INVERTING MULTIVARIATE ANALYTIC CHARACTERISTIC FUNCTIONS WITH FINANCIAL APPLICATIONS

BY

RUNQI HU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Industrial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Liming Feng, Chair
Professor Alexandra Chronopoulou
Professor Justin Sirignano
Professor Runhuan Feng

Abstract

This dissertation is devoted to multivariate analytic characteristic functions inversion and applications in option pricing, option sensitivities estimation, and some electronic engineering problems. We will show that under certain analytic conditions for characteristic functions, the underlying pdfs and cdfs have exponential tails. The inversion from multivariate characteristic functions to the corresponding pdfs and cdfs can be approximated by the trapezoidal rule conveniently with great accuracy. Monte Carlo methods can be applied for option sensitivity analysis. Under multi-dimensional models, acceptance-rejection method is desirable. Simulating from a distribution without explicit pdf or CDF is then transformed to sampling from an easy-to-simulate distribution. Detailed algorithms are provided and comparisons against classical methods in terms of accuracy and efficiency are included.

To My Parents, who are always supportive to me

Acknowledgements

I have to start by expressing my sincere gratitude to my advisor Professor Liming Feng for his support of my research and work, also for his motivation, encouragement, and trust. Professor Feng's mentorship guided me through all the difficulties in research.

Besides, I would like to thank my committee members, Professors Alexandra Chronopoulou, Runhuan Feng, and Justin Sirignano, for providing valuable comments and advices on my research and thesis.

My deepest gratitude also goes to my friends: Liang Ma for giving me suggestions on my PhD life, Pingping Jiang for accompanying me through happiness and sorrow, Yijing, Yue, Shenghan, Jia, Kaihao, Yuan, Shenghui, Professor Tang, and Jane for spending the happiest moments with me, Bo and Fang, for all the fun we had together, Yilun, Lily, and Qi for shouldering my burden in Shenzhen.

In particular, I'm grateful to Zhuo for being supportive on every decision I make, and for sharing and listening.

Finally, I would like to thank my parents and other family members, who are always a great comfort to me and continuously support me both financially and spiritually. Especially, I would like to thank my cousin, Xinzhu. Without her, I wouldn't be able to make it. Their encouragement and recognition has made my life and study enjoyable and meaningful.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

As the most widely used option pricing model, Black-Scholes-Merton model [12], [38] assumes that the asset price process follows a geometric Brownian motion, $S_t = S_0 e^{\mu t + \sigma W_t}$, where the initial asset price is denoted by S_0, and W_t represents the standard Brownian motion. A major reason behind the popularity of the Black-Scholes-Merton model is that it provides a closed-form expression for the price of a European option. Numerical methods can be conveniently implemented in pricing American or exotic options. From the trading and risk management perspective, European option Greeks can be derived explicitly.

However, Black-Scholes-Merton model has its own limitations since it fails to capture some market phenomena. Black-Scholes-Merton model assumes constant volatility, σ, across different strikes and maturities, and thus, leads to a flat volatility curve and surface. In reality, an option's implied volatility varies by moneyness, the distance between the strike and the current underlying asset price, and maturity. Such phenomenon is called volatility smile or skew. Volatility smile is common on the foreign currency options market, while volatility skew appears frequently on index and equity options market.

Non-Gaussian Lévy processes, including pure jump and jump-diffusion asset pricing models based on Lévy processes, have then enjoyed their remarkable popularity. They not only explain volatility smile and skew in a reasonable way [1], but also provide better fit to time series data of financial assets including equities, commodities, and foreign currencies.

To capture extreme price movements in asset prices, Merton's Jump-Diffusion Model [39] models the arrivals of important information with a compound Poisson process and the magnitude of the random jump with a normal distribution. Another Lévy process-based model is Kou's double exponential jump model [31–33], of which the jump component is composed of asymmetric double exponential jumps driven by a Poisson process.

The variance gamma (VG) model of Madan and Seneta [37], Madan and Milne [36] and Madan, Carr, and Chang [35] results from a normal distribution conditional on a variance that is distributed as a gamma variable. Besides, the normal inverse Gaussian (NIG) model [8], the Carr, Geman, Madan and Yor (CGMY) model [13], and the finite moment log-stable model [16] are other infinite activity pure jump models.

By allowing the instantaneous variance of the diffusion part of the price process to be random, many stochastic volatility models have been proposed. In the Heston [29] model, the variance is modeled by a Cox-Ingersoll-Ross (CIR) square-root process. Based on the Heston model, the stochastic volatility jump (SVJ) model proposed by Duffie [9] includes jumps in the asset return process. This model was then extended by the SVCJ model [23] by allowing contemporaneous and correlated jumps in both return and variance processes.

One major reason behind the popularity of non-Gaussian Lévy processes is that even though the majority of these models do not have closed-form expressions for the pdfs and cdfs, by the Lévy-Khintchine formula, they admit explicit characteristic functions

with certain analyticity, leading to the tractability of the processes and exponential tail behaviors of the underlying pdfs and cdfs [34]. One can easily obtain volatility smiles or skews from these models.

The thesis starts with the discussion in univariate cases in Chapter 2, with the classical definition of the characteristic function on the real axis. Then we extend the definition to complex variables since we are interested in its analyticity in the complex plane, or more specifically, its analyticity in a horizontal strip including the real axis . In Section 2.1, we analyze the relations between analyticity of the characteristic function in the horizontal strip and tail behaviors of the underlying pdf and cdf. This lays theoretical foundations for our acceptance-rejection simulation [17] and approximation for option prices. By inverting analytic characteristic functions, one can approximate pdfs and cdfs conveniently with great accuracy by the trapezoidal rule. The total approximation error bounds are provided in Section 2.4.

The analysis for bivariate cases is presented in Chapter 3. We obtain expressions for the bivariate cdf and pdf approximations in Section 3.2 and analyze the total approximation error bounds. Numerical experiments are performed in Section 3.3. In pricing 2d European options, one needs to approximate expectations involving indicator functions. We provide alternative expressions and approximations for such representations. We also price a 2d European call option under the bivariate NIG model in the numerical experiments part.

For pricing non-European options, especially for path-dependent contracts, and in estimating option Greeks, Monte Carlo methods are often attractive. In Chapter 4, we propose a tabulation method based on LRM, which estimates sensitivities by differentiating the probability density function inside the pricing integral. Even though that pdfs of some widely used financial models, including Lévy process models [8], [13], [29], and jump-diffusion models [23], may not be available, their characteristic functions are often

accessible to us. In [27], cumulative distribution functions (cdfs) and their derivatives with respect to certain parameters are inverted and approximated through Laplace transform on selected grid points. Then the approximated cdf and its derivative are constructed by linear interpolation and hence are piecewise linear. Differentiating these functions with respect to a specific parameter gives piecewise constant approximations for both the probability density function and its derivative. One then simulates and estimates sensitivities by the inverse transform method [28]. In this way, the total bias is consisted of three parts: the linear interpolation of order $O(\eta^2)$, the inverse Laplace transform discretization error of the magnitude $O(\exp(-c/h))$ for some positive constant c, and the truncation error incurred by the distribution approximation depending on the decay of the Laplace transform, where η is the grid size in tabulation, and h is the discretization level in approximation. One can re-run the simulation procedure by updating these parameter values until the estimation bias decreases proportionally.

Our method in this thesis combines the method in [28] with an extension from [19]. We tabulate the cdf and the derivative of the pdf with respect to a parameter on discrete grid points. The cdf values are computed through Hilbert transforms of the characteristic function, while the derivative of density is obtained through the inverse Fourier transforms. When the characteristic function is in a certain analytic class, the approximation of the tabulated quantities is highly accurate with exponentially decaying discretization errors, and admits explicit and computable error estimates. Using results from [26], we obtain explicit and computable bounds for the estimation bias in one dimensional problems. This allows us to determine the simulation support, the fineness of the grid, and the numerical parameters for inverting the characteristic function given any bias tolerance level. Random numbers are then generated from the tabulated and linearly interpolated cdf through inverse transform method as in [19]. The derivative of the pdf is then calculated by linear interpolation. In this way, the

bias of the estimator is composed of three parts. The first part, the truncation error, is due to the truncation of the support of the distribution. The second part is caused by linear interpolation, which is quadratic in terms of the step size of the grid. The third part comes from the approximation of the derivative of the density. Consequently, we determine the grid setting, and the approximation parameters in one-dimensional cases. This error analysis can be extended to multi-dimensions.

Section 4.1 introduces the basic idea of likelihood ratio method, how to tabulate from characteristic functions, and how to apply the likelihood ratio method based on the tabulated data. In Section 4.1.4, we derive an explicit upper bound for the estimation bias in one dimensional cases, and explain how to control bias through tuning the fineness of the grid and adjusting the approximation parameter values. We also describe how to control bias in multi-dimensional cases. Section 4.2 implements the likelihood ratio method in estimating European and Asian option deltas under the CGMY model. We verify our theoretical results through numerical experiments.

In many financial models, close-form probability density functions (pdfs) or cumulative distribution functions (cdfs) are not available, while characteristic functions are. Thus, from simulation perspective, the inverse transform method, which needs explicit expression for the cdf, is not applicable directly. One way to simulate Lévy processes is by inverting characteristic functions [20]. Extensions include estimating option sensitivities [28]. When the characteristic function is known, the Gil-Pelaez Formula is widely used in computing the cdf [6, 7], and inverse Fourier transform is a general way in calculating the pdf. But according to [14, 15, 18], for extreme inputs, there arise large pricing errors when the inverse Fourier transform is discretized by Simpson's rule. [26] also applies inverse Fourier transform in pdf calculation and expresses a cdf by means of a Hilbert transform, but it uses the trapezoidal rule in the approximation, which is convenient, accurate and stable. This methodology can then be extended to

multivariate cases.

In this thesis, we focus on inverting analytic characteristic functions for bivariate models to estimate bivariate cdfs and some expectations involving indicator functions. With an explicit characteristic function, to estimate the corresponding bivariate cdf, we represent it by a double integral and discretize it to an infinite double series. We then truncate the infinite series representation. In this way, an explicit error estimate as a function of the discretization stepp size and truncation level is especially important in bounding the total estimation error. Dominating terms in the approximation error bound are decaying exponentially in $1/h_1$, $1/h_2$, $c_1(M_1h_1)^{\nu_1}$, and $c_2(M_2h_2)^{\nu_2}$, where h_1, h_2 are the discretization step sizes for the first and the second dimension respectively, and M_1h_1 and M_2h_2 are the truncation levels. We present ways to choose M_2 based on M_1, and select h_1, h_2 as functions of M_1, M_2, such that all the dominating terms are decaying exponentially at the same rate. Therefore, one just needs to select M_1 to bound the total estimation error to a target level, instead of selecting the four parameters separately.

For many derivative securities, Monte Carlo methods are preferred. One simulation method that is widely used is the inverse transform method [21]. To apply the inverse transform method, one needs to generate a uniform random variable, $U \sim U[0,1]$ and find the corresponding x such that $U = F(x)$. By tabulation, the inverse transform method is quite efficient and convenient in univariate cases. By following this methodology, in simulating multivariate random processes, one needs to generate X_1 from $F(X_1)$, then X_2 from $F(X_2|X_1)$, X_3 from $F(X_3|X_1, X_2)$, and so on. Every time, the simulation grid varies because different values are generated from the previous dimension. Therefore, the inverse transform method is hard to be generalized to higher dimensions if close-form cdfs or conditional cdfs are not accessible. However, when the characteristic function lies in a specific analyticity class such that it is analytic in a

horizontal strip in every dimension with all other dimensions fixed, the underlying conditional pdfs admit exponential tail behavior. Covering the underlying pdf with a scaled multi-variate exponential distribution, instead of simulating from a multi-variate distribution without closed-form pdf, one can generate samples from an easy-to-simulate distribution and apply multi-variate acceptance-rejection method. With careful selection of the parameters in the multi-variate exponential distribution, the acceptance rate can be optimized. In Chapter 5, we apply acceptance-rejection method on univarite and bivariate models with certain analytic characteristic functions. The classical algorithm proposed by [22] is compared with our improvements in terms of running time (acceptance rate) and estimation error in Section 5.1.3.

In addition to traditional financial engineering application, our analysis on models with certain analytic characteristic functions also applies to other fields like electrical engineering.

A defining characteristic of the wireless channel is the variations of the channel strength over time, known as channel fading. The channel fading amplitude at a particular time can be probabilistically modeled as a random variable with Rayleigh, Nakagami-m or Nakagami-n(Rice) distribution. Receiver diversity reception is an effective technology to combat channel fading. The performance of a diversity reception scheme is evaluated by taking average of the conditional error probability (CEP) over the fading amplitudes, which is the average symbol error rate (SER). Equal gain combining (EGC) is one commonly used diversity reception scheme due to its comparable performance to optimal maximal-ratio combining (MRC) but with much greater simplicity and more economical hardware cost [4]. However, research on the performance analysis for EGC receivers is relatively fewer compared to those for other diversity combining schemes, such as MRC.

For the MRC scheme, reseachers proposed an unified approach based on moment

generation function (MGF) technique that can be used for a broad class of modulations under different channel fading [43,44]. However, the MGF approach doesn't work as well for EGC receivers due to the cross-product terms in its output signal-to-noise ration (SNR) . Instead, most of the existing studies on performance analysis for EGC receiver are focusing on a particular distributed channel fading. In [4,10,11], authors analyzed the performance for EGC receivers in Rayleigh, Nakagami-m and Nakagami-n(Rice) channels, respectively, by using Fourier series techniques to approximate the probability density function (pdf) for a sum of independent random variables. The author in [45] provided some closed-form solutions for binary error rate of EGC receivers in Rayleigh fading. Authors in [3,42] studied EGC performance in a Nakagami-m channel. An approximation for SER for EGC receivers in Rice and Hoyt fading channels are proposed in [30,46], by finding an approximation for the MGF of the output SNR. Authors in [4] obtained a rapidly converging series representation for the EGC performance in Nakagami channels.

In Chapter 6, we propose a unified SER closed-form computation approach for EGC receivers under any channel fading with an explicit characteristic function and for any modulations whose CEP can be writen as

$$P_s(\epsilon|\gamma) = a \cdot \mathrm{erfc}(\sqrt{p}\gamma) + b \cdot \mathrm{erfc}^2(\sqrt{p}\gamma), \tag{1.1}$$

by employing a transform-based and exponentially error-decay approximation of $\mathrm{erfc}(\cdot)$, which is a summation of exponential terms and turns out to be the characteristic functions of channel fadings after multiplying by channel fadings' pdf and then taking integral. The approximation error for the proposed method is proved to be bounded and the upper bound is exponentially decaying as a the number of terms considered increase.

There are several existing works studying general methods for the performance analysis of EGC receivers. Authors in [2] derived analytical expressions for the EGC receiver performance in general fading channels in terms of the Appell hypergeometric function, which involves evaluating multiple infinite series. In [2,5], authors proposed a characteristic functions method for calculating SER of a broad class of channel fadings for EGC explicit reciever based on Parsevals theorem, with the requirement of explicit characteristic functions for both the CEP and channel fadings. In this work, the obtained closed-form approximation for SER only involves a finite series calculation, and the approximation error is proved to be bounded and exponentially decay with the number of terms considered in the finite series. The only requirement in our method is the knowledge of the characteristic function for the channel fading distribution, which is very easily obtained, as provided in [5]

Chapter 2

Univariate Analytic Characteristic Functions and Their Inversion

Let $\mathbb{R}$ denote the set of all real numbers, and $\mathbb{C}$ the set of all complex numbers.

Definition 2.1. *The characteristic function of a random variable X with cumulative distribution function (cdf) F is defined to be*

$$\phi(\xi) = \mathbb{E}[e^{i\xi X}] = \int_{-\infty}^{\infty} e^{i\xi x}\, dF(x)$$

for all such $\xi \in \mathbb{C}$ that the above is finite.

We extend the classical definition of the characteristic function to include complex values for ξ since we are interested in its analyticity on the complex plane. The characteristic function is well defined for all $\xi \in \mathbb{R}$. Whether it is well defined for non-real ξ depends on the underlying distribution F. When $\phi(\xi)$ is well defined for complex ξ, we study the relations between the analyticity of ϕ and the tail behavior of $F(x)$.

2.1 Analyticity of Univariate Characteristic Functions

We, first of all, denote the real part of a complex number ξ by $\Re(\xi)$ and the imaginary part by $\Im(\xi)$ and show that the analyticity of the characteristic function in a horizontal strip in the complex plane is equivalent to the existence of exponential moments.

Proposition 2.2. *The characteristic function of a random variable X is analytic in a strip $\mathcal{D}_{(d_-,d_+)} := \{\xi \in \mathbb{C} : d_- < \Im(\xi) < d_+\}$, $-\infty < d_- < 0 < d_+ < +\infty$, if and only if $\mathbb{E}[e^{-bX}]$ is finite for any $b \in (d_-, d_+)$.*

Proof. For the necessity, suppose that ϕ is analytic in the strip $\mathcal{D}_{(d_-,d_+)}$. Then, for any $\xi = a + ib$ with $a \in \mathbb{R}$ and $d_- < b < d_+$, $\phi(\xi)$ is finite. In particular,

$$\mathbb{E}[e^{-bX}] = \phi(ib) < +\infty.$$

For the sufficiency, suppose that $\mathbb{E}[e^{-bX}]$ is finite for any $d_- < b < d_+$. Then for any $\xi = a + ib$ with $a \in \mathbb{R}$ and $d_- < b < d_+$, $\phi(\xi)$ is finite because

$$|\phi(\xi)| = |\mathbb{E}[e^{i(a+ib)X}]| \leq \mathbb{E}[e^{-bX}] < +\infty.$$

Moreover, for any $d_- < b < d_+$, $\mathbb{E}[|X|e^{-bX}] < +\infty$. To see this, let $\epsilon > 0$ be such that $d_- < b \pm \epsilon < d_+$. Then

$$\begin{aligned} \mathbb{E}[|X|e^{-bX}] &= \mathbb{E}[Xe^{-\epsilon X}e^{-(b-\epsilon)X}1_{\{X\geq 0\}}] + \mathbb{E}[-Xe^{-\epsilon(-X)}e^{-(b+\epsilon)X}1_{\{X<0\}}] \\ &\leq \frac{1}{\epsilon}(\mathbb{E}[e^{-(b-\epsilon)X}1_{\{X\geq 0\}}] + \mathbb{E}[e^{-(b+\epsilon)X}1_{\{X<0\}}]) \\ &\leq \frac{1}{\epsilon}(\mathbb{E}[e^{-(b-\epsilon)X}] + \mathbb{E}[e^{-(b+\epsilon)X}]) \ < \ +\infty. \end{aligned} \tag{2.1}$$

To show the analyticity of ϕ in $\mathcal{D}_{(d_-,d_+)}$, it suffices to show that ϕ is differentiable

at any point $\xi_0 = a_0 + ib_0$ with $d_- < b_0 < d_+$. Let $\delta > 0$ be such that $B_\delta(\xi_0) := \{\xi \in \mathbb{C} : 0 < |\xi - \xi_0| < \delta\} \subset \mathcal{D}_{(d_-,d_+)}$. For such a δ, $d_- < b_0 - \delta < b_0 + \delta < d_+$. Note that

$$\lim_{\xi\to\xi_0} \frac{\phi(\xi) - \phi(\xi_0)}{\xi - \xi_0} = \lim_{\xi\to\xi_0} \mathbb{E}\Big[\frac{e^{i\xi X} - e^{i\xi_0 X}}{\xi - \xi_0}\Big].$$

Let $\xi \in B_\delta(\xi_0)$. For any $x \in \mathbb{R}$, by the complex mean value theorem [25], there exist ω_1, ω_2 that are on the line connecting ξ and ξ_0 such that

$$\Re\Big(\frac{e^{i\xi x} - e^{i\xi_0 x}}{\xi - \xi_0}\Big) = \Re(ixe^{i\omega_1 x}),\ \Im\Big(\frac{e^{i\xi x} - e^{i\xi_0 x}}{\xi - \xi_0}\Big) = \Im(ixe^{i\omega_2 x}).$$

Note that $b_0 - \delta < \Im(\omega_1), \Im(\omega_2) < b_0 + \delta$. We thus have

$$\begin{aligned}
\Big|\frac{e^{i\xi x} - e^{i\xi_0 x}}{\xi - \xi_0}\Big| &\leq |x|e^{-\Im(\omega_1)x} + |x|e^{-\Im(\omega_2)x} \\
&\leq 2|x|(e^{-x(b_0-\delta)}1_{\{x\geq 0\}} + e^{-x(b_0+\delta)}1_{\{x<0\}}) \\
&\leq 2|x|(e^{-x(b_0-\delta)} + e^{-x(b_0+\delta)}). \qquad (2.2)
\end{aligned}$$

Since x is arbitrary and the expression in (2.2) doesn't depend on ω_1, ω_2, we have that

$$\Big|\frac{e^{i\xi X} - e^{i\xi_0 X}}{\xi - \xi_0}\Big| \leq 2|X|(e^{-X(b_0-\delta)} + e^{-X(b_0+\delta)})$$

for any $\xi \in B_\delta(\xi_0)$. By (2.1), $2|X|(e^{-X(b_0-\delta)} + e^{-X(b_0+\delta)})$ has a finite expectation. We thus have the following by the dominated convergence theorem:

$$\lim_{\xi\to\xi_0} \frac{\phi(\xi) - \phi(\xi_0)}{\xi - \xi_0} = \lim_{\xi\to\xi_0} \mathbb{E}\Big[\frac{e^{i\xi X} - e^{i\xi_0 X}}{\xi - \xi_0}\Big] = \mathbb{E}\Big[\lim_{\xi\to\xi_0} \frac{e^{i\xi X} - e^{i\xi_0 X}}{\xi - \xi_0}\Big] = \mathbb{E}[iXe^{i\xi_0 X}],$$

where the last expectation is well defined with $|\mathbb{E}[iXe^{i\xi_0 X}]| \leq \mathbb{E}[|X|e^{-b_0 X}] < +\infty$. ϕ is thus differentiable at ξ_0. This finishes the proof.

□

2.2 Analyticity of Characteristic Functions and Tail Behaviors of PDFs

Characteristic function of a distribution can be treated as the Fourier transform of the underlying probability density function (pdf). Characteristic function with certain analyticity indicates how the corresponding pdf behaves at extreme values. The following theorem discusses the relations between analyticity of a function's Fourier transform and tail behavior of this function itself.

Theorem 2.3. *Let $\xi = a + ib$, if*

- *$\hat{f}(\xi)$ is analytic for $d_- < b < d_+$,*
- *$\hat{f}(\cdot + ib) \in L^1(\mathbb{R})$ in the analyticity strip,*
- *$\lim_{a\to\pm\infty} \int_{d_-}^{d_+} |\hat{f}(a+ib)|\, db = 0$,*

then

$$f(x) = \begin{cases} O(e^{(d_- + \epsilon)x}), & x \to \infty \\ O(e^{(d_+ - \epsilon)x}), & x \to -\infty \end{cases}.$$

Proof. Function f can be obtained from $\hat{f}$ by inverse Fourier transform:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(a)e^{-iax}\, da = \lim_{R\to+\infty} \frac{1}{2\pi}\int_{-R}^{+R} \hat{f}(a)e^{-iax}\, da.$$

Consider the following contour integral:

$$\int_{\gamma} \hat{f}(\xi)e^{-i\xi x}\, d\xi = \int_{\gamma_1+\gamma_2+\gamma_3+\gamma_4} \hat{f}(\xi)e^{-i\xi x}\, d\xi,$$

where γ is a closed contour ($d_- < t < d_+$) consisting of:

$$\begin{aligned}
\gamma_1 &= \{a : a \text{ goes from } -R \text{ to } R\}, \\
\gamma_2 &= \{R + ib : b \text{ goes from } 0 \text{ to } t\}, \\
\gamma_3 &= \{a + it : a \text{ goes from } R \text{ to } -R\}, \\
\gamma_4 &= \{-R + ib : b \text{ goes from } t \text{ to } 0\}.
\end{aligned}$$

According to Cauchy's theorem, when $\hat{f}(\xi)e^{-i\xi x}$ is analytic on and inside the closed contour γ, $\int_\gamma \hat{f}(\xi)e^{-i\xi x}\, d\xi = 0$.

Combined with

$$\begin{aligned}
|\int_{\gamma_2} \hat{f}(\xi)e^{-i\xi x}\, d\xi| &= |i\int_0^t \hat{f}(R+ib)e^{-i(R+ib)x}\, db| \le e^{|tx|}\int_0^t |\hat{f}(R+ib)|\, db \\
&\le e^{|tx|}\int_{d_-}^{d_+} |\hat{f}(R+ib)|\, db \to 0 \text{ as } R \to \infty,
\end{aligned}$$

and

$$\begin{aligned}
|\int_{\gamma_4} \hat{f}(\xi)e^{-i\xi x}\, d\xi| &= |i\int_t^0 \hat{f}(-R+ib)e^{-i(-R+ib)x}\, db| \le e^{|tx|}\int_0^t |\hat{f}(-R+ib)|\, db \\
&\le e^{|tx|}\int_{d_-}^{d_+} |\hat{f}(-R+ib)|\, db \to 0 \text{ as } R \to \infty,
\end{aligned}$$

then

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(a)e^{-iax}\, da = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(a+it)e^{-i(a+it)x}\, da, \text{ for } d_- < t < d_+.$$

Since $\hat{f}(\cdot + ib) \in L^1(\mathbb{R})$ for $d_- < b < d_+$,

$$|f(x)| = \frac{1}{2\pi}|\int_{-\infty}^{+\infty} \hat{f}(a+it)e^{-i(a+it)x}\,da| \leq \frac{1}{2\pi}e^{tx}\int_{-\infty}^{+\infty} |\hat{f}(a+it)|\,da = O(e^{tx}).$$

When $x \to +\infty$, let t be arbitrarily close to d_-, then $f(x) = O(e^{(d_-+\epsilon)x})$.

When $x \to -\infty$, let t be arbitrarily close to d_+, then $f(x) = O(e^{(d_+-\epsilon)x})$.

□

According to Theorem 2.3, exponential tail behavior of f can be obtained by certain analyticity of its Fourier transform in a horizontal strip. Specifically, exponential decaying rate of f on the right-hand-side is decided by the lower bound of this analyticity strip, while the decaying rate on the left-hand-side corresponds to the upper bound of the strip. Moreover, certain analyticity together with exponential tail behavior of a function and similar properties of the function's Fourier transform indicates each other, as included in the following theorem.

Theorem 2.4. *(Theorem 26 in [24]) Let $c_+, d_+ > 0$, $c_-, d_- < 0$. Then $\hat{f}(\xi)$ is analytic in the strip $\{\xi = a + ib \in \mathbb{C} : d_- < b < d_+\}$, and*

$$\hat{f}(\xi) = \begin{cases} O(e^{-(c_+-\epsilon)a}), & a \to +\infty \\ O(e^{-(c_-+\epsilon)a}), & a \to -\infty \end{cases}$$

for every positive ϵ, if and only if $f(z)$ is analytic in the strip $\{z = x + iy \in \mathbb{C} : c_- < y < c_+\}$, and

$$f(z) = \begin{cases} O(e^{(d_-+\epsilon)x}), & x \to +\infty \\ O(e^{(d_+-\epsilon)x}), & x \to -\infty \end{cases}$$

for every positive ϵ.

Proof. Here we only prove the necessary condition, while the proof of the sufficient condition follows the same methodology.

First of all, we prove the analyticity of $f(z)$ in the strip by proving its uniform convergence in the same strip.

Let $f_\theta(z) = \frac{1}{2\pi}\int_{-\theta}^{\theta} \hat{f}(a)e^{-iaz}\,da$, where $\theta > 0$, then

$$\begin{aligned} |f_\theta(z) - f(z)| &= \frac{1}{2\pi}\Big|\int_{-\infty}^{-\theta} \hat{f}(a)e^{-iaz}\,da + \int_{\theta}^{+\infty} \hat{f}(a)e^{-iaz}\,da\Big| \\ &\leq \frac{1}{2\pi}\Big(\int_{-\infty}^{-\theta} |\hat{f}(a)|e^{ay}\,da + \int_{\theta}^{+\infty} |\hat{f}(a)|e^{ay}\,da\Big). \end{aligned}$$

Choose $\epsilon > 0$, such that $c_- < y - \epsilon < y + \epsilon < c_+$. By the exponential tail behavior of $\hat{f}(\xi)$, there exist a_1, a_2, C_1, and $C_2 > 0$, such that

$$\begin{cases} |\hat{f}(\xi)| \leq C_1 e^{-(c_+ - \epsilon)a}, & a \geq a_1 \\ |\hat{f}(\xi)| \leq C_2 e^{-(c_- + \epsilon)a}, & a \leq -a_2 \end{cases}.$$

If $\theta \geq \max(a_1, a_2)$,

$$\begin{aligned} |f_\theta(z) - f(z)| &\leq \frac{1}{2\pi}\Big(\int_{-\infty}^{-\theta} C_2 e^{(-c_- - \epsilon + y)a}\,da + \int_{\theta}^{+\infty} C_1 e^{(-c_+ + \epsilon + y)a}\,da\Big) \\ &= \frac{1}{2\pi}\Big(\frac{C_2}{-c_- - \epsilon + y}e^{-\theta(-c_- - \epsilon + y)} + \frac{C_1}{c_+ - \epsilon - y}e^{-\theta(c_+ - \epsilon - y)}\Big) \\ &\leq \frac{1}{2\pi s}(C_1 + C_2)e^{-\theta s}, \end{aligned}$$

where $s = \min(-c_- - \epsilon + y, c_+ - \epsilon - y) > 0$.

Therefore, $\forall \delta > 0$, when $\theta > \max(\frac{1}{s}\ln(\frac{C_1 + C_2}{2\pi\delta s}), a_1, a_2)$, $|f_\theta(z) - f(z)| < \delta$. Then $f(z)$ is uniformly convergent, hence analytic, in the strip $\{z = x + iy \in \mathbb{C} : c_- < y < c_+\}$.

Now we move onto the exponential tail behavior of $|f(z)|$. By inverse Fourier transform, $f(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(a)e^{-iaz}\,da$. Consider the following contour integral:

$$\int_{\gamma} \hat{f}(\xi)e^{-i\xi z}\,dz = \int_{\gamma_1+\gamma_2+\gamma_3+\gamma_4} \hat{f}(z)e^{-i\xi z}\,dz,$$

where γ is defined in Theorem 2.3.

By its analyticity in the strip, $\int_{\gamma} \hat{f}(z)e^{-i\xi z}\,dz = 0$. And it can be easily proved that $\int_{\gamma_2} \hat{f}(z)e^{-i\xi z}\,dz \to 0(R \to \infty)$ and $\int_{\gamma_4} \hat{f}(z)e^{-i\xi z}\,dz \to 0(R \to \infty)$. Then

$$f(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(a+ib)e^{-i(x+iy)(a+ib)}\,da = \frac{e^{bx+iby}}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(a+ib)e^{(ay-iax)}\,da,$$

and

$$|f(z)| \leq \frac{e^{bx}}{2\pi}\int_{-\infty}^{+\infty} |\hat{f}(a+ib)|e^{ay}\,da,$$

Taking b arbitrarily close to d_- or d_+,

$$f(z) = \begin{cases} O(e^{(d_-+\epsilon)x}), & x \to \infty \\ O(e^{(d_+-\epsilon)x}), & x \to -\infty. \end{cases}$$

This finishes the proof. □

2.3 Analyticity of Characteristic Functions and Tail Behaviors of CDFs

By investigating analyticity of univariate characteristic function, we can obtain certain exponential tail behavior of the underlying pdf. Moreover, the analyticity of the characteristic function also indicates certain tail behavior of the corresponding cdf, as in

the following theorem.

Theorem 2.5. *The characteristic function $\phi(\xi)$ of a distribution function $F(x)$ is analytic in the strip $\{\xi = a + ib \in \mathbb{C} : d_- < b < d_+\}$, if and only if*

$$1 - F(x) = O(e^{-rx})$$

as $x \to \infty$, for $0 < r < -d_-$, and

$$F(-x) = O(e^{-rx})$$

as $x \to \infty$, for $0 < r < d_+$.

Proof. We first prove the sufficient condition.

Choose $r, b \in \mathbb{R}$, such that $0 < b < r < d_+$. By $F(-x) = O(e^{-rx})$ as $x \to \infty$, for $0 < r < d_+$, there exists a constant $c_1 > 0$ and a positive integer K_1, such that $F(-x) \leq c_1 e^{-rx}$ when $x \geq K_1$. Then

$$\int_{-\infty}^{-K_1} e^{-bx}\, dF(x) = \sum_{k=K_1}^{\infty} \int_{-k-1}^{-k} e^{-bx}\, dF(x) \leq \sum_{k=K_1}^{\infty} c_1 e^{(b-r)k+b} = \frac{c_1 e^{(b-r)K_1+b}}{1 - e^{b-r}} < +\infty.$$

Combined with the fact that

$$\int_{-K_1}^{+\infty} e^{-bx}\, dF(x) \leq e^{bK_1}(1 - F(-K_1)) < +\infty,$$

$$E[e^{-bX}] = \Big(\int_{-\infty}^{-K_1} + \int_{-K_1}^{+\infty}\Big) e^{-bx}\, dF(x) < +\infty, \text{ for } 0 < b < d_+.$$

Choose $r, b \in \mathbb{R}$, such that $d_- < -r < b < 0$. By $1 - F(x) = O(e^{-rx})$ as $x \to \infty$, for $0 < r < -d_-$, there exists a constant $c_2 > 0$ and a positive integer K_2, such that

$1 - F(x) \leq c_2 e^{-rx}$ when $x \geq K_2$. Then

$$\int_{K_2}^{+\infty} e^{-bx}\, dF(x) = \sum_{k=K_2}^{\infty} \int_{k}^{k+1} e^{-bx}\, dF(x) \leq \sum_{k=K_2}^{\infty} c_2 e^{-(b+r)k} = \frac{c_2 e^{-(b+r)K_2}}{1 - e^{-(b+r)}} < +\infty.$$

And

$$\int_{-\infty}^{K_2} e^{-bx}\, dF(x) \leq e^{-bK_2} F(K_2) \leq e^{-bK_2} < +\infty.$$

Hence,

$$E[e^{-bX}] = \Big(\int_{-\infty}^{K_2} + \int_{K_2}^{+\infty} \Big) e^{-bx}\, dF(x) < +\infty, \text{ for } d_- < b < 0.$$

It is obvious that $E[e^{-bX}] = 1 < +\infty$ for $b = 0$. Consequently, $E[e^{-bX}] < +\infty$ for $d_- < b < d_+$. By Proposition 2.2, $\phi(\xi)$ is analytic in the strip $\{\xi = a + ib \in \mathbb{C} : d_- < b < d_+\}$.

Now we move onto the proof for the necessary condition.

As $\phi(\xi)$ is analytic in the strip $\{\xi = a + ib \in \mathbb{C} : d_- < b < d_+\}$, we have $E[e^{-bX}] < +\infty$ for $d_- < b < d_+$, which gives that:

$$\int_{-\infty}^{-\tilde{x}} e^{-bx}\, dF(x) < c_3 \text{ for some } c_3 > 0 \text{ when } 0 < b < d_+,$$

$$\int_{\tilde{x}}^{+\infty} e^{bx}\, dF(x) < c_4 \text{ for some } c_4 > 0 \text{ when } 0 < b < -d_-,$$

for $\tilde{x} > 0$.

Therefore, $e^{b\tilde{x}} F(-\tilde{x}) \leq c_3 (0 < b < d_+)$ and $e^{b\tilde{x}}(1 - F(\tilde{x})) \leq c_4 (0 < b < -d_-)$.

This finishes the proof.

□

By change of measure, the European option price can be expressed in terms of cdfs of the underlying return process under different measures. The behavior of option price with extreme strikes can be demonstrated by analyticity of the characteristic function.

Define S_t as the price of an underlying asset at time t, T as the maturity of the options, and $C(K)$ and $P(K)$ as European call and put prices as functions of the strike price K. The following corollary studies the tail behaviors of the option prices when K takes extreme values.

Corollary 2.6. *(Corollary 2.2 in [34]) If $\mathbb{E}[S_T^{p+1}] < \infty$, then $C(K) = O(K^{-p})$ as $K \to \infty$. If $\mathbb{E}[S_T^{-q}] < \infty$, then $P(K) = O(K^{1+q})$ as $K \to 0$.*

Proof. $S_T = S_0 e^{X_T}$, where $X_T = \ln(S_T/S_0)$. Let $\hat{f}(z)$ be the characteristic function of X_T. By Proposition 2.2, $\mathbb{E}[S_T^{p+1}] < \infty$ and $\mathbb{E}[S_T^{-q}] < \infty$ imply that $\hat{f}(\xi)$ is analytic in the strip $-(p+1) \leq \Im(\xi) \leq q$.

The option prices can be expressed by linear operations of cdfs under different measures:

$$\begin{aligned} C(K) &= e^{-rT}\mathbb{E}[\max(S_T - K, 0)] \\ &= S_0 e^{-qT}\mathbb{P}^*(X_T \geq \ln(K/S_0)) - Ke^{-rT}\mathbb{P}(X_T \geq \ln(K/S_0)) \\ &= S_0 e^{-qT}(1 - F^*(\ln(K/S_0))) - Ke^{-rT}(1 - F(\ln(K/S_0))), \end{aligned}$$

and

$$P(K) = Ke^{-rT}F(\ln(K/S_0)) - S_0 e^{-qT}F^*(\ln(K/S_0)).$$

$\hat{f}^*(\xi) = \hat{f}(\xi - i)/\hat{f}(-i)$ is the characteristic function of X_T under measure $\mathbb{P}^*$. By

Theorem 2.5, $1 - F(x) = O(e^{-(p+1)x})$, $1 - F^*(x) = O(e^{-px})$, $F(-x) = O(e^{-qx})$, and $F^*(-x) = O(e^{-(q+1)x})$, as $x \to \infty$. Substituting x with $\ln(K/S_0)$ in the first two equations and with $-\ln(K/S_0)$ in the last two, we reach exactly the same conclusion.

□

2.4 Inverting Univariate Analytic Characteristic Functions

In the following, we only discuss models with characteristic function defined in the following group and with exponential tails, $|\phi(a+ib)| \leq \kappa \exp(-c|a|^\nu)$, where κ, c, and ν are independent of a:

Definition 2.7. *A function $g(\omega) \in H(\mathcal{D}_{(d_-,d_+)})$ if:*

- *$g(\omega)$ is analytic in $H(\mathcal{D}_{(d_-,d_+)}) = \{\omega \in \mathbb{C} : \Im(\omega) \in (d_-, d_+)\}$, $d_- < 0$, $d_+ > 0$,*
- *$\int_{d_-}^{d_+} |g(u+iv)|\, dv \to 0, u \to \pm\infty$,*
- *$\|g\|^+ = \int_{\mathbb{R}} |g(u+id_+)|\, du < \infty$, $\|g\|^- = \int_{\mathbb{R}} |g(u+id_-)|\, du < \infty$.*

The underlying pdf and cdf then can be approximated by inverting the analytic characteristic function by trapezoidal rule with great accuracy [26].

The characteristic function can be treated as the Fourier transform of the underlying pdf. Alternatively, the pdf can be obtained by inverting the characteristic function. Therefore, the pdf $f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi x} \phi(\xi)\, d\xi$ is approximated by:

$$f_{h,M,a}(x) = \frac{1}{2\pi} \sum_{m=-M}^{M} e^{-ix(mh+ia)} \phi(mh+ia) h, \tag{2.3}$$

for $a \in (d_-, d_+)$, with discretization level h and truncation level Mh.

The cdf can be represented by a Hilbert transform with parameter zero: $F(x) = \frac{1}{2} - \frac{i}{2}\mathcal{H}(e^{-i\xi x}\phi(\xi)(0)$. And it can approximated similarly by truncating and discretizing the integral in the transform:

$$F_{h,M}(x) = \frac{1}{2} + \frac{i}{2}\sum_{m=-M}^{M} e^{-ix(m-\frac{1}{2})h}\frac{\phi((m-\frac{1}{2})h)}{(m-\frac{1}{2})\pi}, h > 0, M \geq 1. \tag{2.4}$$

The total approximation errors of the pdf and cdf are bounded by:

$$\begin{aligned} |f(x) - f_{h,M,a}(x)| &\leq \frac{\kappa e^{ax}}{\pi\nu c^{(\nu+1)/\nu}}\Gamma(\frac{1}{\nu}, c(Mh)^{\nu}) \\ &+ \frac{e^{2\pi d_-/h}\|\phi\|^-}{2\pi(1-e^{2\pi d_-/h})} + \frac{e^{-2\pi d_+/h}\|\phi\|^+}{2\pi(1-e^{-2\pi d_+/h})}, \end{aligned} \tag{2.5}$$

and

$$\begin{aligned} |F(x) - F_{h,M}(x)| &\leq \frac{4\kappa}{\pi\nu}\Gamma(0, c(Mh)^{\nu}) \\ &+ \frac{e^{2\pi d_-/h + xd_-}\|\phi\|^-}{2\pi|d_-|(1-e^{2\pi d_-/h})} + \frac{e^{-2\pi d_+/h + xd_+}\|\phi\|^+}{2\pi d_+(1-e^{-2\pi d_+/h})}. \end{aligned} \tag{2.6}$$

The first terms in both 2.4 and 2.4 define the approximation error from truncation, and the last two terms combined represent the discretization error of the approximation. With constant $\|\phi\|^{\pm}$, the discretization error is decaying exponentially in terms of $1/h$ and the truncation error is decaying exponentially in the truncation level, Mh. By careful selection of M as a function of h, to control the total approximation error, one only needs to adjust the values of the discretization level h.

Since the European option price can be represented by certain expectation in terms of indicator functions, one can price the European option by applying this approximation representation [26].

Chapter 3

Bivariate Characteristic Functions and Their Inversion

In Chapter 2, we discuss the analyticity of univariate characteristic function and how it is applied in the analysis of the underlying pdf and cdf. Such conclusions can be partially extended to bivariate models, which is the goal of this section. In this section, we begin with defining a special class of bivariate complex functions. For such functions, we approximate their integrals by simple trapezoidal rule. We derive and present the approximation error bound. These results lay the foundations for approximating cdf in Section 3.2 and numerical experiments in Section 3.3.

First of all, for $-\infty < d_j^- < 0 < d_j^+ < +\infty \quad (j = 1, 2)$, we define a horizontal strip in the j-th dimension: $\mathcal{D}_{(d_j^-, d_j^+)} = \{z_j = x_j + iy_j | -\infty < x_j < +\infty, d_j^- < y_j < d_j^+\}$. Then we define a special class of bivariate analytic functions:

Definition 3.1. *A bivariate function g is in $H(\mathcal{D}_{((d_1^-, d_1^+)\times(d_2^-, d_2^+))})$ if it is analytic in $\mathcal{D}_{(d_j^-, d_j^+)}, j = 1, 2$, when we fix the other dimension, and satisfies*

$$\int_{d_1^-}^{d_1^+} |g(u_1 + iv_1, z_2)| \, dv_1 \to 0, \ \ as \ u_1 \to \pm\infty, \forall z_2, \tag{3.1}$$

$$\int_{d_2^-}^{d_2^+} |g(z_1, u_2 + iv_2)|\, dv_2 \to 0, \ \ as\ u_2 \to \pm\infty, \forall z_1, \tag{3.2}$$

$$||g||^{+,+} = \int_{-\infty+id_1^+}^{+\infty+id_1^+} \int_{-\infty+id_2^+}^{+\infty+id_2^+} |g(z_1, z_2)|\, dz_2\, dz_1 < +\infty, \tag{3.3}$$

$$||g||^{+,-} = \int_{-\infty+id_1^+}^{+\infty+id_1^+} \int_{-\infty+id_2^-}^{+\infty+id_2^-} |g(z_1, z_2)|\, dz_2\, dz_1 < +\infty, \tag{3.4}$$

$$||g||^{-,+} = \int_{-\infty+id_1^-}^{+\infty+id_1^-} \int_{-\infty+id_2^+}^{+\infty+id_2^+} |g(z_1, z_2)|\, dz_2\, dz_1 < +\infty, \tag{3.5}$$

$$||g||^{-,-} = \int_{-\infty+id_1^-}^{+\infty+id_1^-} \int_{-\infty+id_2^-}^{+\infty+id_2^-} |g(z_1, z_2)|\, dz_2\, dz_1 < +\infty. \tag{3.6}$$

In this chapter, we only discuss $g \in H(\mathcal{D}_{((d_1^-,d_1^+)\times(d_2^-,d_2^+))})$ satisfying the following inequality:

$$|g(u_1 + iv_1, u_2 + iv_2)| \leq \kappa |u_1|^{n_1} |u_2|^{n_2} \exp(-c_1|u_1|^{\nu_1} - c_2|u_2|^{\nu_2}), \tag{3.7}$$

for some $\kappa > 0, c_1, c_2, \nu_1, \nu_2,\ n_1, n_2 > 0$. In this way, with one dimension fixed, $g(u_1 + iv_1, u_2 + iv_2)$ has exponential tail along the other dimension.

3.1 Trapezoidal Rule for Analytic Bivariate Integrals

For $g \in H(\mathcal{D}_{((d_1^-,d_1^+)\times(d_2^-,d_2^+))})$, we are often interested in its integration along $\mathbb{R}^2$. Applying the Cauchy integral theorem, for any $a_1 \in (d_1^-, d_1^+)$ and $a_2 \in (d_2^-, d_2^+)$, the

integration $\int_{\mathbb{R}^2} g(x_1,x_2)\,dx_2\,dx_1$ can be approximated by the trapezoidal rule:

$$T_{h_1,h_2,M_1,M_2}(g,a_1,a_2)=\sum_{m_1=-M_1}^{M_1}\sum_{m_2=-M_2}^{M_2} g(m_1h_1+ia_1,m_2h_2+ia_2)h_1h_2. \tag{3.8}$$

Denote the total approximation error by

$$E^T_{h_1,h_2,M_1,M_2}(g,a_1,a_2)=\int_{\mathbb{R}^2} g(x_1,x_2)\,dx_2\,dx_1-T_{h_1,h_2,M_1,M_2}(g,a_1,a_2).$$

We then have the following results for $E^T_{h_1,h_2,M_1,M_2}(g,a_1,a_2)$:

Theorem 3.2. *Suppose $g\in H(\mathcal{D}_{((d_1^-,d_1^+)\times(d_2^-,d_2^+))})$ and satisfies (3.7). Then for $a_1\in(d_1^-,d_1^+)$ and $a_2\in(d_2^-,d_2^+)$,*

$$\begin{aligned}
&|E^T_{h_1,h_2,M_1,M_2}(g,a_1,a_2)|\\
\le&\frac{e^{-\pi d_a^1/h_1-\pi d_a^2/h_2}}{(1-e^{-\pi d_a^1/h_1})(1-e^{-\pi d_a^2/h_2})}\Big(||g||^{-,-}+||g||^{-,+}+||g||^{+,-}+||g||^{+,+}\Big)\\
&+\frac{e^{-\pi d_a^1/h_1}}{(1-e^{-\pi d_a^1/h_1})}\Big(||g||^{-,+}+||g||^{+,+}\Big)+\frac{e^{-\pi d_a^2/h_2}}{(1-e^{-\pi d_a^2/h_2})}\Big(||g||^{+,-}+||g||^{+,+}\Big)\\
&+4(M_1+1)h_1\big(\frac{n_1}{c_1\nu_1}\big)^{\frac{n_1}{\nu_1}}\exp(-\frac{n_1}{\nu_1})\frac{1}{\nu_2c_2^{(n_2+1)/\nu_2}}\Gamma\big(\frac{n_2+1}{\nu_2},c_2(M_2h_2)^{\nu_2}\big)\\
&+4(M_2+1)h_2\big(\frac{n_2}{c_2\nu_2}\big)^{\frac{n_2}{\nu_2}}\exp(-\frac{n_2}{\nu_2})\frac{1}{\nu_1c_1^{(n_1+1)/\nu_1}}\Gamma\big(\frac{n_1+1}{\nu_1},c_1(M_1h_1)^{\nu_1}\big)\\
&+\frac{4}{\nu_1\nu_2c_1^{(n_1+1)/\nu_1}c_2^{(n_2+1)/\nu_2}}\Gamma\big(\frac{n_1+1}{\nu_1},c_1(M_1h_1)^{\nu_1}\big)\Gamma\big(\frac{n_2+1}{\nu_2},c_2(M_2h_2)^{\nu_2}\big),
\end{aligned}\tag{3.9}$$

where $d_a^1=2\min(d_1^+-a_1,a_1-d_1^-)$ and $d_a^2=2\min(d_2^+-a_2,a_2-d_2^-)$.

Choose $M_1,M_2>0$, and let $h_j=h_j(M_j)=(\pi d_a^j/c_j)^{\frac{1}{1+\nu_j}}M_j^{-\frac{\nu_j}{1+\nu_j}}$, $j=1,2$, such that $M_1h_1(M_1)\ge(n_1/(c_1\nu_1))^{1/\nu_1}$, and $M_2h_2(M_2)\ge(n_2/(c_2\nu_2))^{1/\nu_2}$, then there exist constant $C_1,C_2,C_3>0$ independent of M_1 and M_2, such that the total error is bounded

by:

$$
\begin{aligned}
&|E^T_{h_1(M_1),h_2(M_2),M_1,M_2}(g,a_1,a_2)| \\
&\le C_1 M_2 M_1^{\frac{n_1+1-\nu_1}{1+\nu_1}} \exp(-c_1^{-\frac{1}{1+\nu_1}} \left(\pi d_a^1 M_1\right)^{\frac{\nu_1}{1+\nu_1}}) + C_2 M_1 M_2^{\frac{n_2+1-\nu_2}{1+\nu_2}} \exp(-c_2^{-\frac{1}{1+\nu_2}} \left(\pi d_a^2 M_2\right)^{\frac{\nu_2}{1+\nu_2}}) \\
&+ C_3 M_1^{\frac{n_1+1-\nu_1}{1+\nu_1}} M_2^{\frac{n_2+1-\nu_2}{1+\nu_2}} \exp(-c_1^{-\frac{1}{1+\nu_1}} \left(\pi d_a^1 M_1\right)^{\frac{\nu_1}{1+\nu_1}} - c_2^{-\frac{1}{1+\nu_2}} \left(\pi d_a^2 M_2\right)^{\frac{\nu_2}{1+\nu_2}}). \qquad (3.10)
\end{aligned}
$$

Proof. First of all, we investigate the discretization error $E^T_{h_1,h_2,\infty,\infty}(g,a_1,a_2)$. Let $N_1, N_2 > 0$ be integers. Construct the following contour:

$$
\begin{aligned}
\gamma_1^1 &= \{u_1 + id_1^-, u_1 \text{ goes from } -(N_1+\frac{1}{2})h_1 \text{ to } (N_1+\frac{1}{2})h_1\}, \\
\gamma_2^1 &= \{(N_1+\frac{1}{2})h_1 + iv_1, v_1 \text{ goes from } d_1^- \text{ to } d_1^+\}, \\
\gamma_3^1 &= \{u_1 + id_1^+, u_1 \text{ goes from } (N_1+\frac{1}{2})h_1 \text{ to } -(N_1+\frac{1}{2})h_1\}, \\
\gamma_4^1 &= \{-(N_1+\frac{1}{2})h_1 + iv_1, v_1 \text{ goes from } d_1^+ \text{ to } d_1^-\}, \\
\gamma_1^2 &= \{u_2 + id_2^-, u_2 \text{ goes from } -(N_2+\frac{1}{2})h_2 \text{ to } (N_2+\frac{1}{2})h_2\}, \\
\gamma_2^2 &= \{(N_2+\frac{1}{2})h_2 + iv_2, v_2 \text{ goes from } d_2^- \text{ to } d_2^+\}, \\
\gamma_3^2 &= \{u_2 + id_2^+, u_2 \text{ goes from } (N_2+\frac{1}{2})h_2 \text{ to } -(N_2+\frac{1}{2})h_2\}, \\
\gamma_4^2 &= \{-(N_2+\frac{1}{2})h_2 + iv_2, v_2 \text{ goes from } d_2^+ \text{ to } d_2^-\}, \\
\gamma^1 &= \gamma_1^1 \cup \gamma_2^1 \cup \gamma_3^1 \cup \gamma_4^1, \\
\gamma^2 &= \gamma_1^2 \cup \gamma_2^2 \cup \gamma_3^2 \cup \gamma_4^2.
\end{aligned}
$$

And consider the following contour integral:

$$\begin{aligned}
&\int_{\gamma^1\times\gamma^2} \frac{g(z_1,z_2)}{(e^{2\pi i(z_1-ia_1)/h_1}-1)(e^{2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1\\
=&\int_{\gamma_1^1+\gamma_2^1+\gamma_3^1+\gamma_4^1}\int_{\gamma_1^2+\gamma_2^2+\gamma_3^2+\gamma_4^2} \frac{g(z_1,z_2)}{(e^{2\pi i(z_1-ia_1)/h_1}-1)(e^{2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1.
\end{aligned}$$

Fix $z_1 \in \gamma^1$, the integrand has poles of order 1 at $m_2h_2 + ia_2$ for $m_2 \in \mathbb{Z}$. The residue for each such point is given by

$$\begin{aligned}
&\lim_{z_2\to m_2h_2+ia_2}(z_2-m_2h_2-ia_2)\frac{g(z_1,z_2)}{(e^{2\pi i(z_1-ia_1)/h_1}-1)(e^{2\pi i(z_2-ia_2)/h_2}-1)}\\
=&\frac{1}{2\pi i}\frac{g(z_1,m_2h_2+ia_2)h_2}{(e^{2\pi i(z_1-ia_1)/h_1}-1)}.
\end{aligned}$$

Extending to the first dimension, we then have the following expression:

$$\begin{aligned}
&\int_{\gamma^1\times\gamma^2} \frac{g(z_1,z_2)}{(e^{2\pi i(z_1-ia_1)/h_1}-1)(e^{2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1\\
=&\int_{\gamma^1}\frac{1}{e^{2\pi i(z_1-ia_1)/h_1}-1}\Big(\sum_{m_2=-N_2}^{N_2} g(z_1,m_2h_2+ia_2)h_2\Big)\,dz_1\\
=&\sum_{m_1=-N_1}^{N_1}\sum_{m_2=-N_2}^{N_2} g(m_1h_1+ia_1,m_2h_2+ia_2)h_1h_2, \qquad (3.11)
\end{aligned}$$

By Assumption (3.1) and (3.2), and taking N_1 and N_2 to $+\infty$, we have:

$$\begin{aligned}
&\Big(\int_{-\infty+id_1^-}^{+\infty+id_1^-}\int_{-\infty+id_2^-}^{+\infty+id_2^-} - \int_{-\infty+id_1^-}^{+\infty+id_1^-}\int_{-\infty+id_2^+}^{+\infty+id_2^+} \\
&- \int_{-\infty+id_1^+}^{+\infty+id_1^+}\int_{-\infty+id_2^-}^{+\infty+id_2^-} + \int_{-\infty+id_1^+}^{+\infty+id_1^+}\int_{-\infty+id_2^+}^{+\infty+id_2^+}\Big) \\
&\frac{g(z_1,z_2)}{(e^{2\pi i(z_1-ia_1)/h_1}-1)(e^{2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1 \\
&= \sum_{m_1=-\infty}^{+\infty}\sum_{m_2=-\infty}^{+\infty} g(m_1h_1+ia_1, m_2h_2+ia_2)h_1h_2. \qquad (3.12)
\end{aligned}$$

Due to the fact that

$$\begin{aligned}
&-\int_{-\infty+id_2^+}^{+\infty+id_2^+} \tfrac{g(z_1,z_2)}{e^{2\pi i(z_2-ia_2)/h_2}-1}\,dz_2 \\
= \;& \int_{-\infty+id_2^+}^{+\infty+id_2^+} g(z_1,z_2)\,dz_2 + \int_{-\infty+id_2^+}^{+\infty+id_2^+} \tfrac{g(z_1,z_2)}{e^{-2\pi i(z_2-ia_2)/h_2}-1}\,dz_2,
\end{aligned}$$

and by applying the results similarly to the first dimension, we thus obtain the discretization error:

$$
\begin{aligned}
&E^T_{h_1,h_2,\infty,\infty}(g,a_1,a_2)(x_1,x_2)\\
=&-\Big(\int_{-\infty+id_1^-}^{+\infty+id_1^-}\int_{-\infty+id_2^-}^{+\infty+id_2^-}\frac{g(z_1,z_2)}{(e^{2\pi i(z_1-ia_1)/h_1}-1)(e^{2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1\\
&+\int_{-\infty+id_1^-}^{+\infty+id_1^-}\int_{-\infty+id_2^+}^{+\infty+id_2^+}\frac{g(z_1,z_2)}{(e^{2\pi i(z_1-ia_1)/h_1}-1)}\,dz_2\,dz_1\\
&+\int_{-\infty+id_1^-}^{+\infty+id_1^-}\int_{-\infty+id_2^+}^{+\infty+id_2^+}\frac{g(z_1,z_2)}{(e^{2\pi i(z_1-ia_1)/h_1}-1)(e^{-2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1\\
&+\int_{-\infty+id_1^+}^{+\infty+id_1^+}\int_{-\infty+id_2^-}^{+\infty+id_2^-}\frac{g(z_1,z_2)}{(e^{2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1\\
&+\int_{-\infty+id_1^+}^{+\infty+id_1^+}\int_{-\infty+id_2^-}^{+\infty+id_2^-}\frac{g(z_1,z_2)}{(e^{-2\pi i(z_1-ia_1)/h_1}-1)(e^{2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1\\
&+\int_{-\infty+id_1^+}^{+\infty+id_1^+}\int_{-\infty+id_2^+}^{+\infty+id_2^+}\frac{g(z_1,z_2)}{(e^{-2\pi i(z_1-ia_1)/h_1}-1)}\,dz_2\,dz_1\\
&+\int_{-\infty+id_1^+}^{+\infty+id_1^+}\int_{-\infty+id_2^+}^{+\infty+id_2^+}\frac{g(z_1,z_2)}{(e^{-2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1\\
&+\int_{-\infty+id_1^+}^{+\infty+id_1^+}\int_{-\infty+id_2^+}^{+\infty+id_2^+}\frac{g(z_1,z_2)}{(e^{-2\pi i(z_1-ia_1)/h_1}-1)(e^{-2\pi i(z_2-ia_2)/h_2}-1)}\,dz_2\,dz_1\Big).
\end{aligned}
\tag{3.13}
$$

It follows that

$$
\begin{aligned}
&|E^T_{h_1,h_2,\infty,\infty}(g,a_1,a_2)(x_1,x_2)|\\
\leq&\frac{e^{-\pi d_a^1/h_1-\pi d_a^2/h_2}}{(1-e^{-\pi d_a^1/h_1})(1-e^{-\pi d_a^2/h_2})}\Big(||g||^{-,-}+||g||^{-,+}+||g||^{+,-}+||g||^{+,+}\Big)\\
&+\frac{e^{-\pi d_a^1/h_1}}{(1-e^{-\pi d_a^1/h_1})}\Big(||g||^{-,+}+||g||^{+,+}\Big)+\frac{e^{-\pi d_a^2/h_2}}{(1-e^{-\pi d_a^2/h_2})}\Big(||g||^{+,-}+||g||^{+,+}\Big).
\end{aligned}
\tag{3.14}
$$

Now we move on to the truncation error part. The truncation error is then bounded

as:

$$
\begin{aligned}
&|T_{h_1,h_2,\infty,\infty}(g,a_1,a_2) - T_{h_1,h_2,M_1,M_2}(g,a_1,a_2)| \\
\leq & 4\Big(\sum_{m_1=M_1}^{\infty} \sum_{m_2=0}^{M_2} |g(m_1h_1+ia_1, m_2h_2+ia_2)|h_1h_2 \\
&+ \sum_{m_1=0}^{M_1} \sum_{m_2=M_2}^{\infty} |g(m_1h_1+ia_1, m_2h_2+ia_2)|h_1h_2 \\
&+ \sum_{m_1=M_1}^{\infty} \sum_{m_2=M_2}^{\infty} |g(m_1h_1+ia_1, m_2h_2+ia_2)|h_1h_2\Big) \\
\leq & 4\kappa\Big(\sum_{m_1=M_1}^{\infty} \sum_{m_2=0}^{M_2} + \sum_{m_1=0}^{M_1} \sum_{m_2=M_2}^{\infty} + \sum_{m_1=M_1}^{\infty} \sum_{m_2=M_2}^{\infty} \Big) \\
&(m_1h_1)^{n_1}(m_2h_2)^{n_2} \exp^{-c_1(m_1h_1)^{\nu_1}-c_2(m_2h_2)^{\nu_2}} h_1h_2, \qquad (3.15)
\end{aligned}
$$

if g satisfies (3.7). $y^n \exp(-cy^\nu)$ reaches its maximum, $(n/(c\nu))^{n/\nu}\exp(-n/\nu)$, when $y = (n/(c\nu))^{1/\nu}$.

When $M_1h_1 \geq (n_1/(c_1\nu_1))^{1/\nu_1}$, and $M_2h_2 \geq (n_2/(c_2\nu_2))^{1/\nu_2}$,

$$
\sum_{m_j=0}^{M_j} (m_jh_j)^{n_j} \exp(-c_j(m_jh_j)^{\nu_j})h_j \leq (M_j+1)h_j\Big(\frac{n_j}{c_j\nu_j}\Big)^{\frac{n_j}{\nu_j}} \exp(-\frac{n_j}{\nu_j}),
$$

and

$$
\begin{aligned}
&\textstyle\sum_{m_j=M_j}^{\infty} (m_jh_j)^{n_j} \exp(-c_j(m_jh_j)^{\nu_j})h_j \leq \tau_{M_jh_j} \\
= \quad & \tfrac{1}{\nu_j c_j^{(n_j+1)/\nu_j}} \Gamma\big(\tfrac{n_j+1}{\nu_j}, c_j(M_jh_j)^{\nu_j}\big), \quad j=1,2.
\end{aligned}
$$

The incomplete Gamma function $\Gamma(s,b)$ is bounded by a multiple of $b^{s-1}e^{-b}$.

Combining the discretization error and the truncation error, the total error bound will be found as (3.9). (3.10) can be obtained consequently by adding selection rules

of $h_1 = h_1(M_1)$ and $h_2 = h_2(M_2)$. $\square$

By selecting h_1 and h_2 by M_1 and M_2, the total approximation error is then bounded by a function of M_1 and M_2. The first term of (3.10) decays exponentially in terms of $M_1^{\frac{\nu_1}{1+\nu_1}}$, while the second decays exponentially in terms of $M_2^{\frac{\nu_2}{1+\nu_2}}$. The last term decays in the fastest way, thus, is dominated by the first two terms. To control the decaying rate of the total error bound, one just needs to adjust values of M_1 and M_2.

3.2 Analytic Bivariate Characteristic Function Inversion

In this section, we present approximation representations for bivariate cdf with characteristic functions in the analytic function class defined in Section 3.1. We then extend our results to expectations associated with indicator functions, which can be applied in option pricing.

3.2.1 Bivariate cdf approximation

Consider a continuous bivariate random variable (X_1, X_2). Denote its cdf by $F(x_1, x_2)$. [26] has proved that for univariate cases with $\phi \in L^1(\mathbb{R})$, cdf can be represented in terms of Hilbert Transform. On the basis of this result, the following theorem shows an alternative representation for $F(x_1, x_2)$:

Theorem 3.3. *Let $F(x_1, x_2)$ and $\phi(\xi_1, \xi_2)$ be the cdf and the characteristic function of a bivariate continuous distribution. Suppose that $\phi \in L^2(\mathbb{R}^2)$ and $\phi \in H(\mathcal{D}_{((d_1^-, d_1^+)\times(d_2^-, d_2^+))})$. Then for any $a_1 \in (0, d_1^+)$ and $a_2 \in (0, d_2^+)$,*

$$F(x_1, x_2) = -\frac{1}{4\pi^2} \int_{-\infty+ia_1}^{+\infty+ia_1} \int_{-\infty+ia_2}^{+\infty+ia_2} \frac{e^{-ix_1\xi_1 - ix_2\xi_2}\phi(\xi_1, \xi_2)}{\xi_1\xi_2} \, d\xi_2 \, d\xi_1 \tag{3.16}$$

Proof. Denote the underlying bivariate pdf as $f(x_1, x_2)$.

$$\begin{aligned}
F(x_1, x_2) &= \int_{\mathbb{R}} \int_{\mathbb{R}} 1_{(-\infty, x_1]}(y_1) 1_{(-\infty, x_2]}(y_2) f(y_1, y_2)\, dy_2\, dy_1 \\
&= \int_{\mathbb{R}} 1_{(-\infty, x_1]}(y_1) \mathcal{F}(1_{(-\infty, x_2]}(\cdot) f(y_1, \cdot))(0)\, dy_1 \\
&= \int_{\mathbb{R}} 1_{(-\infty, x_1]}(y_1) \big(\frac{1}{2} f(y_1) - \frac{i}{2} \mathcal{H}(e^{-i\xi_2 x_2} \phi_2(y_1, \xi_2))(0)\big)\, dy_1 \\
&= \frac{i}{2\pi} \int_{\mathbb{R}} 1_{(-\infty, x_1]}(y_1) \int_{-\infty + ia_2}^{+\infty + ia_2} \frac{e^{-ix_2\xi_2} \phi_2(y_1, \xi_2)}{\xi_2}\, d\xi_2\, dy_1 \\
&= -\frac{1}{4\pi^2} \int_{-\infty + ia_1}^{+\infty + ia_1} \int_{-\infty + ia_2}^{+\infty + ia_2} \frac{e^{-ix_1\xi_1 - ix_2\xi_2} \phi(\xi_1, \xi_2)}{\xi_1 \xi_2}\, d\xi_2\, d\xi_1, \qquad (3.17)
\end{aligned}$$

where $\phi_2(y_1, \xi_2) = \int_{\mathbb{R}} e^{i\xi_2 y_2} f(y_1, y_2)\, dy_2$. The third and the fourth equality follows directly from the following relationship between Fourier transform and Hilbert transform:

$$\mathcal{F}(1_{(-\infty, b)} \cdot f)(\xi) = \frac{1}{2} \hat{f}(\xi) - \frac{i}{2} e^{ib\xi} \mathcal{H}(e^{-ib\cdot} \hat{f}(\cdot))(\xi).$$

□

The bivariate cdf $F(x_1, x_2)$ can then be approximated by:

$$\begin{aligned}
& F_{h_1, h_2, M_1, M_2}(x_1, x_2) \\
= \quad & -\frac{1}{4\pi^2} \sum_{m_1 = -M_1}^{M_1} \sum_{m_2 = -M_2}^{M_2} \frac{h_1 h_2 e^{-is_1 x_1 - is_2 x_2} \phi(s_1, s_2)}{s_1 s_2}, \qquad (3.18)
\end{aligned}$$

where $s_1 = m_1 h_1 + ia_1$, and $s_2 = m_2 h_2 + ia_2$.

By applying the results of Theorem 3.2, the bound for the cdf approximation error can be obtained by the following corollary:

Corollary 3.4. *Suppose $\phi \in H(\mathcal{D}_{((d_1^-, d_1^+) \times (d_2^-, d_2^+))})$ and satisfies (3.7). Then for $a_1 \in (d_1^-, d_1^+)$ and $a_2 \in (d_2^-, d_2^+)$,*
denote $h(\xi_1, \xi_2) = \frac{\phi(\xi_1 + ia_1, \xi_2 + ia_2)}{(\xi_1 + ia_1)(\xi_2 + ia_2)} \in H(\mathcal{D}_{((-a_1, d_1^+ - a_1) \times (-a_2, d_2^+ - a_2))})$,

$$
\begin{aligned}
&|E^F_{h_1,h_2,M_1,M_2}(\phi,a_1,a_2)| \\
\leq &\frac{e^{-\pi d_a^1/h_1-\pi d_a^2/h_2}}{4\pi^2(1-e^{-\pi d_a^1/h_1})(1-e^{-\pi d_a^2/h_2})} \\
&\Big(||h||^{-,-}+e^{d_2^+x_2}||h||^{-,+}+e^{d_1^+x_1}||h||^{+,-}+e^{d_1^+x_1+d_2^+x_2}||h||^{+,+}\Big) \\
&+\frac{e^{-\pi d_a^1/h_1}}{4\pi^2(1-e^{-\pi d_a^1/h_1})}\Big(e^{d_2^+x_2}||h||^{-,+}+e^{d_1^+x_1+d_2^+x_2}||h||^{+,+}\Big) \\
&+\frac{e^{-\pi d_a^2/h_2}}{4\pi^2(1-e^{-\pi d_a^2/h_2})}\Big(e^{d_1^+x_1}||h||^{+,-}+e^{d_1^+x_1+d_2^+x_2}||h||^{+,+}\Big) \\
&+\frac{1}{\pi^2}(M_1+1)h_1\big(\frac{n_1-1}{c_1\nu_1}\big)^{\frac{n_1-1}{\nu_1}}\exp(-\frac{n_1-1}{\nu_1})\frac{e^{a_1x_1+a_2x_2}}{\nu_2c_2^{n_2/\nu_2}}\Gamma\big(\frac{n_2}{\nu_2},c_2(M_2h_2)^{\nu_2}\big) \\
&+\frac{1}{\pi^2}(M_2+1)h_2\big(\frac{n_2-1}{c_2\nu_2}\big)^{\frac{n_2-1}{\nu_2}}\exp(-\frac{n_2-1}{\nu_2})\frac{e^{a_1x_1+a_2x_2}}{\nu_1c_1^{n_1/\nu_1}}\Gamma\big(\frac{n_1}{\nu_1},c_1(M_1h_1)^{\nu_1}\big) \\
&+\frac{e^{a_1x_1+a_2x_2}}{\pi^2\nu_1\nu_2c_1^{(n_1+1)/\nu_1}c_2^{(n_2+1)/\nu_2}}\Gamma\big(\frac{n_1+1}{\nu_1},c_1(M_1h_1)^{\nu_1}\big)\Gamma\big(\frac{n_2+1}{\nu_2},c_2(M_2h_2)^{\nu_2}\big), \qquad (3.19)
\end{aligned}
$$

where $d_a^1=2\min(d_1^+-a_1,a_1)$ *and* $d_a^2=2\min(d_2^+-a_2,a_2)$.

Choose $M_1,M_2>0$, *let* $h_1=h_1(M_1)=(\pi d_a^1/c_1)^{\frac{1}{1+\nu_1}}M_1^{-\frac{\nu_1}{1+\nu_1}}$, $h_2=h_2(M_2)=(\pi d_a^2/c_2)^{\frac{1}{1+\nu_2}}M_2^{-\frac{\nu_2}{1+\nu_2}}$, *such that* $M_1h_1(M_1)\geq(n_1/(c_1\nu_1))^{1/\nu_1}$, *and* $M_2h_2(M_2)\geq(n_2/(c_2\nu_2))^{1/\nu_2}$, *then there exist constant* $C_1,C_2,C_3>0$ *independent of* M_1 *and* M_2, *such that the total error is bounded by:*

$$
\begin{aligned}
&|E^F_{h_1(M_1),h_2(M_2),M_1,M_2}(\phi,a_1,a_2)| \\
&\leq C_1(e^{d_2^+x_2}+e^{d_1^+x_1+d_2^+x_2}+e^{a_1x_1+a_2x_2}M_2M_1^{\frac{n_1+1-\nu_1}{1+\nu_1}})\exp(-c_1^{-\frac{1}{1+\nu_1}}\big(\pi d_a^1M_1\big)^{\frac{\nu_1}{1+\nu_1}}) \\
&+C_2(e^{d_1^+x_1}+e^{d_1^+x_1+d_2^+x_2}+e^{a_1x_1+a_2x_2}M_1M_2^{\frac{n_2+1-\nu_2}{1+\nu_2}})\exp(-c_2^{-\frac{1}{1+\nu_2}}\big(\pi d_a^2M_2\big)^{\frac{\nu_2}{1+\nu_2}}) \\
&+C_3\Big(e^{a_1x_1+a_2x_2}M_1^{\frac{n_1+1-\nu_1}{1+\nu_1}}M_2^{\frac{n_2+1-\nu_2}{1+\nu_2}}+e^{d_2^+x_2} \\
&+e^{d_1^+x_1}+e^{d_1^+x_1+d_2^+x_2})\exp(-c_1^{-\frac{1}{1+\nu_1}}\big(\pi d_a^1M_1\big)^{\frac{\nu_1}{1+\nu_1}}-c_2^{-\frac{1}{1+\nu_2}}\big(\pi d_a^2M_2\big)^{\frac{\nu_2}{1+\nu_2}})\Big). \qquad (3.20)
\end{aligned}
$$

One can also select M_2 by a function of M_1: $M_1^{\nu_1/(1+\nu_1)} = M_2^{\nu_2/(1+\nu_2)}$, $M_2(M_1) = M_1^{\frac{\nu_1(1+\nu_2)}{\nu_2(1+\nu_1)}}$ such that the first two terms in the error bound decay at the same rate. In stead of selecting h_1, h_2, M_1, and M_2 separately, one just needs to choose M_1 to control the total error decaying rate. The total computational cost is then $O(M_1^{1+\frac{\nu_1(1+\nu_2)}{\nu_2(1+\nu_1)}})$.

3.2.2 Expectations involving indicator functions

In the previous section, we present how to approximate $F(x_1, x_2)$, which is basically the expectation of $I(X_1 <= x_1, X_2 <= x_2)$. In this section, we present how to derive representations for some expectations involving indicator functions. The representations lead to the expressions of 2d European option prices, which will be demonstrated in Section 3.3.3.

Theorem 3.5. *Let $\phi(\xi_1, \xi_2)$ be the characteristic function of a bivariate random variable $X = (X_1, X_2)$. Suppose that $\phi \in H(\mathcal{D}_{((d_1^-, d_1^+)\times(d_2^-, d_2^+))})$. Then for any $b_1 \in (d_1^-, d_1^+)$, $b_2 \in (d_2^-, d_2^+)$, $c_1, c_2 \in \mathbb{R}$,*

$$\begin{aligned}
&\mathbb{E}\big[e^{-b_1X_1-b_2X_2}1_{\{X_1\le c_1, X_2\le c_2\}}\big] \\
&= -\frac{\phi(ib_1, ib_2)}{4\pi^2}\int_{-\infty+ia_1}^{+\infty+ia_1}\int_{-\infty+ia_2}^{+\infty+ia_2}\frac{e^{-ic_1\xi_1-ic_2\xi_2}\phi^*_{b_1,b_2}(\xi_1,\xi_2)}{\xi_1\xi_2}\,d\xi_2\,d\xi_1,
\end{aligned} \tag{3.21}$$

where

$$\phi^*_{b_1,b_2}(\xi_1,\xi_2) = \frac{\phi(\xi_1+ib_1, \xi_2+ib_2)}{\phi(ib_1, ib_2)},$$

for $-b_1 < a_1 < d_1^+ - b_1, -b_2 < a_2 < d_2^+ - b_2$.

Proof. Define a new probability measure $\mathbb{P}^*_{b_1,b_2}$ by:

$$Z = \frac{e^{-b_1X_1-b_2X_2}}{\phi(ib_1, ib_2)} = \frac{d\mathbb{P}^*_{b_1,b_2}}{d\mathbb{P}}.$$

The characteristic function associated with measure $\mathbb{P}^*_{b_1,b_2}$ is then $\phi^*_{b_1,b_2}$.

By change of measure, this expectation is converted to a cdf under $\mathbb{P}^*_{b_1,b_2}$.

$$
\begin{aligned}
&\mathbb{E}\big[e^{-b_1X_1-b_2X_2}1_{\{X_1\le c_1,X_2\le c_2\}}\big]\\
&=\phi(ib_1,ib_2)\mathbb{E}\big[Z1_{\{X_1\le c_1,X_2\le c_2\}}\big]=\phi(ib_1,ib_2)\mathbb{E}^*\big[1_{\{X_1\le c_1,X_2\le c_2\}}\big]\\
&=\phi(ib_1,ib_2)\mathbb{P}^*_{b_1,b_2}(X_1\le c_1,X_2\le c_2)\\
&=-\frac{\phi(ib_1,ib_2)}{4\pi^2}\int_{-\infty+ia_1}^{+\infty+ia_1}\int_{-\infty+ia_2}^{+\infty+ia_2}\frac{e^{-ic_1\xi_1-ic_2\xi_2}\phi^*_{b_1,b_2}(\xi_1,\xi_2)}{\xi_1\xi_2}\,d\xi_2\,d\xi_1.
\end{aligned}
$$

The last equality comes from Theorem 3.3. □

The conclusion can be extended to similar expectations involving indicator functions:

$$
\begin{aligned}
&\mathbb{E}\big[e^{-b_1X_1-b_2X_2}1_{\{X_1>c_1\}}\big]\\
=&\phi(ib_1,ib_2)\mathbb{P}^*_{b_1,b_2}(X_1>c_1)=\phi(ib_1,ib_2)\big(1-\mathbb{P}^*_{b_1,b_2}(X_1\le c_1)\big)\\
=&\phi(ib_1,ib_2)\Big(1-\frac{i}{2\pi}\int_{-\infty+ia_1}^{+\infty+ia_1}\frac{e^{-ic_1\xi_1}\phi^*_{b_1,b_2}(\xi_1,0)}{\xi_1}\,d\xi_1\Big),
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\big[e^{-b_1X_1-b_2X_2}1_{\{X_1>c_1,X_2>c_2\}}\big]\\
=&\phi(ib_1,ib_2)\mathbb{P}^*_{b_1,b_2}(X_1>c_1,X_2>c_2)\\
=&\phi(ib_1,ib_2)\big(1-\mathbb{P}^*_{b_1,b_2}(X_1\le c_1)-\mathbb{P}^*_{b_1,b_2}(X_2\le c_2)+\mathbb{P}^*_{b_1,b_2}(X_1\le c_1,X_2\le c_2)\big)\\
=&\phi(ib_1,ib_2)\Big(1-\frac{i}{2\pi}\int_{-\infty+ia_1}^{+\infty+ia_1}\frac{e^{-ic_1\xi_1}\phi^*_{b_1,b_2}(\xi_1,0)}{\xi_1}\,d\xi_1-\frac{i}{2\pi}\int_{-\infty+ia_2}^{+\infty+ia_2}\frac{e^{-ic_2\xi_2}\phi^*_{b_1,b_2}(0,\xi_2)}{\xi_2}\,d\xi_2\\
&-\frac{1}{4\pi^2}\int_{-\infty+ia_1}^{+\infty+ia_1}\int_{-\infty+ia_2}^{+\infty+ia_2}\frac{e^{-ic_1\xi_1-ic_2\xi_2}\phi^*_{b_1,b_2}(\xi_1,\xi_2)}{\xi_1\xi_2}\,d\xi_2\,d\xi_1\Big).
\end{aligned}
\tag{3.22}
$$

Next we derive expressions for some other useful indicator functions related expectations.

Theorem 3.6. *Let $\phi(\xi_1, \xi_2)$ be the characteristic function of a bivariate random variable $X = (X_1, X_2)$. Suppose that $\phi \in H(\mathcal{D}_{((d_1^-, d_1^+)\times(d_2^-, d_2^+))})$. Then for any $b_1 \in (d_1^-, d_1^+)$, $b_2 \in (d_2^-, d_2^+)$, $c_1, c_2 \in \mathbb{R}$,*

$$\begin{aligned}
&\mathbb{E}\big[e^{-b_1X_1-b_2X_2}1_{\{X_1\le c_1, X_2-X_1\le c_2\}}\big] \\
&= -\frac{\phi(ib_1, ib_2)}{4\pi^2}\int_{-\infty+ia_1}^{+\infty+ia_1}\int_{-\infty+ia_2}^{+\infty+ia_2}\frac{e^{-ic_1\xi_1-ic_2\xi_2}\tilde{\phi}^*_{b_1,b_2}(\xi_1,\xi_2)}{\xi_1\xi_2}\,d\xi_2\,d\xi_1,
\end{aligned} \tag{3.23}$$

where

$$\tilde{\phi}^*_{b_1,b_2}(\xi_1,\xi_2) = \frac{\phi(\xi_1-\xi_2+ib_1, \xi_2+ib_2)}{\phi(ib_1, ib_2)},$$

for $-b_1 < a_1 < d_1^+ - b_1, -b_2 < a_2 < d_2^+ - b_2$.

This theorem can be proved by defining:

$$(\tilde{X}_1, \tilde{X}_2) = (X_1, X_2 - X_1).$$

Hence,

$$\tilde{\phi}(\xi_1,\xi_2) = \mathbb{E}[e^{i\xi_1X_1+i\xi_2(X_2-X_1)}] = \phi(\xi_1-\xi_2, \xi_2).$$

Corollary 3.7. *Let $\phi(\xi_1, \xi_2)$ be the characteristic function of a bivariate random variable $X = (X_1, X_2)$. Suppose that $\phi \in H(\mathcal{D}_{((d_1^-, d_1^+)\times(d_2^-, d_2^+))})$. Then for any $b_1 \in (d_1^-, d_1^+)$, $b_2 \in (d_2^-, d_2^+)$, $c_1, c_2 \in \mathbb{R}$,*

$$\begin{aligned}
&\mathbb{E}\big[e^{-b_1X_1-b_2X_2}1_{\{X_1-X_2\le c_1, X_2\le c_2\}}\big] \\
&= -\frac{\phi(ib_1, ib_2)}{4\pi^2}\int_{-\infty+ia_1}^{+\infty+ia_1}\int_{-\infty+ia_2}^{+\infty+ia_2}\frac{e^{-ic_1\xi_1-ic_2\xi_2}\bar{\phi}^*_{b_1,b_2}(\xi_1,\xi_2)}{\xi_1\xi_2}\,d\xi_2\,d\xi_1,
\end{aligned} \tag{3.24}$$

where

$$\bar{\phi}^*_{b_1,b_2}(\xi_1,\xi_2) = \frac{\phi(\xi_1 + ib_1, \xi_2 - \xi_1 + ib_2)}{\phi(ib_1, ib_2)},$$

for $-b_1 < a_1 < d_1^+ - b_1, -b_2 < a_2 < d_2^+ - b_2$.

Expectations above can all be inverted conveniently and efficiently by the trapezoidal rule when the characteristic function lies in our analytic class.

3.3 Numerical Experiments

In this section, we exhibit cdf approximation and 2d European option pricing under a bivariate NIG model through some numerical experiments. The numerical experiments are performed in C++ on a MacBook Pro with 16GB Memory and 2.5GHz CPU.

A bivariate NIG (Normal Inverse Gaussian) process is a Lévy process with characteristic function:

$$\boldsymbol{\phi_{\mathbf{X}}(\xi) = \exp\left(iT\mu^T\xi + \delta T\sqrt{\alpha^2 - \beta^T\Gamma\beta} - \delta T\sqrt{\alpha^2 - (\beta + i\xi)^T\Gamma(\beta + i\xi)}\right),}$$

where $\alpha > 0, \quad \beta \in \mathbb{R}^2, \quad \delta > 0, \quad T > 0, \quad \mu = [\mu_1 \quad \mu_2]^T \in \mathbb{R}^2$, and $\mathbf{\Gamma} \in \mathbb{R}^{2*2}$ is positive semi-definite with determinant 1.

The analyticity strip for $\phi_{\mathbf{X}}(\xi)$ contains the strip in which the real part of $(\beta + i\xi)^T\Gamma(\beta + i\xi) - \alpha^2$ is negative. One sufficient condition is $\{\xi \in \mathbb{C}^2 | (\beta - \Im(\xi))^T\Gamma(\beta - \Im(\xi)) \leq 0\}$.

To find a two-dimensional strip with constant boundaries, one just needs to find a rectangle contained in the stretched ellipse. Each dimension represents the imaginary part of the corresponding dimension in ξ. To be accurate and efficient enough, the rectangle has to be as large as possible.

A reasonable parameter setting is as follows:

$$\mu = [0.149752 \quad 0.227841]^T, \alpha = 15.0, \delta = 0.5, \beta = [-4.0 \quad -6.0]^T, T = 1,$$

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{0.9888} & \mathbf{-0.0025} \\ \mathbf{-0.0025} & \mathbf{1.0113} \end{bmatrix}.$$

Figure 3.1 shows the rectangle we choose for this parameter setting:

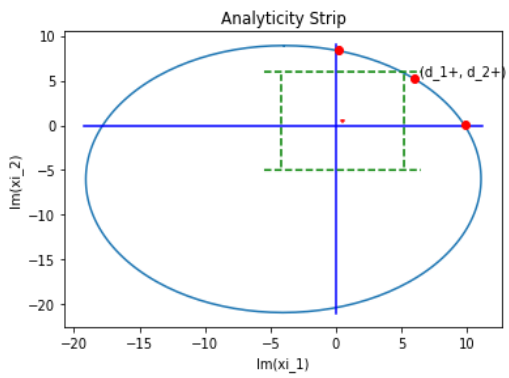

Figure 3.1: Analyticity Strip for the Bivariate NIG Model

3.3.1 Numerical Experiments: Bivariate pdf approximation

By setting small enough discretization parameters, h_1 and h_2, and large enough truncation parameters, M_1 and M_2, the benchmark is computed to be:

$$f_{0.10,0.10,1000,1000}(x_1 = 0.2, x_2 = 0.3) = 0.9820209383008591.$$

We start with the analysis of the truncation error from the first dimension and the discretization error from the first dimension, and then move on to the total approximation error analysis from both dimensions.

- **Truncation error analysis**

We, first of all, fix $h_1 = h_2 = 0.10, M_2 = 1,000$ and increase the square root of M_1 from 6 to 30, with step size 4. The following table records the truncation error from the first dimension:

Bivariate NIG Model pdf Truncation Error			
M_1	$M_1 h_1$	$F_{h_1,h_2,M_1,M_2}(0.2, 0.3)$	abs error
36	1.8000	0.3857564792	5.96E-01
100	5.0000	0.8550298611	1.27E-01
196	9.8000	1.0265893600	4.46E-02
324	16.2000	0.9946052323	1.26E-02
484	24.2000	0.9826618948	6.41E-04
676	33.8000	0.9820292937	8.36E-06
900	45.0000	0.9820209610	2.27E-08

Table 3.1: Truncation Error in Estimating Bivariate NIG pdf: Exact Value = 0.98202094

The following figure plots the truncation error in log scale as a function of M_1:

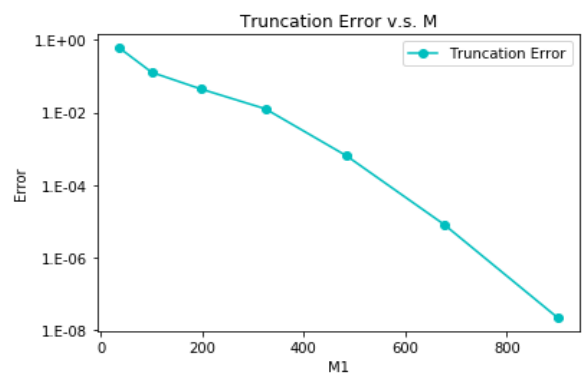

Figure 3.2: Truncation Error in Estimating Bivariate NIG cdf: Exact Value = 0.98202094

As the plot shows, the truncation error from the first dimension decays exponentially in terms of M_1.

- **Discretization error analysis**

  By setting $h_2 = 0.10, M_1 = M_2 = 1000$, and increasing $1/h_1$, we perform an analysis on the discretization error from the first dimension. The numerical experiment results are included in Table 3.2:

Bivariate NIG Model pdf Discretization Error			
$1/h_1$	h_1	$F_{h_1,h_2,M_1,M_2}(0.2, 0.3)$	abs error
0.05	20.0000	2.1671844470	1.19E+00
0.10	10.0000	1.0812294130	9.92E-02
0.15	6.6670	0.9859055288	3.88E-03
0.20	5.0000	0.9821368740	1.16E-04
0.25	4.0000	0.9820240970	3.16E-06
0.30	3.3330	0.9820210209	8.26E-08

Table 3.2: Discretization Error in Estimating Bivariate NIG pdf: Exact Value = 0.98202094

  The following plot shows that the discretization error is decaying exponentially in $1/h_1$.

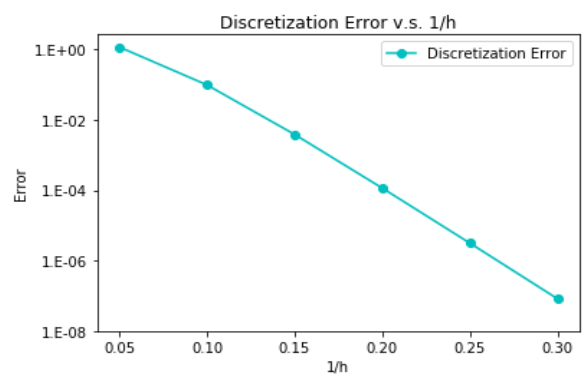

Figure 3.3: Discretization Error in Estimating Bivariate NIG pdf: Exact Value = 0.98202094

- **Total error analysis**

  By choosing M_1, and selecting M_2, h_1 and h_2 by M_1, the total approximation error is included in the following table:

Bivariate NIG Model pdf Total Approximation Error							
M_1	M_2	h_1	h_2	M_1h_1	M_2h_2	$F(0.2, 0.3)$	abs error
9	10	1.2160	1.1300	10.9457	11.2975	1.13580040	1.54E-01
25	27	0.7297	0.6875	18.2428	18.5637	0.99305863	1.10E-02
49	52	0.5212	0.4954	25.5399	25.7623	0.98237331	3.52E-04
81	85	0.4054	0.3875	32.8370	32.9376	0.98202482	3.88E-06
121	127	0.3317	0.3170	40.1341	40.2610	0.98202077	1.66E-07
169	177	0.2807	0.2685	47.4312	47.5302	0.98202093	7.75E-09

Table 3.3: Total Approximation Error in Estimating Bivariate NIG pdf: Exact Value = 0.98202094

The plot 3.4 shows that the total approximation error is decaying exponentially in $M_1^{1/2}$, as suggested by our theoretical convergence rate: $O(e^{-\sqrt{M_1}})$.

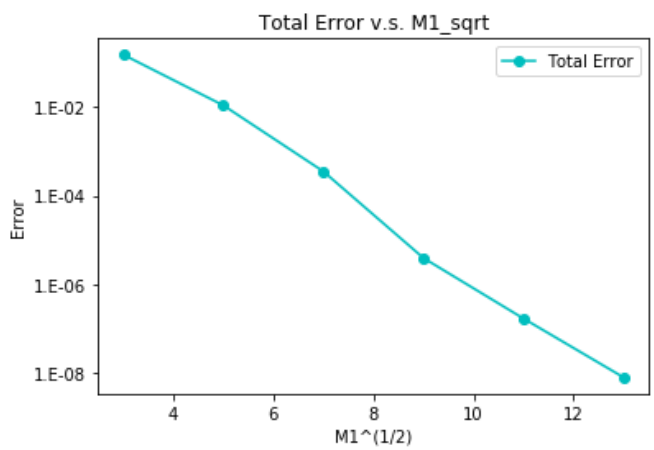

Figure 3.4: Total Approximation Error in Estimating Bivariate NIG pdf: Exact Value = 0.98202094

3.3.2 Numerical Experiments: Bivariate cdf approximation

To investigate accuracy of our approximation, since there is no close-form solution to the cdf, we choose small enough discretization parameters and large enough truncation parameters: $h_1 = h_2 = 0.1, M_1 = M_2 = 1000$. The cdf benchmark is then computed to be:

$$F_{0.1,0.1,1000,1000}(x_1 = 0.2, x_2 = 0.3) = 0.7943850860032705.$$

To have a detailed evaluation of the approximation performance, we analyze truncation error and discretization error separately.

- **Truncation error analysis**

  We first discuss the truncation error from the first dimension by setting a large enough truncation level for the second dimension, and small enough discretization levels for both dimensions: $h_1 = h_2 = 0.1, M_2 = 1000$. Each time we increase square root of M_1 by 5 from 5 to 30. Then the total absolute error is decreasing exponentially, as shown in Table 3.4. Equivalently, the log error is decreasing linearly in terms of M_1 (Figure 3.5).

Bivariate NIG Model cdf Truncation Error			
M_1	$M_1 h_1$	$F_{h_1,h_2,M_1,M_2}(0.2, 0.3)$	abs error
25	1.2500	0.4401345126	3.54E-01
100	5.0000	0.7006439336	9.37E-02
225	11.2500	0.7889790130	5.41E-03
400	20.0000	0.7943703611	1.47E-05
625	31.2500	0.7943854184	3.32E-07
900	45.0000	0.7943850866	5.77E-10

Table 3.4: Truncation Error in Estimating Bivariate NIG cdf: Exact Value = 0.7943850860

- **Discretization error analysis**

  We then investigate the discretization error from the first dimension by setting large M_1 and M_2 and small enough h_2. In increasing $1/h_1$ linearly, we have linearly decreasing log error, as shown in Table 3.5 and Figure 3.6.

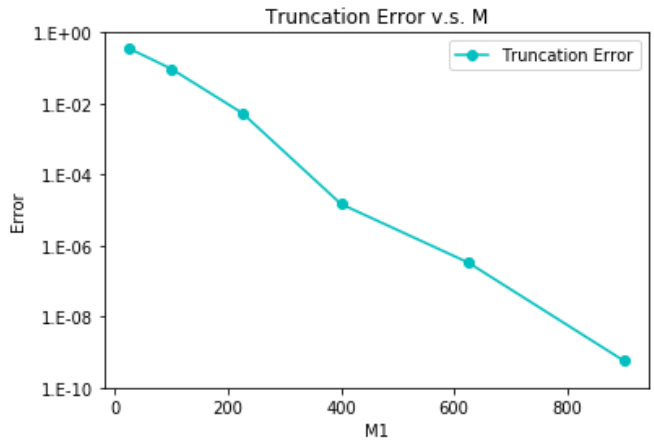

Figure 3.5: Truncation Error in Estimating Bivariate NIG cdf: Exact Value = 0.7943850860

Bivariate NIG Model cdf Discretization Error			
$1/h_1$	h_1	$F_{h_1,h_2,M_1,M_2}(0.2, 0.3)$	abs error
1	1.0000	0.8366819972	4.23E-02
2	0.5000	0.7961371861	1.75E-03
3	0.3333	0.7944606658	7.56E-05
4	0.2500	0.7943883518	3.27E-06
5	0.2000	0.7943852271	1.41E-07
6	0.1667	0.7943850921	6.05E-09

Table 3.5: Discretization Error in Estimating Bivariate NIG cdf: Exact Value = 0.7943850860

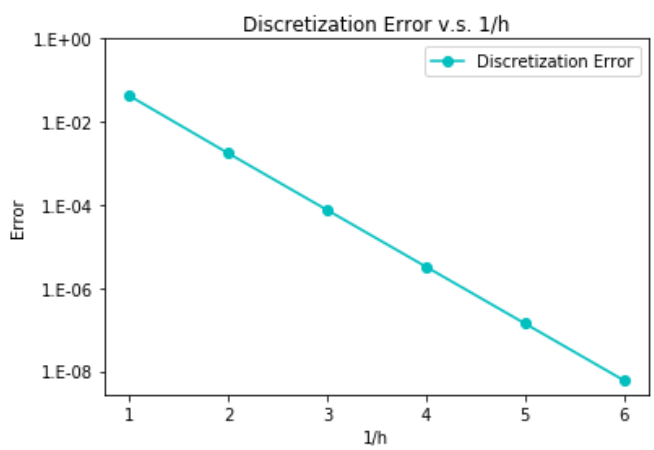

Figure 3.6: Discretization Error in Estimating Bivariate NIG cdf: Exact Value = 0.7943850860

- **Total error analysis**

  Previous sections investigate truncation and discretization error when the second dimension is fixed. In this section, we evaluate the total approximation error by choosing M_2 based on M_1, and selecting h_1, h_2 as functions of M_1 and M_2. The associated total approximation error is :

Bivariate NIG Model cdf Total Approximation Error							
M_1	M_2	h_1	h_2	M_1h_1	M_2h_2	$F(0.2, 0.3)$	abs error
25	27	0.7297	0.6875	18.2428	18.5637	0.81631346	2.19E-02
64	67	0.4561	0.4365	29.1885	29.2429	0.79597802	1.59E-03
121	127	0.3317	0.3170	40.1341	40.2610	0.79449935	1.14E-04
196	205	0.2606	0.2495	51.0798	51.1517	0.79439342	8.33E-06
289	302	0.2146	0.2056	62.0255	62.0850	0.79438569	6.07E-07
400	418	0.1824	0.1747	72.9711	73.0418	0.79438513	4.42E-08

Table 3.6: Total Approximation Error in Estimating Bivariate NIG cdf: Exact Value = 0.79438509

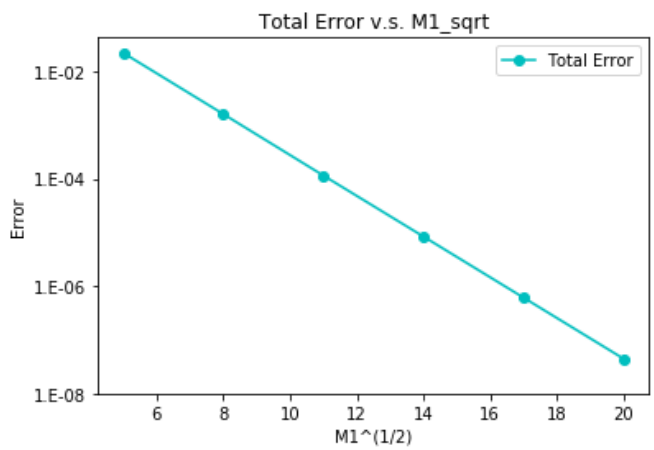

Figure 3.7: Total Approximation Error in Estimating Bivariate NIG cdf: Exact Value = 0.79438509

3.3.3 Numerical Experiments: 2d European option pricing

Another application lies in pricing 2d European options. In this section, we price a 2d European call option with final payoff:

$$(\max(S_T^1, S_T^2) - \mathcal{K})^+,$$

where $S_T^1 = S_0^1 e^{X_T^1}, S_T^2 = S_0^2 e^{X_T^2}$ are the final prices of Stock 1 and Stock 2 at maturity T, S_0^1, S_0^2 are the initial stock prices, and $\mathcal{K}$ is the strike price.

Denote the risk-free interest rate by r and the dividend yields that the underlying assets are paying for by q_1 and q_2. For the discounted gains to be martingales, we need $\mathbb{E}[S_T^1] = S_0^1 e^{(r-q_1)T}$ and $\mathbb{E}[S_T^2] = S_0^2 e^{(r-q_2)T}$, which requires that

$$\mu_1 = r - q_1 - \delta\sqrt{\alpha^2 - \beta^T\Gamma\beta} + \delta\sqrt{\alpha^2 - (\beta^T + [1 \quad 0])\Gamma(\beta + [1 \quad 0]^T)}, \tag{3.25}$$

$$\mu_2 = r - q_2 - \delta\sqrt{\alpha^2 - \beta^T\Gamma\beta} + \delta\sqrt{\alpha^2 - (\beta^T + [0 \quad 1])\Gamma(\beta + [0 \quad 1]^T)}. \tag{3.26}$$

The price of a 2d European call option with maturity $T > 0$ and strike price $\mathcal{K} > 0$ is computed to be:

$$\begin{aligned}
&C(S_0^1, S_0^2, X_1, X_2, \mathcal{K}) \\
&= e^{-rT}\mathbb{E}\big[(\max(S_0^1 e^{X_1}, S_0^2 e^{X_2}) - \mathcal{K})^+\big] \\
&= e^{-rT}\Big(\mathbb{E}[(S_0^2 e^{X_2} - \mathcal{K})1_{\{X_1 \le x_0^1\}}1_{\{X_2 > x_0^2\}}] + \mathbb{E}[(S_0^2 e^{X_2} - \mathcal{K})1_{\{X_1 > x_0^1\}}1_{\{X_2 - X_1 > x_0^2 - x_0^1\}}] \\
&\qquad + \mathbb{E}[(S_0^1 e^{X_1} - \mathcal{K})1_{\{X_2 > x_0^2\}}1_{\{X_1 - X_2 > x_0^1 - x_0^2\}}] + \mathbb{E}[(S_0^1 e^{X_1} - \mathcal{K})1_{\{X_1 > x_0^1\}}1_{\{X_2 \le x_0^2\}}]\Big),
\end{aligned}$$

where $x_0^1 = \ln(\mathcal{K}/S_0^1), x_0^2 = \ln(\mathcal{K}/S_0^2)$. This can be approximated by expanding it and applying Theorem 3.5, Theorem 3.6 and Corollary 3.7.

We then perform numerical experiments by borrowing previous parameter setting. And $r = 0.02, q_1 = q_2 = 0.02$. μ_1 and μ_2 are computed to be 0.149752 and 0.227841, by (3.25) and (3.26).

The following table and plot include the numerical results in estimating the 2d European call option price under Bivariate NIG model (Table 3.7 and Figure 3.8):

2d European Option Pricing Error							
M_1	M_2	h_1	h_2	M_1h_1	M_2h_2	Call Price	abs error
25	27	0.7297	0.6875	18.2428	18.5637	0.15149415	3.81E-05
64	67	0.4561	0.4365	29.1885	29.2429	0.15154069	8.46E-06
121	127	0.3317	0.3170	40.1341	40.2610	0.15153318	9.40E-07
196	205	0.2606	0.2495	51.0798	51.1517	0.15153229	5.86E-08
289	302	0.2146	0.2056	62.0255	62.0850	0.15153224	3.63E-09

Table 3.7: Total Approximation Error in Pricing 2d European Call Options Under Bivariate NIG model: Exact Value = 0.151532236

The benchmark is selected by setting $h_1 = h_2 = 0.1, M_1 = M_2 = 1000$.

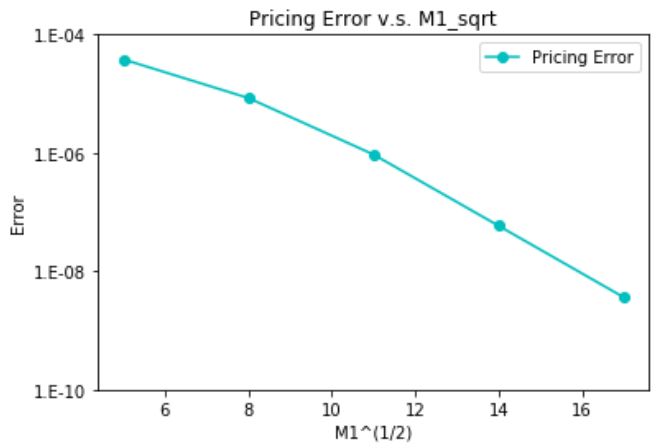

Figure 3.8: 2d European Call Option Pricing Error Analysis

The total approximation error soon converges to $1.0E-07$ level. The plot 3.8 demonstrates that the log error is decreasing linearly in terms of $M_1^{1/2}$.

Chapter 4

Monte Carlo Estimation of Sensitivities from Analytic Characteristic Functions

In financial derivatives trading, sensitivity analysis plays a significant role in risk management. The hedging strategy for a typical financial derivative product relies on the sensitivities of its price with respect to changes in different variables. For some financial models, such as Black-Scholes-Merton (BSM) model, we have closed-form expressions for option sensitivities. When such analytic solutions are not known to us, Monte Carlo simulation is widely used then.

The price of a financial derivative is defined by integrating the product of the discounted payoff function and the corresponding density function. Pathwise derivative method (PDM) and likelihood ratio method (LRM) can be applied in sensitivity estimation by differentiating the integral with respect to a certain parameter of interest. However, pathwise derivative method requires the smoothness of the payoff function, which is often not satisfied. The likelihood ratio method is particularly attractive as

long as the probability density function is differentiable with respect to the parameter. It has no restrictions on the discounted pay-off function.

4.1 The likelihood ratio method

To illustrate the likelihood ratio method, we, first of all, define the likelihood ratio estimator.

4.1.1 The likelihood ratio estimator

Let X be a continuous random variable with probability density function, $f_\theta(x)$, depending on a parameter $\theta \in \mathbb{R}$, which is a parameter of interest. The expectation of a given real-valued function $h(X)$ is:

$$\mathbb{E}_\theta[h(X)] = \int_{\mathbb{R}} h(x) f_\theta(x)\, dx.$$

The sensitivity of the expectation with respect to θ is measured by the derivative of the above expectation with respect to θ. Denote $g_\theta(x) = df_\theta(x)/\, d\theta$. Assuming validity of the interchange of the order of integration and differentiation, the sensitivity is equivalent to the following expectation:

$$\begin{aligned} \frac{d}{d\theta}\mathbb{E}_\theta[h(X)] &= \frac{d}{d\theta}\int_{\mathbb{R}} h(x) f_\theta(x)\, dx = \int_{\mathbb{R}} h(x) g_\theta(x)\, dx \\ &= \int_{\mathbb{R}} h(x)\frac{g_\theta(x)}{f_\theta(x)} f_\theta(x)\, dx = \mathbb{E}_\theta[h(x) S_\theta(X)], \end{aligned} \tag{4.1}$$

where $S_\theta(X) = g_\theta(x)/f_\theta(x)$ is known as a score function. We call $h(X)S_\theta(X)$ a likelihood ratio estimator of the sensitivity.

To be more general, let $X_1, \cdots, X_d$ be d independent continuous random variables,

with probability density functions, $f_\theta^1(x_1), \cdots, f_\theta^d(x_d)$, depending on a common parameter $\theta \in \mathbb{R}$. The expectation of a real-value function $h(X_1, \cdots, X_d)$ is denoted by:

$$\mathbb{E}_\theta[h(X_1, \cdots, X_d)] = \int_{\mathbb{R}^d} h(x_1, \cdots, x_d) \prod_{i=1}^{d} f_\theta^i(x_i)\, dx_d \cdots dx_1.$$

Its sensitivity with respect to parameter θ is:

$$\begin{aligned}
\frac{d}{d\theta}\mathbb{E}_\theta[h(X_1, \cdots, X_d)] &= \frac{d}{d\theta}\int_{\mathbb{R}^d} h(x_1, \cdots, x_d) \prod_{i=1}^{d} f_\theta^i(x_i)\, dx_d \cdots dx_1 \\
&= \int_{\mathbb{R}^d} h(x_1, \cdots, x_d) \sum_{j=1}^{d} \big(g_\theta^j(x_j) \prod_{i=1, i\neq j}^{d} f_\theta^i(x_i)\big)\, dx_d \cdots dx_1 \\
&= \int_{\mathbb{R}^d} h(x_1, \cdots, x_d) \big(\sum_{j=1}^{d} \frac{g_\theta^j(x_j)}{f_\theta^j(x_j)}\big) \prod_{i=1}^{d} f_\theta^i(x_i)\, dx_d \cdots dx_1 \quad (4.2) \\
&= \mathbb{E}_\theta[h(X_1, \cdots, X_d) S_\theta(X_1, \cdots, X_d)],
\end{aligned}$$

assuming that the interchange of the orders of integration and differentiation is valid. Here the score function is $S_\theta(X_1, \cdots, X_d) = \sum_{j=1}^{d} \frac{g_\theta^j(x_j)}{f_\theta^j(x_j)}$, where $g_\theta^j(x_j) = df_\theta^j(x_j)/\, d\theta$. And the likelihood ratio estimator $h(X_1, \cdots, X_d) S_\theta(X_1, \cdots, X_d)$ can be proved to be an unbiased estimator of the sensitivity.

In many financial applications, we do not have access to the closed-form expressions of the probability density functions or the cumulative distribution functions. However, if the characteristic functions of the distributions are available to us, both the density functions and the score functions can be approximated, and hence, the likelihood ratio method can be implemented in estimating the sensitivities. In this thesis, we will be addressing how to evaluate (4.1) and (4.2) through Monte Carlo simulation, and more specifically, likelihood ratio method.

4.1.2 Monte Carlo estimation from tabulated data

Evaluating the sensitivity through Monte Carlo simulation needs approximation and tabulation of the cumulative distribution function. The score function S_θ can also be approximated from tabulated data. Before that, let's review the inverse transform method for simulating X from the tabulated cdf values.

- **The one-dimensional case**

  Let's first discuss the one-dimensional case. Denote the cdf of X by $F_\theta(x)$. Consider an interval $\chi = [x_0, x_K]$, equally divided into K sub-intervals, each of length $\eta = (x_K - x_0)/K$. Denote the k-th nodes of the sub-intervals by $x_k = x_0 + k\eta, 0 \leq k \leq K$. Cumulative distribution functions on these discrete nodes, $F_\theta(x_k)$, can be approximated by $\tilde{F}_k$ by inverting the characteristic function. More details and numerical examples can be found in [26]. Then the cumulative distribution function of X can be approximated by linear interpolation:

$$\tilde{F}_\theta(x) = \frac{1}{\eta}\sum_{k=1}^{K}\big((x_k - x)\tilde{F}_{k-1} + (x - x_{k-1})\tilde{F}_k\big)1_{[x_{k-1},x_k)}(x) + 1_{[x_K,\infty)}(x). \quad (4.3)$$

  $\tilde{F}_\theta(x)$ is a piecewise linear function with a mixed type, continuous on (x_0, x_K) and with probability masses at x_0 and x_K. X can be simulated from it through the inverse transform method. [26] presents explicit and computable upper bounds for the bias introduced by truncating the support of the distribution, linear interpolation, and the errors in approximating the cdf, in particular, when the cdf is approximated from an analytic characteristic function.

  Taking the derivative of $\tilde{F}_\theta(x)$ with respect to x gives us a piecewise constant approximation, $\tilde{f}_\theta(x)$, to $f_\theta(x)$. It is the summation of a piecewise constant

function on (x_0, x_K) and two Dirac Delta functions representing the probability masses:

$$\begin{aligned}\tilde{f}_\theta(x) = \quad & \tilde{F}_0 \cdot \delta(x - x_0) + (1 - \tilde{F}_K) \cdot \delta(x - x_K) \\ & + \frac{1}{\eta} \sum_{k=1}^{K-1} (\tilde{F}_k - \tilde{F}_{k-1}) \cdot 1_{(x_{k-1}, x_k]}(x) + \frac{1}{\eta} (\tilde{F}_K - \tilde{F}_{K-1}) \cdot 1_{(x_{K-1}, x_K)}(x),\end{aligned} \tag{4.4}$$

where $1_A(x)$ is an indicator function which takes value 1 if $x \in A$ and 0 otherwise. Then it remains to approximate $g_\theta(x)$. We accomplish that by a piecewise linear function:

$$\begin{aligned}\tilde{g}_\theta(x) = \quad & \frac{1}{\eta} \sum_{k=1}^{K-1} \big((x_k - x)\tilde{g}_{k-1} + (x - x_{k-1})\tilde{g}_k\big) 1_{[x_{k-1}, x_k)}(x) \\ & + \frac{1}{\eta} \big((x_K - x)\tilde{g}_{K-1} + (x - x_{K-1})\tilde{g}_K\big) 1_{[x_{K-1}, x_K]}(x),\end{aligned} \tag{4.5}$$

where $\tilde{g}_k$ is an approximation of $g_\theta(x_k)$ through an inverse Fourier transform. Details of it is included in Section 4.1.3.

To avoid dividing a Dirac Delta function when approximating the score function $S_\theta(x) = g_\theta(x)/f_\theta(x)$, we only consider $x \in (x_0, x_K)$:

$$\tilde{S}_\theta(x) = \begin{cases} \tilde{g}_\theta(x)/\tilde{f}_\theta(x), & x_0 < x < x_K \\ 0, & \text{otherwise} \end{cases}. \tag{4.6}$$

We are thus estimating the sensitivity $\mathbb{E}_\theta[h(X)S_\theta(X)]$ by the following expectation effectively:

$$\mathbb{E}_\theta[h(\tilde{X})\tilde{S}_\theta(\tilde{X})] = \int_{x_0}^{x_K} h(x)\tilde{g}_\theta(x)\, dx, \tag{4.7}$$

where $\tilde{X}$ is generated from $\tilde{F}_\theta(x)$.

In summary, we tabulate the approximated values of $F_\theta(x)$ and $g_\theta(x)$ in the following way:

$$\begin{pmatrix} x_0 & x_1 & \cdots & x_K \\ \tilde{F}_0 & \tilde{F}_1 & \cdots & \tilde{F}_K \\ \tilde{g}_0 & \tilde{g}_1 & \cdots & \tilde{g}_K \end{pmatrix}. \tag{4.8}$$

We use the second row to construct $\tilde{F}_\theta(x)$ as in (4.3), generate $\tilde{X}$ from it, compute the corresponding $\tilde{f}_\theta(\tilde{x})$ and $\tilde{g}_\theta(\tilde{x})$ by (4.4) and (4.5), and the score function $\tilde{S}_\theta(\tilde{x})$ by (4.6).

In the numerical implementation, we repeat this process for N times, generate $\tilde{X}^n, n = 1, 2, \cdots, N$, and estimate the sensitivity by

$$\frac{1}{N}\sum_{n=1}^{N} h(\tilde{X}^n)\tilde{S}_\theta(\tilde{X}^n). \tag{4.9}$$

- **The multi-dimensional case**

For the multi-dimensional case, since $X_1, \cdots, X_d$ are independent random variables, they should be tabulated separately with distinct $K_i, x_0^i, x_{K_i}^i, \eta_i, \tilde{F}_k^i, \tilde{g}_k^i, 0 \leq k \leq K_i, 1 \leq i \leq d$. Similarly, $\tilde{F}_\theta^i(x_i)$, $\tilde{f}_\theta^i(x_i)$, and $\tilde{g}_\theta^i(x_i)$ can be approximated respectively by (4.3), (4.4), and (4.5). We then approximate the score function by:

$$\tilde{S}_\theta(x_1, \cdots, x_d) = \begin{cases} \sum_{i=1}^d \tilde{g}_\theta^i(x_i)/\tilde{f}_\theta^i(x_i), & x_0^i < x_i < x_{K_i}^i, 1 \leq i \leq d \\ 0, & \text{otherwise} \end{cases}. \tag{4.10}$$

And the sensitivity $\mathbb{E}_\theta[h(X_1,\cdots,X_d)S_\theta(X_1,\cdots,X_d)]$ is then estimated by

$$\begin{aligned} & \mathbb{E}_\theta[h(\tilde{X}_1,\cdots,\tilde{X}_d)\tilde{S}_\theta(\tilde{X}_1,\cdots,\tilde{X}_d)] \\ = \quad & \int_{x_0^1}^{x_{K_1}^1}\cdots\int_{x_0^d}^{x_{K_d}^d} h(x_1,\cdots,x_d)\Big(\sum_{j=1}^{d}\frac{\tilde{g}_\theta^j(x_j)}{\tilde{f}_\theta^j(x_j)}\Big)\prod_{i=1}^{d}\tilde{f}_\theta^i(x_i)\,dx_d\cdots dx_1. \end{aligned} \tag{4.11}$$

Each time for the i-th dimension, we simulate $\tilde{X}_i$ from $\tilde{F}_\theta^i(\tilde{x}_i)$ through the inverse transform method, obtain $\tilde{f}_\theta^i(\tilde{x}_i)$ and $\tilde{g}_\theta^i(\tilde{x}_i)$, and compute the corresponding score function by (4.10). By generating a sample of size $N, \{(\tilde{X}_1^n,\cdots,\tilde{X}_d^n), n = 1,2,\cdots,N\}$, we estimate the sensitivity by

$$\frac{1}{N}\sum_{n=1}^{N} h(\tilde{X}_1^n,\cdots,\tilde{X}_d^n)\tilde{S}_\theta(\tilde{X}_1^n,\cdots,\tilde{X}_d^n). \tag{4.12}$$

In Section 4.1.2 and 4.1.2, we described how to use the tabulated $\tilde{F}_k$ and $\tilde{g}_k$ values to generate samples and to estimate the sensitivity. In the following section, we will elaborate on how to obtain the tabulated values in (4.8) from characteristic functions of the distributions. In approximating (4.7) and (4.11) by (4.9) and (4.12), there are three major sources of bias: truncation of the distributions supports, the linear interpolation, and the approximation of $F_\theta(x_k)$ and $g_\theta(x_k), 0 \le k \le K$. Bias analysis is included in Section 4.1.4.

4.1.3 Inverting characteristic functions

Denote the characteristic function of X by $\phi_\theta(\xi)$ and assume that $\phi_\theta(\xi) \in L^1(\mathbb{R})$. Given $\phi_\theta(\xi)$, we can express $F_\theta(x)$ in terms of the Hilbert transform of $\phi_\theta(\xi)$, according to [26],

$$F_\theta(x) = \frac{1}{2} - \frac{i}{2}\mathcal{H}(e^{-i\xi x}\phi_\theta(\xi))(0),$$

which can be evaluated using a very simple scheme. We then have the following approximation for $F_\theta(x)$:

$$F_{\theta,h,M}(x) = \frac{1}{2} - \frac{i}{2}\sum_{m=-M}^{M} e^{-ix(m-1/2)h}\frac{\phi_\theta((m-1/2)h)}{(m-1/2)\pi}, h > 0, M \geq 1, \tag{4.13}$$

where h and M are respectively the discretization level and the truncation level. We obtain $\tilde{F}_k$ in the second row of (4.8) by letting $\tilde{F}_k = F_{\theta,h,M}(x_k), 0 \leq k \leq K$. According to [19], this approximation turns out to be remarkably accurate when the characteristic function belongs to a certain analytic class. The discretization error decays exponentially in terms of $1/h$ and admits explicit bounds.

Taking derivative of $F_\theta(x)$ with respect to θ and denote it by $\dot{F}_\theta(x)$, we have:

$$\dot{F}_\theta(x) = \frac{1}{2}\psi_\theta(0) - \frac{i}{2}\mathcal{H}(e^{-ix\xi}\psi_\theta(\xi))(0),$$

where $\psi_\theta(\xi) = d\phi_\theta(\xi)/d\theta$. And it can be approximated by

$$\dot{F}_{\theta,h,M}(x) = \frac{1}{2}\psi_\theta(0) + \frac{i}{2}\sum_{m=-M}^{M} e^{-ix(m-1/2)h}\frac{\psi_\theta((m-1/2)h}{(m-1/2)\pi}, h > 0, M \geq 1.$$

To construct (4.8), we need to approximate $g_\theta(x)$ as well. Since pdf is the inverse Fourier transform of characteristic function:

$$f_\theta(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\xi x}\phi_\theta(\xi)\, d\xi,$$

with approximation

$$f_{\theta,h,M}(x) = \frac{1}{2\pi}\sum_{m=-M}^{M} e^{-ixmh}\phi_\theta(mh)h, h > 0, M \geq 1.$$

Assume $\psi_\theta(\xi) \in L^1(\mathbb{R})$, $g_\theta(x)$ is then the inverse Fourier transform of $\psi_\theta(\xi)$:

$$g_\theta(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\xi x}\psi_\theta(\xi)\, d\xi, \tag{4.14}$$

and can be approximated by trapezoidal rule with discretization level h and truncation level M:

$$g_{\theta,h,M}(x) = \frac{1}{2\pi}\sum_{m=-M}^{M} e^{-ixmh}\psi_\theta(mh)h, h > 0, M \geq 1. \tag{4.15}$$

Then $\tilde{g}_k$ in the third row of (4.8) can be obtained by letting $\tilde{g}_k = g_{\theta,h,M}(x_k), 0 \leq k \leq K$. Interestingly, when ψ_θ is in the analytic class mentioned previously, the simple trapezoidal rule is highly accurate, with exponentially decaying discretization error in terms of $1/h$.

If ψ_θ is in this analytic class $H(\mathcal{D}_{(d_-,d_+)})$, we then have the following error bound for approximating $g_\theta(x)$ and $\dot{F}_\theta(x)$. The proof can be adapted from [26] straightforwardly and is omitted here.

Theorem 4.1. *Suppose $\psi_\theta \in L^1(\mathbb{R}) \cap H(\mathcal{D}_{(d_-,d_+)})$.*
If $|\psi_\theta(\xi)| \leq \kappa \exp(-c|\xi|^\nu)$ for any $\xi \in \mathbb{R}$ for some $\kappa, c, \nu > 0$ and $n \geq 0$, then for any $h > 0, M \geq 1$ such that $Mh \geq (\frac{n}{c\nu})^{1/\nu}$,

$$\begin{aligned}|g_\theta(x) - g_{\theta,h,M}(x)| \leq \quad & \frac{e^{-2\pi|d_-|/h+xd_-}}{2\pi(1-e^{-2\pi|d_-|/h})}||\psi_\theta||^- + \frac{e^{-2\pi d_+/h+xd_+}}{2\pi(1-e^{-2\pi d_+/h})}||\psi_\theta||^+ \\ & + \frac{\kappa}{\pi\nu c^{(n+1)/\nu}}\Gamma(\frac{n+1}{\nu}, c(Mh)^\nu),\end{aligned} \tag{4.16}$$

$$|\dot{F}_\theta(x) - \dot{F}_{\theta,h,M}(x)| \leq \quad \frac{e^{-2\pi|d_-|/h+xd_-}}{2\pi|d_-|(1-e^{-2\pi|d_-|/h})}||\psi_\theta||^- + \frac{e^{-2\pi d_+/h+xd_+}}{2\pi d_+(1-e^{-2\pi d_+/h})}||\psi_\theta||^+ + \frac{\kappa}{2\pi}\left(\frac{1}{M} + \frac{4}{\nu c(Mh)^\nu}\right)\exp(-c(Mh)^\nu). \tag{4.17}$$

If $|\psi_\theta(\xi)| \leq \kappa|\xi|^{-\nu-1}$ *for any* $\xi \in \mathbb{R}$ *for some* $\kappa, \nu > 0$*, then for any* $h > 0, M \geq 1$ *such that* $Mh \geq (\frac{n}{c\nu})^{1/\nu}$*,*

$$|g_\theta(x) - g_{\theta,h,M}(x)| \leq \frac{e^{-2\pi|d_-|/h+xd_-}}{2\pi(1-e^{-2\pi|d_-|/h})}||\psi_\theta||^- + \frac{e^{-2\pi d_+/h+xd_+}}{2\pi(1-e^{-2\pi d_+/h})}||\psi_\theta||^+ + \frac{\kappa}{\pi\nu}(Mh)^{-\nu}, \tag{4.18}$$

$$|\dot{F}_\theta(x) - \dot{F}_{\theta,h,M}(x)| \leq \quad \frac{e^{-2\pi|d_-|/h+xd_-}}{2\pi|d_-|(1-e^{-2\pi|d_-|/h})}||\psi_\theta||^- + \frac{e^{-2\pi d_+/h+xd_+}}{2\pi d_+(1-e^{-2\pi d_+/h})}||\psi_\theta||^+ + \frac{\kappa}{2\pi}\left(\frac{1}{M} + \frac{2}{\nu}\right)\frac{1}{(Mh)^\nu}. \tag{4.19}$$

The above results therefore show that when ψ_θ is in the analytic class defined above, the discretization error decays exponentially in $1/h$. Moreover, if ψ_θ has exponential tails, which is the case in many applications, the truncation error also decays exponentially in terms of Mh. It is thus not surprising that rather large h and small M often achieve remarkable accuracy in the above approximation. If ψ_θ has polynomial tails, then the discretization error in $1/h$ still allows one to take relatively large h. Consequently, one usually does not need a tremendous amount of terms in (4.15) to bound the truncation error. Finally, and the best of all, the total approximation error admits explicit bounds. They allow us to determine h and M for any given error tolerance level. This proves to be very useful in controlling the bias of the Monte Carlo estimation of the sensitivities.

Similarly, if ϕ_θ is in this analytic class, we can also find the error bound for the approximation of f_θ, as proved in [26].

Corollary 4.2. *Suppose* $\phi_\theta \in L^1(\mathbb{R}) \cap H(\mathcal{D}_{(d_-,d_+)})$*.*

If $|\phi_\theta(\xi)| \leq \kappa \exp(-c|\xi|^\nu)$ *for any* $\xi \in \mathbb{R}$ *for some* $\kappa, c, \nu > 0$ *and* $n \geq 0$*, then for any* $h > 0, M \geq 1$ *such that* $Mh \geq (\frac{n}{c\nu})^{1/\nu}$*,*

$$|f_\theta(x) - f_{\theta,h,M}(x)| \leq \frac{e^{-2\pi|d_-|/h+xd_-}}{2\pi(1-e^{-2\pi|d_-|/h})}||\phi_\theta||^- + \frac{e^{-2\pi d_+/h+xd_+}}{2\pi(1-e^{-2\pi d_+/h})}||\phi_\theta||^+ + \frac{\kappa}{\pi\nu c^{(n+1)/\nu}}\Gamma(\frac{n+1}{\nu}, c(Mh)^\nu). \tag{4.20}$$

If $|\phi_\theta(\xi)| \leq \kappa|\xi|^{-\nu-1}$ *for any* $\xi \in \mathbb{R}$ *for some* $\kappa, \nu > 0$*, then for any* $h > 0, M \geq 1$ *such that* $Mh \geq (\frac{n}{c\nu})^{1/\nu}$*,*

$$|f_\theta(x) - f_{\theta,h,M}(x)| \leq \frac{e^{-2\pi|d_-|/h+xd_-}}{2\pi(1-e^{-2\pi|d_-|/h})}||\phi_\theta||^- + \frac{e^{-2\pi d_+/h+xd_+}}{2\pi(1-e^{-2\pi d_+/h})}||\phi_\theta||^+ + \frac{\kappa}{\pi\nu}(Mh)^{-\nu}. \tag{4.21}$$

Even though we do not tabulate the approximated data for $f_\theta(x)$ or $\dot{F}_\theta(x)$, the derived estimation error bound will be used in the future proof.

4.1.4 Estimation Bias

As we mentioned in Section 4.1.2, truncation of the distribution supports, linear interpolation, and the errors in approximating $F_\theta(x_k)$ and $g_\theta(x_k), 0 \leq k \leq K$ are the three major sources of bias in sensitivity estimation. In this section, we will analyze the bias, and obtain explicit bounds for it in the one-dimensional cases. This allows us to conveniently determine the width and fineness of the grids in (4.8), as well as numerical parameters for characteristic functions inversion.

4.1.5 The one-dimensional case

First of all, let's assume that g_θ is differentiable with respect to x and introduce the following notations:

$$|\chi| = x_K - x_0,$$

$$||h||_\chi = \sup_{x\in\chi} |h(x)|,$$

$$||g''_\theta|| = \sup_{x\in\chi} |g''_\theta(x)|,$$

$$E^g_\chi = \max_{0\le k\le K} |g_\theta(x_k) - \tilde{g}_k|.$$

Taking derivatives on both sides of (4.14) for twice and assuming validity of changing the orders of integration and differentiation, we obtain

$$g''_\theta(x) = -\frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\xi x}\xi^2\psi_\theta(\xi)\,d\xi, \tag{4.22}$$

which gives an upper bound of $||g''_\theta||_\chi$:

$$||g''_\theta(x)||_\chi \le \frac{1}{2\pi}\int_{\mathbb{R}} |\xi^2\psi_\theta(\xi)|\,d\xi.$$

And we introduce the following notation regrading the error in approximating $\tilde{g}_k, 0 \le k \le K$ by $g_{\theta,h,M}(x_k)$:

$$E^g_{h,M,\chi} = \max_{0\le k\le K} |g_\theta(x_k) - g_{\theta,h,M}(x_k)|.$$

In the one-dimensional case, we have the following estimates for the bias:

Theorem 4.3. *Let $\mathbb{E}_\theta[h(X)S_\theta(X)]$ be defined as in (4.1) and $\mathbb{E}_\theta[h(\tilde{X})\tilde{S}_\theta(\tilde{X})]$ in (4.7).*

Assume that $g_\theta(x)$ is twice differentiable with respect to x. Then we have

$$
\begin{aligned}
& |\mathbb{E}_\theta[h(X)S_\theta(X)] - \mathbb{E}_\theta[h(\tilde{X})\tilde{S}_\theta(\tilde{X})]| && (4.23)\\
\leq \quad & \Big(\int_{-\infty}^{x_0} + \int_{x_K}^{\infty}\Big)|h(x)g_\theta(x)|\,dx \\
& + \frac{1}{2K^2}||g''_\theta||_\chi \cdot ||h||_\chi \cdot |\chi|^3 + ||h||_\chi \cdot E^g_\chi \cdot |\chi|. && (4.24)
\end{aligned}
$$

If $\psi_\theta \in H(\mathcal{D}_{(d_-,d_+)})$, and $\tilde{g}_k = g_{\theta,h,M}(x_k), 0 \leq k \leq K$, then

$$
\begin{aligned}
& |\mathbb{E}_\theta[h(X)S_\theta(X)] - \mathbb{E}_\theta[h(\tilde{X})\tilde{S}_\theta(\tilde{X})]| \\
\leq \quad & \frac{||\psi_\theta||^+}{2\pi}\int_{-\infty}^{x_0}|h(x)|e^{xd_+}\,dx + \frac{||\psi_\theta||^-}{2\pi}\int_{x_K}^{\infty}|h(x)|e^{xd_-}\,dx \\
& + \frac{1}{4\pi K^2}||h||_\chi \cdot |\chi|^3 \cdot \int_{\mathbb{R}}|\xi^2\psi_\theta(\xi)|\,d\xi \\
& + ||h||_\chi \cdot E^g_{h,M,\chi} \cdot |\chi|. && (4.25)
\end{aligned}
$$

Proof. The bias is given by

$$
\begin{aligned}
& \mathbb{E}_\theta[h(X)S_\theta(X)] - \mathbb{E}_\theta[h(\tilde{X})\tilde{S}_\theta(\tilde{X})] \\
= \quad & \int_{\mathbb{R}} h(x)g_\theta(x)\,dx - \int_{x_0}^{x_K} h(x)\tilde{g}_\theta(x)\,dx \\
= \quad & \sum_{k=1}^{K}\int_{x_{k-1}}^{x_k} h(x)(g_\theta(x) - \tilde{g}_\theta(x))\,dx + \Big(\int_{-\infty}^{x_0} + \int_{x_K}^{\infty}\Big)h(x)g_\theta(x)\,dx.
\end{aligned}
$$

Note that over $[x_{k-1}, x_k], \tilde{g}_\theta(x) = \tilde{g}_{k-1} + (\tilde{g}_k - \tilde{g}_{k-1})(x - x_{k-1})/\eta$. Denote $\bar{g}_\theta(x) = g_\theta(x_{k-1}) + (g_\theta(x_k) - g_\theta(x_{k-1}))(x - x_{k-1})/\eta$. That is, $\tilde{g}_\theta(x)$ is the linear interpolation formed using approximated values of $g_\theta(x)$, and $\bar{g}_\theta(x)$ is the one where exact values of $g_\theta(x)$ are used. Since $g_\theta(x_k)$ is differentiable, for any $x \in [x_{k-1}, x_k]$, there exists

$\xi_k(x) \in (x_{k-1}, x_k)$ such that $g_\theta(x) = g_\theta(x_{k-1}) + g'_\theta(\xi_k(x))(x - x_{k-1})$.

$$
\begin{aligned}
& \Big| \int_{x_{k-1}}^{x_k} h(x)(g_\theta(x) - \tilde{g}_\theta(x))\, dx \Big| \\
= \quad & \Big| \int_{x_{k-1}}^{x_k} h(x)(g_\theta(x) - \bar{g}_\theta(x) + \bar{g}_\theta(x) - \tilde{g}_\theta(x))\, dx \Big| \\
= \quad & \Big| \int_{x_{k-1}}^{x_k} h(x)\big(g'_\theta(\xi_k(x)) - \frac{g_\theta(x_k) - g_\theta(x_{k-1})}{\eta}\big)(x - x_{k-1})\, dx \\
& + \int_{x_{k-1}}^{x_k} h(x)\big(g_\theta(x_{k-1}) - \tilde{g}_{k-1} + (g_\theta(x_k) - \tilde{g}_k - (g_\theta(x_{k-1}) - \tilde{g}_{k-1}))\frac{x - x_{k-1}}{\eta}\big)\, dx \Big| \\
= \quad & \Big| \int_{x_{k-1}}^{x_k} h(x)(g'_\theta(\xi_k(x)) - g'_\theta(\xi_k(x_k)))(x - x_{k-1})\, dx \\
& + \int_{x_{k-1}}^{x_k} h(x)\big((g_\theta(x_{k-1}) - \tilde{g}_{k-1})\frac{x_k - x}{\eta} + (g_\theta(x_k) - \tilde{g}_k)\frac{x - x_{k-1}}{\eta}\big)\, dx \Big| \\
\leq \quad & \frac{1}{2}||g''_\theta||_\chi \cdot ||h||_\chi \cdot \eta^3 + ||h||_\chi \cdot E^g_\chi \cdot \eta.
\end{aligned}
$$

By summing K terms up, this leads to (4.23) immediately.

If $\psi_\theta \in H(\mathcal{D}_{(d_-,d_+)})$, by Cauchy's integral theorem, for any $\epsilon > 0$ such that $d_+ - \epsilon > 0$,

$$
\begin{aligned}
g_\theta(x) \quad &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi x} \psi_\theta(\xi)\, d\xi \\
&= \frac{1}{2\pi} e^{x(d_+ - \epsilon)} \int_{\mathbb{R}} e^{-i\xi x} \psi_\theta(\xi + i(d_+ - \epsilon))\, d\xi.
\end{aligned}
$$

Letting $\epsilon \to 0^+$, we obtain $|g_\theta(x)| \leq e^{xd_+}||\psi_\theta||^+/(2\pi), \forall x \in \mathbb{R}$.

Similarly, $|g_\theta(x)| \leq e^{xd_-}||\psi_\theta||^-/(2\pi), \forall x \in \mathbb{R}$. We therefore have

$$
\Big(\int_{-\infty}^{x_0} + \int_{x_K}^{\infty}\Big)|h(x)g_\theta(x)|\, dx \leq \frac{||\psi_\theta||^+}{2\pi} \int_{-\infty}^{x_0} |h(x)|e^{xd_+}\, dx + \frac{||\psi_\theta||^-}{2\pi} \int_{x_K}^{\infty} |h(x)|e^{xd_-}\, dx.
$$

Combining it with (4.23), we obtain the upper bound in (4.25). □

The theorem above is convenient to use in practice in several ways:

- $[x_0, x_K]$ is not required to cover the majority of the support of the distribution.

For example, if $h(x)$ is zero on the negative real line, one can simply take $x_0 = 0$ without introducing any bias.

- The explicit bound in (4.25) is very easy to use. One can choose x_0 according to the first term so that the bias introduced by x_0 is under a desired level. Similarly, the second term allows us to choose x_K. After we obtain x_0 and x_K, we determine K according to the third term, and h and M according to the fourth term and by Theorem 4.1.

- Theorem 4.3 can be used in more general ways for estimating sensitivities when the probability density does not admit an explicit expression. Depending on what information is available about the probability density, bounds for $|g_\theta|$, $|g''_\theta|$ and the values $\tilde{g}_k$ might be obtained in other ways. Then (4.23) can be used to determine $\chi = [x_0, x_K]$, the number of sub-intervals K, and the numerical parameters that affect the accuracy of $\tilde{g}_k$.

In Theorem 4.3, we do not add any constraints on the differentiability of function $h(x)$. Next theorem gives another upper bound in estimating the sensitivity while assuming that $h(x)$ differentiable almost everywhere in χ: $h'(x) = dh(x)/\,dx$. And we denote:

$$||h'||_\chi = ess \sup_{x\in\chi} |h'(x)|,$$

$$||g'_\theta||_\chi = ess \sup_{x\in\chi} |g'_\theta(x)|,$$

$$E_\chi^{\dot{F}} = \max_{0\le k\le K} |\dot{F}_\theta(x_k) - \tilde{\dot{F}}_k|,$$

$$E_{h,M,\chi}^{\dot{F}} = \max_{0\le k\le K} |\dot{F}_\theta(x_k) - \dot{F}_{h,M,\theta}(x_k)|.$$

Theorem 4.4. *Suppose $h(x)$ is differentiable in $\cup_{k=1}^{K}(x_{k-1}, x_k)$ except at n^h points and*

$g_\theta(x)$ is twice differentiable with respect to $x \in \chi$. Then

$$\begin{aligned} & |\mathbb{E}_\theta[h(X)S_\theta(X)] - \mathbb{E}_\theta[h(\tilde{X})\tilde{S}_\theta(\tilde{X})| \\ \leq \quad & \Big(\int_{-\infty}^{x_0} + \int_{x_K}^{\infty}\Big)|h(x)g_\theta(x)|\,dx \\ & + \frac{1}{2K^2}||h'||_\chi \cdot ||g'_\theta||_\chi \cdot |\chi|^3 \\ & + 2\big((n^h + K)\cdot ||h||_\chi + ||h'||_\chi \cdot |\chi|\big)E_\chi^{\dot{F}}. \end{aligned} \tag{4.26}$$

If $\psi_\theta \in H(\mathcal{D}_{(d_-,d_+)})$, and $\tilde{g}_k = g_{\theta,h,M}(x_k), 0 \leq k \leq K$, then

$$\begin{aligned} & |\mathbb{E}_\theta[h(X)S_\theta(X)] - \mathbb{E}_\theta[h(\tilde{X})\tilde{S}_\theta(\tilde{X})]| \\ \leq \quad & \frac{||\psi_\theta||^+}{2\pi}\int_{-\infty}^{x_0}|h(x)|e^{xd_+}\,dx + \frac{||\psi_\theta||^-}{2\pi}\int_{x_K}^{\infty}|h(x)|e^{xd_-}\,dx \\ & + \frac{|\chi|^3}{4\pi K^2}||f'||_\chi \int_{\mathbb{R}}|\xi\psi_\theta(\xi)|\,d\xi \qquad (4.27) \\ & + 2\big((n^h + K)\cdot ||h||_\chi + ||h'||_\chi \cdot |\chi|\big)E_{h,M,\chi}^{\dot{F}}. \qquad (4.28) \end{aligned}$$

Proof. Assuming that in $(x_{k-1}, x_k), 1 \leq k \leq K$, $h(x)$ is indifferentiable except at $n_k^h (\sum_{k=1}^K n_k^h = n^h)$ points, and following similar steps in the proof of Theorem 4.3, the estimation bias can be written as:

$$\begin{aligned} & \mathbb{E}_\theta[h(X)S_\theta(X)] - \mathbb{E}_\theta[h(\tilde{X})\tilde{S}_\theta(\tilde{X})] \\ = \quad & \sum_{k=1}^{K}\int_{x_{k-1}}^{x_k} h(x)(g_\theta(x) - \tilde{g}_\theta(x))\,dx + \Big(\int_{-\infty}^{x_0} + \int_{x_K}^{\infty}\Big)h(x)g_\theta(x)\,dx. \end{aligned}$$

Since $h(x)$ is differentiable in $(x_{k-1}, x_k), 1 \leq k \leq K$ except at n_k^h points, by [20], we have:

$$|\int_{x_{k-1}}^{x_k} h(x)(g_\theta(x)-\tilde{g}_\theta(x))\,dx| \le 2(n_k^h+1)\cdot ||h||_\chi \cdot E_\chi^{\dot{F}} + ||h'||_\chi \cdot \eta \cdot \int_{x_{k-1}}^{x_k} |g_\theta(x)-\tilde{g}_\theta(x)|\,dx,$$

where

$$\begin{aligned}
& \int_{x_{k-1}}^{x_k} |g_\theta(x)-\tilde{g}_\theta(x)|\,dx \\
= \quad & \int_{x_{k-1}}^{x_k} |g_\theta(x) - \frac{\dot{F}_\theta(x_k^-)-\dot{F}_\theta(x_{k-1}^+)}{\eta} + \frac{\dot{F}_\theta(x_k^-)-\dot{F}_\theta(x_{k-1}^+)}{\eta} - \frac{\tilde{\dot{F}}_\theta(x_k^-)-\tilde{\dot{F}}_\theta(x_{k-1}^+)}{\eta}|\,dx \\
= \quad & |\dot{F}_\theta(x_k^-)-\dot{F}_\theta(x_{k-1}^+)+\tilde{\dot{F}}_\theta(x_k^-)-\tilde{\dot{F}}_\theta(x_{k-1}^+)| + \int_{x_{k-1}}^{x_k} |g_\theta(x) - \frac{\dot{F}_\theta(x_k^-)-\dot{F}_\theta(x_{k-1}^+)}{\eta}|\,dx \\
\le \quad & 2E_\chi^{\dot{F}} + \frac{1}{\eta}\int_{x_{k-1}}^{x_k}\int_{x_{k-1}}^{x_k} |g_\theta(x)-g_\theta(y)|\,dy\,dx.
\end{aligned}$$

Due to the fact that $g_\theta(x)$ differentiable in (x_{k-1}, x_k), for any $x, y \in (x_{k-1}, x_k)$, there exists $\xi_k(x,y) \in (x_{k-1}, x_k)$, such that $g_\theta(y) = g_\theta(x) + g'_\theta(\xi_k(x,y))(y-x)$.

Therefore,

$$\begin{aligned}
& |\int_{x_{k-1}}^{x_k} h(x)(g_\theta(x)-\tilde{g}_\theta(x))\,dx| \\
\le \quad & 2\big((n_k^h+1)\cdot ||h||_\chi + ||h'||_\chi \cdot \eta\big)E_\chi^{\dot{F}} + \frac{\eta^3}{2}\cdot ||h'||_\chi \cdot ||q'_\theta||_\chi.
\end{aligned}$$

Naturally, we get the error bound in (4.26). □

Remark 4.5.

- *From Theorem 4.4, similarly, we obtain:*

$$
\begin{aligned}
& |\mathbb{E}_\theta[h(X)] - \mathbb{E}_\theta[h(\tilde{X})]| \\
\leq \quad & \Big(\int_{-\infty}^{x_0} + \int_{x_K}^{\infty}\Big)|h(x)f_\theta(x)|\,dx + |h(x_0)|F_\theta(x_0) + |h(x_K)|(1 - F_\theta(x_K)) \\
& + \frac{1}{2K^2}||h'||_\chi \cdot ||f'_\theta||_\chi \cdot |\chi|^3 \\
& + 2\big((n^h + K + 1)||h||_\chi + ||h'||_\chi \cdot |\chi|\big) \cdot E^F_\chi, \qquad (4.29)
\end{aligned}
$$

where f_θ as a function of x is differentiable in (x_0, x_K), and $h(x)$ is differentiable with respect to x in (x_0, x_K) except at up to n^h points, $||f'_\theta||_\chi = ess\sup_{x\in\chi}|f'_\theta(x)|$, and $E^F_\chi = \max_{0\leq k\leq K}|F_\theta(x_k) - \tilde{F}_k|$.

- *If $\phi_\theta \in H(\mathcal{D}_{(d_-,d_+)})$, and let $\tilde{f}_k = f_{\theta,h,M}(x_k), 0 \leq k \leq K$, then*

$$
\begin{aligned}
|\mathbb{E}_\theta[h(x)] - \mathbb{E}_\theta[h(\tilde{X})]| \leq \quad & \frac{||\phi_\theta||^+}{2\pi}\int_{-\infty}^{x_0}|h(x)|e^{xd_+}\,dx + \frac{||\phi_\theta||^-}{2\pi}\int_{x_K}^{\infty}|h(x)|e^{xd_-}\,dx \\
& + |h(x_0)|F_\theta(x_0) + |h(x_K)|(1 - F_\theta(x_K)) \\
& + \frac{|\chi|^3}{4\pi K^2}||h'||_\chi \cdot \int_{\mathbb{R}}|\xi\phi_\theta(\xi)|\,d\xi \\
& + 2\big((n^h + K + 1)||h||_\chi + ||h'||_\chi \cdot |\chi|\big) \cdot E^F_{h,M,\chi}, \qquad (4.30)
\end{aligned}
$$

where $E^F_{h,M,\chi} = \max_{0\leq k\leq K}|F_\theta(x_k) - F_{h,M,\theta}(x_k)|$.

The multi-dimensional case

For multi-dimensional cases, we examine the absolute difference between (4.12) and (4.11), and provide a theoretical upper bound for the estimation bias. In addition to the notations used earlier in this section, denote

$$
E^f_\chi = \max_{0\leq k\leq K}|f_\theta(x_k) - \tilde{f}_k|,
$$

$$E^f_{h,M,\chi} = \max_{0\le k\le K} |f_\theta(x_k) - f_{\theta,h,M}(x_k)|,$$

where we assume that $f_\theta(x)$ as a function of x is differentiable: $f'_\theta(x) = df_\theta(x)/\,dx$, and h is differentiable almost everywhere: $h'(x) = dh(x)/\,dx$. Approximation of f_θ at x_k can be obtained by similar method to that of F_θ and g_θ. The details are included in [26].

Theorem 4.6. *Let $X_i, 1 \le i \le d$, be d independent random variables with cumulative distribution functions F^i_θ, density functions f^i_θ, and $\dot{F}^i_\theta = \frac{\partial F^i_\theta}{\partial\theta}, g^i_\theta = \frac{\partial f^i_\theta}{\partial\theta}$. For $1 \le i \le d$, let $\chi_i = [x^i_0, x^i_{K_i}]$, with $x^i_k = x^i_0 + k\eta_i (0 \le k \le K_i)$, where $\eta_i = \frac{x^i_{K_i} - x^i_0}{K_i}$. $|h(x_1, x_2, \cdots, x_d)|$ is bounded by $h_1(x_1)h_2(x_2)\cdots h_d(x_d)$ for some $h_i \ge 0$, where $h_i(x_i)$ is differentiable in χ_i except at up to n^{h_i} points. Denote that $||h_i||_{\chi_i} = \sup_{x_i\in\chi_i} |h_i(x_i)|$, $||h'_i||_{\chi_i} = ess\sup_{x_i\in\chi_i} |h'_i(x_i)|$, and assume that they are finite for $1 \le i \le d$. Define*

$$\begin{aligned} B^f_i = \quad & \Big(\int_{-\infty}^{x^i_0} + \int_{x^i_{K_i}}^{\infty}\Big)|h_i(x)f^i_\theta(x)|\,dx + |h_i(x^i_0)|F^i_\theta(x^i_0) + |h_i(x^i_{K_i})|(1 - F^i_\theta(x^i_{K_i})) \\ & + \frac{1}{2K_i^2}||h'_i||_{\chi_i} \cdot ||f^{i\prime}_\theta||_{\chi_i} \cdot |\chi_i|^3 \\ & + 2\big((n^{h_i} + K_i + 1)||h_i||_{\chi_i} + ||h'_i||_{\chi_i} \cdot |\chi_i|\big) \cdot E^{F_i}_{\chi_i}, \end{aligned} \tag{4.31}$$

and define

$$\begin{aligned} B^g_i = \quad & \Big(\int_{-\infty}^{x^i_0} + \int_{x^i_{K_i}}^{\infty}\Big)|h_i(x)g^i_\theta(x)|\,dx + \frac{1}{2K_i^2}||h'_i||_{\chi_i} \cdot ||g^{i\prime}_\theta||_{\chi_i} \cdot |\chi_i|^3 \\ & + 2\big((n^{h_i} + K_i)||h_i||_{\chi_i} + ||h'_i||_{\chi_i} \cdot |\chi_i|\big) \cdot E^{\dot{F}_i}_{\chi_i}. \end{aligned} \tag{4.32}$$

Let $B = \max(B^f_1, B^f_2, \cdots, B^f_d, B^g_1, B^g_2, \cdots, B^g_d)$. Then there exist constants $a_i > 0, 0 \le$

$i \leq d-1$, independent of B such that

$$\begin{aligned} & |\mathbb{E}_\theta[h(X_1, \cdots, X_d) S_\theta(X_1, \cdots, X_d)] - \mathbb{E}_\theta[h(\tilde{X}_1, \cdots, \tilde{X}_d) \tilde{S}_\theta(\tilde{X}_1, \cdots, \tilde{X}_d)]| \\ \leq \quad & d \cdot B(a_0 + a_1 B + \cdots + a_{d-1} B^{d-1}), \end{aligned}$$

where $\tilde{S}_\theta(x_1, x_2, \cdots, x_d)$ is defined by $\sum_{i=1}^d \frac{\tilde{g}_\theta^i(x_i)}{\tilde{f}_\theta^i(x_i)}$.

Proof. Define $||f_i h_i||_1 = \int_{\mathbb{R}} h_i(x) f_\theta^i(x)\, dx$, and $||g_i h_i||_1 = \int_{\mathbb{R}} h_i(x) g_\theta^i(x)\, dx$. By Theorem 4.4 and Remark 4.5, we have the following for any $1 \leq i \leq d$:

$$|\int_{\mathbb{R}} h_i(x) \tilde{g}_\theta^i(x)\, dx| \leq ||g_i h_i||_1 + B_i^g,$$

$$|\int_{\mathbb{R}} h(x_1, \cdots, x_d)(\tilde{g}_\theta^i(x) - g_\theta^i(x))\, dx_i| \leq B_i^g \prod_{j=1, j\neq i}^d h_j(x_j),$$

$$|\int_{\mathbb{R}} h_i(x) \tilde{f}_\theta^i(x)\, dx| \leq ||f_i h_i||_1 + B_i^f,$$

$$|\int_{\mathbb{R}} h(x_1, \cdots, x_d)(\tilde{f}_\theta^i(x) - f_\theta^i(x))\, dx_i| \leq B_i^f \prod_{j=1, j\neq i}^d h_j(x_j).$$

Define $||F_i h_i|| = \max(||f_i h_i||_1, ||g_i h_i||_1)$ and $||\tilde{F}_i h_i|| = \max(||\tilde{f}_i h_i||_1, ||\tilde{g}_i h_i||_1)$, it is easy to see that $||\tilde{F}_i h_i|| \leq ||F_i h_i|| + B$. Then we are ready to find the upper bound of

the error in estimating sensitivity for the multi-dimensional cases:

$$
\begin{aligned}
& |\mathbb{E}_\theta[h(X_1,\cdots,X_d)S_\theta(X_1,\cdots,X_d)] - \mathbb{E}_\theta[h(\tilde{X}_1,\cdots,\tilde{X}_d)\tilde{S}_\theta(\tilde{X}_1,\cdots,\tilde{X}_d)]| \\
= \quad & |\int_{\mathbb{R}} h(x_1,\cdots,x_d)\prod_{i=1}^{d} f_\theta^i(x_i)\sum_{j=1}^{d}\frac{g_\theta^j(x_j)}{f_\theta^j(x_j)}\,dx_1\cdots dx_d \\
& -\int_{\mathbb{R}} h(x_1,\cdots,x_d)\prod_{i=1}^{d} \tilde{f}_\theta^i(x_i)\sum_{j=1}^{d}\frac{\tilde{g}_\theta^j(x_j)}{\tilde{f}_\theta^j(x_j)}\,dx_1\cdots dx_d| \\
= \quad & |\int_{\mathbb{R}} h(x_1,\cdots,x_d)\prod_{i=1}^{d} f_\theta^i(x_i)\sum_{j=1}^{d}\frac{g_\theta^j(x_j)}{f_\theta^j(x_j)}\,dx_1\cdots dx_d \\
& -\int_{\mathbb{R}} h(x_1,\cdots,x_d)\big(\prod_{i=1}^{d-1} f_\theta^i(x_i)\tilde{f}_\theta^d(x_d)\big)\big(\sum_{j=1}^{d-1}\frac{g_\theta^j(x_j)}{f_\theta^j(x_j)}+\frac{\tilde{g}_\theta^d(x_d)}{\tilde{f}_\theta^d(x_d)}\big)\,dx_1\cdots dx_d \\
& +\int_{\mathbb{R}} h(x_1,\cdots,x_d)\big(\prod_{i=1}^{d-1} f_\theta^i(x_i)\tilde{f}_\theta^d(x_d)\big)\big(\sum_{j=1}^{d-1}\frac{g_\theta^j(x_j)}{f_\theta^j(x_j)}+\frac{\tilde{g}_\theta^d(x_d)}{\tilde{f}_\theta^d(x_d)}\big)\,dx_1\cdots dx_d \\
& -\int_{\mathbb{R}} h(x_1,\cdots,x_d)\big(\prod_{i=1}^{d-2} f_\theta^i(x_i)\prod_{i=d-1}^{d}\tilde{f}_\theta^i(x_i)\big)\big(\sum_{j=1}^{d-2}\frac{g_\theta^j(x_j)}{f_\theta^j(x_j)}+\sum_{j=d-1}^{d}\frac{\tilde{g}_\theta^j(x_j)}{\tilde{f}_\theta^j(x_j)}\big)\,dx_1\cdots dx_d \\
& +\int_{\mathbb{R}} h(x_1,\cdots,x_d)\big(\prod_{i=1}^{d-2} f_\theta^i(x_i)\prod_{i=d-1}^{d}\tilde{f}_\theta^i(x_i)\big)\big(\sum_{j=1}^{d-2}\frac{g_\theta^j(x_j)}{f_\theta^j(x_j)}+\sum_{j=d-1}^{d}\frac{\tilde{g}_\theta^j(x_j)}{\tilde{f}_\theta^j(x_j)}\big)\,dx_1\cdots dx_d \\
& \cdots \\
& -\int_{\mathbb{R}} h(x_1,\cdots,x_d)\prod_{i=1}^{d}\tilde{f}_\theta^i(x_i)\sum_{j=1}^{d}\frac{\tilde{g}_\theta^j(x_j)}{\tilde{f}_\theta^j(x_j)}\,dx_1\cdots dx_d| \\
\leq \quad & d\cdot B\prod_{j=1}^{d-1}||F_jh_j|| + d\cdot B\prod_{j=1}^{d-2}||F_jh_j||(||F_dg_d||+B)+\cdots+d\cdot B\prod_{j=2}^{d}(||F_jh_j||+B).
\end{aligned}
$$

Then $a_0, a_1, \cdots, a_{d-1}$ can be found sequentially. □

Theorem 4.6 guarantees that, as B decreases, $\mathbb{E}_\theta[h(\tilde{X}_1,\cdots,\tilde{X}_d)\tilde{S}_\theta(\tilde{X}_1,\cdots,\tilde{X}_d)]$ converges to $\mathbb{E}_\theta[h(X_1,\cdots,X_d)S_\theta(X_1,\cdots,X_d)]$. In practice, we can adjust x_0^i, $x_{K_i}^i$, K_i, and other parameters to decrease B sequentially. And by doing this for several times, we are able to control the bias to our desired tolerance level.

In this theorem, we did not use the stationary increment property for Lévy process.

If we further assume so, then the distribution of $X_1, X_2, \cdots, X_d$ would be the same and so does $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_d$. In that case, the computational effort could be greatly reduced since all the grid points and probability approximation schemes will be the same for all the d dimensions.

4.2 Numerical Experiments

In this section, we illustrate the likelihood ratio method by simulating European vanilla call option and Asian call option deltas under the CGMY model. The European option delta represents a one-dimensional case, while the Asian option delta is a multi-dimensional case.

European vanilla option and Asian option pricing

The asset price is modeled by a geometric Lévy process: $S(t) = e^{X(t)}$, where $X(t) = \ln(S(t))$ is a CGMY model process with $X(0) = \ln(S_0)$. S_0 is the initial asset price. The price of a European call option is given by:

$$V_{S_0} = e^{-rT}\mathbb{E}_{S_0}[\max(0, S(T) - \mathcal{K})] = \mathbb{E}_{S_0}[e^{-rT}\max(0, e^{X(T)} - \mathcal{K})] = \mathbb{E}_{S_0}[h(X(T))],$$

where T is the time-to-maturity, $\mathcal{K}$ is the strike price, r is the risk-free interest rate, and $h(x) = e^{-rT}\max(0, e^x - \mathcal{K})$.

The payoff of an Asian call option with d monitoring periods is $\max(0, A_T^d - \mathcal{K}) = \max(0, \frac{1}{d}\sum_{k=1}^{d} S_{k\delta} - \mathcal{K})$, where δ is the time interval between two consecutive monitoring periods: $\delta = T/d$. Then the price of an Asian call option is:

$$V_{S_0} = e^{-rT}\mathbb{E}_{S_0}[\max(0, A_T^d - \mathcal{K})] = \mathbb{E}_{S_0}[h(X_1, X_2, \cdots, X_d)],$$

where $X_1 = \ln(S_\delta), X_i = \ln(S_{i\delta}/S_{(i-1)\delta}), 2 \leq i \leq d$ are independent Lévy increments, and

$$h(x_1, x_2, \cdots, x_d) = e^{-rT} \max\big(0, \frac{1}{d}(e^{x_1} + e^{x_1+x_2} + \cdots + e^{x_1+x_2+\cdots+x_d}) - \mathcal{K}\big).$$

Hence we have the following inequality:

$$|h(x_1, x_2, \cdots, x_d)| \leq e^{|x_1|}e^{|x_2|}\cdots e^{|x_d|}, 1 \leq i \leq d.$$

Thus $h_i(x_i) = e^{|x_i|}, 1 \leq i \leq d$. Since there is at most one point where $h_i(x_i)$ is not differentiable at, $n^{h_i} = 1$.

The CGMY process

A CGMY process, X_t, is a pure jump Lévy process with drift μ and the following Lévy density:

$$\frac{Ce^{Gx}}{|x|^{1+Y}}1_{x<0} + \frac{Ce^{-\mathbb{M}x}}{|x|^{1+Y}}1_{x>0},$$

for some $C > 0, G > 0, \mathbb{M} > 0, 0 < Y < 2$. By the martingale condition, $\mu = r - q - C\Gamma(-Y)((\mathbb{M}-1)^Y - \mathbb{M}^Y + (G+1)^Y - G^Y)$. Even though we do not have access to explicit expression for pdf and cdf of X_t, the characteristic function of X_t is know explicitly:

$$\phi_t(\xi) = \exp(i\mu t\xi - tC\Gamma(-Y)(\mathbb{M}^Y - (\mathbb{M} - i\xi)^Y + G^Y - (G + i\xi)^Y)),$$

where $\Gamma(\cdot)$ is the gamma function. Moreover, $\phi_t \in H(\mathcal{D}_{(d_-,d_+)})$ for any $-\mathbb{M} \leq d_- < 0 < d_+ \leq G$. When $0 < Y < 1$, we have $|\phi_t(\xi)| \leq \kappa_t e^{-c_t|\xi|^\nu}$, with $\kappa_t = \exp(-tC\Gamma(-Y)(\mathbb{M}^Y + G^Y)), c_t = 2tC|\Gamma(-Y)\cos(\pi Y/2)|$, and $\nu = Y$.

4.2.1 Delta of European call options under the CGMY model

Delta, as an option Greek, is defined as the partial derivative of option price with respect to the underlying asset price.

- **European vanilla call option delta**

  To estimate the delta of a European call option, we first need to construct (4.8) for $X(T) = \ln(S(T))$ by its characteristic function:

$$\tilde{\phi}_{S_0}(\xi) = \exp\big(i(\mu T + \ln(S_0))\xi - TC\Gamma(-Y)(\mathbb{M}^Y - (\mathbb{M} - i\xi)^Y + G^Y - (G + i\xi)^Y)\big).$$

  And the corresponding $\tilde{\psi}_{S_0}(\xi)$ is given by:

$$\tilde{\psi}_{S_0}(\xi) = \frac{i\xi}{S_0}\tilde{\phi}_{S_0}(\xi).$$

  Both $\tilde{\phi}_{S_0}(\xi)$ and $\tilde{\psi}_{S_0}(\xi)$ are in $H(\mathcal{D}_{(d_-,d_+)})$, and satisfy

$$|\tilde{\phi}_{S_0}(\xi)| \leq \kappa_T e^{-c_T|\xi|^\nu}, \quad |\tilde{\psi}_{S_0}(\xi)| \leq \frac{\kappa_T}{S_0}|\xi|e^{-c_T|\xi|^\nu}, \quad \forall \xi \in \mathbb{R}.$$

  The parameters are chosen based on Theorem 4.4 and by constraining the right hand-side of (4.25) to be less than or equal to the desired bias tolerance level ϵ_b. In terms of a European call option, we choose $x_0 = \ln(\mathcal{K})$ since $h(x) = 0$ for $x \in (-\infty, x_0)$. Hence, the first term is 0. Then we determine $x_K > 0$ according to the second term:

$$\frac{||\tilde{\psi}_{S_0}||^-}{2\pi} e^{x_K d_-}\big(\frac{\mathcal{K}}{d_-} - \frac{1}{d_- + 1}e^{x_K}\big).$$

  We find the smallest x_K so that the above is bounded by $\epsilon_b/2$. We then select K

according to the third term and have it less than or equal to $\epsilon_b/2$:

$$K = \lceil \sqrt{\frac{(e^{x_K} - \mathcal{K})(x_K - x_0)^2}{2\pi\epsilon_b} \int_{\mathbb{R}} |\xi^2 \tilde{\psi}_{S_0}(\xi)| \, d\xi} \rceil.$$

The last step is to bound the last term by $0.01\epsilon_b$, which is negligible compared to ϵ_b. In this way, we find h and M accordingly. The current x_0, x_K, K, h, and M setting can be taken as initial values. By adjustments, we can find more appropriate values, which balance the terms in (4.23) better.

- **Arithmetic Asian call option delta**

For an Asian call option, we need to construct (4.8) along each dimension. Due to the fact that a CGMY process is a Lévy process, $X_2, X_3, \cdots, X_d$ have the same characteristic function, and hence share the common sets of grids. The characteristic function of X_1 is:

$$\phi^1_{S_0}(\xi) = \exp(i(\mu\delta + \ln(S_0))\xi - \delta C\Gamma(-Y)(\mathbb{M}^Y - (\mathbb{M} - i\xi)^Y + G^Y - (G + i\xi)^Y)), \tag{4.33}$$

while the characteristic functions of $X_i, 2 \leq i \leq d$ are identical:

$$\phi^2_{S_0}(\xi) = \exp(i\mu\delta\xi - \delta C\Gamma(-Y)(\mathbb{M}^Y - (\mathbb{M} - i\xi)^Y + G^Y - (G + i\xi)^Y)). \tag{4.34}$$

Since there is no S_0 in $\phi^2_{S_0}(\xi)$, $\psi^2_{S_0}(\xi) = 0$, and so are $\tilde{g}^2_k, 0 \leq k \leq K_2$. We then only need two rows in (4.8) for the last $d-1$ dimensions. And the score function

is reduced to:

$$\tilde{S}_\theta(x_1, \cdots, x_d) = \begin{cases} \tilde{g}^1_\theta(x_1)/\tilde{f}^1_\theta(x_1), & x^1_0 < x_1 < x^1_{K_1}, x^2_0 < x_i < x^2_{K_2}, 2 \le i \le d \\ 0, & \text{otherwise} \end{cases}. \tag{4.35}$$

By Theorem 4.6, we fix a desired bias level ϵ_b, and by solving the inequality:

$$d \cdot B\big(||F_2 g_2||^{d-1} + (||F_2 g_2|| + B)^{d-1} + (d-2)||F_1 g_1|| \cdot (||F_2 g_2|| + B)^{d-2}\big) \le \epsilon_b,$$

we find an appropriate value for B. And by adjusting $x^i_0, x^i_{K_i}, K_i$ and other parameter values, we have $B^f_i \le B, B^g_i \le B, 1 \le i \le d$.

4.2.2 Numerical results

The likelihood ratio method in estimating call option deltas is implemented using C++ on a MacBook Pro with Intel Core i5 2.6GHz CPU and 8GB RAM.

For the CGMY process, we use the following parameter setting:

$$C = 4, G = 50, \mathbb{M} = 60, Y = 0.7, r = 0.05, q = 0.02, T = 0.5, S_0 = \mathcal{K} = 1.$$

- **Numerical results for European vanilla call option**

  Since $f(x) = 0$, for $x \le \ln(\mathcal{K})$, we let $x_0 = \ln(\mathcal{K}) = 0$. In Table 4.1, we set the tolerance level of total estimation bias to be $\epsilon_b = 10^{-3}$, and obtain corresponding values for x_0, x_K, K, h, and M in order:

European Call Option Delta in the CGMY model: $\epsilon_b = 10^{-3}$				
x_0	x_K	K	h	M
0	0.66	46	3.64	15

Table 4.1: Parameter Settings for European Call Option Delta in the CGMY model

Using the numbers obtained in Table 4.1, we construct (4.8) and generate Monte Carlo estimates of the European call option delta. Results are reported in Table 4.2:

The exact value of 0.566793 reported in the table is computed through finite

European Call Option Delta in the CGMY Model ($\epsilon_b = 10^{-3}$): Exact Value = 0.566793				
N (* 1000)	Delta	Abs Error	Std Error	CPU (s)
1	0.50552	6.13E-02	3.77E-02	0.0008
4	0.57665	9.85E-03	2.15E-02	0.0023
16	0.57206	5.27E-03	1.04E-02	0.0247
64	0.56238	4.41E-03	5.12E-03	0.0392
256	0.56547	1.33E-03	2.59E-03	0.1542
1024	0.56664	1.58E-04	1.29E-03	0.5749
4096	0.56733	5.34E-04	6.48E-04	2.3927
16384	0.56623	5.62E-04	3.24E-04	9.5731

Table 4.2: European call option delta in the CGMY model

difference method. The first column contains the sample sizes used in the simulation. The second column includes the European call option delta estimated by the likelihood ratio method. The 'Std Error' column and the 'CPU' column report the stand errors and the computational time in seconds. The 'Abs Error' column shows that as the sample size increases, the absolute error approaches to the level of 10^{-4}. Therefore, with the current parameter setting, we achieve the goal of bounding the bias by $\epsilon_b = 0.001$.

- **Numerical results for arithmetic Asian call option**

For the arithmetic Asian call option, we consider monthly monitoring, with $d = 6$. The target tolerance level is $\epsilon_b = 10^{-4}$. Parameter values are included in Table

4.3:

Asian Call Option Delta in the CGMY model: $\epsilon_b = 10^{-4}$				
x_0^1	$x_{K_1}^1$	K_1	h_1	M_1
-0.26	0.22	47	2.062	28
x_0^2	$x_{K_2}^2$	K_2	h_2	M_2
-0.19	0.16	44	1.898	18

Table 4.3: Parameter Setting for Asian Call Option Delta in the CGMY model

Monte Carlo estimates for the Asian call option delta are included in the following table. The exact value is estimated by the finite difference method, with the option prices at different levels of S_0 computed through the Fourier Transform method. To reduce simulation variance, we apply Control Variate method, by using the geometric Asian call option delta as a control variate. The absolute error in the third column decreases by half till bounded by 1.00E-04.

Asian Call Option Delta in the CGMY Model ($\epsilon_b = 10^{-4}$): Exact Value = 0.547941				
N (* 1000)	Delta	Abs Error	Std Error	CPU (s)
1	0.54731	6.35E-04	7.27E-04	0.0058
4	0.54763	3.13E-04	1.43E-04	0.0205
16	0.54780	1.43E-04	1.64E-04	0.0840
64	0.54783	1.08E-04	8.33E-05	0.3420
256	0.54786	8.19E-05	4.09E-05	1.3995
1024	0.54784	9.97E-05	2.04E-05	5.6366
4096	0.54786	8.55E-05	1.02E-05	22.2483
16384	0.54787	6.70E-05	5.11E-06	89.5612

Table 4.4: Asian call option delta in the CGMY model

Chapter 5

Simulating from Characteristic Functions by Acceptance-Rejection

5.1 Acceptance-Rejection Method for Univariate Cases

Steps for acceptance-rejection method for univariate cases are summarized as follows:

- **Step 1.** Find a (easy-to-simulate) probability distribution $G(x)$, with probability density function $g(x)$, such that $\sup_x f(x)/g(x) \leq c$, for some $c \geq 1$.
- **Step 2.** Generate $X \sim G(x)$, and compute $g(x)$.
- **Step 3.** Compute $f(x)$.
- **Step 4.** Generate $U \sim U[0,1]$: if $u \leq f(x)/(cg(x))$, accept x; if $u > f(x)/(cg(x))$, reject x and go back to Step 2.

Instead of simulating from a complicated distribution $f(x)$, acceptance-rejection method generates samples from $g(x)$. When the closed-form $f(x)$ is not accessible, one can approximate it by inverting characteristic function and applying the trapezoidal

rule. The probability of accepting generated samples, in other words, acceptance rate, is $1/c$. To have a higher acceptance rate, one needs to find an appropriate distribution $g(x)$ such that c is minimized. Therefore, selecting a suitable $cg(x)$ is significant in the implementation of the acceptance-rejection method.

5.1.1 Devroye's 2nd-order method

Devroye [22] proposed one way to select $cg(x)$. In the thesis, a 2nd-order polynomial function with a constant bound is used to cover the underlying pdf, $f(x)$:

$$cg(x) = \min(\tilde{c}, \frac{\tilde{k}}{x^2}),$$

where

$$\tilde{c} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\phi(\xi)| \, d\xi \approx \frac{1}{2\pi} \sum_{m=-M}^{M} |\phi(mh)|h,$$

and

$$\tilde{k} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\phi''(\xi)| \, d\xi \approx \frac{1}{2\pi} \sum_{m=-M}^{M} |\phi''(mh)|h.$$

With analyticity in a horizontal strip and exponential tail, $\tilde{c}$ and $\tilde{k}$ are found to be finite. As we proved in previous section, this approximation has an error exponentially decaying.

One disadvantages of Devroye's method is that for every generation, without closed-form pdf, one needs to approximate $f(x)$ by trapezoidal rule, which is time consuming.

To have better performance, we propose several improvements on the Devroye's 2nd-order method.

5.1.2 Improvements on Devroye's 2nd-order Method

- **Devroye's 4th-order Method**

When the magnitude of x gets larger and larger, the 2nd-order polynomial decays much slower than polynomials of higher orders. So we first of all consider covering $f(x)$ with polynoimals of higher orders. $cg(x)$ we select is defined to be:

$$cg(x) = \min(\tilde{c}, \frac{\tilde{k}}{x^2}, \frac{\tilde{\theta}}{x^4}),$$

where

$$\tilde{\theta} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\phi^{(4)}(\xi)| \, d\xi \approx \frac{1}{2\pi} \sum_{m=-M}^{M} |\phi^{(4)}(mh)| h.$$

For every generation, $f(x)$ is computed by trapezoidal rule.

- **Devroye's 2nd-order/4th-order Method with Tabulation**

  Tabulating $f(x)$ on the simulation grid does not improve its acceptance rate, but will theoretically decrease the simulation time. And $f(x)$ is approximated by linear interpolation:

$$f(x) = \frac{x - x_{k-1}}{x_k - x_{k-1}} f(x_k) + \frac{x_k - x}{x_k - x_{k-1}} f(x_{k-1}),$$

  for $x_{k-1} \leq x \leq x_k$, $1 \leq k \leq K$.

- **Coverage with Exponential Functions**

  Due to the fact that the characteristic function is analytic in a horizontal strip including the real axis, its pdf is exponentially decaying on both sides. The lower boundary of the analyticity strip, d_-, determines the decay rate on the positive side, while the upper boundary, d_+, determines the decay rate on the negative side. Based on this property, one can find a two-sided exponential function to cover $f(x)$. And this exponential function does not have to be symmetric since the decay rate on both sides can be different. The structure of $cg(x)$ is then:

$$cg(x) = \begin{cases} \theta_l \exp(-b_l(\hat{x} - x)), & x \leq \hat{x} \\ \theta_r \exp(-b_r(x - \hat{x})), & x > \hat{x} \end{cases}.$$

$\hat{x}$ is the mode of the pdf, $f(x)$.

- **Coverage with Exponential Functions and a Constant Bound**

  By adding a constant bound to the exponential function, $cg(x)$ now becomes:

$$cg(x) = \begin{cases} \min(\theta_l \exp(-b_l(\hat{x} - x)), \tilde{c}), & x \leq \hat{x} \\ \min(\theta_r \exp(-b_r(x - \hat{x})), \tilde{c}), & x > \hat{x} \end{cases}.$$

5.1.3 Numerical Experiment: Acceptance-rejection for Univariate Cases

In this section, we exhibit the performance of Devroye's 2nd-order method, together with our improvements numerically by pricing a future under Double Exponential Jump (Kou's) model.

The asset price is: $S_t = S_0 \exp(X_t)$, where

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i, \quad N_t \sim Poission(\lambda t).$$

The characteristic function of X_t is:

$$\phi(\xi) = \exp\big(-\frac{1}{2}\xi^2\sigma^2 t + i\xi\gamma t + i\xi\gamma t(\frac{p}{\lambda_+ - i\xi} - \frac{1-p}{\lambda_- + i\xi})\big).$$

One reasonable parameter setting for the Kou's model is:

$$r = 0.02, T = 1.00, \sigma = 0.2, \lambda = 7.0, p = 0.2, \lambda_+ = 80, \lambda_- = 10, S_0 = 1.0.$$

For tabulation and approximation, we choose:

$$h = 0.10, M = 1000, \delta_x = 0.002,$$

where δ_x is the tabulation grid size.

By martingale property, the future's theoretical price is computed to be $\mathbb{E}[S_0 e^{rT}] = 1.020200$. In this following table, for each simulation method, we record estimated future price, absolute error, standard error, acceptance rate, and running time for $1,000$ generations.

Acceptance-rejection Methods Comparison ($N = 1,000$)					
Method	Est. Future Price	Abs Error	Std Error	Acceptance Rate	Running Time
Devroye	1.028350	0.008148	0.011828	0.590319	1.886870
Devroye (4th)	1.024150	0.003952	0.011173	0.711744	1.537540
Devroye (tab)	1.026722	0.004522	0.012647	0.599168	0.031286
Devroye (4th, tab)	1.025178	0.003902	0.011624	0.715672	0.026419
Asymm. Exp Dis	1.010780	0.009417	0.012062	0.751880	0.023977
Asymm. Exp Dis with Const.	1.045290	0.025089	0.011914	0.854701	0.030887

Table 5.1: Acceptance-rejection Methods in Pricing Future under Kou's Model: Exact Value = 1.020200

The traditional Devroye's 2nd-order method finishes the simulation within 1.89s, with acceptance rate 59%. The acceptance rate increases to 71% if the 4th-order polynomial is incorporated. The running time is then decreased a little by increased acceptance rate. If we use tabulation for the 2nd-order and the 4th-order methods, acceptance rate does not change much. But the running time decreases to only 2% of the previous level, which is a great improvement. By covering the underlying pdf

with asymmetric exponential functions, the acceptance rate increases to 75%. And if we incorporate a constant bound, it reaches the level 85%, with running time 0.03s.

The comparison between the traditional Devroye's 2nd-order method and our finest model shows that the acceptance-rejection method does improve the simulation efficiency for univariate cases, which is also demonstrated by the following two plots.

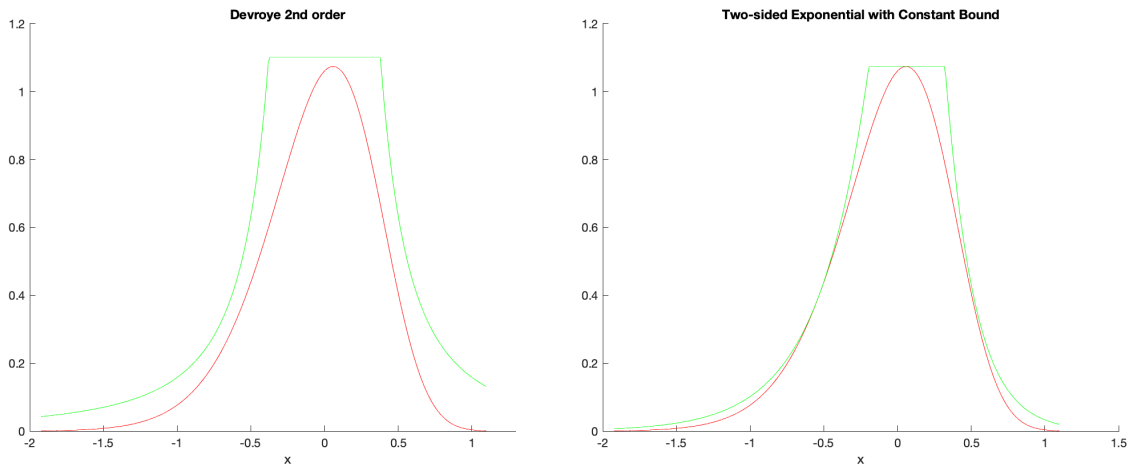

Figure 5.1: Devroye 2nd Order Method v.s. Exponential Function with Constant Bound

The one on the left-hand-side shows how the underlying pdf is covered by $cg(x)$ in Devroye's 2nd-order method, while the one on the right-hand-side exhibits the coverage by the exponential function with a constant bound. As the plot shows, the second coverage is more efficient, with the green and the red curves closer to each other, which is reflected by higher acceptance rate by the last method.

5.2 Acceptance-Rejection Method for Bivariate Cases

The univariate numerical experiment has shown that the exponential coverage is efficient as an improvement. In this section, we investigate if it is applicable for bivariate models.

To begin with, we list the steps for the acceptance-rejection method for bivariate cases:

- **Step 1.** Optimize parameters a_k, b_k, θ_k, $\hat{x}_1$ and $\hat{x}_2$ in:

$$cg(x_1, x_2) = \theta_k \exp\left(-a_k|x_1 - \hat{x}_1| - b_k|x_2 - \hat{x}_2|\right),$$

  $k = 1, 2, 3, 4$ and correspond to four quadrants, such that:

$$f(x_1, x_2) \leq cg(x_1, x_2).$$

- **Step 2.** Tabulate $f(x)$ for approximation and simulation: $x_1 = \hat{x}_1 + n\delta_x^1$, and $x_2 = \hat{x}_2 + n\delta_x^2$, $n = -N, -N+1, \cdots, N$. Calculate $\tilde{f}(x_1^i, x_2^j)$, $i, j = -N, -N+1, \cdots, N$ by inverting characteristic function and trapezoidal rule.

- **Step 3.** Generate $X = (X_1, X_2)$ from $G(x_1, x_2)$, and compute $cg(x_1, x_2)$.

- **Step 4.** Approximate $f(x_1, x_2)$ by linear interpolation:

$$\begin{aligned} & \tilde{f}(x_1, x_2) \\ = \; & \frac{1}{\delta_{X_1}\delta_{X_2}}\big(\tilde{f}_{i,j}(x_{i+1}^1 - x_1)(x_{j+1}^2 - x_2) + \tilde{f}_{i+1,j}(x_1 - x_i^1)(x_{j+1}^2 - x_2) \\ + \; & \tilde{f}_{i,j+1}(x_{i+1}^1 - x_1)(x_2 - x_j^2) + \tilde{f}_{i+1,j+1}(x_1 - x_i^1)(x_2 - x_j^2)\big). \end{aligned}$$

- **Step 5.** Generate $U \sim U[0, 1]$: if $u \leq \tilde{f}(x_1, x_2)/(cg(x_1, x_2))$, accept $x = (x_1, x_2)$; if $u > \tilde{f}(x_1, x_2)/(cg(x_1, x_2))$, reject $x = (x_1, x_2)$ and go back to Step 3.

The approximated pdf and the optimized coverage are displayed in the following two figures:

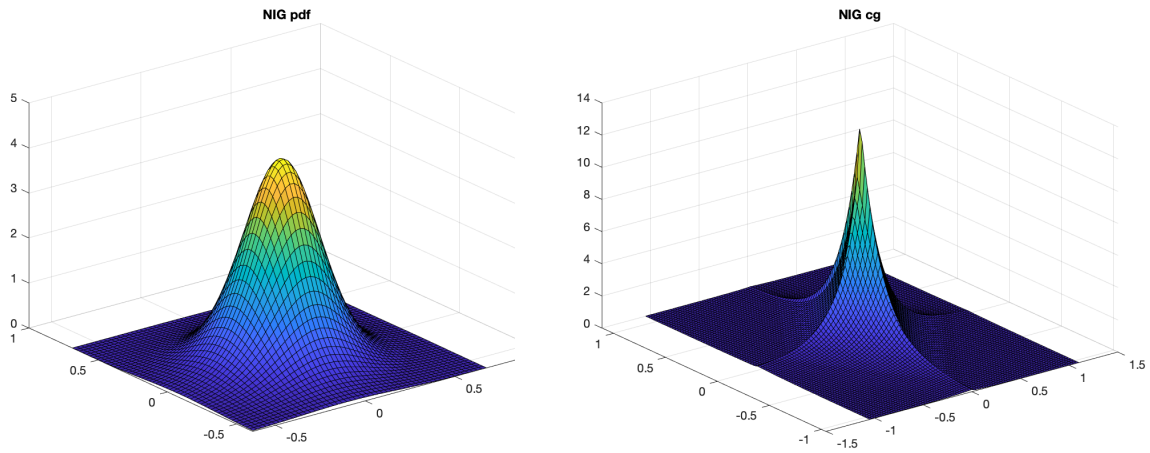

Figure 5.2: Bivariate NIG pdf and the coverage $cg(x_1, x_2)$

The 2d European call option pricing under the bivariate NIG model is again implemented to show the performance of the acceptance-rejection method. In the following table, we summarize the numerical results with sample sizes from 10 to $10,000,000$. We continue to use the previous parameter setting and the benchmark is borrowed from the last section.

Pricing 2d European Pricing by Acceptance-rejection					
Sample size	Est. Call Price	Abs Error	Std. Error	Accept Rate	Running time
10	0.102742	0.048791	0.034228	0.333333	0.0002
100	0.147931	0.003601	0.013377	0.531915	0.0005
1000	0.154931	0.003399	0.004904	0.466636	0.0049
10000	0.148058	0.003474	0.001523	0.476735	0.0469
100000	0.151334	0.000198	0.000486	0.475021	0.4593
1000000	0.151416	0.000117	0.000153	0.474889	4.5813
10000000	0.151278	0.000254	0.000048	0.474975	46.5381

Table 5.2: Acceptance-rejection Methods in Pricing 2d European Call under Bivaraite NIG Model: Exact Value = 0.151532235916, Theoretical Acceptance Rate = 0.4752

The real acceptance rate is close to theoretical acceptance rate, 0.4752.

Chapter 6

Error Probability Analysis of EGC Receivers from Analytic Characteristic Functions

6.1 SER approximation and error analysis

The conditional error probability (CEP) considered in this thesis is:

$$P_s(\epsilon|\gamma) = a \cdot \mathrm{erfc}(\sqrt{p}\beta\gamma) + b \cdot \mathrm{erfc}^2(\sqrt{p}\beta\gamma),$$

where $p > 0, \beta = \sqrt{\frac{E_b}{N_0 L_r}} > 0, a$ and b are constants, and $\mathrm{erfc}(\cdot)$ is the complementary error function. In an EGC system, the parameter, γ, with characteristic function $\phi_\gamma(\cdot)$, is defined as: $\gamma = \sqrt{Y}$, $Y = \left(\sum_{l=1}^{L_r} Y_l\right)^2$, where Y_l's are the fading amplitudes, and can be modeled as Rayleigh, Nakagami-m, or Nakagami-n random variables. L_r is the total number of paths combined in the receiver.

Denote the probability density function of γ as $p(\gamma)$. The average bit or symbol

error rate (SER) is obtained by taking expectation of $P_s(\epsilon|\gamma)$ over $p(\gamma)$:

$$\mathbb{E}[P_s(\epsilon|\gamma)] = \int_0^\infty P_s(\epsilon|\gamma)p(\gamma)\,d\gamma. \tag{6.1}$$

Remark 6.1. *The Bit Error Rate (BER) is a special case of SER. Under Coherent PSK cases,* $a = \frac{1}{2}, b = 0, p = 1$. *Under coherent FSK cases,* $a = \frac{1}{2}, b = 0, p = \frac{1}{2}$.

Remark 6.2. *The Square QAM is a special case of SER with* $q = 1 - 1/\sqrt{\tilde{M}}$, $p = 1.5\log_2 \tilde{M}/(\tilde{M} - 1)$, $a = 2q, b = -q^2$ *(*$\tilde{M}$ *is the modulation order and is of a power of 2).*

To evaluate SER, we rewrite the erfc$(\cdot)$ function in terms of the cumulative distribution function (cdf), $F(\cdot)$, of the standard normal distribution:

$$\text{erfc}(\gamma) = \frac{2}{\sqrt{\pi}} \int_\gamma^\infty e^{-t^2}\,dt = 2(1 - F(\sqrt{2}\gamma)).$$

By [26], $F(x)$ can be approximated efficiently by:

$$\tilde{F}_{h,M}(\gamma) = \frac{1}{2} - \frac{i}{2} \sum_{m=-M}^{M} e^{-i\beta(m-1/2)h} \phi((m-1/2)h) \frac{1}{(m-1/2)\pi}, \tag{6.2}$$

where $\phi(t) = \exp(-\frac{t^2}{2})$ is the characteristic function of the standard normal distribution, h and $U = Mh$ are respectively the discretization level and the truncation level of the approximation.

In this way, $P_s(\epsilon|\gamma)$ is approximated by:

$$\begin{aligned}
\tilde{P}_s^{h,M}(\epsilon|\gamma) &= 2a(1 - \tilde{F}_{h,M}(\sqrt{2p}\beta\gamma)) + 4b(1 - \tilde{F}_{h,M}(\sqrt{2p}\beta\gamma))^2 \\
&= (a+b) - i(a+2b) \sum_{m=-M}^{M} \frac{e^{-i\sqrt{2p}\beta\gamma(m-1/2)h}}{(m-1/2)\pi} \phi((m-1/2)h) \\
&- b \sum_{m=-M}^{M} \sum_{n=-M}^{M} \frac{e^{-i\sqrt{2p}\beta\gamma(m+n-1)h}}{(m-1/2)(n-1/2)\pi^2} \phi((m-1/2)h)\phi((n-1/2)h) \quad (6.3)
\end{aligned}$$

Plugging this into (6.1), and assuming the validity of interchanging the orders of integration and summations, we then approximate the SER by:

$$\begin{aligned}
& \mathbb{E}[\tilde{P}_s^{h,M}(\epsilon|\gamma)] \\
= \; & (a+b) - i(a+2b) \sum_{m=-M}^{M} \frac{\phi((m-1/2)h)}{(m-1/2)\pi} \phi_\gamma(-\sqrt{2p}\beta(m-1/2)h) \\
- \; & b \sum_{m=-M}^{M} \sum_{n=-M}^{M} \frac{\phi((m-1/2)h)\phi((n-1/2)h)}{(m-1/2)(n-1/2)\pi^2} \phi_\gamma(-\sqrt{2p}\beta(m+n-1)h). \quad (6.4)
\end{aligned}$$

If $L_r > 1$ and Y_l's are independent, $\phi_\gamma(\cdot) = \prod_{l=1}^{L_r} \phi_{Y_l}(\cdot)$.

6.1.1 SER approximation error analysis

One major reason of implementing this approximation method is that in many cases, closed-form expressions of the probability density functions are not available to us, but that of the characteristic functions are. Especially, when this characteristic function satisfies some analytic conditions, the total approximation error admits explicit bound, which decays exponentially.

Before the detailed analysis of the error bound, we give the definition of a class of analytic functions on a complex strip. For $-\infty < d_- < 0 < d_+ < +\infty$, $\mathcal{D}_{(d_-,d_+)} = \{z \in \mathbb{C} : \mathcal{I}(z) \in (d_-, d_+)\}$, where $\mathcal{I}(z)$ denotes the imaginary part of a variable z in the complex plane $\mathbb{C}$.

Definition 6.3. *A function f is in $H(\mathcal{D}_{(d_-,d_+)})$ if it is analytic in $\mathcal{D}_{(d_-,d_+)}$ and satisfies*

$$\int_{d_-}^{d_+} |f(x+iy)|\, dy \to 0, x \to \pm\infty, \quad (6.5)$$

Theorem 6.4. *Let γ be a continuous random variable with characteristic function*

$\phi_\gamma(\xi) \in H(\mathcal{D}_{(d_-,d_+)})$. *For any* $0 < d < \min(-\frac{d_-}{2\sqrt{2p}\beta}, \frac{d_+}{2\sqrt{2p}\beta})$, *there exists an error bound in approximating* $\mathbb{E}[P_s(\epsilon|\gamma)]$ *by* $\mathbb{E}[\tilde{P}_s^{h,M}(\epsilon|\gamma)]$*:*

$$\begin{aligned}
& |\mathbb{E}[\tilde{P}_s^{h,M}(\epsilon|\gamma)] - \mathbb{E}[P_s(\epsilon|\gamma)]| \\
\leq\ & \frac{2|b|e^{-4\pi d/h+d^2}}{\pi d^2(1-e^{-2\pi d/h})^2}\big(\phi_\gamma(2\sqrt{2p}\beta di) + \phi_\gamma(-2\sqrt{2p}\beta di) + 2\big) \\
+\ & (2|a| + 16|b| + 4|b|\tau_{Mh})\frac{e^{-2\pi d/h+d^2/2}}{\sqrt{2\pi}d(1-e^{-2\pi d/h})}\big(\phi_\gamma(\sqrt{2p}\beta di) + \phi_\gamma(-\sqrt{2p}\beta di)\big) \\
+\ & \tau_{Mh}(|b|\tau_{Mh} + |a| + 8|b|), \qquad (6.6)
\end{aligned}$$

where $\tau_{Mh} = \frac{1}{\pi M}e^{-\frac{1}{2}(Mh)^2} + \frac{1}{\pi}\Gamma(0, \frac{1}{2}(Mh)^2)$.

If we select h *by* $h = h(M) = (4\pi dM^{-2})^{1/3}$, *then there exists a constant* $C_1, C_2 > 0$ *independent of* M *such that:*

$$\begin{aligned}
& |\mathbb{E}[\tilde{P}_s^{h,M}(\epsilon|\gamma)] - \mathbb{E}[P_s(\epsilon|\gamma)]| \\
\leq\ & C_1|b|\exp(-(4\pi dM)^{2/3})\Big(e^{d^2}\big(\phi_\gamma(2\sqrt{2p}\beta di) + \phi_\gamma(-2\sqrt{2p}\beta di) + 2\big) \\
+\ & M^{-2/3}e^{d^2/2}\big(\phi_\gamma(\sqrt{2p}\beta di) + \phi_\gamma(-\sqrt{2p}\beta di)\big) + M^{-4/3}\Big) \\
+\ & C_2|a|\exp(-(\sqrt{2}\pi dM)^{2/3})\Big(e^{d^2/2}\big(\phi_\gamma(\sqrt{2p}\beta di) + \phi_\gamma(-\sqrt{2p}\beta di)\big) + M^{-2/3}\Big) \qquad (6.7)
\end{aligned}$$

Proof. Corollary 2.9 in [26] gives an upper bound for the absolute error in approximating $F(\sqrt{2p}\beta\gamma)$ by $\tilde{F}_{h,M}(\sqrt{2p}\beta\gamma)$:

$$\begin{aligned}
E_{h,M}^F(\sqrt{2p}\beta\gamma) &:= |\tilde{F}_{h,M}(\sqrt{2p}\beta\gamma) - F(\sqrt{2p}\beta\gamma)| \\
&\leq \frac{e^{-2\pi d/h}}{2\pi d(1-e^{-2\pi d/h})}(e^{\sqrt{2p}\beta d\gamma}||\phi||^+ + e^{-\sqrt{2p}\beta d\gamma}||\phi||^-) + \frac{1}{2}\tau_{Mh} \\
&= \frac{e^{-2\pi d/h+d^2/2}}{\sqrt{2\pi}d(1-e^{-2\pi d/h})}(e^{\sqrt{2p}\beta d\gamma} + e^{-\sqrt{2p}\beta d\gamma}) + \frac{1}{2}\tau_{Mh}. \qquad (6.8)
\end{aligned}$$

d can be selected to be any positive finite real value since $\phi(t) = \exp(-t^2/2)$ is decaying exponentially at the rate that is higher than any linear order.

Denoting $\frac{e^{-2\pi d/h + d^2/2}}{\sqrt{2\pi}d(1-e^{-2\pi d/h})}$ by θ_h, the corresponding approximation error of $\tilde{P}_s^{h,M}(\epsilon|\gamma)$ is bounded by:

$$
\begin{aligned}
& E_{h,M}^{P_s}(\gamma) \\
:= \; & |\tilde{P}_s^{h,M}(\epsilon|\gamma) - P_s(\epsilon|\gamma)| \\
= \; & |(\tilde{F}_{h,M}(\sqrt{2p}\beta\gamma) - F(\sqrt{2p}\beta\gamma))\big(-2a + 4b\big(2 + (\tilde{F}_{h,M}(\sqrt{2p}\beta\gamma) + F(\sqrt{2p}\beta\gamma))\big)\big)| \\
\leq \; & E_{h,M}^{F}(\sqrt{2p}\beta\gamma)\big(2|a| + 16|b| + 4|b|E_{h,M}^{F}(\sqrt{2p}\beta\gamma)\big) \\
\leq \; & 4|b|\theta_h^2(e^{\sqrt{2p}\beta d\gamma} + e^{-\sqrt{2p}\beta d\gamma})^2 + \theta_h(2|a| + 16|b| + 4|b|\tau_{Mh})(e^{\sqrt{2p}\beta d\gamma} + e^{-\sqrt{2p}\beta d\gamma}) \\
+ \; & \tau_{Mh}(|a| + 8|b| + |b|\tau_{Mh}). \qquad (6.9)
\end{aligned}
$$

Taking integration of this over $p(\gamma)$, we then get an upper bound for the approximation error of $\mathbb{E}[\tilde{P}_s^{h,M}(\epsilon|\gamma)]$ in (6.6). $\phi_\gamma(\xi) \in H(\mathcal{D}_{(d_-,d_+)})$ requires that $0 < d < \min(-\frac{d_-}{2\sqrt{2p}\beta}, \frac{d_+}{2\sqrt{2p}\beta})$.

The incomplete gamma function $\Gamma(0, \frac{1}{2}(Mh)^2) \sim \frac{1}{(Mh)^2}e^{-\frac{1}{2}(Mh)^2}$ for Mh large enough. To have the terms in (6.6) decay at the same rate, we set: $2\pi d/h = \frac{1}{2}(Mh)^2$. Therefore, selecting $h = h(M) = (4\pi d M^{-2})^{1/3}$, the error bound is simplified as in (6.7). □

Corollary 6.5. *Let γ be a continuous random variable with pdf $p(\gamma)$. If there exist $-\infty < d_- < 0 < d_+ < +\infty$ such that:*

$$
p(\gamma) = \begin{cases} O(e^{(d_- + \epsilon)\gamma}), & \gamma \to +\infty \\ O(e^{(d_+ - \epsilon)\gamma}), & \gamma \to -\infty \end{cases},
$$

then Theorem 6.4 applies.

According to Theorem 6.4 and Corollary 6.5, once d is chosen, one just needs to

adjust M to bound this approximation error to a desired level.

In multivariate cases where Y_l's are dependent, but γ's pdf satisfies the assumptions in Corollary 6.5, Theorem 6.4 still applies. For the cases that Y_l's are independent, the analyticity strip of γ is the intersection of the analyticity strips of all the Y_l's.

6.2 Commonly used distributions for channel fading amplitude modeling

In this section, we analyze the tail behavior of three widely-used distributions in channel amplitude modeling: Rayleigh, Nakagami-m, and Nakagami-n(Rice) distribution, all of which have characteristic functions satisfying the specific analytic conditions and have exponentially decaying pdfs.

6.2.1 Rayleigh distribution

The pdf of a Rayleigh random variable γ is given as:

$$p(\gamma) = \begin{cases} \frac{\gamma}{\sigma^2} e^{-\frac{\gamma^2}{2\sigma^2}}, & \gamma \geq 0 \\ 0, & \gamma < 0 \end{cases},$$

where $\sigma > 0$ is the scale parameter of the distribution.

Since $p(\gamma)$ is decaying exponentially in the order of $-\frac{\gamma^2}{2\sigma^2}$, the corresponding d_+ and d_- can be chosen to be any finite positive real values.

The characteristic function of a Rayleigh random variable with parameter σ is:

$$\phi_\gamma(t) = 1 - \sigma t e^{-\frac{\sigma^2 t^2}{2}} \sqrt{\frac{\pi}{2}} \big(\text{erfi}(\frac{\sigma t}{\sqrt{2}}) - i\big).$$

Here, erfi is the imaginary error function.

6.2.2 Nakagami-m distribution

Nakagami-m distribution has the probability density function:

$$p(\gamma) = \begin{cases} \frac{2m_l^{m_l}\gamma^{2m_l-1}}{\Omega_l^{m_l}\Gamma(m_l)} \exp\big(-\frac{m_l\gamma^2}{\Omega_l}\big), & \gamma \geq 0 \\ 0, & \gamma < 0 \end{cases},$$

where m_l is the Nakagami-m fading parameter, and $\Omega_l = \mathbb{E}[\gamma^2]$.

The dominating term in $p(\gamma)$, $\exp\big(-\frac{m_l\gamma^2}{\Omega_l}\big)$, determines that it decays at an exponential rate that is higher than any linear order. This determines that d_+ and d_- can be any positive real value.

The characteristic function is:

$$\phi_\gamma(t) = {}_1F_1(m_l; \frac{1}{2}; -\frac{\Omega_l t^2}{4m_l}) + it\sqrt{\frac{\Omega_l}{m_l}}\frac{\Gamma(m_l+\frac{1}{2})}{\Gamma(m_l)}{}_1F_1(m_l+\frac{1}{2}; \frac{3}{2}; -\frac{\Omega_l t^2}{4m_l}),$$

where ${}_1F_1(\cdot;\cdot;\cdot)$ is the confluent hypergeometric function of the first kind.

6.2.3 Nakagami-n(Rice) distribution

A Nakagami-n random variable has the following pdf:

$$p(\gamma) = \begin{cases} \frac{2(1+n_0^2)\gamma e^{-n_0^2}}{\Omega} e^{-\frac{1+n_0^2}{\Omega}\gamma^2} I_0\big(2n_0\gamma\sqrt{\frac{1+n_0^2}{\Omega}}\big), & \gamma \geq 0 \\ 0, & \gamma \leq 0 \end{cases},$$

where $I_0(u) = \sum_{k=0}^{\infty} \frac{(-1)^k u^{2k}}{2^{2k}(k!)^2}$ is the zero-order Bessel Function of the first kind.

We confirm the exponentially decaying tail behavior of $p(\gamma)$. d_+ and d_- can be chosen to be any positive real number.

The corresponding characteristic function of Nakagami-n(Rice) distribution is:

$$\phi_\gamma(t)$$
$$=e^{-n_0^2}\sum_{k=0}^{\infty}\frac{n_0^{2k}}{k!}\big({}_1F_1(k+1;\frac{1}{2};-\frac{t^2\Omega}{4(1+n_0^2)})+\sqrt{\frac{\Omega}{1+n_0^2}}it\frac{\Gamma(k+\frac{3}{2})}{\Gamma(k+1)}{}_1F_1(k+\frac{3}{2};\frac{3}{2};-\frac{t^2\Omega}{4(1+n_0^2)})\big).$$

6.2.4 Correlated Nakagami-m distribution

For the multi-dimensional case, by [41], the characteristic function for the L_r-variate Nakagami-m distributed random variables, $Y_1, Y_2, \cdots, Y_{L_r}$, with exponential correlation coefficient $\rho_{i,j} = \rho^{|i-j|}(i, j = 1, 2, \cdots, L_r)$ is given as:

$$\begin{aligned}
&\phi_{Y_1,Y_2,\cdots,Y_{L_r}}(t_1,t_2,\cdots,t_{L_r})\\
=\quad&\frac{2^{L_r(1-2m_l)}\pi^{\frac{L}{2}}(1-\rho^2)^{m_l}}{\Gamma(m_l)(1+\rho^2)^{m_l(L_r-2)}}\sum_{i_1,i_2,\cdots,i_{L_r-1}=0}^{\infty}(\frac{\rho}{4})^{2\sum_{j=1}^{L_r-1}i_j}\\
&\times\frac{(1+\rho^2)^{-\left(i_1+2\sum_{j=2}^{L_r-2}i_j+i_{L_r-1}\right)\beta}}{\prod_{j=1}L_r-1i_j!\Gamma(m_l+i_j)}\\
&\times\prod_{k=1}^{L_r}\Gamma(2\xi_k)[A_{k1}F_1\big(\xi_k;\frac{1}{2};-c_kt_k^2\big)+it_kB_{k1}F_1\big(\frac{1}{2}+\xi_k;\frac{3}{2};-c_kt_k^2\big)],
\end{aligned}\tag{6.10}$$

where $\beta = 0$ for $L_r = 2$ and $\beta = 1$ for $L_r \geq 3$, $\xi_k = i_{k-1} + i_k + m_l$, and

$$
\begin{aligned}
A_k &= [\Gamma(\frac{1}{2}+\xi_k)]^{-1}, k=1,2,\cdots,L_r \\
B_k &= \begin{cases} \sqrt{2(1-\rho^2)}[\Gamma(\xi_k)]^{-1}, k=1,L_r \\ \sqrt{\frac{2(1-\rho^2)}{1+\rho^2}}[\Gamma(\xi_k)]^{-1}, k=2,3,\cdots,L_r-1 \end{cases} \\
c_k &= \begin{cases} \frac{1-\rho^2}{2}, k=1,L_r \\ \frac{1-\rho^2}{2(1+\rho^2)}, k=2,3,\cdots,L_r-1 \end{cases},
\end{aligned}
$$

where $i_0 = i_{L_r} = 0$.

6.3 Numerical Experiments

In this section, we exhibit the proposed approximation method numerically. The following numerical experiments are done on a MacBook Pro with 8GB RAM and 2.6 GHz Intel Core i5 with C++.

In Section 6.3.1, we display BER approximation results under cases $L_r = 1, 2, 3$. Section 6.3.2 takes one-dimensional Rayleigh distribution as an example by plotting the BER with respect to SNR.

The parameter settings we choose for the numerical experiments are listed in Table 6.1:

Parameter Setting	
Rayleigh	$\sigma = 0.2$
Nakagami-m	$m_l = 2, \quad \Omega_l = 0.25$
Nakagami-n(Rice)	$n_0 = \sqrt{2}, \quad \Omega = 0.25$
(exponentially) correlated Nakagami-m	$m_l = 0.8, \quad \Omega_l = 0.25, \quad \rho = 0.3$

Table 6.1: Parameter Setting

Since the selection of h is based on the choice of d by $h = h(M) = (4\pi d M^{-2})^{1/3}$. Due to the fastly decaying property of the special functions, d should take small values to avoid large approximation error.

6.3.1 BER approximation

In our BER numerical experiment part, we consider only the coherent FSK cases, with $a = \frac{1}{2}, b = 0, p = \frac{1}{2}$.

Table 6.2 gives the BER approximation results under Rayleigh distribution with $\sigma = 0.2$ and $\beta = 50$. d is set to be 0.1. M value ranges from 4 to 28. $h = h(M)$ is taken and values are included in the second column of the table. The approximated BER values are included in the Column 'approx BER'. The benchmark used for computing the approximation error comes from the closed-form BER expression with respect to SNR by [40]: $BER = \frac{1}{2}(1-\sqrt{\frac{SNR}{1+SNR}})$. As we can see, the approximated BER converges to 0.00248141 with approximation error reaching the level of 1.0E-9.

BER Approximation				
M	h	approx BER	approx error	comp time
8	0.2698	0.07032397	6.8E-2	3.0E-6
12	0.2059	0.01237585	9.9E-3	7.0E-6
16	0.1700	0.00312974	1.2E-3	9.0E-6
20	0.1465	0.00259128	1.1E-4	8.0E-6
24	0.1297	0.00249036	9.0E-6	1.2E-5
28	0.1170	0.00248204	6.3E-7	1.4E-5
32	0.1071	0.00248144	3.9E-8	1.4E-5
36	0.0990	0.00248141	2.2E-9	1.5E-5
40	0.0923	0.00248141	1.1E-10	2.7E-5

Table 6.2: BER Approximation: $\gamma \sim$ Rayleigh(0.2), $\beta = 50$, d = 0.1

Table 6.3 shows the approximated BER values under Nakagami-m(2, 0.25) while M takes value from 20 to 40. $d = 0.05$ and $\beta = 50$. Since there is no closed-form solution to BER under Nakagami-m distribution, we use the approximated BER value

with $M = 48$ and $h = h(M)$ correspondingly as the benchmark. The approximated BER approaches to 0.00000760 with approximation error decreases below 1.0E-9 level finally.

BER Approximation				
M	h	approx BER	approx error	comp time
20	0.1162	0.00095006	9.4E-4	1.3E-4
24	0.1029	0.00009887	9.1E-5	1.7E-4
28	0.0929	0.00001500	7.4E-6	2.0E-4
32	0.0850	0.00000811	5.1E-7	2.4E-4
36	0.0786	0.00000763	3.1E-8	2.7E-4
40	0.0732	0.00000760	1.7E-9	3.9E-4
44	0.0687	0.00000760	7.5E-11	4.0E-4

Table 6.3: BER Approximation: $\gamma \sim$ Nakagami-m(2, 0.25), $\beta = 50$, d = 0.05

Similarly, Table 6.4 shows the numerical results for $\gamma \sim$ Nakagami-n(2, 0.25), $d = 0.05, \beta = 50$. Benchmark value is approximated by setting $M = 50$ and $h = h(M)$ correspondingly.

BER Approximation				
M	h	approx BER	approx error	comp time
16	0.1349	0.00878907	8.5E-3	8.4E-4
20	0.1162	0.00120124	8.7E-4	1.2E-3
24	0.1029	0.00039683	6.6E-5	1.5E-3
28	0.0929	0.00033502	3.7E-6	2.0E-3
32	0.0850	0.00033144	1.7E-7	2.4E-3
36	0.0786	0.00033128	5.8E-9	2.6E-3
40	0.0732	0.00033128	1.6E-10	3.5E-3

Table 6.4: BER Approximation: $\gamma \sim$ Nakagami-n($\sqrt{2}$, 0.25), $\beta = 50$, d = 0.05

The numerical results for $\gamma = Y_1 + Y_2, Y_1 \sim$ Rayleigh(0.2), $Y_2 \sim$ Nakagami-m(2, 0.25), $\beta = 10$ are included in Table 6.5. We set $d = 0.1$. The 2-dimensional BER converges faster and stays around 0.00015147 as soon as M reaches 10. The approximation error is below the level of 1.0E-9 afterwards. Here the benchmark is the approximated value when $M = 35$.

BER Approximation				
M	h	approx BER	approx error	comp time
2	0.6798	0.19328928	1.9E-1	6.0E-6
4	0.4282	0.00219236	2.0E-3	1.0E-5
6	0.3268	0.00015608	4.6E-6	1.4E-5
8	0.2698	0.00015148	5.2E-9	1.8E-5
10	0.2325	0.00015147	1.4E-10	2.3E-5
12	0.2059	0.00015147	1.7E-11	2.9E-5

Table 6.5: BER Approximation, $\gamma = Y_1 + Y_2$: $Y_1 \sim$ Rayleigh(0.2), $Y_2 \sim$ Nakagami-m(2, 0.25), $\beta = 10$, d $= 0.1$

Numerical results for BER approximation under 2-d correlated Nakagami-m distribution in Table 6.6 shows exponential convergence. The exact value benchmark is chosen as the approximated value with $M = 90$.

BER Approximation				
M	h	approx BER	approx error	comp time
20	0.1162	0.00294194	2.9E-3	4.0E-3
26	0.0976	0.00031579	2.6E-4	6.2E-3
32	0.0850	0.00007960	1.9E-5	8.3E-3
38	0.0758	0.00006170	1.2E-6	1.1E-2
44	0.0687	0.00006054	6.7E-8	1.3E-2
50	0.0631	0.00006048	2.7E-9	1.4E-2

Table 6.6: BER Approximation: $\gamma = Y_1 + Y_2$, $Y_1, Y_2 \sim$ correlated Nakagami-m(0.8, 0.25), $\rho = 0.3$, $\beta = 10$, d $= 0.05$

As Table 6.7 shows, when $\beta = 10$, d is chosen to be 0.1, and $\gamma = Y_1 + Y_2 + Y_3$, $Y_1 \sim$ Rayleigh(0.2), $Y_2 \sim$ Nakagami-m(2, 0.25), $Y_3 \sim$ Nakagami-n($\sqrt{2}$, 0.25), the approximated BER exhibit fast convergence to the level 0.00000485. We choose the approximated value with $M = 30$ to be the benchmark.

6.3.2 BER vs SNR

This section shows the relationship between BER and SNR numerically under one-dimensional Rayleigh distribution. With increasing value of σ, the value of SNR in-

BER Approximation				
M	h	approx BER	approx error	comp time
4	0.4282	0.17447099	1.7E-1	7.2E-5
6	0.3268	0.00992374	9.9E-3	1.1E-4
8	0.2698	0.00021910	2.1E-4	1.4E-4
10	0.2325	0.00000708	2.2E-6	2.1E-4
12	0.2059	0.00000486	1.3E-8	2.3E-4
14	0.1858	0.00000485	4.6E-11	2.9E-5
16	0.1700	0.00000485	1.0E-13	3.0E-4

Table 6.7: BER Approximation, $\gamma = Y_1 + Y_2 + Y_3$: $Y_1 \sim$ Rayleigh(0.2), $Y_2 \sim$ Nakagami-m(2, 0.25), $Y_3 \sim$ Nakagami-n($\sqrt{2}$, 0.25), $\beta = 10$, d = 0.1

creases, consequently, BER decreases. In Figure 6.3.2, we plot the approximated BER values in log scale with respect to SNR values varying from 0 to 35. Here, the SNR values are transformed into unit dB by SNR $= 10\log_{10}(\beta^2\sigma^2)$.

6.3.3 Square QAM approximation

To show the efficiency in approximating Square QAM by our method. We take three examples: $\gamma \sim$ Rayleigh(0.2), $\gamma = Y_1 + Y_2, Y_1 \sim$ Rayleigh(0.2), $Y_2 \sim$ Nakagami-m(2, 0.25), and $\gamma = Y_1 + Y_2, Y_1, Y_2 \sim$ correlated Nakagami-m(0.8, 0.25), with $\rho = 0.3$.

Under the current parameter setting, Square QAM approximation results in Table 6.8 for the 1-d Rayleigh case reaches the stable level as soon as M gets close to 55.

Square QAM Approximation				
M	h	approx Square QAM	approx error	comp time
25	0.1002	0.00057324	1.4E-8	5.8E-3
30	0.0887	0.00057328	2.8E-8	8.2E-3
35	0.0800	0.00057329	3.0E-8	1.3E-2
40	0.0732	0.00057328	2.3E-8	1.8E-2
45	0.0677	0.00057327	1.5E-8	2.0E-2
50	0.0631	0.00057326	9.5E-9	2.6E-2
55	0.0592	0.00057326	5.3E-9	3.5E-2

Table 6.8: Square QAM Approximation: $\gamma \sim$ Rayleigh(0.2), $\beta = 60$, d = 0.05, $\tilde{M} = 32$

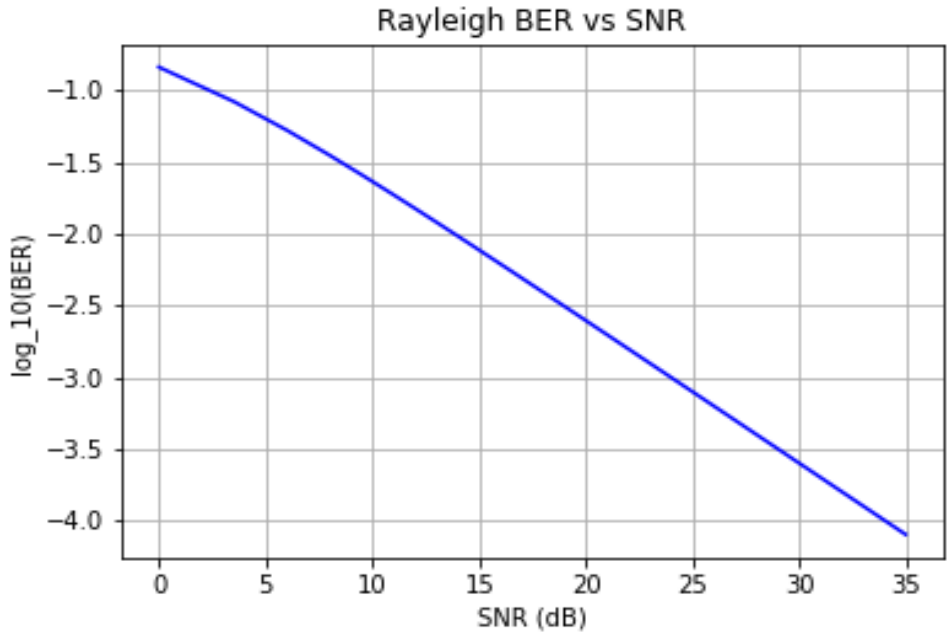

Figure 6.1: 1-d Rayleigh BER vs SNR (dB)

Table 6.9 shows exponential convergence in the Square QAM approximation error.

Square QAM Approximation				
M	h	approx Square QAM	approx error	comp time
10	0.2325	0.02274020	2.3E-2	1.1E-3
14	0.1858	0.00092261	9.2E-4	2.3E-3
18	0.1572	0.00002541	2.0E-5	3.6E-3
22	0.1374	0.00000527	2.9E-7	5.6E-3
26	0.1230	0.00000499	3.0E-9	7.8E-3
30	0.1118	0.00000499	2.7E-11	1.2E-2

Table 6.9: Square QAM Approximation: $\gamma = Y_1 + Y_2$: $Y_1 \sim$ Rayleigh(0.2), $Y_2 \sim$ Nakagami-m(2, 0.25), $\beta = 60$, d $= 0.1$, $\tilde{M} = 32$

Square QAM approximation error in Table 6.10 decays exponentially with increasing M. The approximated value when $M = 66$ is selected as benchmark Square QAM value.

BER Approximation				
M	h	approx Square QAM	approx error	comp time
6	0.2594	0.02641576	2.5E-2	9.7E-3
12	0.1634	0.00126018	9.6E-5	4.2E-2
18	0.1247	0.00116193	2.1E-6	9.8E-2
24	0.1029	0.00116330	7.2E-7	2.0E-1
30	0.0887	0.00116377	2.5E-7	3.3E-1
36	0.0786	0.00116393	8.9E-8	5.2E-1
42	0.0709	0.00116399	3.3E-8	7.5E-1
48	0.0648	0.00116401	1.2E-8	9.8E-1
54	0.0600	0.00116401	4.4E-9	1.3E-0

Table 6.10: Square QAM Approximation: $\gamma = Y_1 + Y_2$, $Y_1, Y_2 \sim$ correlated Nakagami-m(0.8, 0.25), $\rho = 0.3$, $\beta = 15$, d = 0.05

Bibliography

[1] Andersen, L., and Andreasen, J. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research 4*, 3 (2000), 231–262.

[2] Annamalai, A., Su, J., and Tellambura, C. Exact analysis of equal gain diversity systems over fading channels. *VTC2000-Spring. 2000 IEEE 51st Vehicular Technology Conference Proceedings 1* (2000), 612–616.

[3] Annamalai, A., and Tellambura, C. Error rates for nakagami-m fading multichannel reception of binary and m-ary signals. *IEEE Transactions on Communications 49*, 1 (2001), 58–68.

[4] Annamalai, A., Tellambura, C., and Bhargava, V. K. Unified analysis of equal-gain diversity on rician and nakagami fading channels. *IEEE Wireless Communications and Networking Conference 1* (1999), 10–14.

[5] Annamalai, A., Tellambura, C., and Bhargava, V. K. A general method for calculating error probabilities over fading channels. *IEEE Transactions on Communications 53*, 5 (2005), 841–852.

[6] Bakshi, G., Cao, C., and Chen, Z. Empirical performance of alternative option pricing models. *Journal of Finance 52*, 5 (1997), 2003–2049.

[7] Bakshi, G., and Madan, D. Spanning and derivative-security valuation. *Journal of Financial Economics 55* (2000), 205–238.

[8] Barndorff-Nielsen, O. E. Process of normal inverse Gaussian type. *Finance and Stochastics 2*, 1 (1998), 41–68.

[9] Bates, D. S. Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies 9*, 1 (1996), 69–107.

[10] Beaulieu, N. An infinite series for the computation of the complementary probability distribution function of a sum of independent random variables and its application to the sum of rayleigh random variables. *IEEE Transactions on Communications 38*, 9 (1990), 1463–1474.

[11] Beaulieu, N., and Abu-Dayya, A. Analysis of equal gain diversity of nakagami fading channels. *IEEE Transactions on Communications 39*, 2 (1991), 225–234.

[12] Black, F., and Scholes, M. The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 3 (1973), 637–654.

[13] Carr, P., Geman, H., Madan, D. B., and Yor, M. The fine structure of asset returns: an empirical investigation. *Journal of Business 75*, 2 (2002), 305–332.

[14] Carr, P., and Madan, D. Saddlepoint methods for option pricing. *Journal of Computational Finance 13*, 1 (2009), 49–61.

[15] Carr, P., and Madan, D. B. Option valuation using the fast Fourier transform. *Journal of Computational Finance 2*, 4 (1999), 61–73.

[16] Carr, P., and Wu, L. Finite moment log stable process and option pricing. *Journal of Finance 58*, 2 (2003), 753–777.

[17] Casella, G., L.-M., and Robert. Generalized accept-reject sampling schemes. *Institute of Mathematical Statistics 45* (2004), 342–347.

[18] Chacko, G., and Das, S. Pricing interest rate derivatives: a general approach. *Review of Financial Studies 15*, 1 (2002), 195–241.

[19] Chen, Z., Feng, L., and Lin, X. Inverse transform method for simulating Lévy processes and discrete Asian options pricing. In *Proceedings of the 2011 Winter Simulation Conference* (Piscataway, New Jersey, 2011), S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu, Eds., IEEE Press, pp. 444–456.

[20] Chen, Z., Feng, L., and Lin, X. Simulating Lévy processes from their characteristic functions and financial applications. *ACM Transactions on Modeling and Computer Simulation 22*, 3 (2012), 14:1–14:26.

[21] Devroye, L. *Non-Uniform Random Variate Generation.* Springer Verlag New York.

[22] Devroye, L. On the computer generation of random variables with a given characteristic function. *Computers & Mathematics with Applications 7*, 6 (1981), 547–552.

[23] Duffie, D., Pan, J., and Singleton, K. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*, 6 (2000), 1343–1376.

[24] E.C.Titchmarsh. *Introduction to the theory of Fourier integrals.* Claredon Press, Oxford, 1948.

[25] Evard, J., and Jafari, F. A complex rolle's theorem. *American Mathematical Monthly 99*, 9 (1992), 858–861.

[26] Feng, L., and Lin, X. Inverting analytic characteristic functions and financial applications. *SIAM Journal on Financial Mathematics Forthcoming* (2013).

[27] Glasserman, P., and Liu, Z. Sensitivity estimates from characteristic functions. In *Proceedings of the 2007 Winter Simulation Conference*, S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, Eds. 2007, pp. 932–940.

[28] Glasserman, P., and Liu, Z. Sensitivity estimates from characteristic functions. *Operations Research 58*, 6 (2010), 1611–1623.

[29] Heston, S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*, 2 (1993), 327–343.

[30] Karagiannidis, G. Moments-based approach to the performance analysis of equal gain diversity in nakagami-m fading. *IEEE Transactions on Communications 52*, 5 (2004), 685–690.

[31] Kou, S., and Wang, H. Option pricing under a double exponential jump diffusion model. *Management Science 50*, 9 (2004), 1178–1192.

[32] Kou, S. G. A jump-diffusion model for option pricing. *Management Science 48*, 8 (2002), 1086–1101.

[33] Kou, S. G., and Wang, H. First passage times of a jump diffusion process. *Adavances in Applied Probability 35*, 2 (2003), 504–531.

[34] Lee, R. W. Option pricing by transform methods: extensions, unification, and error control. *Journal of Computational Finance 7*, 3 (2004), 51–86.

[35] Madan, D., Carr, P., and Chang, E. The variance gamma process and option pricing. *European Finance Review 2*, 1 (1998), 79–105.

[36] Madan, D., and Milne, F. Option pricing with V.G. martingale components. *Mathematical Finance 1*, 4 (1991), 39–55.

[37] Madan, D., and Seneta, E. The variance gamma (V.G.) model for share market returns. *Journal of Business 63*, 4 (1990), 511–524.

[38] Merton, R. C. Theory of rational option pricing. *Bell Journal of Economics and Management Science 4*, 1 (1973), 141–183.

[39] Merton, R. C. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics 3*, 1-2 (1976), 125–144.

[40] Proakis, J. G., and Salehi, M. *Digital Communications.*

[41] Sahu, P., and Chaturvedi, A. Performance analysis of predetection egc in exponentially correlated nakagami-m fading channel. *IEEE Transactions on Communications 53* (2005), 1252–1256.

[42] Shin, H., and Lee, J. On the error probability of binary and m-ary signals in nakagami-m fading channels. *IEEE Transactions on Communications 52*, 4 (2004), 536–539.

[43] Simon, M. K., and Divsalar, D. A unified approach to performance analysis of digital communication over generalized fading channels. *IEEE Transactions on Communications 46*, 2 (1998), 200–210.

[44] Yilmaz, F., and Alouini, M.-S. Novel mgf-based expressions for the average bit error probability of binary signaling over generalized fading channels. *2014*

IEEE Wireless Communications and Networking Conference (WCNC) (2014), 1121–1125.

[45] Zhang, Q. Probability of error for equal-gain combiners over rayleigh channels: some closed-form solutions. *IEEE Transactions on Communications 45*, 3 (1997), 270–273.

[46] Zogas, D., Karagiannidis, G., and Kotsopoulos, S. Equal gain combining over nakagami-n (rice) and nakagami-q (hoyt) generalized fading channels. *IEEE Transactions on Wireless Communications 4*, 2 (2005), 374–379.