)\geq 0\}}\right] - \partial_y E\left[e^{-rT}C\frac{dh_1(X)}{\partial r}\cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0}. \tag{2.37}$$

The derivative of V_2 with respect to r

$$\frac{dV_2}{dr} = E\left[\frac{dg_2(X)}{dr}\cdot 1_{\{h_2(X)\geq 0\}}\right] - \partial_y\ E\left[g_2(X)\frac{dh_2(X)}{dr}\cdot 1_{\{h_2(X)\geq y\}}\right]\Bigg|_{y=0} \tag{2.38}$$

where the derivative

$$\frac{dh_2(X)}{dr} = -\frac{dh_1(X)}{dr}, \tag{2.39}$$

and

$$\frac{dg_2(X)}{dr} = -Te^{-rT}\left(\Sigma_{i=1}^n X_i(T) - K\right)^+ + e^{-rT}\cdot 1_{\{\Sigma_{i=1}^n X_i(T)-K\geq 0\}}\sum_{i=1}^n \frac{dX_i(T)}{dr}, \tag{2.40}$$

Thus, the derivative (2.38) is written as follows:

$$\frac{dV_2}{dr} = E\left[\frac{dg_2(X)}{dr}\cdot 1_{\{h_2(X)\geq 0\}}\right] + \partial_y\ E\left[e^{-rT}\left(\Sigma_{i=1}^n X_i(T) - K\right)^+\frac{dh_1(X)}{dr}\cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0}. \tag{2.41}$$

The Greek rho equals to the sum of Eqn.(2.37) and Eqn.(2.41):

$$\begin{aligned}
\frac{dV}{dr} &= -CTe^{-rT}E[1_{\{h_1(X)\geq 0\}}] + E\left[\frac{dg_2(X)}{dr}\cdot 1_{\{h_2(X)\geq 0\}}\right] \\
&\quad + \partial_y\ E\left[e^{-rT}\left((\Sigma_{i=1}^n X_i(T) - K)^+ - C\right)\frac{dh_1(X)}{dr}\cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0} \\
&= E\left[-CTe^{-rT}1_{\{h_1(X)\geq 0\}} + \frac{dg_2(X)}{dr}\cdot 1_{\{h_2(X)\geq 0\}}\right] \\
&\quad + \partial_y\ E\left[e^{-rT}\left((\Sigma_{i=1}^n X_i(T) - K)^+ - C\right)\frac{dh_1(X)}{dr}\cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0}.
\end{aligned} \tag{2.42}$$

The expectation in Eqn.(2.42) is estimated by Monte Carlo simulation:

$$\begin{aligned}
\bar{G}_N &= \frac{1}{N}\sum_{l=1}^N\left(-CTe^{-rT}1_{\{h_{1,l}(X)\geq 0\}} + \frac{dg_{2,l}(X)}{dr}1_{\{h_{2,l}(X)\geq 0\}}\right) \\
&\quad - \frac{1}{N\delta_N}\sum_{l=1}^N\left(e^{-rT}((\Sigma_{i=1}^n X_{i,l}(T) - K)^+ - C)\frac{dh_{1,l}(X)}{dr}Z\left(\frac{h_{1,l}(X)}{\delta_N}\right)\right),
\end{aligned}$$

where $h_{1,l}(X)$ and $h_{2,l}(X)$ are the lth sample of $(h_1(X)$ and $h_2(X)$ respectively, with $l = 1, 2, ..., N$. The derivative $\frac{dh_{1,l}(X)}{dr}$ is estimated through Eqn.(2.36) and $\frac{dg_{2,l}(X)}{dr}$ is estimated through Eqn.(2.40). Let Z denote the normal kernel function with the bandwidth $\delta_N > 0$. The optimal value of δ_N is obtained by a pilot simulation.

■

Proposition 14 *The Greek* ***theta*** *of the Altiplano option estimated via kernel method is given as:*

$$\begin{aligned}\bar{G}_N &= \frac{1}{N}\sum_{l=1}^{N}\left(-re^{-rT}C\cdot 1_{\{h_{1,l}(X)\geq 0\}} + \frac{dg_{2,l}(X)}{dT}\cdot 1_{\{h_{2,l}(X)\geq 0\}}\right)\\ &- e^{-rT}\frac{1}{N\delta_N}\sum_{l=1}^{n}\left(\left((\Sigma_{i=1}^{n}X_{i,l}(T)-K)^{+}-C\right)\frac{dh_{1,l}(X)}{dT}Z\left(\frac{\min_i\{\check{X}_{i,l}\}-L}{\delta_N}\right)\right),\end{aligned}$$

where $(h_{1,l}(X), h_{2,l}(X), \frac{dg_{2,l}(X)}{dr}, \frac{dh_{1,l}(X)}{dr})$ *are the* l*th sample of* $(h_1(X),\, h_2(X),\, \frac{dg_2(X)}{dr},\, \frac{dh_1(X)}{dr})$*, with* $l = 1, 2, ..., N$*. Let* Z *denote the normal kernel function with the bandwidth* $\delta_N > 0$*. The optimal value of* δ_N *is obtained by a pilot simulation.*

Proof. We use the same idea to derive Greek theta. The derivative of V_1 with respect to the maturity

$$\frac{dV_1}{dT} = E\left[-rCe^{-rT}\cdot 1_{\{h_1(X)\geq 0\}}\right] - \partial_y\, E\left[e^{-rT}C\frac{dh_1(X)}{dT}\cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0}, \tag{2.43}$$

where $g_1(X) = e^{-rT}C$, $h_1(X) = \min_{1\leq i\leq n}\{\check{X}_i\} - L$, and $\frac{dg_1(X)}{dT} = -rCe^{-rT}$. If X_j reaches the minimum at t_m, the derivatives

$$\begin{aligned}\frac{dh_1(X)}{dT} &= \sum_{j=1}^{n}\frac{d\check{X}_j}{dT}\cdot 1_{\{\min_i \check{X}_i = X_j\}},\\ &= \sum_{j=1}^{n}X_j(t_m)\left((r-\sigma_j^2/2) + \frac{\sigma_j}{2\sqrt{Tm}}L_j\sum_{k=1}^{m}Z_k\right)\cdot 1_{\{\min_i \check{X}_i = X_j\}},\end{aligned} \tag{2.44}$$

where Z_k is the normal distributed n-dimensional vector. L_j is the jth row of the Cholesky decomposition matrix. The derivative of the second part

$$\frac{dV_2}{dT} = E\left[\frac{dg_2(X)}{dT}\cdot 1_{\{h_2(X)\geq 0\}}\right] + \partial_y\, E\left[e^{-rT}(\Sigma_{i=1}^{n}X_i(T)-K)^{+}\frac{dh_1(X)}{dT}\cdot 1_{\{h_2(X)\geq y\}}\right]\Bigg|_{y=0}, \tag{2.45}$$

where $g_2(X) = e^{-rT}\left(\sum_{i=1}^{n}X_i(T)-K\right)^{+}$, $h_2(X) = L - \min_{1\leq i\leq n}\{\check{X}_i\}$ and

$$\frac{dh_2(X)}{dT} = -\frac{dh_1(X)}{dT}.$$

We calculate the derivatives in Eqn.(2.45) and obtain

$$\frac{dg_2(X)}{dT} = e^{-rT}\cdot 1_{\{\sum_{i=1}^{n}X_i(T)-K\geq 0\}}\sum_{i=1}^{n}\frac{dX_i(T)}{dT} - re^{-rT}\left(\Sigma_{i=1}^{n}X_i(T)-K\right)^{+}, \tag{2.46}$$

where

$$\frac{dX_i(T)}{dT} = X_i(T)\left(\left(r - \sigma_i^2/2\right) + \frac{\sigma_i}{2\sqrt{TM}} L_i \sum_{k=1}^{M} Z_k\right),$$

and Z_k and L_i are the same as defined above. The Greek theta is the sum of Eqn.(2.43) and Eqn.(2.45)

$$\begin{aligned}\frac{dV}{dT} &= E\left[-rCe^{-rT}\cdot 1_{\{h_1(X)\geq 0\}} + \frac{dg_2(X)}{dT}\cdot 1_{\{h_2(X)\geq 0\}}\right] \\ &\quad - \lim_{\delta\to 0}\frac{1}{\delta}E\left[e^{-rT}\left((\Sigma_{i=1}^n X_i(T) - K)^+ - C\right)\frac{dh_1(X)}{dT}Z\left(\frac{\min_i \check{X}_i - L}{\delta}\right)\right].\end{aligned} \tag{2.47}$$

The kernel estimator of Eqn.(2.47) is estimated by Monte Carlo simulation as follows:

$$\begin{aligned}\bar{G}_N &= \frac{1}{N}\sum_{l=1}^{N}\left(-re^{-rT}C\cdot 1_{\{h_{1,l}(X)\geq 0\}} + \frac{dg_{2,l}(X)}{dT}\cdot 1_{\{h_{2,l}(X)\geq 0\}}\right) \\ &\quad - e^{-rT}\frac{1}{N\delta_N}\sum_{l=1}^{n}\left(\left((\Sigma_{i=1}^n X_{i,l}(T) - K)^+ - C\right)\frac{dh_{1,l}(X)}{dT}Z\left(\frac{\min_i \check{X}_{i,l} - L}{\delta_N}\right)\right),\end{aligned}$$

where $h_{1,l}(X)$ and $h_{2,l}(X)$ are the lth sample of ($h_1(X)$ and $h_2(X)$ respectively, with $l = 1, 2, ..., N$. The derivative $\frac{dh_{1,l}(X)}{dr}$ is sampled from Eqn.(2.44) and $\frac{dg_{2,l}(X)}{dr}$ is sampled from Eqn.(2.46). Let Z denote the normal kernel function with the bandwidth $\delta_N > 0$. The optimal value of δ_N is obtained by a pilot simulation. ■

Proposition 15 *The kernel estimator of the* ***correlation delta*** *with respect to $\rho_{i,j}$ is given as follows:*

$$\begin{aligned}\bar{G}_N &= e^{-rT}\frac{1}{N}\sum_{l=1}^{N}\left(X'_{j,l}(T)\cdot 1_{\{\sum_{i=1}^n X_{i,l}(T)-K\geq 0\}}\cdot 1_{\{\min_i \check{X}_{i,l}\leq L\}}\right) \\ &\quad - e^{-rT}\frac{1}{N\delta_N}\sum_{l=1}^{n}\left[\left((\Sigma_{i=1}^n X_{i,l}(T) - K)^+ - C\right)\check{X}'_{j,l}\cdot 1_{\{\min_i \check{X}_{i,l}=\check{X}_{j,l}\}}Z\left(\frac{\check{X}_{j,l} - L}{\delta_N}\right)\right],\end{aligned}$$

where $Z(\cdot)$ denotes the standard normal density function and $(X_{j,l}(T), X'_{j,l}(T), \check{X}_{j,l}, \check{X}'_{j,l})$ denotes the lth sample of $(X_j(T), \frac{dX_j(T)}{d\rho_{i,j}}, \check{X}_j, \frac{d\check{X}_j}{d\rho_{i,j}})$, obtained from the Monte Carlo simulation.

Proof. The calculation of the Greek correlation delta is based on how to generate the independent sample performances. We do this in the same way that we calculate the pathwise estimators correlation delta for the Everest option in Section 2.1.3. We simply recap it here. When we are dealing with the correlation delta ($\partial V/\partial\rho_{i,j}$), we reorder the performances in the basket and place the performances X_i and X_j at the last two positions of the performance vector,

$$X = (X_1, ..., X_{i-1}, X_{i+1}, ..., X_{j-1}, X_{j+1}, ..., X_n, X_i, X_j).$$

If X_j reaches its minimum value $\check{X}_j$ at t_m, we have

$$\begin{aligned}\frac{d\check{X}_j}{d\rho_{i,j}} &= X_j(t_m)\sigma_j\frac{d}{d\rho_{i,j}}\left(L_j\cdot\Sigma_{k=1}^m\sqrt{t_k-t_{k-1}}Z_k\right)\\ &= X_j(t_m)\sigma_j\frac{d}{d\rho_{i,j}}\left(l_{n,n-1}\Sigma_{k=1}^m(\sqrt{t_k-t_{k-1}}Z_{k,n-1})+l_{n,n}\Sigma_{k=1}^m(\sqrt{t_k-t_{k-1}}Z_{k,n})\right)\\ &= X_j(t_m)\sigma_j\left(\frac{dl_{n,n-1}}{d\rho_{i,j}}\Sigma_{k=1}^m(\sqrt{t_k-t_{k-1}}Z_{k,n-1})+\frac{dl_{n,n}}{d\rho_{i,j}}\Sigma_{k=1}^m(\sqrt{t_k-t_{k-1}}Z_{k,n})\right)\\ &= \frac{X_j(t_m)\sigma_j}{l_{n-1,n-1}}\left(\Sigma_{k=1}^m(\sqrt{t_k-t_{k-1}}Z_{k,n-1})-\frac{l_{n,n-1}}{l_{n,n}}\Sigma_{k=1}^m(\sqrt{t_k-t_{k-1}}Z_{k,n})\right),\end{aligned} \tag{2.48}$$

where Z_k is the normal distributed n-dimensional vector, and $Z_{k,n}$ is the nth elements of it. L_j is the jth row of Cholesky decomposition matrix. In the same way, we obtain $\frac{dX_j(T)}{d\rho_{i,j}}$ by replacing t_m and m by T and M. The derivative of dV_1 with respect to $\rho_{i,j}$

$$\frac{dV_1}{d\rho_{i,j}} = -\partial_y\, E\left[e^{-rT}C\frac{dh_1(X)}{d\rho_{i,j}}\cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0} \tag{2.49}$$

where $g_1(X) = e^{-rT}C$, $h_1(X) = \min_{1\leq i\leq n}\{\check{X}_i\} - L$, and $\frac{dg_1(X)}{d\rho_{i,j}} = 0$. The derivative of $h_1(X)$ with respect to $\rho_{i,j}$ is given by

$$\frac{dh_1(X)}{d\rho_{i,j}} = \frac{d\check{X}_j}{d\rho_{i,j}}\cdot 1_{\{\min_i \check{X}_i = X_j\}}.$$

The derivative of V_2

$$\frac{dV_2}{d\rho_{i,j}} = E\left[\frac{dg_2(X)}{d\rho_{i,j}}\cdot 1_{\{h_2(X)\geq 0\}}\right] + \partial_y\, E\left[g_2(X)\frac{dh_1(X)}{d\rho_{i,j}}\cdot 1_{\{h_2(X)\geq y\}}\right]\Bigg|_{y=0}, \tag{2.50}$$

given that $g_2(X) = e^{-rT}(\Sigma_{i=1}^n X_i(T) - K)^+$ and its derivative with respect to $\rho_{i,j}$

$$\frac{dg_2(X)}{d\rho_{i,j}} = e^{-rT}\frac{dX_j(T)}{d\rho_{i,j}}\cdot 1_{\{\sum_{i=1}^n X_i(T)-K\geq 0\}}.$$

Since $h_2(X) = -h_1(X)$, we have

$$\frac{dh_2(X)}{d\rho_{i,j}} = -\frac{dh_1(X)}{d\rho_{i,j}}.$$

We add Eqn.(2.49) to Eqn.(2.50) and get the Greek correlation delta,

$$\begin{aligned}\frac{dV}{d\rho_{i,j}} &= E\left[e^{-rT}\frac{dX_j(T)}{d\rho_{i,j}}\cdot 1_{\{\sum_{i=1}^n X_i(T)-K\geq 0\}}1_{\{h_2(X)\geq 0\}}\right]\\ &\quad -\lim_{\delta\to 0}\frac{1}{\delta}E\left[e^{-rT}\left((\Sigma_{i=1}^n X_i(T)-K)^+ - C\right)\frac{d\check{X}_j}{d\rho_{i,j}}\cdot 1_{\{\min_i \check{X}_i = X_j\}}Z\left(\frac{\min_i \check{X}_i - L}{\delta}\right)\right].\end{aligned}$$

In conclusion, the kernel estimator of Greek correlation delta via the Monte Carlo simulation is

$$\bar{G}_N = e^{-rT}\frac{1}{N}\sum_{l=1}^{N}\left(\frac{dX_{j,l}(T)}{d\rho_{i,j}}1_{\{\sum_{i=1}^n X_{i,l}(T)-K\geq 0\}}1_{\{h_{2,l}(X)\geq 0\}}\right)$$
$$-\,e^{-rT}\frac{1}{N\delta_N}\sum_{l=1}^{N}\left(\left((\Sigma_{i=1}^n X_{i,l}(T)-K)^+-C\right)\frac{d\check{X}_{j,l}}{d\rho_{i,j}}\cdot 1_{\{\min_i \check{X}_{i,l}=X_{j,l}\}}Z\left(\frac{\min_i \check{X}_{i,l}-L}{\delta_N}\right)\right),$$

where $h_{2,l}(X)$ is the lth sample of $h_2(X)$, with $l = 1, 2, ..., N$. We generate the samples for $\frac{d\check{X}_{j,l}}{d\rho_{i,j}}$ and $\frac{dX_{j,l}(T)}{d\rho_{i,j}}$ according to Eqn.(2.48). The optimal value of δ_N is obtained by a pilot simulation. ■

2.3.2 Likelihood Ratio Method

The performance vector $X(t) = (X_1(t), X_2(t), ..., X_n(t))'$ follows the geometric Brownian motion. At any time t_k, the performance $X(t_k)$ has the distribution of $\exp(Y(t_k))$, where $Y(t_k)$ follows the normal distribution $\mathcal{N}(\mu(t_k), \tilde{\Sigma}(t_k))$. The vector mean $\mu(t_k)$ has elements $\mu_i(t_k) = \ln(X_i(t_{k-1})) + (r - \sigma_i^2/2)(t_k - t_{k-1})$, $i = 1, ...n$. Covariance matrix has the form $\tilde{\Sigma}(t_k) = (t_k - t_{k-1})A\Sigma A'$, where $A = diag(\sigma_1, ..., \sigma_k)$. The $Y(t)$ is an n-dimensional Brownian motion. By the Markov property of Brownian motion, we factor its density as

$$g(Y(t_1), Y(t_2), ..., Y(t_m)) = g_1(Y(t_1)|Y(t_0))g_2(Y(t_2)|Y(t_1)) \cdot ... \cdot g_m(Y(t_m)|Y_{m-1}), \tag{2.51}$$

where each $g_k(Y(t_k)|Y(t_{k-1}))$ is the transition density from t_{k-1} to time t_k. The probability density function of $Y(t_m)$ is written as

$$g_k(\theta; y(t_k)|y(t_{k-1})) = \frac{1}{\sqrt{(2\pi)^n|\tilde{\Sigma}(\theta; t_k)|}}\exp\left(-\frac{1}{2}(y(t_k) - \mu(\theta; t_k))'\tilde{\Sigma}^{-1}(\theta; t_k)(y(t_k) - \mu(\theta; t_k))\right). \tag{2.52}$$

We generate the sample paths $y(t_k)$ by $\mu(t_k) + \sqrt{t_k - t_{k-1}}ALZ_k$, where L is the Cholesky matrix such that $LL' = \Sigma$, and $Z_k \sim N(0, I)$. Then the density function Eqn.(2.51) can be written as

$$g(\theta; y(t_1), ..., y(t_m)) = \prod_{k=1}^{m}\frac{1}{\sqrt{(2\pi)^n|\tilde{\Sigma}(\theta; t_k)|}}\exp\left(-\frac{1}{2}(y(t_k) - \mu(\theta; t_k))'\tilde{\Sigma}^{-1}(\theta; t_k)(y(t_k) - \mu(\theta; t_k))\right).$$

The score function is

$$\frac{\partial \log g(y(t_1), y(t_2), ..., y(t_m))}{\partial\theta} = \sum_{k=1}^{m}\frac{\partial \log g_k(y(t_k)|y(t_{k-1}))}{\partial\theta}.$$

Rho. We switch the order of the differentiation and integration due to the continuity of the integrand with respect to r, and obtain

$$\begin{aligned}
\frac{d}{dr}E_r[e^{-rT}\phi(Y)] &= \int_{\mathcal{R}^n}\phi(y)\frac{d}{dr}\left(e^{-rT}g(y(t_1),y(t_2),...,y(t_m))\right)dy \\
&= E_r\left[e^{-rT}\phi(y)\left(-T+\frac{\dot{g}_r(y(t_1),y(t_2),...,y(t_m))}{g_r(y(t_1),y(t_2),...,y(t_m))}\right)\right], \\
&= E_r\left[e^{-rT}\phi(y)\left(-T+\sum_{k=1}^{m}\frac{d\log g_r(y(t_k)|y(t_{k-1}))}{dr}\right)\right], \\
&= E_r\left[e^{-rT}\phi(y)\left(-T+\sum_{k=1}^{m}\frac{\dot{g}_r(y(t_k)|y(t_{k-1}))}{g_r(y(t_k)|y(t_{k-1}))}\right)\right],
\end{aligned}$$

where the score function is

$$\sum_{k=1}^{m}\frac{\dot{g}_r(y(t_k)|y(t_{k-1}))}{g_r(y(t_k)|y(t_{k-1}))}=\sum_{k=1}^{m}\frac{\partial\log g(y(t_k)|\mathbf{y}(t_{k-1}))}{\partial r}=\sum_{k=1}^{m}Z_k'L^{-1}A^{-1}\sqrt{\Delta t_k}I_{n\times 1}.$$

Here $\Delta t_k = t_k - t_{k-1}$ and $I_{n\times 1}$ is an n-dimensional vector with all elements equal to 1. The likelihood ratio estimator for rho is

$$e^{-rT}\phi(Y)\cdot\left(-T+\sum_{k=1}^{m}Z_k'L^{-1}A^{-1}\sqrt{\Delta t_k}I_{n\times 1}\right).$$

Theta. The parameter of interest for Greek theta is maturity T. We write the transition density function Eqn.(2.52) is written as a function of T as follows:

$$\begin{aligned}
& g_k(y(t_k)|y(t_{k-1})) \\
= & \frac{1}{\sqrt{(2\pi)^n(\Delta t_k)^n|A\Sigma A'|}}\exp\left(-\frac{1}{2\Delta t_k}(\mathbf{y}(t_k)-\mu(t_k))'(A\Sigma A')^{-1}(y(t_k)-\mu(t_k))\right).
\end{aligned}$$

We take the derivative of the product of the discount factor and the density function with respect to maturity T. For simplicity, consider a uniform discretization where $\Delta t_k = \Delta t$ for all k. The score function is

$$\begin{aligned}
& -r+\frac{\dot{g}(y(t_1),y(t_2),...,y(t_m))}{g(y(t_1),y(t_2),...,y(t_m))} \\
= & -r+\sum_{k=1}^{m}\frac{d\log g(y(t_k)|y(t_{k-1}))}{dT} \\
= & -r+\frac{1}{m}\sum_{k=1}^{m}\left(-\frac{n}{2\Delta t}+\frac{1}{\sqrt{\Delta t}}Z_k'L^{-1}A^{-1}\left(rI_{n\times 1}-\frac{\sigma^2}{2}\right)+\frac{1}{2\Delta t}Z_k'Z_k.\right) \\
= & -r-\frac{n}{2\Delta t}+\frac{1}{m\sqrt{\Delta t}}\left(\sum_{k=1}^{m}Z_k'\right)L^{-1}A^{-1}\left(rI_{n\times 1}-\frac{\sigma^2}{2}\right)+\frac{1}{2T}\sum_{k=1}^{m}Z_k'Z_k.
\end{aligned}$$

The likelihood ratio estimator for theta is

$$e^{-rT}\phi(Y)\cdot\left(-r-\frac{nm}{2T}+\frac{1}{\sqrt{mT}}\left(\sum_{k=1}^{m}Z_k'\right)L^{-1}A^{-1}\left(rI_{n\times 1}-\frac{\sigma^2}{2}\right)+\frac{1}{2T}\sum_{k=1}^{m}Z_k'Z_k\right).$$

Vega. In order to express the parameters σ in the density function explicitly, we write the transition density function (2.52) as

$$g(Y(t_k)|Y(t_k))=\frac{1}{\sigma_1 ... \sigma_n\sqrt{(2\pi)^n\Delta t^n|\Sigma|}}\exp\left(-\frac{1}{2\Delta t}(A^{-1}(y-\mu))'\Sigma^{-1}A^{-1}(y-\mu)\right),$$

where

$$A^{-1}(Y-\mu)=\begin{pmatrix}\frac{1}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_n}\end{pmatrix}\begin{pmatrix}Y_1-\mu_1\\ Y_2-\mu_2\\ \vdots\\ Y_n-\mu_n\end{pmatrix}=\begin{pmatrix}\frac{Y_1-\mu_1}{\sigma_1}\\ \frac{Y_2-\mu_2}{\sigma_2}\\ \vdots\\ \frac{Y_n-\mu_n}{\sigma_n}\end{pmatrix},$$

and the derivative of $A^{-1}(Y-\mu)$ with respect to σ_i is

$$\frac{d}{d\sigma_i}\left(A^{-1}(Y-\mu)\right)=\left(0,...,\Delta t-\frac{1}{\sigma_i^2}(Y_i-\mu_i),...,0\right)'.$$

The score function is given by

$$\begin{aligned}\frac{\dot{g}(y(t_1),...,y(t_m))}{g(y(t_1),...,y(t_m))}&=\sum_{k=1}^{m}\frac{d\log g(y(t_k)|y(t_{k-1}))}{d\sigma_i}\\ &=\sum_{k=1}^{m}\left(-\frac{1}{\sigma_i}-\frac{1}{\Delta t}(A^{-1}(Y-\mu))'\Sigma^{-1}\frac{d}{d\sigma_i}\left(A^{-1}(Y-\mu)\right)\right)\\ &=\sum_{k=1}^{m}\left(-\frac{1}{\sigma_i}-\frac{1}{\Delta t}(A^{-1}(Y-\mu))'\Sigma^{-1}\left(0,...,\Delta t-\frac{1}{\sigma_i^2}(Y_i-\mu_i),...,0\right)'\right)\\ &=\sum_{k=1}^{m}\left(-\frac{1}{\sigma_i}-\frac{1}{\sqrt{\Delta t}}Z'L^{-1}\left(0,...,\Delta t-\frac{\sqrt{\Delta t}}{\sigma_i^2}(ALZ_k)_i,...,0\right)'\right).\end{aligned}$$

The likelihood ratio estimator for vega ith entry

$$e^{-rT}\phi(Y)\cdot\left(-\frac{m}{\sigma_i}-\sqrt{\frac{m}{T}}\sum_{k=1}^{m}Z_k'L^{-1}\left(0,...,\Delta t-\frac{\sqrt{\Delta t}}{\sigma_i^2}(ALZ_k)_i,...,0\right)'\right).$$

Correlation delta. The derivation of the correlation delta estimator involves calculating the derivative of the matrix Σ with respect to its entries $\rho_{i,j}$, for $i,j=1,....,n$ with $i\neq j$. The score function for the correlation delta of assets i and j is given as follows:

$$\begin{aligned}\frac{d\log g(y(t_k)|y(t_{k-1}))}{d\rho_{i,j}}&=-\frac{1}{2}Trace\left(\tilde{\Sigma}^{-1}\frac{\partial\tilde{\Sigma}}{\partial\rho_{i,j}}\right)+\frac{1}{2}(y_k-\mu)'\tilde{\Sigma}^{-1}\frac{\partial\tilde{\Sigma}}{\partial\rho_{i,j}}\tilde{\Sigma}^{-1}(y_k-\mu)\\ &=-\frac{1}{2}Trace\left(L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}\right)+\frac{1}{2}Z_k'L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}Z_k,\end{aligned}$$

where Z_k is a standard n-dimensional normal random vector which is used to generate the sample paths. The matrix $(\partial\Sigma/\partial\rho_{i,j})$ contains entries at (i,j) and (j,i) equal to 1, and 0 otherwise. In summary, the likelihood ratio estimator for the correlation delta between asset i and asset j is

$$e^{-rT}\phi(Y)\cdot\sum_{k=1}^{m}\left(-\frac{1}{2}Tr\left(L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}\right)+\frac{1}{2}Z_k'L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}Z_k\right).$$

2.3.3 Numerical Results

We use the same assets as in Everest option which are Citigroup (C), Freddie Mac (old ticker: FRE), J.P. Morgan Chase (JPM) and Legman Brothers (LEH), and each of the stocks had a zero dividend yield for the period. The maturity $T = 1$ year, and the risk-free rate was $r = 0.04$. The spot prices, volatilities and the correlation coefficient matrix are given in Table. 2.1 and Table. 2.2 respectively.

Table 2.10: Comparison of price and standard deviation for the Altiplano option between Monte Carlo and quasi-Monte Carlo simulation. The sample size is 100,000, with 40 independent runs.

	MC	QMC
Price	0.5794	0.5797
Std.Dev	0.0023	0.0013

Under the Monte Carlo simulation, the Altiplano option price is 0.5794 with a standard deviation of 0.0023; under the quasi-Monte Carlo simulation, the Altiplano option price is 0.5797 with a standard deviation 0.0013. We take them as benchmark value for the price (see Table 2.10). We compare the kernel, likelihood ratio, finite difference estimators for Greeks of the Altiplano option in Table 2.12 and Table 2.13. We observe that the kernel method and the central finite difference method are competitive. Both of them give better results than the likelihood ratio method. However, it is not easy to pick the right perturbation size h for the finite difference method. When h is too large, the estimator is highly biased. When h is small enough, the Greeks estimator is unstable (see Table 2.11).

Table 2.11: The central finite difference estimator for correlation delta using different perturbation size.

h	$\Delta_{\rho_{1,2}}$
0.2	0.0223(0.0079)
0.1	0.0202(0.0101)
0.05	0.0193(0.0158)
0.01	0.0211(0.0304)
0.005	0.0318(0.0469)
0.001	0.0283(0.1313)

Table 2.12: Comparison of Greeks and root mean square errors (RMSE) for an Altiplano option using the kernel method, the likelihood ratio method and the finite difference method under Monte Carlo simulations.

Greeks	Kernel	Likelihood Ratio	Finite Difference
rho	1.1133(0.0189)	1.1161 (0.0319)	1.0669 (0.2339)
theta	0.1242 (0.0112)	0.1159 (0.0157)	0.1194 (0.0129)
vega			
$\partial V/\partial \sigma_1$	-0.1509(0.0069)	-0.1514(0.0465)	-0.1461 (0.0066)
$\partial V/\partial \sigma_2$	-0.1904(0.0108)	-0.1967(0.0595)	-0.1846 (0.0076)
$\partial V/\partial \sigma_3$	-0.1385(0.0090)	-0.1303(0.0425)	-0.1306 (0.0058)
$\partial V/\partial \sigma_4$	-0.2906(0.0078)	-0.2839(0.0288)	-0.2927 (0.0055)
corr. delta			
$\partial V/\partial \rho_{1,2}$	0.0212(0.0016)	0.0210 (0.0582)	0.0206 (0.0101)
$\partial V/\partial \rho_{1,3}$	0.0401(0.0019)	0.0492 (0.0650)	0.0389 (0.0089)
$\partial V/\partial \rho_{1,4}$	0.0390(0.0040)	0.0461 (0.0648)	0.0381 (0.0102)
$\partial V/\partial \rho_{2,3}$	0.0299(0.0018)	0.0383 (0.0713)	0.0295 (0.0083)
$\partial V/\partial \rho_{2,4}$	0.0295(0.0040)	0.0286 (0.0478)	0.0263 (0.0086)
$\partial V/\partial \rho_{3,4}$	0.0557(0.0049)	0.0521 (0.0791)	0.0531 (0.0137)

Table 2.13: Comparison of Greeks and RMSE for an Altiplano option using the kernel, likelihood ratio and finite difference method under quasi-Monte Carlo simulations (scrambled Faure sequences).

Greeks	Kernel	Likelihood Ratio	Finite Difference
rho	1.1109(0.0176)	1.1106(0.0048)	1.1129(0.0260)
theta	0.1138(0.0060)	0.1218(0.0070)	0.1131(0.0110)
vega			
$\partial V/\partial\sigma_1$	-0.1507(0.0069)	-0.1705(0.0264)	-0.1470(0.0063)
$\partial V/\partial\sigma_2$	-0.1881(0.0132)	-0.2050(0.0255)	-0.1864(0.0071)
$\partial V/\partial\sigma_3$	-0.1347(0.0085)	-0.1554(0.0316)	-0.1307(0.0068)
$\partial V/\partial\sigma_4$	-0.2919(0.0083)	-0.2926(0.0181)	-0.2919(0.0067)
corr.delta			
$\partial V/\partial\rho_{1,2}$	0.0212(0.0018)	0.0201(0.0092)	0.0208 (0.0050)
$\partial V/\partial\rho_{1,3}$	0.0397(0.0016)	0.0477(0.0113)	0.0406 (0.0028)
$\partial V/\partial\rho_{1,4}$	0.0406(0.0044)	0.0384(0.0111)	0.0398 (0.0044)
$\partial V/\partial\rho_{2,3}$	0.0302(0.0017)	0.0359(0.0117)	0.0306 (0.0030)
$\partial V/\partial\rho_{2,4}$	0.0303(0.0034)	0.0326(0.0109)	0.0310 (0.0041)
$\partial V/\partial\rho_{3,4}$	0.0570(0.0043)	0.0545(0.0135)	0.0571 (0.0051)

2.4 Himalaya Option

The Himalaya option can be described as a call option on the sum of the best performers of a basket of stocks over a particular time horizon. The unique feature of this option is the withdrawal of the best performer from the basket over different periods. Overhaus [60] gives an extensive discussion of this option. The financial insight behind this product is that if one asset performs well in the first period, it will not perform well over the remaining periods. To get a good return over the entire period, the trader will sell the best performer at the end of the first period. Then at the end of the second period, the trader will sell the asset that has performed best among the remaining ones, etc. In other words, the best-performing stock is withdrawn from the selection after each predefined sampling dates. The payoff is the sum of all the measured returns over the life of the option. Assume there are n different assets in the basket with performances processes X_i, for $i = 1, ..., n$. At discrete time steps $0 = t_0 \leq t_1 \leq t_2 \leq ... \leq t_n = T$, the performances are observed, and the best performing asset will be withdrawn and sold. We write the terminal payoff $\varphi : \mathbb{R}_+^n \to \mathbb{R}$ in terms of performance processes:

$$\varphi(X) = \sum_{i=1}^{n} e^{-rt_i} \left(X_{\pi(i)}(t_i) - 1 \right), \tag{2.53}$$

where $\pi(1), .., \pi(n)$ is the permutation of $1, ..., n$ that describes which asset performed best in each period. For example $\pi(1) = 3$ means the 3rd asset performs the best in the first period, and $\pi(2) = 1$ means the first asset performs best in the second period among the remaining ones after the 3rd asset is removed from the basket. The Himalaya option with the payoff (2.53) will have a small value (in general, negative value). Overhaus [60] observes that the Himalaya option with the payoff (2.53) will have a small value (in general, negative value) and suggests the following modifications to the payoff function:

- Fewer period ($L < n$):

$$\varphi_{few}(X) = \sum_{i=1}^{L} e^{-rt_i} \left(X_{\pi(i)}(t_i) - 1 \right) \tag{2.54}$$

- Local floor:

$$\varphi_{local}(X) = \sum_{i=1}^{n} e^{-rt_i} \max \left(X_{\pi(i)}(t_i) - 1, 0 \right) \tag{2.55}$$

- Global floor

$$\varphi_{global}(X) = \max\left(\sum_{i=1}^{n} e^{-rt_i}\left(X_{\pi(i)}(t_i) - 1\right), 0\right) \tag{2.56}$$

There is no known analytical formulas for the price of Himalaya option, except for the special cases. For the case of the fewer period ($L < n$), the payoff (2.54) degenerates to the simple maximum of all with $L = 1$. Currently, the Monte Carlo simulation is only numerical method to price the Himalaya option in the general case. We estimate the values of the Himalaya option by the Monte Carlo method and the quasi-Monte Carlo method, and compare them in Table 2.14.

In this section, we only consider estimating Greeks with the local floor payoff (2.55). We observe that all above types of payoff functions are discontinuous, and the finite difference estimators can be quite unstable. The payoffs given above have discontinuities since bumping the path of one asset can alter the order in which the assets are removed from the basket. Also, the pathwise estimators are not unbiased, since the interchange of differentiation and integration is no long valid. We use the kernel method [52] to estimate the Greeks for the Himalaya option.

2.4.1 Kernel Method

Here we only consider the Himalaya option written on two assets, i.e. the case $n = 2$. The payoff function is expressed in the form of (1.34) as follows:

$$\begin{aligned}\varphi(S) = &\left(e^{-rt_1}(X_1(t_1) - 1)^+ + e^{-rt_2}(X_2(t_2) - 1)^+\right) \cdot 1_{\{X_1(t_1) \geq X_2(t_1)\}} \\ &+ \left(e^{-rt_1}(X_2(t_1) - 1)^+ + e^{-rt_2}(X_1(t_2) - 1)^+\right) \cdot 1_{\{X_1(t_1) < X_2(t_1)\}}.\end{aligned} \tag{2.57}$$

We break Eqn.(2.57) into two parts. The first part consists of $g_1(X) = e^{-rt_1}(X_1(t_1) - 1)^+ + e^{-rt_2}(X_2(t_2) - 1)^+$ and $h_1(X) = X_1(t_1) - X_2(t_1)$. And the second part consists of $g_2(X) = e^{-rt_1}(X_2(t_1) - 1)^+ + e^{-rt_2}(X_1(t_2) - 1)^+$ and $h_2(X) = -h_1(X)$. Let $V_1 = E\left[g_1(X)1_{\{h_1(X)\geq 0\}}\right]$ and $V_2 = E\left[g_2(X)1_{\{h_2(X)\geq 0\}}\right]$. The Himalaya option price is the sum of V_1 and V_2.

Assumption 1 is typical satisfied, since g_1, h_1 and g_2, h_2 are Lipschitz continuous and differentiable almost everywhere. Assumption 2 is satisfied since the normal density function is infinitely differentiable with respect to the parameter. In conclusion, the kernel method is applicable to the Himalaya option. We derive the estimators for Greeks vega and correlation delta next. Other Greeks can be obtained in a similar way.

Proposition 16 *The kernel estimator of the* ***vega*** *with respect to the first asset is*

$$\begin{aligned}
\bar{G}_N &= \frac{1}{N}\sum_{l=1}^{N}\left(e^{-rt_1}\frac{dX_{1,l}(t_1)}{d\sigma_1}\cdot 1_{\{X_{1,l}(t_1)-X_{2,l}(t_1)\geq 0\}}\cdot 1_{\{X_{1,l}(t_1)>1\}}\right)\\
&+\frac{1}{N}\sum_{l=1}^{N}\left(e^{-rt_2}\frac{dX_{1,l}(t_2)}{d\sigma_1}\cdot 1_{\{X_{1,l}(t_1)-X_{2,l}(t_1)\leq 0\}}\cdot 1_{\{X_{1,l}(t_2)>1\}}\right)\\
&+\frac{1}{N\delta_N}\sum_{l=1}^{N}\left((g_{1,l}(X)-g_{2,l}(X))\frac{dX_{1,l}(t_1)}{d\sigma_1}Z\left(\frac{X_{1,l}(t_1)-X_{2,l}(t_1)}{\delta_N}\right)\right),
\end{aligned}$$

where $Z(\cdot)$ is the normal kernel density function with bandwidth δ_N, and ($X_{1,l}(t_1)$, $X_{1,l}(t_2)$, $X_{2,l}(t_1)$, $g_{1,l}(X)$, $g_{2,l}(X)$, $\frac{dX_{1,l}(t_1)}{d\sigma_1}$, $\frac{dX_{1,l}(t_2)}{d\sigma_1}$) denotes the lth samples obtained from the Monte Carlo simulation. We obtain the vega with respect to the second asset by switching subscript 1 and 2 for all the variables.

Proof. Now we apply the Theorem 1 [52] to get the derivative of V_1 with respect to σ_1

$$\frac{dV_1}{d\sigma_1} = E\left[e^{-rt_1}\frac{dX_1(t_1)}{d\sigma_1}\cdot 1_{\{X_1(t_1)-1>0\}}\cdot 1_{\{h_1(X)\geq 0\}}\right] - \partial_y\, E\left[g_1(X)\frac{dX_1(t_1)}{d\sigma_1}\cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0}, \tag{2.58}$$

where $g_1(X) = e^{-rt_1}(X_1(t_1)-1)^+ + e^{-rt_2}(X_2(t_2)-1)^+$ and $h_1(X) = X_1(t_1)-X_2(t_1)$; The derivatives of $g_1(X)$ and $h_1(X)$ respect to σ_1 are

$$\frac{dg_1(X)}{d\sigma_1} = e^{-rt_1}\frac{dX_1(t_1)}{d\sigma_1}\cdot 1_{\{X_1(t_1)-1>0\}}$$

and

$$\frac{dh_1(X)}{d\sigma_1} = \frac{dX_1(t_1)}{d\sigma_1}.$$

The derivative of V_2 with respect to σ_1 is given by

$$\frac{dV_2}{d\sigma_1} = E\left[e^{-rt_2}\frac{dX_1(t_2)}{d\sigma_1}\cdot 1_{\{X_2(t_2)-1>0\}}\cdot 1_{\{h_1(X)\leq 0\}}\right] + \partial_y\, E\left[g_2(X)\frac{dX_1(t_1)}{d\sigma_1}\cdot 1_{\{-h_1(X)\geq y\}}\right]\Bigg|_{y=0}, \tag{2.59}$$

where $g_2(X) = e^{-rt_1}(X_2(t_1)-1)^+ + e^{-rt_2}(X_1(t_2)-1)^+$ and $h_2(X) = -h_1(X)$. The derivatives of $h_2(X)$ and $g_2(X)$ with respect to σ_1 are

$$\frac{dg_2(X)}{d\sigma_1} = e^{-rt_2}\frac{dX_1(t_2)}{d\sigma_1}\cdot 1_{\{X_1(t_2)-1>0\}}$$

and

$$\frac{dh_2(X)}{d\sigma_1} = -\frac{dX_1(t_1)}{d\sigma_1}.$$

The Greek vega equals to the sum of Eqn.(2.58) and Eqn.(2.59):

$$\begin{aligned}
\frac{dV}{d\sigma_1} &= E\left[e^{-rt_1}\frac{dX_1(t_1)}{d\sigma_1}\cdot 1_{\{X_1(t_1)>1\}}\cdot 1_{\{h_1(X)\geq 0\}} + e^{-rt_2}\frac{dX_1(t_2)}{d\sigma_1}\cdot 1_{\{h_1(X)\leq 0\}}\cdot 1_{\{X_1(t_2)>1\}}\right] \\
&\quad - \partial_y\, E\left[(g_1(X)-g_2(X))\frac{dX_1(t_1)}{d\sigma_1}\cdot 1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0} \\
&= E\left[e^{-rt_1}\frac{dX_1(t_1)}{d\sigma_1}\cdot 1_{\{X_1(t_1)>1\}}\cdot 1_{\{h_1(X)\geq 0\}} + e^{-rt_2}\frac{dX_1(t_2)}{d\sigma_1}\cdot 1_{\{h_1(X)\leq 0\}}\cdot 1_{\{X_1(t_2)>1\}}\right] \\
&\quad + \lim_{\delta\to 0}\frac{1}{\delta}E\left[(g_1(X)-g_2(X))\frac{dX_1(t_1)}{d\sigma_1}Z\left(\frac{X_1(t_1)-X_2(t_1)}{\delta}\right)\right]
\end{aligned} \tag{2.60}$$

with

$$\begin{aligned}
\frac{dX_1(t_1)}{d\sigma_1} &= X_1(t_1)(-\sigma_1 t_1 + L_1\sqrt{t_1}Z_1), \\
\frac{dX_1(t_2)}{d\sigma_1} &= X_1(t_1)(-\sigma_1 t_1 + L_1\sqrt{t_1}Z_1 + L_1\sqrt{t_2-t_1}Z_2).
\end{aligned}$$

The expectations on the right hand side of Eqn.(2.60) are estimated by the Monte Carlo simulation. Let ($X_{1,l}(t_1)$, $X_{1,l}(t_2)$, $X_{2,l}(t_1)$, $g_{1,l}(X)$, $g_{2,l}(X)$, $\frac{dX_{1,l}(t_1)}{d\sigma_1}$, $\frac{dX_{1,l}(t_2)}{d\sigma_1}$) be the lth observations obtained from a simulation. And the optimal bandwidth δ_N could be obtained by pilot simulation.

$$\begin{aligned}
\bar{G}_N &= \frac{1}{N}\sum_{l=1}^{N}\left(e^{-rt_1}\frac{dX_{1,l}(t_1)}{d\sigma_1}\cdot 1_{\{X_{1,l}(t_1)-X_{2,l}(t_1)\geq 0\}}\cdot 1_{\{X_{1,l}(t_1)>1\}}\right) \\
&\quad + \frac{1}{N}\sum_{l=1}^{N}\left(e^{-rt_2}\frac{dX_{1,l}(t_2)}{d\sigma_1}\cdot 1_{\{X_{1,l}(t_1)-X_{2,l}(t_1)\leq 0\}}\cdot 1_{\{X_{1,l}(t_2)>1\}}\right) \\
&\quad + \frac{1}{N\delta_N}\sum_{l=1}^{N}\left((g_{1,l}(X)-g_{2,l}(X))\frac{dX_{1,l}(t_1)}{d\sigma_1}Z\left(\frac{X_{1,l}(t_1)-X_{2,l}(t_1)}{\delta_N}\right)\right).
\end{aligned}$$

■

Proposition 17 *The kernel estimator of the* ***correlation delta*** *with respect to* $\rho_{i,j}$ *is given as follows:*

$$\begin{aligned}
\bar{G}_N &= \frac{1}{N}\sum_{l=1}^{N}\left(\frac{dg_{1,l}(X)}{d\rho}\cdot 1_{\{h_{1,l}(X)\geq 0\}} + \frac{dg_{2,l}(X)}{d\rho}\cdot 1_{\{h_{1,l}(X)\leq 0\}}\right) \\
&\quad + \frac{1}{N\delta_N}\sum_{l=1}^{N}\left((g_{2,l}(X)-g_{1,l}(X))\frac{dX_{2,l}(t_1)}{d\rho}Z\left(\frac{h_{1,l}(X)}{\delta_N}\right)\right),
\end{aligned} \tag{2.61}$$

where $Z(\cdot)$ *denotes the standard normal density function and* ($g_{1,l}(X)$, $g_{2,l}$, $\frac{dg_{1,l}(X)}{d\rho}$, $\frac{dg_{2,l}(X)}{d\rho}$, $\frac{dX_{2,l}(t_1)}{d\rho}$) *denotes the* l*th sample from the Monte Carlo simulation. The optimal value bandwidth* $\delta_N > 0$ *is obtained by a pilot simulation.*

Proof. The derivation of Eqn.(2.61) is the same as above. The derivative with respect to ρ is

$$\begin{aligned}\frac{dV}{d\rho} &= E\left[\frac{dg_1(X)}{d\rho}\cdot 1_{\{h_1(X)\geq 0\}}\right] + E\left[\frac{dg_2(X)}{d\rho}\cdot 1_{\{h_1(X)\leq 0\}}\right] \\ &\quad - \partial_y\, E\left[(g_2(X)-g_1(X))\,\frac{dX_2(t_1)}{d\rho}1_{\{h_1(X)\geq y\}}\right]\Bigg|_{y=0},\end{aligned} \tag{2.62}$$

where the derivatives of g_1 and g_2 are

$$\begin{aligned}\frac{dg_1(X)}{d\rho} &= e^{-rt_2}\frac{dX_2(t_2)}{d\rho}\cdot 1_{\{X_2(t_2)-1>0\}} \\ &= e^{-rt_2}X_2(t_2)\left(\sqrt{t_1}Z_1 - \frac{\sqrt{t_1}\rho Z_2}{\sqrt{1-\rho^2}} + \sqrt{t_2-t_1}Z_3 - \frac{\sqrt{t_2-t_1}\rho Z_2}{\sqrt{1-\rho^4}}\right)\cdot 1_{\{X_2(t_2)-1>0\}}\end{aligned} \tag{2.63}$$

and

$$\begin{aligned}\frac{dg_2(X)}{d\rho} &= e^{-rt_1}\frac{dX_2(t_1)}{d\rho}\cdot 1_{\{X_2(t_1)-1>0\}} \\ &= e^{-rt_1}X_2(t_1)\sigma_2\sqrt{t_1}\left(Z_1 - \frac{\rho}{\sqrt{1-\rho^2}}Z_2\right)\cdot 1_{\{X_2(t_1)-1>0\}}.\end{aligned} \tag{2.64}$$

The standard normal variables Z_1, Z_2, Z_3, Z_4 are used to generate sample processes. Since $h_2(X) = -h_1(X)$, we have

$$\frac{dh_1(X)}{d\rho} = -\frac{dh_2(X)}{d\rho} = -\frac{dX_2(t_1)}{d\rho}.$$

The kernel estimator of Greek correlation delta via the Monte Carlo simulation is

$$\begin{aligned}\bar{G}_N &= \frac{1}{N}\sum_{l=1}^{N}\left(\frac{dg_{1,l}(X)}{d\rho}\cdot 1_{\{h_{1,l}(X)\geq 0\}} + \frac{dg_{2,l}(X)}{d\rho}\cdot 1_{\{h_{1,l}(X)\leq 0\}}\right) \\ &\quad + \frac{1}{N\delta_N}\sum_{l=1}^{N}\left((g_{2,l}(X)-g_{1,l}(X))\frac{dX_{2,l}(t_1)}{d\rho}Z\left(\frac{h_{1,l}(X)}{\delta_N}\right)\right),\end{aligned}$$

where $(g_{1,l}(X),\, g_{2,l},\, \frac{dg_{1,l}(X)}{d\rho},\, \frac{dg_{2,l}(X)}{d\rho},\, \frac{dX_{2,l}(t_1)}{d\rho})$ denotes the lth sample from the Monte Carlo simulation. We generate the samples for $\frac{dg_{1,l}(X)}{d\rho}$, and $\frac{dg_{2,l}(X)}{d\rho}$ from Eqn.(2.63) and Eqn.(2.64). ■

2.4.2 Likelihood Ratio Method

The score functions of the Himalaya option are the same as that of the Altiplano option. We get the likelihood ratio estimators of all the Greeks by replacing the Altiplano payoff with the Himalaya payoff. The results are listed next.

Rho. The likelihood ratio estimator for rho is

$$e^{-rT}\varphi(Y)\left(-T + \sum_{k=1}^{m} Z_k' L^{-1} A^{-1}\sqrt{\Delta t_k} I_{n\times 1}\right), \tag{2.65}$$

Theta. The likelihood ratio estimator for theta is

$$e^{-rT}\varphi(Y)\left(-r-\frac{nm}{2T}+\frac{1}{\sqrt{mT}}\left(\sum_{k=1}^{m}Z_k'\right)L^{-1}A^{-1}\left(rI_{n\times 1}-\frac{\sigma^2}{2}\right)+\frac{1}{2T}\sum_{k=1}^{m}Z_k'Z_k\right),$$

Vega. The likelihood ratio estimator for the ith entry of vega is

$$e^{-rT}\varphi(Y)\left(-\frac{m}{\sigma_i}-\sqrt{\frac{m}{T}}\sum_{k=1}^{m}Z_k'L^{-1}\left(0,...,\Delta t-\frac{\sqrt{\Delta t}}{\sigma_i^2}(ALZ_k)_i,...,0\right)'\right),$$

Correlation delta. The likelihood ratio estimator of the correlation delta between asset i and asset j is

$$e^{-rT}\varphi(Y)\sum_{k=1}^{m}\left(-\frac{1}{2}Tr\left(L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}\right)+\frac{1}{2}Z_k'L^{-1}\frac{\partial\Sigma}{\partial\rho_{i,j}}(L')^{-1}Z_k\right),$$

where L is the Cholesky matrix such that $LL'=\Sigma$, and Z_k is a standard n-dimensional normal random vector which is used to generate the sample paths. The matrix $\partial\Sigma/\partial\rho_{i,j}$ contains entries at (i,j) and (j,i) equal to 1, and 0 otherwise.

2.4.3 Numerical Results

We use the same assets as in the Everest option which are J.P. Morgan Chase (JPM) and Lehman Brothers (LEH), and each of the stocks had a zero dividend yield the period. The maturity is $T=1$ year, and the risk-free rate is $r=0.04$. The volatilities are $\sigma_1=0.113978$ and $\sigma_2=0.316459$, respectively, and the correlation coefficient between them is $\rho=0.6430$. The sample size is set at 10,000 with 40 independent simulations.

Table 2.14: Comparison of price and the standard deviation for the Himalaya option between Monte Carlo and quasi-Monte Carlo simulation.

	MC	QMC
Price	0.3269	0.3268
Std.Dev	0.0018	6.1189e-005

We compare the kernel, likelihood ratio, and finite difference estimators for Greeks of the Himalaya option in Table. 2.15 and Table. 2.16. We observe that kernel method and the central finite difference methods are competitive. Both of them give better results than the likelihood ratio method.

Table 2.15: The Greeks of the Himalaya option by the kernel method, the likelihood ratio method and the finite difference method (central difference) using Monte Carlo simulation. When $\theta = r$, $h = 0.0001$; when $\theta = T$, $h = 1/12$; when $\theta = \sigma$, $h = 0.1$; when $\theta = \rho$, $h = 0.1$;

	Kernel	Likelihood Ratio	Finite Difference
theta($\partial V/\partial T$)	0.1699(0.0031)	0.1668(0.0131)	0.1703(0.0035)
vega			
($\partial V/\partial \sigma_1$)	0.3158(0.0094)	0.3326 (0.0622)	0.3159(0.0082)
($\partial V/\partial \sigma_2$)	0.3226(0.0072)	0.3169 (0.0297)	0.3253(0.0091)
corr. delta			
($\partial V/\partial \rho$)	0.0062(0.0053)	0.0057(0.0200)	0.0069(0.0064)

Table 2.16: The Greeks of the Himalaya option by the kernel method, the likelihood ratio method and the finite difference method (central difference) using quasi-Monte Carlo simulation.

	Kernel	Likelihood Ratio	Finite Difference
theta($\partial V/\partial T$)	0.1701(0.0008)	0.1613(0.0017)	0.1699(0.0012)
vega			
($\partial V/\partial \sigma_1$)	0.3174(0.0025)	0.2998 (0.0116)	0.3137(0.0018)
($\partial V/\partial \sigma_2$)	0.3207(0.0020)	0.3233 (0.0061)	0.3238(0.0016)
corr. delta			
($\partial V/\partial \rho$)	0.0064(9.0627e-004)	0.0058(0.0034)	0.0072(0.0024)

CHAPTER 3

ESTIMATING GREEKS FOR WEATHER DERIVATIVES

Weather is a huge driver of the economy. Weather events and weather prediction are important topics long studied by meteorologists. Even though there exists a variety of models utilized by meteorology departments, weather events cannot be predicted with high precision for long forecast horizons. Many companies face the possibility of significant earnings declines because of adverse or unexpected weather conditions. For example, farmers harvest badly due to failing excessive rain during the growing or harvesting period, high winds in case of plantations or temperature variabilities in case of greenhouse crops. Even in the presence of good weather forecasts the economic risks of weather events cannot be fully eliminated. For example, a natural gas supplier has low residential natural gas demand during warm winters;

Companies might protect themselves against disasters through insurance, which typically covers high-risk, low probability events, as defined in a highly tailored or customized policy. But insurance is not designed to help companies against non-catastrophic weather derivatives. In contrast to the weather insurance, weather derivatives cover low-risk, high probability events. Their economic purpose is to transfer weather risk from one counterparty of the contract to the other. Because of such needs for hedging against weather events, weather derivatives market has been developing in terms of product variety and trade volume rapidly. Weather derivatives gained popularity in the United States in 1997 after the de-regulation of the energy market. Since 1999, standard contracts, i.e., futures and options on futures based on temperature, have been trading on the Chicago Mercantile Exchange (CME), while forward and swap contracts are traded by financial institutions in the over-the-counter market.

3.1 Introduction to Weather Derivatives

Weather derivatives are financial instruments that can be used by organizations or individuals as part of a risk management strategy to reduce risk associated with adverse or unexpected weather conditions. The payoff functions are written on underlyings, such as heating/cooling degree days (HDDs/CDDs), sunshine precipitation, snowfall precipitation, etc. The main difference between weather derivatives and derivatives written on tradable assets is that the weather underlying asset is free of economy risk aversion factors, since it is not the price. Therefore it has no direct value to price the weather derivatives. The components of a basic contract on weather derivatives are listed in Table 3.1.

Table 3.1: Components of a contract

1. Underlying assets	HDD/CDD, wind-speed, snowfall, rainfall
2. Contract period	Periods between the inception date and maturity
3. Location	Business location
4. Unit size	(Currency) per degree or millimetre
5. Limit	The maximum payout under the contract, or notional value
6. Structure	Swap, Collar, Put/Call, Straddle
7. Strike level	The predetermined weather index level

The industry has set up temperature as the common underlying for contracts, and extensive literature on the pricing of weather derivatives written on temperature has been developed. Various approaches such as historical burn analysis, index modelling, and dynamic modelling through s-tochastic differential equations are used in the valuation of weather derivatives. Studies by Oetomo and Stevenson [58] and Schiller et al.[70] suggest that the dynamic modelling of temperatures is superior to the index modelling approach. The dynamic modelling approach often models daily average temperatures with a mean-reverting Ornstein-Uhlenbeck process. These models have the advantage of consistently pricing a wide range of weather derivatives based on different temperature indices. Some examples of such models are provided by Alaton et al.[1], Benth and Benth [4, 5], Campbell and Diebold [16], Brody et al.[15], Richards et al.[66], Cao and Wei [17], Göncü [37], Platen and West [65], Huang et al.[46], Zapranis and Alexandridis [81], and Zeng [83].

3.1.1 Pricing and Hedging Weather Derivatives

The main difference between the pricing of financial derivatives and weather derivatives is in the uniqueness of a risk-neutral probability measure. In financial derivatives, most pricing models are based on the no-arbitrage arguments. However, in weather derivatives, the underlying is a non-tradable weather index and thus there is no unique risk-neutral probability measure. Hull [47] claims that the underlying assets of weather derivatives can be assumed to have zero systematic risk, so that the historical data approach can be used. Historical data approach estimates the expected payoff under the physical probability measure. The expected growth rate of a variable can be assumed to be the same in both the real world and the risk-neutral world if the variable has zero systematic risk so that percentage changes in the variable have zero correlation with stock market returns. Jewson [48] also mentions that, in a diversified portfolio of stocks of many companies, the weather-driven part of the variability of the portfolio performance is likely to disappear almost completely. Thus a company investing in a diversified portfolio of stocks can indeed invest in weather derivatives under the assumption of zero correlation with the market.

Furthermore, in the absence of a tradable underlying there are very few studies on the sensitivities of weather derivatives. In the study by Jewson [48] the sensitivities are calculated based on an index modelling approach. The analytical approximation formula of Alaton et al. [1] is utilized in the study by [37] to derive sensitivities of HDD/CDD options with respect to the mean and standard deviation of the accumulated HDD/CDDs. However, the analytical pricing formula given in [1] assumes that during a winter (summer) HDD (CDD) contract daily temperatures will always be below (above) 65 Fahrenheit degrees. This assumption clearly is not satisfied for large geographical areas, and examples that show the inaccuracy of sensitivities based on this approximation is given in Göncü [37]. In this chapter we will present direct Monte Carlo methods to estimate first and second order sensitivities.

Even though complete hedging opportunities do not exist in weather derivatives, sensitivity information is still useful. There are partial hedging opportunities via tradable assets that are correlated with temperature. A detailed study on partial hedging in incomplete markets can be found in [74]. Some examples of such tradable assets are stocks of energy or utility companies. Hence, there is valuable information to be extracted from the sensitivities of weather derivatives with respect to a variety of parameters that govern the temperature dynamics.

Weather derivatives' sensitivities can be used by a trader to compare a variety of available weather derivatives and choose the most suitable one based on the expectations or requirements of the trader. For example, if the weather forecasts indicate a possible increase in the temperature variation over the short term, then the option with a higher "vega" might be preferred based on this information. To bet on the increasing temperature volatility a trader can long weather derivatives with higher vega and short the ones with a lower vega.

The "delta" of the weather derivative can be used as a measure of sensitivity with respect to forecast errors. For example, to price a January weather contract (i.e. the contract for the period Jan.1st-31th) on December 29th one would use the available weather forecast for January 1st as the initial temperature in the pricing model. In this case the "delta" of the weather derivative can be used to find the sensitivity of the price with respect to the error in temperature forecast. Similarly, the "gamma" can be used as a measure of sensitivity with respect to large forecast errors.

Alternatively, the sensitivities can be used for risk management of weather derivatives portfolios. For example, if a portfolio of weather derivatives has vega equal to $\mathcal{V}$ and there exists a traded weather derivatives with vega equals to $\mathcal{V}_T$, then a position of $-\mathcal{V}/\mathcal{V}_T$ makes the portfolio vega neutral.

Furthermore, the daily temperature models that are based on mean reverting stochastic processes (as examples given in [1] and [5]) are estimated using historical data of daily average temperatures. The parameter estimators involve regression errors and standard errors are used to construct confidence intervals of these estimates. Hence, through the calculation of sensitivities we have more information regarding the effect of regression errors on the estimated option prices.

In Section 3.2 we will introduce the model for daily average temperatures and our data set. Analytical formulas for the estimators of the weather derivatives' sensitivities will be derived in Section 3.3. Numerical results will be presented in Section 3.5.

3.2 Modelling Daily Average Temperatures

Weather derivatives are based on a variety of different weather variables, and the most commonly used one is the temperature. We consider the temperature model proposed by Alaton et al.[1]:

$$dT_t = \left(\frac{dT_t^m}{dt} + a(T_t^m - T_t) \right) dt + \sigma_t dW_t, \tag{3.1}$$

where T_t is the temperature at time t, a is the mean reversion parameter, σ_t is the piecewise constant (constant value during each month) volatility function, W_t is P-Brownian motion (the physical probability measure) and T_t^m represents the long-term mean temperature which can be modelled by

$$T_t^m = A + Bt + C \sin(wt + \phi), \tag{3.2}$$

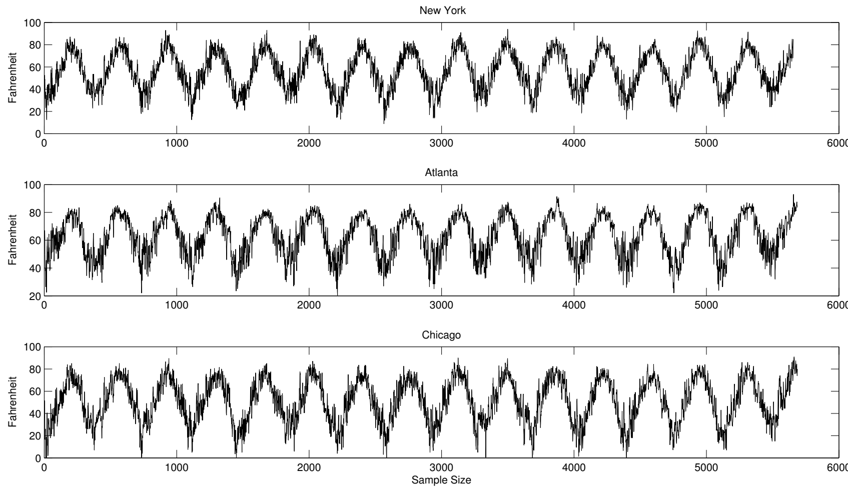

Figure 3.1: Daily average temperatures at New York (La Guardia Airport), Atlanta (Hartsfield-Jackson Airport), and Chicago (O'Hare Airport) from 1997 to 2012.

where $w = 2\pi/365$. The parameters A, B, C, ϕ, and a are estimated from data, using the method discussed in [1]. We will use temperature data from New York, Atlanta and Chicago, to fit the model (3.1) with long-term mean temperature (3.2). The fitted parameters of the model along with the t-statistics are provided in Table 3.2. The daily average temperature T_j at a specific weather station on a given day j is calculated as the average of maximum and minimum temperatures (in degrees Fahrenheit) measured during that day, i.e. $T_j = (T_j^{max} + T_j^{min})/2$. In our data, daily average temperatures are measured by Earth Satellite Corporation. Temperature data for the period between 1st January 1997 and 31th January 2012 is used. The dataset is plotted in Figure 3.1. We observe that seasonality is the main characteristic of daily temperature data.

If we incorporate the market price of risk into the same model then Eqn.(3.1) can be rewritten as

$$dT_t = \left(\frac{dT_t^m}{dt} + a(T_t^m - T_t) - \lambda \sigma_t \right) dt + \sigma_t dW_t^*,$$

where W_t^* is Q-Brownian motion under the risk-neutral probability measure. Without loss of generality we assume the market price of risk λ is equal to zero and thus risk-neutral probability measure reduces to the physical probability measure. Under this assumption the temperature at time t, with an initial value of T_s, for $s < t$, can be written by

$$T_t = (T_s - T_s^m)e^{-a(t-s)} + T_t^m + \int_s^t e^{-a(t-\tau)}\sigma_\tau dW_\tau. \tag{3.3}$$

We provide volatility estimates of daily temperature using a regression method for each month of the year given in Table 3.3.

Table 3.2: Fitted parameters of the model given by Alaton et al.[1] (t-statistics of all the estimated parameters are significant at 95% confidence level)

	A	B	C	ϕ	a
New York	55.7952	0.0003	21.8653	-1.9848	0.3491
Atlanta	61.9267	0.0004	18.1663	-1.8617	0.2620
Chicago	50.0181	0.0004	24.5776	-1.9089	0.3029

Table 3.3: Monthly volatility estimates using the regression method given in Alaton et al.[1]

	New York	Atlanta	Chicago
Jan	6.36	6.56	7.11
Feb	5.84	5.90	6.63
Mar	5.82	5.67	6.71
Apr	5.52	5.06	6.38
May	4.68	3.67	6.12
Jun	4.54	2.57	4.79
Jul	3.62	2.32	4.05
Aug	3.53	2.26	3.85
Sep	4.03	3.17	4.89
Oct	4.67	4.26	5.79
Nov	5.00	5.42	6.26
Dec	5.96	5.87	6.78

3.2.1 Pricing HDD/CDD Options

Temperature contracts structured by options are commonly traded over *Heating Degree Days* (HDDs) and *Cooling Degree Days* (CDDs). Both indices are defined on daily average temperature T_j and a predetermined reference temperture T_{ref}.

Definition 18 ***Heating/Cooling Degree Day (HDD/CDD)***
The heating degree day (HDD) for a given day j is defined as

$$HDD_j = \max(T_{ref} - T_j, 0)$$

and the cooling degree day (CDD) generated on that day is given by

$$CDD_j = \max(T_j - T_{ref}, 0),$$

where T_{ref} often equals to 18° C degrees (or 65° F in the U.S. market).

Definition 19 ***Accumulated HDDs (H_n) and CDDs (C_n)***
The accumulated HDDs and CDDs for a contract period of n days are given by

$$H_n = \sum_{j=1}^{n} HDD_j \quad \text{and} \quad C_n = \sum_{j=1}^{n} CDD_j,$$

respectively.

Suppose an accumulated HDDs is measured during the n-day period denoted by $[t_1, t_n]$, and the discounted payoff of an uncapped HDD call option is given as

$$\varphi_{call}(t) = \gamma e^{-r(t_n - t)} \max(H_n - K, 0), \tag{3.4}$$

where $H_n = \sum_{i=1}^{n} \max(65 - T_{t_i}, 0)$ and T_{t_i} denotes the temperature for day i. K is the strike level; γ is the tick size in monetary units and we set $\gamma = 1$ throughout this section for simplicity.

The call option price on the accumulated HDDs at time $t < t_1$ is given by the following risk-neutral expectation

$$c_{HDD}(t) = e^{-r(t_n - t)} E_Q \left[\max\left(H_n - K, 0\right) | \mathcal{F}_t \right], \tag{3.5}$$

where filtration $\mathcal{F}_t$ contains the information of daily temperatures up to time t. We denote the (remaining) contract period as $[t_1, t_n]$, i.e. there are n days left to the end of the measurement period. If $t \in [t_1, t_n]$, i.e. the HDD/CDD accounting already started we redefine $[t_1, t_n]$ as the period of remaining n-days to the end of the HDD/CDD accounting period. For example, for a January HDD option in out-of-period valuation $n = 31$, whereas for in-period valuation n refers to the remaining number of days in January for the accumulated HDDs. Thus, the same time notation is preserved for both in-period and out-of-period (before measurement period starts) valuation. As discussed in [1] for in-period valuation H_n can be split into two parts as known and unknown parts. By also adjusting the strike price accordingly, we obtain the same call option pricing problem given in Eqn.(3.5).

3.3 Sensitivities of HDD/CDD Options

We derive pathwise estimators for the sensitivities of the model by Alaton et al.[1], with respect to the current temperature T_t, the volatility σ_t, the mean reversion parameter a, and the parameters that appear in the long-term mean temperature model: A, B, C.

3.3.1 Sensitivity with respect to the Temperature

Here we define "delta" is defined as the derivative of the HDD/CDD call and put option price with respect to the current temperature T_t. The pathwise estimator of delta is given by the following proposition.

Proposition 20 *The pathwise estimators of delta for HDD/CDD call and put options are given as follows:*

HDD call

$$\frac{\partial \varphi_{call}(t)}{\partial T_t} = -e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(e^{-a(t_i-t)} 1_{\{T_{t_i}<65\}} \right) \cdot 1_{\{H_n>K\}} \right), \tag{3.6}$$

HDD put

$$\frac{\partial \varphi_{put}(t)}{\partial T_t} = e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(e^{-a(t_i-t)} 1_{\{T_{t_i}<65\}} \right) \cdot 1_{\{H_n<K\}} \right), \tag{3.7}$$

CDD call

$$\frac{\partial \varphi_{call}(t)}{\partial T_t} = e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(e^{-a(t_i-t)} 1_{\{T_{t_i}>65\}} \right) \cdot 1_{\{C_n>K\}} \right), \tag{3.8}$$

CDD put

$$\frac{\partial \varphi_{put}(t)}{\partial T_t} = -e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(e^{-a(t_i-t)} 1_{\{T_{t_i}>65\}} \right) \cdot 1_{\{C_n<K\}} \right). \tag{3.9}$$

Proof. Sufficient conditions for the existence of the pathwise derivative and the interchangeability of the differentiation and expectation operations are listed in Chapter one (see Section 1.4.2). We first verify these conditions. (A1) Note that temperature T_{t_i} follows the normal distribution from Eqn.(3.3) and $\partial T_{t_i}/\partial T_t = e^{-a(t_i-t)}$ for each i. Therefore, at each T_t, $\partial T_{t_i}/\partial T_t$ exists with probability 1. (A2) We have $P(\sum_{i=1}^{n} \max(65 - T_{t_i}, 0) = K) = 0$ since T_{t_i} is normally distributed. Therefore the pathwise derivative exists with probability one, for each T_t. (A3) The payoff function is Lipschitz continuous with respect to H_n, and H_n is Lipschitz continuous with respect to T_i, therefore from the fact that Lipschitz continuity is preserved under composition, we have the payoff function Lipschitz continuous with respect to T_i. The last condition (A4) holds with $K_n = 1$.

Having verified the existence and interchangeability conditions, we derive the pathwise estimator of delta for HDD call options, by simply observing

$$\begin{aligned}\frac{\partial \varphi_{call}(t)}{\partial T_t} &= e^{-r(t_n-t)}\left(\frac{\partial H_n}{\partial T_t}\cdot 1_{\{H_n>K\}}\right)\\ &= -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{\partial T_{t_i}}{\partial T_t}1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n>K\}}\right)\end{aligned}$$

and $\partial T_{t_i}/\partial T_t = e^{-a(t_i-t)}$. The pathwise estimators for the other options are derived similarly. ■

3.3.2 Sensitivity with respect to the Volatility

Vega is the derivative of the option price with respect to the volatility. In Alaton's model [1], the volatility is assumed to be constant during each month. The monthly constant volatility is denoted by σ_j, where $j = 1, ..., 12$ from January to December. Throughout the following derivations, we need to work with a volatility function whose domain is the time variable, so we introduce the notation $\sigma(t_i)$ for the volatility during day i.

Consider a contract of n days, and define sets $I_n = \{1, 2, ..., n\}$ and $\chi_j = \{t_i : \sigma(t_i) = \sigma_j, i \in I_n\}$ for $j = 1, ..., 12$. The set I_n is simply the duration of the contract where the time variable is in days. The set χ_j consists of days, within the contract duration, when the volatility is equal to σ_j. For example, for a given day i during January, the volatility is constant, i.e. $\sigma(t_i) = \sigma_1$, and thus $t_i \in \chi_1$.

Proposition 21 *The pathwise estimators of vega for HDD/CDD call and put options are given as follows:*

HDD call

$$\frac{\partial \varphi_{call}(t)}{\partial \sigma_j} = -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{dT_{t_i}}{d\sigma_j}\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n>K\}}\right),$$

HDD put

$$\frac{\partial \varphi_{put}(t)}{\partial \sigma_j} = e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{dT_{t_i}}{d\sigma_j}\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n<K\}}\right),$$

CDD call

$$\frac{\partial \varphi_{call}(t)}{\partial \sigma_j} = e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{dT_{t_i}}{d\sigma_j}\cdot 1_{\{T_{t_i}>65\}}\right)\cdot 1_{\{C_n>K\}}\right),$$

CDD put

$$\frac{\partial \varphi_{put}(t)}{\partial \sigma_j} = -e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(\frac{dT_{t_i}}{d\sigma_j} \cdot 1_{\{T_{t_i}>65\}} \right) \cdot 1_{\{C_n<K\}} \right),$$

where

$$\frac{dT_{t_i}}{d\sigma_j} = \sum_{k=1}^{i} \sqrt{\frac{1-e^{-2a(t_k-t_{k-1})}}{2a}} e^{-a(t_i-t_k)} Z_k \cdot 1_{\{t_k \in \chi_j\}}. \tag{3.10}$$

We specify $t_0 = t$ in (3.10) as the initial time when $k = 1$.

Proof. The existence of pathwise derivative and the justification of interchange of operations can be established similar to the proof of Proposition 20. The derivative for the HDD call option is

$$\begin{aligned} \frac{\partial \varphi_{call}(t)}{\partial \sigma_j} &= e^{-r(t_n-t)} \left(\frac{\partial H_n}{\partial \sigma_j} \cdot 1_{\{H_n>K\}} \right) \\ &= -e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(\frac{dT_{t_i}}{d\sigma_j} \cdot 1_{\{65>T_{t_i}\}} \right) \cdot 1_{\{H_n>K\}} \right). \end{aligned}$$

The solution of the stochastic differential equation (3.1) over the interval $[t_{i-1}, t_i]$ is given by

$$T_{t_i} = (T_{t_{i-1}} - T^m_{t_{i-1}})e^{-a(t_i-t_{i-1})} + T^m_{t_i} + \sigma(t_i)\sqrt{\frac{1-e^{-2a(t_i-t_{i-1})}}{2a}} Z_i, \tag{3.11}$$

where $\sigma(t_i)$ is the volatility over one day $[t_{i-1}, t_i]$ and $\{Z_i\}_{i=1}^n$ are independent standard normal random variables. For a fixed j, we have $dT_{t_i}/d\sigma_j = 0$ when $t_i < \min\{\chi_j\}$, which follows from the definition of the set χ_j.

By differentiating both sides of (3.11), we get the following recursion for the derivatives:

$$\frac{dT_{t_i}}{d\sigma_j} = \frac{dT_{t_{i-1}}}{d\sigma_j} e^{-a(t_i-t_{i-1})} + \sqrt{\frac{1-e^{-2a(t_i-t_{i-1})}}{2a}} Z_i \cdot 1_{\{t_i \in \chi_j\}}, \tag{3.12}$$

with the initial condition $dT_{t_0}/d\sigma_j = 0$. Starting with Eqn.(3.12) and applying the recursion, we obtain

$$
\begin{aligned}
\frac{dT_{t_i}}{d\sigma_j} &= \frac{dT_{t_{i-1}}}{d\sigma_j} e^{-a(t_i - t_{i-1})} + \sqrt{\frac{1 - e^{-2a(t_i - t_{i-1})}}{2a}} Z_i \cdot 1_{\{t_i \in \chi_j\}} \\
&= \left(\frac{dT_{t_{i-2}}}{d\sigma_j} e^{-a(t_{i-1} - t_{i-2})} + \sqrt{\frac{1 - e^{-2a(t_{i-1} - t_{i-2})}}{2a}} Z_{i-1} \cdot 1_{\{t_{i-1} \in \chi_j\}} \right) e^{-a(t_i - t_{i-1})} \\
&\quad + \sqrt{\frac{1 - e^{-2a(t_i - t_{i-1})}}{2a}} Z_i \cdot 1_{\{t_i \in \chi_j\}} \\
&= \frac{dT_{t_{i-2}}}{d\sigma_j} e^{-a(t_i - t_{i-2})} + \sqrt{\frac{1 - e^{-2a(t_{i-1} - t_{i-2})}}{2a}} Z_{i-1} \cdot 1_{\{t_{i-1} \in \chi_j\}} e^{-a(t_i - t_{i-1})} \\
&\quad + \sqrt{\frac{1 - e^{-2a(t_i - t_{i-1})}}{2a}} Z_i \cdot 1_{\{t_i \in \chi_j\}} \\
&= \frac{dT_{t_{i-3}}}{d\sigma_j} e^{-a(t_i - t_{i-3})} + \sqrt{\frac{1 - e^{-2a(t_{i-2} - t_{i-1})}}{2a}} Z_{i-2} \cdot 1_{\{t_{i-2} \in \chi_j\}} e^{-a(t_i - t_{i-2})} \\
&\quad + \sqrt{\frac{1 - e^{-2a(t_{i-1} - t_{i-2})}}{2a}} Z_{i-1} \cdot 1_{\{t_{i-1} \in \chi_j\}} e^{-a(t_i - t_{i-1})} + \sqrt{\frac{1 - e^{-2a(t_i - t_{i-1})}}{2a}} Z_i \cdot 1_{\{t_i \in \chi_j\}} \\
&\quad \ldots \\
&= \frac{dT_t}{d\sigma_j} e^{-a(t_1 - t)} + \sum_{k=1}^{i} \sqrt{\frac{1 - e^{-2a(t_k - t_{k-1})}}{2a}} Z_k \cdot 1_{\{t_k \in \chi_j\}} e^{-a(t_i - t_k)}.
\end{aligned}
$$

With the initial condition $\frac{dT_t}{d\sigma_j} = 0$, the above expression simplifies to

$$
\frac{dT_{t_i}}{d\sigma_j} = \sum_{k=1}^{i} \sqrt{\frac{1 - e^{-2a(t_k - t_{k-1})}}{2a}} e^{-a(t_i - t_k)} Z_k \cdot 1_{\{t_k \in \chi_j\}},
$$

where we set $t_0 = t$ when $k = 1$. ■

3.3.3 Sensitivity with respect to the Mean Reversion Parameter

We will denote the derivative of the option price with respect to the mean reversion parameter a by $\Upsilon := \partial c(t; t_1, t_n)/\partial a$. To derive the pathwise estimator, we will first discretize our model, Eqn.(3.1), using the Euler scheme, and then use a recursion to compute derivatives. This technique introduces a bias, due to discretization, however we will slightly abuse language and call the resulting estimator still the "pathwise" estimator.

Proposition 22 *The pathwise estimators of* Υ *for HDD/CDD call and put options are given as follows:*

HDD call

$$\frac{\partial \varphi_{call}(t)}{\partial a} = -e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(\frac{dT_{t_i}}{da} \cdot 1_{\{T_{t_i}<65\}} \right) \cdot 1_{\{H_n>K\}} \right),$$

HDD put

$$\frac{\partial \varphi_{put}(t)}{\partial a} = e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(\frac{dT_{t_i}}{da} \cdot 1_{\{T_{t_i}<65\}} \right) \cdot 1_{\{H_n<K\}} \right),$$

CDD call

$$\frac{\partial \varphi_{call}(t)}{\partial a} = e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(\frac{dT_{t_i}}{da} \cdot 1_{\{T_{t_i}>65\}} \right) \cdot 1_{\{C_n>K\}} \right),$$

CDD put

$$\frac{\partial \varphi_{put}(t)}{\partial a} = -e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(\frac{dT_{t_i}}{da} \cdot 1_{\{T_{t_i}>65\}} \right) \cdot 1_{\{C_n<K\}} \right),$$

where

$$\frac{dT_{t_i}}{da} = \sum_{k=1}^{i} (1-a)^{k-1} (T^m_{t_{i-k}} - T_{t_{i-k}}), \quad i = 1, ..., n, \tag{3.13}$$

and we set $t_0 = t$ *in (3.13) when* $k = i$.

Proof. Existence and interchangeability of operators are proved similar to Proposition 20. We derive the pathwise estimator of Υ for HDD call; derivations for the other options are similar. Note that

$$\begin{aligned} \frac{\partial \varphi_{call}(t)}{\partial a} &= e^{-r(t_n-t)} \left(\frac{\partial H_n}{\partial a} \cdot 1_{\{H_n>K\}} \right) \\ &= -e^{-r(t_n-t)} \left(\sum_{i=1}^{n} \left(\frac{\partial T_{t_i}}{\partial a} \cdot 1_{\{T_{t_i}<65\}} \right) \cdot 1_{\{H_n>K\}} \right). \end{aligned}$$

Discretize the SDE (3.1) using the backwards Euler scheme to obtain

$$T_{t_i} = T^m_{t_i} - T^m_{t_{i-1}} + aT^m_{t_{i-1}} + (1-a)T_{t_{i-1}} + \sigma(t_i) Z_i, \quad i = 1, ..., n, \tag{3.14}$$

where $\{Z_i\}_{i=1}^n$ are i.i.d. standard normal random variables. Differentiating both sides of Eqn.(3.14) recursively, and solving for dT_{t_i}/da, we obtain Eqn.(3.13).

■

3.3.4 Sensitivity with respect to the Long-term Mean Model Parameters

The long-term mean temperature, which is modeled by

$$T_t^m = A + Bt + C\sin(wt + \phi)$$

has three important parameters; the constant term A, the trend B, and the amplitude C. Alaton et al.[1] suggest that meteorological weather forecasts can be used to update these parameters. Therefore, the sensitivity of the option price with respect to these parameters will be useful in measuring the impact of weather forecasts on the option price.

Proposition 23 *The pathwise estimators of the derivative of the option price with respect to the long-term mean model parameters A, B, and C, for HDD/CDD call and put options, are given as follows:*

Sensitivity with respect to constant A:

HDD call

$$\frac{\partial \varphi_{call}(t)}{\partial A} = -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left((1-e^{-a(t_i-t)})\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n>K\}}\right),$$

HDD put

$$\frac{\partial \varphi_{put}(t)}{\partial A} = e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left((1-e^{-a(t_i-t)})\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n<K\}}\right),$$

CDD call

$$\frac{\partial \varphi_{call}(t)}{\partial A} = e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left((1-e^{-a(t_i-t)})\cdot 1_{\{T_{t_i}>65\}}\right)\cdot 1_{\{C_n>K\}}\right),$$

CDD put

$$\frac{\partial \varphi_{put}(t)}{\partial A} = -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left((1-e^{-a(t_i-t)})\cdot 1_{\{T_{t_i}>65\}}\right)\cdot 1_{\{C_n<K\}}\right).$$

Sensitivity with respect to trend B:

HDD call

$$\frac{\partial \varphi_{call}(t)}{\partial B} = -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left((t_i-te^{-a(t_i-t)})\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n>K\}}\right),$$

HDD put

$$\frac{\partial \varphi_{put}(t)}{\partial B} \;=\; e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left((t_i - te^{-a(t_i-t)})\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n<K\}}\right),$$

CDD call

$$\frac{\partial \varphi_{call}(t)}{\partial B} \;=\; e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left((t_i - te^{-a(t_i-t)})\cdot 1_{\{T_{t_i}>65\}}\right)\cdot 1_{\{C_n>K\}}\right),$$

CDD put

$$\frac{\partial \varphi_{put}(t)}{\partial B} \;=\; -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left((t_i - te^{-a(t_i-t)})\cdot 1_{\{T_{t_i}>65\}}\right)\cdot 1_{\{C_n<K\}}\right).$$

Sensitivity with respect to amplitude C***:***

HDD call

$$\frac{\partial \varphi_{call}(t)}{\partial C} = -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{dT_{t_i}}{dC}\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n>K\}}\right),$$

HDD put

$$\frac{\partial \varphi_{put}(t)}{\partial C} = e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{dT_{t_i}}{dC}\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n<K\}}\right),$$

CDD call

$$\frac{\partial \varphi_{call}(t)}{\partial C} = e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{dT_{t_i}}{dC}\cdot 1_{\{T_{t_i}>65\}}\right)\cdot 1_{\{C_n>K\}}\right),$$

CDD put

$$\frac{\partial \varphi_{put}(t)}{\partial C} = -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{dT_{t_i}}{dC}\cdot 1_{\{T_{t_i}>65\}}\right)\cdot 1_{\{C_n<K\}}\right),$$

where dT_{t_i}/dC *is given by*

$$\frac{dT_{t_i}}{dC} = \sin(\omega t_i + \phi) - \sin(\omega t + \phi)e^{-a(t_i-t)}.$$

Proof. Existence and interchangeability of operators are proved similar to Proposition 20. Differentiate the discounted payoff for the HDD call option, $\varphi_{call}(t)$ (see Eqn.(3.4)), with respect to θ (θ=A, B, or C) to get

$$\frac{\partial \varphi_{call}(t)}{\partial \theta} = -e^{-r(t_n-t)}\left(\sum_{i=1}^{n}\left(\frac{dT_{t_i}}{d\theta}\cdot 1_{\{T_{t_i}<65\}}\right)\cdot 1_{\{H_n>K\}}\right).$$

We then compute $dT_{t_i}/d\theta$ by differentiating both sides of Eqn.(3.11) with respect to θ, and then solving the resulting recursion. We obtain:

$$\begin{aligned}
\frac{\partial T_{t_i}}{\partial A} &= 1 - e^{-a(t_i - t)}, \\
\frac{\partial T_{t_i}}{\partial B} &= t_i - t e^{-a(t_i - t)}, \\
\frac{\partial T_{t_i}}{\partial C} &= \sin(\omega t_i + \phi) - \sin(\omega t + \phi) e^{-a(t_i - t)}.
\end{aligned}$$

These results yield the sensitivity of HDD call with respect to the parameters A, B, and C. The sensitivity of the other options are obtained similarly. ■

Remark 24 *The sensitivity with respect to the trend coefficient B given in Proposition 23 depends on how time index is defined in Eqn.(3.1) and Eqn.(3.2). For example, without changing the fitted long-term mean temperature function one can equivalently start the time index from another initial value, say $t = 1000, 1001, 1002, \ldots$ instead of $t = 1, 2, \ldots$. Thus, for the sensitivity with respect to the trend coefficient B, we use the "normalized" estimator given by $\frac{\partial \varphi^*_{call}(t)}{\partial B} := \frac{\partial \varphi_{call}(t)}{\partial B} / \sum_{i=1}^{n} t_i$. In the numerical results in Section 3.5, we will be estimating this normalized estimator. For the coefficients A and C, we do not use any normalization since the dependence is on the increments of time. Following [1] and the standard notation in the literature, we use daily time increments.*

3.4 Second Order Sensitivities of HDD/CDD Options

We define "gamma" as the second derivative of the option price with respect to the current temperature T_t. We cannot apply pathwise method twice directly to obtain gamma, since the delta estimators (3.6)-(3.9) are not Lipschitz continuous with respect to T_t. To compute an estimator of gamma we use the kernel method introduced by Liu and Hong [52].

The next proposition uses the kernel method to derive an estimator to gamma for HDD call options. Estimators for the other options can be obtained similarly.

Proposition 25 *The kernel estimator of gamma for HDD call option is given by*

$$\frac{\partial^2 \varphi_{call}(t)}{\partial T_t^2} = -e^{-r(t_n - t)} \sum_{i=1}^{n} e^{-a(t_i - t)} (I_{1,i} + I_{2,i}),$$

where

$$I_{1,i} = -\frac{1}{L\delta_L} \sum_{l=1}^{L} e^{-a(t_i - t)} \cdot 1_{\{H_{n,l} - K > 0\}} f\left(\frac{65 - T_{t_{i,l}}}{\delta_L}\right)$$

and

$$I_{2,i} = -\frac{1}{L\delta_L}\sum_{l=1}^{L}\left(\sum_{j=1}^{n} e^{-a(t_j-t)} \cdot 1_{\{65>T_{t_{j,l}}\}}\right) f\left(\frac{H_{n,l}-K}{\delta_L}\right).$$

Here $f(\cdot)$ is the standard normal density function; L is the number of independent Monte Carlo simulations of the temperature paths $(T_{t_{1,l}}, T_{t_{2,l}}, ..., T_{t_{n,l}})_{l=1}^{L}$; $(H_{n,l},\ T_{t_{i,l}})$ is the lth observations of $(H_n,\ T_{t_i})$; δ_L is the bandwidth parameter of the kernel.

Proof. We recall the delta pathwise estimator Eqn.(3.6), and then take the expectation under the risk-neutral measure. Delta is given by

$$\frac{\partial c(t)}{\partial T_t} = -e^{-r(t_n-t)} E_Q\left[\sum_{i=1}^{n}\left(e^{-a(t_i-t)} \cdot 1_{\{65>T_{t_i}\}}\right) \cdot 1_{\{H_n>K\}}\middle|\, \mathcal{F}_t\right],$$

The derivative of delta with respect to the current temperature is

$$\frac{\partial^2 c(t)}{\partial T_t^2} = -e^{-r(t_n-t)} \frac{\partial}{\partial T_t} E_Q\left[\sum_{i=1}^{n}\left(e^{-a(t_i-t)} \cdot 1_{\{65>T_{t_i}\}}\right) \cdot 1_{\{H_n>K\}}\middle|\, \mathcal{F}_t\right]. \tag{3.15}$$

We bring the expectation inside the summation in Eqn.(3.15) and rewrite it as

$$\begin{aligned}\frac{\partial^2 c(t)}{\partial T_t^2} &= -e^{-r(t_n-t)} \frac{\partial}{\partial T_t} \sum_{i=1}^{n} e^{-a(t_i-t)} E_Q\left[1_{\{65>T_{t_i}\}} \cdot 1_{\{H_n>K\}}\middle|\, \mathcal{F}_t\right] \\ &= -e^{-r(t_n-t)} \frac{\partial}{\partial T_t} \sum_{i=1}^{n} e^{-a(t_i-t)} E_Q\left[1_{\{\min\{65-T_{t_i}, H_n-K\}\geq 0\}}\middle|\, \mathcal{F}_t\right].\end{aligned}$$

Now we verify Assumption 1 in Theorem 1 of Liu and Hong [52]. Let $\phi_i = \min\{65-T_{t_i}, H_n-K\}$. We claim that ϕ_i is Lipschitz continuous and differentiable almost everywhere, because both H_n-K and $65-T_{t_i}$ are Lipschitz continuous and differentiable almost everywhere, and Lipschitz continuity is preserved under composition. Taking the derivative with respect to T_t, we get

$$\begin{aligned}\frac{d\phi_i}{dT_t} &= -\frac{dT_{t_i}}{dT_t} \cdot 1_{\{65-T_{t_i}<H_n-K\}} + \frac{dH_n}{dT_t} \cdot 1_{\{65-T_{t_i}>H_n-K\}} \\ &= -\frac{dT_{t_i}}{dT_t} \cdot 1_{\{65-T_{t_i}<H_n-K\}} - \left(\sum_{j=1}^{n} e^{-a(t_i-t)} \cdot 1_{\{65>T_{t_j}\}}\right) \cdot 1_{\{65-T_{t_i}\geq H_n-K\}}.\end{aligned}$$

Since the temperature T_t at fixed t follows the normal distribution, Assumption 2 holds according to Lemma 1 in [52]. By Theorem 1 in [52], we have the gamma of HDD call given as

$$\begin{aligned}
&\frac{\partial^2 c(t)}{\partial T_t^2} \\
= \;& e^{-r(t_n-t)} \sum_{i=1}^{n} e^{-a(t_i-t)} \cdot \frac{\partial}{\partial y} E_Q\left[\frac{d\phi_i}{dT_t} \cdot 1_{\{\phi_i > y\}} \middle| \mathcal{F}_t \right]\Bigg|_{y=0} \\
= \;& -e^{-r(t_n-t)} \sum_{i=1}^{n} e^{-a(t_i-t)} \left\{ \frac{\partial}{\partial y} E_Q\left[\frac{dT_{t_i}}{dT_t} 1_{\{65-T_{t_i} < H_n - K\}} \cdot 1_{\{65-T_{t_i} \geq y\}} \middle| \mathcal{F}_t \right]\Bigg|_{y=0} \right. \\
+ \;& \left. \frac{\partial}{\partial y} E_Q\left[\left(\sum_{j=1}^{n} \frac{dT_{t_j}}{dT_t} 1_{\{65-T_{t_j} \geq 0\}} \right) \cdot 1_{\{65-T_{t_i} \geq H_n - K\}} \cdot 1_{\{H_n - K \geq y\}} \middle| \mathcal{F}_t \right]\Bigg|_{y=0} \right\} \\
= \;& -e^{-r(t_n-t)} \sum_{i=1}^{n} e^{-a(t_i-t)} \left\{ \frac{\partial}{\partial y} E_Q\left[\frac{dT_{t_i}}{dT_t} 1_{\{H_n - K > 0\}} \cdot 1_{\{65-T_{t_i} \geq y\}} \middle| \mathcal{F}_t \right]\Bigg|_{y=0} \right. \\
+ \;& \left. \frac{\partial}{\partial y} E_Q\left[\left(\sum_{j=1}^{n} \frac{dT_{t_j}}{dT_t} 1_{\{65-T_{t_j} \geq 0\}} \right) \cdot 1_{\{65-T_{t_i} \geq 0\}} \cdot 1_{\{H_n - K \geq y\}} \middle| \mathcal{F}_t \right]\Bigg|_{y=0} \right\}. \qquad (3.16)
\end{aligned}$$

The kernel estimator of the derivative of the first expectation in Eqn.(3.16), denoted by $I_{1,i}$, is given by

$$I_{1,i} \;=\; -\frac{1}{L\delta_L} \sum_{l=1}^{L} e^{-a(t_i-t)} \cdot 1_{\{H_{n,l} - K > 0\}} f\left(\frac{65 - T_{t_{i,l}}}{\delta_L} \right), \qquad (3.17)$$

where $H_{n,l}$ and $T_{t_{i,l}}$ are the lth independent sample of H_n and T_{t_i}, $l = 1, 2, ..., L$, and L is the number of Monte Carlo simulations. Here we choose the standard normal density function $f(\cdot)$ as the kernel. δ_L is the bandwidth parameter in the kernel which depends on L.

In the same way, we have the kernel estimator $I_{2,i}$ of the derivative of the second expectation in Eqn.(3.16)

$$I_{2,i} = -\frac{1}{L\delta_L} \sum_{l=1}^{L} \left(\sum_{j=1}^{n} e^{-a(t_j-t)} \cdot 1_{\{65 > T_{t_{j,l}}\}} \right) f\left(\frac{H_{n,l} - K}{\delta_L} \right). \qquad (3.18)$$

Finally, we write the estimator of gamma of HHD call by adding Eqn.(3.17) and Eqn.(3.18) together

$$\hat{\Gamma} = -e^{-r(t_n-t)} \sum_{i=1}^{n} e^{-a(t_i-t)} (I_{1,i} + I_{2,i}).$$

In order to the minimize the mean square error of $\hat{\Gamma}$, the asymptotically optimal bandwidth is chosen as $\delta_L^* = c \cdot L^{-1/5}$, where c could be obtained by a pilot simulation. We start with $\delta_L = L^{-1/5}$

and approximate c by

$$\hat{c} = \left(\frac{\hat{V}_L}{4(\hat{\Gamma}''(0))^2} \right)^{\frac{1}{5}}. \tag{3.19}$$

where $\hat{V}_L$ is the sample variance of $\hat{\Gamma}$, and $\hat{\Gamma}''(0)$ is the central difference of the second derivative of $\hat{\Gamma}(0)$. The function $\hat{\Gamma}(u)$ is defined as

$$\hat{\Gamma}(u) = -e^{-r(t_n-t)} \sum_{i=1}^{n} e^{-a(t_i-t)} (I_{1,i}(u) + I_{2,i}(u))$$

with

$$\begin{aligned}
I_{1,i}(u) &= -\frac{1}{L\delta_L} \sum_{l=1}^{L} e^{-a(t_i-t)} \cdot 1_{\{H_{n,l}-K>0\}} \cdot f\left(\frac{65 - T_{t_{i,l}} + u}{\delta_L} \right), \\
I_{2,i}(u) &= -\frac{1}{L\delta_L} \sum_{l=1}^{L} \left(\sum_{j=1}^{n} e^{-a(t_j-t)} \cdot 1_{\{65>T_{t_{j,l}}\}} \right) \cdot f\left(\frac{H_{n,l} - K + u}{\delta_L} \right).
\end{aligned}$$

Then we have the new $\delta_L = \hat{c} \cdot L^{-1/5}$ and use it to estimate $\hat{V}_L$ and $\hat{\Gamma}''(0)$. Repeating this procedure for a number of times, we remove the dependence on the initial choice of δ_L when estimating $\hat{c}$ by Eqn.(3.19). We could achieve this by setting a desired error tolerance to get a stable value of $\hat{c}$ and multiply it by $L^{-1/5}$ as the optimal choice of bandwidth. See details about selection of bandwidth δ in the electronic companion of Liu and Hong [52]. ■

3.5 Numerical Results and Discussions

In Tables 3.4 and 3.5, we consider HDD call and put options for New York, Chicago, and Atlanta, for a contract period of January 1st to 31st, 2006. The strike prices are given as 898, 586, and 1148, for New York, Chicago, and Atlanta, respectively. The tables display the option prices, and their sensitivities, using the estimators derived in Sections 3.3 and 3.4. The numbers in the tables are the sample means of forty independent Monte Carlo estimates, where each Monte Carlo estimate is based on 10,000 sample paths. Each sample path is a simulation of the model, i.e., Eqn.(3.1) and Eqn.(3.2). The numbers in parenthesis are the sample standard deviations of these forty estimates.

Numerical results in Tables 3.4 and Tables 3.5 show that the sensitivity with respect to the mean reversion parameter is very large relative to the rest of the sensitivities, whereas the sensitivities with respect to the constant and seasonality terms (A and C) are of similar magnitude. The

Table 3.4: HDD call options for New York, Chicago and Atlanta for the contract period of January 1st to 31st, 2006 with strike prices given as $K_{NYC} = 898$, $K_{ATL} = 586$, and $K_{CHI} = 1148$, respectively. The sensitivities are estimated via 40 independent Monte Carlo simulations, where each simulation has a sample size 10000. Sample standard deviation of 40 estimates is given in parenthesis. The function $c(\cdot)$ is the call option price.

Strike	New York 898	Atlanta 586	Chicago 1148
Call Price	37.7709	49.7990	44.7511
delta	-1.1900	-1.6412	-1.4112
	(0.0142)	(0.0157)	(0.0151)
gamma	0.0247	0.0350	0.0283
	(0.0069)	(0.0075)	(0.0080)
vega	5.9739	7.9138	6.8044
	(0.0982)	(0.0832)	(0.1047)
$\frac{\partial c(\cdot)}{\partial a}$	-100.4250	-173.8490	-133.8651
	(1.7121)	(1.8993)	(1.9567)
$\frac{\partial c(\cdot)}{\partial A}$	-14.2230	-13.5927	-14.0672
	(0.1693)	(0.1310)	(0.1505)
$\frac{\partial c(\cdot)}{\partial B}$	-0.4590	-0.4387	-0.4540
	(0.0055)	(0.0042)	(0.0042)
$\frac{\partial c(\cdot)}{\partial C}$	14.0294	13.4818	13.9484
	(0.1670)	(0.1299)	(0.1492)

sensitivities with respect to the long-term mean temperature parameters are almost identical for all three cities. Therefore, from this aspect there is no differentiation between these contracts. However, there is more of a differentiation between contracts with respect to their delta, gamma, vega, and mean reversion parameter. For example, suppose the trader has a temperature forecast that shows increasing temperature volatility or deviations from long-term mean levels. Given this belief or forecast the trader might choose to long the Atlanta HDD option, which has the highest vega out of these three cities. The delta of the New York HDD option is smallest in absolute terms, suggesting a smaller sensitivity with respect to possible measurement or forecast errors in the initial temperature. In case of partial hedging with a correlated asset, a lower gamma means less re-balancing of the partial-hedging portfolio and possibly lower transaction costs. Results show that the gamma of all three options are low with New York HDD call/put options having the lowest value.

Table 3.5: HDD put options for New York, Chicago and Atlanta for the contract period of January 1st to 31st, 2006 with strike prices given as $K_{NYC} = 898$, $K_{ATL} = 586$, and $K_{CHI} = 1148$, respectively. The sensitivities are estimated via 40 independent Monte Carlo simulations, where each simulation has a sample size 10000. Sample standard deviation of 40 estimates is given in parenthesis. The function $p(\cdot)$ is the put option price.

Strike	New York 898	Atlanta 586	Chicago 1148
Put Price	38.3321	50.5335	44.6165
delta	1.2017	1.6539	1.4131
	(0.0122)	(0.0173)	(0.0133)
gamma	0.0247	0.0350	0.0283
	(0.0069)	(0.0075)	(0.0080)
vega	5.9816	6.1302	6.8081
	(0.0810)	(0.1061)	(0.0931)
$\frac{\partial p(\cdot)}{\partial a}$	-100.6036	-153.7170	-134.0439
	(1.3990)	(2.5794)	(1.8398)
$\frac{\partial p(\cdot)}{\partial A}$	14.3619	13.5108	14.0859
	(0.1454)	(0.1412)	(0.1330)
$\frac{\partial p(\cdot)}{\partial B}$	0.4635	0.4387	0.4540
	(0.0055)	(0.0042)	(0.0049)
$\frac{\partial p(\cdot)}{\partial C}$	-14.1665	-13.4005	-13.9670
	(0.1434)	(0.1401)	(0.1319)

APPENDIX A

ARBITRAGE PRICING AND STOCHASTIC CALCULUS

We summarize some definitions and theorems from Brigo and Mercurio [12] and Shreve [72]. We consider a stochastic economy, where the uncertainty is represented by a probability space $(\Sigma,\ \mathcal{F}, P)$. The flow of information accruing is represented by a filtration $\mathcal{F}_t$. The filtration $\{\mathcal{F}_t : t \in [0,\ \infty)\}$ is right continuous, and $\mathcal{F}_0$ contains all the P null sets of $\mathcal{F}$ (See Protter [62]).

Consider a portfolio that consists of underlying assets $S_1,\ S_2,\ \ldots,\ S_n$. Its value $V(t)$ associated with the strategy $\phi_1(t)\ ,\ \ldots,\ \phi_n(t)$ is given by

$$V_\phi(t) = \sum_{i=1}^{n} \phi_i(t) S_k(t), \tag{A.1}$$

where the processes $\phi_i(t)$ are locally bounded and predictable. The gain process associated with a strategy ϕ is given by

$$G_\phi(t) = \sum_{i=1}^{n} \int_0^t \phi_i(u) dS_i(u)\ ,\ 0 \le t < T. \tag{A.2}$$

Definition 26 *A trading strategy ϕ is self-financing or a portfolio is self-financing, if $V_\phi(t) \ge 0$ and*

$$V_\phi(t) = V_\phi(0) + G_\phi(t)\ ,\ 0 \le t < T. \tag{A.3}$$

Intuitively, we say a strategy is self-financing or a portfolio is self-financing, if its value changes only due to changes in the underlying asset prices. In other words, no additional cash inflows or outflows occur after the initial time. A key result in Harrison and Kreps [41], and Harrison and Pliska ([42] and [43]) established connection between the economic concept of absence of arbitrage and the mathematical property of existence of a probability measure, the equivalent martingale measure (or risk-neutral measure, or risk-adjusted measure), whose definition is given in the following.

Definition 27 *An equivalent martingale measure Q is a probability measure on the space $(\Omega, \mathcal{F})$ such that*

1. *P and Q are equivalent measures, that is $P(A) = 0$ if and only if $Q(A) = 0$, for every $A \in \mathcal{F}$;*

2. *The Radon-Nikodym derivative dQ/dP belongs to $L^2(\Omega, \mathcal{F}, P)$;*

3. *The discounted process $D(0, \cdot)S$ is an $(\mathcal{F}, Q)$-martingale.*

An arbitrage opportunity is defined, in mathematical terms, as a self-financing strategy ϕ *such* that $V_\phi(0) = 0$ but $P(V_\phi(T)) > 0$. Harrison and Pliska [27] then proved the fundamental result that the existence of an equivalent martingale measure implies the absence of arbitrage opportunities.

Proposition 28 *A contingent claim H is attainable if there exists some self-financing ϕ such that $V_\phi(T) = H(T)$.*

If Ω is finite, there is no restriction on the class of portfolios, and the no arbitrage assumption is equivalent to the existence of "risk-neutral measure". (See Harrison and Kreps [41], Harrison and Pliska [42]). The following proposition, proved in [42], provides with mathematical characterization of the unique no arbitrage price associated with any attainable contingent claim.

Proposition 29 *Assume there exists an equivalent martingale measure Q and let H be an attainable contingent claim. Then, for each time t, $0 \leq t \leq T$, there exists a unique price $V(t)$ associated with H, i.e.,*

$$V(t) = \mathrm{E}_Q\left[\frac{B(t)}{B(T)} H | \mathcal{F}_t\right], \tag{A.4}$$

where $B(t) = \exp(\int_0^t r(s)ds)$ is the money-market account and $r(s)$ is the short rate.

Definition 30 *A financial market is complete if and only if every contingent claim is attainable.*

Theorem 31 ***(Girsanov Theorem, one dimension)*** *[72] We first define a new probability measure Q by the formula*

$$Q(A) = \int_A Z(\omega)dP(\omega), \quad for\ all\ A \in \mathcal{F}, \tag{A.5}$$

then Q is equivalent to P. Let $W(t)$, $0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}$, $0 \leq t \leq T$, be a filtration for this Brownian motion. Let $\gamma(t)$, $0 \leq t \leq T$, be an adapted process. Define

$$\begin{aligned} Z(t) &= \exp\left(-\int_0^t \gamma(u)dW(u) - \frac{1}{2}\int_0^t \gamma^2(u)du\right), \\ \tilde{W}(t) &= W(t) + \int_0^t \gamma(u)du, \end{aligned}$$

and assume that

$$E\left[\int_0^T \gamma^2(u)Z^2(u)du\right] < \infty.$$

Set $Z = Z(T)$. *Then* $EZ = 1$ *and under the probability* Q *defined in (A.5), the process* $\tilde{W}_t$, $0 \le t \le T$, *is a Brownian motion.*

Theorem 32 ***Lévy Theorem*** *[72]: Let* $M(t)$, $t \ge 0$, *be a martingale relative to filtration* $\mathcal{F}(t), t > 0$. *Assume that* $M(0) = 0$, $M(t)$ *has continuous paths, and* $[M, M](t) = t$ *for* $t \ge 0$. *Then* $M(t)$ *is a Brownian motion.*

Definition 33 ***Radon-Nikodym Derivative Process****: Suppose we have a probability space* $(\Omega, \mathcal{F}, P)$ *and a filtration* $\mathcal{F}(t)$, *defined for* $0 \le t \le T$, *where* T *is a fixed final time. Suppose further that* Z *is an almost surely positive random variable satisfying* $EZ = 1$, *and we define a new probability measure* $\tilde{P}$ *by the formula*

$$\tilde{P}(A) = \int_A Z(\omega)dP(\omega), \quad for\ all\ A \in \mathcal{F}. \tag{A.6}$$

We say Z *is the Radon-Nikodym derivative of* $\tilde{P}$ *with respect to* P, *and we write* $Z = d\tilde{P}/dP$. *We can then defined the Radon-Nikodym derivative process*

$$Z(t) = E[Z|\mathcal{F}(t)], \qquad 0 \le t \le T. \tag{A.7}$$

Theorem 34 ***Lebesgue's Dominated Convergence Theorem*** *[6]: If* $|f_n| \le g$ *almost everywhere, where* g *is integrable, and if* $f_n \to f$ *almost everywhere, then* f *and* f_n *are integrable and* $\int f_n d\mu \to \int f d\mu$.

APPENDIX B

CHOLESKY DECOMPOSITION

We summarize the Cholesky decomposition in this section (see [51] for details). Cholesky decomposition is a special version of LU decomposition tailored to handle symmetric matrices more efficiently. For a symmetric matrix A, by definition, $a_{i,j} = a_{j,i}$. LU decomposition is not efficient enough for symmetric matrices. The computational load can be halved using Cholesky decomposition. Using the fact that A is symmetric, write $A = LL'$, where L is a lower triangular matrix. That is,

$$L = \begin{pmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{pmatrix}.$$

Note that the diagonal elements of L are not ls as in the case of LU decomposition. With Cholesky decomposition, the elements of L are evaluated as follows:

$$l_{kk} = \sqrt{a_{kk} - \sum_{j--1}^{k-1} l_{k,j}^2}, \quad k = 1, 2, \ \ldots, \ n \tag{B.1}$$

$$l_{kl} = \frac{1}{l_{ii}} \left(a_{kl} - \sum_{j=1}^{j-1} l_{ij} l_{kj} \right), \quad k = 1, 2, \ \ldots, \ k-1 \tag{B.2}$$

where the first subscript is the row index and the second one is the column index.

Cholesky decomposition is evaluated column by column (starting from the first column) and, in each row, the elements are evaluated from top to bottom. That is, in each column the diagonal element is evaluated first using (B.1) (the elements above the diagonal are zero) and then the other elements in the same row are evaluated next using (B.2). This is carried out for each column starting from the first one. Note that if the value within the square root in (B.1) is negative, Cholesky decomposition will fail. However, this will not happen for positive semidefinite matrices, which are encountered commonly in many engineering systems (e.g., circuits, covariance matrix). Thus, Cholesky decomposition is a good way to test for the positive semi-definiteness of symmetric matrices.

APPENDIX C

FAURE SEQUENCE

We describe how to generate quasi-random sequence and describe a particular class of quasi-random sequences known as Faure sequences (see [25, 50, 38, 59] for details). Faure sequences have certain advantages for the valuation of high dimensional integrals. This leads to efficient algorithms for the computation of prices and price sensitivities for complex derivative securities.

To compute the nth element of the Faure sequence, we start by representing any integer n in terms of the base b,

$$n = \sum_{j=0}^{m} a_j^1 b^j.$$

The first element of Faure sequence is given by reflection about the decimal point as before,

$$\phi_b^1(n) = \sum_{j=0}^{m} a_j^1 b^{-j-1}.$$

The remaining elements of the sequence can be found recursively. As also explained in Joy et al. [50], given $a_j^{k-1}(n)$ the next term can be written as

$$a_j^k(n) = \sum_{j=1}^{m} C_j^k a_i^{k-1}(n) \ (mod\ b),$$

where $C_j^i = \dfrac{i!}{j!(i-j)!}$. Hence, to generate the next dimension in the sequence we multiply with the generator matrix, which is the upper triangular matrix with entries as the binomial coefficients,

$$\begin{pmatrix} C_0^0 & C_0^1 & C_0^2 & C_0^3 & \cdots \\ 0 & C_1^1 & C_1^2 & C_1^3 & \cdots \\ 0 & 0 & C_2^2 & C_2^3 & \cdots \\ 0 & 0 & 0 & C_3^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Hence, the rest of the points can be generated by,

$$\phi_b^k(n) = \sum_{j=0}^{m} a_j^k(n) b^{-j-1}, \quad 2 \le k \le d.$$

We can represent the nth element of the d dimensional Faure sequence by,

$$\phi_n = \ (\phi_n^1,\ \phi_n^2,\ \ldots,\ \phi_n^d).$$

There are variations for the construction of the Faure sequence, such as the Generalized Faure Sequence. Generalized Faure sequence is formulated by Tezuka [77], which is based on the reordering of the Halton sequence by polynomials. In the literature Generalized Faure Sequence was used by Traub and Papageorgiou [61], and claimed to have faster convergence than the simple Faure sequence.

Linear scrambled Faure sequence. The Faure sequence is an example of a *digital* $(t,\ s)$ *sequence* with $t = 0$. The digital sequences comprise a big portion of low-discrepancy sequences used in practice. Consider an s-dimensional digital sequence $q_1,\ q_2, \ldots$, in the s-dimensional unit hypercube $[0, 1)^s$. To compute the nth term $q_n = (q_n^{(1)},\ \ldots,\ q_n^{(s)})$ of the sequence, we first write n in its base p expansion: $n = a_l p^{l-1} + \ldots + a_2 p + a_1$. Here p is called the base of the digital sequence. In the case of the Faure sequence, the base is the smallest prime number greater than or equal to s. Let $d_n = (a_1,\ \ldots,\ a_l)'$ be the (column) vector of the digits of n. Similarly, let $\phi(q_n^{(k)})$ denote the (column) vector of the digits of the real number $q_n^{(k)} \in [0, 1)$ in base p. For example, if the base p expansion of a real number is

$$\tilde{a}_{\tilde{l}}/p^{\tilde{l}} + \ldots + \tilde{a}_2/p^2 + \tilde{a}_1/p,$$

then

$$\phi(\cdot) = (\tilde{a}_1,\ \ldots,\ \tilde{a}_{\bar{l}})'.$$

The kth component of the nth term q_n is defined through its digit vector:

$$\phi(q_n^{(k)}) = \mathbf{C}^{(k)} d_n, \quad k = 1,\ \ldots,\ s \tag{C.1}$$

where $\mathbf{C}^{(k)}$ is the kth *generator matrix* of the digital sequence, and the matrix-vector multiplication is done in modular arithmetic mod p. The generator matrix for the Faure sequence is

$$\mathbf{C}^{(k)} = \mathbf{P}^{k-1}, \quad k = 1,\ \ldots,\ s \tag{C.2}$$

where $\mathbf{P}$ is the $l \times l$ Pascal matrix mod p: this is an upper triangular matrix where the ijth entry is the binomial term C_{j-1}^{i-1} mod p. For example, the 3×3 Pascal Matrix is

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Powers of the Pascal matrix can be computed without matrix multiplication as follows. Define $\mathbf{P}(a)$ as the upper triangular matrix with the ijth entry equal to $a^{j-i}C_{j-1}^{i-1}$. Then, we have the following useful identity : $\mathbf{P}^k = \mathbf{P}(k)$ for any integer k. (adopting the convention $0^0 = 1$, so that $\mathbf{P}^0 = \mathbf{P}(0) = \mathbf{I}$).

We now describe how to obtain randomly scrambled "copies" of the Faure sequence. Generate s lower triangular square matrices $\mathbf{L}^{(k)},\ k = 1,\ \ldots,\ s$, whose non-diagonal entries are random integers between 0 and $p-1$, and diagonal entries are random integers between 1 and $p-1$. Also generate s vectors $g^{(k)}$ whose entries are random integers between 1 and $p-1$. The dimension of the square matrices and the vectors is equal to l; the dimension of the vector d_n. A scrambled Faure sequence is obtained by replacing the generator matrices of the Faure sequence (C.2).

$$\mathbf{C}^{(k)} = \mathbf{L}^{(k)}\mathbf{P}^{k-1},\ k = 1,\ \ldots,\ s$$

and replacing (C.1) by

$$\phi(q_n^{(k)}) = \mathbf{C}^{(k)}d_n + g^{(k)},\ k = 1,\ \ldots,\ s.$$

As before, all operations are done in mod p. This scrambling method is called linear scrambling by Matoušek [54], who discusses this as well as other scrambling techniques. A detailed discussion of scrambled sequences and their implementation is also given by Hong and Hickernell [45].

BIBLIOGRAPHY

[1] P. Alaton, B. Djehiche, and D. Stillberger. On modelling and pricing weather derivatives. *Applied Mathematical Finance*, 9(1):1–20, 2002.

[2] S. Asmussen and P. Glynn. *Stochastic Simulation: Algorithms and Analysis.* New York, 2007.

[3] D. Bayazit and C. Nolder. Sensitivities of options via Malliavin calculus: applications to markets of exponential variance gamma and normal inverse gaussian processes. *Quantitative Finance*, 13(8):1257–1287, 2013.

[4] F. Benth and J. Benth. Stochastic modelling of temperature variations with a view towards weather derivatives. *Applied Mathematical Finance*, 12(1):53–85, 2005.

[5] F. Benth and J. Benth. The volatility of temperature and pricing of weather derivatives. *Applied Mathematical Finance*, 7(5):553–561, 2007.

[6] P. Billingsley. *Probability and Measure, 3rd Edition.* Chicago, 1995.

[7] J. Birge. *Quasi-Monte Carlo approaches to option pricing.* Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, 1994.

[8] D. Bosq. *Nonparametric Statistics for Stochastic Processes.* Springer; 2nd edition.

[9] P. Boyle. Options: A Monte Carlo approach. *Journal of Financial Economics*, 4:323–338, 1977.

[10] P. Boyle, M. Broadie, and P. Glasserman. Monte Carlo mehtod for security pricing. *Journal of Economic Dynamics and Control*, 21:1267–1321, 1997.

[11] P. Boyle and Y. Tse. An algorithm for computing values of options on the maximum or minimum of several assets. *Journal of Financial and Quantitative Analysis*, 2(25).

[12] D. Brigo and F. Mercurio. *Interest Rate Models-Theory and Practice.* Springer Finance, Germany, 2001.

[13] M. Broadie and P. Glasserman. Estimating security price derivatives using simulation. *Management Science*, 42(2):269–285, 1996.

[14] M. Broadie and P. Glasserman. Pricing American-style securities using simulation. *Journal of Economic Dynamics and Control*, 21(8-9):1323–1352, 1997.

[15] D. Brody, J. Syroka, and M. Zervos. Dynamical pricing of weather derivatives. *Quantitative Finance*, 2(3):189–198, 2002.

[16] S. Campbell and F. Diebold. Weather forecasting for weather derivatives. *Journal of the American Statistical Association*, 100(469):6–16, 2005.

[17] M. Cao and J. Wei. Weather derivatives valuation and market price of weather risk. *The Journal of Future Markets*, 24(11):1065–1089, 2004.

[18] X. Cao and Y. Ho. Optimization and perturbation analysis of queueing networks. *Journal of Optimization: Theory and Applications*, 40:559–582, 1983.

[19] Y. Cao and M. Fu. Estimating greeks for variance gamma. *In Proceedings of the 2010 Winter Simulation Conference*, 40:243–250, 2010.

[20] J. Chen and M. Fu. Efficient sensitivity analysis of mortgage backed securites. *12 th Annual Derivative Securities Conference*, 2002.

[21] N. Chen and P. Glasserman. Malliavin greeks without Malliavin calculus. *Stochastic Processes and their Applications*, 117(2):1689–1723, 2007.

[22] N. Chen and Y. Liu. American option sensitivities estimation via a generalized infinitesimal perturbation analysis approach. *Operations Research*, 62(3):616–632, 2014.

[23] P. Embrechts, A. McNeil, and D. Straumann. Correlation: Pitfalls and alternatives. *Risk*, 5, 1999.

[24] K. Eriksson, D. Estep, and C. Johnson. *Applied Mathematics Body and Soul, Volume 1: Derivatives and Geometry in R3*. Springer-Verlag, New York, 2004.

[25] H. Faure. Discrépence de suites associées à un système de numération (en dimensions). *Acta Arithmetica*, 41:337–351, 1982.

[26] E. Fournie, L-M. Lasry, J. Lebuchoux, P-L. Lions, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3:391–412, 1999.

[27] B. Fox and P. Glynn. Replication schemes for limiting expectations. *Probability in the Engineering and Information Sciences*, 3:299–318, 1989.

[28] A. Frolov and N. Chentsov. On the calculation of definite integrals dependent on the parameter by the Monte Carlo method. *USSR Journal of Computational Mathematics and Mathematical Physics*, 4:802–808, 1963.

[29] M. Fu and J. Hu. *Conditional Monte Carlo, gradient estimation and optimization applications.* Kluwer Academic Publishs, Boston, 1997.

[30] A. Genz. Numerical computation of rectangular bivariate and trivariate normal and probabilities. *Statistics and Computing*, 14(3):251–260, 2004.

[31] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Cambridge University Press, Spring-Verlag, 2005.

[32] P. Glasserman and Z. Liu. Estimating greeks in simulating Lévy-driven models. *Journal of Computational Finance*, 14(2):3–56, 2010.

[33] P. Glasserman and X. Zhao. Fast greeks by simulation in forward libor models. *Journal of Computational Finance*, 3:5–39, 1999.

[34] P. Glynn. Likelihood ratio gradient estimation: an overview. *Proceedings of the 1987 Winter Simulation Conference*, 8:366–374, 1987.

[35] P. Glynn. Optimization of stochastic systems via simulation. *Proceedings of the 1989 Winter Simulation Conference, IEEE Process, New York*, pages 90–105, 1989.

[36] E. Gobet and A. Kohatsu-Higa. Computation of greeks for barrier and lookback options using Malliavin calculus. *Electronic Communications in Probability*, 8:5162, 2003.

[37] A. Göncü. Pricing temperature-based weather derivatives in China. *Journal of Risk Finance*, 13:32–44, 2011.

[38] A. Göncü. Monte Carlo simulation of a two-factor stochastic volatility model. *Proceedings of the International MultiConference of Engineers and Computer Scientists IMECS*, Vol II:IMECS March 12–16, HongKong, 2012.

[39] A. Hall. Gradient estimation and mountain range options. *PhD thesis, University of Maryland, College Park*, 2003.

[40] J. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals. *Numerische Mathematik*, 2:84–90, 1960.

[41] J. Harrision and D. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

[42] J. Harrision and S. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:251–260, 1981.

[43] J. Harrision and S. Pliska. A stochastic calculus model of continuous trading: Complete markets. *Stochastic Processes and their Applications*, 15:313–316, 1983.

[44] D. Hertz. Risk analysis in capital investment. *Harvard Business Review*, 4(2):327–343, 1964.

[45] H. Hong and F. Hickernell. Algorithm 823: Implementing scrambled digital sequences. *ACM Transactions on Mathematical Software*, 29:95–109, 2003.

[46] H. Huang, Y. Shiu, and P. Lin. HDD and CDD option pricing with market price of weather risk for Taiwan. *Journal of Futures Markets*, 28(8):790–814, 2008.

[47] J. Hull. *Option, Futures and Other Derivative Securities,6th ed.* Prentice-Hall, Englewood Cliffs, NJ, 2006.

[48] S. Jewson. *Weather Derivative Valuation: The Meteorological, Statistical, Financial and Mathematical Foundations.* Cambridge University Press, Englewood Cliffs, NJ, 2005.

[49] H. Johnson. Options on the maximum or minimum of several assets. *Journal of Financial and Quantitative Analysis*, 22:277–283, 1987.

[50] C. Joy, P. Boyle, and K. Tan. Quasi-Monte Carlo methods in numerical finance. *Management Science*, 42(6):926–938, 1996.

[51] T. Kirubarajan. Lecture notes of computer-aided engineering. *www.ece.mcmaster.ca/ kiruba /3sk3/lecture7.*

[52] G. Liu and L. Hong. Kernel estimation of greeks for options with discontinuous payoffs. *Operations Research*, 59(1):96–108, 2011.

[53] F. Longstaff and E. Schwartz. Valuing American options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 14:113–147, 2001.

[54] J. Matouek. On the L2-discrepancy for anchored boxes. *Journal of Complexity*, 14:527–556, 1998.

[55] B. Moskowitz and R. Caflisch. Smoothness and dimension reduction in quasi-Monte Carlo methods. *Math. Comp. Modeling*, 23:37–54, 1996.

[56] H. Niederreiter. Low discrepancy and low dispersion sequences. *Journal of Number Theory*, 30:51–70, 1988.

[57] H. Niederreiter. Random number generation and quasi-Monte Carlo methods. *CBMS-NSF 63, SIAM*, pages Philadelphia, PA, 1992.

[58] T. Oetomo and M. Stevenson. Hot or cold? a comparison of different approaches to the pricing of weather derivatives. *Journal of Emerging Market Finance*, 4(2):101–133, 2005.

[59] G. Ökten. *Lecture notes of Monte Carlo simulation in Finance.*

[60] M. Overhaus. Himalaya options, technical report. *www.risk.net*, March, 2002.

[61] A. Papageorgiou and J. Traub. Faster valuation of multidimensional integrals. *Computers in Physics*, 11(6):574–578, 1997.

[62] P. Protter. *Stochastic Integration and Differential Equations, the second Edition*. Springer-Verlag, New York, 2004.

[63] R. Quessette. New products, new risks, technical report. *www.risk.net*, March, 2002.

[64] M. Reiman and A. Weiss. Sensitivity analysis for simulations via likelihood ratios. *Operations Research*, 37:830–844, 1989.

[65] M. Reiman and A. Weiss. A fair pricing approach to weather derivatives. *Asian-Pacific Financial Markets*, 11(1):23–53, 2005.

[66] T. Richards, M. Manfredo, and D. Sanders. Pricing weather derivatives. *American Journal of Agricultural Economics*, 86(4):1005–1017, 2004.

[67] P. Ritchken. On pricing barrier options. *Journal of Derivatives*, 3:19–28, 1995.

[68] R. Rubinstein. Sensitivity analysis and performance extrapolation for computer simulation models. *Operations Research*, 37:72–81, 1989.

[69] R. Rubinstein and A. Shapiro. *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization*. Wiley, New York, 1993.

[70] F. Schiller, G. Seidler, and M. Wimmer. Temperature models for pricing weather derivatives. *Quantitative Finance*, 12(3):489–500, 1995.

[71] M. Scholes and F. Black. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

[72] S. Shreve. *Stochastic Calculus for Finance II*. Springer-Verlag, New York, 2004.

[73] I. Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4):86–112, 1967.

[74] S. Stojanovic. *Neutral and Indifference Portfolio Pricing, Hedging and Investing:With applications in Equity and FX*. Springer, New York, 2012.

[75] R. Stulz. Options on the miminum or maximum of two risky assets: analysis and application. *Journal of Financial Economics*, 10:161–185, 1982.

[76] R. Suri and M. Zazanis. Perturbation analysis gives strongly consistent estimates for the M/G/1 queue. *Management Sci.*, 34:39–64, 1988.

[77] S. Tezuka. A generalization of Faure sequences and its efficient implementation. *IBM Tokyo Research Labratory*, 1994.

[78] S. Tezuka. *Uniform Random Numbers: Theory and Practice.* Kluwer Academic Publishers, Boston, 1995.

[79] J. Tilley. Valuing American options in a path simulation model. *Transactions of Society of Actuaries*, 45:83–104, 1993.

[80] Y. Wang and R. Caflisch. Pricing and hedging American-style options: A simple simulation-based approach. *The Journal of Computational Finance*, 13:95–125, 2010.

[81] A. Zapranis and A. Alexandridis. Modelling the temperature time-dependent speed of mean reversion in the context of weather derivatives pricing. *Applied Mathematical Finance*, 15(4):355–386, 2008.

[82] M. Zazanis. Statistical properties of perturbation analysis estimates for discrete event systems, doctoral dissertation. 1987.

[83] L. Zeng. Pricing weather derivatives. *Journal of Risk Finance*, 1(3):72–78, 2000.

BIOGRAPHICAL SKETCH

Wei Yuan

Wei Yuan was born in Shenyang, China. She is currently in her doctoral study in the Financial Mathematics program at Florida State University under the advisement of Professor Giray Ökten and Professor Kyounghee Kim. In the summer of 2015, she is expecting to graduate with a PhD degree in Financial Mathematics. Her research interests include estimating sensitivities using Monte Carlo methods, financial derivative pricing and fixed income research.

Before Wei enrolled in the doctoral program at Florida State University, she completed her Bachelor degree in Computational Mathematics at Northeastern University in China, and obtained a Master degree in Applied Mathematics at Memorial University of Newfoundland in Canada.