APPROXIMATION OF α-STABLE PROBABILITY DENSITIES USING FINITE GAUSSIAN MIXTURES

*Ercan E. Kuruoğlu**, *Christophe Molina***, *and William J. Fitzgerald**
(*) Signal Processing Laboratory,
Department of Engineering, University of Cambridge,
Trumpington Street CB2 1PZ, Cambridge, UK
(**) Computer Science Department
Anglia Polytechnic University
East Road CB1 1PT, Cambridge, UK.
Tel: 44 1223 363271 ext 2330, e-mail: cmolina@csd.anglia.ac.uk

ABSTRACT

In this paper, we introduce a new analytical model for the α-stable probability density function (p.d.f.). The new model is based on a corollary of the mixing theorem for symmetric α-stable ($S\alpha S$) random variables (r.v.) [1] which states that a $S\alpha S$ r.v. can be expressed as the product of a Gaussian r.v. and a positive-stable r.v. We also extend this model to provide an analytical approximation for a subclass of multivariate α-stable p.d.f.s, namely the sub-Gaussian α-stable p.d.f.s. Simulation results indicate the success of our technique. The new analytical representation opens path to the application of maximum likelihood and Bayesian techniques for problems involving α-stable random variables. The paper is concluded with the examples of possible application areas.

1 INTRODUCTION

Recently, there has been great interest in α-stable distributions for modelling impulsive noise. This interest has been motivated by the experimental evidence that various real-life impulsive noise processes such as atmospheric noise can be successfully modelled with the α-stable distribution [2] and by the generalised central limit theorem which tells that the α-stable distributions are the only possible limits of the sum of infinitely many small processes possibly with infinite variance [3]. Moreover, the Gaussian distribution is a special case of α-stable distributions and they share many nice properties such as *stability*: α-stable random variables are closed under addition. Finally, the α-stable distribution family is parametrised in a very convenient way to cover a wide range of impulsiveness and also skewness.

Despite being such an attractive model for impulsive noise, the α-stable distribution family has received limited attention in the literature since there are no explicit compact expressions for the probability density function except for a few special cases. Therefore, the stable distributions are most conveniently represented by their characteristic functions (which are related to the p.d.f. through a Fourier transform):

$$\varphi(z) = \exp\{j\mu z - \gamma |z|^\alpha [1 + j\beta \,\mathrm{sign}(z)\, w(z,\alpha)]\} \quad (1)$$

where

$$w(z,\alpha) = \begin{cases} \tan\frac{\alpha\pi}{2}, & \text{if } \alpha \neq 1 \\ \frac{2}{\pi}\log|z|, & \text{if } \alpha = 1 \end{cases} \quad (2)$$

and $-\infty < \mu < \infty,\ \gamma > 0,\ 0 < \alpha \leq 2,\ -1 \leq \beta \leq 1$. α, the *characteristic exponent*, is a measure of the thickness of the tails of the distribution. β is the *symmetry parameter*. $\beta = 0$ corresponds to a distribution that is symmetric around μ, in which case the distribution is called *Symmetric α-Stable* ($S\alpha S$). μ is the *location parameter* and for $S\alpha S$ distributions it is the symmetry axis. γ, the *dispersion*, similar to the variance of the Gaussian distribution, is a measure of the deviation around the mean.

Given the characteristic function in Eq.(1), the α-stable p.d.f. can be obtained by taking the inverse Fourier transform of the characteristic function numerically. Although for a large number of samples this is an efficient method, it does not provide an analytic form and it is not suitable for real-time applications due to the extensive numerical integrations involved. Only in the cases of the Gaussian ($\alpha = 2$), the Cauchy ($\alpha = 1, \beta = 0$) and the Pearson ($\alpha = 1/2, \beta = -1$) distributions can the transform be carried out analytically to obtain closed form expressions for the p.d.f. Other than these cases, the α-stable p.d.f. can be expressed only as infinite power series expansions [4]:

$$p_\alpha(x) = \begin{cases} \displaystyle\sum_{k=0}^{\infty} \frac{1}{\pi\alpha}\frac{(-1)^k}{(2k)!}\Gamma\left(\frac{2k+1}{\alpha}\right)x^{2k}, \text{if } 1 \leq \alpha \leq 2, \\ \displaystyle\sum_{k=1}^{\infty} -\frac{1}{\pi}\frac{(-1)^k}{k!}\Gamma(\alpha k+1)\frac{\sin(k\alpha\pi/2)}{x|x|^{\alpha k}}, \\ \qquad \text{if } 0 < \alpha < 1. \end{cases} \quad (3)$$

Asymptotic series are available for $S\alpha S$ density functions with $\alpha > 1$ [4], however, it was shown in [5] that these asymptotic series expansions are good only in the tails and the origin of the p.d.f. and deviate from the actual p.d.f. for the intermediate values.

Tsihrintzis and Nikias suggest an alternative method based on polynomial interpolation [5]. However, our simulations have shown that the expansions calculated with this method are numerically unstable due to significant-bit cancellation caused by the presence of terms with large amplitudes and opposite signs [6]. Ilow suggests using LePage series expansion instead [7], however, it was shown by Janicki and Weron that the convergence of LePage series is extremely slow [8].

In this paper, we introduce a new method, which is numerically very stable, for obtaining analytic expressions for some classes of α-stable p.d.f.s.

2 FINITE-MIXTURE OF GAUSSIANS APPROXIMATION FOR α-STABLE P.D.F.S

2.1 Univariate Symmetric α-Stable Distributions

Our method is based on a corollary of the mixing property of α-stable r.v.s, which states that any $S\alpha S$ r.v. can be represented as the product of a Gaussian r.v. and a positive stable r.v. [1]:

Theorem 1: (Scale Mixtures of Gaussians)

- Let X be distributed with the Gaussian distribution, $X \sim \mathcal{N}(0, 2\gamma_x)$. Also let Y be a positive stable random variable, $Y \sim S_{\alpha_z/2}\left(\gamma = \left(\cos(\frac{\pi\alpha_z}{4})\right)^{2/\alpha_z}, \beta = -1, \mu = 0\right)$ and be independent from X. Then,

$$Z = Y^{1/2}\ X \sim S_{\alpha_z}(\gamma_x, 0, 0). \tag{4}$$

Given that Z is a *compound r.v.*, $Z = Y^{1/2}X$, we can express the p.d.f. of Z in the following way:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{Z|V}(z|v)\ f_V(v)\ J(z,v)\ dv \tag{5}$$

where $f_Z(.)$ and $f_V(.)$ represent the p.d.f.s of the r.v.s Z and $V = Y^{1/2}$ respectively and $J(z,v)$ represents the Jacobian of Z with respect to V. Considering that X is distributed with the standard Gaussian distribution, for a given realisation of $V = v$, $f_{Z|V}(z|v)$ is conditionally distributed with the Gaussian distribution and Eq. (5) can be re-expressed as [9]

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-\frac{z^2}{2v^2}) f_V(v) v^{-1} dv. \tag{6}$$

It follows from Theorem 1 that $Y = V^2$ in Eq. (6) is distributed as $Y \sim S_{\alpha_z/2}\left(\left(\cos(\frac{\pi\alpha_z}{4})\right)^{2/\alpha_z}, -1, 0\right)$.

Probability density functions of Z that can be expressed as in Eq. (6) are called *scale mixtures* of normal distributions; accordingly $f_V(v)$ is called a *mixing function* [9]. Eq. (6) can be sampled uniformly at discrete points to obtain an approximate finite mixture model for an $S\alpha S$ p.d.f. with arbitrary parameters:

$$p_{\alpha,\gamma,\beta=0,\mu}(z) = \frac{\sum_{i=1}^{N} \frac{1}{\sqrt{2\pi} v_i} \exp(-\frac{(z-\mu)^2}{2v_i^2}) f_V(v_i)}{\sum_{i=1}^{N} f_V(v_i)}, \tag{7}$$

where the term in the denominator comes from the normalization to keep the p.d.f. a proper probability function. It should be noted that this analytic expression for the $S\alpha S$ p.d.f. is only an approximation, since the continuous integral was approximated by a finite sum. Therefore, Eq. (6) should be sampled at a large number of points for a good approximation. To reduce the complexity of the model in Eq. (7), one might prefer to use only a small number of components and to sample Eq. (6) at a few points only. In this case, the approximation is coarse and we suggest using the *Expectation-Maximization* (EM) algorithm [10] to fine tune the components to obtain a better approximation.

Hence, our suggested algorithm can be summarized as follows:

Finite-Mixture Approximation to $S\alpha S$ P.D.F.

1. Given the parameters $(\alpha, \gamma, \beta = 0, \mu)$ of the desired stable p.d.f., generate the characteristic function of Y which is positive stable distributed with parameters $\left(\alpha/2, \gamma = \left(\cos(\frac{\pi\alpha_z}{4})\right)^{2/\alpha_z}, \beta = -1, \mu = 0\right)$.
2. Evaluate the positive stable p.d.f. $f_Y(.)$ at N equally spaced points taking the FFT of the characteristic function generated above, where N is the prespecified number of components in the mixture.
3. The mixing function is the p.d.f. of the random variable $V = Y^{1/2}$, which is obtained by
$$f_V(v) = 2v f_Y(v^2). \tag{8}$$
4. Substitute the mixing function samples calculated by Eq. (8) in Eq. (7) as the coefficients of the Gaussian kernels to obtain an approximate analytic expression for the α-stable p.d.f.
5. Use the mixing function samples as initial values of the coefficients of the Gaussian kernels and execute the *expectation-maximization* (EM) algorithm [10] to fine-tune the coefficients.

2.2 Univariate Skewed α-Stable Distributions

Unfortunately, the mixing theorem for representing $S\alpha S$ r.v.s as a product of a Gaussian r.v. and a positive stable r.v. does not extend to the skewed α-stable p.d.f.s, which is also clear from the observation that we cannot express a unimodal non-symmetric function as the summation of a number of unimodal symmetric functions.

2.3 Multivariate α-Stable Distributions

Similar to the univariate case, multivariate stable distributions do not possess a compact analytic form for their probability density functions. Moreover, unlike the characteristic function of univariate α-stable distributions, the characteristic function of the non-Gaussian multivariate stable distributions do not have a simple form as can be seen from:

$$\phi(\mathbf{t}) = \begin{cases} \exp(j\mathbf{t}^T\mathbf{a} - \mathbf{t}^T\mathbf{A}\mathbf{t}), & \text{if } \alpha = 2, (\text{Gaussian}) \\ \exp(j\mathbf{t}^T\mathbf{a} - \int_S |\mathbf{t}^T s|^\alpha \, m(d\mathbf{s}) + j\beta_\alpha(t)), & \\ \qquad \text{if } 0 < \alpha < 2, \end{cases} \tag{9}$$

where

$$\beta_\alpha = \begin{cases} \int_S \mathbf{t}^T \, \mathbf{s} \log|\mathbf{t}^T\mathbf{s}| \; m(d\mathbf{s}), & \text{if } \alpha = 1, \\ \tan(\frac{\alpha\pi}{2}) \; \int_S \; |\mathbf{t}^T \; \mathbf{s}|^\alpha \text{sign}(\mathbf{t}^T\mathbf{s}) \; m(d\mathbf{s}), & \\ \qquad \text{if } \alpha \neq 1, \quad 0 < \alpha < 2, \end{cases} \tag{10}$$

S is the k-dimensional unit sphere, $m(.)$ is a finite Borel measure on S and $\mathbf{A}$ is a positive semi-definite symmetric matrix.

Unlike the Gaussian case, most of the properties of univariate stable variables do not carry to the multivariate stable vectors. One such property is the mixing property.

Although, the mixing theorem is not valid for the general multivariate stable distribution case, a very similar relation exists for a subclass of multivariate stable distributions, namely the sub-Gaussian stable (α-SG) distributions, which are defined with their characteristic function given as:

$$\phi(\mathbf{t}) = \exp(-\frac{1}{2}(\mathbf{t}^T\mathbf{R}\mathbf{t})^{\alpha/2}) \tag{11}$$

where the matrix $\mathbf{R}$ is positive-definite and corresponds to the covariance matrix of the underlying Gaussian multivariate distribution. This relation can be stated as follows [1]:

Theorem 2:

- If $\mathbf{X} \sim \alpha\text{-SG}(\mathbf{R})$ then

$$\mathbf{X} = \eta^{1/2}\mathbf{G} \tag{12}$$

  where η is a scalar positive $\frac{\alpha}{2}$-stable random variable and $\mathbf{G}$ is a Gaussian random vector with mean zero and covariance matrix $\mathbf{R}$. η and $\mathbf{G}$ are independent.

Theorem 2 suggests that a finite-mixture of multivariate Gaussians can be used to approximate the α-SG($\mathbf{R}$) p.d.f. in a similar manner to the univariate case. In the multivariate case, the main difference is that the Gaussian components in the mixture are represented by their covariance matrices which may be non-diagonal. The mixture p.d.f. which determines the amplitudes of the Gaussian mixture components is evaluated in exactly the same way as the univariate case, that is, it is distributed as $\eta \sim S_{\alpha/2}(\gamma = (\cos\frac{\pi\alpha}{4})^{2/\alpha}, \beta = -1, \mu = 0)$.

Then, in analogy to Eq. (6), the following integral can be written for the p.d.f. of an α-SG($\mathbf{R}$) random vector, this time employing multivariate Gaussian kernels in the integral:

$$\begin{aligned} f_{\mathbf{Z}}(\mathbf{z}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{v} \frac{1}{(2\pi)^{dim(\mathbf{z})/2} det(\mathbf{R})} \\ &\quad \times \exp(-\frac{1}{2v^2}\mathbf{z}^T\mathbf{R}\mathbf{z}) f_V(v) dv \end{aligned} \tag{13}$$

where $dim(.)$ stands for the dimension of the vector, $det(.)$ for the determinant and $f_V(v)$ is the mixing function which is the p.d.f. of $\eta^{1/2}$.

Eq. (13) can be sampled uniformly at discrete points to obtain an approximate finite mixture model for a sub-Gaussian α-stable p.d.f. with arbitrary parameters:

$$p_{\alpha,\gamma,\beta=0,\mu}(\mathbf{z}) = \frac{\sum_{i=1}^{N} \frac{1}{v_i} \frac{\exp(-\frac{1}{2v_i^2}\mathbf{z}^T\mathbf{R}\mathbf{z}) f_V(v_i)}{(2\pi)^{dim(\mathbf{z})/2} det(\mathbf{R})}}{\sum_{i=1}^{N} f_V(v_i)}. \tag{14}$$

As in the univariate case, this is only an approximation and when N is restricted to be small, one needs to run the EM algorithm to fine-tune the coefficients of the mixture model. In the multidimensional case, however, EM algorithm is much slower than the one-dimensional case and therefore the computational complexity is much higher. It is therefore, very desirable to convert the problem into a form that we can employ the one-dimensional EM algorithm rather than the multi-dimensional EM.

The covariance matrix of a multivariate Gaussian p.d.f. can be easily diagonalised by a linear transformation and the p.d.f. can be expressed as a product of univariate Gaussian p.d.f.s. One is tempted to look for such a decomposition for the multivariate α-stable case, so that we can express each univariate α-stable p.d.f. in the product as a mixture of Gaussians with our method and end up with a mixture model for multivariate α-stables. Unfortunately, this is not possible in general: firstly, the covariance matrix is not defined for non-Gaussian α-stable r.v.s; secondly, although there exists an analogue in the α-stable case which is the covariation matrix, it does not uniquely determine the p.d.f. Therefore, diagonalising the covariation matrix does not necessarily imply independence and we cannot decompose a multivariate α-stable p.d.f. into a product of univariate α-stable p.d.f.s by a linear transformation.

3 SIMULATION RESULTS AND CONCLUSIONS

To demonstrate the success of our scale mixture of Gaussians model for the $S\alpha S$ p.d.f. we provide here simulation results obtained for an α-stable p.d.f. with parameters $\alpha = 1.0, \beta = 0, \gamma = 1, \mu = 0$ using the mixture approximation (Fig. (2)) and the polynomial-asymptotic series expansion (Fig. (1)). The actual p.d.f.s calculated by directly taking the inverse FFT of the characteristic functions are also provided in the same

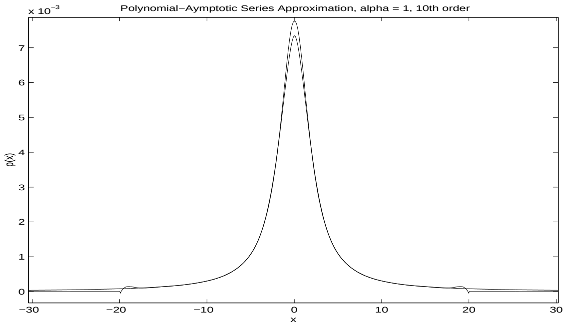

Figure 1: polynomial approximation for standard S(1.0)S vs actual p.d.f.

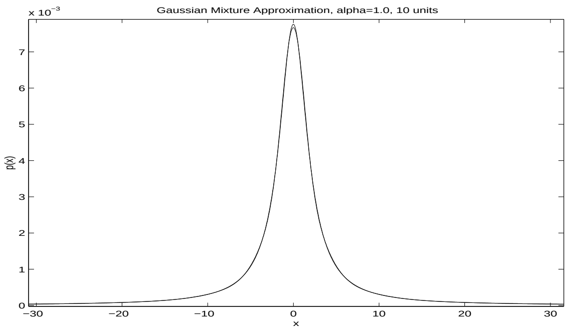

Figure 2: Gaussian mixtures approximation for standard S(1.0)S vs actual p.d.f.

graphs for comparison. For the mixtures approximation, only 10 Gaussian terms are used, similarly, for the polynomial-asymptotic series approximation, a polynomial of order 10 is used and only the first 10 terms of Eq. (3) are taken.

Simulation results show that with a small number of components it is possible to obtain very accurate construction of the α-stable p.d.f. using the scale mixtures of Gaussians approximation while the polynomial-asymptotic series approximation show imperfect fitting at the origin and the tails of the α-stable p.d.f.s. It has been observed that the asymptotic series expansion always leads to poor approximation in the tails. Polynomial fit is good only below a cut-off, if this cut-off point is made bigger for better fit in the tails, the fit around the origin is worsened.

It should be stressed that our algorithm has none of the drawbacks of the Tsihrintzis and Nikias' method: It does not require the predetermination of various parameters and it is numerically very stable. Moreover, some techniques for the solution of some signal processing problems for the case of noise modelled as finite-mixtures of Gaussians have been developed previously and our model enables the signal processing problems involving α-stable data to be addressed in the same framework. Modelling α-stable p.d.f.s as scale mixtures of Gaussians also lends itself to the application of popular signal processing techniques such as maximum-likelihood estimation and Bayesian estimation in problems involving α-stable distributed noise. Examples of such problems are:

- parameter estimation of AR or ARMA models with α-stable innovations,
- Kalman filtering where the process noise and/or the observation noise are α-stable distributed,
- estimation of sinusoids in α-stable noise,
- direction of arrival estimation and bearing estimation in α-stable noise,
- detection of signals in α-stable noise,
- the estimation of the parameters of the α-stable distribution especially in the case of coloured data.

References

[1] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes,* New York, NY: Chapman& Hall, 1994.

[2] C. L. Nikias and M. Shao, *Signal Processing with α-Stable Distributions and Applications,* John Wiley & Sons, 1995.

[3] W. Feller, *An Introduction to Probability Theory and Its Applications,* vol. II, John Wiley & Sons, 1966.

[4] H. Bergström, "On some expansions of stable distribution functions," *Ark. Math.,* vol. 2, pp. 375-378, 1952.

[5] G. A. Tsihrintzis and C. L. Nikias, "Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an α-stable process," *IEEE Trans. Communications,* vol. 43, no. 4, pp. 900-914, April 1995.

[6] E. E. Kuruoğlu, *Signal Processing in α-Stable Noise Environments: A Least L_p-Norm Approach,* PhD Thesis, University of Cambridge, 1998.

[7] J. Ilow, *Signal Processing in α-Stable Noise Environments : Noise Modeling, Detection and Estimation,* PhD Thesis, University of Toronto, 1995.

[8] A. Janicki and A. Weron, *Simulation and Chaotic Behavior of α-Stable Stochastic Processes,* Marcel Dekker, New York, 1994.

[9] D. F. Andrews and C. L. Mallows, "Scale mixtures of normal distributions," *J. Royal Statistical Society,* Vol. B 36, pp. 99-102, 1974.

[10] E. J. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via EM algorithm," *Annals Royal Statistical Society,* Vol. 39, pp. 1-38, Dec. 1977.