INDEPENDENT COMPONENT ANALYSIS USING THE SPECTRAL MEASURE FOR ALPHA-STABLE DISTRIBUTIONS

Preben Kidmose[*]

Informatics and Mathematical Modelling, Richard Petersens Plads, Building 321,
Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark.

ABSTRACT

A method for identifying the independent components of an alpha-stable random vector is proposed. The method is based on an estimate of the spectral measure for the characteristic function. Simulations with both synthetic and speech signals demonstrate that the proposed method can identify independent components in the so-called over-complete case of more sources than mixtures.

Keywords: Alpha-stable distributions, spectral measure, independent component analysis, blind identification, over-complete representation.

1. INTRODUCTION

Independent Component Analysis (ICA) addresses the problem of reconstruction of N sources from the observation of M instantaneous linear combinations thereof.

Stable distributions is suitable for modelling random variables with tails of the probability density function that are *heavier* than the tails of the Gaussian density function. Stable distributions have found applications in signal processing [1], and also in processing of audio signals [2, 3].

This work consider the ICA on stable vectors, and the paper is organized as follows. Stable distributions are introduced in Sec.2 and it is shown that a fundamental property of stable random vectors with independent components is that the so-called spectral measure is discrete (established result). In Sec.3 the estimation of the spectral measure is discussed, and in Sec.4 the spectral measure is used for identifying the number of independent components and the non-orthogonal bases of these components. The results of a simulation study, of both synthetic and speech signals, are presented in Sec.5.

2. STABLE DISTRIBUTIONS

Closed-form expressions do not in general exist for Alpha-Stable (αS) density and distribution functions, therefore the αS distributions are usually described by characteristic functions (ch.f.).

[*]This work is supported by the Danish Research Councils through the THOR Center for Neuroinformatics.

2.1. Characteristic Function for αS Variable

The ch.f. of a random variable x in $\mathbb{R}$ is defined as $\Phi(t) = \mathrm{E}\ \exp(\mathrm{i}tx)$, where E denotes the expectation. The random variable is αS with characteristic exponent α, if the ch.f. can be written on the form

$$\Phi(t) = \exp\left(-I(t) + i\mu t\right) \tag{1}$$

where

$$I(t) = \begin{cases} \gamma|t|^{\alpha}(1 - i\beta\tan(\frac{\alpha\pi}{2})\mathrm{sign}(t)) & , \quad \alpha \neq 1 \\ \gamma|t|(1 + i\beta\frac{\pi}{2}\mathrm{sign}(t)\ln(t)) & , \quad \alpha = 1 \end{cases}$$

and $0 < \alpha \leq 2$, $\gamma > 0$, $-1 \leq \beta \leq 1$, and $(\alpha, \beta, \gamma, \mu) \in \mathbb{R}$. The parameter β determines the symmetry of the distribution, γ is the scale parameter, and μ is the location parameter. For $\beta = 0$ and $\mu = 0$ the distribution is symmetric and is denoted $\mathrm{S}\alpha\mathrm{S}(\alpha, \gamma)$.

2.2. Characteristic Function for αS Vector

The ch.f. for a real random vector $\mathbf{x}$ is defined as $\Phi(\mathbf{t}) = \mathrm{E}\exp(i\mathbf{t}^T\mathbf{x})$. A random vector is an αS vector if a finite measure Γ exists on the unit sphere S_M of $\mathbb{R}^M$ such that the ch.f. can be written on the form

$$\Phi(\mathbf{t}) = \exp\left(-I(\mathbf{t}) + i\mathbf{t}^T\boldsymbol{\mu}\right) \tag{2}$$

where

$$I(\mathbf{t}) = \int_{S_M} \psi_{\alpha}(\mathbf{t}^T\mathbf{s})\ d\Gamma(\mathbf{s})$$

and

$$\psi_{\alpha}(u) = \begin{cases} |u|^{\alpha}(1 - i\tan(\frac{\alpha\pi}{2})\mathrm{sign}(u)) & , \quad \alpha \neq 1 \\ |u|(1 + i\frac{\pi}{2}\mathrm{sign}(u)\ln(u)) & , \quad \alpha = 1 \end{cases}$$

The measure Γ is called the *spectral measure* of the αS random vector, and $\boldsymbol{\mu}$ is called the shift vector. The spectral measure is a real function, and the spectral representation $(\Gamma, \boldsymbol{\mu})$ is unique. See [4, 1].

2.3. αS Vector with Discrete Spectral Measure

Consider the real vector $\mathbf{v} = [v_1 \dots v_N]^T$, where $v_n, n = 1, \dots, N$ are independent $\alpha\mathrm{S}(\alpha, \beta_n, \gamma_n, \mu_n)$ random variables. Let $\mathbf{x} = [x_1 \dots x_M]^T$ be a random vector in $\mathbb{R}^M$, and $\mathbf{A}$ be a $\mathbb{R}^{M \times N}$ matrix. The vector $\mathbf{x} = \mathbf{A}\mathbf{v}$ is then an αS random vector with ch.f.

$$\Phi(\mathbf{t}) = \mathrm{E} \exp\left(i\mathbf{t}^T\mathbf{A}\mathbf{v}\right) = \prod_{n=1}^{N} \mathrm{E} \exp\left(i\mathbf{t}^T\mathbf{a}_{.n}v_n\right)$$

where $\mathbf{a}_{.n}$ is the nth column of $\mathbf{A}$. This is essentially the product of all the ch.f.s of v_n, and can be written as

$$\Phi(\mathbf{t}) = \exp\left(-I_N(\mathbf{t}^T\mathbf{a}_{.n}) + i\sum_{n=1}^{N} \mu_n \mathbf{t}^T\mathbf{a}_{.n}\right) \qquad (3)$$

where

$$I_N(\mathbf{t}^T\mathbf{a}_{.n}) = \sum_{n=1}^{N} \gamma_n \psi_\alpha\left(\mathbf{t}^T\mathbf{a}_{.n}\right) \qquad (4)$$

Rewriting (3) to the form of (2) yields

$$\begin{aligned}\Gamma(\mathbf{s}) = \quad & \sum_{n=1}^{N} \frac{1+\beta_n}{2}\gamma_n(\mathbf{a}_{.n}^T\mathbf{a}_{.n})^{\alpha/2}\delta(\mathbf{s}-\mathbf{s}_n) + \\ & \sum_{n=1}^{N} \frac{1-\beta_n}{2}\gamma_n(\mathbf{a}_{.n}^T\mathbf{a}_{.n})^{\alpha/2}\delta(\mathbf{s}+\mathbf{s}_n)\end{aligned} \qquad (5)$$

where $\mathbf{s}_n = \mathbf{a}_{.n}/(\mathbf{a}_{.n}^T\mathbf{a}_{.n})^{1/2}$ and

$$\boldsymbol{\mu} = \begin{cases} \displaystyle\sum_{n=1}^{N} \mathbf{a}_{.n}\mu_n & , \quad \alpha \neq 1 \\ \displaystyle\sum_{n=1}^{N} \mathbf{a}_{.n}\left(\mu_n - \frac{1}{\pi}\sigma_n\beta_n \ln(\mathbf{a}_{.n}^T\mathbf{a}_{.n})\right) & , \quad \alpha = 1 \end{cases}$$

See [4, Example 2.3.6]. Hence the spectral measure $\Gamma(\mathbf{s})$ of the αS random vector $\mathbf{x}$ is discrete and concentrated on N symmetric pairs of points $(\mathbf{s}_n, -\mathbf{s}_n)$, $n = 1, \dots, N$ of S_N. This result holds in general, thus the spectral measure $\Gamma(\mathbf{s})$ of an αS random vector $\mathbf{x}$ is discrete on the unit sphere S_N if, and only if, the αS random vector can be expressed as a linear transformation of independent αS random variables.

3. ESTIMATION OF THE SPECTRAL MEASURE

We describe a method for estimation of the spectral measure, examine the statistical characteristics of the estimator, and discus the sampling radius of the ch.f..

3.1. Discrete Approximation of the Spectral Measure

From [5, Theorem 1] we have that for a d-dimensional αS variable with spectral measure $\Gamma(\mathbf{s})$ and density function $p(\mathbf{x})$, there is a discrete spectral measure $\Gamma_a(\mathbf{s})$ with corresponding density function $p_a(\mathbf{x})$ satisfying

$$\sup_{\mathbf{x}\in\mathbb{R}^d} |p(\mathbf{x}) - p_a(\mathbf{x})| \leq \epsilon$$

where $\epsilon > 0$. Thus an arbitrary spectral measure $\Gamma(\mathbf{s})$ can be approximated by a discrete measure $\Gamma_a(\mathbf{s})$ such that the corresponding densities are arbitrarily close. The only requirement is that sampling of $\mathbf{s}$ is sufficiently dense.

The ch.f. corresponding to the approximated discrete spectral measure $\Gamma_a(\mathbf{s})$ sampled in L points, can be written as

$$\Phi_a(\mathbf{t}) = \exp\left(-\sum_{n=1}^{L} \psi_\alpha(\mathbf{t}^T\mathbf{s}_n)\Gamma_a(\mathbf{s}_n)\right)$$

Now define a vector $\boldsymbol{\Gamma} = [\Gamma_a(\mathbf{s}_1) \dots \Gamma_a(\mathbf{s}_L)]^T$ containing the L values of the approximated discrete spectral measure. If we evaluate the approximated ch.f. for L values of $\mathbf{t}$ and define the function $I_a(\mathbf{t}) = -\ln \Phi_a(\mathbf{t})$ then we can formulate the set of linear equations

$$\mathbf{I} = \boldsymbol{\Psi}\boldsymbol{\Gamma} \qquad (6)$$

where

$$\mathbf{I} = \begin{bmatrix} I_a(\mathbf{t}_1) \\ \vdots \\ I_a(\mathbf{t}_L) \end{bmatrix}, \quad \Psi = \begin{bmatrix} \psi_\alpha(\mathbf{t}_1^T\mathbf{s}_1) & \dots & \psi_\alpha(\mathbf{t}_1^T\mathbf{s}_L) \\ \vdots & \ddots & \vdots \\ \psi_\alpha(\mathbf{t}_L^T\mathbf{s}_1) & \dots & \psi_\alpha(\mathbf{t}_L^T\mathbf{s}_L) \end{bmatrix}$$

If the L values of $\mathbf{t}$ are chosen so that Ψ^{-1} exists, then the approximated spectral measure is given exact by the solution to (6).

3.2. Estimation of the Discrete Spectral Measure

In [6] a method for estimation of the stable spectral measure is proposed. The principle behind the estimation is based on (6). From this point, without loss of generality and to simplify the presentation, we will assume that $\mathbf{x}$ is SαS. For estimation of the spectral measure in the non-symmetric case see the more general treatment in [6].

In the case of symmetric density function the spectral measure and the ch.f. is real valued and symmetric. From the definition of the ch.f. an estimate based on samples of the random vector $\mathbf{x}$ can be obtained as

$$\hat{\Phi}(\mathbf{t}) = \frac{1}{K}\sum_{k=1}^{K} \cos(\mathbf{t}^T\mathbf{x}_k) \qquad (7)$$

for K samples of $\mathbf{x}$. By introducing the function $\hat{I}(\mathbf{t}) = -\ln(\hat{\Phi}_\alpha(\mathbf{t}))$ and substitute $\mathbf{I}$ by $\hat{\mathbf{I}} = [\hat{I}(\mathbf{t}_1) \dots \hat{I}(\mathbf{t}_L)]^T$, an estimate of the approximate discrete spectral measure is directly obtained from (6).

The spectral measure is defined on the d-dimensional unit sphere. If no a priori knowledge about the spectral measure is available and if all directions are of equal importance, it is natural to sample $\mathbf{s}$ uniformly on the unit sphere. In the SαS case it suffice to sample on the half d-dimensional unit sphere. In the $d = 2$ case the nth sample point in $\mathbf{s}$ is then $\mathbf{s}_n = [\cos\left(\pi\frac{n-1}{L}\right) \quad \sin\left(\pi\frac{n-1}{L}\right)]^T$.

The natural choice of the sampling grid of the ch.f. is to sample symmetrically on a d-dimensional sphere, again in the SαS case it suffices to sample on a half d-dimensional sphere. Thus the nth sample point in $\mathbf{t}$ is $\mathbf{t}_n = r\ \mathbf{s}_n$, where r is the sampling radius. The sampling radius is decisive for the performance of the estimator of the spectral measure, this is discussed in section 3.4.

3.3. Estimation Procedure

To avoid negative values in the estimate of the spectral measure $\mathbf{\Gamma}$, the solution in (6) is restated as

$$\min \left\| \hat{\mathbf{I}} - 2\mathrm{Re}(\mathbf{\Psi})\mathbf{\Gamma} \right\|_2 \quad \text{s.t.} \ \ \mathbf{\Gamma} \geq \mathbf{0} \tag{8}$$

The estimation procedure for the spectral measure is now: determine the sampling radius of $\mathbf{t}$, calculate $\mathbf{\Psi}$, estimate $\hat{\mathbf{I}}$, and solve the constrained least square problem in (8).

3.4. Estimation Considerations

This section examine the statistical characteristics of the estimator subject to the sampling radius of the ch.f., and the choice of the characteristic exponent. To evaluate the performance of the estimation of the spectral measure we define the performance measure

$$\Upsilon = \mathrm{E}\ (\mathbf{I} - \hat{\mathbf{I}})^{\mathrm{T}}(\mathbf{I} - \hat{\mathbf{I}})/(\mathbf{I}^{\mathrm{T}}\mathbf{I})$$

where $\mathbf{I}$ is defined in (6) and $\hat{\mathbf{I}}$ is the estimate of $\mathbf{I}$ that corresponds to the estimate of the spectral measure from (8).

Applying the estimation procedure from section 3.3, we estimate the spectral measure for the SαS random vector $\mathbf{x} \in \mathbb{R}^2$ with the discrete spectral measure

$$\Gamma(\mathbf{s}_n) = \begin{cases} \frac{\gamma}{2} & \text{for } \ n = 11, 36\ (101, 126) \\ 0 & \text{else} \end{cases}$$

where $\mathbf{s_n}$ is distributed uniformly on the half unit circle, and $L = 90$.

A desirable property is to have a scale independent estimator, and from (2) one directly obtain that if $r \propto \gamma_x^{-1/\alpha}$ then the estimator of the ch.f. is scale independent[1].

[1] In [6] the sampling radius is fixed to $r = 1$, and the influence of the sampling radius is not addressed.

In Fig.1 the performance measure versus the sampling radius is shown for $\gamma = 0.01, 0.1, 1, 10, 100$. It appears that

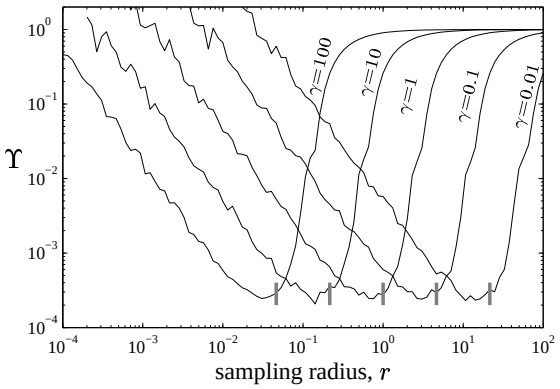

Fig. 1. Performance measure versus sampling radius, $\mathbf{x} \in \mathrm{S}\alpha\mathrm{S}(1.5, \gamma)$.

the sampling radius of the ch.f. is of crucial importance. A natural choice of the sampling radius is

$$r = \gamma_x^{-1/\alpha} \tag{9}$$

and the vertical bars in Fig.1 indicate the sampling radius that corresponds to (9).

In Fig.2 the performance measure versus the choice of the characteristic exponent is shown for $\mathbf{x} \in \mathrm{S}\alpha\mathrm{S}(\alpha, 1)$, and $\alpha = 1.2, 1.4, 1.6, 1.8$. Obviously the best performance is

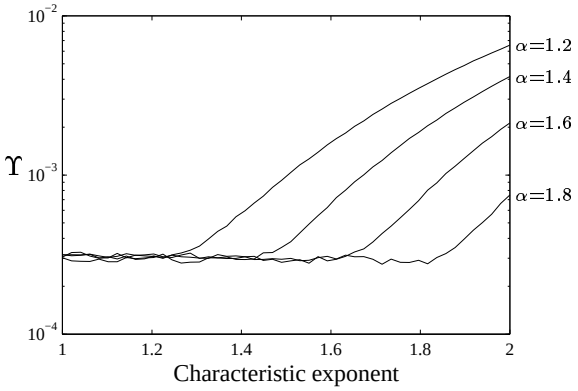

Fig. 2. Performance measure versus the choice of characteristic exponent in the estimation. $\mathbf{x} \in \mathrm{S}\alpha\mathrm{S}(\alpha, 1)$.

obtained when the choice of characteristic exponent in the estimation is equal to the characteristic exponent of the SαS vector. It is surprising that the characteristic exponent has only minor influence on the performance of the estimator, and notice that the characteristic exponent in the estimation should be chosen conservatively low.

4. BLIND IDENTIFICATION OF BASIS VECTORS

Consider the mixture $\mathbf{x} = \mathbf{A}\mathbf{v}$, where $\mathbf{v} = [v_1 \dots v_N]^T$ is a random vector with independent $\alpha\mathrm{S}(\alpha, \beta_n, \gamma_n, \mu_n)$ components, and $\mathbf{A} \in \mathbb{R}^{M\times N}$. The objective is to estimate the basis vectors of $\mathbf{A}$ based on observations of $\mathbf{x}$.

Fundamentally blind identification has a scale ambiguity between γ_n of v_n and the norm of the columns $\|\mathbf{a}_{.n}\|_2$, thus without loss of generality we can assume that $\|\mathbf{a}_{.n}\|_2 = 1$. Moreover, blind identification provides no ordering of the basis vectors, and for the case of symmetric densities there is a sign ambiguity of the basis vectors.

Now considering (5) leads to the conclusion that the identification of the basis vectors $\mathbf{a}_{.n}$ is simply to determine the directions in which the spectral measure has mass[2]. In general the spectral measure has mass in $2N$ directions (N directions if we only sample the half d-dimensional sphere). The method will work for both the under-complete ($M > N$) and the over-complete ($M < N$) case.

4.1. Practical Considerations

Due to a finite number of samples, observation noise and possible deviations from the theoretical distributions, there will be some noise in the estimated spectral measure. The basis vectors should be determined as the directions in which the estimated spectral measure has dominating mass.

The number of sample points L in $\mathbf{t}$ and $\mathbf{s}$ determines the angular resolution of the basis vectors. In the $d = 2$ case with sampling on the half 2-dimensional sphere the resolution is $\pm\pi/(2L)$. A disadvantage of the proposed method is that the optimization problem in (8) scales exponentially in the dimension of the mixture.

5. SIMULATION STUDY

The basis vectors for over-complete mixtures ($M < N$) are identified for both synthetic SαS signals and speech signals.

5.1. Identification of Basis Vectors, SαS Signal

Let $\mathbf{v}$ be a random vector with independent $\mathrm{S}\alpha\mathrm{S}(1.2, 1)$ random variables, $\mathbf{x} = \mathbf{A}\mathbf{v}$ be the observable random vector, and

$$\mathbf{A} = \begin{bmatrix} \cos(\theta_1) & \cos(\theta_2) & \cos(\theta_3) & \cos(\theta_4) \\ \sin(\theta_1) & \sin(\theta_2) & \sin(\theta_3) & \sin(\theta_4) \end{bmatrix}$$

with $\theta_1 = (\frac{\pi}{9})$, $\theta_2 = (\frac{2\pi}{9})$, $\theta_3 = (\frac{3\pi}{9})$, and $\theta_4 = (\frac{4\pi}{9})$.

The basis vectors are identified as the directions in which the estimated spectral measure has significant mass. In Fig.3 a typical example of an estimated spectral measure, using the estimation procedure in Sec.3, is shown. Observe that the distribution of masses for the spectral measure is very specific in the 4 directions corresponding to the directions of the basis of $\mathbf{A}$. In Fig.4 the scatter plot for the observations of $\mathbf{x}$ is depicted, and the basis vectors of the mixing matrix and the directions in which the spectral measure has dominating mass are shown. Three basis vectors are correctly identified, and the fourth basis vector is one point away from the right value.

[2] On the contrary hereto, conventional ICA algorithms is based on minimization of contrast functions.

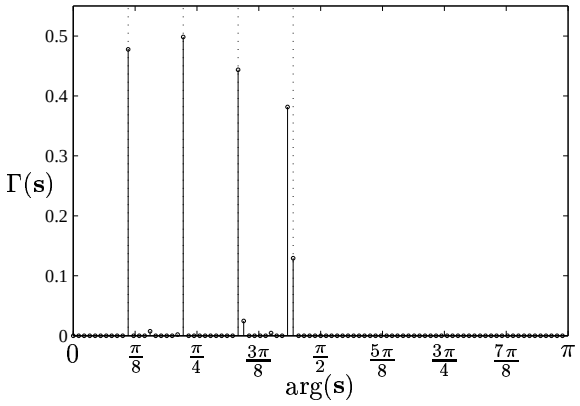

Fig. 3. Empirical spectral measure, $\Gamma(\mathbf{s})$, for the SαS random vector $\mathbf{x}$. The estimation is based on $K = 10000$ observations of $\mathbf{x}$, sampling radius $r = \frac{1}{2}$, and $L = 90$. The direction of the basis vectors of $\mathbf{A}$ are indicated by dotted lines.

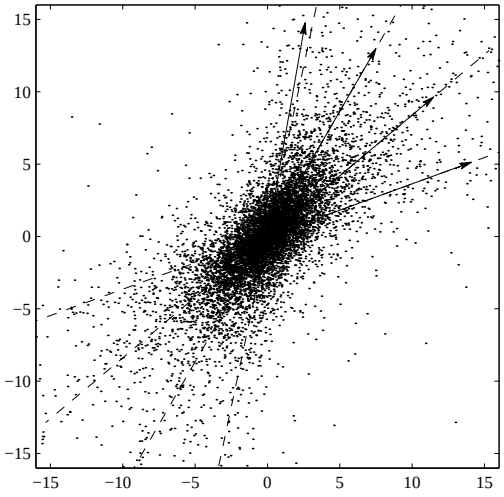

Fig. 4. Scatter plot of $\mathbf{x}$. The basis vectors of $\mathbf{A}$, are depicted by vectors, and the directions in which the spectral measure has significant masses, are depicted by dashed lines.

5.2. Identification of Basis Vectors, Speech Signals

The proposed method for identification of bases in a mixture is applied to a mixture of three speech signals. In [3] it

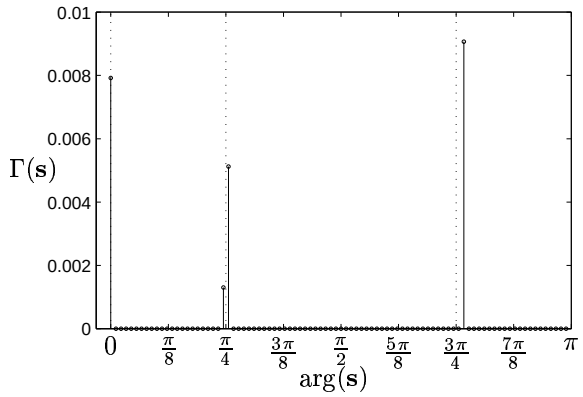

Fig. 5. Empirical spectral measure, $\Gamma(\mathbf{s})$, for the mixture $\mathbf{x}$. The ch.f. is estimated based on $N = 23000$ observations, sampling radius $r = 50$, $\alpha = 1.2$, and $L = 90$.

is demonstrated that SαS distributions are suitable for modelling a broad class of acoustical signals, including speech signals, and this is the theoretical justification for the example. For the purpose of comparison the method is applied to the same mixing matrix and the same speech signals as used in [7]. The speech signals are from the TIMIT database, and the mixing matrix is $\mathbf{A} = [1 \quad 1/\sqrt{2} \quad 1/\sqrt{2}\,;\; 0 \quad 1/\sqrt{2} \quad -1/\sqrt{2}]$. In Fig.5 the estimated spectral measure is depicted. The distribution of masses is very specific in the directions corresponding to directions of the basis vectors of $\mathbf{A}$. In Fig.6 the scatter plot for the observations is depicted. The spectral measure has dominant masses in $(2\times)3$ symmetric directions, and these directions are, very close to the directions of the basis vectors of $\mathbf{A}$.

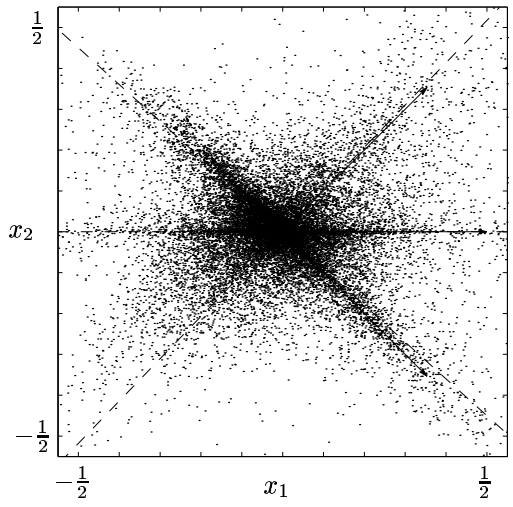

Fig. 6. Scatter plot of $\mathbf{x}$. The directions of the basis vectors of $\mathbf{A}$ are depicted by vectors, and the directions in which the spectral measure has masses, are depicted by dashed lines

6. CONCLUSION

We proposes an ICA method based on the observation that the spectral measure is discrete for stable random vectors with independent components. The method identifies the number of independent components and the non-orthogonal bases of the mixture.

We examine the properties of the estimator of the spectral measure, and discus the sampling radius for the ch.f..

Simulations on synthetic and speech signals demonstrates that the method can identify the number of independent components and the bases of the overcomplete mixtures.

Acknowledgements

Thanks to J.P.Nolan for making [6] available prior to publication, and to Lars Kai Hansen for discussions and comments on the manuscript.

7. REFERENCES

[1] Chrysostomos L. Nikias and Min Shao, *Signal Processing with Alpha-Stable Distribution and Applications*, Wiley, 1 edition, 1995.

[2] P. G. Georgiou, P. Tsakalides, and C. Kyriakakis, "Alpha-Stable Modeling of Noise and Robust Time-Delay Estimation in the Presence of Impulsive Noise," *IEEE Transactions on Multimedia*, vol. 1, no. 3, pp. 291–301, September 1999.

[3] P. Kidmose, "Alpha-Stable Distributions in Signal Processing of Audio Signals," in *41st Conference on Simulation and Modelling, Scandinavian Simulation Society*, 2000, pp. 87–94.

[4] Gennady Samorodnitsky and Murad S. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman&Hall, 1994.

[5] T. Byczkowski, J. P. Nolan, and B. Rajput, "Approximation of Multidimensional Stable Densities," *Journal of Multivariate Analysis*, vol. 46, no. 1, pp. 13–31, 1993.

[6] J. P. Nolan, A. K. Panorska, and J. H. McCulloch, "Estimation of Stable Spectral Measures," *To appear in Mathematical and Computer Modelling*, 2001, Available from http://www.cas.american.edu/˜jpnolan/.

[7] T. W. Lee, M. S. Lewicki, M. Girolami, and T. J. Sejnowski, "Blind Source Separation of More Sources Than Mixtures Using Overcomplete Representations," *IEEE Signal Processing Letters*, vol. 6, pp. 87–90, April 1999.