Stable Distribution of Multivariate Data

Vo Thi Truc Giang and Ho Dang Phuc

***Abstract* — The main theorem of the paper states that every stable random vector with marginal skewness parameters different from ±1 can be turned into a sub-Gaussian random vector by using an appropriately tailored transformation in multidimensional space. The theorem is used to create a formula on probability density function of stable random vector and to perform a procedure of testing the stable distribution of multivariate data. A dataset collected from the Nasdaq stock market is used to illustrate the proposed procedure.**

***Keywords* — Data analysis; Heavy-tailed distribution; Portfolio selection; Stock market.**

I. INTRODUCTION

Most of the traditional statistical analysis methods were developed under normality assumptions. The important class of normal distributions established itself in many years as a cornerstone of the most successful models in various areas of modern quantitative finance. These include, for instance, Markowitz' pioneering model for portfolio selection and asset allocation [1], Ross' arbitrage pricing theory [2], the famous capital asset pricing model of Sharpe [3], Treynor, Lintner and Mossin. In applications, however, normality is only a poor approximation of reality. Namely, whilst normal distributions are always symmetric around their mean, most of quantities usually concerned in empirical studies do not have symmetric distributions. For instance, observable returns mostly exhibit asymmetry in favor of large negative return deviations. Besides, normal distributions do not allow heavy tails, which are so common, especially in finance and risk management studies [4]–[9].

Stable distributions are asymmetric heavy-tailed extensions of normal distributions and have attracted a lot of attention in applied research [7], [9]–[13]. The univariate stable distributions are actually accessible by methods to estimate stable parameters and reliable programs to compute stable probability density functions, cumulative distribution functions, and quantiles for stable random variables [10], [14]–[16]. However, the use of the heavy-tailed models in practice has been restricted by the lack of tools for multivariate stable distributions.

The distribution function calculation problem of stable random vectors remains open in the general cases. Besides, in studies on portfolio selection and asset allocation, analysts must determine the joint probability density function and the cumulative distribution function of a linear combination of several stable random variables. It is worthy to notice that a random vector has stable distribution when and only when all linear combinations of its marginal have stable distribution with a common stable index. Thus, the problems of finding the multivariate stable probability density function and of testing the stable distribution of a random vector plays an important role in application. That convinces the aim of this paper to find a formula of multivariate stable probability density function estimation and to create a procedure of goodness-of-fit testing for a broader family of multivariate stable distributions.

Let $\mathbf{X} = (X_1, \ldots, X_d)$ be a *random vector* (rv hereafter) taking values in $\mathbb{R}^d$, its *cumulative distribution function* (cdf) and *probability density function* (pdf) are denoted by $F_{\mathbf{X}}$ and $f_{\mathbf{X}}$. The coordinates $X_1, \ldots, X_d$ are called *marginal's*, simultaneously $F_{X_1}, \ldots, F_{X_d}$ and $f_{X_1}, \ldots, f_{X_d}$ are *called marginal cdf's* and *marginal pdf's* of $\mathbf{X}$. The rv $\mathbf{X}$ is said to have *stable distribution* if for every pair $(\mathbf{X}', \mathbf{X}'')$ of independent rv's identically distributed as $\mathbf{X}$, for every pair (a, b) of positive numbers, there exist a positive number c and a vector $\mathbf{d} \in \mathbb{R}^d$such that $a\mathbf{X}' + b\mathbf{X}''$ has the same distribution as $c\mathbf{X} + \mathbf{d}$. It is well known (Theorem 2.3.1 [17]) that the stable distribution of $\mathbf{X}$ is determined by a spectral measure Λ (a finite Borel measure on the unit sphere $\mathbb{S}_d$in $\mathbb{R}^d$) and a shift vector $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_d) \in \mathbb{R}^d$ through the representation

$$\varphi_{\mathbf{X}}(\mathbf{t}) = exp\left(-\int_{S_d} \psi_\alpha\left(\langle \mathbf{s}, \mathbf{t}\rangle\right)\Lambda(ds) + i\langle \boldsymbol{\delta}, \mathbf{t}\rangle\right) \quad (1)$$

where $\varphi_{\mathbf{X}}$ is the *characteristic function* of $\mathbf{X}$ defined by $\varphi_{\mathbf{X}}(\mathbf{t}) := \mathrm{E}\ exp\{i\langle \mathbf{X}, \mathbf{t}\rangle\}$,
for $\mathbf{t} = (t_1, \ldots, t_d) \in \mathbb{R}^d, \langle \mathbf{x}, \mathbf{t}\rangle = x_1t_1 + \ldots + x_dt_d$, and

Published on August 31, 2023.
Vo Thi Truc Giang, Tien Giang University, Vietnam.
(e-mail: vothitrucgiang@tgu.edu.vn).
Ho Dang Phuc, Institute of Mathematics, Vietnam Academy of Science and Technology, Vietnam.
(corresponding e-mail: hdphuc@math.ac.vn)

$$\psi_\alpha(u)=\begin{cases}|u|^\alpha\left(1-isign(u)tan\frac{\pi\alpha}{2}\right) & \alpha\neq 1\\ |u|\left(1+i\frac{2}{\pi}sign(u)ln|u|\right) & \alpha=1.\end{cases}$$

We denote $\mathbf{X}\sim S(\alpha;\Lambda;\boldsymbol{\delta})$ to mark (1) is valid. Moreover, $\mathbf{X}$ is said to be *α-stable*. Especially, characteristic function of α-stable random variable has the form

$$\varphi_X(u)=\mathrm{E}\exp(iuX)=\begin{cases}\exp\left(-\gamma^\alpha|u|^\alpha\left[1-i\beta\left(tan\frac{\pi\alpha}{2}\right)sign(u)+i\delta u\right]\right) & \alpha\neq 1\\ \exp\left(-\gamma|u|\left[1+i\beta\frac{2}{\pi}sign(u)ln|u|+i\delta u\right]\right) & \alpha=1.\end{cases}$$

with fixed $\beta\in[-1;1],\gamma>0$ and $\delta\in\mathbb{R}$. Then the distribution of X is uniquely determined by the parameters α,β,γ and δ, we write $X\sim S(\alpha;\beta;\gamma;\delta)$. Usually, α is called the *stable index*, meanwhile β, γ and δ are named as the *skewness*, the *scale* and the *location* parameters of X, respectively.

II. STABLE RANDOM VECTOR AND SUB-GAUSSIAN RANDOM VECTOR

For fixed $\alpha\in(0,2]$, let $A\sim S(\alpha/2;\ 1;\ [\cos(\pi\alpha/4)]^{2/\alpha};0)$ be a positive α/2-stable random variable and $G=(G_1,\dots,G_d)$ be a zero-mean Gaussian vector independent of A. Then $\mathbf{X}=(A^{1/2}G_1,\dots,A^{1/2}G_d)$ is called a *sub-Gaussian* rv. By virtue of Theorems 1.3.1 and 2.1.5 [17], $\mathbf{X}$ is an α-stable rv. Moreover, the following result is an immediate consequence of Proposition 2.5.5 [17].

LEMMA 2.1. *For $\alpha\in(0,2]$, let $\mathbf{X}\sim S(\alpha;\boldsymbol{\Lambda};\mathbf{0})$ be an α-stable rv in $\mathbb{R}^d$. Suppose that the distribution of $\mathbf{X}$ is isotropic, which means $\mathbf{X}\overset{d}{=}M\mathbf{X}$ for all orthonormal $d\times d$ - matrices M (corresponding to linear rotations around the origin $\mathbf{0}$), where $\overset{d}{=}$ denotes the equality in distribution. Then $\mathbf{X}$ is sub-Gaussian with Gaussian vector $\boldsymbol{G}$ having iid components $G_i, i=1,\dots,d$.*

In the sequel, we use repeatedly the couple of the polar representation mappings $B=(B_0,\dots,B_{d-1})$ and $D=(D_1,\dots,D_d)$ defined by (2) and (3) as the follows. For $\mathbf{x}=(x_1,\dots,x_d)\in\mathbb{R}^d\backslash\{\mathbf{0}\}$, let

$$r=B_0(\mathrm{x})=\|\mathrm{x}\|=\sqrt{x_1^2+\dots+x_d^2},$$
$$\theta_i=B_i(\mathrm{x})=\operatorname{arccot}\left(x_i/\sqrt{x_{i+1}^2+\dots+x_d^2}\right), i=1,2,\dots,d-2 \tag{2}$$
$$\theta_{d-1}=B_{d-1}(\mathbf{x})=2\operatorname{arccot}\left(\left[x_{d-1}+\sqrt{x_{d-1}^2+x_d^2}\right]/x_d\right).$$

Then $B=(B_0,\dots,B_{d-1})\colon\mathbb{R}^d\backslash\{\mathbf{0}\}\to\mathbb{R}^+\times\mathbb{I}^{d-1}$ is a bijective mapping,

$$\mathrm{I}^{d-1}:=\underbrace{[0,\pi)\times\dots\times[0,\pi)}_{d-2}\times[0,2\pi)\subset{}^{\sim d-1}.$$

Inversely, each point $\mathbf{x}\in\mathbb{R}^d\backslash\{\mathbf{0}\}$ is of the form

$$\begin{aligned}x_1&=D_1(r,\theta_1,\dots,\theta_{d-1})=r\cos\theta_1,\\ x_2&=D_2(r,\theta_1,\dots,\theta_{d-1})=r\sin\theta_1\cos\theta_2,\\ &\dots\\ x_{d-1}&=D_{d-1}(r,\theta_1,\dots,\theta_{d-1})=r\sin\theta_1\sin\theta_2\dots\sin\theta_{d-2}\cos\theta_{d-1},\\ x_d&=D_{d-1}(r,\theta_1,\dots,\theta_{d-1})=r\sin\theta_1\sin\theta_2\dots\sin\theta_{d-2}\sin\theta_{d-1},\end{aligned}\tag{3}$$

with $0\le\theta_1<\pi,\dots,0\le\theta_{d-2}<\pi, 0\le\theta_{d-1}<2\pi$. Then the mapping $D=(D_1,\dots,D_d)\colon\mathbb{R}^+\times\mathbb{I}^{d-1}\to\mathbb{R}^d\backslash\{0\}$ is the inverse transformation of B.

LEMMA 2.2. *Given $\alpha\in(0,2]$ and a symmetric $\alpha-stable\ rv$ $\mathbf{Z}=(Z_1,\dots,Z_d)$, let $B(\mathbf{Z})=(R,\Theta_1,\dots,\Theta_{d-1})$. Suppose that there exists an invertible differentiable transformation $Q\colon\mathbb{I}^{d-1}\to\mathbb{I}^{d-1}$ such that $\mathbf{Y}=D(R,Q(\Theta_1,\dots,\Theta_{d-1}))$ is symmetric. Then $\mathbf{Y}$ has an α-stable distribution.*

PROOF. From the symmetry of $\mathbf{Z}$ and $\mathbf{Y}$ we see their characteristic function $\varphi_{\mathbf{Z}}$ and $\varphi_{\mathbf{Y}}$ taking only real values and

$$\varphi_{\mathbf{Z}}(\mathbf{t})=\mathrm{E}\cos\langle\mathbf{Z},\mathbf{t}\rangle,\ \varphi_{\mathbf{Y}}(\mathbf{t})=\mathrm{E}\cos\langle\mathbf{Y},\mathbf{t}\rangle \tag{4}$$

for $\mathbf{t}=(t_1,\dots,t_d)\in\mathbb{R}^d$, where $\langle\mathbf{z},\mathbf{t}\rangle=z_1t_1+\dots+z_dt_d$ for $\mathbf{z}=(z_1,\dots,z_d)$. Meantime, the spectral representation of symmetric stable multivariate distributions (see e.g. Theorem 2.4.3 [17]) confirms that

$$\varphi_{\mathbf{Z}}(\mathbf{t})=\exp\left(-\int_{\mathrm{S}_d}|\langle\mathbf{s},\mathbf{t}\rangle|^\alpha\Lambda(ds)\right) \tag{5}$$

with a spectral measure Λ on $\mathbb{S}_d$. Looking at the polar representations

$$
\begin{aligned}
z_1 &= r\cos\theta_1, & t_1 &= u\cos\eta_1,\\
z_2 &= r\sin\theta_1\cos\theta_2, & t_2 &= u\sin\eta_1\cos\eta_2,\\
&\ldots & &\ldots.\\
z_{d-1} &= r\sin\theta_1\ldots\sin\theta_{d-2}\cos\theta_{d-1}, & t_{d-1} &= u\sin\eta_1\ldots\sin\eta_{d-2}\cos\eta_{d-1},\\
z_d &= r\sin\theta_1\ldots\sin\theta_{d-2}\sin\theta_{d-1}, & t_d &= u\sin\eta_1\ldots\sin\eta_{d-2}\sin\eta_{d-1},
\end{aligned}
$$

where $0 \le \theta_1 < \pi, \ldots, 0 \le \theta_{d-2} < \pi, 0 \le \theta_{d-1} < 2\pi; 0 \le \eta_1 < \pi, \ldots, 0 \le \eta_{d-2} < \pi, 0 \le \eta_{d-1} < 2\pi$ and $r = \|\mathbf{z}\|$; $u = \|\mathbf{t}\|$, we see

$$\langle \mathbf{z}, \mathbf{t} \rangle = ru \cdot \upsilon(\boldsymbol{\theta}, \boldsymbol{\eta}), \tag{6}$$

the function υ depends only of the arguments $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{d-1}), \boldsymbol{\eta} = (\eta_1, \ldots, \eta_{d-1})$ from $\mathbb{I}^{d-1}$. The existence of pdf $f_{\mathbf{Z}}$ of the stable rv $\mathbf{Z}$ and the assumption that Q is differentiable imply the pdf $f_{\mathbf{Y}}$ of the rv $\mathbf{Y}$ exists, that yields

$$f_{\mathbf{Y}}^*(r, \theta_1, \ldots, \theta_{d-1}) = f_{\mathbf{Z}}^*(r, Q(\boldsymbol{\theta}))J_Q,$$

where f^* means the polar coordinates form of a multivariate function f and J_Q denotes the Jacobian of Q, that is dependent only on the arguments $\theta_1, \ldots, \theta_{d-1}$. Therefore (4) can be rewritten as

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = \int_{\mathrm{I}^{d-1}} \int_0^\infty \cos(ru \cdot \upsilon(\boldsymbol{\theta}, \boldsymbol{\eta})) f_{\mathbf{Z}}^*(r, \boldsymbol{\theta}) dr d\boldsymbol{\theta} \tag{7}$$

and

$$\varphi_{\mathbf{Y}}(\mathbf{t}) = \int_{\mathrm{I}^{d-1}} \int_0^\infty \cos(ru \cdot \upsilon(Q(\boldsymbol{\theta}), \boldsymbol{\eta})) f_{\mathbf{Z}}^*(r, Q(\boldsymbol{\theta})) J_Q dr d\boldsymbol{\theta} \tag{8}$$

Simultaneously, (5) implies

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = \exp\left(-\int_{\mathrm{I}^{d-1}} \left|u \cdot \upsilon(\boldsymbol{\theta}, \boldsymbol{\eta})^\alpha\right| \Lambda_*(d\boldsymbol{\theta})\right),$$

where Λ_* denotes the polar representation form of Λ. Combining the above equality with (6), (7), and (8), it is easily pointed out that

$$\varphi_{\mathbf{Y}}(\mathbf{t}) = \exp\left(-\int_{\mathrm{I}^{d-1}} \left|u \cdot \upsilon(\boldsymbol{\theta}, \boldsymbol{\eta})^\alpha\right| \Lambda_*^1(d\boldsymbol{\theta})\right) = \exp\left(-\int_{\mathrm{S}_d} \left|\langle \mathbf{s}, \mathbf{t} \rangle\right|^\alpha \Lambda^1(ds)\right)$$

for some spectral measure Λ^1 defined on $\mathbb{S}_d$. Consequently, the rv $\mathbf{Y}$ is α-stable by virtue of Theorem 2.4.3, [17]. □

For a given cdf $G: \mathbb{R} \to [0; 1]$, let $G^{\leftarrow}(y) = \inf\{x: G(x) \ge y\}$ be its generalized inverse. The *copula* of rv $\mathbf{X}$, denoted by $C_{\mathbf{X}}$, can be defined by

$$C_{\mathbf{X}}(t_1, \ldots, t_d) = F_{\mathbf{X}}(F_{X_1}^{\leftarrow}(x_1), \ldots, F_{X_d}^{\leftarrow}(x_d)),$$

for $0 \le t_1, \ldots, t_d \le 1$. Then we have (see also Sklar's Theorem [18])

$$F_{\mathbf{X}}(x_1, \ldots, x_d) = C_{\mathbf{X}}(F_{X_1}(x_1), \ldots, F_{X_d}(x_d)), \tag{9}$$

for $x_1, \ldots, x_d \in \overline{\mathbb{R}} = [-\infty; +\infty]$. Moreover, if $\mathbf{X}$ is continuous then $F_{X_k}^{\leftarrow} = F_{X_k}^{-1}$ for $k = 1, \ldots, d$. By virtue of (9), if C is a copula of any stable rv and $X_1, \ldots, X_d$ are stable then

$$F_{\mathbf{X}}(x_1, \ldots, x_d) = C(F_{X_1}(x_1), \ldots, F_{X_d}(x_d))$$

is a cdf of a stable rv.

THEOREM 2.3. *Let* **X** *be a rv with stable distribution such that its marginals have skewness parameters different from* ± 1. *Then there exists for* **X** *an invertible differentiable transformation* $K: \mathbb{R}^d \to \mathbb{R}^d$ *such that the rv* $\mathrm{Y} = K(\mathbf{X})$ *is a sub-Gaussian rv.*

PROOF. Let $X_1, \ldots, X_d$ be a stable rv of marginals $X_k \sim S(\alpha; \beta_k; \gamma_k; \delta_k)$, $0 < \alpha < 2$; $-1 < \beta_k < 1$; $0 < \gamma_k < \infty, k = 1, \ldots, d$. Let $X_0 \sim S(\alpha; 0; 1; 0)$ denote a standard stable random variable. From the well known-fact (Property 1.2.14 [17]) that if a stable random variable has skewness parameter different from ±1 then its cdf is positive on whole $\mathbb{R}$, the pdf's $f_{X_1}, \ldots, f_{X_d}$and f_{X_0} are positive on whole $\mathbb{R}$, the cdf's $F_{X_1}, \ldots, F_{X_d}$ and F_{X_0} are strictly increasing. Then the functions $T_k: \mathbb{R} \to \mathbb{R}, k = 1, \ldots, d$, defined by $T_k(u) = F_{X_0}^{-1}(F_{X_k}(u))$, are strictly increasing functions. Besides,

$$T_k'(u) = f_{X_k}(u) / f_{X_0}(T_k(u)), \quad k = 1, \ldots, d$$

are positive functions. This implies $f_{X_0}(T_k(u))T'_k(u)du = f_{X_k}(u)du$, that yields

$$F_{X_0}(T_k(u)) = \int_{-\infty}^{T_k(u)} f_{X_0}(T_k(u))T_k'(u)du = \int_{-\infty}^{u} f_{X_k}(u)du = F_{X_k}(u). \tag{10}$$

On the other hand, for every $t \in \mathbb{R}$ we have

$$F_{T_k \circ X_k}(t) = \mathrm{P}\{\omega : T_k(X_k(\omega)) \le t\} = \mathrm{P}\{\omega : X_k(\omega) \le T_k^{-1}(t)\} = F_{X_k}(T_k^{-1}(t)).$$

Compared the above with (10), by putting $t = T_k(u)$, we get $F_{X_0}(t) = F_{T_k \circ X_k}(t)$. This confirms the two random variables X_0 and $T_k \circ X_k$ have the same distribution. Then, by virtue of Proposition 5.6 [19] and the remark after (9), the rv $\mathbf{Z} = (T_1(X_1), \ldots, T_d(X_d))$ is stable with symmetric $S(\alpha; 0; 1; 0)$-distributed marginals.

Let $\mathbf{e}_1 = (1,0,0,\ldots,0), \mathbf{e}_2 = (0,1,0,\ldots,0), \ldots, \mathbf{e}_d = (0,0,0,\ldots,1)$ be the unit vectors in the basis of the Euclidean space $\sim^d$. Let $\mathcal{P}_{(\mathbf{e}_i, \mathbf{e}_j)}$ denote the two-dimensional plan generated by $\{\mathbf{e}_i, \mathbf{e}_j\}$ for all $1 \le i, j \le d-1$. It is clear that every shift (modulo 2π) of θ_{d-1} in the interval [0;2π) corresponds to one rotation of $D(r, \theta_1, \ldots, \theta_{d-1})$ around the origin in $\mathcal{P}_{(\mathbf{e}_{d-1}, \mathbf{e}_d)}$.

Considering $(R, \Theta_1, \ldots, \Theta_{d-1}) = B(\mathbf{Z})$, we see the random variable $F_{\Theta_{d-1}} \circ \Theta_{d-1}$uniformly distributed on [0;1). Consequently, the random variable $2\pi F_{\Theta_{d-1}} \circ \Theta_{d-1}$ is uniformly distributed on [0; 2π), its distribution is invariant against every shift (modulo 2π) in the interval [0; 2π). Then, the distribution of the new defined rv

$$U^{(1)}(\mathbf{Z}) = \mathbf{Z}^{(1)} = \left(Z_1^{(1)}, Z_2^{(1)}, \ldots, Z_d^{(1)}\right) := D\left(R, \Theta_1, \ldots, \Theta_{d-2}, 2\pi F_{\Theta_{d-1}} \circ \Theta_{d-1}\right)$$

is symmetric and invariant under all rotations around the origin in $\mathcal{P}_{(\mathbf{e}_{d-1}, \mathbf{e}_d)}$.

In the second step, we change the coordinates of $(Z_1^{(1)}, Z_2^{(1)}, \ldots, Z_d^{(1)})$ by moving the first coordinate to the end and shifting the others ahead by one place (that means the basis $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{d-1}, \mathbf{e}_d)$ in $\mathbb{R}^d$ is replaced by the basis $(\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_d, \mathbf{e}_1)$). With the new rv, we get

$$\left(R, \Theta_1^{(1)}, \ldots, \Theta_{d-1}^{(1)}\right) = B\left(Z_2^{(1)}, Z_3^{(1)}, \ldots, Z_d^{(1)}, Z_1^{(1)}\right).$$

The second transformation $U^{(2)}(\mathbf{Z}^{(1)}) = \mathbf{Z}^{(2)}$is defined by

$$\mathbf{Z}^{(2)} = \left(Z_1^{(2)}, Z_2^{(2)}, \ldots, Z_d^{(2)}\right) := D\left(R, \Theta_1^{(1)}, \ldots, \Theta_{d-2}^{(1)}, 2\pi F_{\Theta_{d-1}^{(1)}} \circ \Theta_{d-1}^{(1)}\right).$$

The random variable $2\pi F_{\Theta_{d-1}^{(1)}} \circ \Theta_{d-1}^{(1)}$is uniformly distributed on [0;2π), then by a similar argument as the above, we confirm that the distribution of the rv $\mathbf{Z}^{(2)}$ is symmetric and invariant under all rotations around the origin in $\mathcal{P}_{(\mathbf{e}_d, \mathbf{e}_1)}$. Simultaneously, $\mathbf{Z}^{(2)}$is also invariant under all rotations around the origin in $\mathcal{P}_{(\mathbf{e}_{d-1}, \mathbf{e}_d)}$ because $\mathbf{Z}^{(1)}$ is invariant under all rotations around the origin in $\mathcal{P}_{(\mathbf{e}_{d-1}, \mathbf{e}_d)}$.

Continuing the above process to the *d*-th step, the basis $(\mathbf{e}_{d-1}, \mathbf{e}_d, \ldots, \mathbf{e}_{d-3}, \mathbf{e}_{d-2})$ is replaced by the basis $(\mathbf{e}_d, \mathbf{e}_1, \ldots, \mathbf{e}_{d-2}, \mathbf{e}_{d-1})$, the rv obtained after the (*d*-1)-th step $(Z_1^{(d-2)}, Z_2^{(d-2)}, \ldots, Z_{d-1}^{(d-2)}, Z_d^{(d-2)})$ is rearranged into the rv $(Z_2^{(d-2)}, Z_3^{(d-2)}, \ldots, Z_d^{(d-2)}, Z_1^{(d-2)})$. Then the polar representation of the new rv is

$$B\left(Z_2^{(d-2)}, \ldots, Z_d^{(d-2)}, Z_1^{(d-2)}\right) = \left(R, \Theta_1^{(d-2)}, \ldots, \Theta_{d-2}^{(d-2)}, \Theta_{d-1}^{(d-2)}\right).$$

By applying *D* to the polar representation rv replaced the last coordinate $\Theta_{d-1}^{(d-2)}$ by $2\pi F_{\Theta_{d-1}^{(d-2)}} \circ \Theta_{d-1}^{(d-2)}$, we have $U^{(d-1)}(\mathbf{Z}^{(d-2)}) = \mathbf{Z}^{(d-1)}$, where

$$\mathbf{Z}^{(d-1)} = \left(Z_1^{(d-1)}, Z_2^{(d-1)}, \ldots, Z_d^{(d-1)}\right) := D\left(R, \Theta_1^{(d-2)}, \ldots, \Theta_{d-2}^{(d-2)}, 2\pi F_{\Theta_{d-1}^{(d-2)}} \circ \Theta_{d-1}^{(d-2)}\right).$$

For the same reason as presented above, we can conclude the distribution of the rv $\mathbf{Z}^{(d-1)}$ has symmetric distribution and is invariant under every two-dimensional rotation around the origin in $\mathcal{P}_{(\mathbf{e}_{d-2}, \mathbf{e}_{d-1})}$. Consequently, the distribution is invariant under every two-dimensional rotation around the origin in each of the two-dimensional planes $\mathcal{P}_{(\mathbf{e}_{d-1}, \mathbf{e}_d)}$,. This confirms the fact that the distribution of the rv $\mathbf{Z}^{(d-1)}$ is invariant under all linear rotations around the origin **0** in the whole $\mathbb{R}^d$, which means the rv has isotropic distribution.

Let $V(x_d, x_1, \ldots, x_{d-2}, x_{d-1}) = (x_1, x_2, \ldots, x_{d-1}, x_d)$ and define $U = V \circ U^{(d-1)} \circ U^{(d-2)} \circ \ldots \circ U^{(2)} \circ U^{(1)}$. Then we see all the transformations $U^{(1)}, U^{(2)}, \ldots, U^{(d-2)}, U^{(d-1)}$ are essentially based on the random variables defined in $\mathbb{I}^{d-1}$. That implies the existence of an invertible differentiable transformation $Q: \mathbb{I}^{d-1} \to \mathbb{I}^{d-1}$ such that $U = D \circ Q \circ B$. Besides, the above argument ensures that the rv $\mathbf{Y} = U(\mathbf{Z}) = U(T(\mathbf{X}))$ is an isotropic rv. This together with Lemma 2.1 and Lemma 2.2 show **Y** is a sub-Gaussian rv. Therefore, we can confirm $K = U \circ T$ is the desired transformation. ☐

Theorem 2.3 ensures $\mathbf{Y} = K(\mathbf{X}) = (A^{1/2}G_1, \ldots, A^{1/2}G_d)$ is a sub-Gaussian rv, where $\mathbf{G} = (G_1, \ldots, G_d)$ is a zero-mean Gaussian vector in $\mathbb{R}^d$ independent of *A*, with iid components $G_i, i = 1, \ldots, d$, and $A \sim S(\alpha/2; 1; [\cos(\pi\alpha/4)]^{2/\alpha}; 0)$. We get $R^2 = A(G_1^2 + \ldots G_d^2) = AH$, where *H* is chi-squared with *d* degrees of freedom, independent of *A*, and $R = \|\mathbf{Y}\| = (Y_1^2 + \ldots + Y_d^2)^{1/2}$. Then the density function of *R* (see (5) [16]) can be expressed as

$$f_R(r) = 2r\int_0^\infty f_A\left(\frac{r^2}{t}\right)\frac{f_H(t)}{t}dt, \tag{11}$$

Where f_A and f_H are density functions of random variables A and H, respectively. Further, the argument of Subsection 2.1 [16] yields the follows.

LEMMA 2.4. *The density function of* **Y** *has the following form*

$$f_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} \left(\Gamma(d/2)/\left(2\pi^{d/2}\right)\right)\|\mathbf{y}\|^{1-d} f_R(\|\mathbf{y}\|) & \mathbf{y} \neq \mathbf{0} \\ \Gamma(d/\alpha)/\left(\alpha 2^{d-1}\pi^{d/2}\Gamma(d/2)^2\right) & \mathbf{y} = \mathbf{0}, \end{cases} \tag{12}$$

with the density function f_R given in (11).

Due to $\mathbf{Y} = K(\mathbf{X}) = U(T(\mathbf{X}))$, it is clear that

$$f_{\mathbf{X}}(\mathbf{x}) = J_U(\mathbf{x}) \times J_T(\mathbf{x}) \times f_{\mathbf{Y}}(K(\mathbf{x})), \tag{13}$$

where J_U and J_T are respectively the Jacobian's of the operators U and $T = (T_1, \dots, T_d)$ defined in the proof of Theorem 2.3. Meanwhile, it is easy to see

$$J_T(x_1,\dots,x_d) = \frac{f_{X_1}(x_1)\times\dots\times f_{X_d}(x_d)}{f_{S(\alpha;0;1;0)}\left(F_{X_1}(x_1)\right)\times\dots\times f_{S(\alpha;0;1;0)}\left(F_{X_d}(x_d)\right)} \tag{14}$$

and

$$J_U(x_1,\dots,x_d) = (2\pi)^d \times f_{\Theta_{d-1}}\left(2\arctan\frac{z_d}{z_{d-1}+\sqrt{z_{d-1}^2+z_d^2}}\right) \times f_{\Theta_{d-1}^{(1)}}\left(2\arctan\frac{z_1}{z_d+\sqrt{z_d^2+z_1^2}}\right)$$

$$\times\dots\times f_{\Theta_{d-1}^{(d-2)}}\left(2\arctan\frac{z_{d-1}}{z_{d-2}+\sqrt{z_{d-2}^2+z_{d-1}^2}}\right), \tag{15}$$

where $z_k = F^{-1}_{S(\alpha;0;1;0)}(F_{X_k}(x_k))$, for $k = 1, \dots, d$, and $\Theta_{d-1}, \Theta_{d-1}^{(1)}, \dots, \Theta_{d-1}^{(d-2)}$ are defined in the proof of Theorem 2.3. The equations (12) - (15) together with Theorem 2.3 provide the following theorem.

THEOREM 2.5. *Let* **X** *be a given stable rv in* $\mathbb{R}^d$. *Suppose the skewness parameters of all marginals of* **X** *are different from* $\pm$ 1. *Then the probability density function of* **X** *can be expressed as*

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} h(\mathbf{x}) \cdot g(\mathbf{x})^{1-d} f_R(g(\mathbf{x})) \times \frac{\Gamma(d/2)}{2\pi^{d/2}} & \mathbf{x} \neq (med_1, \dots, med_d) \\ \Gamma(d/\alpha)/\left[\alpha 2^{d-1}\pi^{d/2}\Gamma(d/2)^2\right] & \mathbf{x} \neq (med_1, \dots, med_d) \end{cases} \tag{16}$$

where the density function f_R given in (11),

$$g(\mathbf{x}) = \sqrt{\left[F^{-1}_{S(\alpha;0;1;0)}\left(F_{X_1}(x_1)\right)\right]^2 + \dots + \left[F^{-1}_{S(\alpha;0;1;0)}\left(F_{X_d}(x_d)\right)\right]^2}$$

for $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d, h(\mathbf{x}) = J_U(\mathbf{x}) \times J_T(\mathbf{x})$, *with* $J_U(\mathbf{x})$ *and* $J_T(\mathbf{x})$ *expressed in* (14) *and* (15).

PROOF. It is evident in the proof of Theorem 2.3 that the rv $\mathbf{Z} = T(\mathbf{X})$ is stable with symmetric marginals $S(\alpha; 0; 1; 0)$-distributed. Besides,

$$T(med_1, \dots, med_d) = (0, \dots, 0) \tag{17}$$

Meantime, Theorem 2.3 ensures the rv $\mathbf{Y} = U(\mathbf{Z})$ to be a sub-Gaussian rv. Therefore, Lemma 2.4 provides the probability density function $f_{\mathbf{Y}}$ of **Y** expressed as (12). Simultaneously, it is clear that the transformation U does not change the amplitude of any vector $\mathbf{x} \in \mathbb{R}^d$, that confirms

$$\|\mathbf{y}\| = \|T(\mathbf{x})\| = \sqrt{\left[F^{-1}_{S(\alpha;0;1;0)}\left(F_{X_1}(x_1)\right)\right]^2 + \dots + \left[F^{-1}_{S(\alpha;0;1;0)}\left(F_{X_d}(x_d)\right)\right]^2}$$

For $\mathbf{x} = (x_1, \dots, x_d)$. This together with (13) - (15) and (17) yield (16). □

III. TEST ON THE STABLE DISTRIBUTION OF MULTIVARIATE DATA

Let $\mathbf{x} = \{x_{ij}: i = 1,2, \dots, d; j = 1,2, \dots, n\}$ be a dataset collected from a rv $\mathbf{X} = (X_1, \dots, X_d)$. The transformation K in Theorem 2.3 turns **x** into $\mathbf{y} = K(\mathbf{x}) = \{y_{ij}: i = 1,2, \dots, d; j = 1,2, \dots, n\}$, that can be considered as a dataset extracted from the rv $\mathbf{Y} = (Y_1, \dots, Y_d) = K(\mathbf{X})$. According to Theorem 2.3, the rv **X** has stable distribution with all marginal skewness parameters different from ± 1 iff **Y** is a sub-Gaussian rv. Simultaneously, in spirit of (9), it is easy to see that the rv **Y** is a sub-Gaussian rv iff its copula is the copula of sub-Gaussian rv and all its marginals have α-stable distribution with common stable index α. Then the stable distribution of **X** can be checked by conducting the following two actions.

a) Use the non-parametric test based on the Kendall functions (see [13], [20], [21]) to check the goodness-of-fit of the copula of **Y** to sub-Gaussian copula;

b) Use the Kolmogorov - Smirnov test to check the goodness-of-fit of all marginals $Y_1, \dots, Y_d$ to α_*-

stable distribution $S(\alpha_*; 0; 1; 0)$, where $\alpha_* = (\alpha_1 + \alpha_2 + \ldots + \alpha_d)/d$, and α_i is the estimated stable index of the data marginal $\mathbf{x}_i := \{x_{ij} : j = 1, \ldots, n\}$ for $i = 1, \ldots, d$.

The functions *McCullochParametersEstim* and *ks.test* in the R software package can be used to estimate the stable parameters of each data marginal $\mathbf{x}_i := \{x_{ij} : j = 1, \ldots, n\}$, $i = 1, \ldots, d$, and then to verify the hypothesis of stable distribution goodness-of-fit of all marginals $i = 1, \ldots, d$ in the task b). Meantime, the task a) can be completed by conducting the following steps.

Step 1. Generate a hypothetical dataset $\mathbf{y}^{(0)} = \{y_{ij}^{(0)} : i = 1,2, \ldots, d; j = 1,2, \ldots, n\}$ that is the dataset extracted from the rv with sub-Gaussian distribution $\mathbf{Y}^* = (Y_1^*, Y_2^* \ldots, Y_d^*) = A^{1/2}\mathbf{G}$ (see Subsection 2.2 [22]), where $\mathbf{G} \sim N(\mathbf{0}; \mathbf{I})$ is a Gaussian rv with expectation $\mathbf{0}$ and covariance matrix $\mathbf{I}$, the unit matrix of the size $d \times d$, $A \sim S(\alpha_*/2; 1; [\cos(\pi\alpha_*/4)]^{2/\alpha_*}; 0)$, A and $\mathbf{G}$ are independent.

Step 2. Let $M_k = \#\{j \neq k : y_{1j}^{(0)} < y_{1k}^{(0)}, y_{2j}^{(0)} < y_{2k}^{(0)}, \ldots, y_{dj}^{(0)} < y_{dk}^{(0)}\}/n$ for $k = 1,2,\ldots,n$. Based on the formula of the Kendall functions [23] of the rv

$$K_{\mathbf{Y}^*}(t) = \mathrm{P}\left(F_{\mathbf{Y}^*}\left(Y_1^*, Y_2^*, \ldots, Y_d^*\right) \le t\right),$$

to estimate the values of the Kendall function at M_k for the hypothetical dataset

$$K_{\mathbf{y}^{(0)}}(M_k) = \#\left\{j : \bar{F}_{\mathbf{y}^{(0)}}\left(y_{1j}^{(0)}, y_{2j}^{(0)}, \ldots, y_{dj}^{(0)}\right) \le M_k\right\}/n,$$

and for the transformed sample dataset

$$K_{\mathbf{y}}(M_k) = \#\left\{j : \bar{F}_{\mathbf{y}}\left(y_{1j}, y_{2j}, \ldots, y_{dj}\right) \le M_k\right\}/n,$$

where $\bar{F}$ denotes the empirical distribution function. To determine the test statistic

$$\mathbf{d} = \sum_{k=1}^{n}\left(K_{\mathbf{y}^{(0)}}(M_k) - K_{\mathbf{y}}(M_k)\right)^2$$

Step 3. Apply the Monte-Carlo sampling procedure by repeating Step 1 to construct 1000 hypothetical datasets $\mathbf{y}^{(m)} = \{y_{ij}^{(0)} : i = 1,2, \ldots, d; j = 1,2, \ldots, n\}, m = 1,2, \ldots, 1000$. Then repeatedly conduct Step 2 with the transformed sample dataset $\mathbf{y}$ replaced by each of the new obtained hypothetical datasets to get the values $\mathbf{d}_m = \sum_{k=1}^{n}(K_{\mathbf{y}^{(0)}}(M_k) - K_{\mathbf{y}^{(m)}}(M_k))^2$. With a given probability value $p \in (0; 1)$, the critical value L_p of test for the hypothesis H "**Y** has sub-Gaussian copula" is taken equal to the $(1 - p)$-quantile of the set $\{d_m, m = 1,2, \ldots, 1000\}$. Reject H if $\mathbf{d} \ge L_p$ and accept H if $\mathbf{d} < L_p$.

In the following we present two examples of the application of the above goodness-of-fit test procedure to examine the multivariate stable distribution of the daily return data of Nasdaq Finance. The daily return data from the 10 stocks FA (Facebook); AMC (AMC Entertainment Holdings); AXP (American Express); NFLX (Netflix); ZM (Zoom); JNJ (Johnson & Johnson); XOM (Exxon Mobil Corporation Common); FB (Meta Platforms); HD (Home Depot); and PPG (PPG Industries), contain a sample from 22/4/2019 to 31/12/2020 to imply 430 observations. Continuously compounded percentage returns are considered, i.e. daily returns are measured by the log-differences of closing pricing multiplied by 100.

Example 1. With the 5-dimensional data of the stocks NFLX (X_1); ZM (X_2); AMC (X_3); AXP (X_4); and FA (X_5), the function *McCullochParametersEstim* in the R software package is used to estimate the stable parameters of each coordinate, giving the results in Table I.

TABLE I: STABLE PARAMETERS OF NFLX, ZM, AMC, AXP, AND FA

Coordinate	α	α_*	β	γ	δ	p-value
X_1 (NFLX)	1.771	1.478	-0.149	1.550	0.006	0.5699
X_2 (ZM)	1.461	1.478	0.000	2.090	-0.169	0.3648
X_3 (AMC)	1.392	1.478	0.046	2.719	-0.264	0.6843
X_4 (AXP)	1.388	1.478	-0.152	1.128	0.029	0.6843
X_5 (FA)	1.378	1.478	0.134	1.229	1.229	0.5699

The R function *ks.test* is used to verify the hypotheses of stable distribution goodness-of-fit $X_i \sim S(\alpha_*; \beta_i; \gamma_i; \delta_i)$, $i = 1,\ldots,5$, with $\alpha_* = (\alpha_1 + \ldots + \alpha_5)/5 = 1.478$, giving the p-values 0.5699; 0.3648; 0.6843; 0.6843; and 0.5699, indicate all coordinate data fit to the stable distributions with the corresponding parameters.

Then we conduct Step 1 followed by Step 2 and Step 3 to get the test statistic **d** = 26.4866 and the test critical values 87.6010; 126.6208; 188.3399; 236.3742; and 302.8049, corresponding to the significance levels 0.10; 0.05; 0.01; 0.005; and 0.001, respectively. Because the test statistic **d** = 26.4866 is smaller than the critical values, the hypothesis stating that the transformed rv **Y** = *K*(**X**) has sub-Gaussian copula is accepted. Simultaneously, the R function *ks.test* gives the p-values 0.4115; 0.3648; 0.8900; 0.8455; and

0.4115, showing the coordinate y_1, y_2, y_3, y_4 and y_5 of the transformed data **y** = K(**x**) fit to the stable distribution $S(1.478;0;1;0)$. The above two facts confirm **Y** is a sub-Gaussian rv. Consequently, Theorem 2.3 allows to conclude that the returns' dataset of NFLX, ZM, AMC, AXP, and FA fits to the 5-dimensional 1.478-stable distribution.

Example 2. Regarding the 5-dimensional data of the stocks JNJ; XOM; FB; HD; and PPG, we proceed with the same procedure as that of Example 1. The function *McCullochParametersEstim* gives the estimated stable parameters in Table II.

TABLE II: STABLE PARAMETERS OF THE MARGINALS JNJ, XOM, FB, HD, PPG

Coordinate	α	α_*	β	γ	δ	p-value
X_1 (JNJ)	1.385	1.4444	-0.313	0.642	0.098	0.2462
X_2 (XOM)	1.329	1.4444	0.061	1.146	-0.129	0.1847
X_3 (FB)	1.580	1.4444	-0.264	1.261	0.127	0.2821
X_4 (HD)	1.465	1.4444	-0.150	0.808	0.110	0.6843
X_5 (PPG)	1.463	1.4444	-0.003	0.953	0.083	0.8900

The R function *ks.test* is used to verify the hypotheses of stable distribution goodness-of-fit $X_i \sim S(\alpha_*; \beta_i; \gamma_i; \delta_i)$, $i = 1,...,5$, with $\alpha_* = (\alpha_1 + \ldots + \alpha_5)/5 = 1.4444$, giving the p-values 0.2462; 0.1847; 0.2821; 0.6843; and 0.8900, confirming the α_*-stability of all marginal distributions of the concerned data.

Then Step 1, Step 2, and Step 3 of the proposed procedure are sequentially proceeded to provide the test statistic **d** = 342.6804 and the test critical values 115.2128; 131.7771; 168.7637; 173.2055; and 186.3036 of the test on sub-Gaussian 1.4444-stable copula in $\sim^{5}$, corresponding to the significance levels 0.10; 0.05; 0.01; 0.005; and 0.001, respectively. Since the test statistic **d** is greater than all the critical values, we can realize that the hypothesis stating that the transformed rv **Y** = K(**X**) has sub-Gaussian copula is not accepted. This rejects the sub-Gaussian distribution of the transformed rv **Y**. Therefore, Theorem 2.3 yields the conclusion that the returns' dataset of JNJ, XOM, FB, HD, and PPG does not fit to 5-dimensional stable distribution, although all its marginals have stable distribution.

IV. Conclusion and Discussion

The results obtained in this study provide useful tools to investigate thoroughly multivariate data those have stable distribution. The bijective transformation in multidimensional space created on the base of Theorem 2.3, that turns a given stable random vector into a sub-Gaussian random vector, provides a theoretical basis of the testing procedure on the stable distribution of multivariate data. Examples of datasets collected from the Nasdaq stock market are used to illustrate the practicability of the proposed testing procedure. Theorem 2.5, an immediate consequence of Theorem 2.3, represents a method of probability density function estimation for multivariate data with stable distribution.

Let **X** be a random vector, its symmetrization is defined as **X*** = **X'** – **X''**, where **X'** and **X''** are two independent copies of **X**. It is clear that if **X** is a stable random vector then **X*** is also a stable random vector with all marginal skewness parameters equal to 0. Consequently, the rejection of stable distribution hypothesis for **X*** implies the rejection of stable distribution hypothesis for **X**. Therefore, although Theorem 2.3 is valid only for stable random vector with marginal's having all skewness parameters different from ± 1, the proposed stability testing procedure can be partially extended to the class of all stable distribution.

References

[1] Markowitz HM. Portfolio selection, *The Journal of Finance* 1952; 7 (1): 77-91.

[2] Ross S. The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 1976; 13 (3): 341-360.

[3] Sharpe WF. Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 1964; 19 (3): 425-442.

[4] Fama E. The behavior of stock prices, *Journal of Business* 1965; 38: 34-105.

[5] Samuelson P. Efficient portfolio selection for Pareto - Lévy investments, *Journal of Financial and Quantitative Analysis* 1967; 2: 107-117.

[6] Kunst RM. Apparently stable increments in finance data: Could ARCH effects be the cause?, *J. Statist. Comput. Simulation* 1993; 45: 121-127.

[7] McCulloch JH. Financial applications of stable distributions. In: *Handbook of Statistics* 14. Maddala G, Rao C. Eds. North-Holland: Elsevier Science Publishers; 1996, 393-425.

[8] Taqqu MS. The modeling of Ethernet data and of signals that are heavy-tailed with infinite variance, *Scand. J. Statist.* 2002; 829: 273-295.

[9] Palmer KJ, Ridout MS, Morgan BJT. Modeling cell generation times using the tempered stable distribution, *Journal of the Royal Statistical Society* 2008; 57, Series C, 379-397.

[10] Adler RJ, Feldman RE, Taqqu MS. *A Practical Guide to Heavy Tailed Data*. Boston: Birkhauser; 1998.

[11] Lamantia F, Ortobelli S, Rachev S. VaR, CVaR and Time Rules with Elliptical and Asymmetric Stable Distributed Returns, *Investment Management and Financial Innovations* 2006; 3(4): 19-39.

[12] Resnik SI. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer, 2007.
[13] Majoros S, Zempléni A. Multivariate stable distributions and their application in modeling returns. Stat. AP 2018, arXiv:1810.09521.
[14] McCulloch JH. Simple consistent estimators of stable distribution parameters, *Commun. Statist. Simul.* 1986; 15: 1109-1136.
[15] Kogon SM, Williams DB. Characteristic function based estimation of stable parameters. In: *A Practical Guide to Heavy Tailed Data*. Adler R, Feldman R, Taqqu M Eds. Boston, MA: Birkhauser; 1998, 311-335.
[16] Nolan JP. *Stable Distributions - Models for Heavy Tailed Data*. Boston, MA, USA: Birkhauser, 2016.
[17] Samorodnitsky G, Taqqu MS. *Stable Non-Gaussian Random Processes*, New York, NY: Chapman & Hall, 1994.
[18] Sklar A. Fonctions de repartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Universitè de Paris* 1959; 8: 229-231.
[19] Embrechts P, Lindskog F, McNeil A. Modeling Dependence with Copulas and Applications to Risk Management. In: *Handbook of Heavy Tailed Distributions in Finance*. Rachev S Ed. Elsevier, 2003, Chapter 8, pp. 329--84.
[20] Genest C, Rivest LP. Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.* 1993; 88: 1034-1043.
[21] Genest C, Quessy JF, Rémillard B. Goodness-of-fit procedures for copula models based on the probability integral transformation, *Scand. J. Statist.* 2006; 33: 337-66.
[22] Omelchenko V. Parameter estimation of sub-Gaussian stable distributions, *Kybernetika* 2014; 50 (6): 929—949.
[23] Nelsena RB, Quesada-Molina JJ, Rodriguez-Lallena JA, Ubeda-Flores M. Kendall distribution functions, *Statistics & Probability Letters* 2003; 65 (3): 263-268.