ESTIMATION AND SIMULATION FOR MULTIVARIATE TEMPERED STABLE DISTRIBUTIONS WITH APPLICATIONS TO FINANCE

by

Yunfei Xia

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2022

Approved by:

Dr. Michael Grabchak

Dr. Stanislav Molchanov

Dr. Isaac Sonin

Dr. Weidong Tian

ABSTRACT

YUNFEI XIA. Estimation and Simulation for Multivariate Tempered Stable Distributions with Applications To Finance. (Under the direction of DR. MICHAEL GRABCHAK)

In this thesis, we introduce a methodology for the simulation and parameter estimation of multivariate tempered stable distributions. Using the fact that tempered stable distributions can be specified indirectly by a Lévy measure, our approach is based on an approximation due to a discretization of the Lévy measure. We derive this discretization in general and give an explicit construction of the discretization in the bivariate case. Also, our approximation results hold for a wide class of multivariate infinitely divisible distributions.

Based on our main approximation, we develop a method for simulations, which we call the discretization and simulation (DS) method. To demonstrate how well the method works, we perform a series of simulations in the bivariate cases and compare it with another approximate simulation method developed by Rosiński's. Further, we use our discretization for parameter estimation by minimizing the distance between the characteristic function of the multivariate tempered stable distribution and the empirical characteristic function. We then apply our methodology to two bivariate financial datasets related to exchange rates. The first is comprised of exchange rates between standard currencies, while the second is based on exchange rates related to cryptocurrencies. We also perform goodness-of-fit tests to show that the multivariate tempered stable model does a good job fitting the model.

Further, we apply our model for the pricing of the bivariate basket option. Toward this end, we provide theoretical results on the existence of equivalent martingale measures. Then combining this with our model for parameter estimation and the DS method for simulation, we develop a Monte Carlo based method for option pricing. We apply it to the pricing of European call options with different strikes and the

pricing of the Multi-asset rainbow option.

ACKNOWLEDGEMENTS

First and foremost I would like to express my sincere gratitude to my advisor, Dr. Michael Grabchak for his continuous support during my graduate life. His immense knowledge and passion for research have inspired me all the time in my academic research. His patient guidance and insightful feedback pushed me to sharpen my thinking and brought my work to a higher level. It was a great fortune to learn from his plentiful experience, which inspired me a lot in my research and my future career. I could not have imagined having a better advisor and mentor for my Ph.D. study.

I would like to extend my sincere thanks to the rest of my thesis committee: Dr. Stanislav Molchanov, Dr. Isaac Sonin, and Dr. Weidong Tian for their time and insightful comments. Moreover, I would like to offer my special thanks to Dr. Shaozhong Deng and Dr. Mohammad A. Kazemi for their kind help and support throughout my Ph.D. study and teaching assistant role at the Department of Mathematics and Statistics. Also, I gratefully acknowledge the financial support received for my Ph.D. from the Graduate School and Mathematics department. Moreover, I would like to extend my thanks to my loving friends and all the people who help me throughout my Ph.D. journey.

I am deeply thankful to my parents Rencai Xia and Ming Sun for their love and sacrifices. Without their tremendous understanding and endless support, it would be impossible for me to complete my study. This last word of acknowledgment I dedicate this dissertation in loving memory of my uncle Yun Sun, whose role in my life was, and remains, immense. His belief in me has made this journey possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Tempered stable (TS) distributions form a large and flexible class of models and they are popular in real-world applications and have been used in computer science, see [1], [2], actuarial science [3], physics [4], dynamical system [5], biostatistics [6], mathematical finance, see [7], [8], [9] and so on.

Especially in the finance area, TS distribution has an important role and this dissertation is motivated by this. Over the past decades, academic researchers and market practitioners have developed and investigated many different models and techniques for modeling financial returns. As early as 1900, Louis Bachelier proposed Brownian motion to model financial return in his Ph.D thesis, and then Brownian motion is still widely used and is the framework. The most famous model is Black-Scholes model [10], which assumed that the asset price dynamics are driven by a Geometric Brownian Motion (GBM). But Brownian motion is based on Normal distribution, and its tails are too light and cannot fit financial data well. To improve this, stable distributions were proposed, and they allow skewness and fat tails. Stable distributions are widely applied in financial returns, and they could provide a good fit in financial data, except on the extreme tails [11]. Hence, both normal and stable distributions have limitations to model financial return, see [12]. To know the limitation of stable distributions further, we give a brief introduction to them. Stable distribution belongs to infinitely divisible distributions. We give the definition of infinitely divisible distributions as follows,

Definition 1. *A probability measure μ on $\mathbb{R}^d$ is called* ***infinitely divisible*** *if for any n there exists a probability measure μ_n such that if $X \sim \mu$ and $Y_1^{(n)}, \ldots, Y_n^{(n)} \overset{\text{iid}}{\sim} \mu_n$*

then

$$X \stackrel{d}{=} Y_1^{(n)} + Y_2^{(n)} + \cdots + Y_n^{(n)}.$$

The characteristic function of an infinitely divisible distribution μ on $\mathbb{R}^d$ can be written in the form

$$\hat{\mu}(z) \;\; = \;\; \mathrm{E}[e^{i\langle z,X\rangle}] = e^{\psi(z)}, \;\; z \in \mathbb{R}^d$$

where

$$\psi(z) = -\frac{1}{2}\langle z, Az\rangle + i\langle b, z\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle x,z\rangle} - 1 - i\frac{\langle x, z\rangle}{1+|x|^2} \right) L(\mathrm{d}x),$$

A is a $d \times d$ covariance matrix, $b \in \mathbb{R}^d$, L is a Borel measure, called the Lévy measure satisfying

$$L(\{0\}) = 0 \text{ and } \int_{\mathbb{R}^d} \left(1 \wedge |x|^2\right) L(\mathrm{d}x) < \infty$$

The parameters (A, L, b) uniquely determine the distribution. According that, we give the definition of stable distribution, as follow,

Definition 2. *Let X_1, X_2, ...$X_n \stackrel{\text{iid}}{\sim} \mu$ be d-dimensional random vectors. We say that μ is **stable** if for any n there are nonrandom constants $a_n > 0$ and $b_n \in \mathbb{R}^d$ such that*

$$a_n \left(X_1 + X_2 + ... + X_n - b_n\right) \stackrel{d}{=} X_1.$$

It turns out that $a_n = n^{-1/\alpha}$ for some $\alpha \in (0, 2]$, and we say that μ is α-stable. There is a fact that a distribution is Gaussian if and only if it is 2-stable. Let's consider the case $d = 1$. For $\alpha \in (0, 2)$ the Lévy measure of a stable distribution is of the form

$$L_0(\mathrm{d}x) = c_-|x|^{-1-\alpha}1_{x<0}\mathrm{d}x + c_+x^{-1-\alpha}1_{x>0}\mathrm{d}x$$

for $c_-, c_+ \geq 0$. Since $\int_{|x|>1} |x|^2 L_0(\mathrm{d}x) = \infty$, all stable distributions with $\alpha \in (0, 2)$ have an infinite variance. Their infinite variance and extreme heavy tails let stable distributions be often unrealistic in practices, there may be real-world obstacles limiting the size of random phenomena. This has led researchers to find new models that are similar to stable distributions but with a lighter tail. TS distributions were first introduced in Tweedie (1984) [13]. Under $d = 1$, the Lévy measure of a TS distribution is given by

$$L(\mathrm{d}x) = c_-|x|^{-1-\alpha}e^{-b_-|x|}1_{x<0}\mathrm{d}x + c_+x^{-1-\alpha}e^{-b_+x}1_{x>0}\mathrm{d}x.$$

This leads to the exponential decay of the distribution. So, TS distributions are obtained by modifying the tails of infinite variance stable distributions to make them lighter. And they have

$$\int_{|x|>1} |x|^\gamma L(\mathrm{d}x) < \infty,$$

thus, all moments of TS distributions with $\alpha \in (0, 2)$ are finite. Also, TS distributions are similar to stable distributions in some central regions. Based on these nice properties, obviously, TS distributions are more realistic for financial markets.

In the applications in finance, we are interested in how to provide a reasonable model of financial returns and focus on option pricing. Under univariate case, there are lots work of TS distributions on financial derivative, see, e.g [14], [15], [16], [17], [18], and so on. These literature show that TS tends to provide a good fit for financial data. For multivariate cases, Rosiński provided a more general class of multivariate TS distributions in [19] and set up a general theoretical framework for them. And many theoretical properties can be found in [20], [21], and [22]. But there are almost no applications to finance based on the multivariate TS distributions, and the existing models related to normal TS distributions (NTS), see, [23], [24], [25], which are different with generalized TS distributions. And they simulated TS subordinator of

NTS by Theorem 5.3 in [19], which is the only existing simulation methodology of multivariate TS distributions, although it is approximate. To apply the generalized TS distributions in the multi-asset option pricing problem, we introduce a methodology for simulation and parameter estimation of multivariate TS distributions, with an emphasis on the bivariate case.

However, working with TS distributions is nontrivial in the univariate case, as these distributions do not have a closed form for their probability density functions, which must be evaluated numerically by inverting the characteristic function. See e.g. [26] for a discussion of various numerical methods. The situation is even more complicated in the multivariate case, where there is not even a closed form for the characteristic function and where the model is only semiparametric as it relies on the so-called spectral measure, and spectral measure is an infinite dimensional parameter. We realized that little work is on the multivariate case in terms of computation and data analysis, and this problem becomes one of the most important parts of this dissertation.

This dissertation is organized as follows. In Chapter 2, we introduce the class of multivariate TS distributions and give some properties of their mean, variance, correlation, skewness, and kurtosis's formula. Also, we state our main approximation results. And we introduce a methodology to simulate the approximate multivariate TS distributions by discretizing the spectral measure. This approach is inspired by a similar discretization for multivariate stable distributions, which was derived in [27]. That discretization was then used for simulation and parameter estimation of stable distributions in [28], [29], and [30]. Further, we give a construction of the discretization in the bivariate case.

In Chapter 3, based on our main approximation, we develop a method for simulations in the bivariate case and call it the discretization and simulation (DS) method. To demonstrate how well the method works, we perform a series of simulations in the

bivariate cases and compare it with another approximate simulation method called Rosińskis method in [19].

In Chapter 4, we use the DS method to do parameter estimation by minimizing the distance between the characteristic function of the multivariate tempered stable distribution and the empirical characteristic function. And we give a couple of simulation examples to show the performance of the methodology. Further, we apply it to two bivariate financial datasets related to exchange rates and also perform goodness-of-fit tests to show that the multivariate tempered stable model does a good job fitting the model.

In Chapter 5, we consider another related problem which is how to find a suitable risk-neutral measure in order to discuss option pricing based on arbitrage theory, some related work in [31], [32]. We provide the theoretical result of the existence of equivalent martingale measures. Then combing with the DS method, we apply them to bivariate financial datasets to get the prices of European call options with different strikes and the price of the Multi-asset rainbow option by performing a Monte Carlo study.

In Chapter 6 we extend our theoretical results to a wide class of multivariate infinitely divisible models. And the Chapter 7 gives the proofs of Chapter 6.

This dissertation is based on the published paper: Estimation and simulation for multivariate tempered stable distributions published in the *Journal of Statistical Computation and Simulation*, 2022, Volume 92, no.3, page 451-475, [33], but with more details and extend its application to finance, specific in option pricing.

we introduce some notation. We write $\mathbb{R}^d$ to denote the space of d-dimensional column vectors equipped with the usual inner product $\langle \cdot, \cdot \rangle$ and the usual norm $|\cdot|$. We write $\mathbb{S}^{d-1} = \left\{ s \in \mathbb{R}^d : |s| = 1 \right\}$ to denote the unit sphere in $\mathbb{R}^d$.For a Borel set $\mathbb{H} \subset \mathbb{R}^d$, we write $\mathfrak{B}(\mathbb{H})$ to denote the collection of Borel sets contained in $\mathbb{H}$. If μ is a probability distribution on $\mathbb{R}^d$, we write $X \sim \mu$ to denote that X is a d-dimensional

random vector with distribution μ. We write $U(a, b)$ to denote a uniform distribution on (a, b) and $\text{Exp}(\lambda)$ to denote an exponential distribution with rate λ. We write $\mathbb{C}$ to denote the set of complex numbers. For $z \in \mathbb{C}$ we write $\Re z$ to denote its real part and $\Im z$ to denote its imaginary part.

CHAPTER 2: MULTIVARIATE TEMPERED STABLE DISTRIBUTIONS

2.1 Introduction of multivariate tempered stable distributions

Let's formally introduce multivariate tempered stable distributions first. Fix $\alpha \in (0,1)$, $\gamma \in \mathbb{R}^d$, let $b : \mathbb{S}^{d-1} \mapsto (0,\infty)$ be a Borel function, and let σ be a finite Borel measure on $\mathbb{S}^{d-1}$. Consider a distribution μ on $\mathbb{R}^d$, whose characteristic function is given, for any $z \in \mathbb{R}^d$, by

$$\begin{aligned} \hat{\mu}(z) &= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\int_0^\infty (e^{i\langle s,z\rangle x} - 1)\frac{e^{-b(s)x}}{x^{1+\alpha}}\mathrm{d}x\sigma(\mathrm{d}s)\right] \\ &= \exp\left[i\langle\gamma, z\rangle + \Gamma(-\alpha)\int_{\mathbb{S}^{d-1}}((b(s) - i\langle s, z\rangle)^\alpha - b^\alpha(s))\sigma(\mathrm{d}s)\right], \end{aligned} \tag{2.1}$$

where the equality follows by (3.38) in [22]. We call this a tempered stable (TS) distribution and write $\mu = \mathrm{TS}_\alpha(\sigma, b, \gamma)$. Here α is called the index of stability, σ is called the spectral measure, $b(\cdot)$ is called the exponent function, and γ is called the drift. This terminology is influenced by the corresponding terms in the context of stable distributions, which correspond to the limiting case when $b \equiv 0$, see [34]. Let's see the univariate case of TS distribution first.

Example 1. *When $d = 1$, the unit sphere is $\mathbb{S}^0 = \{-1, 1\}$. Perhaps the simplest class of TS distributions corresponds to the case where $d = 1$, $\sigma(\{-1\}) = 0$, and $\gamma = 0$. In this case, taking $a = \sigma(\{1\}) > 0$ and $b = b(1) > 0$, the characteristic function reduces to*

$$\hat{\mu}(z) = \exp\left[\int_0^\infty (e^{izx} - 1)\frac{ae^{-bx}}{x^{1+\alpha}}\mathrm{d}x\right] = \exp\left[\Gamma(-\alpha)\left((b - iz)^\alpha - b^\alpha\right)\right] \tag{2.2}$$

for $z \in \mathbb{R}$. We refer to such distributions as simple tempered stable (STS) distributions

and we write $\mu = \mathrm{STS}_\alpha(a, b)$. *These distributions were first introduced in Tweedie (1984) [13] and are sometimes called Tweedie distributions.*

We now give several useful facts about TS distributions. The distribution $\mathrm{TS}_\alpha(\sigma, b, \gamma)$ is infinitely divisible with Lévy measure

$$L(B) = \int_{\mathbb{S}^{d-1}} \int_0^\infty 1_B(sx) x^{-1-\alpha} e^{-b(s)x} \mathrm{d}x \sigma(\mathrm{d}s), \quad B \in \mathcal{B}(\mathbb{R}^d). \tag{2.3}$$

See [35] for an overview of the theory of infinitely divisible distributions. And they have the following property.

Proprtion 1. *If* $X \sim \mathrm{TS}_\alpha(\sigma, b, \gamma)$ *and* $Y = cX + d$ *for some* $c > 0$ *and* $d \in \mathbb{R}^d$, *then*

$$Y \sim \mathrm{TS}_\alpha(c^\alpha \sigma, c^{-1} b, c\gamma + d).$$

Proof. Let's begin with the characteristic function of Y,

$$\begin{aligned}
\mathrm{E}[e^{i\langle z,\ Y\rangle}] &= \mathrm{E}[e^{i\langle z, cX+d\rangle}] = \mathrm{E}[e^{i\langle cz+d,\ X\rangle}] \\
&= \exp\left[i \langle \gamma, cz + d\rangle + \int_{\mathbb{S}^{d-1}} \int_0^\infty \left(e^{it\langle cz,\ s\rangle} - 1\right) t^{-1-\alpha} e^{-tb(s)} \mathrm{d}t \sigma(\mathrm{d}s) \right] \\
&= \exp\left[i \langle c\gamma + d, z\rangle + \int_{\mathbb{S}^{d-1}} \int_0^\infty \left(e^{iu\left\langle z,\ \frac{cs}{|cs|}\right\rangle} - 1\right) u^{-1-\alpha} e^{\frac{-ub(s)}{|cs|}} \mathrm{d}u\ |cs|^\alpha\ \sigma(\mathrm{d}s) \right] \\
&= \exp\left[i \langle c\gamma + d, z\rangle + \int_{\mathbb{S}^{d-1}} \int_0^\infty \left(e^{iu\langle z,\ s\rangle} - 1\right) u^{-1-\alpha} e^{\frac{-ub(s)}{c}} \mathrm{d}u\ c^\alpha\ \sigma(\mathrm{d}s) \right],
\end{aligned}$$

where the third line by change of variable, let $u = t|cs|$, then $\mathrm{d}u = \mathrm{d}t|cs|$. By the last line, we get $Y \sim \mathrm{TS}_\alpha(c^\alpha \sigma, c^{-1} b(s), c\gamma + d)$.

□

Next, if $X \sim \mathrm{TS}_\alpha(\sigma, b, \gamma)$ and $\inf_{s \in \mathbb{S}^{d-1}} b(s) > 0$, then, by Corollary 25.8 in [35], all moments of the distribution exist. Followed that, let's $\hat{\mu}(z) = \exp\{C_\mu(z)\}$, where

$$C_\mu(z) = i \langle \gamma, z\rangle + \int_{S^{d-1}} \Gamma(-\alpha)((b(s) - i\langle s, z\rangle)^\alpha - b^\alpha(s))\sigma(ds),$$

and we call C_μ be cumulant generating function.

Since $X \sim \mu$ and the cumulant determine the moments, we let c_1 represent the first cumulant, which is $\mathrm{E}(X_j)$; c_2 represent the second cumulant, which is $Var(X_j)$; c_3 represent the third cumulant, which is the third central moment of $\mathrm{E}[(X_j - \mathrm{E}(X_j))^3]$; c_4 represent the fourth cumulant, which is $\mathrm{E}[(X_j - \mathrm{E}(X_j))^4] - 3\mathrm{E}^2[(X_j - \mathrm{E}(X_j))^2]$ and c_{11} represent the covariance between X_i and X_j, where $i, j \in \{1, 2, \ldots, d\}$. By [36], skewness $\gamma_1(X_j) = \dfrac{c_3}{c_2^{3/2}}$ and the kurtosis $\gamma_2(X_j) = \dfrac{c_4}{c_2^2}$. We then give the general formula of them.

Proprtion 2. *Let $X \sim \mu$, $i, j \in \{1, 2, \ldots, d\}$ then*

- *The expectation of X_j*

$$c_1 = \mathrm{E}[X_j] = \gamma_j + \Gamma(1-\alpha) \int_{\mathbb{S}^{d-1}} b^{\alpha-1}(s) s_j \sigma(\mathrm{d}s). \tag{2.4}$$

- *The variance of X_j*

$$c_2 = \mathrm{Var}[X_j] = \Gamma(2-\alpha) \int_{S^{d-1}} b^{\alpha-2}(s) s_j^2 \sigma(\mathrm{d}s). \tag{2.5}$$

- *The third central moment of X_j*

$$c_3 = \mathrm{E}[(X_j - \mathrm{E}(X_j))^3] = \Gamma(3-\alpha) \int_{S^{d-1}} b^{\alpha-3}(s) s_j^3 \sigma(\mathrm{d}s) \tag{2.6}$$

- *The fourth cumulan of X_j*

$$\begin{aligned} c_4 &= \mathrm{E}[(X_j - \mathrm{E}(X_j))^4] - 3\mathrm{E}^2[(X_j - \mathrm{E}(X_j))^2] \\ &= \Gamma(4-\alpha) \int_{S^{d-1}} b^{\alpha-4}(s) s_j^4 \sigma(\mathrm{d}s) \end{aligned} \tag{2.7}$$

- *The Covariance between X_i and X_j*

$$c_{11} = \mathrm{Cov}(X_i, X_j) = \Gamma(2-\alpha) \int_{\mathbb{S}^{d-1}} b^{\alpha-2}(s) s_i s_j \sigma(\mathrm{d}s) \tag{2.8}$$

- *The skewness of X_j*

$$\begin{aligned} \gamma_1(X_j) &= \frac{c_3}{c_2^{3/2}} \\ &= \frac{2-\alpha}{[\Gamma(2-\alpha)]^{1/2}} \frac{\int_{S^{d-1}} b^{\alpha-3}(s) s_j^3 \sigma(\mathrm{d}s)}{\left[\int_{S^{d-1}} b^{\alpha-2}(s) s_j^2 \sigma(\mathrm{d}s)\right]^{3/2}} \end{aligned} \tag{2.9}$$

- *The kurtosis of X_j*

$$\begin{aligned} \gamma_2(X_j) &= \frac{c_4}{c_2^2} \\ &= \frac{(2-\alpha)(3-\alpha)}{\Gamma(2-\alpha)} \frac{\int_{S^{d-1}} b^{\alpha-4}(s) s_j^4 \sigma(\mathrm{d}s)}{\left[\int_{S^{d-1}} b^{\alpha-2}(s) s_j^2 \sigma(\mathrm{d}s)\right]^2} \end{aligned} \tag{2.10}$$

In principle, if there are parametric forms for σ and b, then one could use the method of cumulant matching to fit the parameters. Instead, we propose a methodology based on a discretization of σ. We will give the detailed information in the following sections.

Let's begin by recalling the following definition. We say that a Borel measure σ on $\mathbb{S}^{d-1}$ is full if there are d linearly independent vectors in its support. We now give several statements that are equivalent to this.

Lemma 1. *Let σ be a finite Borel measure on $\mathbb{S}^{d-1}$ and, for $\beta > 0$, let $u_\beta = \inf_{\xi \in \mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |\langle s, \xi \rangle|^\beta \sigma(\mathrm{d}s)$. The following statements are equivalent:*

1. σ is full;

2. $u_\beta > 0$ for every $\beta > 0$;

3. $u_\beta > 0$ for some $\beta > 0$.

We now give our main approximation result.

Theorem 1. *Fix $\alpha \in (0,1)$ and let $\mu = \mathrm{TS}_\alpha(\sigma, b, \gamma)$. Assume that the function b has an upper bound and that the set of its discontinuities has σ measure 0. If σ is full, then μ has a pdf p. Further, in this case, for any $\varepsilon > 0$, there is a finite measure σ^* with a finite support such that $\mu^* = \mathrm{TS}_\alpha(\sigma^*, b, \gamma)$ has a pdf p^* satisfying*

$$\sup_{x \in \mathbb{R}^d} |p(x) - p^*(x)| \leq \varepsilon.$$

This is a special case of a more general result, see Section 6. The theorem suggests that, for practical purposes, one just needs to work with TS distributions where the measure σ has finite support. In this case, assuming that the support has k elements, there exist $a_1, a_2, \ldots, a_k > 0$ and $s_1, s_2, \ldots, s_k \in \mathbb{S}^{d-1}$ such that

$$\sigma = \sum_{j=1}^{k} a_j \delta_{s_j}, \tag{2.11}$$

where δ_{s_j} is a point mass at s_j. We now give a simple interpretation of the corresponding TS distributions.

Theorem 2. *Fix $\alpha \in (0,1)$ and let $\mu = \mathrm{TS}_\alpha(\sigma, b, \gamma)$, where σ is as in (2.11). If $X_1, X_2, ..., X_k$ are independent random variables with $X_j \sim \mathrm{STS}_\alpha(a_j, b_j)$, where $b_j = b(s_j)$, and*

$$Y = \gamma + X_1 s_1 + X_2 s_2 + ... + X_k s_k, \tag{2.12}$$

then $Y \sim \mu$.

Proof. To get Y's characteristic function, let's $X_1, X_2..., X_n$ are independent random variables following $\mu_{X_j}(z) = \mathrm{STS}_\alpha(a_j, b_j)$, and $s_1, s_2, ..., s_n \in S^{d-1}$. Then, $Y = \gamma + X_1 s_1 + X_2 s_2 + ... + X_n s_n$, and its characteristic is

$$\hat{\mu}_Y(z) = E[e^{iz\langle Y, s\rangle}] = E[e^{i\langle \gamma, z\rangle + \sum_{j=1}^{n} \langle X_j s_j, z\rangle}]$$

$$
\begin{aligned}
&= e^{i\langle\gamma,z\rangle}\prod_{j=1}^{n} E\left[e^{i\langle X_j s_j,z\rangle}\right] \\
&= \exp\left[i\langle\gamma,z\rangle + \int_0^\infty (e^{i\langle s_1,z\rangle x}-1)\frac{a_1 e^{-b_1 x}}{x^{1+\alpha}}\mathrm{d}x + \ldots\right. \\
&\left. + \int_0^\infty (e^{i\langle s_n,z\rangle x}-1)\frac{a_n e^{-b_n x}}{x^{1+\alpha}}\mathrm{d}x\right] \\
&= exp\left[[i\langle\gamma,z\rangle + \sum_{j=1}^{n}\int_0^\infty (e^{i\langle s_j,z\rangle x}-1)\frac{a_j e^{-b_j x}}{x^{1+\alpha}}\mathrm{d}x\right] \\
&= exp\left[i\langle\gamma,z\rangle + \int_{\mathbb{S}^{d-1}}\int_0^\infty (e^{i\langle s,z\rangle x}-1)\frac{e^{-b(s)x}}{x^{1+\alpha}}\mathrm{d}x\sigma(\mathrm{d}s)\right],
\end{aligned}
$$

which is equalivent to equation (2.1) with the given σ, as required. □

2.2 Discretization in the bivariate base

In practice, to use the result of Theorem 1, we need a systematic approach for discretizing the spectral measure. In this section, we discuss such an approach in the bivariate case where $d = 2$. In this case $\mathbb{S}^1$, the unit circle in $\mathbb{R}^2$, is isomorphic to the interval $[0, 2\pi)$ through the bijection $\eta : [0, 2\pi) \mapsto \mathbb{S}^1$ given by $\eta(\theta) = (\cos(\theta), \sin(\theta))^T$. Further, every finite Borel measure σ on $\mathbb{S}^1$ is in one-to-one correspondence with a finite Borel measure σ' on $[0, 2\pi)$. Specifically, for $A \in \mathfrak{B}([0, 2\pi))$ and $B \in \mathfrak{B}(\mathbb{S}^1)$

$$
\sigma'(A) = \sigma\left(\eta(A)\right) \text{ and } \sigma(B) = \sigma'\left(\eta^{-1}(B)\right),
$$

where $\eta^{-1}(B)$ is the inverse image of B and $\eta(A)$ is the image of A (or, equivalently, the inverse image of the inverse function of η). For any complex valued Borel function $f : \mathbb{S}^1 \mapsto \mathbb{C}$, which is integrable with respect to σ the following change of variables formula holds

$$
\int_{\mathbb{S}^1} f(s)\sigma(\mathrm{d}s) = \int_{[0,2\pi)} f(\cos(\theta), \sin(\theta))\sigma'(\mathrm{d}\theta), \tag{2.13}
$$

Now, the characteristic function of $\mu = \mathrm{TS}_\alpha(\sigma, b, \gamma)$ can be written, for any $z =$

$(z_1, z_2) \in \mathbb{R}^2$, as

$$\exp\left[i\langle\gamma, z\rangle + \Gamma(-\alpha)\int_{[0,2\pi)}\left((b_*(\theta) - i\big(\cos(\theta)z_1 + \sin(\theta)z_2\big)\right)^\alpha - b_*^\alpha(\theta))\sigma'(\mathrm{d}\theta)\right],$$

where $b_*(\theta) = b(\eta(\theta)) = b(\cos(\theta), \sin(\theta))$.

This leads to a simple approach for discretization. First, fix k and divide the interval $[0, 2\pi)$ into k subintervals as follows. Fix $0 = d_0 < d_1 < \cdots < d_k = 2\pi$, let $\theta_1, \theta_2, \ldots, \theta_k \in (0, 2\pi)$ be any numbers with $d_{j-1} < \theta_j < d_j$, and let

$$a_j = \sigma'([d_{j-1}, d_j)) = \sigma(\eta([d_{j-1}, d_j))).$$

It follows that

$$\sigma'_k = \sum_{j=1}^{k} a_j \delta_{\theta_j} \text{ and } \sigma_k = \sum_{j=1}^{k} a_j \delta_{\eta(\theta_j)} \tag{2.14}$$

are discretization of σ' and σ, respectively. In this dissertation, we take the subintervals of $[0, 2\pi)$ to be evenly spaced and we take θ_j to be the midpoint of the subinterval. Thus, we take $d_j = 2\pi j/k$ and $\theta_j = (d_{j-1} + d_j)/2 = (2j-1)\pi/k$. Of course, in practice, if σ' has point masses, then it would make sense to keep at least the ones with larger weights. Then, after these have been extracted, one can discretize the rest using the procedure described above.

2.3 Example

Now, we give two specific examples in this section to show the results in Section 2.1 and 2.2.

Proprtion 3. *Let $X = (X_1, X_2) \sim \mathrm{TS}_\alpha(\sigma, b, \gamma)$, where $\sigma'(\mathrm{d}s) = f(\theta)\mathrm{d}\theta$ such that for*

$$f(\theta) = \frac{1}{2\pi}, \quad \theta \in [0, 2\pi).$$

This means that σ' is a uniform distribution on $[0, 2\pi)$. Under this case, we have the means are

$$\begin{aligned} \mathrm{E}(X_1) &= \gamma_1 + \frac{1}{2\pi}\Gamma(1-\alpha)\int_0^{2\pi} b_*^{\alpha-1}(\theta)\cos\theta \mathrm{d}\theta \\ \mathrm{E}(X_2) &= \gamma_2 + \frac{1}{2\pi}\Gamma(1-\alpha)\int_0^{2\pi} b_*^{\alpha-1}(\theta)\sin(\theta)\mathrm{d}\theta, \end{aligned}$$

the variance are

$$\begin{aligned} \mathrm{Var}(X_1) &= (1-\alpha)\Gamma(1-\alpha)\frac{1}{2\pi}\int_0^{2\pi} b_*^{\alpha-2}(s)\cos^2\theta \mathrm{d}\theta \\ \mathrm{Var}(X_2) &= (1-\alpha)\Gamma(1-\alpha)\frac{1}{2\pi}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\sin^2\theta \mathrm{d}\theta, \end{aligned}$$

the covariance between X_1 and X_2 is,

$$\mathrm{Cov}(X_1, X_2) = (1-\alpha)\Gamma(1-\alpha)\frac{1}{4\pi}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\sin 2\theta \mathrm{d}\theta.$$

Proof. Following Proprtion 2 and equation(2.13), we easy to get means, variances and covariance of X. Let's see the means first.

$$\begin{aligned} \mathrm{E}(X_1) &= \gamma_1 - \tfrac{1}{2\pi}\Gamma(-\alpha)\alpha\textstyle\int_0^{2\pi} b_*^{\alpha-1}(\theta)\cos(\theta)\mathrm{d}\theta \\ &= \gamma_1 + \tfrac{1}{2\pi}\Gamma(1-\alpha)\textstyle\int_0^{2\pi} b_*^{\alpha-1}(\theta)\cos(\theta)\mathrm{d}\theta, \end{aligned}$$

where the last equation follows $\Gamma(-\alpha) = -\frac{\Gamma(1-\alpha)}{\alpha}$. To get $\mathrm{E}(X_2)$ is similar.

Then, let's show the variances as follow,

$$\begin{aligned} \mathrm{Var}(X_1) &= \int_{\mathbb{S}^1} \Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(s)s_1^2\sigma(\mathrm{d}s) \\ &= \int_0^{2\pi} \Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(s)\cos^2\theta\frac{1}{2\pi}\mathrm{d}\theta \\ &= (1-\alpha)\Gamma(1-\alpha)\frac{1}{2\pi}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\cos^2\theta \mathrm{d}\theta \end{aligned}$$

Similar with $\mathrm{Var}(X_1)$, for $\mathrm{Var}(X_2)$, we have

$$\begin{aligned}\mathrm{Var}(X_2) &= \int_{\mathbb{S}^1} \Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(s)s_2^2\sigma(\mathrm{d}s)\\ &= \int_0^{2\pi} \Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(\theta)\sin^2\theta\frac{1}{2\pi}\mathrm{d}\theta\\ &= (1-\alpha)\Gamma(1-\alpha)\frac{1}{2\pi}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\sin^2\theta\mathrm{d}\theta.\end{aligned}$$

And then $\mathrm{Cov}(X_1, X_2)$ is

$$\begin{aligned}\mathrm{Cov}(X_1,X_2) &= \int_{S^{d-1}} \Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(s)s_1s_2\sigma(\mathrm{d}s)\\ &= \int_0^{2\pi} \Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(\theta)\cos\theta\sin\theta\frac{1}{2\pi}\mathrm{d}\theta\\ &= (1-\alpha)\Gamma(1-\alpha)\frac{1}{4\pi}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\sin 2\theta\mathrm{d}\theta,\end{aligned}$$

where the last equation follow $\Gamma(-\alpha) = -\frac{\Gamma(1-\alpha)}{\alpha}$ and $\sin 2\theta = 2\sin\theta\cos\theta$. □

Under this case, when we discretize measure σ, we get

$$\begin{aligned}a_j &= \sigma'([d_{j-1}, d_j)) = \sigma'([2\pi(j-1)/k, 2\pi j/k))\\ &= \int_{2\pi(j-1)/k}^{2\pi j/k} f(\theta)\mathrm{d}\theta = \int_{2\pi(j-1)/k}^{2\pi j/k}\frac{1}{2\pi}\mathrm{d}\theta\\ &= \frac{1}{k}.\end{aligned}$$

Next, let's see the other case.

Proprtion 4. *Let $X = (X_1, X_2) \sim \mathrm{TS}_\alpha(\sigma, b, \gamma)$, where $\sigma'(\mathrm{d}\theta) = f(\theta)\mathrm{d}\theta$ such that for some $\alpha^*, \beta^* > 0$*

$$f(\theta) = \frac{1}{C}\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}, \quad \theta\in[0,2\pi),$$

where $C = B(\alpha^*, \beta^*)(2\pi)^{\alpha^*+\beta^*-1}$ and $B(\cdot,\cdot)$ is the beta function. This means that σ' is a beta distribution on $[0, 2\pi)$. We denote this distribution by $\mathrm{Beta}(\alpha^*, \beta^*, 2\pi)$. *In*

this case, the means are

$$\begin{aligned}
\mathrm{E}(X_1) &= \gamma_1 + \frac{\Gamma(1-\alpha)}{C}\int_0^{2\pi} b_*^{\alpha-1}(\theta)\cos(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}d\theta \\
\mathrm{E}(X_2) &= \gamma_2 + \frac{\Gamma(1-\alpha)}{C}\int_0^{2\pi} b_*^{\alpha-1}(\theta)\sin(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}d\theta,
\end{aligned}$$

the variances are

$$\begin{aligned}
\mathrm{Var}(X_1) &= \frac{\Gamma(2-\alpha)}{C}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\cos^2(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}d\theta \\
\mathrm{Var}(X_2) &= \frac{\Gamma(2-\alpha)}{C}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\sin^2(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}d\theta,
\end{aligned}$$

and the covariance is

$$\mathrm{Cov}(X_1, X_2) = \frac{\Gamma(2-\alpha)}{2C}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\sin(2\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta.$$

Proof. Again, following Proprtion 2 and equation (2.13), we get means, variances and covariance of X. First, the means of X as follow,

$$\begin{aligned}
\mathrm{E}(X_1) &= \gamma_1 - \int_0^{2\pi}\Gamma(-\alpha)\alpha b_*^{\alpha-1}(\theta)\cos(\theta)\frac{\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}}{C}\mathrm{d}\theta \\
&= \gamma_1 - \frac{\Gamma(-\alpha)\alpha}{C}\int_0^{2\pi} b_*^{\alpha-1}(\theta)\cos(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta \\
&= \gamma_1 + \frac{\Gamma(1-\alpha)}{C}\int_0^{2\pi} b_*^{\alpha-1}(\theta)\cos(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta,
\end{aligned}$$

where the last equation follows $\Gamma(-\alpha) = -\frac{\Gamma(1-\alpha)}{\alpha}$. Similar to the $\mathrm{E}(X_1)$, $\mathrm{E}(X_2)$ is easy to get. Then, for the variances of X,

$$\begin{aligned}
\mathrm{Var}(X_1) &= \int_{\mathbb{S}^1}\Gamma(-\alpha)\alpha(\alpha-1)b^{\alpha-2}(s)s_1^2\sigma(\mathrm{d}s) \\
&= \int_0^{2\pi}\Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(s)\cos^2\theta\frac{\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}}{C}\mathrm{d}\theta
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\Gamma(-\alpha)\alpha(\alpha-1)}{C}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\cos^2(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta \\
&= \frac{\Gamma(2-\alpha)}{C}\int_0^{2\pi} b^{\alpha-1}(\theta)\cos^2(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta.
\end{aligned}
$$

Similar with $\mathrm{Var}(X_1)$, for $\mathrm{Var}(X_2)$,

$$
\begin{aligned}
\mathrm{Var}(X_2) &= \int_{\mathbb{S}^1}\Gamma(-\alpha)\alpha(\alpha-1)b^{\alpha-2}(s)s_2^2\sigma(\mathrm{d}s) \\
&= \int_0^{2\pi}\Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(s)\sin^2\theta\frac{\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}}{C}\mathrm{d}\theta \\
&= \frac{\Gamma(-\alpha)\alpha(\alpha-1)}{C}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\sin^2(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta \\
&= \frac{\Gamma(2-\alpha)}{C}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\sin^2(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta.
\end{aligned}
$$

Last, for the covariance of X_1 and X_2,

$$
\begin{aligned}
\mathrm{Cov}(X_1,X_2) &= \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)b^{\alpha-2}(s)s_1s_2\sigma(\mathrm{d}s) \\
&= \int_0^{2\pi}\Gamma(-\alpha)\alpha(\alpha-1)b_*^{\alpha-2}(\theta)\cos(\theta)\sin(\theta)\frac{\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}}{C}\mathrm{d}\theta \\
&= \frac{\Gamma(-\alpha)\alpha(\alpha-1)}{C}\int_0^{2\pi} b_*^{\alpha-2}(\theta)\cos(\theta)\sin(\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta \\
&= \frac{\Gamma(2-\alpha)}{2C}\int_0^{2\pi} b^{\alpha-2}(\theta)\sin(2\theta)\theta^{\alpha^*-1}(2\pi-\theta)^{\beta^*-1}\mathrm{d}\theta
\end{aligned}
$$

as required. □

Under this case, when we discretize, similarly, we get

$$
\begin{aligned}
a_j &= \sigma'([d_{j-1},d_j)) = \sigma'([2\pi(j-1)/k, 2\pi j/k)) \\
&= \int_{2\pi(j-1)/k}^{2\pi j/k} f(\theta)\mathrm{d}\theta = F\left(\frac{j}{k}\right) - F\left(\frac{j-1}{k}\right),
\end{aligned}
$$

where F is the cumulative distribution function (cdf) of the Beta distribution on $[0,1)$ with shape parameters α^* and β^*.

CHAPTER 3: SIMULATION METHODOLOGY

3.1 Simulation methodology of tempered stable variable

In this section, we discuss our methodology for simulation. When σ has finite support, following Theorem 2, an exact simulation method is provided by (2.12). At this time, we only need to have an efficiently method to simulate $\mathrm{STS}_\alpha(a_j, b_j)$. In the literature, one approach is the algorithm on page 11 of [37] or a simple rejection sampling algorithm, which is given in e.g. Algorithm 0 of [38]. Simple rejection sampling algorithm is easy to achieve, however, the probability of rejecting on a given iteration is $e^{-a_j\Gamma(1-\alpha)b_j^\alpha/\alpha}$, which is not efficient when a_j or b_j is large. To deal with this, a double rejection sampling approach was developed in [37], where the probability of rejection is bounded away from 0 for all choices of the parameters. This algorithm was further optimized in [39]. This optimized version is implemented in the `retstable` method of the "copula" package for the statistical software R. In this dissertation, we choose the double rejection sampling approach to simulate $\mathrm{STS}_\alpha(a_j, b_j)$.

Now, we have known how to simulate $\mathrm{STS}_\alpha(a_j, b_j)$ efficiently, and then based on (2.12), when σ has a finite support, an exact simulation method is provided. When σ does not have a finite support, Theorem 1 implies that there exists a σ^* with a finite support such that $\mathrm{TS}_\alpha(\sigma, b, \gamma) \approx \mathrm{TS}_\alpha(\sigma^*, b, \gamma)$. Thus, we can approximately simulate $\mathrm{TS}_\alpha(\sigma, b, \gamma)$ by using (2.12) to simulate $\mathrm{TS}_\alpha(\sigma^*, b, \gamma)$. We refer to this methodology as the discretization and simulation (DS) method.

The only other simulation approach that we have seen in the literature is an approximate method, which is based on truncating an infinite series representation. This representation is given, for a more general class of models, in Theorem 5.1 of [19]. Let's see that how to apply this algorithm in our case. According to another

definition in [19], the Levy measure could write as

$$M(B) = \int_{\mathbb{S}^{d-1}} \int_0^\infty 1_B(sx) q(x,s) x^{-1-\alpha} \mathrm{d}x \sigma(\mathrm{d}s), \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Comparing with equation (2.3), we get $q(x,s) = e^{-b(s)x}$ and

$$q(x,s) = \int_0^\infty e^{-xu} Q(\mathrm{d}u|s), \quad Q(\mathrm{d}u|s) = \delta_{b(s)}(\mathrm{d}u),$$

where $\{Q(.|s)\}_{s\in\mathbb{S}^{d-1}}$ is a measure on $(0,\infty)$. Following equation (2.4) in [19], we defined Q be a Borel measure on $\mathbb{R}^d$ given by

$$\begin{aligned} Q(B) &= \int_{\mathbb{S}^{d-1}} \int_0^\infty 1_B(sx) Q(\mathrm{d}x|s) \sigma(\mathrm{d}s) \\ &= \int_{\mathbb{S}^{d-1}} \int_0^\infty 1_B(sx) \delta_{b(s)}(\mathrm{d}u) \sigma(\mathrm{d}s) \\ &= \int_{\mathbb{S}^{d-1}} 1_B(sb(s)) \sigma(\mathrm{d}s), \end{aligned}$$

where $\sigma(\mathrm{d}s) = c\sigma_1(\mathrm{d}s)$, $\sigma_1(\mathrm{d}s)$ is a probability measure and $c > 0$. Hence, we have

$$Q(B) = c \int_{\mathbb{S}^{d-1}} 1_B(sb(s)) \sigma_1(\mathrm{d}s). \tag{3.1}$$

Now, let $\{v_j\}$ be an independent sequences of independent and identically distributed (iid) random vectors in $\mathbb{R}^d$ with the distribution of $\dfrac{Q}{|\sigma|}$, where $|\sigma| = |c\sigma_1| = c\sigma(\mathbb{S}^{d-1}) = c$. By equation (3.1), we have

$$\begin{aligned} \frac{Q(B)}{|\sigma|} &= \frac{c \int_{\mathbb{S}^{d-1}} 1_B(sb(s)) \sigma_1(\mathrm{d}s)}{c} \\ &= \int_{\mathbb{S}^{d-1}} 1_B(sb(s)) \sigma_1(\mathrm{d}s). \end{aligned} \tag{3.2}$$

Further, $|v_j| = |b(s_j)s_j| = |b(s_j)|$ and $\dfrac{v_j}{|v_j|} = S_j \sim \dfrac{\sigma}{\sigma(\mathbb{S}^{d-1})}$, which is iid random variable. Also, let $\{e_j\}$, $\{e'_j\}$, $\{u_j\}$ be iiid sequence of random variables such that

$e_j \sim \text{Exp}(1)$, $e'_j \sim \text{Exp}(1)$, $u_j \sim U(0,1)$ and let $\gamma_j = e'_1 + e'_2 + \cdots + e'_j$. If $X \sim \text{TS}_\alpha(\sigma, b, 0)$, then

$$\begin{aligned} X &\stackrel{d}{=} \sum_{j=1}^{\infty} \min\left\{ \left(\frac{\alpha\gamma_j}{|\sigma|}\right)^{-1/\alpha}, \frac{e_j u_j^{1/\alpha}}{|v_j|} \right\} \frac{v_j}{|v_j|} \\ &= \sum_{j=1}^{\infty} \min\left\{ \left(\frac{\alpha\gamma_j}{\sigma(\mathbb{S}^{d-1})}\right)^{-1/\alpha}, \frac{e_j u_j^{1/\alpha}}{|b(s_j)|} \right\} S_j. \end{aligned} \tag{3.3}$$

Truncating this infinite sum at some large value k gives an approximate simulation technique. We refer to this methodology as Rosiński's method after the author of [19]. Unlike the DS method, this method is never exact.

We now use simulation to compare the performance of the DS method and Rosiński's method in the bivariate case. Our approach is as follows. First, we choose an integer k. For the DS method, we use the approach described in Section 2.2 to discretize σ into σ_k which has k elements in its support. From the corresponding TS distribution, we simulate N observations. Using these, we estimate the means m_1, m_2 of both components, the variances σ_1^2, σ_2^2 of both components, and the covariance between the components σ_{12} by using the empirical means $\bar{x}_1, \bar{x}_2$, empirical variances s_1^2, s_2^2, and the empirical covariance s_{12}. Also, followed by section 2.3, we can get the value of $m_1,\ m_2,\ \sigma_1^2,\ \sigma_2^2$ and σ_{12}. We then calculate the absolute errors between theoretical values and empirical values, which include mean, variance in each component, and covariance. Further, we calculate the error in the approximation by

$$E_k = \sqrt{|m_1 - \bar{x}_1|^2 + |m_2 - \bar{x}_2|^2 + |\sigma_1^2 - s_1^2|^2 + |\sigma_2^2 - s_2^2|^2 + |\sigma_{12} - s_{12}|^2}.$$

We follow the same procedure for Rosiński's method, except that now, for simulations we use the infinite sum in (3.3) truncated at k, i.e. we only take the first k terms in that sum. We similarly calculate E_k in this case.

3.2 Simulation results

In this section, we use each method to generate a dataset of bivariate TS variables with $N = 10000$. Then we compare Rosinski's method with the DS method by the absolute error between theoretical values and empirical values, which include mean, variance in each component, and covariance. Also, we compare with their E_k. We are interested in how large k will let errors converge and approach to 0. We set $\alpha = 0.6$, $b(s) \equiv 1$, $\gamma = 0$, and $\sigma' = \text{Beta}(\alpha^*, \beta^*, 2\pi)$ for several choices of α^* and β^*. The formulas for the means, variances, and covariance are given in Section 2.3. In this case, for Rosiński's method, we can simulate the S_i's by taking $S_i = (\cos(\theta_i), \sin(\theta_i))$, where $\theta_i \sim \text{Beta}(\alpha^*, \beta^*, 2\pi)$. We present results for four cases: $\alpha^* = \beta^* = 1$, $\alpha^* = 0.5$ and $\beta^* = 1$, $\alpha^* = 2$ and $\beta^* = 2$, and $\alpha^* = 2$ and $\beta^* = 5$. We also considered other choices for the parameters, but the results were similar. In all cases, the results of k from 1 to 100 are plotted in each plot. And in all the figures, the first row gives the absolute error of their mean, variance in the first component, and the covariance between two components, and the second row gives their mean, variance in the second component, and E_k. Then, let's see these cases one by one.

The first case is under $\alpha^* = \beta^* = 1$, which corresponds to the uniform distribution. In Figure 3.1, the performance of DS method and Rosinski's method are similar when we focus on the absolute error of mean and covariance, while, under the first component, DS method has larger errors for small values of k. For the absolute error of variances and E_k, DS method is better than Rosinski's as DS method has a faster speed of convergence. And both of these two methods' amount of error is close to zero at $k = 100$.

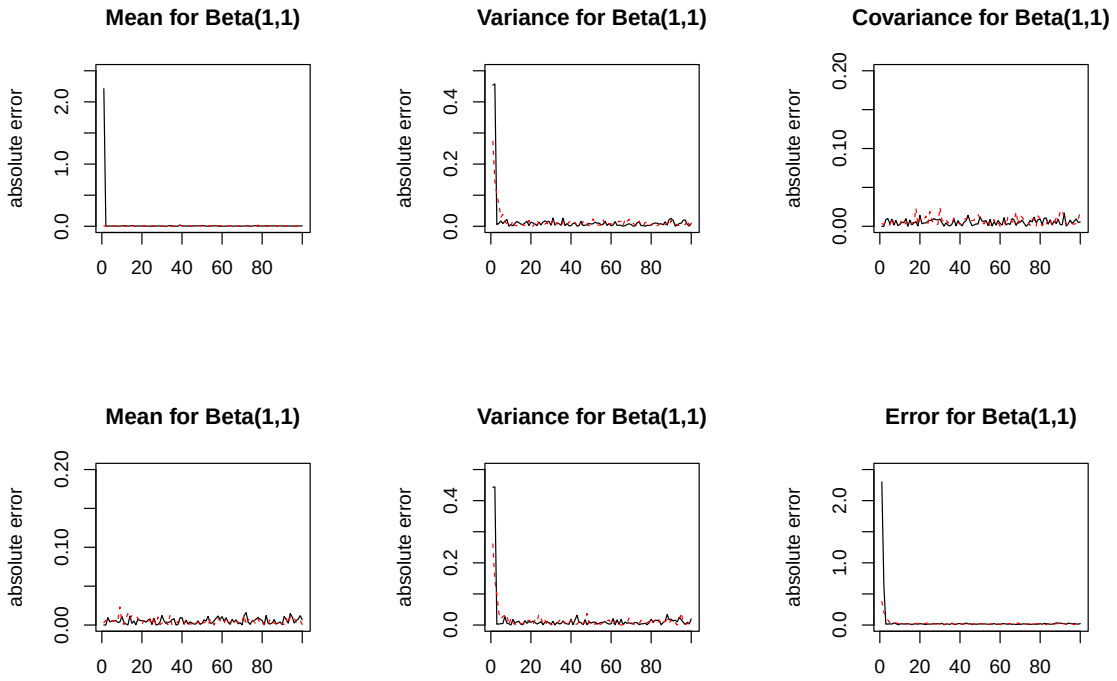


Figure 3.1: The solid (black) line is the error in the DS method, while the dashed (red) line is the error in Rosiński method. The x-axis represents k, the number of terms in the sum

For the second case, we consider is $\alpha^* = 0.5$ and $\beta^* = 1$. In Figure 3.2, the performance of DS method is better than Rosinski's under E_k and both two components of mean. Although their speed of convergence looks similar, Rosiński method has less error for small values of k, but more for large values. However, when $k = 100$, the amount of error is quite small in the DS method and is quite a bit higher for Rosiński's method. For the absolute error of covariance, DS method is a little worse than Rosiński method at the beginning, and they are similar after $k = 30$. And both these two methods have the similar results of variances in each component.

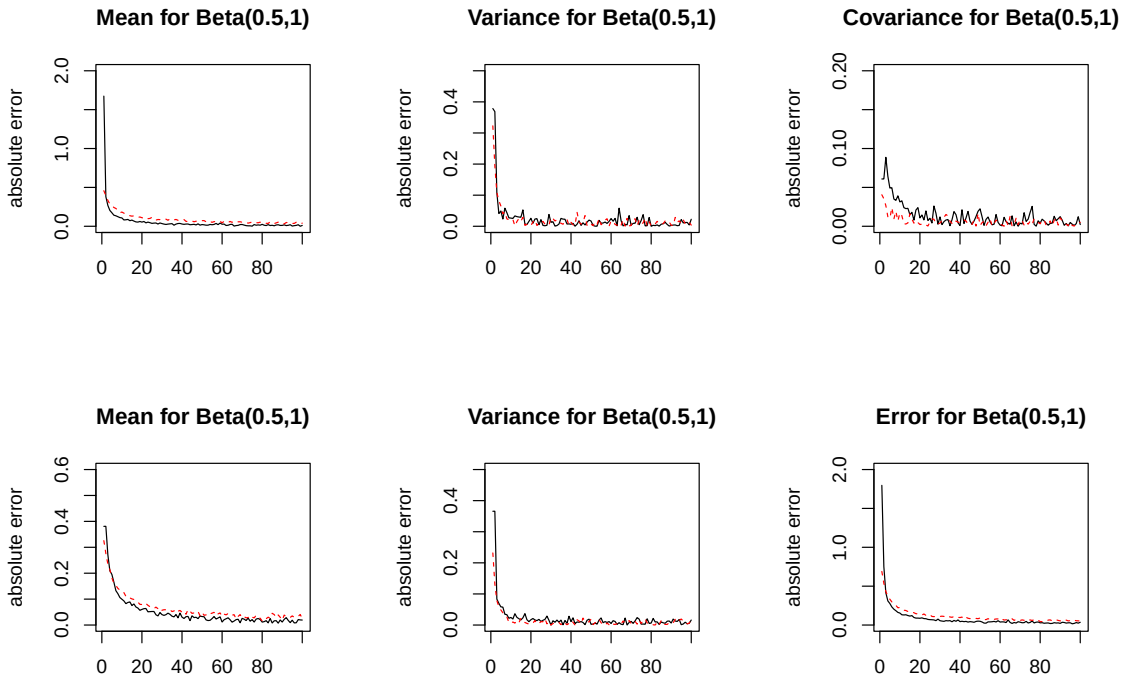


Figure 3.2: The solid (black) line is the error in the DS method, while the dashed (red) line is the error in Rosiński method. The x-axis represents k, the number of terms in the sum

For the third case, we consider $\alpha^* = 2$ and $\beta^* = 2$. From Figure 3.3, the plot of E_k and the absolute error of mean which is in the first component are similar to the second case, but DS method is a little worse in the second component at the beginning. And then they have the similar performance until $k = 60$, at which the error is almost close to 0. The absolute error of covariance is similar to the mean in the second component. And the absolute error of variances, DS method provides a slightly faster speed of convergence than Rosinski's and their other performances are similar in each component.

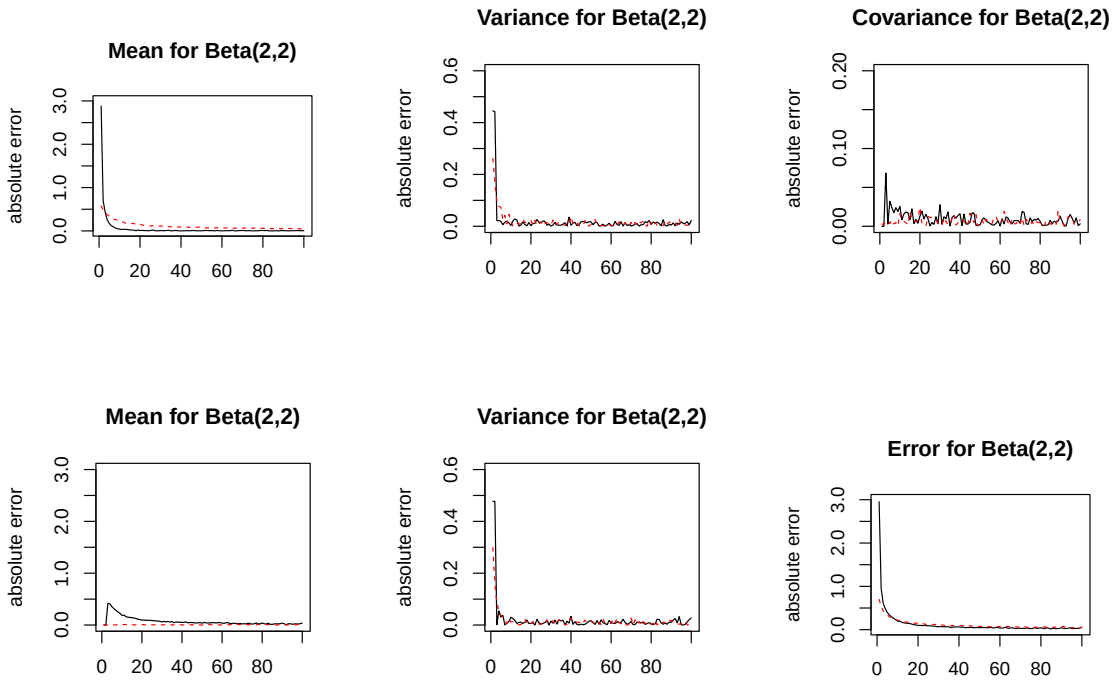


Figure 3.3: The solid (black) line is the error in the DS method, while the dashed (red) line is the error in Rosiński method. The x-axis represents k, the number of terms in the sum

The last case is under $\alpha^* = 2$ and $\beta^* = 5$. By the Figure 3.4, the performance of DS is much better than Rosiński method for the absolute error of mean in the second component since the solid (black) line has a faster speed of convergence and less error for large k. But the contrary performances are shown in the first component of the absolute error of mean. While DS method has the similar trend as the second component, Rosiński method gives error is less than DS at the beginning, which has closed to 0. Next, let's see the absolute error of variance. In the second component, both two methods have similar results. But in the first component, less error of Rosiński method at the beginning and its speed of convergence is faster than DS. Further, both two methods have the similar results for the absolute error of covariance except for the beginning. Although we cannot distinguish which method's performance is

better by means, variances, and covariance, from E_k, we could get the same result as before, which is that DS method has less error than Rosiński method for large k.

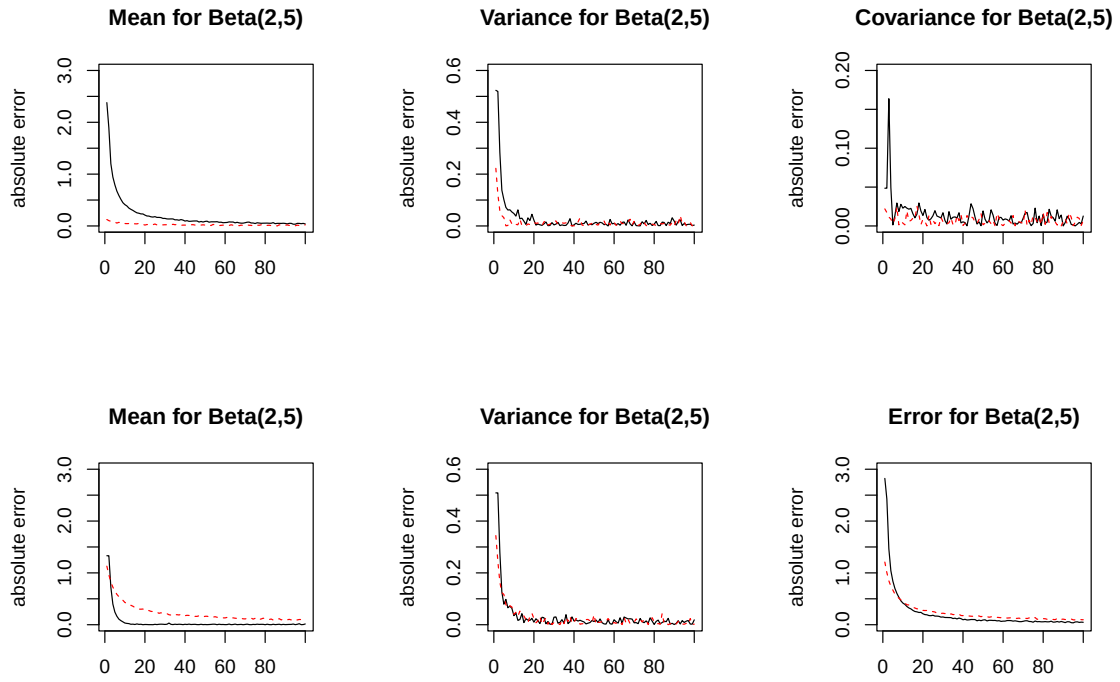


Figure 3.4: The solid (black) line is the error in the DS method, while the dashed (red) line is t he error in Rosiński method. The x-axis represents k, the number of terms in the sum

Overall, we could get similar patterns of E_k for all parameter values. Rosiński's method has less error for small values of k, but more for large values. We can also see that at $k = 100$ the amount of error is fairly small in the DS method for all four distributions, and is quite a bit higher for Rosiński's method in all cases except when $\alpha^* = 1, \beta^* = 1$, i.e. the uniform case. Of course, E_k is a limited way of understanding the error in the approximation, but it nevertheless gives a baseline for comparison.

CHAPTER 4: PARAMETER ESTIMATION AND DATA ANALYSIS

4.1 Methodology

In this section, we give the methodology to solve the parameter estimation problem. Our approach is based on the method of characteristic functions similar to that developed for fitting multivariate stable distributions in [29]. Note that $\mathrm{TS}_\alpha(\sigma, b, \gamma)$ is a semiparametric model since σ is an infinite dimensional parameter. To deal with this, we approximate σ by a finitely supported measure σ_k^* of the form (2.11). Here k is a tuning parameter and the directions $s_1, s_2, \ldots, s_k$ are chosen from the unit sphere $\mathbb{S}^{d-1}$ or some relevant subset of it. For instance if all components of the data are known to be nonnegative, say if we are directly modeling prices, then the direction would be chosen to lie in the first quadrant. The remaining parameters $a = (a_1, a_2, \ldots, a_k)$, $b = (b_1, b_2, \ldots, b_k)$, $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_d)$, and α need to be fit using the data. Thus we must estimate $2k + d + 1$ parameters. For simplicity, we sometimes denote the $(2k + d + 1)$-dimensional vector of all of these parameters by θ.

Our approach to parameter estimation is to find the parameters that minimize the distance between the characteristic function of the multivariate TS distribution and the empirical characteristic function. Specifically, let $x_1, x_2, \ldots, x_n$ be a random sample and note that each $x_i \in \mathbb{R}^d$. The empirical characteristic function is given by

$$\hat{\mu}_E(z) = \frac{1}{n}\sum_{j=1}^{n} e^{i\langle z, X_j\rangle} \quad = \quad \frac{1}{n}\sum_{i=1}^{n} \cos(\langle z, X_j\rangle) + i\frac{1}{n}\sum_{i=1}^{n} \sin(\langle z, X_j\rangle), \quad z \in \mathbb{R}^d.$$

And then we need to find the real and imaginary parts of the multivariate TS distribution's characteristic function. From (2.1) and Lemma 13, we get another from of this characteristic function.

Proprtion 5. *Let $z \in \mathbb{R}^d$, $X \sim \mathrm{TS}_\alpha(\sigma, b, \gamma)$, and there $\exists$ a finitely support measure σ^* shuch that $\mathrm{TS}_\alpha(\sigma, b, \gamma) \approx \mathrm{TS}_\alpha(\sigma^*, b, \gamma)$, then by equation (2.11), we have $\theta = (\alpha,\ b,\ a,\ \gamma)$ and the multivariate TS distribution's characteristic function becomes*

$$\hat{\mu}_\theta(z) = \exp\left\{A_{\alpha,b,a}(z)\right\}\Big(\cos\left(B_{\alpha,b,a,\gamma}(z)\right) + i\sin\left(B_{\alpha,b,a,\gamma}(z)\right)\Big),$$

where

$$A_{\alpha,b,a}(z) = \Gamma(-\alpha)\sum_{j=1}^{k} a_j\left((b_j^2 + \langle s_j, z\rangle^2)^{\alpha/2}\cos\left(\alpha\arctan\left(\frac{\langle s_j, z\rangle}{b_j}\right)\right) - b_j^\alpha\right)$$

and

$$B_{\alpha,b,a,\gamma}(z) = \langle\gamma, z\rangle - \Gamma(-\alpha)\sum_{j=1}^{k} a_j(b_j^2 + \langle s_j, z\rangle^2)^{\alpha/2}\sin\left(\alpha\arctan\left(\frac{\langle s_j, z\rangle}{b_j}\right)\right).$$

Proof. Based on (2.1), we have

$$\begin{aligned}
\hat{\mu}_\theta(z) &= \exp\left\{i\langle\gamma, z\rangle + \Gamma(-\alpha)\int_{\mathbb{S}^{d-1}}\left((b(s) - i\langle s, z\rangle)^\alpha - b^\alpha(s)\right)\sigma_k^*(\mathrm{d}s)\right\} \\
&= \exp\left\{i\,\langle\gamma, z\rangle + \Gamma(-\alpha)\sum_{j=1}^{k} a_j((b_j - i\langle s_j, z\rangle)^\alpha - b_j^\alpha)\right\} \\
&= \exp\Bigg(i\,\langle\gamma, z\rangle + \Gamma(-\alpha)\sum_{j=1}^{k} a_j\Bigg((b_j^2 + \langle s_j, z\rangle^2)^{\frac{\alpha}{2}}\cos\left(\alpha\arctan\left(\frac{\langle s_j, z\rangle}{b_j}\right)\right) \\
&\quad -b_j^\alpha - i(b_j^2 + \langle s_j, z\rangle^2)^{\frac{\alpha}{2}}\sin\left(\alpha\arctan\left(\frac{\langle s_j, z\rangle}{b_j}\right)\right)\Bigg)\Bigg) \\
&= \exp\Bigg(\Gamma(-\alpha)\sum_{j=1}^{k} a_j(b_j^2 + \langle s_j, z\rangle^2)^{\frac{\alpha}{2}}\Big(\cos\left(\alpha\arctan\left(\frac{\langle s_j, z\rangle}{b_j}\right)\right) - b_j^\alpha\Big) \\
&\quad +i\left(\langle\gamma, z\rangle - \Gamma(-\alpha)\sum_{j=1}^{k} a_j(b_j^2 + \langle s_j, z\rangle^2)^{\frac{\alpha}{2}}\sin\left(\alpha\arctan\left(\frac{\langle s_j, z\rangle}{b_j}\right)\right)\right)\Bigg) \\
&= \exp\left(\Gamma(-\alpha)\sum_{j=1}^{k} a_j(b_j^2 + \langle s_j, z\rangle^2)^{\frac{\alpha}{2}}\Big(\cos\left(\alpha\arctan\left(\frac{\langle s_j, z\rangle}{b_j}\right)\right) - b_j^\alpha\Big)\right)
\end{aligned}$$

$$* \left(\cos \left[\langle \gamma, z \rangle - \Gamma(-\alpha) \sum_{j=1}^{k} a_j (b_j^2 + \langle s_j, z \rangle^2)^{\frac{\alpha}{2}} \sin \left(\alpha \arctan \left(\frac{\langle s_j, z \rangle}{b_j} \right) \right) \right] \right.$$
$$\left. + i \sin \left[\langle \gamma, z \rangle - \Gamma(-\alpha) \sum_{j=1}^{k} a_j (b_j^2 + \langle s_j, z \rangle^2)^{\frac{\alpha}{2}} \sin \left(\alpha \arctan \left(\frac{\langle s_j, z \rangle}{b_j} \right) \right) \right] \right),$$

where the third equation is followed Lemma 13. Then, we let

$$A_{\alpha,b,a}(z) \;=\; \Gamma(-\alpha) \sum_{j=1}^{k} a_j (b_j^2 + \langle s_j, z \rangle^2)^{\frac{\alpha}{2}} \left(\cos \left(\alpha \arctan \left(\frac{\langle s_j, z \rangle}{b_j} \right) \right) - b_j^{\alpha} \right)$$

and

$$B_{\alpha,b,a,\gamma}(z) = \langle \gamma, z \rangle - \Gamma(-\alpha) \sum_{j=1}^{k} a_j (b_j^2 + \langle s_j, z \rangle^2)^{\frac{\alpha}{2}} \sin \left(\alpha \arctan \left(\frac{\langle s_j, z \rangle}{b_j} \right) \right).$$

Finally, we get

$$\hat{\mu}_\theta(z) = \exp \left\{ A_{\alpha,b,a}(z) \right\} \Big(\cos \left(B_{\alpha,b,a,\gamma}(z) \right) + i \sin \left(B_{\alpha,b,a,\gamma}(z) \right) \Big).$$

This completes the proof. □

Next we choose $z_1, z_2, \ldots, z_m \in \mathbb{R}^d$ with $m \geq 2k+d+1$ and estimate the parameters by

$$\begin{aligned} &\operatorname*{argmin}_{\theta} \sum_{\ell=1}^{m} |\hat{\mu}_E(z_\ell) - \hat{\mu}_\theta(z_\ell)|^2 \\ &\quad = \operatorname*{argmin}_{\theta} \sum_{\ell=1}^{m} \left(|\Re \hat{\mu}_E(z_\ell) - \Re \hat{\mu}_\theta(z_\ell)|^2 + |\Im \hat{\mu}_E(z_\ell) - \Im \hat{\mu}_\theta(z_\ell)|^2 \right), \end{aligned} \tag{4.1}$$

where $\hat{\mu}_\theta$ is the characteristic function of a TS distribution with parameter vector θ.

Example 2. *When $d = 1$, the characteristic function of TS is in equation 2.2. Then*

by the propsition 5, we have

$$A_{\alpha,b,a}(z) = \Gamma(-\alpha)a(b^2+z^2)^{\frac{\alpha}{2}}\left(\cos\left(\alpha\arctan\left(\frac{z}{b}\right)\right)-b^{\alpha}\right),$$

and

$$B_{\alpha,b,a,\gamma}(z) = \gamma z-\Gamma(-\alpha)a(b^2+z^2)^{\frac{\alpha}{2}}\sin\left(\alpha\arctan\left(\frac{z}{b}\right)\right),$$

Then, we could easily get from of characteristic function in propsition 5.

4.2 Data analysis for the bivariate case

We first give a small-scale simulation study for methodology in section 4.1. We consider the case where σ has $k = 5$ equally spaced direction, $\alpha = 0.6$, the a_j's are $(1, 0.5, 1, 0.5, 1)$, the b_j's are $(1, 2, 1, 2, 1)$, and the drift $\gamma = (1, 1)$. We simulated 100 datasets each comprised of 5000 observations. We use the relative error (RE) to measure the accuracy and it is a ratio of absolute error and the expected value, where is the following equation

$$RE = \frac{|estimated - true|}{true}.$$

The following tables give the means and standard deviations of the relative errors for each parameter.

Table 4.1: The results of siumlation study.

Parameter	a_1	a_2	a_3	a_4	a_5
Mean	0.1002	0.0832	0.1206	0.0977	0.1253
Standard Deviation	0.0928	0.0765	0.0908	0.0829	0.1283

Parameter	b_1	b_2	b_3	b_4	b_5
Mean	0.0642	0.0446	0.0608	0.0371	0.0590
Standard Deviation	0.0623	0.0598	0.0473	0.0401	0.0459

Parameter	α	γ_1	γ_2
Mean	0.1270	0.1491	0.1319
Standard Deviation	0.0864	0.1149	0.1141

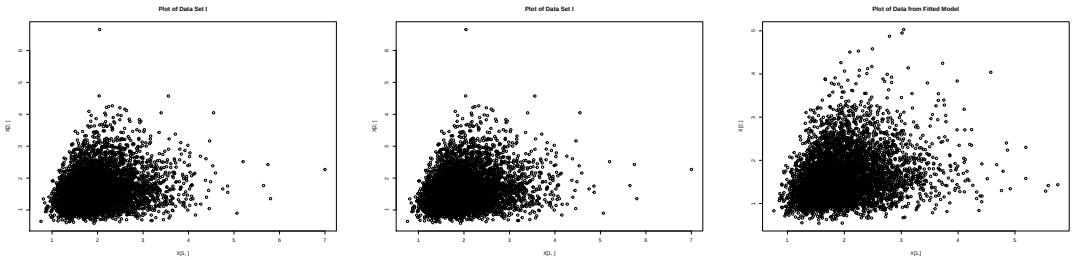
In this table, the mean of RE for all parameters is less than 0.15, and their standard deviation of RE is less than 0.13. We say this table performs good results for all parameters, which means our methodology of parameter estimation works well. In the following part, we give four examples of data analysis based on a bivariate case. For the first three examples, we consider σ to be discrete and set up the same true values of all parameters and directions, then provide different choices of directions in the unit circle when we simulate the fitted data. Our goal is to see the ability of our model to capture the true directions and are interested in seeing their performances when the directions chosen in the fitted data have a small angle and a large angle different from the true directions. Under the bivariate case, as we mentioned in Section 2.2, $s = (\cos(\theta), \sin(\theta))$ is a bijection function between s and θ. In these examples, we consider the direction based on θ. The last example is based on the continuous σ, our goal is to see the methodology's performance of approximation by discretizing the spectral measure.

4.2.1 Examples of data analysis with discrete sigma

This section includes three cases based on the same true model. We begin by building the true model first. Let's focus on the first quadrant on the unit sphere $\mathbb{S}^1$, where $0 = d_0 < d_1 < \cdots < d_k = \pi/2$ and $\theta_1, \ldots, \theta_k \in (0, \frac{\pi}{2})$, and also consider the

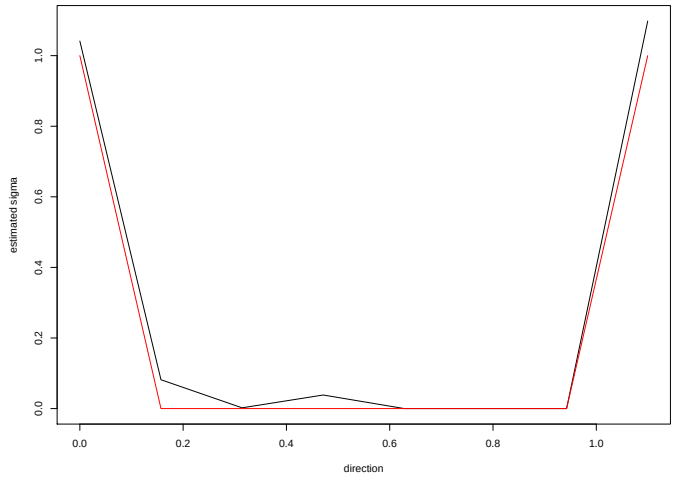
case where σ has $k = 8$ equally spaced directions. we assume that σ is of the form given in (2.14), hence $d_j = \dfrac{\pi j}{2k}$ and $\theta_j = \dfrac{\pi(2j-1)}{4k}$. Next we choose the first direction and the last direction to be the true directions to build the dataset, which is s_1 and s_8. Then we simulate $X \sim \mathrm{TS}_\alpha(b, \sigma, \gamma)$ with parameters $\alpha = 0.6$, $\gamma = 0$, $b(s) \equiv 2$ and a_j's are $(1, 0, 0, 0, 0, 0, 0, 1)$. In each case, we use the same way to simulate the training data and the testing data with sample size $n = 5000$.

The first case is using the same θ_j as the true model when we do the parameter estimation and generate the fitted data. The direction in fitted model denoted by θ'_j, which is $\theta'_j = \theta_j$. The plot of training data and the testing data for the first case plot are in Figure 4.1(a) and Figure 4.1(b) separately. Based on the true model, we need to estimate 17 parameters, so we took $m = 17$ as the number of z_ℓ's. Same as θ_j, we chose these to be evenly spaced in the first quadrant. We then fit the parameters using the training data by minimizing the objective function given in (4.1). To perform the optimization we first used differential evolution [40], which is a global optimization method based on randomly exploring the parameter space by randomly modifying pairs of vectors, as implemented in the `DEoptim` function of the "DEoptim" R package, to get initial values. These were then plugged into the `optim` function in R with the L-BFGS-B option. When we do the optimization, we set up the lower bound for a_j is 0, for b_j, α are 0.1 , also set up the upper bound for a_j, b_j are 5 and α is 0.95. After optimization, we get the objective function was $2.582 * 10^{-6}$. The $\hat{a} = (1.041, 0.082, 0.002, 0.038,\ 0, 0, 0, 1.098)$, $\hat{b} = (2.117, 2.457, 1.971, 2.507, 1.989, 2.502, 1.967, 2.085)$ and $\hat{\alpha} = 0.567$. Plots of the estimate values of the a_j's are given in Figure 4.2(a). Comparing with the true value, we find that estimated parameters are close to the true value and the distance also approaches zero. Based on these estimated parameters and θ'_j, we simulated the 5000 observations from the fitted model, and it is plotted in Figure 4.1(c). Let's call it fitted data.



(a) (b) (c)

Figure 4.1: Plot of three different 2 dimensional data sets. (a) is training data, (b) is testing data and (c) is fitted data. All of these three data sets have the sample size, which is 5000. In figures, xlab represents the first component and ylab represents the second component.

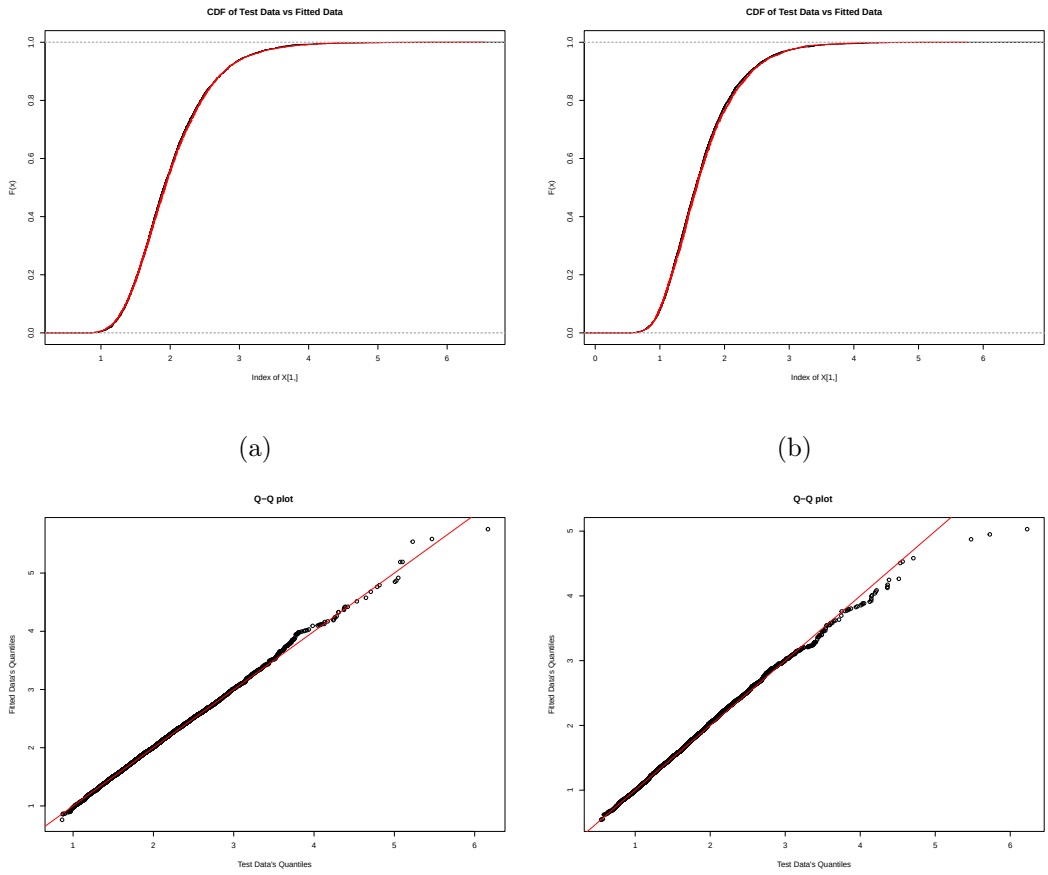


(a)

Figure 4.2: Xlab represents the direction, since in the first quadrant of unit circle, the direction is from 0 to $\pi/2$ and ylab represents $\hat{a}$.

We move to check the goodness-of-fit and try to show the fitted data have the same distribution as the test data. Figure 4.3(a) and 4.3(b) are the plot of CDF of testing data set vs fitted data in each component. And Figure 4.4(a) is the plot of join CDF of testing and fitted data. All of these three figures show that the plot of CDF between testing and fitted data is almost the same. Otherwise, Figure 4.3(c) and 4.3(d) are the Q-Q plots, which compare the quantiles of the testing and fitted data

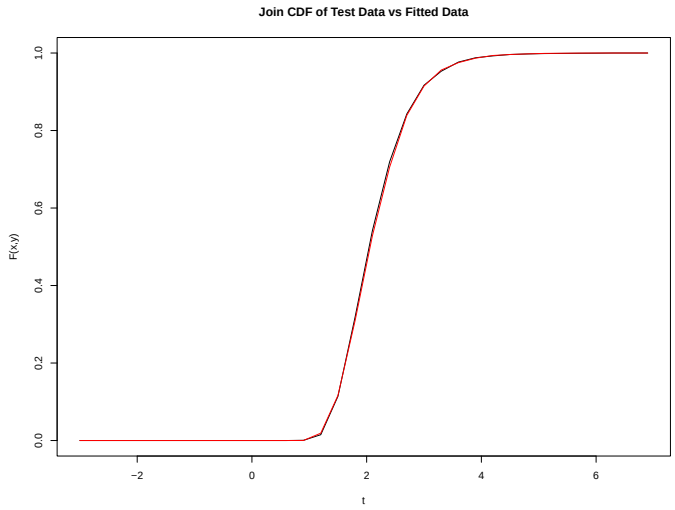
in each component. In these two plots, almost points are one the straight line $y = x$, which indicates two empirical distributions coincide. Also, we performed Kolmogorov-Smirnov (KS) test in each component of the testing and fitted data separately and the hypothesis is H_0 : Fitted data follow Testing data's distribution and H_a : Fitted data doesn't follow Testing data's distribution. For the first component, we get test statistic $D = 0.0138$ and p-value $= 0.728$ and for the second component the test statistic $D = 0.019$ and p-value $= 0.340$. Both two components' in the KS test fail to reject null hypothesis at 10% level. Next, we tested both components together, using the kernel consistent density equality test, which was introduced in [41] and is implemented in the `npdeneqtest` function of the R package "np". This test is, essentially, based on using kernel density estimators to estimate a certain distance between two multivariate densities and see if this estimated distance is significantly different from zero. In this case, the test statistic is $Tn = -8.812$ and the p-value $=$ 0.556. Thus, we have the same results as Kolmogorov-Smirnov test in each component and conclude the fitted data have the same distribution as the test data.



(a) (b)

(c) (d)

Figure 4.3: Figure(a) is CDF plots between fitted data and testing data in the first component. Figure(b) is CDF plots in the second component. The red line represents fitted data, and black line represents testing data. Figure(c) is a Q-Q plot of testing data and fitted data in the first component and Figure(d) is in the second component. In these two plots, xlab is quantiles of test data and ylab is quantiles of fitted data, also red line is $y = x$.

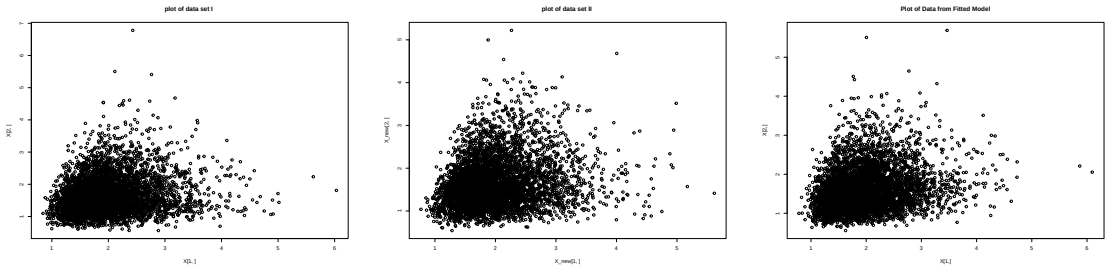


(a)

Figure 4.4: ylab is the value of $F(x = t, y = t)$, xlab is the value of t. Red line represents joint CDF of fitted data and black line represents joint CDF of testing data.

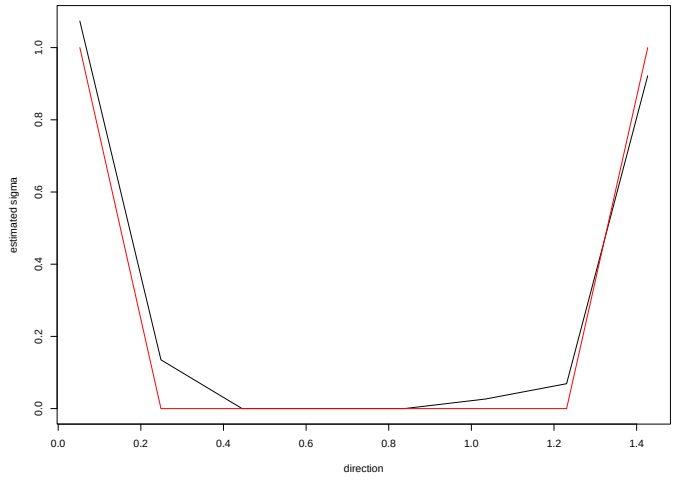
Then, let's move to the second case. In this case, we begin to simulate the training data and the testing data, which are plotted in Figure 4.5(a) and Figure 4.5(b) separately. Different from the first case, we added a small angle on the original θ_j when we did the optimization and set up the fitted model. Let's define the small angle is 3° or $\frac{\pi}{60}$, then $\theta'_j = \theta_j + \frac{\pi}{60} = \frac{\pi(2j-1)}{4k} + \frac{\pi}{60}$ and $\theta'_1, \ldots, \theta'_k \in (0, \frac{\pi}{2} + \frac{\pi}{60})$. Same as Case 1, we need to estimate 17 parameters, so we took $m = 17$ as the number of z_ℓ's. And we chose these based on θ'_j to get z_ℓ. Although the true model and fitted model have different directions, and the difference is small, we hope that the estimated parameters could be close to the true parameters and the small angle will not influence the results too much. We then fit the parameters using the same method. Also, when we did the optimization, we set up the same lower bound and upper bound for a_j, b_j and α. After optimization, we get the objective function was $9.098 * 10^{-5}$. The $\hat{a} = (1.073, 0.135, 0.000, 0.000, 0.000, 0.026, 0.069, 0.922)$, $\hat{b} = (2.301, 2.178, 2.230, 1.820, 2.095, 2.457, 1.813, 2.102)$ and $\hat{\alpha} = 0.573$. Plots of the

estimate values of the a_j's is given in Figure 4.6(a). Comparing with the true value, the estimated parameters are not influenced by the small angle, our model still can capture the correct directions. Based on these estimated parameters and θ'_j, we siumlated the 5000 observations from the fitted model, and it is plotted in Figure 4.5(c). And let's call it fitted data.



(a) (b) (c)

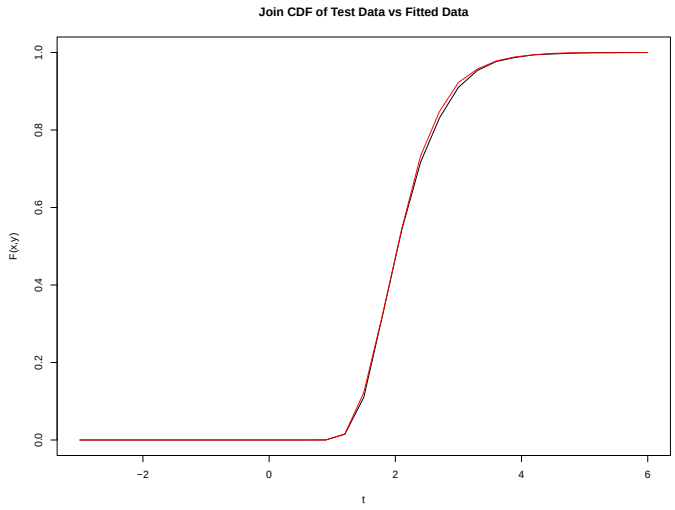
Figure 4.5: Plot of three different 2 dimensional data sets. (a) is training data, (b) is testing data and (c) is fitted data. All of these three data sets have the same sample size, which is 5000. In figures, xlab represents the first component and ylab represents the second component.



(a)

Figure 4.6: Here xlab represents the direction, and the direction is from 0 to $\frac{\pi}{2} + \frac{\pi}{60}$ and ylab represents $\hat{a}$.

We then check the goodness of fit. Figure 4.8(a) and 4.8(b) are the plot of CDF of testing data set vs fitted data in each component. And Figure 4.7(a) is the plot of joint CDF of testing and fitted data. Figure 4.8(c) and 4.8(d) are the Q-Q plots, which compare the quantiles of the testing and fitted data in each component separately. Also, we performed KS test in each component of the testing and fitted data separately. For the first component, we get test statistic $D = 0.0134$ and p-value $= 0.744$ and for the second component the test statistic $D = 0.022$ and p-value $= 0.163$. Next, we tested both components together, using the kernel consistent density equality test and the test statistic is $Tn = 2.754$ and the p-value $= 0.111$. Thus, we have the same results as KS test in each component and conclude the fitted model has the same distribution as the true model, and the small angle cannot influence the result.



(a)

Figure 4.7: ylab is the value of $F(x = t, y = t)$, xlab is the value of t. Red line represents joint CDF of fitted data and black line represents joint CDF of testing data

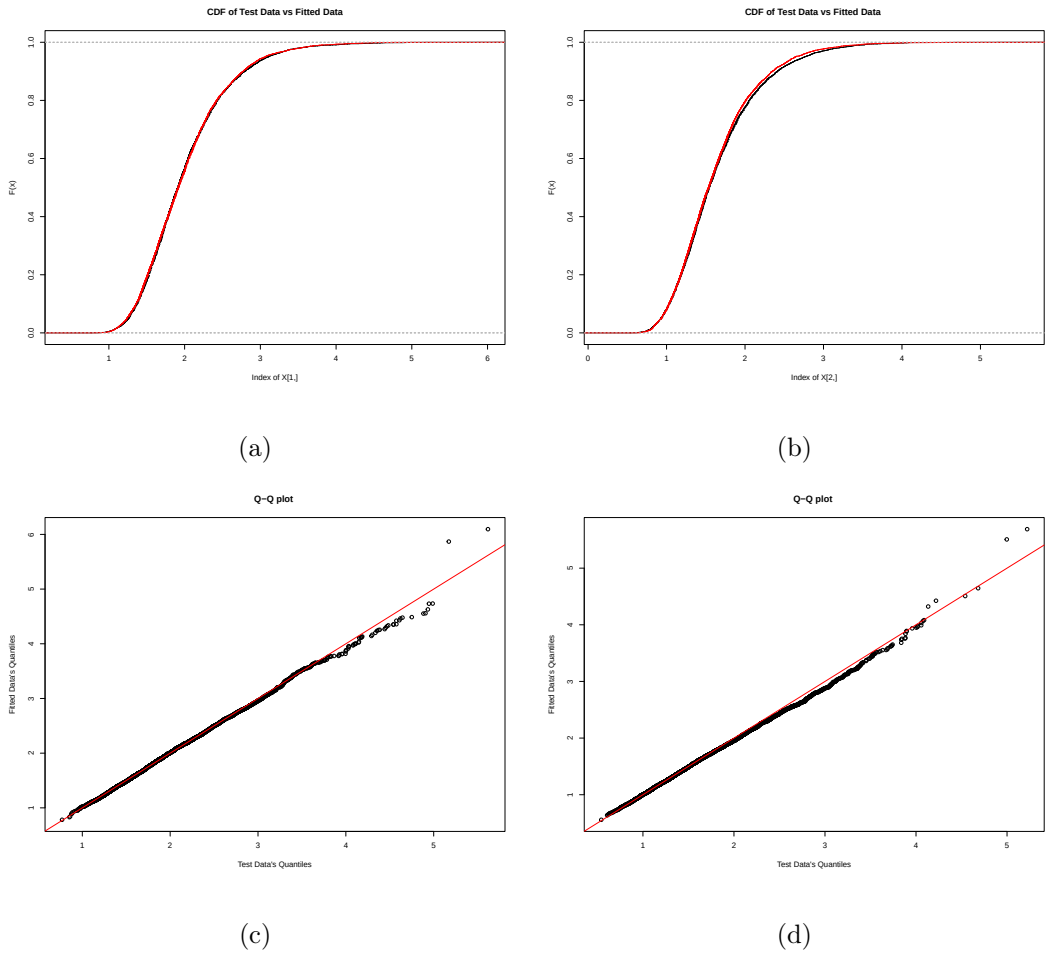
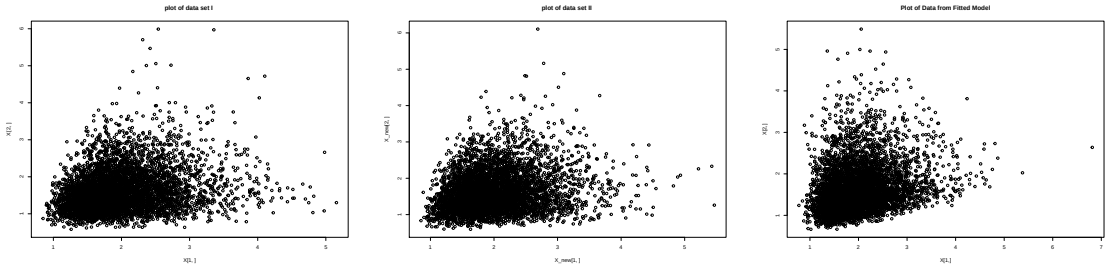


Figure 4.8: Figure(a) is CDF plots between fitted data and testing data in the first component. Figure(b) is CDF plots in the second component. The red line represents fitted data, and black line represents testing data. Figure(c) is a Q-Q plot of testing data and fitted data in the first component and Figure(d) is in the second component. In these two plots, xlab is quantiles of test data and ylab is quantiles of fitted data, also red line is $y = x$.

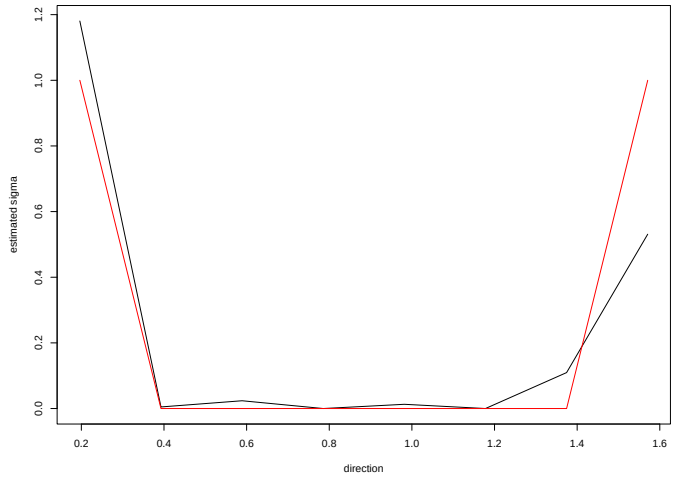
Next, let's focus on the third case of discrete σ. In this case, again we simulate the training data and the testing data first, which are plotted in Figure 4.9(a) and Figure 4.9(a) separately. Then we considered to add a large angle on original θ_j when we did the optimization and set up the fitted model. Let's define the large angle is 11.25° or $\frac{\pi}{16}$, then $\theta'_j = \theta_j + \frac{\pi}{16} = \frac{\pi(2j-1)}{4k} + \frac{\pi}{16}$ and $\theta'_1, \ldots, \theta'_k \in (0, \frac{\pi}{2} + \frac{\pi}{16})$. Same as last two casses, we took $m = 17$ as the number of z_ℓ's and chose these based on θ'_j to get z_ℓ. In this case, we interested in the fitted model performance when there existed

a large different direction with the true model. We then fit the parameters using the same method. Also, when we did the optimization, we set up the same lower boundand upper bound for a_j, b_j and α. After optimization, we get the objective function was $1.108 * 10^{-4}$. The $\hat{a} = (1.181, 0.005, 0.024, 0, 0.013, 0.000, 0.109, 0.531)$, $\hat{b} = (2.215, 2.153, 1.970, 1.841, 2.186, 1.596, 1.947, 1.485)$ and $\hat{\alpha} = 0.609$. Plots of the estimate values of the a_j's is given in Figure 4.10(a). From Figure 4.10(a), we find the $\hat{a}$ are not as good as before and the large angle may influence the model to capture the correct directions as accurate as before cases. Based on these estimated parameters and θ'_j, we simulated the 5000 observations from the fitted model, and it is plotted in Figure 4.9(c). Let's call it fitted data.



(a) (b) (c)

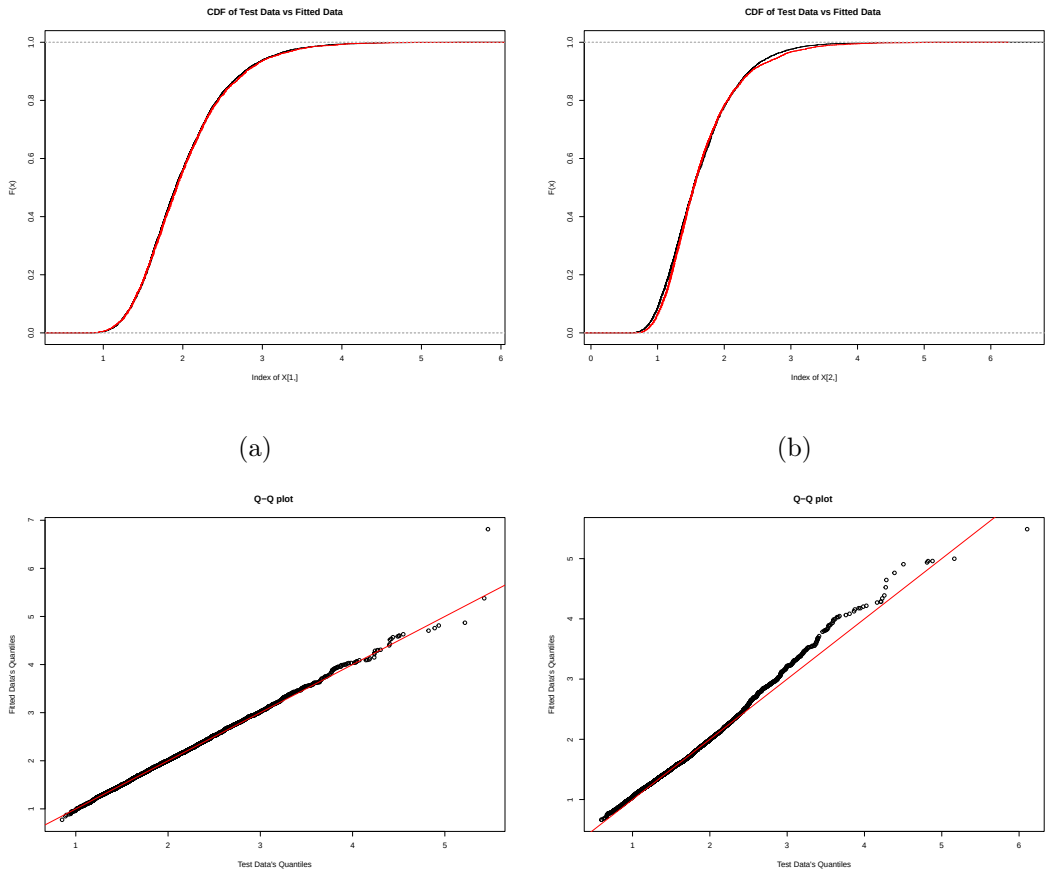
Figure 4.9: Plot of three different 2 dimensional data set. (a) is training data, (b) is testing data and (c) is fitted data. All of these three data sets have the sample size is 5000. In figures, xlab represents the first component and ylab represents the second component.



(a)

Figure 4.10: Xlab represents the direction, and the direction is from 0 to $\frac{\pi}{2} + \frac{\pi}{16}$ and ylab represents $\hat{a}$.

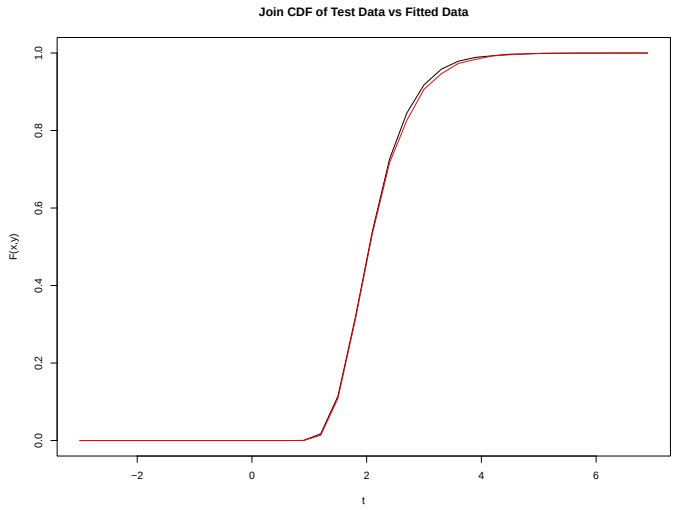
We then check the goodness of fit. Figure 4.11(a) and 4.11(b) are the plot of CDF of testing data set vs fitted data in each component separately. And Figure 4.12(a) is the plot of joint CDF of testing and fitted data. Figure 4.11(c) and 4.11(d) are the Q-Q plots, which compare the quantiles of the testing and fitted data in each component separately. Also, we performed KS test in each component of the testing and fitted data separately. For the first component, we get test statistic $D = 0.0142$ and p-value $= 0.695$ and for the second component the test statistic $D = 0.035$ and p-value $= 0.004$, but the null hypothesis is rejected at 10% level. Next, we tested both components together, using the kernel consistent density equality test and the test statistic is $Tn = 12.730$ and the p-value $< 2.22 \times 10^{-16}$. This also implies that we reject to null hypothesis at 10% level. Thus we conclude that the fitted model has different distribution than the true model, and the large angle influences our model to capture the correct directions.



(a) (b)

(c) (d)

Figure 4.11: Figure(a) and Figure(b) are CDF plots between fitted data and testing data in the first component and the second component separately. The red line represents fitted data, and black line represents testing data. Figure(c) is a Q-Q plot of testing data and fitted data in the first component and Figure(d) is in the second component. In these two plots, xlab is quantiles of test data and ylab is quantiles of fitted data, also red line is $y = x$.



(a)

Figure 4.12: ylab is the value of $F(x = t, y = t)$, xlab is the value of t. Red line represents joint CDF of fitted data and black line represents joint CDF of testing data

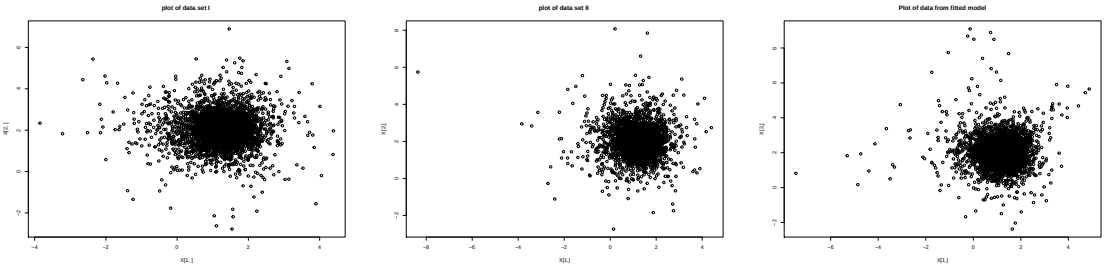
4.2.2 Examples of data analysis with continuous sigma

In this section, we let $X = (X_1, X_2) \sim TS_\alpha(\sigma, b, \gamma)$, where $\sigma'(d\theta) = f(\theta)\mathrm{d}\theta$, such that for some $\alpha^*, \beta^* > 0$

$$f(\theta) = \frac{1}{C}\theta^{\alpha^*-1}(2\pi - \theta)^{\beta^*-1}, \quad \theta \in [0, 2\pi),$$

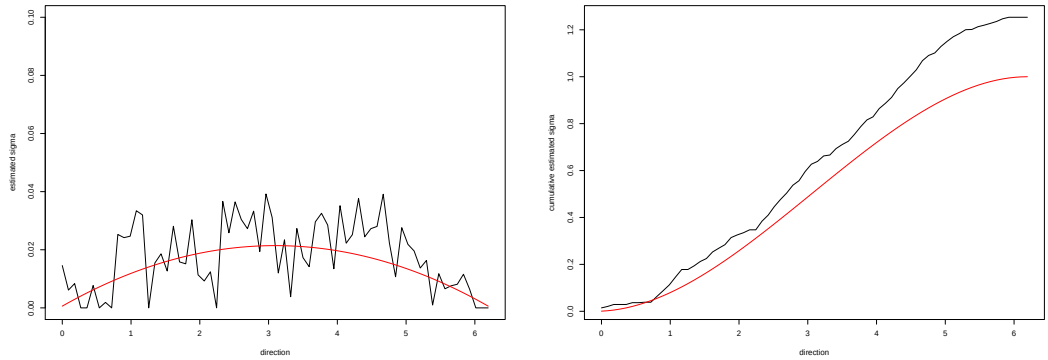
where $C = B(\alpha^*, \beta^*)(2\pi)^{\alpha^*+\beta^*-1}$ and $B(\cdot, \cdot)$ is the beta function. Then σ' is a beta distribution on $[0, 2\pi)$. Here, we chose $\alpha^* = \beta^* = 2$ and then denote this distribution by Beta$(2, 2, 2\pi)$. In this case, we assume all directions are on the unit sphere $\mathbb{S}^1$, so the σ is of the form given in (2.14) and the θ_j's are as described just after that equation. Since σ' follows a continuous distribution, we need to choose a large value of k to make sure all characteristics in this distribution could be captured. Hence, we set $k = 70$, which means we discretized σ into 70 pieces evenly on $\mathbb{S}^1$. Next, we must fit 143 parameters, so took $m = 143$ as the number of z_ℓ's and chose them to be evenly

spaced on $\mathbb{S}^1$. To build the true model, all directions are considered in the model, where true directions are $s_1, s_2, ..., s_{70}$. Them we simulate $X \sim \mathrm{TS}_\alpha(b, \sigma, \gamma)$ with parameter $\alpha = 0.6$, $b(s) \equiv 1$, $\gamma = (2, 2)$ and $\sigma' \sim \mathrm{Beta}(2, 2, 2\pi)$. Similar as before, we simulate the train dataset with and both of their sample size are 5000. One is train dataset in Figure 4.13(a) and the other is testing dataset.in Figure 4.13(b). We then fit the parameters using the same method. Also, when we did the optimization, we set up the lower bounder for a_j is 0, for b_j, α are 0.1 , also set up the upper bound for a_j, b_j are 5 and α is 0.95. After optimization, we get the objective function was $8.643 * 10^{-3}$. The $\hat{\alpha} = 0.654$, drift was $\hat{\gamma} = (2.142, 2.115)$, the 70 $\hat{a_j}$'s were all in the interval $(0.001, 0.040)$ and the 70 $\hat{b_j}$'s were all in the interval $(0.127, 2.100)$. Plots of the estimated values for the $\hat{a_j}$'s and of the cumulative values for the estimated σ are given in Figure 4.14(a) and 4.14(b). From these two plots, the performance of them is not perfect since there exists obviously difference in some directions, while the differences are accepted. Then we simulated the 5000 observations from the fitted model, and it is plotted in Figure 4.13(c). Let's call it fitted data.



(a) (b) (c)

Figure 4.13: Plot of three different 2 dimensional data sets. (a) is training data, (b) is testing data and (c) is fitted data. All of these three data sets have the sample size is 5000. In figures, xlab represents the first component and ylab represents the second component.



(a) (b)

Figure 4.14: Plots of the estimated a_j's is given in (a). A plot of the cumulative function of the estimated spectral measure is given in (b). The x-axis in all plots represents the direction from 0 to 2π.

Further, we check the goodness of fit. Figure 4.15(a) and 4.15(b) are the plot of CDF of testing data set vs fitted data in each component separately. And Figure 4.16 is the plot of joint CDF of testing and fitted data. Figure 4.15(c) and 4.15(d) are the Q-Q plots, which compare the quantiles of the testing and fitted data in each component separately. Also, we performed KS test in each component of the testing and fitted data separately. For the first component, we get test statistic $D = 0.023$ and p-value $= 0.156$ and for the second component the test statistic $D = 0.019$ and p-value $= 0.315$. Next, we tested both components together, using the kernel consistent density equality test and the test statistic is $Tn = -24.931$ and the p-value $= 0.273$. All tests imply that we fail to reject the null hypothesis at 10% level. Thus, we conclude that the fitted model has the same distribution as the true model and our methodology has good performance in both discrete σ and continuous σ.

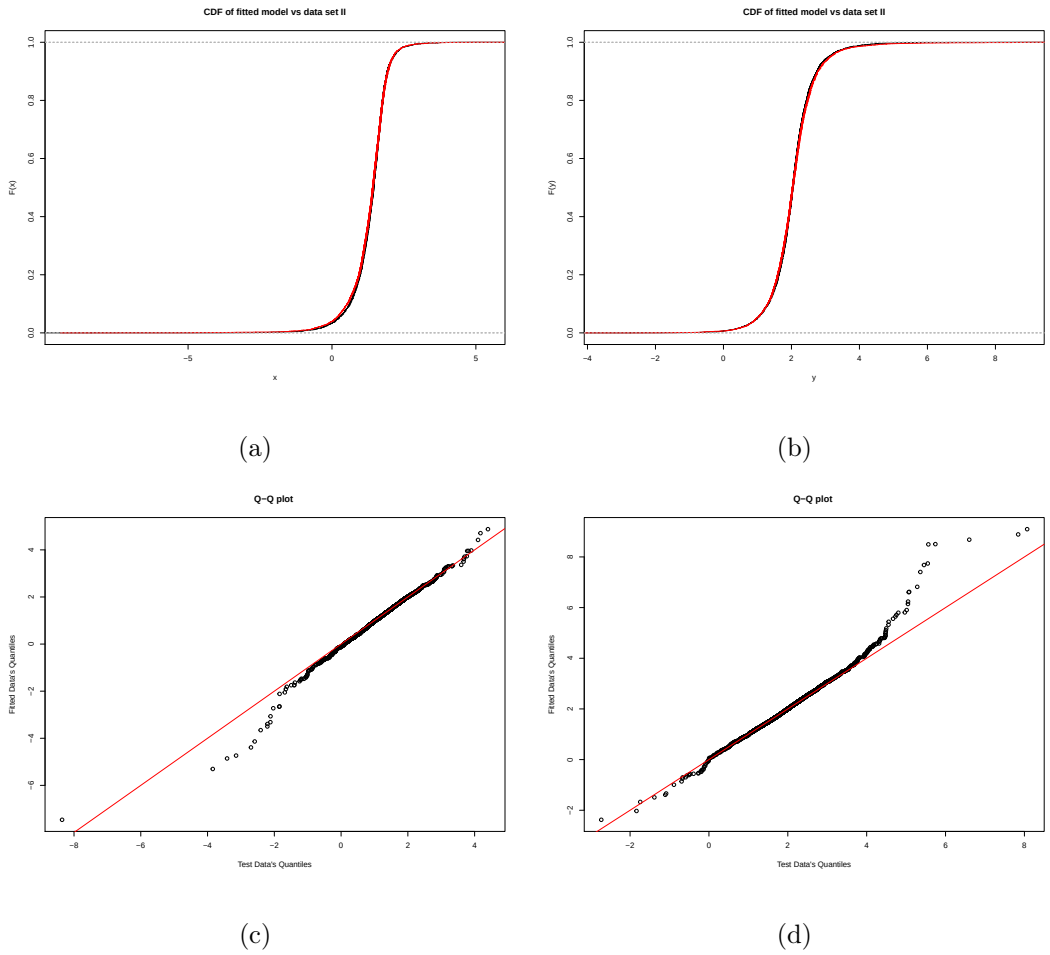
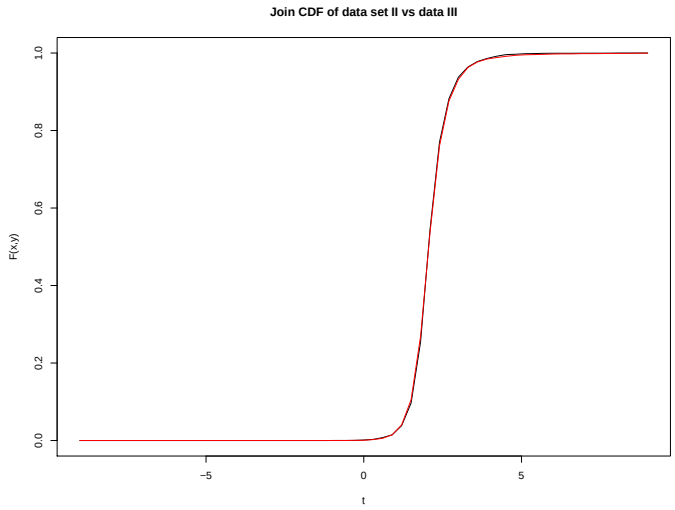


Figure 4.15: In Figure(a) and Figure(b), red line represents CDF of fitted data, and black line represents CDF of test data. Figure(c) and (d) are Q-Q plots of test data set and fitted data in each component. In these two plots, xlab is quantiles of test data and ylab is quantiles of fitted data, also red line is $y = x$.



(a)

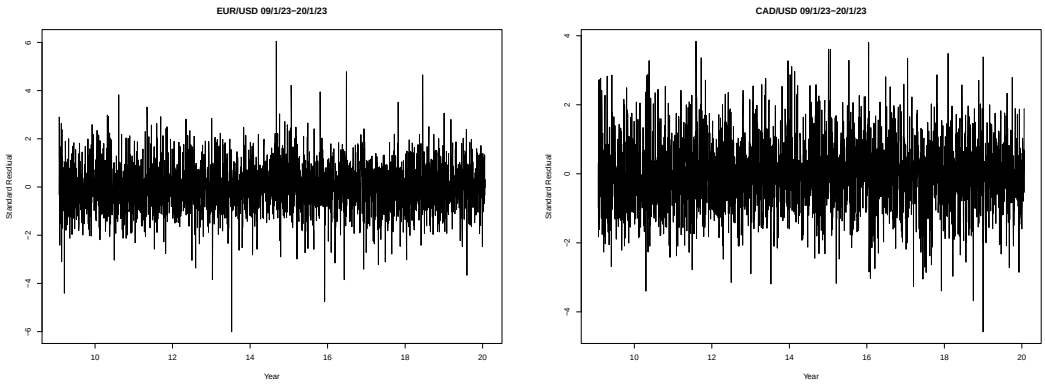
Figure 4.16: ylab is the value of $F(x = t, y = t)$, xlab is the value of t. Red line represents joint CDF of fitted data and black line represents joint CDF of test data.

4.2.3 Real data analysis for the bivariate case

In this section, we apply our methodology to the modeling of two real-world bivariate financial datasets related to exchange rates. The first is comprised of the exchange rates between the US Dollar (USD) and two standard currencies: the Canadian Dollar (CAD) and the Euro (EUR). The second is comprised of the exchange rates between USD and two common cryptocurrencies: Bitcoin (BTC) and Ethereum (ETH). All of the data was downloaded from Yahoo Finance, `http://finance.yahoo.com`.

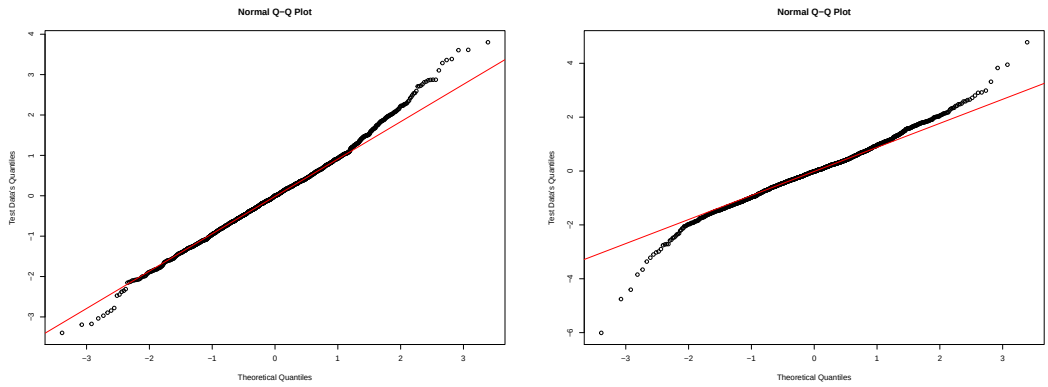
We begin by jointly modeling the exchange rates CAD/USD and EUR/USD. Our data consists of the daily closing prices for the period from January 23, 2009 to January 23, 2020 for each exchange rate. The prices are converted to log returns and, to model dependence within each time series, an ARMA$(1,1)$-GARCH$(1,1)$ filter is applied to the log returns for each exchange rate separately. This is implemented using the R package "rugarch". The standardized residuals are then considered as ordered pairs with the first component corresponding to CAD/USD and the second to EUR/USD for the same day, and their time series plot of standard residuals for

daily are shown in Figure 4.17. In total, the data consists of 2863 ordered pairs. This data was randomly split into a training data set consisting of 1432 observations and a testing dataset consisting of 1431 observations, which is shown in Figure 4.18. Figure 4.19 gives normal Q-Q plots for each component based on the testing data. We also performed an adjusted Jarque-Bera test for normality (using the function `ajb.norm.test` in the R package "normtest") on each component of the testing data. In both cases, the p-value was less than 10^{-5}. Clearly, the normal distribution is not reasonable for either component. Instead, we follow our methodology for fitting a bivariate TS distribution to the training data.



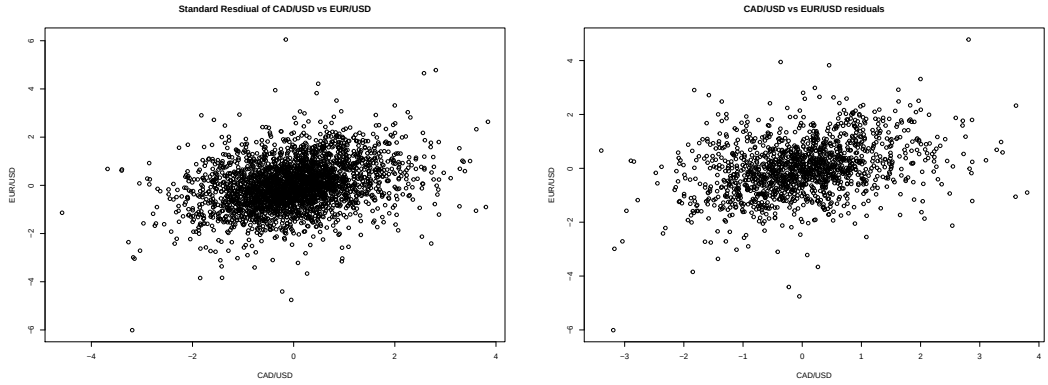
(a) (b)

Figure 4.17: Both of these plots have the same sample size 2863. The left is time serise plot of CAD/USD standard residuals for daily. The right is time serise plot of EUR/USD standard residuals for daily. In these two figures, xlab represents the time and ylab represents the standard residuals.



(a) (b)

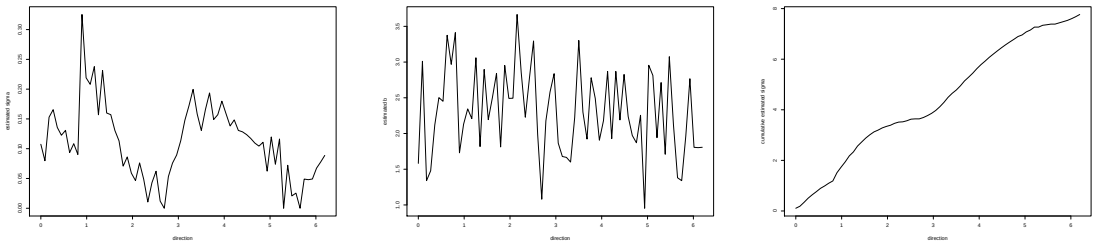
Figure 4.19: Normal Q-Q plots for the testing data. The plot in a) is for the first component (CAD/USD) and the plot in b) is for the second component (EUR/USD). The diagonal (red) line is the line passing through the first and third quartiles.



(a) (b)

Figure 4.18: The left plot shows train data set and its sample size is 1432. The right plot shows test data and its sample size is 1431. These two figures' xlab represents the standard residuals of CAD/USD and ylab represents the standard residuals of EUR/USD.

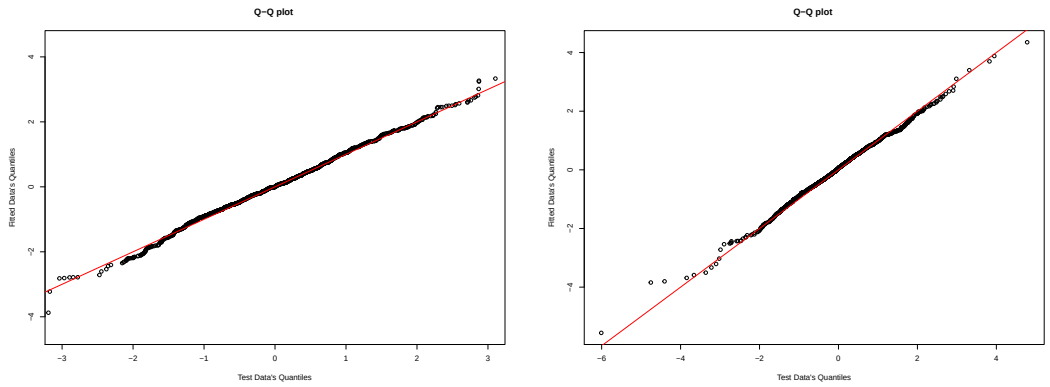
We discretized σ into $k = 70$ equally spaced directions. Thus, we assume that σ is of the form given in equation (2.14) where the θ_j's are as described just after that equation. This means that we must fit 143 parameters. For this reason, we took $m = 143$ as the number of z_ℓ's. We chose these to be evenly spaced on the unit sphere $\mathbb{S}^1$. We then fit the parameters using the training data by minimizing



(a) (b) (c)

Figure 4.20: Estimated values for CAD/USD and EUR/USD data. Plots of the estimated a_j's and b_j's are given in (a) and (b), respectively. A plot of the cumulative function of the estimated spectral measure is given in (c). The x-axis in all plots represents the direction from 0 to 2π.

the objective function given in equation (4.1). To perform the optimization, we first used differential evolution to get initial values. These were then plugged into the `optim` function in R with the L-BFGS-B option. After optimization, the value of the objective function was 1.245×10^{-4}. The estimated value of α was $\hat{\alpha} = 0.240$, the 70 estimated values of the a_j's were all in the interval $(0.000, 0.325)$, the 70 estimated values of the b_j's were all in the interval $(0.953, 3.666)$, and the estimated drift was $\hat{\gamma} = (0.004, -0.153)$. Plots of the estimated values of the a_j's, the b_j's, and of the cumulative values for the estimated σ are given in Figure 4.20. The plots suggest that, in this case, the spectral measure is approximately a constant times a uniform distribution and the values of the b_j are approximately all equal.



(a) (b)

Figure 4.22: Fitted TS Q-Q plots for the (a) CAD/USD and (b) EUR/USD data. These compare the quantiles of the testing data with the fitted data. The diagonal (red) line is the line $y = x$.

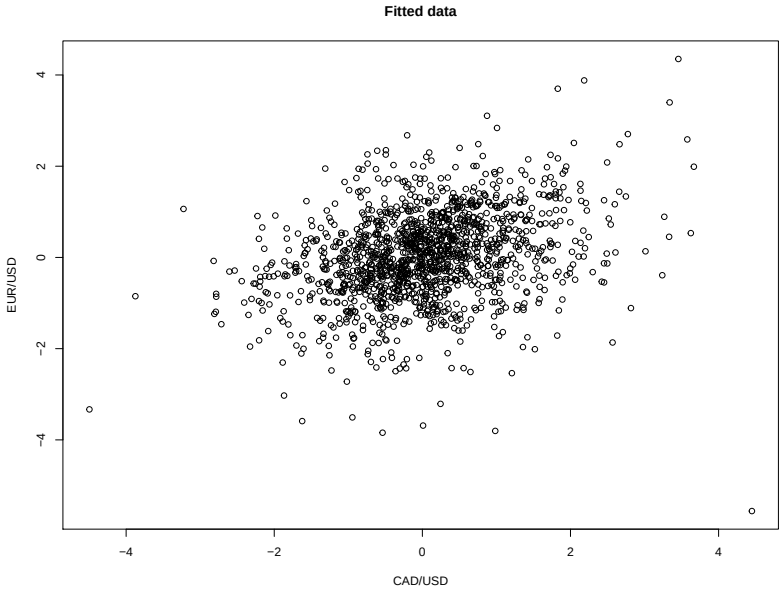
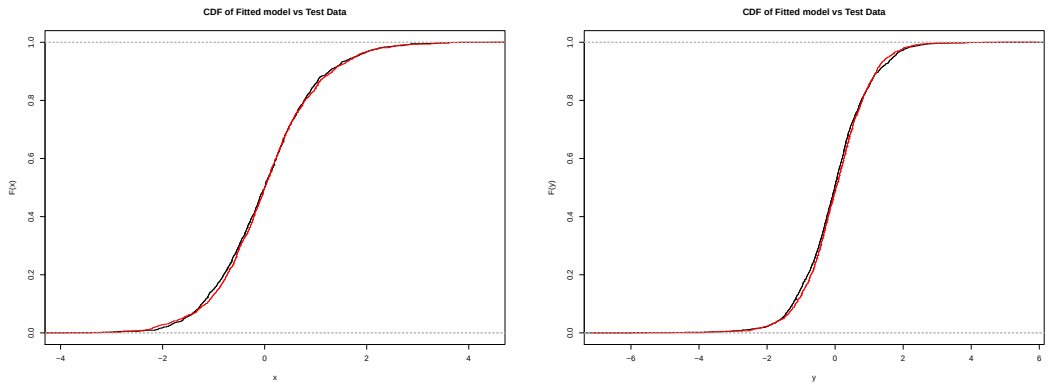


Figure 4.21: The sample size is 1432. In figures, x-axis represents the first component CAD/USD and y-axis represents the second component EUR/USD.

We next check the goodness of fit. Since it is computationally intractable to evaluate the cdfs and pdfs of bivariate TS distributions, we instead use simulation-based approaches. We simulated 1431 observations from the fitted model, which is the same as the number of observations in the testing data. We call this the fitted data, and its plot is shown in Figure 4.21. In Figure 4.22, we give Q-Q plots, which compare the



(a) (b)

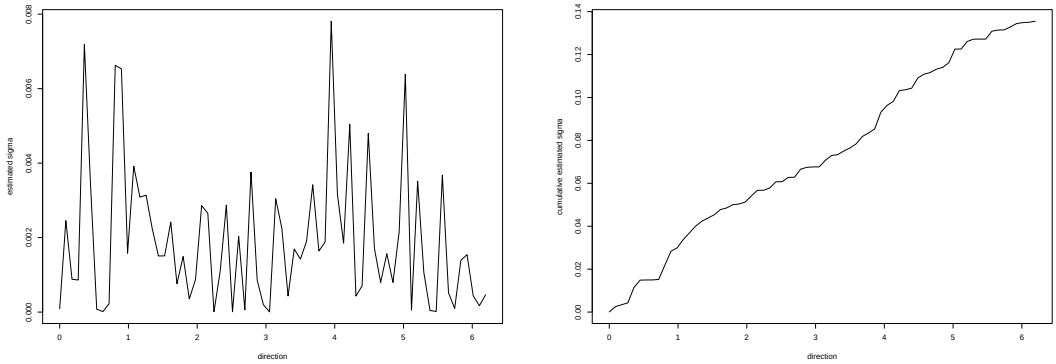
Figure 4.23: Fitted CDF for the (a) CAD/USD and (b) EUR/USD data. These compare the cdf of the testing data with the fitted data. Red line is Fitted TS CDF and Black line is Test CDF.

quantiles of the testing and fitted data. In both (a) and (b), almost points lie on the red line $y = x$, which means test data set and fitted data set's empirical distribution coincide. Figure 4.23, we give CDF of test data set and fitted data set, from these two plots, we got the same result. We also performed formal goodness-of-fit testing. First, we performed KS tests comparing each component of the testing and fitted data separately. For the first component we obtained the test statistic $D = 0.029$ and the p-value $= 0.600$ and for the second component we obtained the test statistic $D = 0.034$ and the p-value $= 0.397$. Next, we tested both components together, using the kernel consistent density equality test, the test statistic is $Tn = 10.396$ and the p-value $= 0.111$. Note that the samples compared by the hypothesis tests are independent as the fitting was done using the training data, but the tests are performed using the testing data. The results of our goodness-of-fit plots and tests suggest that the bivariate TS distribution is a reasonable model for this data.

We note that our estimates for the b_j's are all above 0.95 and many are above 2, which suggests that a fairly large amount of tempering is needed. For comparison, we also fit a multivariate stable distribution. The methodology is the same except that now when we perform the optimization in (4.1), we do not optimize over the b_j's and

instead fix these to all be 0. In this case, the minimum value of the objective function was $= 5.627 \times 10^{-3}$. The estimated value of α was $\hat{\alpha} = 0.914$, the estimated drift was $\hat{\gamma} = (0.243, -0.064)$, and the 70 estimated values of the a_j's were all in the interval $(10^{-6}, 2 \times 10^{-2})$. Plots of the estimated values of the a_j's and of the cumulative values for the estimated σ are given in Figure 4.24. It is interesting to note that the shape of the spectral measure is similar to what was found when fitting the tempered stable distribution, which suggests that the models find similar dependence structures. However, we can see that the magnitude of the spectral measure is much smaller, which suggests that the model is trying to compensate for the heavy tails of the stable distribution by using, what is essential, a small scaling factor. This is further emphasized by the fact that the estimated value of α is high and, of course, the higher the value of α the lighter the tails for a stable distribution.

We again simulated 1431 observations from the fitted model and used it to check the goodness of fit. From Figure 4.25, comparing with the test data set, we find there exist some large-value points, which implies the tail may heavier than test data's empirical distribution. In Figure 4.26, we give Q-Q plots comparing the quantiles of the testing and the fitted data. Also in Figure 4.27, we compare CDF of test data and the fitted data. Next, we perform a KS tests for each component separately. For the first component we obtained the test statistic $D = 0.081$ and the p-value $< 10^{-3}$ and for the second component we obtained the test statistic $D = 0.077$ and the p-value $< 10^{-3}$. We also performed the kernel consistent density equality test. However, the test seemed to fail in this case, which may be due to the very heavy tails of the stable distribution. However, from the Q-Q plots , the plot of CDF and the KS test, it is clear that the tails of the stable distribution are too heavy and that it does not provide a good fit.



(a) (b)

Figure 4.24: Stable fit. The plot of the estimated a_j's is given in (a), while the plot of the cumulative function of the estimated σ is given in (b). The x-axis represents the direction from 0 to 2π.

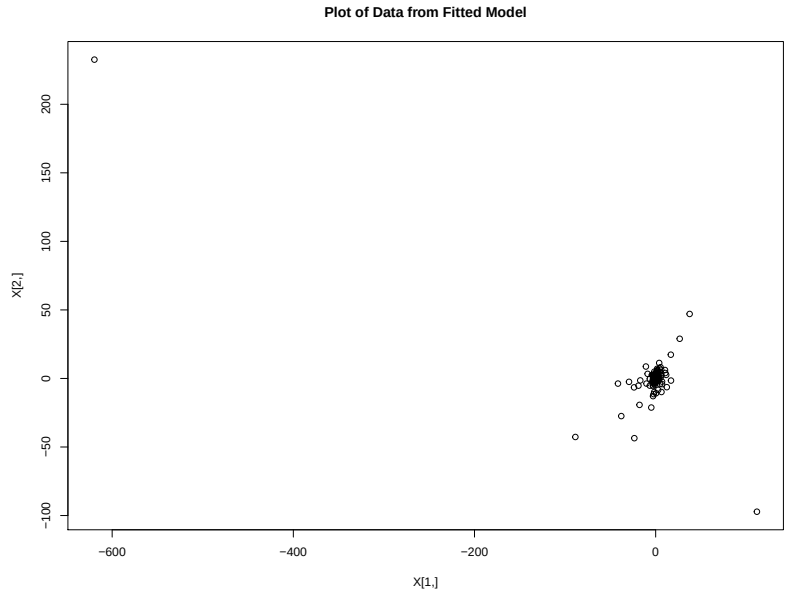
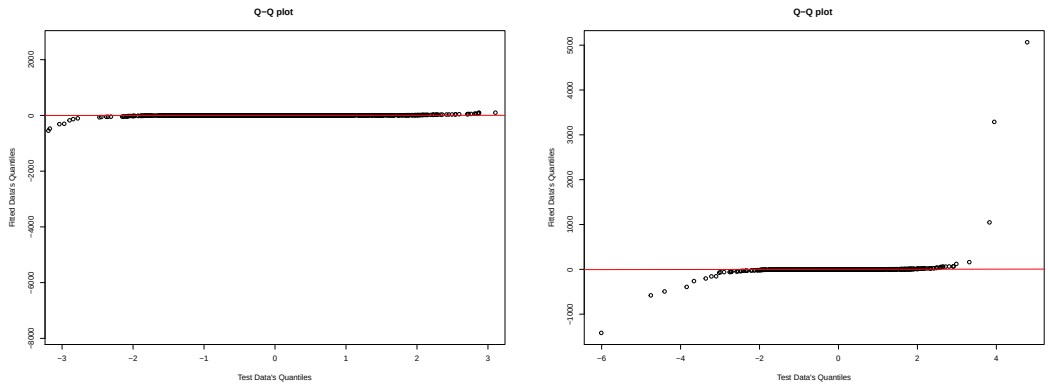
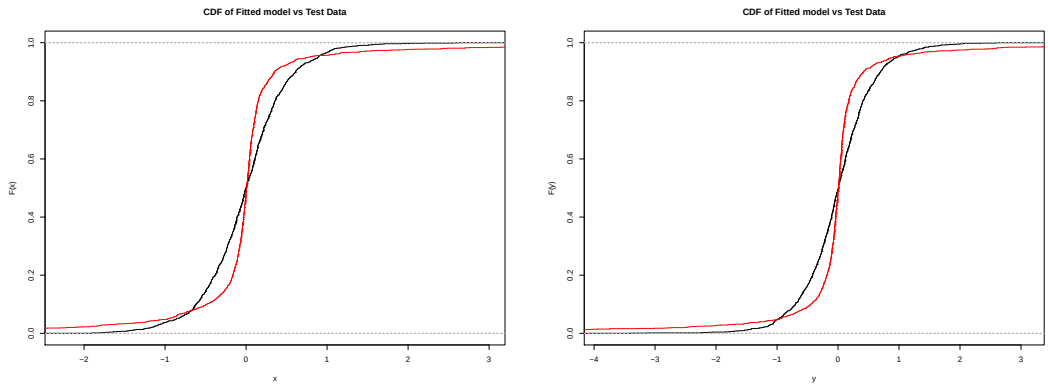


Figure 4.25: The sample size is 1432. In figures, x-axis represents the first component and y-axis represents the second component.



(a) (b)

Figure 4.26: Stable fit. These are the Q-Q plots for the (a) CAD/USD and (b) EUR/USD data. These compare the quantiles of the testing data with the fitted data. The diagonal (red) line is the line $y = x$.

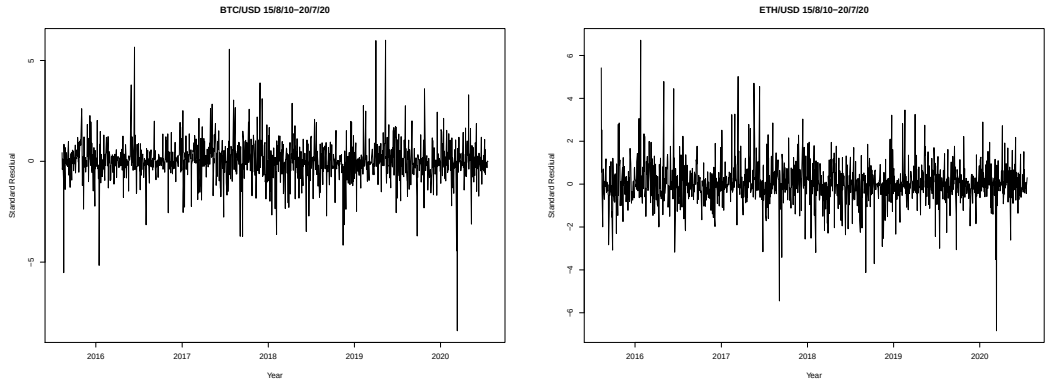


(a) (b)

Figure 4.27: Stable fit. These are CDF for the (a) CAD/USD and (b) EUR/USD data. Black lines represents test data and Red line represents fitted data.

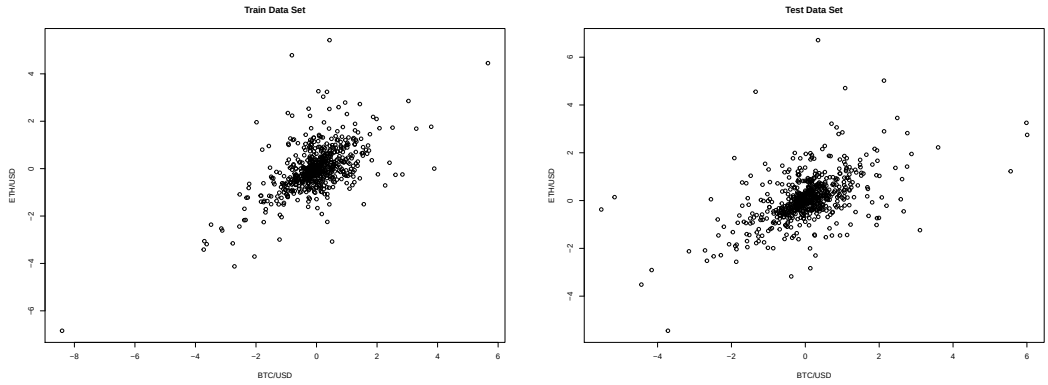
We now turn to the joint modeling of the exchange rates BTC/USD and ETH/USD. In both cases, we consider the daily closing prices from August 10, 2015 to July 20, 2020. We then convert these to log-returns, apply an ARMA$(1, 1)$-GARCH$(1, 1)$ filter, and pair the standardized residuals as before. And their time series plot of standard residuals for daily are shown in Figure 4.28 We get 1290 bivariate observations, which we randomly split into a training and a testing dataset, each with 645 observations,

and the plots are in Figure 4.29.



(a) (b)

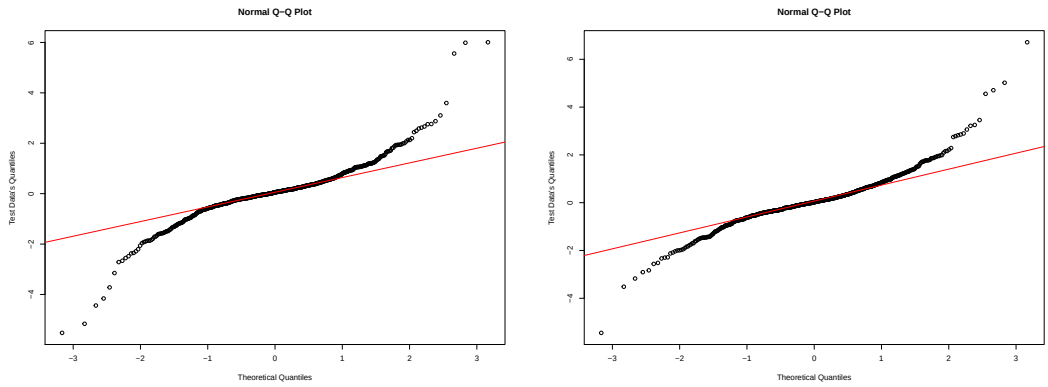
Figure 4.28: Both of these plots have the same sample size 1290. The left is time serise plot of BTC/USD standard residuals for daily. The right is time serise plot of ETH/USD standard residuals for daily. In these two figures, xlab represents the time and ylab represents the standard residuals.



(a) (b)

Figure 4.29: Both of these plots have the same sample size 645. The left plot shows train data set. The right plot shows test data. These two figures' xlab represents the standard residual of BTC/USD and ylab represents the standard residual s of ETH/USD.

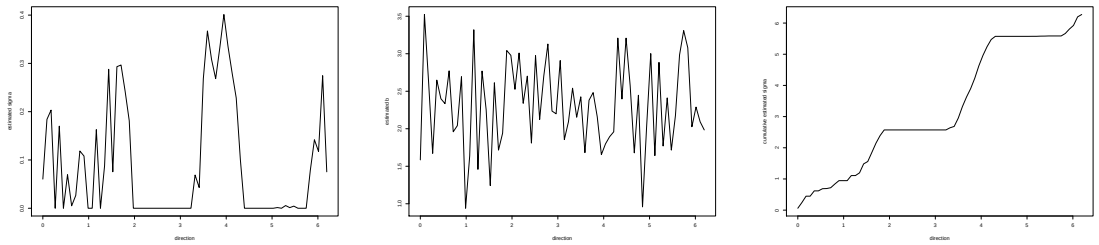
To check for normality, normal Q-Q plots for each component of the testing data are given in Figure 4.30 and an adjusted Jarque-Bera test was performed on each component of the testing data separately. In both cases the p-value was less than 10^{-5}.



(a) (b)

Figure 4.30: Normal Q-Q plots for the testing data. The plot in a) is for the first component (BTC/USD) and the plot in b) is for the second component (ETH/USD). The diagonal (red) line is the line passing through the first and third quartiles.

Clearly, the normal distribution is not reasonable for either component. Instead, we follow our methodology for fitting a bivariate TS distribution to the training data. We chose the same discretization and the same values for the z_ℓ's as in the previous example and we apply the same procedure for fitting the data. After optimization, the value of the objective function was 4.612×10^{-3}. The estimated value of α was $\hat{\alpha} = 0.010$, the 70 estimated values of the a_j's were all in the interval $(0.000, 0.402)$, the 70 estimated values of the b_j's were all in the interval $(0.958, 3.524)$, and the estimated drift was $\hat{\gamma} = (0.466, 0.144)$. Plots of the estimated a_j's, the estimated b_j's, and the cumulative curve for the estimated σ are given in Figure 4.31.



(a) (b) (c)

Figure 4.31: Estimated values for BTC/USD and ETH/USD data. Plots of the estimated a_j's and b_j's are given in (a) and (b), respectively. A plot of the cumulative function of the estimated spectral measure is given in (c). The x-axis in all plots represents the direction from 0 to 2π.

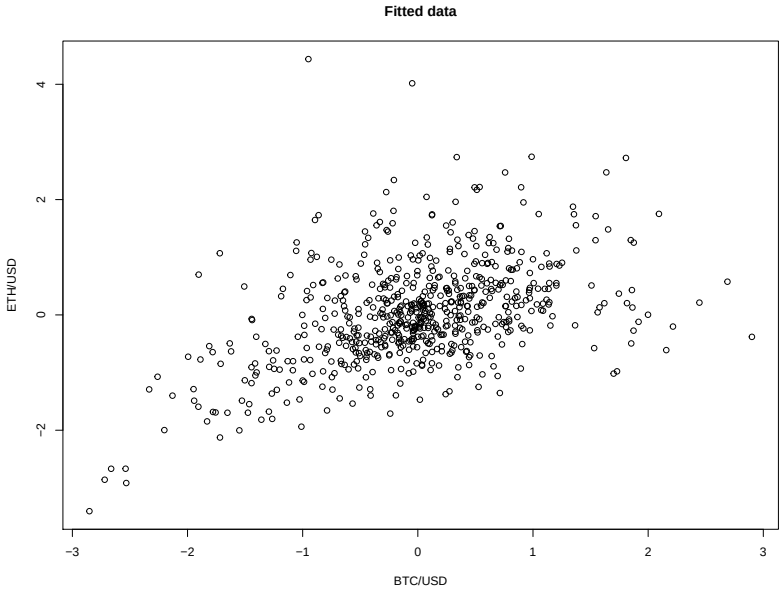
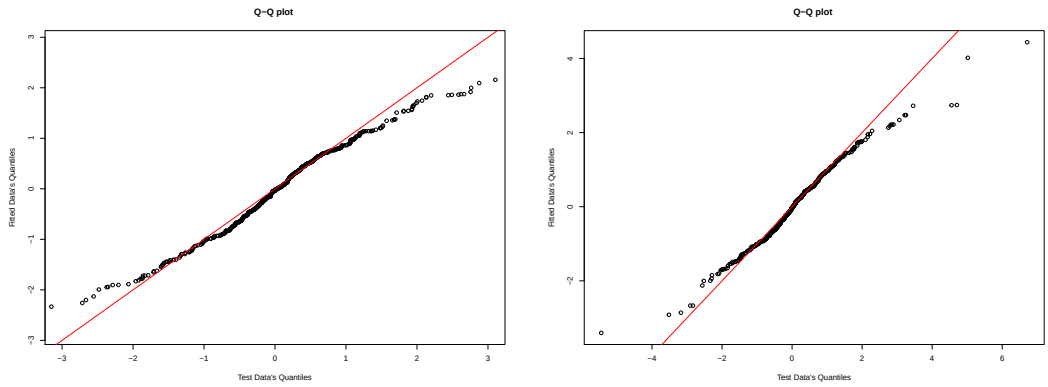


Figure 4.32: In figures, x-axis represents the first component BCT/USD and y-axis represents the second component ETH/USD.

Next, we check the goodness of fit. We begin by simulating 645 observations from the fitted model in Figure 4.32. In Figure 4.33, we give Q-Q plots comparing the quantiles of the testing and fitted data. And the plots of CDF of test data and fitted data in each component are shown in Figure 4.34. Next, we perform a KS test for each component separately. For the first component we get the test statistic $D = 0.057$ and the p-value $= 0.239$. For the second component we get the test statistic $D = 0.065$

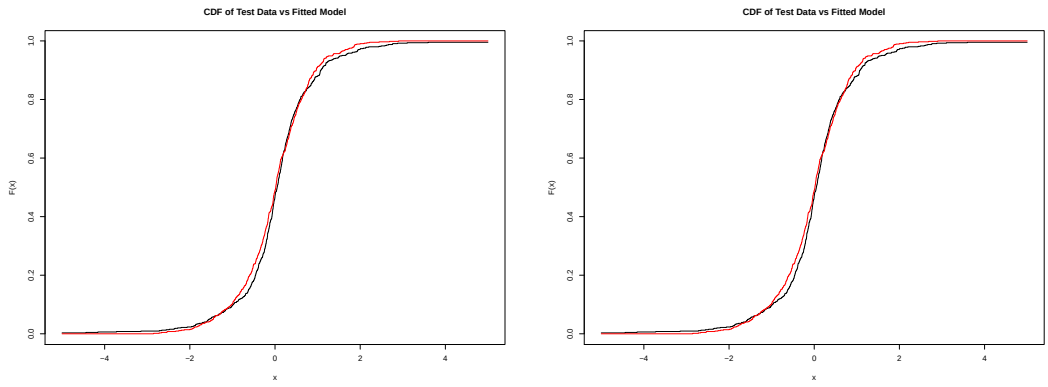


(a) (b)

Figure 4.33: Fitted TS Q-Q plots for the (a) BTC/UDS and (b) ETH/USD data. These compare the quantiles of the testing data and the fitted data. The diagonal (red) line is the line $y = x$.

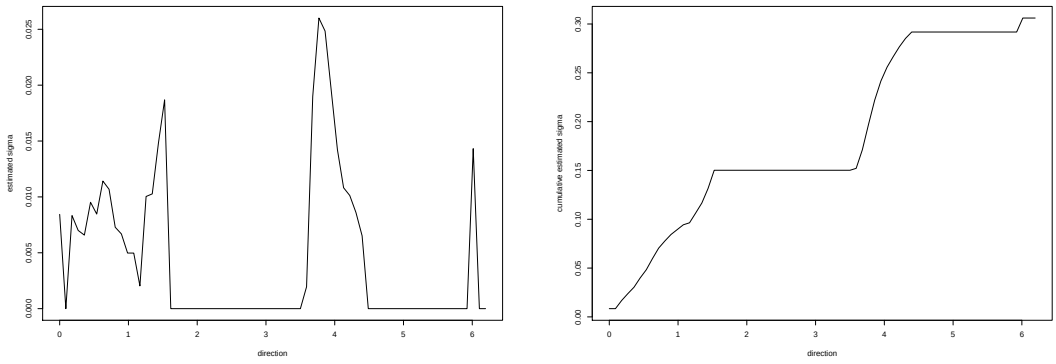
and the p-value $= 0.130$. We then test the two components together with the kernel consistent density equality test, which gives the test statistic $Tn = -10.120$ and the p-value $= 0.111$. While the fit is not as good as in the previous example, it is much better than the normal distribution and we cannot reject the assumption that the data comes from a bivariate tempered stable distribution.

For comparison, we again fit a multivariate stable distribution. Here, the minimum value of the objective function was $= 7.436 \times 10^{-3}$, the estimated value of α was $\hat{\alpha} = 0.916$, the estimated drift was $\hat{\gamma} = (-0.031, 0.010)$, and the 70 estimated values of the a_j's were all in the interval $(0, 3 \times 10^{-2})$. Plots of the estimated values of the a_j's and of the cumulative values for the estimated σ are given in Figure 4.35. The shape of the spectral measure is again similar to what was found when fitting the tempered stable distribution, but with a much smaller magnitude. The estimated value of α is again much higher than for the TS distribution.



(a) (b)

Figure 4.34: These are CDF for the (a) BTC/USD and (b) ETH/USD data. Black lines represents test data and Red line represents fitted data.



(a) (b)

Figure 4.35: Stable fit. The plot of the estimated a_j's is given in (a), while the plot of the cumulative function of the estimated σ is given in (b). The x-axis represents the direction from 0 to 2π.

We again simulated 645 observations from the fitted stable model, which is shown in Figure 4.36, and used it to check the goodness-of-fit. In Figure 4.37 we give Q-Q plots, which compare the quantiles of the testing and the fitted data. Also, in Figure 4.38, we compare CDF of test data and the fitted data. Next, we perform a KS test for each component separately. For the first component we obtained the test statistic $D = 0.092$ and the p-value 0.007 and for the second component we obtained the test statistic $D = 0.114$ and the p-value $< 10^{-3}$. We also performed the kernel consistent

density equality test, but it again seems to fail possibly due to the very heavy tails of the stable distribution. From the Q-Q plots, the plot of CDF and the KS test it is clear that the tails of the stable distribution are too heavy and that the model does not provide a good fit.

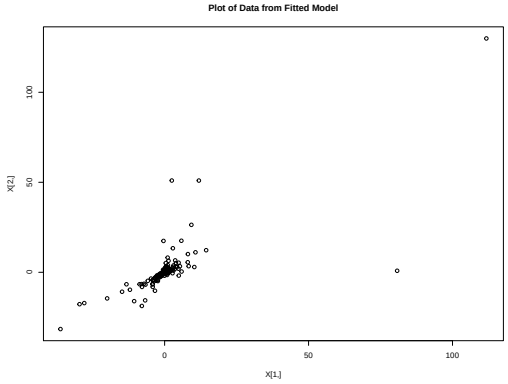
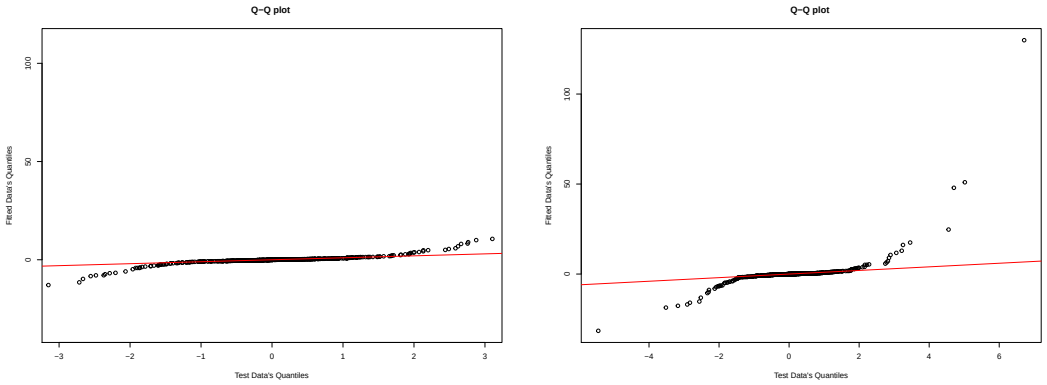
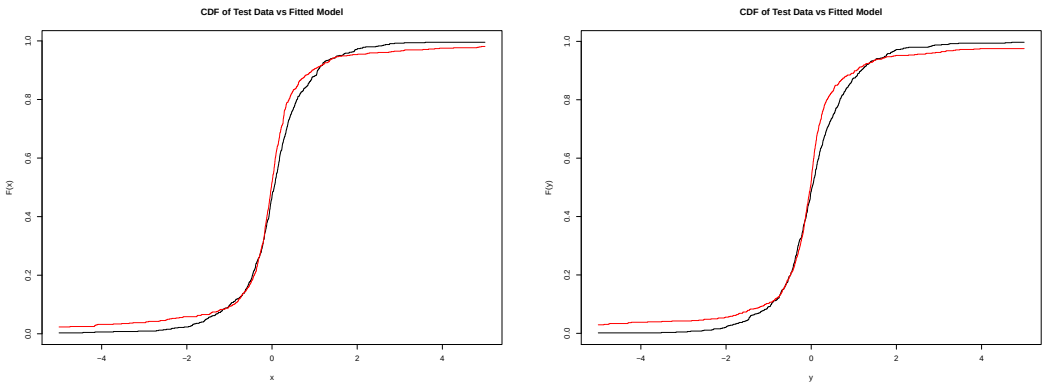


Figure 4.36: Stable fit. The plot of the estimated a_j's is given in (a), while the plot of the cumulative function of the estimated σ is given in (b). The x-axis represents the direction from 0 to 2π.



(a) (b)

Figure 4.37: Stable fit. These are the Q-Q plots for the (a) BTC/USD and (b) ETH/USD data. These compare the quantiles of the testing data with the fitted data. The diagonal (red) line is the line $y = x$.



(a) (b)

Figure 4.38: Stable fit. These are CDF for the (a) BTC/USD and (b) ETH/USD data. Black lines represents test data and Red line represents fitted data.

CHAPTER 5: OPTION PRICE

In this chapter, we discuss the application of multivariate TS distribution in the basket option. We consider the return from stocks to follow TS distribution and use TS process to derive the efficient pricing procedures for European call option under arbitrage-free. Hence, we show the existence of equivalent martingale measure first, then use the diagonal model to improve our methodology of parameter estimation, which leads to more effective computation. Then, we combine this with our model for the DS method for simulation to model real-world bivariate financial datasets related to the stock market. Towards the end, we give the pricing of European call options with different strikes and the pricing of the Multi-asset rainbow option. Here, we choose Google (GOOGL) and Facebook (FB) in here. All of the data was downloaded from Yahoo Finance, `http://finance.yahoo.com`.

5.1 Risk-neutral Measure

Fixed $d \geq 1$, and let $\Omega = D([0,T], \mathbb{R}^d)$ be a canonical probability space of càdlàg function from $[0,T]$ into $\mathbb{R}^d$ and it is right continuous with left limits, where T is a fixed final time. Let X_t is a lévy process as $\{X_t : t \geq 0\}$ define on a this space, let $\{S_t : t \geq 0\}$ be the price process of a dividend paying stock with dividend rate $q \geq 0$. We assume the stock process is $S_t = S_0 e^{X_t}$, where $S_0 > 0$, and it is the stock price at time $t = 0$. Then let $\mathscr{F} = \sigma(X_t : t \in [0,T])$, further $\mathscr{F}_t = \sigma(X_s : s \in [0,t])$, and $(\mathscr{F}_t)_{t \geq 0}$ is a filtration, which means the information based on S_t, so there is no information in $\mathscr{F}_0$ and $\mathscr{F}_T = \mathscr{F}$.

We assume probability measure $\mathbb{P}$ on the space $(\Omega, \mathscr{F})$, for here the $\mathbb{P}$ is estimated using historical return data from stock, we call it market measure or physical measure.

Under physical measure $\mathbb{P}$, the process $\{X_t : t \geq 0\}$ is a tempered stable process with $X_1 \sim \mathrm{TS}_\alpha(\sigma, b, \gamma)$. Since $X_{t+1} - X_t \overset{d}{=} X_1$, then log-returns $\log(\frac{S_{t+1}}{S_t}) = X_{t+1} - X_t$ which also follows $\mathrm{TS}_\alpha(\sigma, b, \gamma)$. According to Theorem 2 in [31], and also arbitrage-free claim prices and market completeness, there exists a probability measure $\mathbb{Q}$ on $(\Omega, \mathscr{F})$, which is equivalent to $\mathbb{P}$. Let r be the interest rate with $r \geq 0$, then under assumption of $r \geq q \geq 0$, if the discounted price process $(e^{-(r-q)t}S_t)_{t\in[0,T]}$ under measure $\mathbb{Q}$ is a martingale, like

$$e^{-(r-q)t}S_t = E_{\mathbb{Q}}[e^{-(r-q)u}S_u|\mathscr{F}_t], \qquad 0 \leq t \leq u \leq T, \tag{5.1}$$

we call measure $\mathbb{Q}$ is a risk-neutral probability measure.

We know probability measure $\mathbb{Q} \sim \mathbb{P}$ is equivalent, next we need to show how to use $\mathbb{P}$ to represent $\mathbb{Q}$. Here, we consider Esscher transfer to build Radon-Nikodym derivatives process to change $\mathbb{Q}$ to $\mathbb{P}$, its univariate form is defined in [31], and then extend it to high dimensional, which is

$$\left.\frac{\mathrm{d}\mathbb{Q}^\eta}{\mathrm{d}\mathbb{P}}\right|_{\mathscr{F}_t} = \frac{e^{\langle \eta, X_t\rangle}}{E[e^{\langle \eta, X_t\rangle}]}$$

where $\eta = (\eta_1, \eta_2, ..., \eta_d) \in \mathbb{R}^d$. Following this, we show that X_t, which under probability measure $\mathbb{Q}$, also is Tempered stable process.

Theorem 3. *Under $\mathbb{Q}^\eta$ measure, let $S = (s_1, s_2, ..., s_d) \in \mathbb{S}^{d-1}$ and $\eta = (\eta_1, \eta_2, ..., \eta_d) \in \mathbb{R}^d$ s.t $b(s) - \langle s, \eta\rangle \geq 1$ $\forall s \in \mathbb{S}^{d-1}$. Then $\{X_t : t \geq 0\}$ is a Tempered stable with $X_1 \sim \mathrm{TS}_\alpha(\sigma, b(s) - \langle s, \eta\rangle, \gamma)$.*

Proof. We begin to change of measure, and have $E_{\mathbb{Q}^\eta}[g(X_t)] = E_{\mathbb{P}}[g(X_t)Z_t]$, hence

$$\begin{aligned} E_{\mathbb{Q}^\eta}(e^{\langle z, X_t\rangle}) &= E_{\mathbb{P}}(e^{\langle z, X_t\rangle}Z_t) \\ &= E_{\mathbb{P}}\left(e^{\langle z, X_t\rangle}\frac{e^{\langle \eta, X_t\rangle}}{E_{\mathbb{P}}(e^{\langle \eta, X_t\rangle}}\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{E_{\mathbb{P}}(e^{\langle \eta, X_t\rangle})} E_{\mathbb{P}}(e^{\langle z, X_t\rangle + \langle \eta, X_t\rangle}) \\
&= \exp\left\{ t \left[\langle \gamma, z+\eta\rangle + \int_{\mathbb{S}^{d-1}} \int_0^\infty (e^{\langle s, z+\eta\rangle x} - 1) \frac{e^{-b(s)x}}{x^{1+\alpha}} \mathrm{d}x\sigma(\mathrm{d}s) \right.\right. \\
&\left.\left. - \langle \gamma, \eta\rangle - \int_{\mathbb{S}^{d-1}} \int_0^\infty (e^{\langle s, \eta\rangle x} - 1) \frac{e^{-b(s)x}}{x^{1+\alpha}} \mathrm{d}x\sigma(\mathrm{d}s) \right] \right\} \\
&= \exp\left\{ t \left[\langle \gamma, z\rangle + \int_{\mathbb{S}^{d-1}} \int_0^\infty (e^{\langle s, z+\eta\rangle x} - e^{\langle s, \eta\rangle x}) \frac{e^{-b(s)x}}{x^{1+\alpha}} \mathrm{d}x\sigma(\mathrm{d}s) \right] \right\} \\
&= \exp\left\{ t \left[\langle \gamma, z\rangle + \int_{\mathbb{S}^{d-1}} \int_0^\infty (e^{\langle s, \eta\rangle x + \langle s, z\rangle x} - e^{\langle s, \eta\rangle x}) \right.\right. \\
&\left.\left. \frac{e^{-b(s)x}}{x^{1+\alpha}} \mathrm{d}x\sigma(\mathrm{d}s) \right] \right\} \\
&= \exp\left\{ t \left[\langle \gamma, z\rangle + \int_{\mathbb{S}^{d-1}} \int_0^\infty (e^{\langle s, z\rangle x} - 1) \frac{e^{-(b(s)x - \langle s, \eta\rangle x)}}{x^{1+\alpha}} \mathrm{d}x\sigma(\mathrm{d}s) \right] \right\},
\end{aligned}
$$

gives the result. □

Next, we give the condition to always keep the probability measure $\mathbb{Q}$ to be the risk neutral measure since the risk neutral measure may not exists.

Theorem 4. *Let $X(t) = (X_{1,t}, X_{2,t}, ..., X_{d,t})$ be a d-dimensional random variable, which follows $\mathrm{TS}_\alpha(\sigma, b(s) - \langle s, \eta\rangle, \gamma)$. Under probability measure $\mathbb{Q}^\eta$, $\mathbb{Q}^\eta$ is a risk-neutral measure if and only if*

$$
r - q = \gamma_j + \int_{\mathbb{S}^{d-1}} \Gamma(-\alpha)[((b(s) - \langle s, \eta\rangle) - s_j)^\alpha - (b(s) - \langle s, \eta\rangle)^\alpha]\sigma(\mathrm{d}s),
$$

where $j = 1, 2, ..., d$.

Proof. Let $S_t = (S_{1,t}, S_{2,t}, ..., S_{d,t})$ be $d-$dimensional vector, which represent d different stocks' price at time t. Since $S_t = S_0 e^{X_t}$, considering the jth stock price $S_{j,t}$, which is $S_{j,t} = S_0 e^{X_{j,t}}$. So,

$$
\begin{aligned}
E_{\mathbb{Q}^\eta}(e^{-(r-q)u} S_{j,u} | \mathscr{F}_t) &= e^{-(r-q)u} E_{\mathbb{Q}^\eta}(S_{j,u} | \mathscr{F}_t) = e^{-(r-q)u} E_{\mathbb{Q}^\eta}(S_0 e^{X_{j,u}} | \mathscr{F}_t) \\
&= e^{-(r-q)u} S_0 E_{\mathbb{Q}^\eta}(e^{X_{j,u} - X_{j,t} + X_{j,t}} \big| \mathscr{F}_t)
\end{aligned}
$$

$$= e^{-(r-q)u} S_0 e^{X_t} E_{\mathbb{Q}^\eta}(e^{X_{j,u}-X_{j,t}}|\mathscr{F}_t)$$
$$= e^{-(r-q)u} S_{j,t} E_{\mathbb{Q}^\eta}(e^{X_{j,u}-X_{j,t}})$$

If probability measure $\mathbb{Q}^\eta$ is a risk-neutral measure, for any $0 \leq t \leq u \leq T$, the equation (5.1) exists, which indicates that $e^{-(r-q)t} S_{j,t} = E_{\mathbb{Q}^\eta}[e^{-(r-q)u} S_{j,u}|\mathscr{F}_t]$ exists. Furthermore, $E_{\mathbb{Q}^\eta}(e^{-(r-q)u} S_{j,u}|\mathscr{F}_t) = e^{-(r-q)u} S_{j,t} E_{\mathbb{Q}^\eta}(e^{X_{j,u}-X_{j,t}}) = e^{-(r-q)t} S_{j,t}$, which implies that $E(e^{X_{j,u}-X_{j,t}}) = e^{(r-q)(u-t)}$.

Since $X_{j,u} - X_{j,t} \overset{d}{=} X_{j,u-t}$ and,

$$\hat{\mu}_{X_{j,u-t}}(z) = \exp\Bigg[(u-t)\big(i\gamma_j z_j + \int_{\mathbb{S}_{d-1}} \Gamma(-\alpha)((b(s) - \langle s, \eta\rangle) - is_j z_j)^\alpha$$
$$-(b(s) - \langle s, \eta\rangle)^\alpha\big)\sigma(\mathrm{d}s)\Big].$$

We have

$$E(e^{X_{j,u-t}}) = \hat{\mu}_{X_{j,u-t}}(-i),$$

then

$$E(e^{X_{j,u-t}}) = \exp\Bigg[(u-t)\big(\gamma_j + \int_{\mathbb{S}_{d-1}} \Gamma(-\alpha)((b(s) - \langle s, \eta\rangle) - s_j)^\alpha$$
$$-(b(s) - \langle s, \eta\rangle)^\alpha\big)\sigma(\mathrm{d}s)\Big],$$

which is equal to $e^{(r-q)(u-t)}$. Finally, we get,

$$r - q = \gamma_j + \int_{\mathbb{S}^{d-1}} \Gamma(-\alpha)[((b(s) - \langle s, \eta\rangle) - s_j)^\alpha - (b(s) - \langle s, \eta\rangle)^\alpha]\sigma(\mathrm{d}s),$$

which is the condition of $\mathbb{Q}^\eta$ is a martingale measure. □

5.2 Methodology of Option Pricing

After getting the equivalent martingale measure, we move to the model of parameter estimations. The real data analysis in Section 4.2, we choose the tuning parameter $k = 70$, which exists a high computation burden, since the cost is very expensive when the number of directions is large. Hence, we want to improve our methodology by reducing the number of required estimated parameters. We consider the Diagonal model here and combine this with our moethodology mentioned in Chapter 4.1 to do the parameter estimation. Let's see the Diagonal model first.

> "Its assumption is that the common movement in all assts is due to one common factor only in the market " *Value at risk: the new benchmark for managing financial risk* (page 169 in [42])

Diagonal model is proposed by Sharpe [43] in 1963. Formally, the model is

$$R_i = \alpha_i + \lambda_i R_m + \varepsilon_i,$$

where R_i is the return on asset i, R_m is the return to the market portfile, α_i is the abnormal return and λ_i is responsiveness to the market return and ε_i is residual returns. The model means that the return on asset i is driven by the market return R_m and ε_i, which is not correlated with the market. The diagonal model is a simple asset pricing model to measure both the risk and the return of assets. Also, it could be vastly reduced the number of estimated parameters and lead to the pricing model more efficiently.

To set up our model, we let $S_1, ..., S_d$ are return process, $X_1 - X_2, X_3 - X_4, \ldots,$ $X_{2d-1} - X_{2d}$ are individual risk factors, and $X_{2d+1} - X_{2d+2}$ is a market risk factor. Since the random variable from TS distributions are always positive, to let the model

more flexible, we then extend the Diagonal model to a negative part, as follows

$$\begin{aligned} S_1 &= \gamma_1 + (X_1 - X_2) + \beta_1(X_{2d+1} - X_{2d+2}) \\ S_2 &= \gamma_2 + (X_3 - X_4) + \beta_2(X_{2d+1} - X_{2d+2}) \\ \vdots &\ \vdots\ \ \vdots\ \ \vdots\ \ \vdots\ \ \vdots\ \ \vdots\ \ \vdots \\ S_d &= \gamma_d + (X_{2d-1} - X_{2d}) + \beta_d(X_{2d+1} - X_{2d+2}) \end{aligned}$$

where $X_i - X_{i+1}$ are the excess return on the asset i and $X_{2d+1} - X_{2d+2}$ are the excess return on the market, γ_i is a constant depend on the asset i , β_i is a coefficient of $X_{2d+1} - X_{2d+2}$.

Then based on the diagonal model and theorem 2, given $Y = SX + \gamma$, let Y be a $d-$dimensional TS variable. Let S be a $d \times (2d+2)$ matrix and

$$S = \begin{pmatrix} 1 & -1 & \dots & 0 & 0 & \beta_1 & -\beta_1 \\ 0 & 0 & \dots & 0 & 0 & \beta_2 & -\beta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 & \beta_d & -\beta_d \end{pmatrix}$$

S is composed of two parts. There are diagonal matrix and parameters $\beta = (\beta_1, \dots, \beta_d)^T$. For the aspect of diagonal matrix part, there is a little different with the traditional diagonal matrix, the matrix S extends to the negative, as the random variables of tempered stable could up and down like stock price's path. The other part, parameter $\beta \in \mathbb{S}^{d-1}$ and β should satisfy $|\beta| = 1$. Also, let $X = (X_1, \dots, X_{2d+2})^T$, where $X_1, \dots, X_{2d+2}$ are independent random variables with $X_j \sim \mathrm{STS}_\alpha(a_j, b_j)$ and $b_j = b(s_j)$. By Theorem 2, we then have $Y \sim \mathrm{TS}_\alpha(\sigma, b, \gamma)$, where $\gamma = (\gamma_1, \dots, \gamma_d)^T$ is drift. We give an example of bivariate case.

Example 3. *When $d = 2$, let* $S = \begin{pmatrix} 1 & -1 & 0 & 0 & \beta_1 & -\beta_1 \\ 0 & 0 & 1 & -1 & \beta_2 & -\beta_2 \end{pmatrix}$, $X = (X_1, X_2, \dots, X_6)^T$,

$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$. *Then we have*

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 & \beta_1 & -\beta_1 \\ 0 & 0 & 1 & -1 & \beta_2 & -\beta_2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ : \\ X_6 \end{pmatrix} = \begin{pmatrix} X_1 - X_2 + \beta_1(X_5 - X_6) \\ X_3 - X_4 + \beta_2(X_5 - X_6) \end{pmatrix}.$$

If we consider that Y_1 and Y_2 are return of assets, $X_1 - X_2$, $X_3 - X_4$ and $X_5 - X_6$ are independent and $\beta_1(X_5 - X_6)$ and $\beta_2(X_5 - X_6)$ correspond on the market.

Followed by Diagonal model, our directions has already reduced to $6d+5$, included $(2d+2)$'s a_j and b_j, α, d's γ_j and d's β_j. To further reduce the number of estimated parameters, we consider using following equation to estimate the drift.

$$\mathrm{E}[X] \;=\; \gamma + \Gamma(1-\alpha) \int_{\mathbb{S}^{d-1}} (b(s))^{\alpha-1} s \sigma(\mathrm{d}s), \tag{5.2}$$

If we normalized data, then the dataset has zero mean and variance of one, then based on equation (5.2) we get

$$\gamma = -\Gamma(1-\alpha) \int_{\mathbb{S}^{d-1}} (b(s))^{\alpha-1} s \sigma(\mathrm{d}s), \tag{5.3}$$

hence, we only need to estimate $5d+5$ parameters. And in bivariate case, regularly, we consider the $\beta = (\cos\theta, \sin\theta)$, indicated that we only need to find the θ. In the end, under the bivariate case, there are only reminders 14 parameters needed to estimate, which is much less than 143.

Since we consider normalizing the data first during data analysis, here we provide the theoretical result about how to transfer between the normalized data and the original data set. We assume $Y \sim \mathrm{TS}_\alpha(\sigma, b, \gamma)$, then $X = \Sigma_X Y + \mu_X$, where Σ_X is a

covariance matrix of X and μ_X is the mean of X. To find X's distribution, we have

$$\begin{aligned}
\mathrm{E}[e^{i\langle z,\ X\rangle}] &= \exp\left[i\left\langle\gamma, \Sigma_X z+\mu_X\right\rangle + \int_{\mathbb{S}^{d-1}}\int_0^\infty \left(e^{it\langle \Sigma_X z,\ s\rangle}-1\right) t^{-1-\alpha} e^{-tb(s)}\mathrm{d}t\sigma(\mathrm{d}s)\right] \\
&= \exp\left[i\left\langle\Sigma_X\gamma+\mu_X, z\right\rangle\right. \\
&\quad + \left.\int_{\mathbb{S}^{d-1}}\int_0^\infty \left(e^{iu\left\langle z,\ \frac{\Sigma_X s}{|\Sigma_X s|}\right\rangle}-1\right) u^{-1-\alpha} e^{\frac{-ub(s)}{|\Sigma_X s|}}\mathrm{d}u\ |\Sigma_X s|^\alpha\ \sigma(\mathrm{d}s)\right] \\
&= \exp\left[\int_{\mathbb{S}^{d-1}}\int_0^\infty \left(e^{iu\langle z,\ s'\rangle}-1\right) u^{-1-\alpha} e^{\frac{-ub(s)}{|\Sigma_X s|}}\mathrm{d}u\ |\Sigma_X s|^\alpha\ \sigma'(\mathrm{d}s')\right. \\
&\quad \left. + i\left\langle\Sigma_X\gamma+\mu_X, z\right\rangle\right]
\end{aligned}$$

where in the second line, let $u = t|\Sigma_X s|$, then $\mathrm{d}u = \mathrm{d}t|\Sigma_X s|$. In the last line, let $s' = \dfrac{\Sigma_X s}{|\Sigma_X s|}$, since $\sigma(\mathrm{d}s) = \sum_{n=1}^k \delta_{s_i}(ds)$, then we have $\sigma'(\mathrm{d}s') = \sum_{n=1}^k \delta_{\frac{\Sigma_X s}{|\Sigma_X s|}}(ds')$. Hence, we get $X \sim \mathrm{TS}_\alpha(|\Sigma_X s|^\alpha\sigma', \frac{b(s)}{\Sigma_X s}, \Sigma_X\gamma+\mu_X)$.

Next, we will give a brief introduction to our model for the pricing of the bivariate basket option. We first get the data set of log returns from real data and randomly split raw data into two halves. The first is the training data, which will be used to fit the model, while the second is the testing data, which will be used for analyzing the goodness of fit. To easily and effectively get the parameters, let X be the training data or testing data. Then we normalize the training data and testing data separately in each component by $\dfrac{X_i-\mu_{X_i}}{\sigma_{X_i}}$ to get the new training data and testing data, where σ_{X_i} is a variance of X_i and μ_{X_i} is the mean of X_i with $i = 1, 2$. Let's call them transferred training data and transferred testing data. Based on the transferred datasets, combing the Diagonal model with our model for parameter estimation in Section 4.1 and the DS method for simulation, we get the estimated parameters and do the goodness-of-fit. Further, followed results $X \sim \mathrm{TS}_\alpha(|\Sigma_X s|^\alpha\sigma', \frac{b(s)}{\Sigma_X s}, \Sigma_X\gamma+\mu_X)$, we get the distribution of the training data. After that, we do the goodness of fit with the testing data to verify whether TS distribution is a reasonable model or not. If the estimated parameters work, according to the Theorem 4, we get the parameter η to

find another equivalent TS distribution under the measure $\mathbb{Q}$ and keep the measure $\mathbb{Q}$ is a risk-neutral measure. Followed that, we generate X_t and develop a Monte Carlo based method for option pricing and apply it to the pricing of European call options with different strikes by equation (5.4). The formula of European call option is as follows,

$$\pi = e^{-rT}\mathbb{E}_{\mathbb{Q}}[S_T - K^+], \tag{5.4}$$

where $S_T = S_0 e^{X_t}$ and S_0 is the stock price at $t = 0$. And we can estimate equation (5.4) using the following algorithm [44],

Monte Carlo Algorithm for Option Pricing.

for $i = 1, \ldots, n$

generate $X_T^{(i)} \overset{\text{iid}}{\sim} \text{TS}_{\hat{\alpha}}(T\hat{\sigma}, \hat{b_{est}}, T\hat{\gamma}')$ with $i = 1, 2, \ldots, n$

Set $S_T^{(i)} = S_0 e^{X_T^{(i)}}$

Set $\pi_i = e^{-rT}(S_T^{(i)} - K)^+$

$\pi = \dfrac{1}{n}\sum_{i=1}^{n} \pi_i$

Also, we do the pricing of the Multi-asset rainbow option. We consider the call option on the minimum of two assets payoff, based on Equation (5.5) and the call option on the maximum of two assets payoff, based on Equation (5.6)

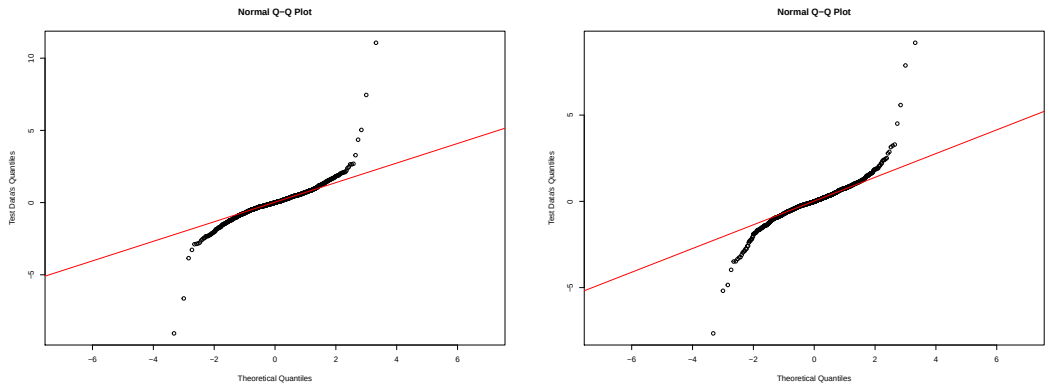
$$\pi_{\min} = \max(\min(S_T^{[1]}, S_T^{[2]}) - K, 0) \tag{5.5}$$

$$\pi_{\max} = \max(\max(S_T^{[1]}, S_T^{[2]}) - K, 0), \tag{5.6}$$

where $S_T^{[i]}$ with $i = 1, 2$ is the stock price at time T and K is a strike price.

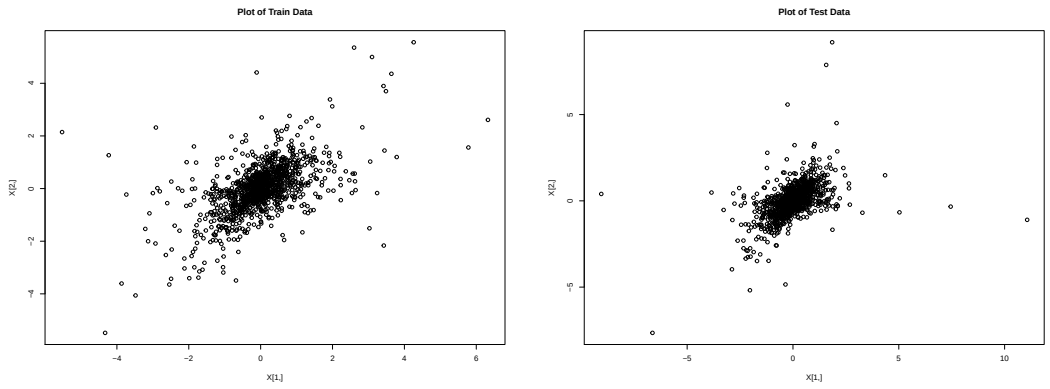
5.3 Real data analysis for the bivariate case

Similar to Section 4.2.3, we did the parameter estimation first. We begin by jointly modeling FB and GOOGL. The data consists of the daily closing prices for the period from May 31, 2012 to March 25, 2021 for each stock. The prices are converted to log returns by taking R_t, where $R_t = \log(P_t/P_{t-1})$ and P_t is the price at time t, are then considered as ordered pairs with the first component corresponding to FB and the second to GOOGL for the same day. In total, the data consists of 2516 ordered pairs. Then we randomly split raw data into two halves: training data and testing data, each consisting of 1109 data points. The first question is to see if a multivariate normal model will work for this data. For this reason, Figure 5.1(a) and 5.1(b) give normal Q-Q plots for each component separately in the testing data. We also performed an adjusted Jarque-Bera test for normality on each component of the testing data. In both cases, the p-value was less than 10^{-16}. Clearly, the normal distribution is not reasonable for either component. Instead, we follow our methodology for fitting a bivariate TS distribution. After normalized in each component, we have the transferred training data, which plots in Figure 5.2(a) and transferred testing data, which plots in Figure 5.2(b).



(a) (b)

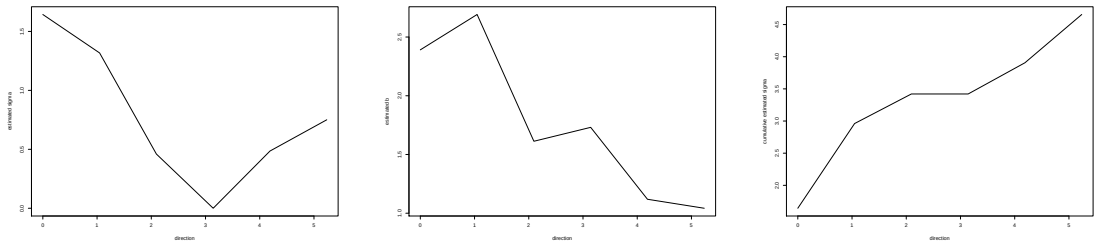
Figure 5.1: Normal Q-Q plots for the testing data. The plot in a) is for the first component FB and the plot in b) is for the second component GOOGL.



(a) (b)

Figure 5.2: Both of these plots have the same sample size 1258. The left plot shows train data set. The right plot shows test data. These two figures' xlab represents the normalized log returns of FB and ylab represents the normalized log returns of GOOGL.

Based on diagonal model, we discretized σ into $k = 6$ pieces on $\mathbb{S}^1$, which implies $S = \begin{pmatrix} 1 & -1 & 0 & 0 & \beta_1 & -\beta_1 \\ 0 & 0 & 1 & -1 & \beta_2 & -\beta_2 \end{pmatrix}$, and $\beta_1 = \cos 2\pi\theta$, $\beta_2 = \sin 2\pi\theta$, then $S = \begin{pmatrix} 1 & -1 & 0 & 0 & \cos 2\pi\theta & -\cos 2\pi\theta \\ 0 & 0 & 1 & -1 & \sin 2\pi\theta & -\sin 2\pi\theta \end{pmatrix}$ and we must fit 14 parameters. For this reason's, we took $m = 14$ to be the number of z_l's. We chose these to be evenly spaced on $\mathbb{S}^1$. Next, we fitted the parameters using the transferred training data by minimizing the objective function given in (4.1). The perform the optimization we first used Particle Swarm Optimization [45] as implemented in the hydroPSO function of the "hydroPSO" R package, to get initial values. These were then plugged into the optim function in R with the L-BFGS-B option. During the optimization, we set up the lower bound for a, b are 0, for α is 0.01, and for θ is 0, also set up the upper bound for a, b, are 6, α is 0.99, and θ is 1. After optimization, the value of the objective function was $9.330 * 10^{-5}$. The estimated value of a_j is $\hat{a} = (1.644,\ 1.317,\ 0.460,\ 0,\ 0.485,\ 0.750)$, the estimated value of b_j is $\hat{b} = (2.391,\ 2.692,\ 1.613,\ 1.731,\ 1.119,\ 1.042)$, $\hat{\alpha}$ is 0.014 and $\hat{\theta} = 0.992$. Plots of the estimated values of the a_j's, the b_j's, and of the cu-



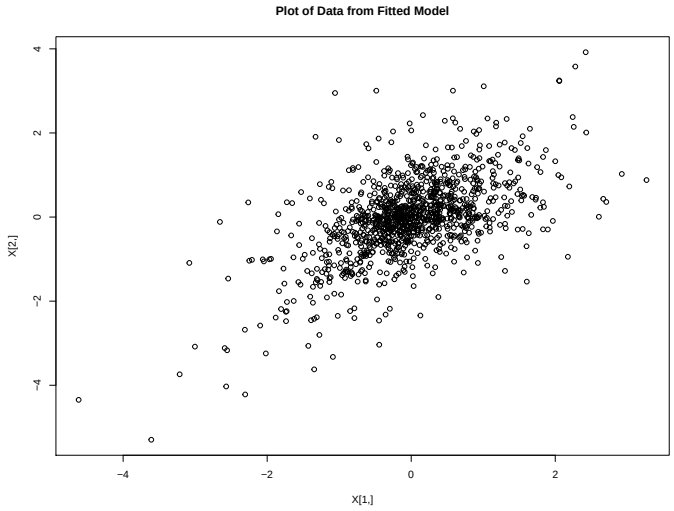
(a) (b) (c)

Figure 5.3: Estimated values for FB and GOOGL data. Plots of the estimated a_j's and b_j's are given in (a) and (b), respectively. A plot of the cumulative function of the estimated spectral measure is given in (c). The x-axis in all plots represents the direction from 0 to 2π.

mulative values for the estimated σ are given in Figure 5.3. Followed by equation (5.3), then we get $\hat{\gamma} = (-0.044,\ -0.048)$. Besides, based on the value of $\hat{\theta}$, we get $\beta_1 = 0.547,\ \beta_2 = 0.837$. Hence,

$$S = \begin{pmatrix} 1 & -1 & 0 & 0 & 0.547 & -0.547 \\ 0 & 0 & 1 & -1 & 0.837 & -0.837 \end{pmatrix}$$

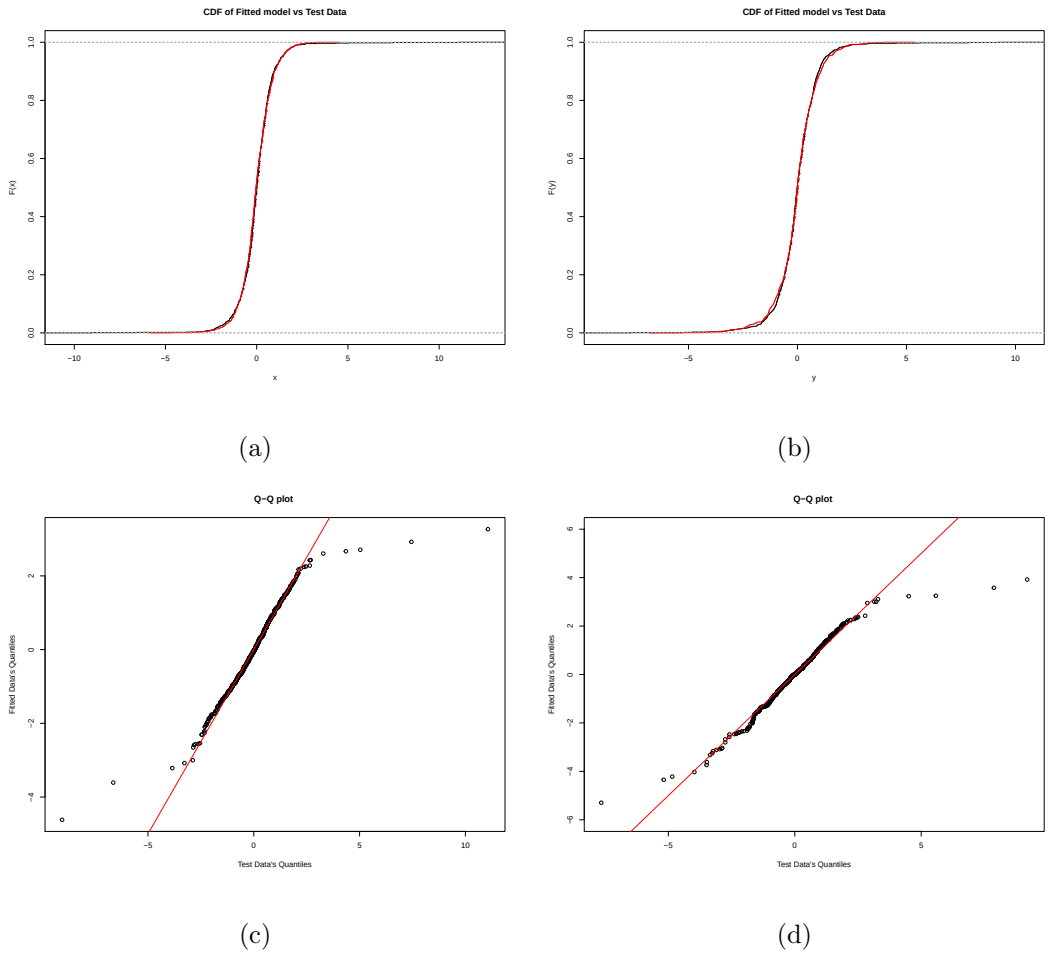
.



(a)

Figure 5.4: The same sample size 1258. In figures, x-axis represents the first component FB and y-axis represents the second component GOOGL.

Next, we check the goodness of fit. We simulated 1258 observations from the fitted model, which is the same as the number of observations in the transferred testing data. We call this the transferred fitted data, which is plotted in Figure 5.4. In Figure 5.5 we gave several diagnostic plots, which compare cdfs and quantiles of the transferred testing and transferred fitted data. We also performed formal goodness-of-fit testing. First, we performed KS tests comparing each component of the transferred testing and transferred fitted data separately. For the first component we obtained the test statistic $D = 0.039$ and the p-value $= 0.375$ and for the second component we obtained the test statistic $D = 0.026$ and the p-value $= 0.843$. Next, we tested both components together, using the kernel consistent density equality test and the test statistic is $Tn = -35.534$ and the p-value $= 0.778$. The results of our goodness of fit plots and tests suggest that the bivariate TS distribution is a reasonable model for this data.

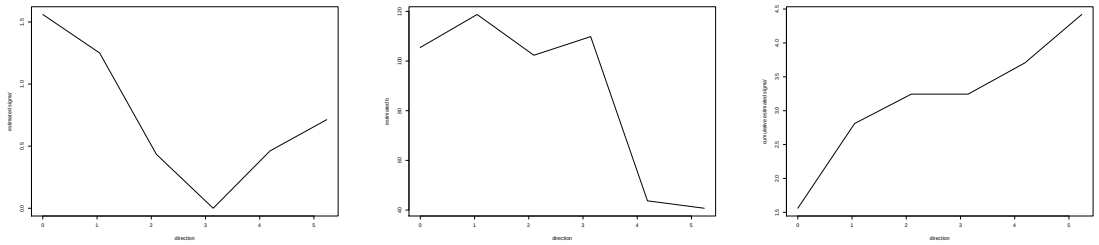


(a) (b)

(c) (d)

Figure 5.5: TS CDF plots for the (a) FB and (b) GOOGL data, red line represents CDF of transferred fitted data, and black line represents CDF of the transferred testing data. Figure(c) and (d) are fitted TS Q-Q plots for the (a) FB and (b) GOOGL data. These compare the quantiles of the transferred testing data with the transferred fitted data. The diagonal (red) line is the line $y = x$.

To find FB and GOOGL log-returns' distribution, we need to transfer the transferred testing data and training data back to log return's status. Since we have $X \sim \mathrm{TS}_\alpha(|\Sigma_X s|^\alpha \sigma', \frac{b(s')}{\Sigma_X s}, \Sigma_X \gamma + \mu_X)$, and knew the estimated value of α, b, σ, and γ, which denoted by $\hat{\alpha}$, $\hat{a}$, $\hat{b}$ and $\hat{\gamma}$, we get the estimated parameter of X's distribution. We notate $X \sim \mathrm{TS}_{\hat{\alpha}}(\hat{\sigma'}, \hat{b'}, \hat{\gamma'})$, $\hat{a'} = |\Sigma_X s|^{\hat{\alpha}} \hat{a}$, $\hat{b'} = \dfrac{\hat{b}}{\Sigma_X s}$ and $\hat{\gamma'} = \Sigma_X \hat{\gamma} + \mu_X$, then we get the parameter $\hat{a'} = (1.560,\ 1.250,\ 0.434,\ 0.000,\ 0.461,\ 0.713)$; $\hat{b'} = (105.465,\ 118.740,\ 102.327,\ 109.843,\ 43.709,\ 40.720)$ and $\hat{\gamma'} = (-9.470 * 10^{-5},\ 3.816 * 10^{-5})$. The cumulative estimated $\hat{a'}$ is shown in Figure 5.6, comparing with Figure



(a) (b) (c)

Figure 5.6: Estimated values for FB and GOOGL data. Plots of $\hat{a}'$ and $\hat{b}'$ are given in (a) and (b), respectively. A plot of the cumulative function of the estimated spectral measure is given in (c). The x-axis in all plots represents the direction from 0 to 2π.

5.3 they are really similar. And currently, $S' = \dfrac{\Sigma_X s}{|\Sigma_X s|}$, then

$$S' = \begin{pmatrix} 1 & -1 & 0 & 0 & 0.485 & -0.485 \\ 0 & 0 & 1 & -1 & 0.515 & -0.515 \end{pmatrix}$$

Using these new parameters, we simulated 1258 observations from the fitted model, which is the same as the number of observations in the testing data, which is plotted in Figure 5.8. Then we move on to check the goodness-of-fit. Again, in Figure 5.7, we gave several diagnostic plots, which compare cdfs and quantiles of the testing and fitted data. While the Q-Q plots exist some extreme points formed heavier tails, in contrast to normal distribution, bivariate TS distribution more lightly underestimates the chance of extreme value. We also performed formal goodness-of-fit testing. For the first component we obtained the test statistic $D = 0.050$ and the p-value $= 0.118$ and for the second component we obtained the test statistic $D = 0.038$ and the p-value $= 0.404$. Next, we tested both components together, using the kernel consistent density equality test. In this case, the test statistic is $Tn = -33.823$ and the p-value $= 0.778$. The results of our goodness-of-fit plots and tests are the same as before and have the same suggestion that the bivariate TS distribution is a reasonable model for this data.

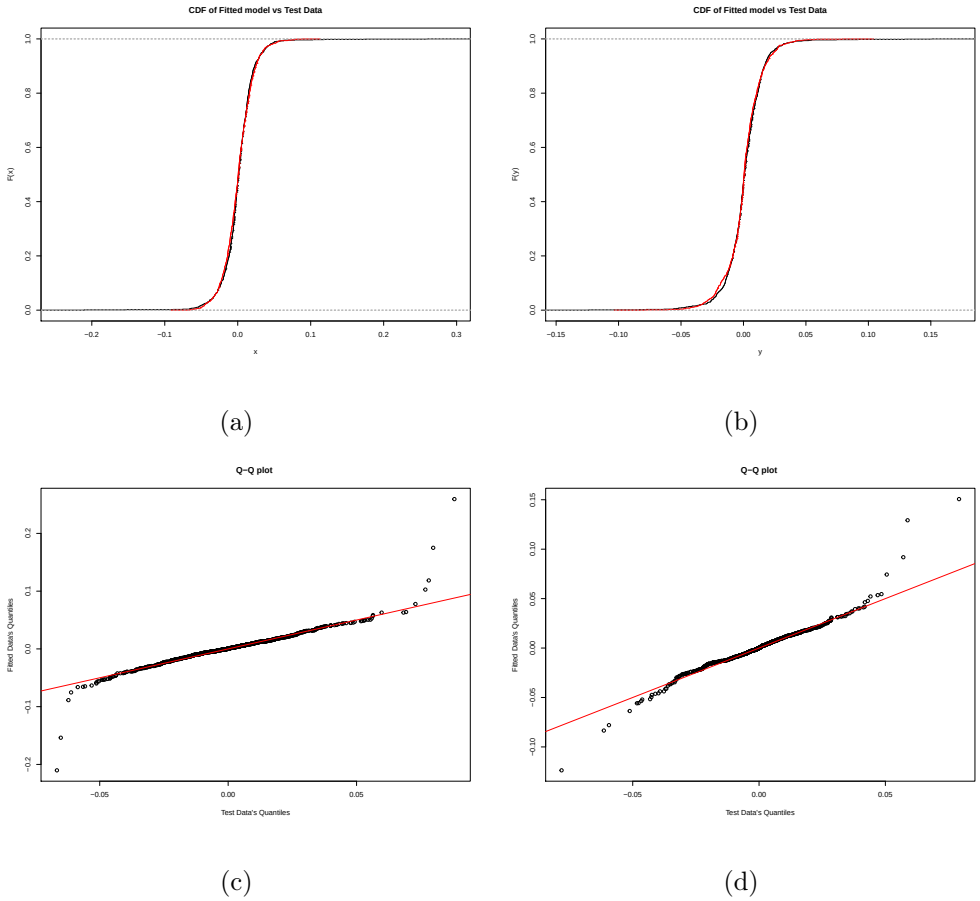


Figure 5.7: TS CDF plots for the (a) FB and (b) GOOGL data, red line represents CDF of fitted data, and black line represents CDF of the testing data. Figure(c) and (d) are the fitted TS Q-Q plots for the (a) FB and (b) GOOGL data. These compare the quantiles of the testing data with the fitted data. The diagonal (red) line is the line $y = x$.

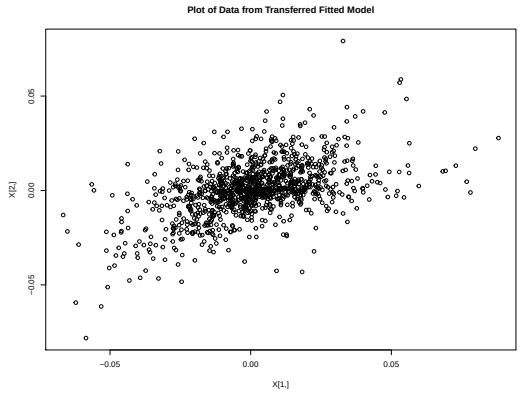


Figure 5.8: The same sample size 1258. In figures, x-axis represents the first component FB and y-axis represents the second component GOOGL.

Now we have got a suitable model to fit the R_t, then to price the options, we need to set up the risk-neutral measure first. We have known that under the real-world probability measure $\mathbb{P}$ the process R_t is TS process with the estimated parameter $\mathrm{TS}_{\alpha_{est}}(\hat{\sigma}, \hat{b'}, \hat{\gamma'})$. Also, we let the time t is measured in trading days, r is the daily interest rate and q is the dividend rate. Since we choose Facebook and Google in this case, both of their dividend rates are 0.

Then, we followed the Theorem 4 to transform the real-world probability measure $\mathbb{P}$ to risk-neutral measure $\mathbb{Q}$. To satisfy the equation existence, we set a new parameter η_{est} and $\hat{b_{est}}$, which is $\hat{b_{est}} = \hat{b'} - \langle \eta_{est}, S' \rangle$. Hence, the process $X_t \sim \mathrm{TS}_{\alpha_{est}}(\hat{\sigma}, \hat{b_{est}}, \hat{\gamma})$ which is under the risk-neutral measure.

Before we find the parameter η_{est}, we chose Oct 22, 2021 to Oct 29, 2021 to be the period of option, so the maturity date $T = 5$. And we use 13-week treasury bill to be annualized interest rate, the closed price on Oct 22, 2021 is \$0.05. Due to require daily interest rate in our experiment, $r = 0.05/252 = 1.984 \times 10^{-4}$. Now, we use the Theorem 4 again to calibrate the parameter η_{est}. Let's rewrite the equation, then the objective function becomes

$$\mathrm{argmin} \sum_{j=1}^{d} \left| r - \left\{ \gamma_j + \int_{\mathbb{S}^{d-1}} \Gamma(-\alpha)[((b(s) - \langle s, \eta \rangle) - s_j)^{\alpha} - (b(s) - \langle s, \eta \rangle)^{\alpha}] \sigma(ds) \right\} \right|.$$

In this equation, $\alpha = \hat{\alpha},\ b(s) = \hat{b'},\ \gamma = \hat{\gamma'},\ s = S',\ \sigma = \hat{\sigma}$. We used hydroPSO to optimize it, and set the boundary of η is from -50 to 50. The final result is $\eta_{est} = (-2.546, -2.488)$, at the same time, the value of objective function is $2.229 * 10^{-15}$, which means that the equation exists, also we think that the risk-neutral measure exists. Then the parameter b became $\hat{b_{est}} = (108.012,\ 116.194,\ 104.814,\ 107.355,\ 46.225,\ 38.204)$. For here, we know all parameters of $\mathrm{TS}_{\hat{\alpha}}(\hat{\sigma}, \hat{b_{est}}, \hat{\gamma'})$, and then drive process X_t, which is under the martingale measure $\mathbb{Q}$. In this process, by [35] Definition 1.6, we get $X_1 \sim \mathrm{TS}_{\hat{\alpha}}(\hat{\sigma}, \hat{b_{est}}, \hat{\gamma'})$, then $X_t \sim \mathrm{TS}_{\hat{\alpha}}(t\hat{\sigma}, \hat{b_{est}}, t\hat{\gamma'})$. We simulate 5000 obser-

Table 5.1

Case 1					
S_0 / Stock	Real time	Closed Price	Low Price	High Price	True Value
FB (K=210)	\$115.90	\$113.72	\$110.67	\$119.11	\$114.3
GOOGL (K=2250)	\$507.63	\$522.04	\$492.98	\$581.18	\$508

Table 5.2

Case 2					
S_0 / Stock	Real time	Closed Price	Low Price	High Price	True Value
FB (K=250)	\$73.74	\$73.76	\$70.71	\$79.15	\$72.17
GOOGL (K=2500)	\$277	\$272.5	\$243.64	\$331.47	\$279

vations to be X_t. Same as the training data, the first component corresponds to FB and the second component corresponds to GOOGL. Since from yahoo finance, each option has the last trade date and it is accurate to time, to find the fair option price, we consider different stock prices S_0 to calculate. The choices are the real time stock quotes at t_0, the closed price, the lowest price and the highest price, where $t = 0$ is the first day in period which is Oct 22, 2021. Under the same strike price K, we use Algorithm of Monte Carlo in 5.2 to get the price of European call option π.

We consider one dimensional case first, and then get each stock's option price. Our final results are in Table 5.1 and 5.2. The real time stock quotes are $S_0 = (\$326.35, \$2758.07)$; the closed price are $S_0 = (\$324.16, \$2772.5)$; the low price are $S_0 = (\$321.11, \$2743.41)$ and the high price are $S_0 = (\$329.56, \$2831.17)$ of FB and GOOGL.

Comparing with the marketing price, our option price is closed to the marketing value. Both of FB and GOOGL, the marketing price between the option price based on low price and high price of stocks. Also, we find that our option price based the real time stock quotes is the closest to the marketing price. Next, let's turn to price the multivariate assets. Followed equation (5.5) and (5.6), we provide the pricing of

Table 5.3

Min Rainbow Option								
K	\$270.98	\$275.98	\$280.98	\$285.98	\$290.98	\$295.98	\$ 300.98	\$305.98
Price	\$54.98	\$49.98	\$44.99	\$40.01	\$35.04	\$30.11	\$ 25.26	\$20.57
K	\$310.98	\$315.98	\$320.98	\$325.98	\$330.98	\$335.98	\$340.98	\$345.98
Price	\$16.13	\$12.11	\$8.61	\$5.80	\$3.71	\$2.26	\$1.32	\$0.73
K	\$350.98	\$355.98	\$360.98	\$365.98	\$370.98	\$375.98	\$380.98	\$385.98
Price	\$0.39	\$0.19	\$ 0.09	\$0.047	\$0.024	\$ 0.01	\$0.005	\$0.003

Table 5.4

Max Rainbow Option							
K	\$2393.32	\$2433.32	\$2473.32	\$2513.32	\$2553.32	\$2593.32	\$ 2633.32
Price	\$364.37	\$324.45	\$284.61	\$244.93	\$205.48	\$166.57	\$129.07
K	\$2673.32	\$2713.32	\$2753.32	\$2793.32	\$2833.32	\$2873.32	\$ 2913.32
Price	\$94.19	\$63.65	\$39.06	\$21.57	\$10.85	\$5.07	\$2.25
K	\$2953.32	\$2993.32	\$3033.32	\$3073.32	\$ 3113.32	\$ 3153.32	\$ 3193.32
Price	\$0.98	\$0.44	\$0.21	\$0.09	\$0.05	\$0.031	\$0.015

rainbow option. We consider the real time stock quotes to be S_0, and give 24 different strike prices, then the results are in Table 5.3 and 5.4.

CHAPTER 6: GENERAL MULTIVARIATES TS DISTRIBUTIONS

The class of multivariate TS distributions introduced in Chapter 2 has a simple, yet flexible structure. It is a generalization to the multivariate setting of the so-called class of classical TS distributions, see [16]. However, there are many other classes of TS distributions available in the literature see, e.g. [16], [19], [20], [21], [22], [46], [47], and the references therein. Our theoretical results hold for many of these other classes and much of our methodology can be modified to work with them as well. For this reason, we extend our main results to a more general context. We begin with a definition.

A distribution μ on $\mathbb{R}^d$ is said to be a generalized tempered stable (GTS) distribution if its characteristics function can be written in the form

$$\hat{\mu}(z) = \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\int_0^\infty l_\ell(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\mathrm{d}x\sigma(\mathrm{d}s)\right],\ z\in\mathbb{R}^d, \tag{6.1}$$

where

$$l_\ell(s,x,z) = e^{i\langle s,z\rangle x} - 1 - i\,\langle s,z\rangle\, x h_\ell(x),$$

$\alpha \in (-\infty, 2)$, $\gamma \in \mathbb{R}^d$, σ is a finite Borel measure on $\mathbb{S}^{d-1}$, and $q : \mathbb{S}^{d-1}\times(0,\infty) \mapsto [0,\infty)$ is a Borel function satisfying conditions that are described below. The parameter $\ell \in \{0,1,2\}$ determines which parametrization we are using. Specifically, it determines the choice of function $h_\ell : \mathbb{R}^d \mapsto \mathbb{R}$, where $h_0 \equiv 0$, $h_1 \equiv 1$, and h_2 is given by $h_2(x) = 1_{[|x|\le 1]}$. Depending on the properties of q we may be able to use one or more of these parametrizations. We denote the distribution with characteristic function given in (6.1) by $\mu = \mathrm{GTS}_\alpha(\sigma, q, \gamma)_\ell$. This is an infinitely divisible distribution

with Lévy measure

$$L(B) = \int_{\mathbb{S}^{d-1}} \int_0^\infty 1_B(sx) x^{-1-\alpha} q(s,x) \mathrm{d}x \sigma(\mathrm{d}s), \quad B \in \mathcal{B}(\mathbb{R}^d).$$

We refer to q as the tempering function. Many parametric forms for q have appeared in the literature, see e.g. [22], [48] or [49]. We assume that q satisfies the property:

Q1. *There exist an $A \in \mathfrak{B}((0,\infty))$ with Lebesgue measure 0 and a $B \in \mathfrak{B}(\mathbb{S}^{d-1})$ with $\sigma(B) = 0$ such that for all $x \in A^c$ the function $q(\cdot, x)$ is continuous for all $s \in B^c$.*

We further assume that there exists a Borel function $q_U : (0,\infty) \mapsto [0,\infty)$ satisfying:

U1. *$q(s,x) \leq q_U(x)$ for every s and Lebesgue a.e. x;*

U2. $\int_0^1 x^{1-\alpha} q_U(x) \mathrm{d}x < \infty$;

U3. $\int_1^\infty x^{-1-\alpha} q_U(x) \mathrm{d}x < \infty$.

If q_U satisfies U1-U3, then we can take $\ell = 2$. Further, if, instead of U2, q_U satisfies the stronger condition

U2'. $\int_0^1 x^{-\alpha} q_U(x) \mathrm{d}x < \infty$,

then we can take $\ell = 0$. Similarly, if, instead of U3, q_U satisfies the stronger condition

U3'. $\int_1^\infty x^{-\alpha} q_U(x) \mathrm{d}x < \infty$,

then we can take $\ell = 1$. We call q_U the upper bounding function. When $\alpha \in (0,2)$, if q is a bounded function, i.e. if there exists a $K > 0$ such that $q(s,x) \leq K$, then we can take $q_U(x) \equiv K$. We now give our first main result for GTS distributions.

Theorem 5. *Fix $\alpha \in (-\infty, 2)$, let $\mu = \mathrm{GTS}_\alpha(\sigma, q, \gamma)_\ell$, assume that Q1 holds and that an upper bounding function q_U satisfying appropriate conditions to use parametrization ℓ exists. Then, there exists a sequence of Borel measures σ_n on $\mathbb{S}^{d-1}$ with finite support, such that, if $\mu_n = \mathrm{GTS}_\alpha(\sigma_n, q, \gamma)_\ell$, then $\mu_n \xrightarrow{w} \mu$ as $n \to \infty$.*

For certain special cases, related results are given in [22] and [50]. Next, we give a local version of this result. In this case, we need the distribution to have a bounded density. When $\alpha \leq 0$, this does not hold in general even when the distribution

is full, see [22]. For this reason, we focus on the case where $\alpha \in (0,2)$. Further, we assume that q has a lower bounding function q_L. Specifically, we assume that $q_L : (0,\infty) \mapsto (0,\infty)$ is a Borel function satisfying:

L1. $q(s,x) \geq q_L(x)$ for σ-a.e. s and Lebesgue a.e. x;

L2. $\int_0^1 x^{1-\alpha} q_L(x)\mathrm{d}x > 0$;

L3. $\int_0^\infty (1-\cos(x)) x^{-1-\alpha} q_L(x)\mathrm{d}x > 0$; and

L4. q_L is bounded and monotonely decreasing.

Theorem 6. *Fix $\alpha \in (0,2)$, let $\mu = \mathrm{GTS}_\alpha(\sigma, q, \gamma)_\ell$, and assume that Q1 holds. Assume further that there exists an upper function q_U satisfying appropriate conditions to use parametrization ℓ and a lower bounding q_L satisfying L1-L4. If σ is full, then μ has pdf p and for any $\varepsilon > 0$, there exists a finite measure σ^* on $\mathbb{S}^{d-1}$, having a finite support, such that the distribution $\mu^* = \mathrm{GTS}_\alpha(\sigma^*, q, \gamma)_\ell$ has a density p^*, which satisfies*

$$\sup_{x\in\mathbb{R}^d} |p(x) - p^*(x)| \leq \varepsilon.$$

For stable distributions (i.e. when $q \equiv 1$) a version of this result is given in [27].

Remark 1. *The TS distributions introduced in Section 2 are an important class of GTS distributions. Specifically, for $\alpha \in (0,1)$, $\mathrm{TS}_\alpha(\sigma, b, \gamma) = \mathrm{GTS}_\alpha(\sigma, q, \gamma)_0$, where $q(s,x) = e^{-b(s)x}$ and $b(\cdot)$ is continuous except, perhaps, on a set of σ measure 0. In this case we can take $q_U(x) \equiv 1$ and if there exists an $M > 0$ with $b(s) \leq M$, then we can take $q_L(x) = e^{-Mx}$. Thus, Theorem 1 is a special case of Theorem 6.*

Further, we verify L1-L4 under $q(s,x) = e^{-b(s)x}$, which we really interesting in this dissertation, with $b(s) \geq 0$. Let's assume there $\exists M$, such that $b(s) \leq M$, then $q(s,x) = e^{-b(s)x} \geq e^{-Mx} = q_L(x)$, which satisfies L1. For L2, we consider $\int_0^1 x^{1-\alpha} q_L(x)\mathrm{d}x = \int_0^1 x^{1-\alpha} e^{-Mx}\mathrm{d}x$. By changing variable $u = Mx$ and following

gamma function

$$\begin{aligned}\int_0^1 x^{1-\alpha}e^{-Mx}\mathrm{d}x &= M^{\alpha-2}\int_0^1 u^{1-\alpha}e^{-u}\mathrm{d}u \\ &\geq M^{\alpha-2}\int_0^1 u^{1-\alpha}(1-u)\mathrm{d}u \\ &= M^{\alpha-2}\frac{1}{(2-\alpha)(3-\alpha)} > 0,\end{aligned}$$

where the second linefollows 4.2.29 in [51]. So L2 is satisfied. Next, for L3

$$\begin{aligned}\int_0^\infty (1-\cos(x))x^{-1-\alpha}q_L(x)\mathrm{d}x &= \int_0^\infty (1-\cos(x))x^{-1-\alpha}e^{-Mx}dx \\ &= -\int_0^\infty (\cos(x)-1)x^{-1-\alpha}e^{-Mx}\mathrm{d}x \\ &\geq \frac{11}{24}\int_0^1 x^{1-\alpha}e^{-Mx}\mathrm{d}x > 0,\end{aligned}$$

where the third line follows by Lemma 4.13 [22] and the fourth line follows by L2. For L4, we consider $q_L'(x)$*, where* $q_L'(x) = (e^{-Mx})' = -Me^{-Mx}$*. Since* $x \in (0,\infty)$*,* $0 \leq b(s) \leq M$*, we have* $q_L'(x) < 0$*. Hence, L4 is satisfied. Hence, function of* $q(s,x) = e^{-b(s)x}$ *satisfies all 4 conditions.*

We use the same function of $q(s,x)$ *to verify U1, U2' and U3 under* $\ell = 2$*, such that* $q(s,x) = e^{-b(s)x}$*. Here, we take* $q_U(x) \equiv 1$*, then* $e^{-b(s)x} \leq 1 = q_U(x)$*, which satisfies U1. For U2',* $\int_0^1 x^{-\alpha}q_U(x)\mathrm{d}x = \int_0^1 x^{-\alpha}\mathrm{d}x = \frac{1}{1-\alpha} < \infty$*, so U2' is satisfied. For U3,* $\int_1^\infty x^{-1-\alpha}q_U(x)\mathrm{d}x = \int_1^\infty x^{-1-\alpha}\mathrm{d}x = \frac{1}{\alpha} < \infty$*. Then, function of* $q(s,x) = e^{-b(s)x}$ *satisfies U1, U2' and U3. The verification of* $\ell = 1$ *and* $\ell = 0$ *are the similar.*

CHAPTER 7: PROOFS of MAIN THEOREMS

7.1 Proof of Theorem 5

Before giving the proof, we recall a definition. If σ and σ_n are finite measures on $\mathbb{S}^{d-1}$, then we write $\sigma_n \xrightarrow{w} \sigma$ if for any continuous and bounded function $f : \mathbb{S}^{d-1} \mapsto \mathbb{R}$,

$$\lim_{n\to\infty} \int_{\mathbb{S}^{d-1}} f(x)\sigma_n(\mathrm{d}x) = \int_{\mathbb{S}^{d-1}} f(x)\sigma(\mathrm{d}x). \tag{7.1}$$

Lemma 2. *Let $f : \mathbb{S}^{d-1} \mapsto \mathbb{R}$ be a bounded Borel function and let D_f be its set of discontinuities. If $\sigma_n \xrightarrow{w} \sigma$ and $\sigma(D_f) = 0$, then (7.1) holds.*

Proof. Let $c_n = \sigma_n(\mathbb{S}^{d-1}) = \int_{\mathbb{S}^{d-1}} \sigma_n(\mathrm{d}x)$ and $c = \sigma(\mathbb{S}^{d-1}) = \int_{\mathbb{S}^{d-1}} \sigma(\mathrm{d}x)$. Note that (7.1) implies that $c_n \to c$. If $c = 0$ then the result follows easily from the fact that f is bounded. Henceforth assume $c > 0$. Let $\sigma'_n = \sigma_n/c_n$, let $\sigma' = \sigma/c$, and note that these are probability measures. The fact that $c_n \to c$ implies that $\sigma'_n \xrightarrow{w} \sigma'$. Thus, from a version of the Portmanteau Theorem, see e.g. Theorem 13.16 in [52], it follows that $c_n^{-1} \int_{\mathbb{S}^{d-1}} f(x)\sigma_n(\mathrm{d}x) \to c^{-1} \int_{\mathbb{S}^{d-1}} f(x)\sigma(\mathrm{d}x)$. The fact that $c_n \to c$ completes the proof. □

Lemma 3. *When $\ell = 2$, then $h_2(x) = 1_{[x\le 1]}$. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that if $s, s' \in \mathbb{S}^{d-1}$ with $|s - s'| \le \delta$, then for each x and fixed $z \in \mathbb{R}^d$,*

$$|l_2(s, x, z) - l_2(s', x, z)| \le \varepsilon.$$

Proof. Let's begin with

$$|l_2(s, x, z) - l_2(s', x, z)| \;=\; |e^{i\langle s,z\rangle x} - 1 - i\,\langle s, z\rangle\, x 1_{x\le 1}$$

$$
\begin{aligned}
& -(e^{i\langle s',z\rangle x} - 1 - i\langle s', z\rangle x 1_{x\leq 1})| \\
= \; & |e^{i\langle s,z\rangle x} - i\langle s, z\rangle x 1_{[x\leq 1]} - e^{i\langle s',z\rangle x} + i\langle s', z\rangle x 1_{[x\leq 1]}| \\
= \; & |e^{i\langle s,z\rangle x} - e^{i\langle s',z\rangle x} + i\langle s' - s, z\rangle x 1_{[x\leq 1]}| \\
= \; & |1 - e^{i\langle s'-s,z\rangle x} + e^{-i\langle s,z\rangle x} i\langle s' - s, z\rangle x 1_{[x\leq 1]}| \qquad (7.2)
\end{aligned}
$$

If under the case of $0 \leq x \leq 1$,

$$
\begin{aligned}
|l_2(s,x,z) - l_2(s',x,z)| &= |1 - e^{i\langle s'-s,z\rangle x} + e^{-i\langle s,z\rangle x} i\langle s' - s, z\rangle x| \\
&= |1 - e^{i\langle s'-s,z\rangle x} + i\langle s' - s, z\rangle x \\
&\quad + i\langle s' - s, z\rangle x(e^{-i\langle s,z\rangle x} - 1)| \\
&\leq |1 - e^{i\langle s'-s,z\rangle x} + i\langle s' - s, z\rangle x| \\
&\quad + |\langle s' - s, z\rangle x||e^{-i\langle s,z\rangle x} - 1| \\
&\leq \frac{1}{2}|x|^2|\langle s' - s, z\rangle| + |\langle s' - s, z\rangle||x|^2|\langle s, z\rangle| \qquad (7.3) \\
&= |x|^2|\langle s' - s, z\rangle|\left(\frac{1}{2}|\langle s' - s, z\rangle| + |\langle s, z\rangle|\right) \\
&\leq |x|^2||s' - s||z|\left(\frac{1}{2}|\langle s' - s, z\rangle| + |\langle s, z\rangle|\right), \qquad (7.4)
\end{aligned}
$$

where equation (7.3) follows 26.4_0 and 26.4_1 in [53].

Since $|s - s'| \leq \delta$, by equation (7.4), we have

$$
|l_2(s,x,z) - l_2(s',x,z)| \leq |x|^2|\delta|z|\left(\frac{1}{2}|\langle s' - s, z\rangle| + |\langle s, z\rangle|\right) = \varepsilon.
$$

If under the case of $x > 1$,

$$
\begin{aligned}
|l_2(s,x,z) - l_2(s',x,z)| &= |1 - e^{i\langle s'-s,z\rangle x}| \\
&\leq \min\{|\langle s' - s, z\rangle||x|, 2\} \\
&\leq \min\{|s' - s||z||x|, 2\}
\end{aligned}
$$

$$= \quad \min\{\delta|z||x|, 2\} = \varepsilon,$$

where the second line follows 26.4_0 in [53]. From here the result follows. □

Lemma 4. *Assume that q satisfies Q1 and that an upper bounding function q_U satisfying conditions **U1.-U3.**, then $\int_0^\infty l_\ell(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\mathrm{d}x$ is continuous in s for each x and bounded for each fixed $z \in \mathbb{R}^d$.*

Proof. It is enough to show the proof when $\ell = 2$. Other cases' proof is similar. So, let's consider $\int_0^\infty l_2(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\mathrm{d}x$ is continuous in s for each x. By equation (8.9) in [35], we have

$$|l_2(s,x,z)| \leq \frac{1}{2}|z|^2|x|^2 1_{[x\leq 1]} + 21_{[x>1]}.$$

Thus,

$$\left|l_2(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\right| \leq \left(\frac{1}{2}|z|^2|x|^2 1_{[x\leq 1]} + 21_{[x>1]}\right)\frac{q(s,x)}{x^{1+\alpha}}.$$

Since $q(s,x) \leq q_U(x)$,

$$\left|l_2(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\right| \leq \left(\frac{1}{2}|z|^2|x|^2 1_{[x\leq 1]} + 21_{[x>1]}\right)\frac{q_U(x)}{x^{1+\alpha}} = g(x,z), \tag{7.5}$$

where $g(x,z)$ is bounded for each fixed $z \in \mathbb{R}^d$ and integrable.

From the Lemma 3, we have $l_2(s,x,z)$ is continuous in s, and under Q1 $q(s,x)$ is also continuous in s, then we get $l_2(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}$ is continuous in s. Thus, following Theorem 16.8 in [53], $\int_0^\infty l_2(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\mathrm{d}x$ is continuous in s for each x.

Next, let's show $\int_0^\infty l_2(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\mathrm{d}x$ is bounded for each fixed $z \in \mathbb{R}^d$.

$$\left|\int_0^\infty l_2(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\mathrm{d}x\right| \quad \leq \quad \int_0^\infty \left|l_2(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\right|\mathrm{d}x$$

$$
\begin{aligned}
&\leq \int_0^\infty \left(\frac{1}{2}|z|^2|x|^2 1_{x\leq 1} + 21_{[x>1]}\right)\frac{q_U(x)}{x^{1+\alpha}}\mathrm{d}x \\
&= \int_0^1 \frac{1}{2}|z|^2|x|^2\frac{q_U(x)}{x^{1+\alpha}}\mathrm{d}x + 2\int_1^\infty \frac{q_U(x)}{x^{1+\alpha}}\mathrm{d}x \\
&= \frac{1}{2}|z|^2\int_0^1 x^{1-\alpha}q_U(x)\mathrm{d}x + 2\int_1^\infty x^{-1-\alpha}q_U(x)\mathrm{d}x < \infty,
\end{aligned}
$$

where the last inequality based on the condition**U2.** and **U3.** of $q_U(x)$. From here the results follows. □

Lemma 5. *Fix $\alpha \in (-\infty, 2)$, let σ and σ_n be finite Borel measures on $\mathbb{S}^{d-1}$, let $\mu = \mathrm{GTS}_\alpha(\sigma, q, \gamma)_\ell$, and let $\mu_n = \mathrm{GTS}_\alpha(\sigma_n, q, \gamma)_\ell$. Assume that q satisfies Q1 and that an upper bounding function q_U satisfying appropriate conditions to use parametrization ℓ exists. If $\sigma_n \xrightarrow{w} \sigma$, then $\mu_n \xrightarrow{w} \mu$.*

Proof. It suffices to show that the characteristic function of μ_n converges to that of μ. Toward this end let $\psi(s,z) = \int_0^\infty l_\ell(s,x,z)\frac{q(s,x)}{x^{1+\alpha}}\mathrm{d}x$. From (6.1), the fact that $\sigma_n \xrightarrow{w} \sigma$, and Lemma 2, it suffices to show that for each fixed $z \in \mathbb{R}^d$, the function $\psi(\cdot, z)$ is bounded and that the set of its discontinuities has σ measure 0.

We begin with boundedness. By (26.4) in [53] and (8.9) in [35] we have

$$
|l_\ell(s,x,z)| \leq 2\begin{cases} |z|x1_{[x\leq 1]} + 1_{[x>1]} & \text{if } \ell = 0 \\ |z|^2x^2 1_{[x\leq 1]} + |z|x1_{[x>1]} & \text{if } \ell = 1 \\ |z|^2x^2 1_{[x\leq 1]} + 1_{[x>1]} & \text{if } \ell = 2 \end{cases}.
$$

Letting $g_\ell(x,z)$ be this upper bound on $|l(s,x,z)|$, it follows that for every s, every z, and Lebesgue a.e. $x > 0$

$$
\left| l(s,x,z)\frac{q(s,x)}{x^{1+\alpha}} \right| \leq g_\ell(x,z)\frac{q_U(x)}{x^{1+\alpha}}. \tag{7.6}
$$

This is integrable by the assumptions on q_U needed to use representation ℓ. Thus, for each s and z we have $|\psi(s,z)| \leq \int_0^\infty g_\ell(x,z)\frac{q_U(x)}{x^{1+\alpha}}\mathrm{d}x < \infty$. We showed the detailed

proof in Lemma 4.

Next, note that for fixed z and x, $l_\ell(\cdot, x, z)$ is continuous. It follows that for all z and all $x \in A^c$ the set of discontinuities of the function mapping s to $l_\ell(s, x, z)\frac{q(s,x)}{x^{1+\alpha}}$ is contained in set B. Now applying the dominated convergence theorem for continuity, see e.g. Theorem 16.8 in [53] shows that for fixed z the set of discontinuities of $\psi(\cdot, z)$ is contained in B, which has σ measure 0. □

Lemma 6. *Fix $\beta > 0$. For every $\varepsilon > 0$ there exists a $\delta > 0$ such that, if $s, s' \in \mathbb{S}^{d-1}$ with $|s - s'| \leq \delta$, then for every $\xi \in \mathbb{S}^{d-1}$,*

$$\left| |\langle s, \xi \rangle|^\beta - |\langle s', \xi \rangle|^\beta \right| \leq \varepsilon.$$

Proof. We begin with the case when $\beta \in (0, 1]$. Let $\delta = \varepsilon^{\frac{1}{\beta}}$ and note that

$$\begin{aligned} \left| |\langle s, \xi \rangle|^\beta - |\langle s', \xi \rangle|^\beta \right| &\leq \left| |\langle s, \xi \rangle| - |\langle s', \xi \rangle| \right|^\beta \\ &= |\langle s - s', \xi \rangle|^\beta \\ &\leq |s - s'|^\beta |\xi|^\beta \leq \delta^\beta = \varepsilon, \end{aligned}$$

where the first inequality follows from the fact that, for such β, $|x^\beta - y^\beta| \leq |x - y|^\beta$ for $x, y \in [0, 1]$, see e.g. the proof of Lemma 2 in [**?**]. We now turn to the case when $\beta \in (1, \infty)$. Let $\delta = \dfrac{\varepsilon}{\beta}$ and note that

$$\begin{aligned} \left| |\langle s, \xi \rangle|^\beta - |\langle s', \xi \rangle|^\beta \right| &\leq \beta |\langle s, \xi \rangle - \langle s', \xi \rangle| \\ &= \beta |\langle s - s', \xi \rangle| \\ &\leq \beta |s - s'||\xi| \leq \beta\delta = \varepsilon, \end{aligned}$$

where the first inequality follows from the fact that, for such β, $|x^\beta - y^\beta| \leq \beta |x - y|^\beta$ for $x, y \in [0, 1]$, which itself follows from a standard application of the mean value

theorem. □

Lemma 7. *For any finite Borel measure σ on $\mathbb{S}^{d-1}$, there exists a sequence $\{\sigma_n\}$ of finite Borel measure on $\mathbb{S}^{d-1}$, each having finite support, such that $\sigma_n \xrightarrow{w} \sigma$ and for any $\beta > 0$*

$$\lim_{n\to\infty} \inf_{\xi\in\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^\beta \sigma_n(\mathrm{d}s) = \inf_{\xi\in\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^\beta \sigma(\mathrm{d}s).$$

Further, σ_n is full for each n.

Note that a part of the result is that we can take each σ_n to be full even if σ is not full.

Proof. Fix $n \in \mathbb{N}$ and consider the open cover of $\mathbb{S}^{d-1}$ comprised of open balls of radius $1/n$ centered at each point of $\mathbb{S}^{d-1}$. Since $\mathbb{S}^{d-1}$ is compact, there exists a finite subcover $B_1, B_2, ..., B_{M_n}$, where M_n is the number of ball in the subcover. Let $s_1, s_2, \ldots, s_{M_n} \in \mathbb{S}^{d-1}$ be the centers of these balls. Without loss of generality, we assume that these vectors span $\mathbb{R}^d$, as otherwise we can add a finite number of balls to the subcover to insure thus. Now construct a disjoint cover $A_1, A_2, \ldots, A_{M_n}$, where $A_j = (B_j \setminus \bigcup_{i<j} A_i) \cap \mathbb{S}^{d-1}$ for $j = 1, .., M_n$. Next, define

$$\sigma_n = \sum_{j=1}^{M_n} \left(\sigma(A_j) \vee \frac{1}{nM_n}\right) \delta_{s_j},$$

where δ_{s_j} is the pointmass at s_j. Note that $s_j \in \mathbb{S}^{d-1}$ may not be in A_j. Note further that since $s_1, s_2, \ldots, s_{M_n}$ span $\mathbb{R}^d$ and $\left(\sigma(A_j) \vee \frac{1}{nM_n}\right) > 0$, this measure if full.

We now show that $\sigma_n \xrightarrow{w} \sigma$. Toward this end, let $f : \mathbb{S}^{d-1} \mapsto \mathbb{R}$ be a continuous function that is bounded by some $K > 0$. Since $\mathbb{S}^{d-1}$ is compact, f is uniformly continuous. Thus, for any $\varepsilon > 0$, there exists a $\delta > 0$ such that if $s, s' \in \mathbb{S}^{d-1}$ and $|s - s'| \leq \delta$, then

$$|f(s) - f(s')| \leq \frac{\varepsilon}{2\sigma(\mathbb{S}^{d-1})}.$$

Next, note that for any $n \geq \left(\frac{1}{\delta} \vee \frac{2K}{\varepsilon}\right)$

$$\begin{aligned}
&\left| \int_{\mathbb{S}^{d-1}} f(s)\sigma(\mathrm{d}s) - \int_{\mathbb{S}^{d-1}} f(s)\sigma_n(\mathrm{d}s) \right| \\
&\qquad \leq \left| \int_{\bigcup_{j=1}^{M_n} A_j} f(s)\sigma(\mathrm{d}s) - \sum_{j=1}^{M_n} f(s_j)\sigma(A_j) \right| + \sum_{j=1}^{M_n} \frac{|f(s_j)|}{nM_n} \\
&\qquad \leq \left| \sum_{j=1}^{M_n} \int_{A_j} (f(s) - f(s_j))\sigma(\mathrm{d}s) \right| + \varepsilon/2 \\
&\qquad \leq \sum_{j=1}^{M_n} \int_{A_j} |f(s) - f(s_j)|\sigma(\mathrm{d}s) + \varepsilon/2 \\
&\qquad \leq \sum_{j=1}^{M_n} \int_{A_j} \frac{\varepsilon}{2\sigma(\mathbb{S}^{d-1})}\sigma(\mathrm{d}s) + \varepsilon/2 = \varepsilon, \qquad (7.7)
\end{aligned}$$

where the last line follows from the fact that if $s \in A_j$, then $s, s_j \in B_j$ and hence $|s - s_j| < 1/n \leq \delta$.

We now turn to the last part. Fix $\varepsilon > 0$. By Lemma 6 there exists a $\delta > 0$, such that if $s, s' \in \mathbb{S}^{d-1}$ with $|s - s'| \leq \delta$ then for every $\xi \in \mathbb{S}^{d-1}$,

$$\left| |\langle s, \xi\rangle|^\beta - |\langle s', \xi\rangle|^\beta \right| \leq \frac{\varepsilon}{2\sigma(\mathbb{S}^{d-1})}.$$

Note that for any $s, \xi \in \mathbb{S}^{d-1}$, $|\langle s, \xi\rangle|^\beta \leq 1$. Thus, if $n \geq \left(\frac{1}{\delta} \vee \frac{2}{\varepsilon}\right)$, then by (7.7), for every $\xi \in \mathbb{S}^{d-1}$, we get

$$\left| \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^\beta \sigma_n(\mathrm{d}s) - \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^\beta \sigma(\mathrm{d}s) \right| \leq \varepsilon$$

and hence

$$\int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^\beta \sigma(\mathrm{d}s) - \varepsilon \leq \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^\beta \sigma_n(\mathrm{d}s) \leq \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^\beta \sigma(\mathrm{d}s) + \varepsilon.$$

Taking infimums over this inequality gives the result. □

Proof of Theorem 5. The result follows immediately from Lemmas 7 and 5. □

7.2 Proof of Theorem 6

Lemma 8. *Fix $\alpha \in (0,2)$ and assume that q has a lower bounding function q_L satisfying L1-L4. There exists a constant $C = C(\alpha, q_L) > 0$, such that for σ-a.e s,*

$$\int_0^\infty (\cos(\langle s,z\rangle x) - 1)x^{-1-\alpha} q(s,x)\mathrm{d}x \leq -C(|\langle s,z\rangle|^2 \wedge |\langle s,z\rangle|^\alpha).$$

Proof. When $|\langle s,z\rangle| \leq 1$, arguments similar to those in the proof of Lemma 4.13 in [22] imply that for σ-a.e s

$$\begin{aligned}\int_0^\infty (\cos(\langle s,z\rangle x) - 1)x^{-1-\alpha} q(s,x)\mathrm{d}x &\leq -\frac{11}{24}|\langle s,z\rangle|^2 \int_0^1 x^{1-\alpha} q(s,x)\mathrm{d}x \\ &\leq -\frac{11}{24}|\langle s,z\rangle|^2 \int_0^1 x^{1-\alpha} q_L(x)\mathrm{d}x.\end{aligned}$$

When $|\langle s,z\rangle| > 1$, then for σ-a.e s

$$\begin{aligned}&\int_0^\infty (\cos(\langle s,z\rangle x) - 1)x^{-1-\alpha} q(s,x)\mathrm{d}x \\ &\qquad \leq -\int_0^\infty (1-\cos(\langle s,z\rangle x))x^{-1-\alpha} q_L(x)\mathrm{d}x \\ &\qquad \leq -|\langle s,z\rangle|^\alpha \int_0^\infty (1-\cos(x))x^{-1-\alpha} q_L\left(\frac{x}{|\langle s,z\rangle|}\right)\mathrm{d}x \\ &\qquad \leq -|\langle s,z\rangle|^\alpha \int_0^\infty (1-\cos(x))x^{-1-\alpha} q_L(x)\mathrm{d}x,\end{aligned}$$

where the second line follows by change of variables and the third from the fact that $q_L(x)$ is monotonously decreasing. Now, taking

$$C = C(\alpha, q_L) = \min\left\{\frac{11}{24}\int_0^1 x^{1-\alpha} q_L(x)\mathrm{d}x, \int_0^\infty (1-\cos(x))x^{-1-\alpha} q_L(x)\mathrm{d}x\right\}$$

gives the result. □

Lemma 9. *Let $\alpha \in (0,2)$, $\mu = \mathrm{GTS}_\alpha(\sigma, q, \gamma)_\ell$, $u_2 = \inf_{\xi \in \mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |\langle s, \xi \rangle|^2 \sigma(\mathrm{d}s)$, and assume that a lower bounding function q_L satisfying L1-L4 exists. If σ is full, then there exists a constant $C = C(\alpha, q_L) > 0$ such that*

$$|\hat{\mu}(z)| \leq \exp\left[-C(|z|^2 \wedge |z|^\alpha)u_2\right] \tag{7.8}$$

and

$$\int_{\mathbb{R}^d} |\hat{\mu}(z)| \mathrm{d}z \leq \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \int_0^\infty \exp\left[-C(r^2 \wedge r^\alpha)u_2\right] r^{d-1} \mathrm{d}r < \infty.$$

Proof. From equation (6.1) and Lemma 8 it follows that there exists a $C > 0$ with

$$\begin{aligned}
|\hat{\mu}(z)| &= \exp\left[\Re\left(i\langle \gamma, z\rangle + \int_{\mathbb{S}^{d-1}} \int_0^\infty l_\ell(s, x, z) \frac{q(s,x)}{x^{1+\alpha}} \mathrm{d}x \sigma(\mathrm{d}s)\right)\right] \\
&= \exp\left[\int_{\mathbb{S}^{d-1}} \int_0^\infty (\cos(\langle s, z\rangle x) - 1) \frac{q(s,x)}{x^{1+\alpha}} \mathrm{d}x \sigma(\mathrm{d}s)\right] \\
&\leq \exp\left[-C \int_{\mathbb{S}^{d-1}} (|\langle s, z\rangle|^2 \wedge |\langle s, z\rangle|^\alpha) \sigma(\mathrm{d}s)\right].
\end{aligned}$$

Next, for $z \neq 0$, let $\xi_z = \dfrac{z}{|z|}$, and note that, for such z

$$\begin{aligned}
|\hat{\mu}(z)| &\leq \exp\left[-C \int_{\mathbb{S}^{d-1}} \left((|z|^2 |\langle s, \xi_z\rangle|^2) \wedge (|z|^\alpha |\langle s, \xi_z\rangle|^\alpha)\right) \sigma(\mathrm{d}s)\right] \\
&\leq \exp\left[-C \int_{\mathbb{S}^{d-1}} \left(\left((|z|^2 \wedge |z|^\alpha)|\langle s, \xi_z\rangle|^2\right) \wedge \left((|z|^\alpha \wedge |z|^2)|\langle s, \xi_z\rangle|^\alpha\right)\right) \sigma(\mathrm{d}s)\right] \\
&\leq \exp\left[-C(|z|^2 \wedge |z|^\alpha) \inf_{\xi \in \mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} (|\langle s, \xi\rangle|^2 \wedge |\langle s, \xi\rangle|^\alpha) \sigma(\mathrm{d}s)\right] \\
&= \exp\left[-C(|z|^2 \wedge |z|^\alpha) \inf_{\xi \in \mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^2 \sigma(ds)\right] \\
&= \exp\left[-C(|z|^2 \wedge |z|^\alpha) u_2\right],
\end{aligned}$$

where the last two linse follows from the fact that $s, \xi \in \mathbb{S}^{d-1}$ and $\alpha \in [0, 2)$, then $|\langle s, \xi\rangle| \leq 1$, and $|\langle s, \xi\rangle|^\alpha > |\langle s, \xi\rangle|^2$. This gives (7.8) for $z \neq 0$, and since $\hat{\mu}(0) = 1$, the result holds for $z = 0$ as well. Now, converting to polar coordinates (see e.g.

Theorem 5.1.8 in [54]) gives

$$\begin{aligned}\int_{\mathbb{R}^d} |\hat{\mu}(z)|\mathrm{d}z &\leq \int_{\mathbb{R}^d} \exp\left[-C(|z|^2 \wedge |z|^\alpha)u_2\right] \mathrm{d}z \\ &= \int_{\mathbb{S}^{d-1}} \int_0^\infty \exp\left[-C(r^2 \wedge r^\alpha)u_2\right] r^{d-1}\mathrm{d}r\lambda_{\mathbb{S}^{d-1}}(\mathrm{d}\omega) \\ &= \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \int_0^\infty \exp\left[-C(r^2 \wedge r^\alpha)u_2\right] r^{d-1}\mathrm{d}r < \infty,\end{aligned}$$

where the finiteness follows by the fact that σ is full and thus that $u_2 > 0$ by Lemma 1. Here $\lambda_{\mathbb{S}^{d-1}}$ is the surface measure on $\mathbb{S}^{d-1}$ and $\lambda_{\mathbb{S}^{d-1}}(\mathbb{S}^{d-1}) = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}$, see [54]. □

Lemma 10. *Let σ_n be a discretization of σ as defined in Lemma 7 and μ follows $\mathrm{GTS}_\alpha(\sigma, q, \gamma)_\ell$ with $\alpha \in (0, 2)$. If σ is full, then for large enough n*

$$\int_{\mathbb{R}^d} |\hat{\mu}_n(z)|\mathrm{d}z < \infty. \tag{7.9}$$

Proof. Applying Lemma 7 with $\beta = 2$ gives

$$\lim_{n\to\infty} \inf_{\xi\in\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^2 \sigma_n^*(\mathrm{d}s) \longrightarrow \inf_{\xi\in\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^2 \sigma(\mathrm{d}s) > 0.$$

Thus, for large enough n, σ_n satisfies Lemma 1. From here the result follows by Lemma 9. □

Proof of Theorem 6. Let $\hat{\mu}(z)$ be the characteristic function of $\mathrm{GTS}_\alpha(\sigma, q, \gamma)$ and let $\hat{\mu}_n(z)$ be the characteristic function of $\mathrm{GTS}_\alpha(\sigma_n, q, \gamma)$. By Lemma 10, there is an $N_0 \geq 1$ such that if $n \geq N_0$, then (7.9) holds. For such n, by the inversion formula, see e.g. Proposition 2.5(xii) in [35], we get

$$\begin{aligned}|p(x) - p_n(x)| &= (2\pi)^{-d} \left| \int_{\mathbb{R}^d} e^{-i\langle x,z\rangle} (\hat{\mu}(z) - \hat{\mu}_n(z))\mathrm{d}z \right| \\ &\leq (2\pi)^{-d} \int_{\mathbb{R}^d} |e^{-i\langle x,z\rangle} (\hat{\mu}(z) - \hat{\mu}_n(z))|\mathrm{d}z\end{aligned}$$

$$= (2\pi)^{-d} \int_{\mathbb{R}^d} |\hat{\mu}(z) - \hat{\mu}_n(z)| \mathrm{d}z. \tag{7.10}$$

Lemma 7 implies that there exists an $N_1 \geq N_0$ such that $u_n \geq u/2$. Taking $n \geq N_1$ and applying (7.8) gives

$$\begin{aligned} |\hat{\mu}(z) - \hat{\mu}_n(z)| &\leq \exp[-C(|z|^2 \wedge |z|^\alpha)u] + \exp[-C(|z|^2 \wedge |z|^\alpha)u_n] \\ &\leq \exp[-C(|z|^2 \wedge |z|^\alpha)u] + \exp[-.5C(|z|^2 \wedge |z|^\alpha)u_2] \\ &\leq 2\exp[-.5C(|z|^2 \wedge |z|^\alpha)u_2]. \end{aligned}$$

Thus,

$$\lim_{L\to\infty} \sup_{n\geq N_1} \int_{|z|>L} |\hat{\mu}(z) - \hat{\mu}_n(z)| \mathrm{d}z \leq \lim_{L\to\infty} 2 \int_{|z|>L} \exp[-.5C(|z|^2 \wedge |z|^\alpha)u] \mathrm{d}z = 0,$$

where the final equality follows by the dominated convergence theorem. It follows that there is an $L' > 0$ such that $\sup_{n\geq N_1} \int_{|z|>L'} |\hat{\mu}(z) - \hat{\mu}_n(z)| \mathrm{d}z < \frac{(2\pi)^d \varepsilon}{2}$.

Since $\hat{\mu}_n \xrightarrow{w} \hat{\mu}$, following Proposition 2.5(vi) in [35], we have $\hat{\mu}_n(z) \longrightarrow \hat{\mu}(z)$ uniformly on any compact set. This implies that, $\forall \varepsilon > 0, \forall k > 0, L = L' > 0, \exists N_2$ s.t $n \geq N_2$, then $\sup_{|z|\leq L'} |\hat{\mu}(z) - \hat{\mu}_n(z))| < \frac{\varepsilon}{k}$.

Hence, for $n \geq N_2$ and $L = L'$,

$$\begin{aligned} \int_{\Re^d} |\hat{\mu}(z) - \hat{\mu}_n(z)| \mathrm{d}z &= \int_{|z|\leq L'} |\hat{\mu}(z) - \hat{\mu}_n(z)| \mathrm{d}z + \int_{|z|>L'} |\hat{\mu}(z) - \hat{\mu}_n(z)| \mathrm{d}z \\ &\leq \frac{\varepsilon}{k} \int_{|z|\leq L'} \mathrm{d}z + \int_{|z|>L'} |\hat{\mu}(z) - \hat{\mu}_n(z)| \mathrm{d}z \qquad (7.11) \\ &\leq \frac{\varepsilon}{k} \int_{|z|\leq L'} \mathrm{d}z + \frac{(2\pi)^d \varepsilon}{2} \\ &= \frac{\varepsilon}{k} * \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)} * (L')^d + \frac{(2\pi)^d \varepsilon}{2} \\ &= \frac{(2\pi)^d \varepsilon}{2} + \frac{(2\pi)^d \varepsilon}{2} = (2\pi)^d \varepsilon \end{aligned}$$

For the fouth line, $\frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$ is the volume of $\{z \in R^d : |z| \leq L'\}$, and let $k = \frac{2^{1-d}\pi^{-\frac{d}{2}}(L')^d}{\Gamma(\frac{d}{2}+1)}$.
Thus, going back to the equation (7.10), we have

$$\begin{aligned} |p(x) - p_n(x)| &\leq (2\pi)^{-d} \int_{\mathbb{R}^d} |\hat{\mu}(z) - \hat{\mu}_n(z)| \mathrm{d}z \\ &= (2\pi)^{-d} * (2\pi)^d \varepsilon = \varepsilon. \end{aligned}$$

From here the results follows from equation (7.10) by taking $p^* = p_n$ for any $n \geq N_2$. □

REFERENCES

[1] L. Cao and M. Grabchak, "Smoothly truncated levy walks: Toward a realistic mobility model," in *2014 IEEE 33rd International Performance Computing and Communications Conference (IPCCC)*, pp. 1–8, IEEE, 2014.

[2] S. Thananjeyan, C. A. Chan, E. Wong, and A. Nirmalathas, "Deployment and resource distribution of mobile edge hosts based on correlated user mobility," *IEEE Access*, vol. 7, pp. 148–159, 2018.

[3] P. S. Griffin, R. A. Maller, and D. Roberts, "Finite time ruin probabilities for tempered stable insurance risk processes," *Insurance: Mathematics and Economics*, vol. 53, no. 2, pp. 478–489, 2013.

[4] M. L. Zuparic and A. C. Kalloniatis, "Analytic solution to space-fractional fokker–planck equations for tempered-stable lévy distributions with spatially linear, time-dependent drift," *Journal of Physics A: Mathematical and Theoretical*, vol. 51, no. 3, p. 035101, 2017.

[5] X. Yu, Y. Zhang, and H. Sun, "Modeling covid-19 spreading dynamics and unemployment rate evolution in rural and urban counties of alabama and new york using fractional derivative models," *Results in Physics*, vol. 26, p. 104360, 2021.

[6] K. J. Palmer, M. S. Ridout, and B. J. Morgan, "Modelling cell generation times by using the tempered stable distribution," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 57, no. 4, pp. 379–397, 2008.

[7] Y. Liu, P. M. Djurić, Y. S. Kim, S. T. Rachev, and J. Glimm, "Systemic risk modeling with lévy copulas," *Journal of Risk and Financial Management*, vol. 14, no. 6, p. 251, 2021.

[8] S. Asmussen and M. Bladt, "Gram–charlier methods, regime-switching and stochastic volatility in exponential lévy models," *Quantitative Finance*, pp. 1–15, 2021.

[9] J. Choi, Y. S. Kim, and I. Mitov, "Reward-risk momentum strategies using classical tempered stable distribution," *Journal of Banking & Finance*, vol. 58, pp. 194–213, 2015.

[10] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *The Journal of political economy*, vol. 81, no. 3, pp. 637–654, 1973.

[11] J. P. Nolan, "Fitting data and assessing goodness-of-fit with stable distributions," *Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics, Washington DC*, 1999.

[12] M. Grabchak and G. Samorodnitsky, "Do financial returns have finite or infinite variance? a paradox and an explanation," *Quantitative Finance*, vol. 10, no. 8, pp. 883–893, 2010.

[13] M. C. Tweedie, "An index which distinguishes between some important exponential families," *Statistics: Applications and new directions: Proc. Indian statistical institute golden Jubilee International conference*, vol. 579, pp. 579–604, 1984.

[14] J. Poirot and P. Tankov, "Monte carlo option pricing for tempered stable (cgmy) processes," *Asia-Pacific Financial Markets*, vol. 13, no. 4, pp. 327–344, 2006.

[15] Y. S. Kim, S. T. Rachev, D. M. Chung, and M. L. Bianchi, "The modified tempered stable distribution, garch-models and option pricing," *Probability and Mathematical Statistics*, vol. 29, no. 1, pp. 91–117, 2009.

[16] S. T. Rachev, Y. S. Kim, M. L. Bianchi, and F. J. Fabozzi, *Financial models with Lévy processes and volatility clustering*, vol. 187. John Wiley & Sons, 2011.

[17] J. Li, C. Favero, and F. Ortu, "A spectral estimation of tempered stable stochastic volatility models and option pricing," *Computational Statistics & Data Analysis*, vol. 56, no. 11, pp. 3645–3658, 2012.

[18] U. Küchler and S. Tappe, "Exponential stock models driven by tempered stable processes," *Journal of Econometrics*, vol. 181, no. 1, pp. 53–63, 2014.

[19] J. Rosiński, "Tempering stable processes," *Stochastic processes and their applications*, vol. 117, no. 6, pp. 677–707, 2007.

[20] J. Rosinski and J. L. Sinclair, "Generalized tempered stable processes," *Stability in Probability*, vol. 90, pp. 153–170, 2010.

[21] M. Grabchak, "On a new class of tempered stable distributions: moments and regular variation," *Journal of Applied Probability*, vol. 49, no. 4, pp. 1015–1035, 2012.

[22] M. Grabchak, *Tempered Stable Distributions Stochastic Models for Multiscale Processes*. Springer, 2016.

[23] Y. S. Kim, "The fractional multivariate normal tempered stable process," *Applied Mathematics Letters*, vol. 25, no. 12, pp. 2396–2401, 2012.

[24] Y. S. Kim, "Multivariate tempered stable model with long-range dependence and time-varying volatility," *Frontiers in Applied Mathematics and Statistics*, vol. 1, p. 1, 2015.

[25] Y. S. Kim, H. Kim, J. Choi, and F. J. Fabozzi, "Multi-asset option pricing using normal tempered stable processes with stochastic correlation," *Available at SSRN 3927399*, 2021.

[26] M. Scherer, S. T. Rachev, Y. S. Kim, and F. J. Fabozzi, "Approximation of skewed and leptokurtic return distributions," *Applied Financial Economics*, vol. 22, no. 16, pp. 1305–1316, 2012.

[27] T. Byczkowski, J. P. Nolan, and B. Rajput, "Approximation of multidimensional stable densities," *Journal of Multivariate Analysis*, vol. 46, no. 1, pp. 13–31, 1993.

[28] R. Modarres and J. P. Nolan, "A method for simulating stable random vectors," *Computational Statistics*, vol. 9, no. 1, pp. 11–19, 1994.

[29] J. P. Nolan, A. K. Panorska, and J. H. McCulloch, "Estimation of stable spectral measures," *Mathematical and Computer Modelling*, vol. 34, no. 9-11, pp. 1113–1122, 2001.

[30] M. Mohammadi, A. Mohammadpour, and H. Ogata, "On estimating the tail index and the spectral measure of multivariate α-stable distributions," *Metrika*, vol. 78, no. 5, pp. 549–561, 2015.

[31] R. Cont and P. Tankov, *Financial modelling with jump processes*. Chapman and Hall/CRC, 2004.

[32] F. Hubalek and C. Sgarra, "Esscher transforms and the minimal entropy martingale measure for exponential lévy models," *Quantitative finance*, vol. 6, no. 02, pp. 125–145, 2006.

[33] Y. Xia and M. Grabchak, "Estimation and simulation for multivariate tempered stable distributions," *Journal of Statistical Computation and Simulation*, vol. 92, no. 3, pp. 451–475, 2022.

[34] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance: Stochastic Modeling*. Routledge, 2017.

[35] S. Ken-Iti, *Lévy processes and infinitely divisible distributions*. Cambridge university press, 1999.

[36] D. N. Joanes and C. A. Gill, "Comparing measures of sample skewness and kurtosis," *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 47, no. 1, pp. 183–189, 1998.

[37] L. Devroye, "Random variate generation for exponentially and polynomially tilted stable distributions," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 19, no. 4, pp. 1–20, 2009.

[38] R. Kawai and H. Masuda, "On simulation of tempered stable random variates," *Journal of Computational and Applied Mathematics*, vol. 235, no. 8, pp. 2873–2887, 2011.

[39] M. Hofert, "Sampling exponentially tilted stable distributions," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 22, no. 1, pp. 1–11, 2011.

[40] K. Price, R. M. Storn, and J. A. Lampinen, *Differential evolution: a practical approach to global optimization*. Springer Science & Business Media, 2006.

[41] Q. Li, E. Maasoumi, and J. S. Racine, "A nonparametric test for equality of distributions with mixed categorical and continuous data," *Journal of Econometrics*, vol. 148, no. 2, pp. 186–200, 2009.

[42] J. Philippe, *Value at risk: the new benchmark for managing financial risk.* NY: McGraw-Hill Professional, 2001.

[43] W. F. Sharpe, "A simplified model for portfolio analysis," *Management science*, vol. 9, no. 2, pp. 277–293, 1963.

[44] P. Glasserman, *Monte Carlo methods in financial engineering*, vol. 53. Springer, 2004.

[45] R. Eberhart and J. Kennedy, "Particle swarm optimization," in *Proceedings of the IEEE international conference on neural networks*, vol. 4, pp. 1942–1948, Citeseer, 1995.

[46] M. L. Bianchi, S. T. Rachev, Y. S. Kim, and F. J. Fabozzi, "Tempered infinitely divisible distributions and processes," *Theory of Probability & Its Applications*, vol. 55, no. 1, pp. 2–26, 2011.

[47] M. Grabchak, "An exact method for simulating rapidly decreasing tempered stable distributions in the finite variation case," *Statistics & Probability Letters*, vol. 170, p. 109015, 2021.

[48] G. Terdik and W. Woyczynski, "Rosinski measures for tempered stable and related ornstein-uhlenbeck processes," *Probability and Mathematical Statistics*, vol. 26, no. 2, p. 213, 2006.

[49] A. Dassios, J. W. Lim, and Y. Qu, "Exact simulation of a truncated lévy subordinator," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 30, no. 3, pp. 1–17, 2020.

[50] O. E. Barndorff-Nielsen, M. Maejima, and K.-i. Sato, "Some classes of multivariate infinitely divisible distributions admitting stochastic integral representations," *Bernoulli*, pp. 1–33, 2006.

[51] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series 55. Tenth Printing.* ERIC, 1972.

[52] A. Klenke, *Probability theory: a comprehensive course.* Springer Science & Business Media, 2014.

[53] P. Billingsley, *Measure and probability.* John Wiley & Sons, 1995.

[54] D. W. Stroock *et al.*, *Essentials of integration theory for analysis.* Springer, 2011.

[55] G. B. Folland, *Real analysis: modern techniques and their applications*, vol. 40. John Wiley & Sons, 1999.

APPENDIX A: PROOFS of THEOREMS IN CHAPTER 2

A.1 Proof of Proprtion 2

Proof of Proprtion 2. Let's begin to show the expectation of X_j,

$$\begin{aligned} c_1 &= \mathrm{E}[X_j] = (-i)\frac{\partial C_\mu(z)}{\partial z_j}\Big|_{z=0} \\ &= (-i)\left[i\gamma_j + \int_{\mathbb{S}^{d-1}} \Gamma(-\alpha)\alpha(b(s) - i\sum_{i=1}^{d} s_j z_j)^{\alpha-1}(-is_j)\sigma(\mathrm{d}s)\right]\Big|_{z=0} \\ &= \gamma_j - \int_{\mathbb{S}^{d-1}} \Gamma(-\alpha)\alpha b^{\alpha-1}(s)s_j\sigma(\mathrm{d}s) \\ &= \gamma_j + \Gamma(1-\alpha)\int_{\mathbb{S}^{d-1}} b^{\alpha-1}(s)s_j\sigma(\mathrm{d}s) \end{aligned}$$

where the last line followed by $\Gamma(-\alpha) = -\dfrac{\Gamma(1-\alpha)}{\alpha}$.

Next, we derive the variance of X_j as follow

$$\begin{aligned} c_2 = \mathrm{Var}(X_j) &= (-i)^2\frac{\partial^2 C_\mu(z)}{\partial^2 z_j}\Big|_{z=0} \\ &= -\frac{\partial\left[i\gamma_j + \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha b^{\alpha-1}(s)(-is_j)\sigma(\mathrm{d}s)\right]}{\partial z_j}\Big|_{z=0} \\ &= -\left[\int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)(b(s) - i\sum_{j=1}^{d}s_j z_j)^{\alpha-2}(is_j)^2\sigma(\mathrm{d}s)\right]\Big|_{z=0} \\ &= \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)b^{\alpha-2}(s)s_j^2\sigma(\mathrm{d}s) \\ &= \Gamma(1-\alpha)(1-\alpha)\int_{\mathbb{S}^{d-1}} b^{\alpha-2}(s)s_j^2\sigma(\mathrm{d}s) \\ &= \Gamma(2-\alpha)\int_{\mathbb{S}^{d-1}} b^{\alpha-2}(s)s_j^2\sigma(\mathrm{d}s), \end{aligned}$$

where the laset line since $\Gamma(n) = (n-1)\Gamma(n-1)$.

Similiarly, we get the third central moment of X_j

$$c_3 = \mathrm{E}[(X_j - \mathrm{E}(X_j))^3] = (-i)^3\frac{\partial^3 C_\mu(z)}{\partial^3 z_j}\Big|_{z=0}$$

$$
\begin{aligned}
&= i\frac{\partial^2\left[i\gamma_j + \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha b^{\alpha-1}(s)(-is_j)\sigma(\mathrm{d}s)\right]}{\partial^2 z_j}\Big|_{z=0}\\
&= i\frac{\partial\left[\int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)(b(s)-i\sum_{j=1}^{d}s_jz_j)^{\alpha-2}(is_j)^2\sigma(\mathrm{d}s)\right]}{\partial z_j}\Big|_{z=0}\\
&= -\int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)(\alpha-2)b^{\alpha-3}(s)s_j^3\sigma(\mathrm{d}s)\\
&= \int_{\mathbb{S}^{d-1}}\Gamma(1-\alpha)(\alpha-1)(\alpha-2)b^{\alpha-3}(s)s_j^3\sigma(\mathrm{d}s)\\
&= \Gamma(3-\alpha)\int_{\mathbb{S}^{d-1}}b^{\alpha-3}(s)s_j^3\sigma(\mathrm{d}s).
\end{aligned}
$$

And for the fourth cumulan of X_j,

$$
\begin{aligned}
c_4 &= \mathrm{E}[(X_j-\mathrm{E}(X_j))^4] - 3\mathrm{E}^2[(X_j-\mathrm{E}(X_j))^2] = (-i)^4\frac{\partial^4 C_\mu(z)}{\partial^4 z_j}\Big|_{z=0}\\
&= \frac{\partial^2\left[\int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)b^{\alpha-2}(s)(-is_j)^2\sigma(\mathrm{d}s)\right]}{\partial^2 z_j}\Big|_{z=0}\\
&= \frac{\partial\left[\int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)(\alpha-2)(b(s)-i\sum_{j=1}^{d}s_jz_j)^{\alpha-3}i(s_j)^3\sigma(\mathrm{d}s)\right]}{\partial z_j}\Big|_{z=0}\\
&= \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)(\alpha-2)(\alpha-3)b^{\alpha-4}(s)s_j^4\sigma(\mathrm{d}s)\\
&= \Gamma(4-\alpha)\int_{\mathbb{S}^{d-1}}b^{\alpha-4}(s)s_j^4\sigma(\mathrm{d}s).
\end{aligned}
$$

Then, let's see the covariance between X_i and X_j

$$
\begin{aligned}
c_{11} &= \mathrm{Cov}(X_i,X_j) = (-i)^2\frac{\partial^2 C_\mu(z)}{\partial z_i\partial z_j}\Big|_{z=0}\\
&= -\frac{\partial\left[i\gamma_i + \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha b^{\alpha-1}(s)(-is_i)\sigma(\mathrm{d}s)\right]}{\partial z_j}\Big|_{z=0}\\
&= \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)(b(s)-i\sum_{i=1}^{d}s_iz_i)^{\alpha-2}s_is_j\sigma(\mathrm{d}s)\Big|_{z=0}\\
&= \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)\alpha(\alpha-1)b^{\alpha-2}(s)s_is_j\sigma(\mathrm{d}s)\\
&= \Gamma(2-\alpha)\int_{\mathbb{S}^{d-1}}b^{\alpha-2}(s)s_is_j\sigma(\mathrm{d}s)
\end{aligned}
$$

Last, we could easy to get the skewness and kurtosis by equation (2.5), (2.6) and (2.7). □

A.2 Proof of Lemma 1

Lemma 11. *Let $\{u_n\}$ be a sequence in $\mathbb{R}^d$. If for every $v \in \mathbb{R}^d$, $\langle v, u_n\rangle \to 0$ as $n \to \infty$, then $u_n \to 0$ as $n \to \infty$.*

Proof. First, we show there no subsequences diverges to infinite. Assume for the sake of contradiction that there is a subsequence $\{u_{n_k}\}$, with $|u_{n_k}| \to \infty$. For large enough k, $u_{n_k} \neq 0$ and we can write $u_{n_k} = |u_{n_k}|\frac{u_{n_k}}{|u_{n_k}|}$. Since $\frac{u_{n_k}}{|u_{n_k}|} \in \mathbb{S}^{d-1}$, every component of this vector is bounded, and thus there exists a further subsequence $\{u_{n_{k_i}}\}$ and a $\xi \in \mathbb{S}^{d-1}$ with $\frac{u_{n_{k_i}}}{|u_{n_{k_i}}|} \to \xi$. It follows that for every $v \in \mathbb{R}^d$, $\left\langle v, u_{n_{k_i}}/|u_{n_{k_i}}|\right\rangle \to \langle v, \xi\rangle$, which equals 0 since $|u_{n_{k_1}}| \to \infty$ while $|u_{n_{k_i}}|\left\langle v, u_{n_{k_i}}/|u_{n_{k_1}}|\right\rangle = \left\langle v, u_{n_{k_i}}\right\rangle \to 0$. This means that ξ is orthogonal to every $v \in \mathbb{R}^d$, and hence that that $\xi = 0$, but $\xi \in \mathbb{S}^{d-1}$, which gives the contradiction. Now, consider any convergent subsequence $\{u_{n_k}\}$. Thus, there is a $u \in \mathbb{R}^d$ with $u_{n_k} \to u$. It follows that for every $v \in \mathbb{R}^d$, $\langle v, u\rangle = \lim_{k\to\infty}\langle v, u_{n_k}\rangle = 0$. Thus u is orthogonal to every vector in $\mathbb{R}^d$ and, hence, $u = 0$. Thus every subsequence converges to 0 and the result holds. □

Proof of Lemma 1. For $\beta > 0$ and $\xi \in \mathbb{S}^{d-1}$, let $u_\beta(\xi) = \int_{\mathbb{S}^{d-1}} |\langle s, \xi\rangle|^\beta \sigma(\mathrm{d}s)$. Note that $u_\beta = \inf_{\xi\in\mathbb{S}^{d-1}} u_\beta(\xi)$. We begin by showing that the first condition implies the second. Assume, for the sake of contradiction, that σ is full and that $u_\beta = 0$ for some $\beta > 0$. It follows that there is a sequence $\xi_1, \xi_2, \cdots \in \mathbb{S}^{d-1}$ with $u_\beta(\xi_n) \to 0$. By a version of Chebyshev's inequality (see e.g. Section 6.3 in [55]) we have, for any $h > 0$,

$$\sigma(\{s \in \mathbb{S}^{d-1} : |\langle s, \xi_n\rangle| > h\}) \leq h^{-\beta} u_\beta(\xi_n) \to 0.$$

It follows that the sequence of function $f_n : \mathbb{S}^{d-1} \mapsto \mathbb{R}$ given by $f_n(s) = \langle s, \xi_n\rangle$ converges to 0 in measure σ. From here, Theorem 2.30 in [55] implies that there is a

subsequence with

$$\lim_{k\to\infty} \langle s, \xi_{n_k} \rangle = 0 \quad \text{for } \sigma - a.e.\ s$$

Let $A \in \mathfrak{B}(\mathbb{S}^{d-1})$ be the set on which this convergence holds and note that $\sigma(A^c) = 0$. Let $\bar{A}$ be the closure of A and note that the support of σ is contained in $\bar{A}$. Since the support of σ is assumed to contain d-linearly independent vectors, it follows that $\bar{A}$ does as well, and by properties of closures so does A. Thus, A is not contained in a proper subspace of $\mathbb{R}^d$. This implies that there are distinct vectors $s_1, s_2, \ldots, s_d \in A$ that span $\mathbb{R}^d$. It follows that for any vector $y \in \mathbb{R}^d$, we have $y = \sum_{j=1}^{d} a_j s_j$ for some $a_1, a_2, \ldots, a_d \in \mathbb{R}$. Hence

$$\langle y, \xi_{n_k} \rangle = \sum_{j=1}^{d} a_j \langle s_j, \xi_{n_k} \rangle \to 0.$$

From here, Lemma 11 implies that $\xi_{n_k} \to 0$, which contradicts the fact that $\xi_{n_k} \in \mathbb{S}^{d-1}$ for each k.

It is immediate that the second condition implies the third. We now show that the third implies the first. Our proof is by contrapositive. Assume that there are less than d linearly independent vectors in the support of σ. It follows that there exists a $\xi^* \in \mathbb{S}^{d-1}$ with $\langle s, \xi^* \rangle = 0$ for each s in the support of σ. Hence, $u_\beta(\xi^*) = 0$ and $u_\beta = 0$. □

A.3 Proof of Lemma 12 and Lemma 13

Lemma 12. *Fix $\alpha \in (0,1)$, $\gamma \in \mathbb{R}^d$, let $b : \mathbb{S}^{d-1} \mapsto (0,\infty)$ be a Borel function, and let σ be a finite Borel measure on $\mathbb{S}^{d-1}$. Let's $\mu \sum \mathrm{TS}_\alpha(\sigma, b, \gamma)$ and its characteristic function is given, for any $z \in \mathbb{R}^d$,*

$$\hat{\mu}(z) = \exp\left[i\langle \gamma, z \rangle + \int_{\mathbb{S}^{d-1}} \Gamma(-\alpha)((b(s) - i\langle s, z \rangle)^\alpha - b^\alpha(s))\sigma(ds) \right]$$

.

Proof. By equation (1), we know

$$\hat{\mu}(z) = \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\int_0^\infty (e^{i\langle s,z\rangle x} - 1)\frac{e^{-b(s)x}}{x^{1+\alpha}}dx\sigma(\mathrm{d}s)\right],$$

so for this lemma, we will show these two expressions are equivalent. We have

$$\begin{aligned}
\hat{\mu}(z) &= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\int_0^\infty (e^{i\langle s,z\rangle x} - 1)\frac{e^{-b(s)x}}{x^{1+\alpha}}dx\sigma(\mathrm{d}s)\right] \\
&= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\int_0^\infty \left(\sum_{k=0}^\infty \frac{(i\langle s, z\rangle x)^k}{k!} - 1\right)\frac{e^{-b(s)x}}{x^{1+\alpha}}dx\sigma(\mathrm{d}s)\right] \\
&= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\int_0^\infty \sum_{k=1}^\infty \frac{(i\langle s, z\rangle x)^k}{k!}\frac{e^{-b(s)x}}{x^{1+\alpha}}dx\sigma(\mathrm{d}s)\right] \\
&= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\sum_{k=1}^\infty \frac{(i\langle s, z\rangle)^k}{k!}\int_0^\infty x^{k-1-\alpha}e^{-b(s)x}dx\sigma(\mathrm{d}s)\right],
\end{aligned}$$

where the second line follows by $e^x = \sum_{k=0}^\infty \frac{x^k}{k!}$. By change of variable,

$$\begin{aligned}
\hat{\mu}(z) &= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\sum_{k=1}^\infty \frac{(i\langle s, z\rangle)^k}{k!}\int_0^\infty \left(\frac{x}{b(s)}\right)^{k-1-\alpha} e^{-x}\frac{1}{b(s)}dx\sigma(\mathrm{d}s)\right] \\
&= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\sum_{k=1}^\infty \frac{(i\langle s, z\rangle)^k}{k!}b^{\alpha-k}(s)\Gamma(k-\alpha)\sigma(\mathrm{d}s)\right] \\
&= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\sum_{k=1}^\infty \frac{(i\langle s, z\rangle)^k}{k!}b^{\alpha-k}(s)\Gamma(-\alpha)\prod_{u=0}^{k-1}(u-\alpha)\sigma(\mathrm{d}s)\right] \\
&= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)b^\alpha(s)\sum_{k=1}^\infty \left(\frac{(i\langle s, z\rangle)}{b(s)}\right)^k (-1)^k\prod_{u=1}^{k}\frac{\alpha-u+1}{u}\sigma(\mathrm{d}s)\right],
\end{aligned}$$

where the second line follows by the definition of the gamma function, the third line follows by $\Gamma(1+a) = a\Gamma(a)$. Then,

$$\begin{aligned}
\hat{\mu}(z) &= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)b^\alpha(s)\sum_{k=1}^\infty \binom{\alpha}{k}\left(\frac{-i\langle s, z\rangle}{b(s)}\right)^k\sigma(\mathrm{d}s)\right] \\
&= \exp\left[i\langle\gamma, z\rangle + \int_{\mathbb{S}^{d-1}}\Gamma(-\alpha)b^\alpha(s)\left(\left(1 - \frac{i\langle s, z\rangle}{b(s)}\right)^\alpha - 1\right)\sigma(\mathrm{d}s)\right]
\end{aligned}$$

$$= \exp\left[i\langle \gamma, z\rangle + \int_{\mathbb{S}^{d-1}} \Gamma(-\alpha)((b^{\alpha}(s) - i\langle s, z\rangle)^{\alpha} - b^{\alpha}(s))\sigma(\mathrm{d}s)\right],$$

where the second line follows by the binomial theorem. □

Lemma 13. *Let* $w = (b(s) - i\langle s, z\rangle)^{\alpha}$, *then*

$$\Re(w) = (b^2(s) + \langle s, z\rangle^2)^{\frac{\alpha}{2}} \cos\left(\alpha \arctan\left(\frac{\langle s, z\rangle}{b(s)}\right)\right) \tag{5}$$

and

$$\Im(w) = -(b^2(s) + \langle s, z\rangle^2)^{\frac{\alpha}{2}} \sin\left(\alpha \arctan\left(\frac{\langle s, z\rangle}{b(s)}\right)\right) \tag{6}$$

Proof.

$$\begin{aligned}
w &= (b(s) - i\langle s, z\rangle)^{\alpha} = \exp[\alpha \log(b(s) - i\langle s, z\rangle)] \\
&= \exp\left[\alpha\left(\log(|b(s) - i\langle s, z\rangle|) + i \arctan\left(\frac{-\langle s, z\rangle}{b(s)}\right)\right)\right] \\
&= \exp\left[\frac{\alpha}{2}\log\left(b^2(s) + \langle s, z\rangle^2\right)\right] * \exp\left[i * \alpha \arctan\left(\frac{-\langle s, z\rangle}{b(s)}\right)\right] \\
&= (b^2(s) + \langle s, z\rangle^2)^{\frac{\alpha}{2}}\left(\cos\left(\alpha \arctan\left(\frac{-\langle s, z\rangle}{b(s)}\right)\right) + i \sin\left(\alpha \arctan\left(\frac{-\langle s, z\rangle}{b(s)}\right)\right)\right) \\
&= (b^2(s) + \langle s, z\rangle^2)^{\frac{\alpha}{2}} \cos\left(\alpha \arctan\left(\frac{\langle s, z\rangle}{b(s)}\right)\right) - i(b^2(s) \\
&\quad + \langle s, z\rangle^2)^{\frac{\alpha}{2}} \sin\left(\alpha \arctan\left(\frac{\langle s, z\rangle}{b(s)}\right)\right),
\end{aligned}$$

where the third equality follows by the 4.1.2 and 4.1.3 in [51], and the last equality follows by the fact that the function arctan(.) and sin(.) are odd functions and cos(.) is an even function. □