Comput Stat (2013) 28:2067–2089
DOI 10.1007/s00180-013-0396-7

ORIGINAL PAPER

Multivariate elliptically contoured stable distributions: theory and estimation

John P. Nolan

Received: 23 June 2011 / Accepted: 27 December 2012 / Published online: 30 January 2013

Abstract Stable distributions with elliptical contours are a class of distributions that are useful for modeling heavy tailed multivariate data. This paper describes the theory of such distributions, presents formulas for calculating their densities, and methods for fitting the data and assessing the fit. Efficient numerical routines are implemented and evaluated in simulations. Applications to data sets of a financial portfolio with 30 assets and to a bivariate radar clutter data set are presented.

Keywords Stable distribution · Elliptical contours · Heavy tails

1 Introduction

Stable distributions are a class of probability distributions that generalize the normal law, allowing heavy tails and skewness that make them attractive in modeling financial data. While there are many attractive theoretical properties of stable laws, the use of these models in practice has been restricted by the lack of formulas for stable densities and distribution functions. The univariate stable distributions are now accessible: there are reliable programs to compute stable densities, distribution functions, and quantiles. And there are fast methods to simulate stable r.v.s and several methods of estimating stable parameters based on maximum likelihood, quantiles, empirical characteristic functions, and fractional moments, see Nolan (2001).

On the other hand, multivariate stable laws are only partially accessible. This is a function of the lack of closed form expressions for densities, and the possible complexity of the dependence structures. Byczkowski et al. (1993) and Abdul-Hamid and Nolan (1998) give expressions for general multivariate stable densities. In the

J. P. Nolan (✉)
American University, Washington, DC, USA
e-mail: jpnolan@american.edu

bivariate case, there are some methods of computing densities and estimating, but these are difficult to implement in higher dimensions. This paper focuses on a computationally tractable case—elliptically contoured stable laws. It is shown that one can compute densities, approximate cumulative probabilities for elliptical stable distributions in dimension $d \leq 100$. Maximum likelihood estimation for the parameters is feasible in low dimensions, and a pairwise estimation method is applicable to arbitrary dimensions.

If $\mathbf{X}$ is α-stable and elliptically contoured, then it has joint characteristic function

$$E \exp(i\mathbf{u}^T\mathbf{X}) = \exp(-(\mathbf{u}^T Q\mathbf{u})^{\alpha/2} + i\mathbf{u}^T\boldsymbol{\delta}) \tag{1}$$

for some $d \times d$ positive definite matrix Q and shift vector $\boldsymbol{\delta} \in \mathbb{R}^d$. Here $\mathbf{x}^T\mathbf{y} = \sum_{k=1}^{d} x_i y_i$ is the inner product in $\mathbb{R}^d$. The spectral measure of this stable law is known, but complicated; see Proposition 2.5.8 of Samorodnitsky and Taqqu (1994). We will call the matrix Q the *shape matrix* or *dispersion matrix* of the elliptical distribution. Note that it is not the covariance matrix of the distribution, which does not exist. It is also not exactly the covariation matrix, because for $\mathbf{X}$ given by (1), the pairwise covariations are $[X_i, X_j]_\alpha = q_{ij}q_{jj}^{\alpha/2-1}$. (Example 2.7.4 in Samorodnitsky and Taqqu 1994 shows this, with a constant factor that does not arise in our formulation.)

We assume throughout that $\mathbf{X}$ is nonsingular, which is equivalent to Q being strictly positive definite, i.e. for every $\mathbf{u} \neq 0$, $\mathbf{u}^T Q\mathbf{u} > 0$. All elliptically contoured stable distributions are scale mixtures of multivariate normal distributions, see Proposition 2.5.2 of Samorodnitsky and Taqqu (1994). Let $\mathbf{G} \sim N(\mathbf{0}, Q)$ be a d-dimensional multivariate normal r. vector and $A \sim \mathbf{S}(\alpha/2, 1, \gamma, 0)$ be an independent univariate positive $(\alpha/2)$-stable r. v. with $0 < \alpha < 2$. Then $\mathbf{X} = A^{1/2}\mathbf{G}$ is α-stable and elliptically contoured with joint characteristic function

$$\exp(-(\gamma/2)^{\alpha/2}(\sec \pi\alpha/4)(\mathbf{u}^T Q\mathbf{u})^{\alpha/2}).$$

For this reason, elliptically contoured stable distributions are sometimes called sub-Gaussian stable. This gives a formula for simulating elliptical stable distributions. In particular, if $0 < \alpha < 2$, $A \sim \mathbf{S}(\alpha/2, 1, 2(\cos \pi\alpha/4)^{2/\alpha}, 0)$ and $\mathbf{G} \sim N(0, Q)$, then

$$\mathbf{X} = A^{1/2}\mathbf{G} + \boldsymbol{\delta}$$

has characteristic function (1). An algorithm to simulate A is in Chambers et al. (1976) and $\mathbf{G}$ can be simulated by standard methods.

The isotropic/radially symmetric cases arise when Q is a multiple of the identity matrix, in which case the characteristic function simplifies to

$$E \exp(i\mathbf{u}^T\mathbf{X}) = \exp(-\gamma_0^\alpha |\mathbf{u}|^\alpha + i\mathbf{u}^T\boldsymbol{\delta}) \tag{2}$$

where $\gamma_0 > 0$ is a scale parameter and $\boldsymbol{\delta} \in \mathbb{R}^d$ is a location parameter. The spectral measure in this case is a uniform distribution on the unit sphere $\mathbb{S} = \{\mathbf{x}^T\mathbf{x} = 1\} \subset \mathbb{R}^d$. If $A \sim \mathbf{S}(\alpha/2, 1, 2\gamma_0^2(\cos \pi\alpha/4)^{2/\alpha}, 0)$ and independent $\mathbf{G} \sim N(\mathbf{0}, I)$, then $\mathbf{X} = A^{1/2}\mathbf{G} + \boldsymbol{\delta}$ has characteristic function (2).

We briefly comment on recent related work on this problem which the referees identified. Amengual and Sentana (2010) estimate elliptical models, but assume that fourth moments exist, so their methods do not handle the heavier tailed models considered here. In Bonato (2011), financial data is modeled by a multivariate elliptical stable model. The differences between that work and this one are that they use a much more numerically intensive method of numerically inverting the characteristic function on a grid and they use a dynamic conditional correlation GARCH model to deal with volatility. While they state that they fit the full model in one step (as we do with maximum likelihood estimation below), they only do this in two dimensions. For dimension higher than 2, they also resort to a pairwise model. Lombardi and Veredas (2009) use indirect inference to fit a multivariate stable elliptical model. Finally, Dominicy et al. (2010) fit multivariate elliptical stable models using sample quantiles to estimate the scale of certain terms. It is known that using sample quantiles to estimate the scale of univariate stable data sets is less efficient than the two methods used here. (See the discussion in Sect. 3.2.)

The organization of this paper is as follows. Section 2 focuses on a special case: the radially symmetric or isotropic case. Here the radial symmetry allows one to characterize the joint distribution in terms of the amplitude $R = |\mathbf{X}|$. The density and cdf of this univariate random variable can be numerically evaluated, and provides a way of evaluating the multivariate isotropic stable densities. Section 3 treats the elliptically contoured stable laws, shows how to compute their multivariate densities, and discusses estimation of this model. We then present simulations and a practical application, where the 30 stocks in the Dow Jones index are jointly analyzed as an elliptical stable model with $\alpha = 1.71$. A brief discussion of engineering and other applications in Sect. 4 is followed by an appendix with more facts about the amplitude distribution.

2 Isotropic stable distributions

2.1 The amplitude distribution

Let $\mathbf{X}$ be a centered d-dimensional isotropic stable random vector with characteristic function $\exp(-\gamma_0^\alpha |\mathbf{u}|^\alpha)$. The *amplitude* of $\mathbf{X}$ is defined by

$$R = |\mathbf{X}| = \sqrt{X_1^2 + \cdots + X_d^2}.$$

Our primary interest here is in using the distribution of univariate R to get expressions for the density of multivariate isotropic and elliptical stable distributions. However, in some problems the amplitude arises directly, so it is worthwhile exploring it's properties. This section derives expressions for its density and d.f. for any dimension. In dimension $d = 1$, isotropic is equivalent to symmetric, so the cumulative distribution function of $R = |X|$ is $F_R(r) = P(|X| \leq r) = F_X(r) - F_X(-r) = 2F_X(r) - 1$ and the density is $f_R(r) = 2f_X(r)$. For the rest of this paper we assume $d \geq 2$.

When $0 < \alpha < 2$, $\mathbf{X} \overset{d}{=} A^{1/2}\mathbf{Z}$, where $A \sim \mathbf{S}(\alpha/2, 1, 2\gamma_0^2(\cos \pi\alpha/4)^{2/\alpha}, 0)$ is positive stable and $\mathbf{Z} \sim N(\mathbf{0}, I)$, A and $\mathbf{Z}$ independent. Thus

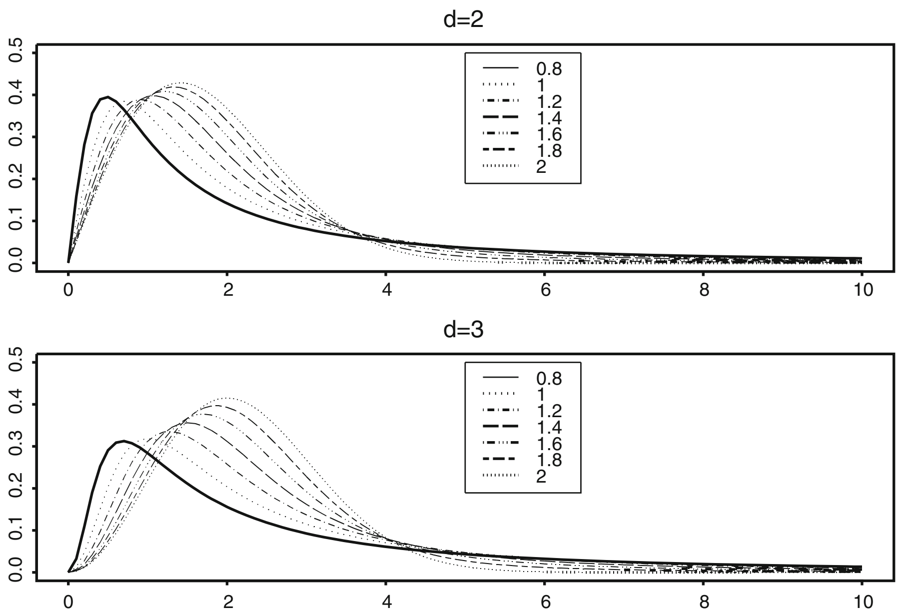

Fig. 1 The density of the standardized ($\gamma_0 = 1$) amplitude in 2 dimensions (*top*) and 3 dimensions (*bottom*) for $\alpha = 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2$

$$R^2 \stackrel{d}{=} A(Z_1^2 + \cdots + Z_d^2) = AT, \tag{3}$$

where T is chi-squared with d degrees of freedom, and independent of A. Using the standard formula for products of independent r.v., the d.f. of R can be expressed as

$$F_R(r) = F_R(r|\alpha, \gamma_0, d) = P(R \le r) = P(AT \le r^2) = \int_0^\infty F_A(r^2/t) f_T(t)dt, \tag{4}$$

and the density as

$$f_R(r) = f_R(r|\alpha, \gamma_0, d) = \frac{d}{dr} F_R(r) = 2r \int_0^\infty f_A(r^2/t) \frac{f_T(t)}{t} dt. \tag{5}$$

A scaling argument shows $F_R(r|\alpha, \gamma_0, d) = F_R(r/\gamma_0|\alpha, 1, d)$ and $f_R(r|\alpha, \gamma_0, d) = f_R(r/\gamma_0|\alpha, 1, d)/\gamma_0$. Figure 1 shows the graph of the density in two and three dimensions.

Equation (3) gives a way of simulating the amplitude distribution directly, without having to generate multivariate $\mathbf{X}$. It also gives an alternative way of simulating radially symmetric stable random vectors in d dimensions: let $A \sim \mathbf{S}(\alpha/2, 1, 2\gamma_0^2(\cos \pi\alpha/4)^{2/\alpha}, 0)$, $T \sim \chi^2(d)$, and $\mathbf{S}$ uniform on $\mathbb{S}$, then $\mathbf{X} \stackrel{d}{=} \sqrt{AT}\mathbf{S}$ is radially symmetric α-stable with scale γ_0. In particular, in two dimensions, T is exponential and can be generated by $-2\log U_1$ and $\mathbf{S} = (\cos(2\pi U_2), \sin(2\pi U_2))$, where U_1 and U_2 are independent U(0,1).

There are other expressions for the amplitude distribution. One is a simple change of variables: setting $s = r^2/t$ transforms (4) and (5) to

$$F_R(r) = r^2 \int_0^\infty s^{-2} F_A(s) f_T(r^2/s) ds = \frac{r^d}{2^{d/2}\Gamma(d/2)} \int_0^\infty F_A(s) s^{-d/2-1} e^{-r^2/(2s)} ds \tag{6}$$

$$f_R(r) = 2r \int_0^\infty s^{-1} f_A(s) f_T(r^2/s) ds = \frac{2r^{d-1}}{2^{d/2}\Gamma(d/2)} \int_0^\infty f_A(s) s^{-d/2} e^{-r^2/(2s)} ds \tag{7}$$

A third expression is from Zolotarev (1981):

$$f_R(r) = \frac{2}{2^{d/2}\Gamma(d/2)} \int_0^\infty (rt)^{d/2} J_{d/2-1}(r\,t) e^{-\gamma_0^\alpha t^\alpha} dt, \tag{8}$$

where $J_\nu(\cdot)$ is the Bessel function of order ν.

We have written a C program to compute f_R and F_R by evaluating (4) and (5) using numerical integration. It is based on existing routines to calculate the univariate stable d.f. F_A or the univariate stable density f_A, respectively. The current program works for $\alpha \geq 0.8$ and dimensions $1 \leq d \leq 100$. The integral in (8) is more difficult to evaluate reliably, because the integrand oscillates infinitely many times, whereas the integrands in (4) and (5) do not. (For $d \geq 100$, there are numerical difficulties integrating these expressions and the current algorithms are unreliable.)

More facts about the amplitude density and d.f. are given in the Appendix. The series expansions for the amplitude d.f. and density from there show behavior on the tail and near the origin:

$$\lim_{r\to\infty} r^\alpha (1 - F_R(r)) = \lim_{r\to\infty} r^\alpha P(R > r) = k_1 \gamma_0^\alpha \tag{9}$$

$$\lim_{r\to 0} r^{-d} F_R(r) = k_2 \gamma_0^{-d} \tag{10}$$

$$\lim_{r\to\infty} r^{\alpha+1} f_R(r) = \alpha k_1 \gamma_0^\alpha \tag{11}$$

$$\lim_{r\to 0} r^{1-d} f_R(r) = d k_2 \gamma_0^{-d} \tag{12}$$

for positive constants

$$k_1 = k_1(\alpha, d) = 2^\alpha \frac{\sin(\pi\alpha/2)}{\pi\alpha/2} \frac{\Gamma((\alpha+2)/2)\Gamma((\alpha+d)/2)}{\Gamma(d/2)},$$

$$k_2 = k_2(\alpha, d) = \frac{4\Gamma(d/\alpha)}{\alpha 2^d \Gamma(d/2)^2}.$$

We note that R is not stable, but (9) shows R is in the domain of attraction of a univariate α-stable law with $\beta = 1$.

2.2 Densities of isotropic stable distributions

Let $\mathbf{X}$ be any radially symmetric (around $\mathbf{0}$) r. vector, not necessarily stable, with density $f_{\mathbf{X}}(x)$ and amplitude $R = |\mathbf{X}|$. The d.f. of R, $F_R(r) = P(|\mathbf{X}| \leq r)$, directly gives circular probabilities. The following argument gives an expression for the density of X in terms of the density of R. Using polar coordinates and radially symmetry for $r > 0$,

$$\begin{aligned} F_R(r) = P(|\mathbf{X}| \leq r) &= \int_{|\mathbf{x}|\leq r} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_0^r \int_{\mathbb{S}} f_{\mathbf{X}}(u\mathbf{s}) u^{d-1} d\mathbf{s} du \\ &= \int_0^r \int_{\mathbb{S}} f_{\mathbf{X}}(u, 0, 0, \ldots, 0) u^{d-1} d\mathbf{s} du \\ &= \text{Area}(\mathbb{S}) \int_0^r f_{\mathbf{X}}(u, 0, 0, \ldots, 0) u^{d-1} du. \end{aligned}$$

Differentiating shows $f_R(r) = \text{Area}(\mathbb{S}) f_{\mathbf{X}}(r, 0, 0, \ldots, 0) r^{d-1}$. Hence for $\mathbf{x} \neq \mathbf{0}$, radial symmetry shows

$$f_{\mathbf{X}}(\mathbf{x}) = f(|\mathbf{x}|, 0, \ldots, 0) = \frac{f_R(|\mathbf{x}|)|\mathbf{x}|^{1-d}}{\text{Area}(\mathbb{S})} = \frac{\Gamma(d/2)}{2\pi^{d/2}} |\mathbf{x}|^{1-d} f_R(|\mathbf{x}|). \tag{13}$$

The key fact here is that calculating the density of multivariate $\mathbf{X}$ only requires calculating the univariate function $f_R(r)$.

Therefore, when $\mathbf{X}$ is α-stable with characteristic function (2), the above reasoning shows

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} \left(\Gamma(d/2)/\left(2\pi^{d/2}\right)\right) |\mathbf{x} - \boldsymbol{\delta}|^{1-d} f_R(|\mathbf{x} - \boldsymbol{\delta}|\,|\alpha, \gamma_0, d) & \mathbf{x} \neq \boldsymbol{\delta} \\ \Gamma(d/\alpha)/\left(\alpha 2^{d-1} \pi^{d/2} \Gamma(d/2)^2 \gamma_0^d\right) & \mathbf{x} = \boldsymbol{\delta}. \end{cases}$$

The value at $\mathbf{x} = \boldsymbol{\delta}$ uses (12). It is useful to consider the *radial function* $h(r|\alpha, d) = f_{\mathbf{X}}(r, 0, \ldots, 0|\alpha, \gamma_0 = 1, \boldsymbol{\delta} = \mathbf{0})$, which is given by

$$h(r|\alpha, d) = \begin{cases} \Gamma(d/2)/\left(2\pi^{d/2}\right) r^{1-d} f_R(r|\alpha, \gamma_0 = 1, d) & r > 0 \\ \Gamma(d/\alpha)/\left(\alpha 2^{d-1} \pi^{d/2} \Gamma(d/2)^2\right) & r = 0. \end{cases}$$

Then for a general isotropic α-stable $\mathbf{X}$ with scale γ_0 and location $\boldsymbol{\delta}$,

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\gamma_0^d} h\left(\frac{|\mathbf{x} - \boldsymbol{\delta}|}{\gamma_0} \,\middle|\, \alpha, d\right). \tag{14}$$

A different approach is to use Abdul-Hamid and Nolan (1998). Assuming $\gamma_0 = 1$ and $|\mathbf{x}| > 0$,

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-d} \int_{\mathbb{S}} g_{\alpha,d}(\mathbf{x}^T\mathbf{s})d\mathbf{s},$$

where $g_{\alpha,d}(t) = \int_0^\infty \cos(tr)r^{d-1}e^{-r^\alpha}dr$. Then using the radial symmetry of $\mathbf{X}$ and the fact that the projection of $\mathbf{s} = (s_1, \ldots, s_d)$ onto the first component s_1 has density $c_1(1-t^2)^{(d-3)/2}$,

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-d} \int_{-1}^{1} g_{\alpha,d}(|\mathbf{x}|\, s_1)d\mathbf{s} = c_2 \int_{0}^{1} g_{\alpha,d}(|\mathbf{x}|\, t)(1-t^2)^{(d-3)/2}dt,$$

where $c_1 = 2\pi^{(d-1)/2}/\Gamma((d-1)/2)$ and $c_2 = 2c_1(2\pi)^{-d}$. Using (13) gives another expression for the amplitude density: $f_R(r) = c_3 r^{d-1}\int_0^1 g_{\alpha,d}(rt)(1-t^2)^{(d-3)/2}dt$.

3 Elliptically contoured stable distributions

3.1 Densities of elliptically contoured stable laws

Let $\mathbf{Y}$ be d-dimensional α-stable elliptically contoured random vector with shape matrix Q and shift vector $\boldsymbol{\delta}$. Then $\mathbf{Y} \overset{d}{=} A^{1/2}\mathbf{G} + \boldsymbol{\delta}$, where positive $A \sim \mathbf{S}(\alpha/2, 1, 2(\cos\pi\alpha/4)^{2/\alpha}, 0)$ and $\mathbf{G} \sim$N$(0, Q)$ as above. It is well known that $\mathbf{G} \overset{d}{=} Q^{1/2}\mathbf{G}_1$, where $Q^{1/2}$ is from the Cholesky decomposition of Q and $\mathbf{G}_1 \sim$N(0,I) has independent standard normal components. Hence $\mathbf{Y} \overset{d}{=} A^{1/2}Q^{1/2}\mathbf{G}_1 + \boldsymbol{\delta} = Q^{1/2}A^{1/2}\mathbf{G}_1 + \boldsymbol{\delta} := Q^{1/2}\mathbf{X} + \boldsymbol{\delta}$, where $\mathbf{X}$ is radially symmetric α-stable. So $\mathbf{Y}$ is an affine transformation of $\mathbf{X}$, and (14) shows

$$f_{\mathbf{Y}}(\mathbf{y}) = |\det Q|^{-1/2} f_{\mathbf{X}}(Q^{-1/2}(\mathbf{y}-\boldsymbol{\delta})) = |\det Q|^{-1/2} h\left(|Q^{-1/2}(\mathbf{y}-\boldsymbol{\delta})| \,\middle|\, \alpha, d\right). \tag{15}$$

We note that this is true for any elliptical distribution: the amplitude density of the isotropic case gives an expression for multivariate densities of the corresponding elliptically contoured distribution.

3.2 Statistical analysis of data as elliptical stable

We first describe ways of assessing a d-dimensional data set to see if it is approximately sub-Gaussian and then discuss two methods of estimating the parameters. These methods are evaluated in simulations and then illustrated using the 30 stocks that make up the Dow Jones index.

To assess for elliptical stability, first perform a one dimensional stable fit to each component of the data using a univariate estimation method to get estimates $\widehat{\boldsymbol{\theta}}_i = (\widehat{\alpha}_i, \widehat{\beta}_i, \widehat{\gamma}_i, \widehat{\delta}_i)$. If the $\widehat{\alpha}_i$'s are significantly different, then the data is not jointly $\widehat{\alpha}$-stable, so it cannot be sub-Gaussian. Likewise, if the $\widehat{\beta}_i$'s are not close to 0, then the distribution is not symmetric and it cannot be sub-Gaussian. If the $\widehat{\alpha}_i$'s are all close, form a pooled estimate of $\widehat{\alpha} = (\sum_{i=1}^{d} \widehat{\alpha}_i)/d =$ average of the indices of each component.

Next, assess for elliptical behavior. This can be approached by examining two dimensional projections because if $\mathbf{X}$ is a d-dimensional elliptical α-stable random vector, then every two dimensional projection

$$\mathbf{Y} = (Y_1, Y_2) = (\mathbf{a}_1^T \mathbf{X}, \mathbf{a}_2^T \mathbf{X}) \tag{16}$$

($\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^d$) is a 2-dimensional elliptical α-stable random vector. Conversely, suppose $\mathbf{X}$ is a d-dimensional α-stable random vector with the property that every two dimensional projection of form (16) is non-singular elliptical, then d-dimensional $\mathbf{X}$ is non-singular elliptical α-stable. Thus it suffices to assess multivariate data by looking at two dimensional distributions. While one cannot do this for all projections, one can do some basic checking. A partial solution is to check pairs of assets visually by looking at scatter plots.

Full estimation for an elliptical stable model requires estimating the stable index α, a shift $\boldsymbol{\delta}$, and the upper triangular part of the matrix Q, i.e. $1 + d + d(d+1)/2 = (d^2 + 3d + 2)/2$ parameters. This is one more parameter than is needed in a Gaussian fit, which is equivalent to setting $\alpha = 2$. Estimating these parameters can be done in two ways. The first is full maximum likelihood (ML), and the second we call a projection method. While maximum likelihood is theoretically the most efficient, it is computationally very expensive with the existing algorithms. Below we give simulation results that show that it is prohibitive to use ML with the existing numerical algorithms for dimensions $d > 4$, and that the projection method gives approximately the same efficiency.

The maximum likelihood scheme is conceptually straightforward: the log-likelihood of $(\alpha, Q, \boldsymbol{\delta})$ given the data is

$$\ell(\alpha, Q, \boldsymbol{\delta} | \mathbf{X}_1, \ldots, \mathbf{X}_n) = \sum_{i=1}^{n} \log f(\mathbf{X}_i | \alpha, Q, \boldsymbol{\delta}).$$

Use (15) to evaluate the right hand side above for a given $(\alpha, Q, \boldsymbol{\delta})$ and then numerically maximize to get maximum likelihood (ML) estimators:

$$(\widehat{\alpha}_{ML}, \widehat{Q}_{ML}, \widehat{\boldsymbol{\delta}}_{ML}) = \arg \min_{\alpha, Q, \boldsymbol{\delta}} \ \ell(\alpha, Q, \boldsymbol{\delta} | \mathbf{X}_1, \ldots, \mathbf{X}_n).$$

The second method is based on linear combinations of components and any method of estimating univariate stable parameters. First we estimate an index $\widehat{\alpha}_i$ and a shift $\widehat{\delta}_i$ for each component. As above, we use a pooled estimate for the index of stability: $\widehat{\alpha}_{PROJ} = (\sum_{i=1}^{d} \widehat{\alpha}_i)/d$. Then shift the data by $\hat{\boldsymbol{\delta}}_{PROJ} := (\hat{\delta}_1, \hat{\delta}_2, \ldots, \hat{\delta}_d)$ so the data is centered at the origin. In what follows, we assume $\mathbf{X}$ has been centered in this way.

To estimate the shape matrix $Q = (q_{ij})$, note that for any $\mathbf{u}$, the linear combination $\mathbf{u}^T\mathbf{X}$ is univariate α-stable with (squared) scale

$$\gamma^2(\mathbf{u}) := [-\ln E \exp(i\mathbf{u}^T\mathbf{X})]^{2/\alpha} = \mathbf{u}^T Q\mathbf{u} = \sum_i u_i^2 q_{ii} + 2\sum_{i<j} u_i u_j q_{ij}. \quad (17)$$

In particular if we let $\mathbf{e}_i$ be the standard ith unit basis vector, $\gamma^2(\mathbf{e}_i) = q_{ii}$ and $\gamma(\mathbf{e}_i + \mathbf{e}_j) = q_{ii} + q_{jj} + 2q_{ij}$. We use these population identities to obtain sample estimates of the shape parameters $\widehat{q}_{ij}$. For the diagonal elements, set $\widehat{q}_{ii} = \widehat{\gamma}_i^2 = \widehat{\gamma}^2(\mathbf{e}_i)$, i.e. the square of the estimated scale parameter of the ith coordinate. Likewise, estimate q_{ij} by analyzing the pair (X_i, X_j) and set $\widehat{q}_{ij} = (\widehat{\gamma}^2(\mathbf{e}_i + \mathbf{e}_j) - \widehat{q}_{ii} - \widehat{q}_{jj})/2$, where $\widehat{\gamma}(\mathbf{e}_i + \mathbf{e}_j)$ is the estimated scale parameter of the projection $(\mathbf{e}_i + \mathbf{e}_j)^T\mathbf{X} = (1, 1)^T(X_i, X_j) = X_i + X_j$. This involves estimating $d(d+1)/2$ one dimensional scale parameters. Since this method projects all pairs of the data to a sequence of one dimensional data sets, we call this a projection estimator and denote the resulting estimate by $\widehat{Q}_{PROJ}$.

There are a couple variations of the projection method. Using (17) again, $\gamma^2(\mathbf{e}_i - \mathbf{e}_j) = q_{ii} + q_{jj} - 2q_{ij}$, and combining with the expression above, $\gamma^2(\mathbf{e}_i + \mathbf{e}_j) - \gamma^2(\mathbf{e}_i - \mathbf{e}_j) = 4q_{ij}$. Plugging in sample estimates of $\gamma(\mathbf{e}_i + \mathbf{e}_j) =$ scale of $X_i + X_j$ and $\gamma(\mathbf{e}_i - \mathbf{e}_j) =$ scale of $X_i - X_j$ give a different projection estimate of q_{ij}. Note that it involves twice as many univariate estimation steps for the off-diagonal elements of Q as the method above. A second variation is to include even more projections. Pick a grid of $\mathbf{u}$ points, then for each $\mathbf{u}$ form a univariate data set by projecting each data vector $\mathbf{u}^T\mathbf{X}_i$, and estimating the scale for each of these univariate data sets. Using these estimates on the left hand side of (17), and the known $\mathbf{u}$ values on the right hand side, we get a linear system in the q_{ij} that can be estimated by regression.

In the projection methods, the shape matrix is estimated term-by-term and there is a possibility that the resulting matrix $\widehat{Q}$ is not positive definite. This can be corrected by methods like those of Higham (2002). The function `nearPD` in the R package `Matrix` implements this algorithm, and we use it to guarantee that $\widehat{Q}$ is a positive definite matrix.

While we do not consider the non-stable case here, this projection approach works for any elliptical model, whether heavy tailed or not. Specifically, the multivariate shape model can be estimated by looking at univariate projections of pairs of (centered) components. This requires a way to estimate the scale for the particular distributional model being used.

We note that it is complicated to directly search over positive definite matrices Q. Our implementation of the full maximum likelihood method uses the fact that any positive definite Q can be written as $S^T S$, where S is an upper triangular matrix (Cholesky decomposition). The ML search is initialized by using a projection estimator and then searching over the space of upper triangular matrices S.

3.3 Simulation results

Simulation was used to compare full maximum likelihood estimation with the projection method. For the projection method, we can use any one of several univariate

Table 1 Execution times for the different estimation methods when $n = 1,000$

d	Execution time (s)		
	PROJ_ECF	PROJ_ML	FULL_ML
2	<0.01	0.22	1,417
3	<0.01	0.42	2,365
4	0.01	0.77	3,551
5	0.02	1.03	n/a
10	0.05	3.90	n/a
25	0.22	25.27	n/a
50	0.91	93.49	n/a
100	3.62	372.61	n/a
1,000	347.31	n/a	n/a

Times are in seconds, when run on a 2.4 GHz Intel processor using a single core. Note that the clock resolution is approximately 0.01 s, so small number are approximate. n/a indicates not available—timing results were not done because of the long execution times

estimation techniques. In principle, univariate maximum likelihood is asymptotically the most efficient, and can now be computed relatively quickly and reliably using the method of Nolan (2001). However, other methods can be almost as efficient and much faster. Below both ML and the empirical characteristic function (ECF) method of Kogon and Williams (1998) will be used. Separate simulations have shown that the relative efficiency of the ECF method, defined as MSE($\widehat{\theta}_{ECF}$)/MSE(($\widehat{\theta}_{ML}$), is approximately 1.3 for estimating α, 1.4 for estimating β, 1.05 for estimating γ (which is of most interest to us here), and 1.03 for estimating δ. (These relative efficiency values are based on 1,000 simulations of size $n = 2,000$ with $\alpha = 1.7$ and $\beta = 0$, motivated by the values in the Dow Jones portfolio example analyzed below.) The timing results in Table 1 show that the ECF method is about 100 times as fast as ML. Other univariate estimation methods, e.g. fractional moment estimators or quantile based estimators, can be quick, but are much less accurate for small and moderate sized samples.

So in the simulations below, we will compute three estimators: full maximum likelihood—denoted by $\widehat{\alpha}_{FULL\_ML}$, $\widehat{Q}_{FULL\_ML}$ and $\widehat{\boldsymbol{\delta}}_{FULL\_ML}$; projection method with empirical characteristic function method of estimating the univariate parameters—denoted by $\widehat{\alpha}_{PROJ\_ECF}$, $\widehat{Q}_{PROJ\_ECF}$ and $\widehat{\boldsymbol{\delta}}_{PROJ\_ECF}$; and projection method with maximum likelihood method of estimating the univariate parameters—denoted by $\widehat{\alpha}_{PROJ\_ML}$, $\widehat{Q}_{PROJ\_ML}$ and $\widehat{\boldsymbol{\delta}}_{PROJ\_ML}$;

The first simulation we did was to compare efficiency of the three different estimators in dimension $d = 3$. For this, we used $\alpha = 1.86$, $\boldsymbol{\delta} = (0, 0, 0)$ and

$$Q = 10^{-5} \times \begin{pmatrix} 6.293 & 3.289 & 3.643 \\ 3.289 & 9.133 & 3.921 \\ 3.643 & 3.921 & 7.871 \end{pmatrix}.$$

(The values of α and Q came from a real data: the daily returns from three stocks—IBM, Google, and Microsoft were analyzed over a 1 year period. A fit was done using the projection ECF method.) Figure 2 shows the results from analyzing $M = 215$ simulations, each of $n = 250$ days, i.e. one trading year. Each set of boxplots shows

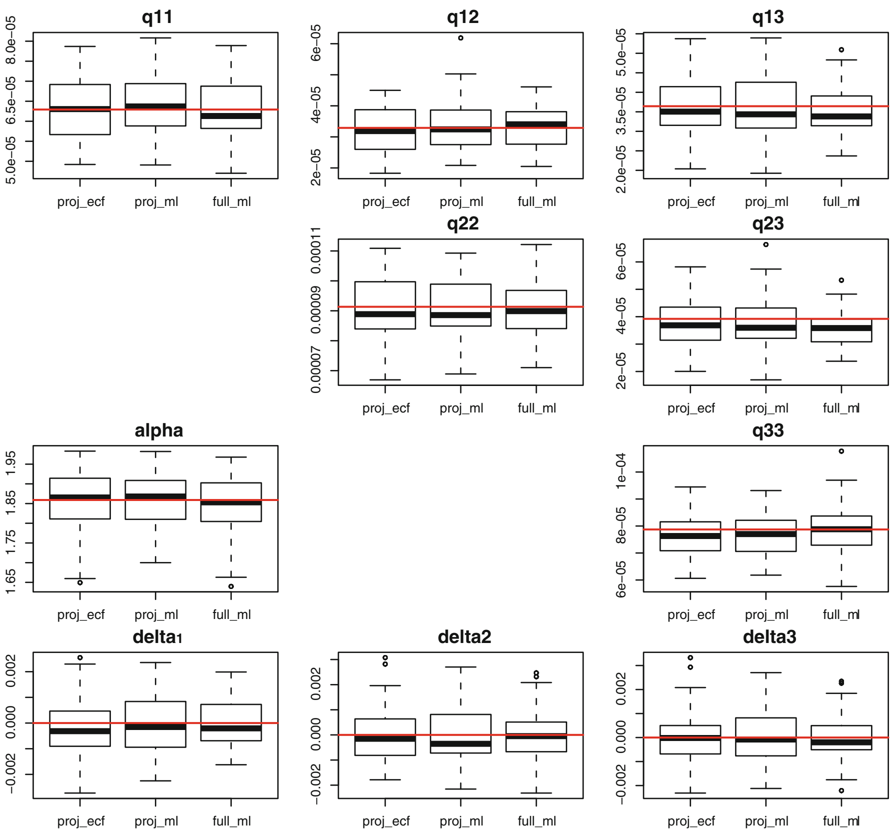

Fig. 2 *Boxplots* of the different estimators. The order is arranged so that the upper triangular entries q_{ij} of Q are in the *upper right*. Then α and $\boldsymbol{\delta}$ are *below*

the spread of the three estimators over the 215 simulations for each parameter, with the horizontal line showing the exact value. Our conclusion is that the three methods yield approximately the same accuracy.

Next we performed timing results. Here for different dimensions d and number of data vectors n, we set $\alpha = 1.7$, simulated a random positive definite matrix Q, and a random shift $\boldsymbol{\delta}$. Then we fit the data set using the three estimation methods described above. The results are shown in Table 1. There are huge differences in timing, for different reasons. The two projection methods both have to estimate $d(d+1)/2$ parameters in the shape matrix Q, but this is done using existing univariate estimation routines that have been optimized for speed and programmed in C. The ECF method is much faster than ML, and seems to give almost the same accuracy. In contrast, the full ML method is programmed in R and is very slow. Even in dimension 4, it took almost 10 h of computing time. Since the computations involve $\sim nd^2$ operations, this will get extremely costly for higher dimensions. While the full ML method could be improved by translating from R to C, we do not see much reason to do this as the other methods are much, much faster and give comparable accuracy.

We conclude that full ML is prohibitive if $d > 4$. The projection ECF method is already feasible for matrices up to size 100×100 in <4 s. Furthermore, the projection methods are easily implemented in a parallel processing environment—separate threads only need read access to data vectors, no sequential steps or complicated control is needed. While parts of full ML estimation could be parallelized, the slowest part is the minimization step, which is a sequential procedure. Hence parallelization is not likely to greatly increase the speed in the ML case.

3.4 Analysis of the Dow Jones 30 portfolio

Adjusted daily closing prices for the 30 stocks in the Dow Jones index (DJ30) were collected between January 3, 2000 and December 31, 2004. Days with missing prices for one or more stock were deleted—this occurred 8 times in the 2,256 trading days. Log-ratios of consecutive prices were computed separately for each stock, with the resulting data set having 2,247 returns for 30 stocks.

The results of the analysis of each component of the Dow Jones data set is given in Table 2, Fig. 3 shows plots of the estimated α and β for each of the 30 components from the Table, and Fig. 4 shows one pairwise plot. While there is some variability in the α's and in the βs around 0, we will proceed with the analysis here, and discuss these issues in more detail below.

The 30×30 shape matrix Q was estimated for this set using the first method above. For space reasons we do not show the numeric values of this large matrix, instead a heat map of Q is displayed in Fig. 5. The color shows the magnitude of the entries in the shape matrix. The estimation of the individual stable fits and the shape matrix estimation using the projection method with maximum likelihood estimation for the 30 component example took 36.5 s on a desktop PC. The projection method with ECF univariate estimations took 0.42 s, and the results from the two estimation methods were highly correlated.

Because it is possible to quickly simulate from an elliptical stable distribution of high dimension, Monte Carlo estimates of tail probabilities can be computed. Figure 6 compares the probability $P(|X_i| < a, i = 1, \ldots, 30)$ for (a) the observed data of size 1,247, (b) an MC estimate from a simulated stable sample of size $n = 10{,}000$ generated from the elliptical stable fit, and (c) an MC estimate from a simulated normal sample of size $n = 10{,}000$ generated from the normal fit. Note that the normal fit severely underestimates the tail, while the stable fit is more accurate above the 0.95 level. In particular, if a multivariate value at risk is computed for this data, the normal model will significantly underestimate the risk. For example, using the bottom plot of Fig. 6 gives the 0.995 VaR as approximately 0.10 for the normal model, whereas the empirical value is approximately 0.25 and the stable model value is approximately 0.28. Note that the time frame of this data set does not include the 2008 economic crisis, where more extreme behavior was present.

We briefly discuss the appropriateness of an elliptical stable model. While the fluctuations in the αs from the different components in Fig. 3 and the fluctuations of the βs around 0 argue that an elliptical stable model won't be perfect, we think that there is reason to use these models. First of all, a normal model does a worse

Table 2 Maximum likelihood estimates of stable parameters for the 30 stocks in the Dow Jones index

Symbol	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$
MMM	1.69	0.27	0.00972	−0.000657
AA	1.86	0.16	0.01623	−0.000576
MO	1.53	−0.05	0.01067	0.001335
AXP	1.72	−0.00	0.01383	0.000284
AIG	1.69	0.05	0.01168	−0.000313
BA	1.80	−0.05	0.01379	0.000720
CAT	1.82	0.27	0.01334	−0.000213
C	1.70	0.02	0.01280	0.000210
KO	1.62	0.01	0.00956	−0.000222
DD	1.69	0.26	0.01134	−0.001330
XOM	1.79	−0.25	0.00966	0.000862
GE	1.73	0.11	0.01268	−0.000670
GM	1.71	0.09	0.01319	−0.000709
HPQ	1.68	0.05	0.01805	−0.000970
HD	1.65	0.01	0.01436	−0.000229
HON	1.64	0.08	0.01460	−0.000469
INTC	1.75	0.05	0.02048	−0.000495
IBM	1.59	0.02	0.01179	−0.000361
JNJ	1.73	0.02	0.00939	0.000376
JPM	1.67	−0.00	0.01452	−0.000313
MCD	1.69	−0.03	0.01128	0.000115
MRK	1.72	−0.10	0.01127	0.000218
MSFT	1.65	0.02	0.01386	−0.000498
PFE	1.75	0.00	0.01216	0.000095
PG	1.55	0.08	0.00816	0.000170
SBC	1.71	0.01	0.01339	−0.000474
UTX	1.77	−0.01	0.01263	0.000721
VZ	1.76	0.11	0.01274	−0.000637
WMT	1.66	0.08	0.01185	−0.000689
DIS	1.79	0.16	0.01485	−0.000644
	$\overline{\alpha} = 1.71$			

Note that δ is in the continuous 0-parameterization. α and β across stocks are plotted in Fig. 3

job of modeling the data: all of the estimated αs are below 2, so a normal model will underestimate the tails in the data. A normal model implicitly assumes symmetry, so an elliptical stable model won't do any worse than this. Within the family of stable models, we can compute a likelihood ratio test using the computations of the 30 dimensional density function for both the Gaussian model and the non-Gaussian stable model. For the Dow Jones data, the stable log-likelihood is $\ell_1 = 96,307$. In contrast, if the data is fit with a $\mathrm{N}(\boldsymbol{\mu}, Q)$ model, the log-likelihood is $\ell_2 = 97,549$. The likelihood ratio test is $\exp(\ell_1 - \ell_2) \approx 10^{539}$, strongly favoring the stable model.

Here G. Box's quote seems relevant: "All models are wrong, but some are useful." An elliptical multivariate stable model allows for heavy tails, captures some of the

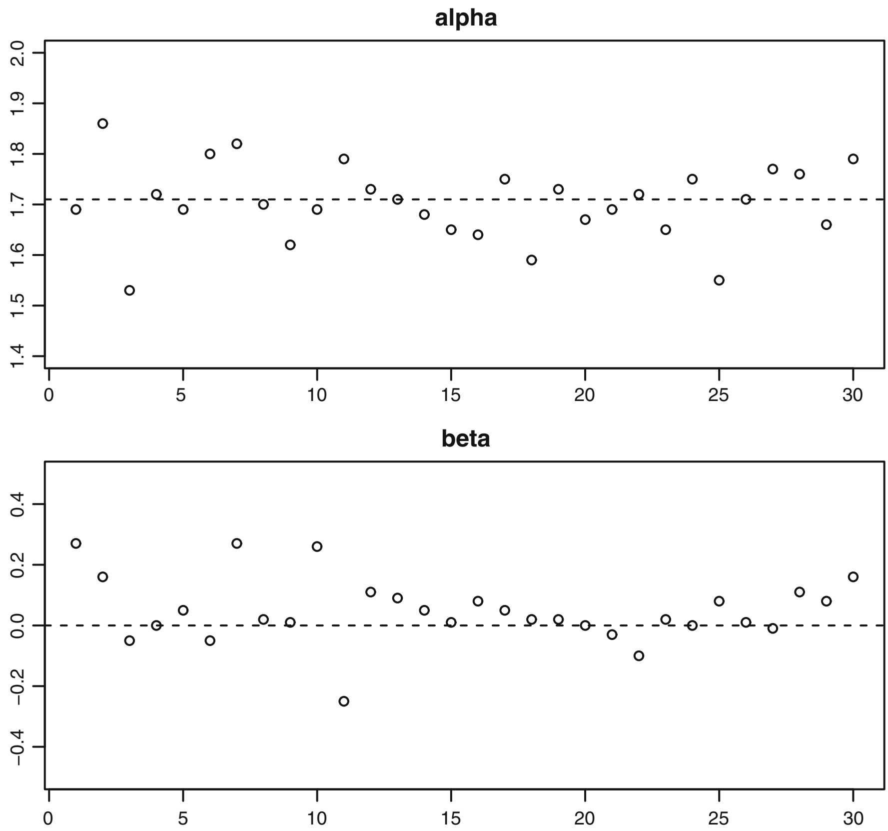

Fig. 3 $\widehat{\alpha}$ and $\widehat{\beta}$ for the 30 components of the Dow Jones data. The *horizontal* axis is the index of the stock in the order listed in Table 2. The *dashed horizontal lines* show the values used in the joint analysis: $\widehat{\alpha}_{PROJ} = 1.71$ and $\beta = 0$, respectively

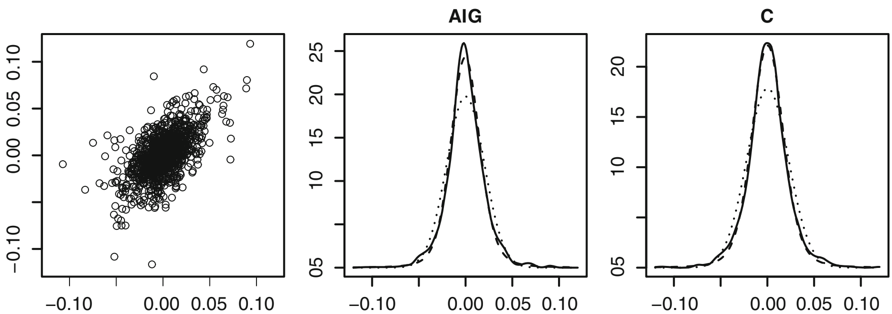

Fig. 4 Comparison of returns for AIG and Citigroup (*symbol C*). The *scatterplot on the left* shows an approximate elliptical pattern (AIG on the *horizontal* axis, Citigroup on the *vertical* axis). The other *plots* show the marginals for each asset: the *solid curve* is smoothed data, *dashed curve* is stable fit, *dotted curve* is normal fit

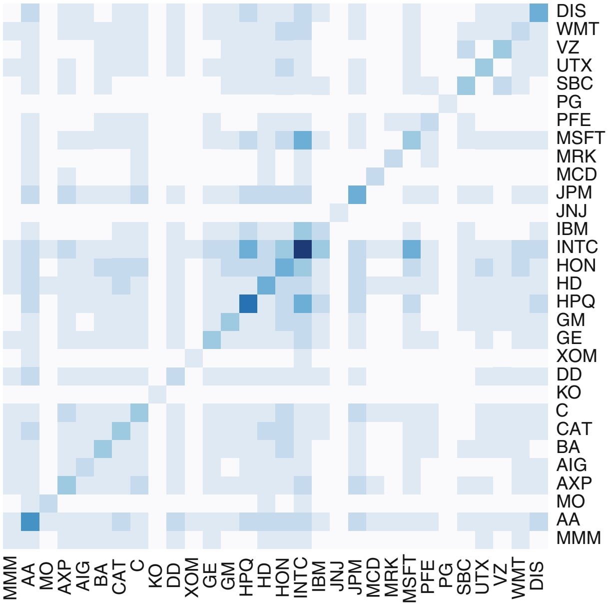

Fig. 5 Heat map of the shape matrix Q for the DJ30 portfolio. *Shading* ranges from *white* for lowest values (min = 0.0000177), through increasing *darkness* to highest values (max = 0.000419)

dependence, and retains the property of accumulated returns having the same type of distribution as daily returns: cumulative sums of stable terms are stable.

Another advantage of the joint stable model is that linear combinations are automatically univariate stable. Hence for a portfolio of d assets with returns $(X_1, \ldots, X_d)$ modeled by elliptical stable model (1) and weights $\mathbf{w} = (w_1, \ldots, w_d)$, the distribution of the weighted returns $w_1 X_1 + \cdots + w_d X_d$ is univariate stable $S(\alpha, 0, \gamma(\mathbf{w}), \delta(\mathbf{w}))$, where $\gamma(\mathbf{w}) = (\mathbf{w}^T Q \mathbf{w})^{1/2}$ and $\delta(\mathbf{w}) = \mathbf{w}^T \boldsymbol{\delta}$.

In large practical problems, there may be many assets that exhibit different tail behaviors. It is possible to adapt the ideas of grouping data as in Daul et al. (2003). They considered grouping data with similar tail behavior and then modeling each group with a multivariate elliptical t-distribution, with different groups having different degrees of freedom. One can extend this to allow the groups to have elliptical stable models with different indices α, some Gaussian components, and some components with multivariate t distributions, potentially with different degrees of freedom.

3.5 Serial dependence and tail dependence

Our primary focus in this paper is on the dependence structure among the components of a multivariate heavy tailed data set without serial dependence. This was not an issue for the simulations results discussed above, where we simulated independent observations, but it may be relevant for the Dow Jones 30 analysis and other applications. Here we summarize results of estimation for that data set with various types of serial dependence modeling: AR(1), ARMA(1,1) and GARCH(1,1) combined with

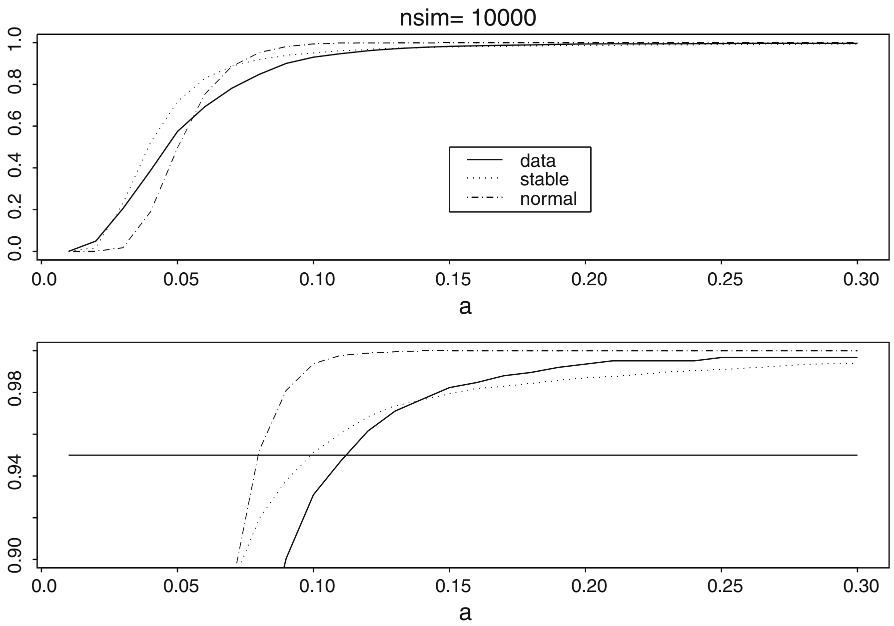

Fig. 6 Comparison of empirical estimate of $P(|X_i| < a, i = 1, \ldots, 30)$ (*solid line*) for 30 dimensional data set to Monte Carlo estimates for stable fit (*dotted line*) and normal fit (*dash-dot line*). The *top plot* uses the full range of [0,1] for the vertical scale, the *bottom plot* restricts the vertical scale to [0.9,1]. A *horizontal line* at $p = 0.95$ is added to the *bottom graph* for reference

the projection estimator. In all cases, the R package `tseries` was used to fit the time series model.

We first filtered each component of the DJ30 data with an AR(1) model and then fit the residuals with the elliptical model. When this was done, the residuals had a lighter tailed distribution with $\widehat{\alpha}_{PROJ} = 1.79$ than the raw returns ($\widehat{\alpha}_{PROJ} = 1.71$). This was repeated with an ARMA(1,1) model, and the elliptical fit to the residuals gave $\widehat{\alpha}_{PROJ} = 1.79$ again. Finally, we repeated the process with a GARCH(1,1) model, this time the estimated index of stability was raised to $\widehat{\alpha}_{PROJ} = 1.89$. Comparing the shape matrices is a bit more complicated. The GARCH filtering changes the scale of residuals, and the shape matrices are not directly comparable. To deal with this, we plot the normalized shape matrices: $Q^* = q^*_{ij}$, where $q^*_{ij} = q_{ij}/\sqrt{q_{ii}q_{jj}}$. This correspond to the correlation matrices of the Gaussian covariance. Figure 7 shows the heatmaps of the four models, with the rows and columns are in the same order as in Fig. 5. Figure 8 shows scatterplots of the off-diagonal elements of the normalized shape matrices from the four different models for serial dependence.

Visually, it appears that the AR(1) and ARMA(1,1) model change little from the independent case. The GARCH(1,1) model changes a bit. To quantify this, the mean absolute difference of the entries q^*_{ij} from the independent case was computed: it was 0.019 for the AR(1) model, 0.020 for the ARMA(1,1) model, and 0.036 for the GARCH(1,1) model. While there is some change in the different cases, the serial dependence has only a small effect on the overall shape matrix for this data set. This

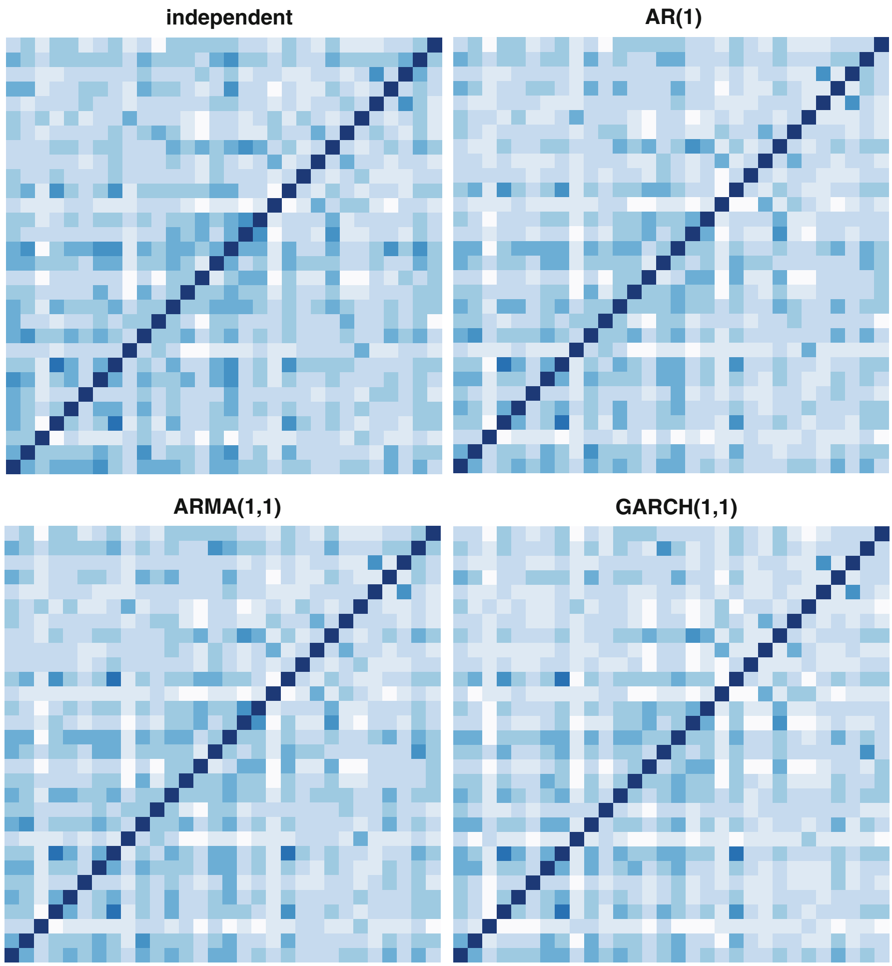

Fig. 7 Heatmaps for the normalized shape matrices Q^* for the four different models of serial dependence: independent, AR(1), ARMA(1,1) and GARCH(1,1). The *shading* scale uses *white* for lowest value ($\min q^*_{ij} = 0.123$) through increasing *darkness* to the highest values ($q^*_{ii} = 1.0$)

suggests that the co-movements of the assets are generally the same whether volatility is high or low.

It is interesting to note that unlike Gaussian models, elliptical stable models with $\alpha < 2$ have positive tail dependence. Equation (5.2) of Schmidt (2002) gives the following formula (correcting a sign misprint) for the tail dependence λ of a elliptical α-stable vector (X_1, X_2):

$$\lambda(\alpha, q^*) := \lim_{v \to 1^-} P(X_2 > G_2^{-1}(v) | X_1 > G_1^{-1}(v)) = \frac{\int_0^{\sqrt{(1+q^*)/2}} \frac{u^\alpha}{\sqrt{1-u^2}}\, du}{\int_0^1 \frac{u^\alpha}{\sqrt{1-u^2}}\, du},$$

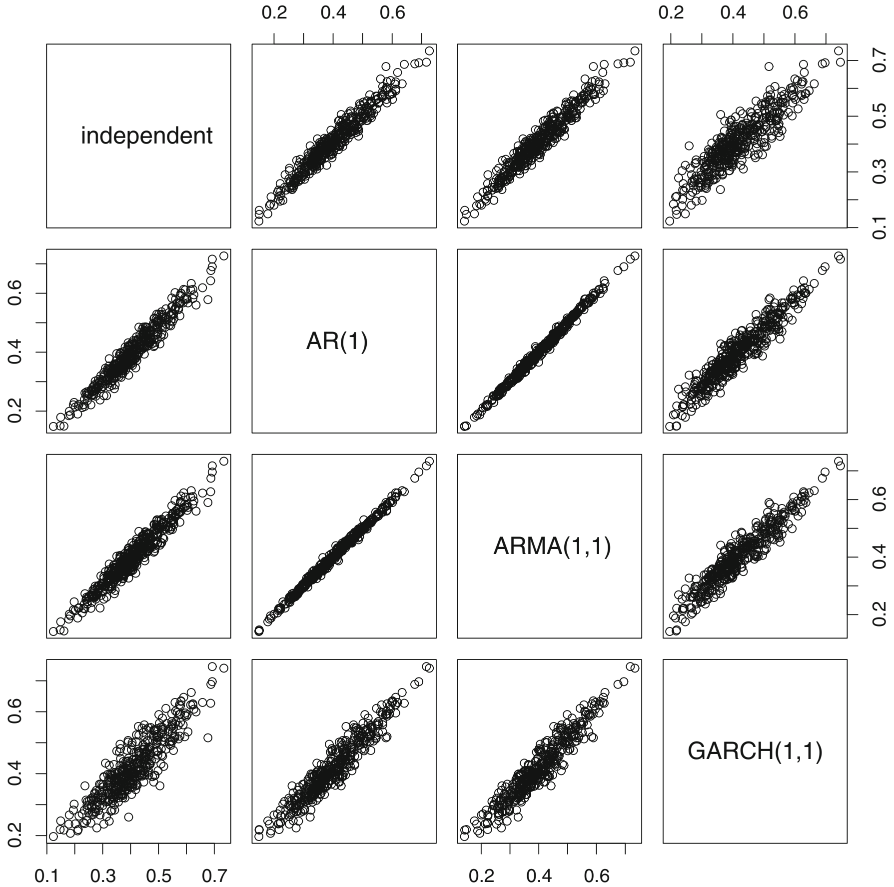

Fig. 8 *Scatterplots* of off-diagonal elements of the normalized shape matrices Q^*

where $G_i^{-1}(\cdot)$ is the inverse cdf of X_i and $q^* = q_{12}/\sqrt{q_{11}q_{22}}$ is the normalized shape coefficient. These expressions for $\lambda(\alpha, q^*)$ were evaluated numerically and plotted in Fig. 9. The observed q^* values for the DJ30 data were in the range 0.10–0.75, with α in the range 1.71 (when no adjustment is made for serial dependence) to 1.89 (when a GARCH(1,1) model is used), this for this data, $\lambda(\alpha, q^*)$ is in the range 0.3–0.7.

4 Signal processing and other applications

There is considerable interest in handling heavy tailed noise in engineering problems. In signal processing applications, one wants to filter out impulsive noise from some signal. Standard linear filters perform poorly in the presence of extreme values. Some basic references to these problems are Nikias and Shao (1995), Kuruoglu and Zerubia (2004), and Nolan et al. (2010). In particular, when there is a radar signal with in-phase and quadrature components, one is explicitly dealing with an isotropic bivariate

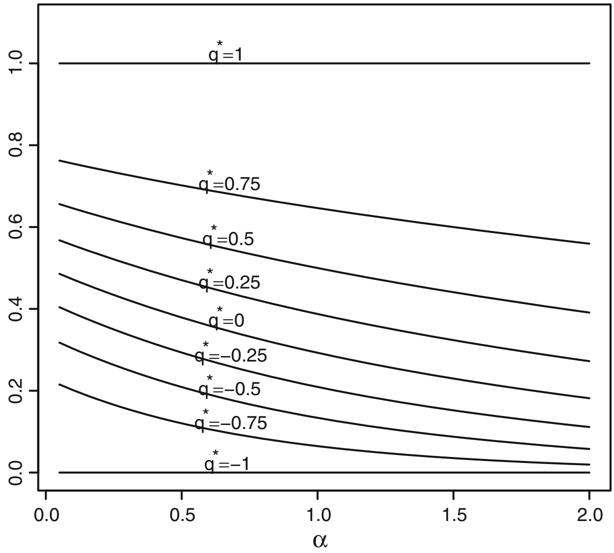

Fig. 9 Tail dependence parameter $\lambda(\alpha, q^*)$ for α-stable elliptical distributions

stable law. The envelope of the noise distribution is precisely the amplitude distribution discussed above. The density of the amplitude derived in (5) is used to derive robust filters for signals in the presence of heavy tailed noise.

A radar clutter data set from the authors of Tsakalides and Nikias (1998) consists of $n = 320{,}000$ pairs of in-phase and quadrature components. Using the programs above to fit a bivariate elliptical model using the PROJ_ML method gave $\widehat{\alpha}_{PROJ} = 1.78$ and

$$Q = \begin{pmatrix} 0.1667 & 0.0009 \\ 0.0009 & 0.16702 \end{pmatrix}.$$

This is strong empirical evidence for an isotropic heavy tailed stable model for the data.

Other examples are in sonar data, where environmental conditions generate impulsive noise that is well modeled by a stable law. In either a radar setting or a sonar setting, if a sensor array with multiple sensors is used to beamform as in Tsakalides and Nikias (1998), the spatial positions of the sensors lead to a multi-dimensional elliptically stable model.

Elliptical stable laws arise in other applications. In astronomy, the Holtsmark distribution is an isotropic stable law with $\alpha = 3/2$ in $\mathbb{R}^3$. And Boldyrev and Gwinn (2003) use stable laws to describe fluctuations in observations of interstellar radiation. Lévy flights in $\mathbb{R}^d$ are typically random walks with steps having a multivariate isotropic stable distribution, e.g. Schlesinger et al. (1995). The methods developed here can be used to work with these and other multi-dimensional problems.

Acknowledgments This work is based on a presentation at a Deutsche Bundesbank Conference at Eltville, Germany on 10–12 November 2005. The author would like to thank Robin Lumsdaine, Robert Jernigan, and Alan Isaac for discussions on computational questions. The projections methods described above are now part of the STABLE program, Robust Analysis Inc (2010).

Appendix A: Additional facts about the amplitude distribution

There are many other facts about the amplitude density and cdf. Since they are useful in finance applications, in signal processing, and astronomy, we collect them here.

Using the Bergstrom series expansions for stable densities in Eqs. (4) and (5) leads to series expansions for $f_R(r)$ and $F_R(r)$: when $0 < \alpha < 1$

$$F_R(r) = 1 - \frac{2}{\pi\alpha\Gamma(d/2)} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}\Gamma\left(\frac{k\alpha+2}{2}\right)\Gamma\left(\frac{k\alpha+d}{2}\right)\sin\left(\frac{k\alpha\pi}{2}\right)}{k\,k!} \left(\frac{r}{2\gamma_0}\right)^{-k\alpha} \tag{18}$$

$$f_R(r) = \frac{1}{\pi\gamma_0\Gamma(d/2)} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}\Gamma\left(\frac{k\alpha+2}{2}\right)\Gamma\left(\frac{k\alpha+d}{2}\right)\sin\left(\frac{k\alpha\pi}{2}\right)}{k!} \left(\frac{r}{2\gamma_0}\right)^{-k\alpha-1} \tag{19}$$

When $1 < \alpha < 2$,

$$F_R(r) = \frac{4}{\alpha\Gamma(d/2)} \sum_{k=0}^{\infty} \frac{(-1)^k\Gamma\left(\frac{2k+d}{\alpha}\right)}{(2k+d)\,k!\,\Gamma\left(\frac{2k+d}{2}\right)} \left(\frac{r}{2\gamma_0}\right)^{2k+d} \tag{20}$$

$$f_R(r) = \frac{2}{\alpha\gamma_0\Gamma(d/2)} \sum_{k=0}^{\infty} \frac{(-1)^k\Gamma\left(\frac{2k+d}{\alpha}\right)}{k!\,\Gamma\left(\frac{2k+d}{2}\right)} \left(\frac{r}{2\gamma_0}\right)^{2k+d-1} \tag{21}$$

When $\alpha < 1$, (18) and (19) converges absolutely for any $r > 0$; when $\alpha > 1$, they are asymptotic series as $r \to \infty$. Likewise, (20) and (21) are absolutely convergent for $\alpha > 1$ and an asymptotic series for $\alpha < 1$ for r near 0.

Let $f_d(r) = f_{R,d}(r)$ be the amplitude density and $F_d(r) = F_{R,d}(r)$ be the amplitude d.f. in d dimensions. An argument using (6) and (7) shows

$$F_{d+2}(r) = F_d(r) - \frac{r}{d} f_d(r) \quad \text{and} \quad f_{d+2}(r) = \frac{d-1}{d} f_d(r) - \frac{r}{d} f_d'(r). \tag{22}$$

One consequence of the latter expression is that the score function for R can be computed without explicitly differentiating:

$$-\frac{d}{dr}\log f_d(r) = -\frac{f_d'(r)}{f_d(r)} = \frac{d-1}{r} - \frac{d\,f_{d+2}(r)}{r f_d(r)}.$$

When $\alpha = 2$, $R^2 = X_1^2 + \cdots + X_d^2 = 2\gamma_0^2 T$, where T is chi-squared with d degrees of freedom. The d.f. and density are $F_R(r) = F_T(r^2/(2\gamma_0^2)) =$

$1 - \Gamma(d/2, r^2/(4\gamma_0^2))/\Gamma(d/2)$ and $f_R(r) = (r/\gamma^2)\ f_T(r^2/(2\gamma_0^2))$. In two dimensions, $R = \sqrt{2}\gamma_0\sqrt{T}$ is a Rayleigh distribution with density and d.f.

$$f_R(r) = \frac{1}{2\gamma_0^2} r e^{-r^2/(4\gamma_0^2)} \quad \text{and} \quad F_R(r) = 1 - e^{-r^2/(4\gamma_0^2)}. \tag{23}$$

(Note that this is not the customary scaling for the Rayleigh, which is based on $\mathbf{X} \sim N(0, \gamma_0^2 I)$ and has density $r/\gamma_0^2 \exp(-r^2/(2\gamma_0^2))$ and d.f. $1 - \exp(-r^2/(2\gamma_0^2))$.)

When $\alpha = 1$, the amplitude density and d.f. have explicit formula in all dimensions. The expressions in dimensions 1, 2 and 3 are:

$$\begin{array}{lll} d=1 & f_R(r) = \frac{2}{\pi}\gamma_0/(\gamma_0^2 + r^2) & F_R(r) = \frac{2}{\pi}\arctan(r/\gamma_0) \\ d=2 & f_R(r) = \gamma_0 r/(\gamma_0^2 + r^2)^{3/2} & F_R(r) = 1 - \gamma_0/\sqrt{\gamma_0^2 + r^2} \\ d=3 & f_R(r) = \dfrac{4\gamma_0}{3\pi}\dfrac{\gamma_0^2 + 2r^2}{(\gamma_0^2 + r^2)^2} & F_R(r) = \dfrac{2}{\pi}\left[\arctan(r/\gamma_0) - \dfrac{\gamma_0 r}{3(\gamma_0^2 + r^2)}\right] \end{array}$$

Expressions for higher dimensions can be found using the recursion relations (22).

The fractional moments of R can be found using (3): if $-d < p < \alpha$,

$$\begin{aligned} E(R^p) &= E|\mathbf{X}|^p = E(AT)^{p/2} = (EA^{p/2})(ET^{p/2}) \\ &= (2\gamma_0)^p \frac{\Gamma(1 - p/\alpha)}{\Gamma(1 - p/2)} \frac{\Gamma((d+p)/2)}{\Gamma(d/2)}, \end{aligned} \tag{24}$$

where the first expectation (which is finite for all for all $p < \alpha$) is from Section 2.1 of Zolotarev (1986); a short calculation is used for the second expectation (which is finite for all $p > -d$). This expression holds for complex p in the strip $-d < \Re p < \alpha$, giving the Mellin transform of R.

The above expression for moments combined with Markov's inequality gives a uniform upper bound on tail probabilities of R and isotropic $\mathbf{X}$:

$$\sup_{r>0} r^p(1 - F_R(r)) = \sup_{r>0} r^p P(|\mathbf{X}| > r) \le E(R^p), \quad 0 < p < \alpha \tag{25}$$

Let X be univariate strictly stable, e.g. $X \sim \mathbf{S}(\alpha, \beta, \gamma, 0)$ with $\alpha \neq 1$ or $X \sim \mathbf{S}(1, 0, \gamma, 0)$. Section 3.6 of Zolotarev (1986) shows $\log|X|$ has mean and variance

$$\begin{aligned} E(\log|X|) &= \gamma_{Euler}\left(\frac{1}{\alpha} - 1\right) + \log\left(\frac{\gamma}{(\cos\alpha\theta_0)^{1/\alpha}}\right) \\ \text{Var}(\log|X|) &= \frac{\pi^2(1 + 2/\alpha^2)}{12} - \theta_0^2 \end{aligned}$$

where $\gamma_{Euler} \approx 0.57721$ is Euler's constant and $\theta_0 = \arctan(\beta\tan(\pi\alpha/2))/\alpha$. (Note the constant θ_0 arises in our expression because Zolotarev uses a different parameterization.) The following is a multivariate generalization of this result, it uses the digamma function $\psi(z) = \Gamma'(z)/\Gamma(z)$.

Lemma 1 $\log R$ *has moment generating function* $E \exp(u \log R) = E\, R^u$ *given by (24) for* $-d < u < \alpha$. *The mean and variance of* $\log R$ *are*

$$E(\log R) = \log(2\gamma_0) + \gamma_{Euler}\left(\frac{1}{\alpha} - \frac{1}{2}\right) + \frac{1}{2}\psi(d/2)$$
$$\mathrm{Var}(\log R) = \frac{\pi^2}{6}\left(\frac{1}{\alpha^2} - \frac{1}{4}\right) + \frac{1}{4}\psi'(d/2).$$

We will not pursue it here, but there are several ways of estimating γ_0 and α from amplitude data: (a) maximum likelihood estimation using $f_R(r)$, (b) fractional moment methods using (24), and (c) using the first and second sample moments of $\log R$ and Lemma 1.

References

Abdul-Hamid H, Nolan JP (1998) Multivariate stable densities as functions of one dimensional projections. J Multivar Anal 67:80–89

Amengual D, Sentana E (2010) Inference in multivariate dynamic models with elliptical innovations. Available at https://www.researchgate.net/publication/228923497 (Preprint)

Boldyrev S, Gwinn C (2003) Scintillations and Lévy flights through the interstellar medium. Phys Rev Lett 91:131101

Bonato M (2011) Modeling fat tails in stock returns: a multivariate stable-GARCH approach. Comput Stat 27(3):499–521

Byczkowski T, Nolan JP, Rajput B (1993) Approximation of multidimensional stable densities. J Multivar Anal 46:13–31

Chambers J, Mallows C, Stuck B (1976) A method for simulating stable random variables. J Am Stat Assoc 71(354), 340–344 (Correction in JASA 82, 704 (1987))

Daul S, DeGiorgi E, Lindskog F, McNeil A (2003) The grouped t-copula with an application to credit risk. RISK 16(11):73–76

Dominicy Y, Ogata H, Veredas D (2010) Quantile-based inference for elliptical distributions. Available at http://ssrn.com/abstract=1006749

Higham NJ (2002) Computing the nearest correlation matrix—a problem from finance. IMA J Numer Anal 22:329–343

Kogon SM, Williams DB (1998) Characteristic function based estimation of stable parameters. In: Adler R, Feldman R, Taqqu M (eds) A practical guide to heavy tailed data. Birkhäuser, Boston, MA, pp 311–338

Kuruoglu E, Zerubia J (2004) Modeling SAR images with a generalization of the Rayleigh distribution. IEEE Trans Image Process 13(4):527–533

Lombardi MJ, Veredas D (2009) Indirect inference of elliptical fat tailed distributions. Comput Stat Data Anal 53:2309–2324

Nikias CL, Shao M (1995) Signal processing with alpha-stable distributions and applications. Wiley, New York

Nolan JP (2001) Maximum likelihood estimation of stable parameters. In: Barndorff-Nielsen OE, Mikosch T, Resnick SI (eds) Lévy processes: theory and applications. Birkhäuser, Boston, pp 379–400

Nolan JP, Gonzalez JG, Núñez RC (2010) Stable filters: a robust signal processing framework for heavy-tailed noise. In: Proceedings of the 2010 IEEE radar conference, pp 470–473

Robust Analysis Inc (2010) User Manual for STABLE 5.0. Software and user manual. Available online at www.RobustAnalysis.com

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

Schlesinger MF, Zaslavsky GM, Frisch U (1995) Lévy flights and related topics in physics. Lecture Notes in physics, No. 450. Springer, Berlin

Schmidt R (2002) Tail dependence for elliptically contoured distributions. Math Methods Oper Res 55(2):301–327

Tsakalides P, Nikias C (1998) A practical guide to heavy tails, chapter deviation from normality in statistical signal processing: parameter estimation with alpha-stable distributions. Birkhäuser, Boston, pp 379–404

Zolotarev VM (1981) Integral transformations of distributions and estimates of parameters of multidimensional spherically symmetric stable laws. In: Contributions to probability. Academic Press, New York, pp 283–305

Zolotarev VM (1986) One-dimensional stable distributions, volume 65 of Translations of mathematical monographs. American Mathematical Society (Translation from the original 1983 Russian edition)