Bayesian analysis of multivariate stable distributions using one-dimensional projections

Mike G. Tsionas*

Abstract

In this paper we take up Bayesian inference in general multivariate stable distributions. We exploit the representation of Matsui and Takemura (2009) for univariate projections, and the representation of the distributions in terms of their spectral measure. We present efficient MCMC schemes to perform the computations when the spectral measure is approximated discretely or, as we propose, by a normal distribution. Appropriate latent variables are introduced to implement MCMC. In relation to the discrete approximation, we propose efficient computational schemes based on the characteristic function.

Key words: Multivariate stable distributions; spectral measure; Markov Chain Monte Carlo; Bayesian inference.

Acknowledgments: The paper benefited from the comments of two anonymous referees. Additional technical material is presented in the working paper version of the study, arXiv: http://arxiv.org/abs/1507.07323

*Economics Department, Lancaster University Management School, LA1 4YX UK, Tel: +44 (0)1524 592668, email: m.tsionas@lancaster.ac.uk

1 Introduction

Univariate stable distributions have been thoroughly studied in econometrics, statistics and finance over the past few decades (Samorodnitsky and Taqqu, 1994). Their empirical application is still hampered by the fact that their density is not available in analytic form, despite advances in Bayesian computation using MCMC. Buckle (1995) and Tsionas (1999) provided Gibbs sampling schemes for general and symmetric stable distributions, respectively. The problem is that the conditional posterior distributions of certain latent variables are cumbersome to work with and require careful tuning. The analogous problem in the multivariate case is exceedingly difficult although a few attempts have been made to solve it. The impediment is that multivariate stable distributions, unlike the univariate case, are defined through their spectral measures which, in practice, are unknown. Ravishanker and Qiou (1999) for example, proposed an EM algorithm based on Buckle (1995) in the case of symmetric isotropic stable distributions but this class is too narrow to be of empirical importance. It is defined by the transformation $X = \mu + C\xi$, where ξ is a vector of independent random variables each one distributed as standard symmetric stable, μ is a vector of location parameters, Σ is a scale matrix, and $C^\top C = \Sigma$. It is known that the class of elliptical stable distributions can be defined through the transformation $X = \mu + RCu$ where u is uniformly distributed on the unit sphere $\mathbb{S}^{d-1} = \{\boldsymbol{x} \in \mathbb{R}^d | \, ||\boldsymbol{x}|| = 1\}$, C is a $d \times d$ scale matrix of full rank, and $R = \sqrt{VS_{\alpha/2}}$ where, independently, $V \sim \chi^2_d$ and $S_{\alpha/2}$ follows a stable distribution with parameter $\alpha/2$ and maximal skewness $\beta = 1$. Of course not all multivariate stable distributions are elliptical. See Lombardi and Veredas (2009). When $V \sim \chi^2_1$ the distribution of X is in the class of elliptically contoured stable distributions (Nolan, 2006, p.2).

In connection with multivariate stable Paretian distributions, even the computation of the characteristic functions becomes complicated because they are only defined through their spectral measure, an object that is needed in order to retain the equivalence between the density and the characteristic function. The estimation of the spectral measure itself has proved itself to be quite cumbersome even for bivariate distributions (see the seminal works of McCulloch, 1994, 2000, Nolan et al., 2001, and Nolan and Rajput, 1995).

The present paper is related to recent advances in the econometrics of stable distributions. Dominicy and Veredas (2013) propose a method of quantiles to fit symmetric stable distributions. Since the quantiles are not available in closed form they are obtained using simulation resulting in the method of simulated quantiles or MSQ. Hallin, Swan, Verdebout and Veredas (2013) propose an easy-to-implement R-estimation procedure which remains -consistent contrary to least squares with stable disturbances. Broda, Haas, Krause, Paolella and Steude (2013)

propose a new stable mixture GARCH model that encompasses several alternatives and can be extended easily to the multivariate asset returns case using independent components analysis. Ogata (2013) uses a discrete approximation to the spectral measure of multivariate stable distributions and proposes estimating the parameters by equating the theoretical and empirical characteristic function in a generalized empirical likelihood / GMM framework.

Relative to this work, we show how to implement Bayesian inference for multivariate stable distributions by providing statistical inferences about the spectral measure jointly with the other parameters of the model. For numerical analysis via MCMC we employ a novel data augmentation technique for stable distributions. We use a discrete approximation of the measure where the configuration and the number of points are unknown. We also propose a novel approximation to the spectral measure based on a multivariate normal distribution.

2 Stable distributions

A random variable X is called strictly (univariate) stable if for all n, $\sum_{i=1}^{n} X_i \sim c_n X$ for some constant c_n, where $X_1, ..., X_n$ are independently distributed with the same distribution as X. It is known that the only possible choice is to have $c_n = n^{1/\alpha}$ for some $\alpha \in (0, 2]$. General non-symmetric stable distributions are defined via the log characteristic function which is given by the following expression (Samorodnitsky and Taqqu, 1994, and Zolotarev, 1986):

$$\log \varphi(\tau) = \log E \exp(\iota \tau X) =$$
$$\begin{cases} \iota\mu\tau - |\sigma\tau|^{\alpha}\{1 - \iota\beta \mathrm{sgn}(\tau) \tan\frac{\pi\alpha}{2}\}, \ \alpha \neq 1 \\ \iota\mu\tau - \sigma|\tau|\{1 + \iota\beta \mathrm{sgn}(\tau)\frac{2}{\pi}\log|\tau|\}, \ \alpha = 1, \end{cases} \tag{1}$$

where $\tau \in \mathbb{R}$, μ and σ are location and scale parameters, α is the characteristic exponent, $\beta \in [-1, 1]$ is the skewness parameter, and $\iota = \sqrt{-1}$. In this paper we are interested in multivariate stable distributions, that is distributions of a random variable in $\mathbb{R}^d$. Suppose X is a vector of random variables with characteristic exponent $\alpha \in (0, 2]$. Its characteristic function is $\varphi_X(\boldsymbol{\tau}) = E \exp\{\iota < \boldsymbol{\tau}, X >) = \exp\left(-I_X(\boldsymbol{\tau}) + \iota < \tau, \mu >\right)$ where $< \boldsymbol{\tau}, X >= \boldsymbol{\tau}^{\top} X$ denotes inner product, and

$$I_X(\boldsymbol{\tau}) = \int_{\mathbb{S}^{d-1}} \psi_{\alpha}\left(< \boldsymbol{\tau}, \boldsymbol{s} >\right) \Gamma(d\boldsymbol{s}), \tag{2}$$

where $\mathbb{S}^{d-1} = \left\{ \boldsymbol{u} \in \mathbb{R}^d | < \boldsymbol{u}, \boldsymbol{u} >= 1 \right\}$ is the boundary of the unit ball in $\mathbb{R}^d$, Γ is a finite Borel measure of the vector X, called the spectral measure, $\mu \in \mathbb{R}^d$ is a vector of location parameters, and the complex function ψ is defined as follows:

$$\psi_\alpha(u) = \begin{cases} |u|^\alpha \{1 - \iota \text{sgn}(u) \tan \frac{\pi\alpha}{2}\}, & \alpha \neq 1, \\ |u| \{1 + \iota \frac{2}{\pi} \text{sgn}(u) \log |u|\}, & \alpha = 1 \cdot \end{cases} \tag{3}$$

See seminal work by Nolan (1998), Nolan and Rajput (1995), Abdul-Hamid and Nolan (1998), and also Cambanis and Miller (1981), and Nagaev (2000). Notably the parameters (α, Γ) fully define all centered multivariate stable distributions, and a skewness parameter β is not needed[1] in this case, since we have the full measure, Γ. We denote the class by $X \sim \mathscr{S}_{\alpha,d}(\mu, \Gamma)$. Press (1972) attempted to define a multivariate α-stable distribution without using the spectral measure Γ. Later on Paulauskas (1976) provided some corrections as not all α-stable distributions can be represented using Press' (1972) characteristic function. Chen and Rachev (1995) is an interesting paper where the authors provided estimates of the spectral measure as well as applications to a stable portfolio. It is notable that the projection of X on τ, viz. $< \boldsymbol{\tau}, X >$ has a univariate stable distribution. The characteristic exponent α remains the same but scale, location and skewness depend on $\boldsymbol{\tau}$.The multivariate characteristic function is not easy to work with as in the univariate case because of the dependence on the spectral measure. As this can rarely be specified in advance, it is necessary to provide posterior inferences about it, in the context of Bayesian analysis.

One approach (Byczkowski et al., 1993) is to assume that Γ can be approximated by a discrete measure, in which case we have:

$$\Gamma(d\boldsymbol{s}) = \sum_{j=1}^{J} \gamma_j \delta_{\{\boldsymbol{s}^{(j)}\}}(d\boldsymbol{s}), \tag{4}$$

where $\gamma_j > 0$, $\boldsymbol{s}^{(j)} \in \mathbb{S}^d, j = 1, ..., J$, δ denotes Dirac's function and the approximation is made at J points of the unit sphere in $\mathbb{R}^d$. The meaning of (4) is that we define a finite partition $A_1, ..., A_J$ of $\mathbb{S}^{d-1}$, points $\boldsymbol{s}^{(1)}, ..., \boldsymbol{s}^{(J)} \in \mathbb{S}^{d-1}$ and construct Γ by placing mass $\Gamma(A_j)$ at $\boldsymbol{s}^{(j)}$ so that:

$$\Gamma(d\boldsymbol{s}) = \sum_{j=1}^{J} \Gamma(A_j) \delta_{\{\boldsymbol{s}^{(j)}\}}(d\boldsymbol{s}).$$

[1] Actually, there are skewness parameters $\beta(\boldsymbol{\tau})$ that depend on the particular projection $\boldsymbol{\tau}$.

Suppose θ denotes the parameters. Since $\varphi_X(\boldsymbol{\tau};\theta) = \exp\left\{-\int_{\mathbb{S}^d} \psi_\alpha(<\boldsymbol{\tau},\boldsymbol{s}>)\Gamma(d\boldsymbol{s})\right\}$ we obtain:

$$\varphi_X(\boldsymbol{\tau};\theta) = \exp\left(-\sum_{j=1}^{J}\gamma_j\psi_\alpha(<\boldsymbol{\tau},\boldsymbol{s}^{(j)}>)\right), \tag{5}$$

or alternatively:

$$-\log\varphi_X(\boldsymbol{\tau}^{(i)},\boldsymbol{s}^J;\theta) = \sum_{j=1}^{J}\gamma_j\psi_\alpha(<\boldsymbol{\tau},\boldsymbol{s}^{(j)}>), i = 1,...,J, \tag{6}$$

where $\boldsymbol{\tau}^K = (\boldsymbol{\tau}^{(1)},...,\boldsymbol{\tau}^{(J)})$ denotes a set of points where the log-characteristic function is evaluated and $\boldsymbol{s}^J = (\boldsymbol{s}^{(1)},...,\boldsymbol{s}^{(J)})$. If we define:

$$\mathbb{Y}(\boldsymbol{\tau}^K,\boldsymbol{s}^J) = [-\log\varphi_X(\boldsymbol{\tau}^{(1)},\boldsymbol{s}^J;\theta),...,-\log\varphi_X(\boldsymbol{\tau}^{(K)},\boldsymbol{s}^J;\theta)]' =$$

$$[I_X(\boldsymbol{\tau}^{(1)},\boldsymbol{s}^J;\theta),...,I_X(\boldsymbol{\tau}^{(K)},\boldsymbol{s}^J;\theta)]'$$

$$\mathbb{X}(\boldsymbol{\tau}^K,\boldsymbol{s}^J) = [\mathscr{X}_{ij}(\boldsymbol{\tau}^{(i)},\boldsymbol{s}^{(j)})]$$

$$\mathscr{X}_{ij}(\boldsymbol{\tau}^{(i)},\boldsymbol{s}^{(j)}) = \psi_\alpha(<\boldsymbol{\tau}^{(i)},\boldsymbol{s}^{(j)}>), i = 1,...,K, j = 1,...,J,$$

we can write (6) as a system of linear equations[2]:

$$\mathbb{Y}(\boldsymbol{\tau}^K,\boldsymbol{s}^J) = \mathbb{X}(\boldsymbol{\tau}^K,\boldsymbol{s}^J)\gamma, \tag{7}$$

from which, in principle, we can obtain approximations to the spectral weights, γ_j which, in (7), we collect in vector γ. In practice, the system of equations suffers from singularities and the estimates of γ are not always non-negative. The reason that $\mathbb{X}^\top\mathbb{X}$ is often singular is that when a full grid is used, we encounter points where $I_X(-\boldsymbol{\tau}) = \overline{I_X(\boldsymbol{\tau})}$. Nolan, Panorska and McCulloch (2001) propose such symmetric grids around the basic directions (1,0), (0,1), (-1,0), and (0,-1) corresponding to independent components in the bivariate case. Generally we would need $J \propto 2^d$ in the $d-$dimensional case *if we need* to explore the measure around coordinates corresponding to independent components. This is manageable in dimensions up to 10. Of course the possibility arises that we *may* actually need a value of J that is *significantly lower* since the spectral measure can be embedded in a much smaller subspace. McCulloch (1994, 2000) has proposed the use of quadratic programming imposing the non-negativity and Nolan, Panorska and McCulloch (1997) report that, at least in small dimensions, the procedure

[2]The system is in general complex-valued so we need to take the real and imaginary parts of $\mathbb{Y}$ and $\mathbb{X}$.

works well. The problem is challenging in that a double grid has to be specified, $\boldsymbol{\tau}^K$ for the set of points to evaluate the log-characteristic function and $\boldsymbol{s}^J$ for the support of the discrete measure. Apparently in all but very low dimensions ($d = 2$ specifically) if we were to use a full grid in (7) we would face the curse of dimensionality as matrix $\mathbb{X}$ would be huge. Therefore, the only choice is to place the points $\boldsymbol{\tau}^K, \boldsymbol{s}^J$ in a "wise" manner without sacrificing computational ease.

3 A hierarchical model for multivariate stable distributions

Matsui and Takemura (2009) extended the work of Abdul-Hamid and Nolan (1998) and provided semi-closed expressions for the class of multivariate stable distributions. Abdul-Hamid and Nolan (1998) used higher order derivatives of one-dimensional densities. The role of the following function is important:

$$g_{\alpha,d}(v,\beta) = \begin{cases} (2\pi)^{-d} \int_0^\infty \cos\left(vu - \left\{\beta \tan\frac{\pi\alpha}{2}\right\} u^\alpha\right) u^{d-1} \exp(-u^\alpha) du, \ \alpha \neq 1, \\ (2\pi)^{-d} \int_0^\infty \cos\left(vu + \frac{2}{\pi}\beta u \log u\right) u^{d-1} \exp(-u) du, \ \alpha = 1. \end{cases} \tag{8}$$

By theorem 1.1 in Matsui and Takemura (2009) due to Theorem 1 in Nolan (1998), the density of $X \sim \mathscr{S}_{\alpha,d}(\zeta, \Gamma)$, where $\zeta \in \mathbb{R}^d$ is a shift parameter, can be expressed as:

$$f_{\alpha,d}(x) = \begin{cases} \int_{\mathbb{S}^{d-1}} g_{\alpha,d}\left(\frac{<x-\zeta,\boldsymbol{S}>}{\sigma(\boldsymbol{S})}, \beta(\boldsymbol{S})\right) \sigma(\boldsymbol{S})^{-d} d\boldsymbol{S}, \ \alpha \neq 1, \\ \int_{\mathbb{S}^{d-1}} g_{1,d}\left(\frac{<x-\zeta,\boldsymbol{S}> - \mu(\boldsymbol{S}) - \frac{2}{\pi}\beta(\boldsymbol{S})\sigma(\boldsymbol{S})\log\sigma(\boldsymbol{S})}{\sigma(\boldsymbol{S})}, \beta(\boldsymbol{S})\right) \sigma(\boldsymbol{S})^{-d} d\boldsymbol{S}, \ \alpha = 1. \end{cases} \tag{9}$$

Here, the parameters

$$\sigma(\boldsymbol{S}) = \left(\int_{\mathbb{S}^{d-1}} |<\boldsymbol{t}, \boldsymbol{S}>|^\alpha \Gamma(d\boldsymbol{t})\right)^{1/\alpha}, \tag{10}$$

$$\beta(\boldsymbol{S}) = \sigma(\boldsymbol{S})^{-\alpha} \int_{\mathbb{S}^{d-1}} \operatorname{sgn}(<\boldsymbol{t}, \boldsymbol{S}>) \cdot |<\boldsymbol{t}, \boldsymbol{S}>|^\alpha \Gamma(d\boldsymbol{t}), \tag{11}$$

and in these expressions $\boldsymbol{t} \in \mathbb{R}^d$ is the angle, see Samorodnitsky and Taqqu (1994, example 2.3.4) and Matsui and Takemura (2009). Notably, the location or shift parameter is zero when $\alpha \neq 1$ and a certain constant independent of ζ when $\alpha = 1$. We can certainly assume $\zeta = 0$. The role of ζ arises mainly when the distribution is symmetric around ζ, see Corollary 4 in Abdul-Hamid and Nolan (1998). For the definition of location when $\alpha = 1$ see equation (6) in Matsui and Takemura (2009) or equation (9) in Abdul-Hamid and Nolan (1998). The advantage of the expressions is that function $g_{\alpha,d}$ "is a function of two real variables no matter what the dimension d is, and that this function is the same for every α-stable r.v. X; i.e., it is independent of the spectral measure" (Abdul-Hamid and Nolan, 1998).

It is possible to represent the density in terms of functions similar to those used by Zolotarev (1986, equation 2.2.18, p.74) to convert the range of integration to a finite interval (whose upper bound is one) and avoid the infinite oscillations caused by the trigonometric terms (Buckle, 1995). It is clear, however, that latent variables $\boldsymbol{s}$ can be defined so that when $X_t \sim \text{iid} \mathscr{S}_{\alpha,d}(\zeta, \Gamma)$, $t = 1, ..., n$, then from (8) and (9) we have:

$$p\left(X_t | \boldsymbol{S}_t, \theta\right) \propto g_{\alpha,d}\left(v_t; \beta(\boldsymbol{S}_t)\right), t = 1, ..., n, \tag{12}$$

$$v_t \equiv \frac{< X_t - \zeta, \boldsymbol{S}_t >}{\sigma(\boldsymbol{S}_t)}, \tag{13}$$

$$\boldsymbol{S}_t \sim \mathscr{U}\left(\mathbb{S}^{d-1}\right), t = 1, ..., n, \tag{14}$$

and $\mathscr{U}(\mathbb{S}^{d-1})$ denotes the uniform distribution over the unit sphere. We remark that:

$$\sigma(\boldsymbol{S}_t) = \left(\int_{\mathbb{S}^{d-1}} | < \boldsymbol{S}_t, \boldsymbol{c} > |^{\alpha} \Gamma(d\boldsymbol{c}) \right)^{1/\alpha}, \tag{15}$$

$$\beta(\boldsymbol{S}_t) = \sigma(\boldsymbol{S}_t)^{-\alpha} \int_{\mathbb{S}^{d-1}} \text{sgn}(< \boldsymbol{S}_t, \boldsymbol{c} >) | < \boldsymbol{S}_t, \boldsymbol{c} > |^{\alpha} \Gamma(d\boldsymbol{c}) \tag{16}$$

The representation in (10)-(12) is a hierarchical model for a general multivariate stable distribution. Of course, the hierarchical model involves the spectral measure, $\Gamma(d\boldsymbol{S})$, through the functions $\beta(\boldsymbol{S})$ and $\sigma(\boldsymbol{S})$. These functions can be computed either via direct integration over the unit sphere in $\mathbb{R}^d$ or by simulation (when $\boldsymbol{S} \sim \Gamma$ in their

computation). It is possible to use further augmentation of the model by latent variables which is, however, not recommended since it will affect seriously the mixing properties of MCMC. See for example the integrals over a bounded interval used by Matsui and Takemura (2009) to obtain the $g_{\alpha,d}$ function. These correspond to similar expressions in Zolotarev (1986, pp. 76-77) and Buckle (1995). Although numerical integration is facilitated it does not avoid the problem of the presence of the spectral measure, Γ. Eventually, the density (9) has to be evaluated using v_t in (14) and as Matsui and Takemura (2009) note the singularities are not completely avoided and further work is needed to deal with them as in section 3 of Nolan (1997) or the work reported here.

Computation of (8) is a subtle matter. We proceed as follows. Since the cosine function $\cos(x) = 0$ at $x = \frac{k\pi}{2}$, for $k = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, ...$ we locate the roots of the equation $vu - \left\{\beta(\boldsymbol{s}_t) \tan \frac{\pi\alpha}{2}\right\} u^\alpha = \frac{k\pi}{2}$, for $k = 1, \frac{3}{2}, 2, ...$ with respect to u. Denote these roots by $r_1, r_2, ...$ The largest value of k is determined by the weight factor in (8), viz. $f(u) = u^{d-1}\exp(-u^\alpha)du$ so that the value of the weight factor at a certain $\bar{u}$ is less than $\varepsilon = 10^{-7}$ relative to its mode. The weight factor can be modified using the transformation $w = u^\alpha$ in which case w follows a standard *gamma* distribution with shape parameter $\frac{d}{\alpha}$. The mode is $m = \left(\frac{d}{\alpha} - 1\right)^{1/\alpha}$ so in practice we can set $\bar{u} = Cm$ and determine the constant C so that the weight factor is sufficiently small. Therefore the integral in (15) is computed in the intervals $[0, r_1]$, $[r_1, r_2]$, etc as follows:

$$\begin{gathered} g_{\alpha,d}(v, \beta(\boldsymbol{S}_t)) = \\ (2\pi)^{-d} \sum_{k=1}^{I} \int_{r_{k-1}}^{r_k} \cos\left(vw^{1/\alpha} - \left\{\beta(\boldsymbol{S}_t) \tan \frac{\pi\alpha}{2}\right\} w\right) w^{\frac{d}{\alpha}-1} \exp(-w) dw \end{gathered} \tag{17}$$

after a change of variables to $w = u^\alpha$, where $r_0 = 0$ and $r_I > \bar{w} = 10\left(\frac{d}{\alpha} - 1\right)$. This is always possible since the cosine function has an infinite number of roots. We use 20-point Gaussian quadrature to compute the integrals in the intervals $[r_{k-1}, r_k]$ determined by the roots of $\chi_k(u) = vu - \left\{\beta(\boldsymbol{s}_t) \tan \frac{\pi\alpha}{2}\right\} u^\alpha - \frac{k\pi}{2} = 0$. An alternative stopping criterion we used is when the roots $r_I - r_{I-1} < \epsilon = 10^{-4}$ so that the contribution to the integral in (8) is trivial. We have found that locating the roots is extremely easy provided we can locate the root of $\chi_1(u) = 0$ which requires a good starting value. Then the root r_k is an excellent starting value to locate r_{k+1} using a standard Newton algorithm with analytical derivatives. The root r_1, viz. $\chi_1(r_1) = 0$ can be located using bisection. Moreover, Gaussian quadrature was found to work well. This procedure takes account of the oscillations of $g_{\alpha,d}(v, \beta(\boldsymbol{s}_t))$ and is quite efficient in computing its values. In practice, too many zero points are needed and, therefore, it is better to utilize the finite integral representation in Matsui and Takemura (2009) and Abdul-Hamid and Nolan (1986). Our results in artificial and actual data were compared with 30- and 40-point quadrature and we found no essential

differences. We were not able to find different results when the finite integral representation mentioned above was used. However, the form of the integrand suggests an adaptive quadrature scheme in a finite interval, until specified tolerance is achieved. This method is comparable in terms of accuracy, but less efficient compared to Matsui and Takemura (2009) and Abdul-Hamid and Nolan (1986).

Here, *we propose to approximate the spectral measure by*

$$\boldsymbol{c}_t^* \sim \mathscr{N}_d(\mu \cdot 1_d, \omega^2 I),\ \boldsymbol{c}_t = \frac{\boldsymbol{c}_t^*}{|\boldsymbol{c}_t^*|},$$

viz. *a multivariate normal over the unit sphere*, where μ and ω are unknown parameters and 1_d is the unit vector in $\mathbb{R}^d$. The augmented posterior of the model conditionally on Γ can be written in the form:

$$\begin{gathered} p(\alpha, \{\boldsymbol{S}_t\}, \zeta | \Gamma, X) \propto \\ \prod_{t=1}^n g_{\alpha,d}\left(\frac{<X_t - \zeta, \boldsymbol{S}_t>}{\sigma_{\alpha,\Gamma}(\boldsymbol{S}_t)}, \beta_{\alpha,\Gamma}(\boldsymbol{S}_t)\right) \sigma_{\alpha,\Gamma}(\boldsymbol{S}_t)^{-d} \cdot \mathbb{I}(\boldsymbol{S}_t \in \mathbb{S}^{d-1}) \cdot p(\alpha, \zeta), \end{gathered} \tag{18}$$

where $p(\alpha, \zeta)$ denotes the prior and the notation $\sigma_{\alpha,\Gamma}(\boldsymbol{S}_t), \beta_{\alpha,\Gamma}(\boldsymbol{S}_t)$ makes explicit the dependence on α and the measure Γ. The variables $\boldsymbol{S}_t$ are treated as latent variables with a uniform prior over $\mathbb{S}^{d-1}$ and, provided the measure Γ is approximated with a normal distribution the only unknowns parameter are μ and ω for which we adopt a prior of the form: $p(\mu, \omega) \propto \omega^{-1}$. Then, from (13) and (14) we have:

$$\sigma_{\alpha,\Gamma}(\boldsymbol{S}_t)^\alpha = E_{\boldsymbol{c}\sim\Gamma} | < \boldsymbol{S}_t, \boldsymbol{c} > |^\alpha. \tag{19}$$

We notice that:

$$\beta_{\alpha,\Gamma}(\boldsymbol{S}_t) = \begin{cases} \frac{-\iota I_X(\boldsymbol{S}_t)}{\sigma(\boldsymbol{S}_t)^\alpha \tan\frac{\pi\alpha}{2}}, \alpha \neq 1, \\ \frac{\iota\{I_X(2\boldsymbol{S}_t) - 2I_X(\boldsymbol{S}_t)\}}{4\sigma(\boldsymbol{S}_t)\ln\frac{2}{\pi}}, \alpha = 1. \end{cases} \tag{20}$$

4 MCMC scheme

The great advantage of (16) is that the measure enters implicitly through the functions $\sigma_{\alpha,\Gamma}(\boldsymbol{S}_t), \beta_{\alpha,\Gamma}(\boldsymbol{S}_t)$ only. With the Discrete Approximation of the spectral measure, the computation of functions in (17) and (18) is trivial. When the measure Γ is approximated with a normal distribution the only unknown parameters are μ and ω for

which we adopt a prior of the form: $p(\mu,\omega) \propto \omega^{-1}$. The expectations can be approximated as follows:

$$\sigma_{\alpha,\omega}(\boldsymbol{S}_t)^\alpha = M^{-1}\sum_{m=1}^{M} |<\boldsymbol{S}_t, \boldsymbol{c}_m>|^\alpha, \tag{21}$$

where $\boldsymbol{c}_m \sim \text{iid}\mathcal{N}(\mu \cdot 1_d, \omega^2 I_d), i = 1, ..., M$. The expectation in (19) can be computed easily once we have (4) or the solution to (7) and $\beta_{\alpha,\omega}(\boldsymbol{S}_t)$ can be obtained from (18).

For the Discrete Approximation denote the parameters by $\theta = (\alpha, \zeta, \gamma)$. For the Normal Approximation the parameter vector is $\theta = (\alpha, \zeta, \mu, \omega)$ and we also have the latent variables $\{\mathbf{s}_t, t = 1, ..., n\}$ which are absent from the discrete approximation. In the Discrete Approximation, our prior has the form:

$$\gamma|\phi \sim \mathcal{N}_J(0, \tfrac{\phi^2}{\varpi} I),$$

subject to the constraints that $\gamma \geq 0$ where $\varpi > 0$ is a constant and ϕ is defined below. To proceed, we write (7) in the following form:

$$\mathbb{Y}(\bar{\boldsymbol{\tau}}^K, \boldsymbol{s}^J) = \mathbb{X}(\bar{\boldsymbol{\tau}}^K, \boldsymbol{s}^J)\gamma + \mathbb{U}, \tag{22}$$

where[3] $\mathbb{U} \sim N_{KJ}(0, \phi^2)$ is an error term that denotes the deviation between the empirical and theoretical log-characteristic functions. The bar denotes that the $\boldsymbol{\tau}^K$ are fixed. Combining with (16) we have the posterior resulting from the Discrete Approximation:

$$\begin{aligned} p(\alpha,\zeta,\{\boldsymbol{S}_t\},\Gamma|X) \propto \textstyle\prod_{t=1}^n g_{\alpha,d}\left(\frac{<X_t-\zeta,\boldsymbol{S}_t>}{\sigma_{\alpha,\Gamma}(\boldsymbol{S}_t)}, \beta_{\alpha,\Gamma}(\boldsymbol{S}_t)\right)\sigma_{\alpha,\Gamma}(\boldsymbol{S}_t)^{-d}\cdot p(\alpha,\zeta)\cdot \\ \phi^{-(n+1)}\exp\left\{-\tfrac{1}{2\phi^2}\left[(\mathbb{Y}-\mathbb{X}\gamma)'(\mathbb{Y}-\mathbb{X}\gamma)+\varpi\gamma'\gamma\right]\right\}\cdot\mathbb{I}(\gamma\geq 0). \end{aligned} \tag{23}$$

The posterior resulting from the Normal Approximation is the following:

$$\begin{aligned} p(\alpha,\zeta,\{\boldsymbol{S}_t\},\mu,\omega|X) \propto \textstyle\prod_{t=1}^n g_{\alpha,d}\left(\frac{<X_t-\zeta,\boldsymbol{S}_t>}{\sigma_{\alpha,\omega}(\boldsymbol{S}_t)}, \beta_{\alpha,\omega}(\boldsymbol{S}_t)\right)\sigma_{\alpha,\omega}(\boldsymbol{S}_t)^{-d}\cdot p(\alpha,\zeta)\cdot \\ \omega^{-(n+1)}\exp\left\{-\tfrac{1}{2\omega^2}\textstyle\sum_{t=1}^n \boldsymbol{S}_t^\top \boldsymbol{S}_t\right\}\cdot\textstyle\prod_{t=1}^n\mathbb{I}(\boldsymbol{S}_t\in\mathbb{S}^{d-1}). \end{aligned} \tag{24}$$

[3]Matrices $\mathbb{Y}$ and $\mathbb{X}$ are redefined so that they contain their real and imaginary parts.

We assume throughout that $p(\alpha, \zeta) \propto \mathbb{I}(0 < \alpha \leq 2)$. It appears that there are at least three novelties in the formulation of these posterior distributions.

1. The formal treatment of (7) in the context of (22) which facilitates considerably the posterior estimation of spectral weights, γ.

2. The introduction of latent variables $\{\boldsymbol{S}_t\}$ to avoid integration over $\mathbb{S}^{d-1}$ in (9).

3. The normal approximation to the measure Γ in connection with (9).

The proper prior on γ facilitates the regularization of the troublesome matrix $\mathbb{X}'\mathbb{X}$ through the parameter ϖ, and positive values of γ are guaranteed through the truncation. It is well known that the matrix is ill-conditioned in the univariate case when J is moderately large, due partly to the fact that placing optimally the support points is a difficult problem. See Koutrouvelis (1980, 1981) and Madan and Seneta (1987) among others. It seems equally difficult to find "optimal" placements in the multivariate case. In this paper we examine sensitivity of posterior results to alternative configurations of $\boldsymbol{\tau}^K$. Nolan, Panorska and McCulloch (2001) report that the matrix is well-behaved in the two-dimensional case even with fine grids. Our experience is similar and implies that one needs to avoid points where the real and imaginary parts of $I_X(\tau)$ have (numerically) the same value.

Of course, drawing γ from the second term in (24) is straightforward. It is required to draw from:

$$\begin{gathered} \gamma \sim N_J(\hat{\gamma}, V), \gamma \geq 0, \\ \hat{\gamma} = (\mathbb{X}'\mathbb{X} + \varpi I)^{-1}\mathbb{X}'\mathbb{Y}, V = \phi^2(\mathbb{X}'\mathbb{X} + \varpi I)^{-1}. \end{gathered} \tag{25}$$

This proposal is accepted with certain probability given by the first term of (24) to maintain the correct posterior distribution. Our MCMC algorithms (whose details are presented in Appendix A available in the working paper version of the study in arXiv, http://arxiv.org/abs/1507.07323) are as follows.

`MCMC − Discrete Approximation`

i. Draw $\gamma = \gamma^*$ using the normal linear model in (24). Denote the normal proposal by $q_\gamma(\gamma)$.

ii. Compute (19) and (20) and therefore (22).

iii. Accept the draw with certain probability.

iv. Update θ using (22) with given $\boldsymbol{s}_t = \bar{\boldsymbol{s}}_t$ using proposal $q_\theta(\theta)$. Accept the candidate with certain probability.

v. Update $\boldsymbol{s}_t$.

`MCMC − Normal Approximation`

i. Propose $\boldsymbol{S}_t^* \sim \mathscr{N}_d(\mu \cdot 1_d, \omega^2 I)$. Denote the proposal by $q(\boldsymbol{s}_t^*|\theta, X)$.

ii. Compute (19) and (20).

iii. Accept the draw with certain probability based on (23).

iv. Draw $\theta^* \sim q_\theta(\theta)$. Accept the candidate with certain probability based on (23).

There are certain numerical issues to resolve. *First*, the choice of the number of support points, J, is made formally using the log-marginal likelihood, for the MCMC-DA. *Second*, for the parameters θ we use as proposals Student-t densities, centered at the maximum likelihood quantities for α, ζ derived from fitting univariate stable symmetric distributions. For parameter ω we use as proposal: $\frac{\sum_{t=1}^n \boldsymbol{S}_t^\top \boldsymbol{S}_t}{\omega^2} \sim \chi_n^2$. The univariate stable symmetric density is computed using McCulloch's (1998) method. ML estimates of ζ and a diagonal covariance matrix times a constant h is used in constructing the proposal for ζ. For α we take the average of ML estimates with variance h times the median variance from ML estimates. We adjust h during the transient or "burn-in" phase to get acceptance rates between 20% and 30%.

Third, To construct a proposal $q_\gamma(\gamma)$ for the (normalized) spectral measure, we begin with draws from standard uniform, let the MCMC algorithm run through its transient phase and run it again for another $S_o = 50,000$ iterations. Then we use uniform proposals in intervals of the form $[a, b]$ where a and b are determined from the 99% probability intervals during the S_o phase. The termination of transient phase is determined using Geweke's (1992) diagnostics every 10,000 passes by comparing the first and last 2,500 draws. In artificial samples, depending on the parameters, we needed 50,000 to 150,000 draws. The results are not reported to save space but a separate appendix is available on request. All reported results are based on the final 100,000 draws by thinning every other tenth draw.

Fourth, the number of simulations, M, to approximate the functions $\beta(\mathbf{s})$ and $\sigma(\boldsymbol{s})$ is set to $M = 5,000$ and we check in preliminary numerical work whether this is sufficient. Depending on the values of α and $\boldsymbol{s}$ reasonable precision can be achieved with M ranging from 500 to 2,500. For some computational details see the first paragraph of Section A1 in the Appendix.

Fifth, in (22) and (6) or (7) we need to pre-select the configuration $\left\{\bar{\boldsymbol{\tau}}^K\right\}$ where the log-characteristic function is computed. We can use, again, points in $\mathbb{S}^{d-1}$ as in McCulloch (1994). We opt for

$$\bar{\tau}^{(i)} \sim \mathscr{N}_d(0, I), < \tau^{(i)}, \tau^{(i)} >= 1, i = 1, ..., K,$$

instead of a uniform distribution. The reason is that we need to concentrate such points near the origin (Madan and Seneta, 1987, Koutrouvelis, 1980, 1981, Xu and Knight, 2010, and Yu, 2007). We set $K = 10J$ so we have ten times as many points to evaluate the log-characteristic functions than the number of spectral weights. *Sixth*, the prior parameter $\varpi = 0.01$ which produces a sufficiently diffuse prior for the spectral weights although we also examine alternative values for this parameter.

5 Empirical application

We consider ten currencies against the US dollar over the period July 3 1996 to May 21 2012 (see Tsionas, 2012). The currencies are Canadian dollar, Euro, Japanese yen, British pound, Swiss franc, Australian dollar, Hong-Kong dollar, New Zealand dollar, South Korean won and Mexican peso. The data is daily and was converted to log differences. The data are filtered using an AR(1)-GARCH(1,1) model. MCMC is implemented using a preliminary, transient phase of length 20,000. Then we take another 100,000 draws from the posteriors of Discrete and Normal Approximations. The number of support points J for each $\boldsymbol{s}_t$ is determined by running different MCMC chains and computing the approximate log-marginal likelihood using the method of Lewis and Raftery (1997) and DiCiccio et al. (1997). For computational details see paragraph 3 in Section A1 of the Appendix.

Student-t proposal densities (with 10 degrees of freedom) were tuned to provide acceptance rates between 20% and 30% for the latent variables. From the results in Table 1 it turns out that Bayes factors favor $J = 20$ points for several values of ϖ so we choose this value to proceed further with Bayesian analysis with $\varpi = 0.01$. The matrix $\mathbb{X}^\top\mathbb{X}$ was not found singular in all but exceptional cases where approximately (numerically) grids were generated.

In such cases the parameter ϖ resolves the problem. This computational experience is consistent with the results in Nolan, Panorska and McCulloch (2001). In their paper they mention that one needs to scale the data by the median of $|X_t|$. Similar observations were made by Meerschaert and Scheffler (1999) and Tsionas (2012b). Here we followed the same approach.

Table 1. Bayes factors relative to $J = 5$

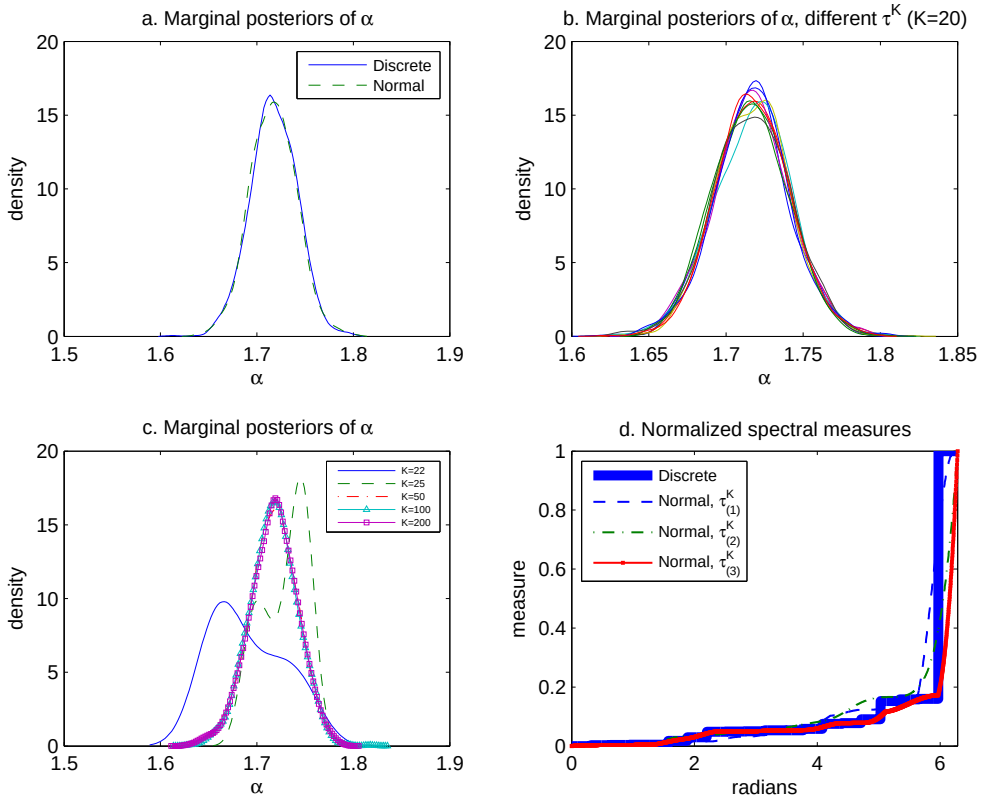
$J \rightarrow$	10	15	20	30	40	50
$\varpi = 10^{-4}$	27.12	33.87	71.23	40.01	22.1	12.67
$\varpi = 10^{-3}$	29.03	35.55	81.59	32.33	10.01	6.50
$\varpi = 0.01$	32.33	37.10	82.03	20.93	9.12	4.33
$\varpi = 0.1$	31.44	40.32	95.44	14.32	1.83	2.16

In panel (a) of Figure 1 we report marginal posterior densities of α using the Discrete and Normal approximation for a fixed configuration $\boldsymbol{\tau}^K$. Sensitivity of marginal posteriors with respect to ten alternative configurations[4] is examined in panel (b). Marginal posteriors for different values of K are given in panel (c) up to $K = 10J$. Clearly, as K increases marginal posteriors of α behave in the same way.

Finally, in panel (d) we report posterior means of the Discrete Approximation to the spectral measure (thick line) and typical posterior means from the Normal Approximation under various configurations of $\boldsymbol{\tau}^K$. In fact, the plot underestimates the Normal's ability to approximate closely the discrete measure (which we take as the "true" measure or its the best approximation) because the discrete measure also changes with the configuration. Overall, the results are quite robust reasonable and it is, indeed, encouraging that the Normal Approximation behaves so well provided, of course, that the multivariate characteristic function is evaluated at a large enough number of points. Moreover, the Bayes factor resolves successfully the problem of selecting the number of support points, J, for the latent $\boldsymbol{s}_t$.

[4]Autocorrelation of parameter draws is non-trivial as expected. The maximum autocorrelation ranged from 0.20 to 0.50 at lag 50 for the latent variables but was significantly lower for the structural parameters θ ranging between 0.10 and 0.30. So there is enough evidence that the chains mix well.

Figure 1. Empirical results for multivariate stable distribution, exchange rate data



To study the association between multivariate stable random variables, the role of the spectral (Levy) measure (Γ) has been found critical (Mittnik, Rachev and Rüschendorf, 1999). The dependence-at-the tails function can be estimated non-parametrically or using the posterior mean of the Levy measure which has been computed. Technical details on computing the Mittnik, Rachev and Rüschendorf (1999) function, $m(x)$, are too many to reproduce here, so we refer instead to their paper. Specifically, the dependence function is defined in their (2.7) using (2.6) with the spectral measure involved directly in (2.7) and (2.4). As they mention: "If an explicit parametric model is assumed, it would be more natural and efficient to use the rank-order process or the dependence function itself" (p. 184) which is precisely what we do here.

Our empirical results are reported in Table 2. We report posterior means of dependence at the tails. The tail is defined as the upper 5% percentile. Each dependence measure is the posterior mean from MCMC simulation. The tail dependence measures that include zero in the Bayes HPD are not reported. Diagonal elements are not reported.

Table 2. Empirical results: Tail-Dependence measures

We report the Mittnik, Rachev and Rüschendorf (1999) dependence measure at the tails. The tail is defined as the upper 5% percentile. Each dependence measure is the posterior mean from MCMC simulation. Tail dependence measures that include zero in the Bayes HPD are not reported. Diagonal elements are not reported.

	CAD	EUR	JPY	GBP	CHF	AUD	HKD	NZD	KRW	MXN
CAD		0.625		0.742						
EUR			0.645	0.815	0.713	0.613	0.511	0.662	0.772	0.341
JPY				0.740	0.855	0.410			0.815	
GBP					0.603		0.722	0.825	0.649	0.328
CHF						0.515			0.501	
AUD	0.610							0.717		0.332
HKD									0.423	
NZD										0.756
KRW										0.336
MXN										

Dependence at the tails is, obviously, quite large with the Euro associated with most currencies followed by the GBP and JPY. Non-parametric measures computed as in Mittnik, Rachev and Rüschendorf (1999) are somewhat different, showing that relying explicitly on multivariate stability delivers some gains in terms of efficiency assuming, of course, the model is a better description of reality. To our knowledge this is the first application of the tail-dependence measure provided an explicit Levy measure Γ is used. This measure is computed here as the posterior mean from Bayes MCMC simulation.

Given the empirical results it does not appear possible to remove any currency from the multivariate vector due to its weak dependence or no dependence at all to other currencies. This is despite the fact that we have allowed for an AR(1)-GARCH(1,1) scheme. Removing GARCH effects which are prevalent in many financial time series is essential in order to satisfy, at least approximately, the i.i.d. assumption involved in the analysis of multivariate stable distributions.

Concluding remarks

In this paper we break new ground in the treatment of multivariate stable distributions along the following lines. *First*, we propose a normal approximation to their spectral measure that seems to work very well in practice. *Second*, we propose efficient MCMC techniques by introducing appropriate latent variables. These are estimated from the data along with the spectral measure. *Third*, in connection with the important per se discrete approximation of the measure, we estimate it in the context of the simple normal linear model based on the log-characteristic function. The normal approximation reduces considerably the computational burden without sacrificing, as it seems, the quality of the approximation to the benchmark provided by the discretization of the spectral measure. The fact that it works well in a data set with ten variables and almost 4,000 observations is quite encouraging in terms of applications of multivariate stable distributions.

Posterior inferences seem to be quite robust with respect to the configurations of $\boldsymbol{\tau}^K$. Provided these are normally distributed over the unit sphere in $\mathbb{R}^d$, our MCMC schemes mix well with respect to the structural parameters and latent $\boldsymbol{s}^J$.

References

Abdul-Hamid, H., and J.P. Nolan, 1998, Multivariate stable densities as functions of one-dimensional projections, Journal of Multivariate Analysis 67, 80-89.

Bazant, Z.P., and B.H. Oh, 1986, Efficient numerical integration on the surface of a sphere, Zeitschrift für Angewandte Mathematik und Mechanik 1, 37-49.

Belisle, C., H. Romeijn, and R. Smith, R., 1993, Hit and run algorithms for generating multivariate distributions, Mathematics of Operations Research 18, 255-266.

Broda, S.A., Haas, M., Krause, J., Paolella, M.S., and S.C. Steude, 2013, Stable mixture GARCH models, Journal of Econometrics 172 (2), 292-306.

Buckle, D.J., 1995, Bayesian inference for stable distributions, Journal of the American Statistical Association 90, 605-613.

Byczkowski, T., J.P. Nolan, and B. Rajput, 1993, Approximation of multidimensional stable distributions, Journal of Multivariate Analysis 46, 13-31.

Cambanis, S., and G. Miller, 1981, Linear problems in pth order and stable processes, SIAM Journal of Applied Mathematics 41, 43-69.

Cheng, B.N., and S.T. Rachev, 1995, Multivariate stable futures prices, Mathematical Finance 5, 133-153.

Chib,S. and E. Greenberg, 1995, Understanding the Metropolis-Hastings algorithm, The American Statistician 49, 327-335.

DiCiccio, T. J., R.E. Kass, A. Raftery, Adrian, L. Wasserman, 1997, Computing Bayes Factors by Combining Simulation and Asymptotic Approximations 92, 903-915.

Dominicy, Y., and D. Veredas, 2013, The method of simulated quantiles, Journal of Econometrics 172 (2), 235-247.

Garcia, R., E. Renault, and D. Veredas, 2011, Estimation of stable distributions by indirect inference, Journal of Econometrics 161, 325-337.

Geweke, J.. 1992, Evaluating the accuracy of sampling based approaches to the calculation of posterior moments, in J.O. Berger, et al. (eds.), Bayesian Statistics, Vol. 4, pp. 169-194. Oxford: Oxford University Press.

Hallin, M., Swan, Y., Verdebout, T., and D. Veredas, 2013, One-step R-estimation in linear models with stable errors, Journal of Econometrics 172 (2), 195-204.

Koutrouvelis, I. A., 1980, Regression-type estimation of the parameters of stable laws, Journal of the American Statistical Association 75, 918–928.

Koutrouvelis, I. A., 1981, An iterative procedure for estimation of the parameters of stable laws, Communications in Statististcs - Simulation and Computation 10, 17–28.

Lewis, S.M. and A.E. Raftery, 1997, Estimating Bayes' factors via posterior simulation with the Laplace- Metropolis estimator, Journal of the American Statistical Association, 92, 648-655.

Lombardi, M. J., and D. Veredas, 2007, Indirect estimation of elliptical stable distributions, Computational Statistics and Data Analysis 53, 2309-2324.

Madan, D.B., and E. Seneta, 1987, Simulation of estimates using the empirical characteristic function, International Statistical Review 55, 153-161.

Matsui, M., and A. Takemura, 2009, Integral representations of one-dimensional projections for multivariate stable densities, Journal of Multivariate Analysis 100, 334-344.

McCulloch, J. H. , 1994, Estimation of bivariate stable spectral densities, Technical Report, Department of Economics, Ohio State University.

McCulloch, J. H., 1998, Numerical Approximation of the Symmetric Stable Distribution and Density, in R. Adler, R. Feldman, and M. Taqqu (eds), A practical guide to heavy tails: Statistical techniques for analyzing heavy tailed data, Boston, Birkhauser, 489-500.

McCulloch, J. H. , 2000, Estimation of the bivariate stable spectral representation by the projection method, Computational Economics 16, 47-62.

Meerschaert, M.M., and H.-P. Scheffler, 1999, Moment estimator for random vectors with heavy tails, Journal of Multivariate Analysis 71, 145-159.

Mittnik, S., S.T. Rachev and L. Rüschendorf, 1999, Test of association between multivariate stable vectors, Mathematical and Computer Modelling 29 (10-12), 181-195.

Nagaev, A., 2000, On non-parametric estimation of the Poisson spectral measure of a stable law, Journal of Mathematical Sciences 106, 2854-2859.

Nolan, J., 1997, Numerical calculation of stable densities and distribution functions, Communications in Statistics – Stochastic Models 13, 759-774.

Nolan, J. P., 1998, Multivariate stable distributions: approximation, estimation, simulation and identification. In R. J. Adler, R. E. Feldman, and M. S. Taqqu (Eds.), A Practical Guide to Heavy Tails: Statistical techniques for analyzing heavy tailed data, pp. 509-526. Boston: Birkhauser.

Nolan, J. P., A. Panorska, and J. H. McCulloch, 2001, Estimation of stable spectral measures, Mathematical and Computer Modelling 34, 1113-1122.

Nolan, J. P. and B. Rajput, 1995, Calculation of multidimensional stable densities, Communications in Statistics - Simulation 24, 551-556.

Ogata, H., 2013, Estimation for multivariate stable distributions with generalized empirical likelihood, Journal of Econometrics 172 (2), 248-254.

Paulauskas, V. I., 1976, Some remarks on multivariate stable distributions. Journal of Multivariate Analysis 6, 356-368.

Pourahmadi, M., 1987, Some properties of empirical characteristic functions viewed as harmonizable processes, Journal of Statistical Planning and Inference 17, 345-359.

Press, S. J., 1972, Estimation in univariate and multivariate stable distributions, Journal of the American Statistical Association 67, 842-846.

Ravishanker, N., and Z. Qiou, 1999, Monte Carlo EM estimation for multivariate stable distributions, Statistics & Probability Letters 45, 335-340.

Roose, D., and E. De Doncker, 1981, Automatic integration over a sphere, Journal of Computational and Applied Mathematics 7, 203-224.

Samorodnitsky, G., and M. S. Taqqu, 1994, Stable non-Gaussian random processes, Chapman and Hall, New York.

Stroud, A., 1971, Approximate Calculation of Multiple Integrals, Prentice Hall.

Tierney, L., 1994, Markov chains for exploring posterior distributions, Annals of Statistics 22 (4), 1701-28.

Tsionas, E.G., 1999, Monte Carlo inference in econometric models with symmetric stable disturbances, Journal of Econometrics 88, 365-401.

Tsionas, E.G., 2012, Estimating tail indices and principal directions easily, Statistics & Probability Letters 82, 1986-89.

Zolotarev, V.M., 1986, One-dimensional stable distributions, AMS Translations of Mathematical Monographs, vol. 65, AMS, Providence, Long Island.

Xu, D., and J. Knight, 2010, Continuous empirical characteristic function estimation of mixtures of normal parameters, Econometric Reviews 30, 25-50.

Yu, J., 2007, Empirical characteristic function and its applications, Econometric Reviews 23, 93-123.