Identifiability of Finite Mixtures - with Applications to Circular Distributions

Hajo Holzmann, Axel Munk and Bernd Stratmann
Institute for Mathematical Stochastics
Georg-August-University Göttingen
Germany

Abstract

A general result about identifiability and strong identifiability of finite mixtures of a family of distributions is obtained via tail conditions on the corresponding characteristic functions. This is applied to location-scale families on the real line and to circular distributions. Particular cases include circular wrapped distributions of location-scale families, stable distributions and the d-dimensional wrapped normal distribution. Finally, counter examples are given which highlight differences between identifiability on the real line and on the circle.

Running title: Identifiability of finite mixtures

AMS 2000 subject classifications. Primary 62E10; secondary 62F03, 62F05.

Key words and phrases: Circular distributions; finite mixture; wrapped distributions; identifiability; characteristic function; location-scale family; Fourier transform; stable distributions

1 Introduction

Identifiability is a task of general interest in the theory of mixture models, see e.g. Teicher (1961) and Lindsay (1995), but also has applications to other fields (Leroux 1992 and Bickel et al. 1998). Finite mixtures of continuous densities were first studied by Teicher (1963). Most results for this problem are obtained for distributions on the real line or the d-dimensional Euclidean space. For distributions on the circle, only few examples are known.
Identifiability of finite mixtures of von Mises distributions was first proved by Fraser et al. (1981). Kent (1983) extended this result to certain generalized von Mises distributions. While Fraser et al. (1981) used the real part

of the characteristic functions in its proof, Kent (1983) argued directly via the tail behaviour of the densities.
Here we extend the result of Fraser et al. (1981) in another direction which allows us to treat wrapped distributions, among others. Our method is based on a condition on the tail behaviour of the Fourier transform which is adapted from Teicher (1963). The Fourier transform turns out to be particularly useful for circular distributions. For example, wrapping leaves the Fourier transform almost unchanged and hence identifiability results from the real line can be directly transferred to wrapped distributions.
Our results apply to certain location-scale families on the real line as well as to the corresponding wrapped families on the circle. We also obtain identifiability of arbitrary mixtures of some location families on the circle. Counter examples are given which highlight differences between identifiability of distributions on the circle and on the real line. If only either location or scale families on the real line are considered, identifiability (even of arbitrary mixtures) was established by Teicher (1961). However, let us stress that general location-scale families have not been treated in the literature so far. Exceptions are normal and Cauchy mixtures (cf. Teicher (1963) and Yakowitz and Spragins (1968)), however, the technique used there cannot be transfered to general location-scale families.
We also consider strong identifiability as introduced by Chen (1995), and show that location-scale mixtures from the normal and the Cauchy distribution are strongly identifiable. Therefore Chen's (1995) result on the speed of convergence in estimating the mixing distribution and the asymptotics for the modified likelihood ratio statistic of Chen, Chen and Kalbfleisch (2001) for testing homogeneity apply to such mixtures.
The paper is organized as follows. In Section 2 the main theorem is proved. In Section 3.1 this is applied to several distributions on the real line and on the circle. Extensions to multivariate distributions are briefly discussed in Section 3.2. This includes the multidimensional wrapped normal distribution and a distribution on the cylinder as introduced by Johnson and Wehrly (1978). Finally, in Section 3.3. strong identifiability is considered.

2 Identifiability of mixtures - the main theorem

Finite mixtures of a class of continuous densities $\{f_\alpha : \alpha \in \Omega\}$ are identifiable if for any distinct set of parameters $\alpha_1, \ldots, \alpha_m,\ m \geq 1$, the functions $f_{\alpha_j},\ j = 1, \ldots, m$, are linearly independent (Yakowitz and Spragins (1968), p. 210). Throughout the following let $\{f_\alpha : \alpha \in \Omega\}$ be either a two parameter family of continuous densities on the real line, $f_\alpha(x),\ \alpha = (\mu, a) \in \Omega = \mathbb{R} \times \mathbb{R}_+,\ x \in \mathbb{R}$, or on the circle, $f_\alpha(\theta),\ \alpha = (\mu, a) \in \Omega = [0, 2\pi) \times \mathbb{R}_+,\ \theta \in [0, 2\pi)$. For densities $f_\alpha(x)$ on the real line (respectively $f_\alpha(\theta)$ on the circle) the characteristic function will be denoted by $\Phi_\alpha(t),\ t \in \mathbb{R}$ (respectively

$\phi_\alpha(n),\ n \in \mathbb{Z})$. Now we are in the position to state our main result.

Theorem 1. *Let $\{f_\alpha : \alpha \in \Omega\}$ be a two-parameter family of continuous densities on the real line or on the circle, respectively. Suppose that the characteristic function of f_α is of the form*

$$\Phi_\alpha(t) = \exp(i\mu t) \cdot h(t,a), \quad t \in \mathbb{R}, \tag{1}$$

or

$$\phi_\alpha(n) = \exp(i\mu n) \cdot h(n,a), \quad n \in \mathbb{Z},$$

where $h(\cdot,a)$ is a function which satisfies

$$\lim_{t\to\infty} h(t,a_2)/h(t,a_1) = 0, \quad a_2 > a_1, \tag{2}$$

or

$$\lim_{n\to\infty} h(n,a_2)/h(n,a_1) = 0, \quad a_2 > a_1. \tag{3}$$

Then finite mixtures of $\{f_\alpha,\ \alpha \in \Omega\}$ are identifiable.

For the proof we will require the following simple lemma. Note that it will be important to exclude 0 from the support of the measure ν.

Lemma 1. *Let ν be a finite signed Borel measure on $(0,2\pi)$. Then for $b_n = \int_{(0,2\pi)} e^{inx}\, d\nu(x),\ n \geq 0$, we have that $n^{-1}\sum_{l=0}^{n-1} b_l \to 0$.*

Proof. Since

$$n^{-1}\sum_{l=0}^{n-1}\int_{(0,2\pi)} e^{ilx}\, d\nu(x) = \int_{(0,2\pi)} n^{-1}\frac{1-e^{inx}}{1-e^{ix}}\, d\nu(x)$$

and $|\sum_{l=0}^{n-1} e^{ilx}/n| \leq 1$, we can apply the dominated convergence theorem to obtain the result. □

In particular, if $\nu = \sum_{j=1}^{k} \lambda_j\, \delta_{\mu_j}$, where $0 < \mu_1, \ldots, \mu_k < 2\pi,\ \lambda_j \in \mathbb{R},\ j = 1, \ldots, k$ and δ_μ denotes the Dirac measure at μ, then

$$n^{-1}\sum_{l=0}^{n-1}\sum_{j=1}^{k} \lambda_j \exp(i\mu_j l) \to 0, \quad n \to \infty. \tag{4}$$

Proof of Theorem 1. First let us consider the slightly more simple case of densities on the circle. Fix $m \in \mathbb{N}$ and suppose that $\sum_{j=1}^{m} \lambda_j f_{\alpha_j} = 0$ and hence

$$\sum_{j=1}^{m} \lambda_j \exp(i\mu_j n) \cdot h(n,a_j) = 0 \quad \forall\ n \in \mathbb{Z}, \tag{5}$$

where $\alpha_j = (\mu_j, a_j)$ are distinct. We may arrange the α_j such that $a_1 \leq a_2 \leq \ldots \leq a_m$ and $a_1 = a_2 = \ldots = a_k < a_{k+1}$. The case $a_1 = \ldots = a_m$ is treated implicitly. Upon multiplying (5) by $e^{-in\mu_1}$ we can assume $\mu_1 = 0$. From (5) it follows that

$$\sum_{j=1}^{k} \lambda_j \exp(i\mu_j n) + \sum_{j=k+1}^{m} \lambda_j \exp(i\mu_j n) \cdot h(n, a_j)/h(n, a_1) = 0.$$

Since by (3) the second sum tends to 0 as $n \to \infty$, so must the first one. If $k = 1$, we immediately conclude $\lambda_1 = 0$ (remember $\mu_1 = 0$). If $k \geq 2$, setting

$$b_n = \sum_{j=2}^{k} \lambda_j \exp(i\mu_j n)$$

we have that $b_n \to -\lambda_1$ as $n \to \infty$. Then also $n^{-1} \sum_{l=0}^{n-1} b_l \to -\lambda_1$. Since the α_j are all distinct, at most one μ_j, $1 \leq j \leq k$, can be 0, this being μ_1. Thus $\mu_j \neq 0$, $2 \leq j \leq k$, and using (4) we can conclude $\lambda_1 = 0$. The claim for densities on the circle follows by induction.
As for densities on the real line, using (2) we argue similarly and may choose k in the above way. If $k \geq 2$, $\sum_{j=2}^{k} \lambda_j \exp(i\mu_j t) \to -\lambda_1$, $t \to \infty$, where now $\mu_j \in \mathbb{R} \backslash \{0\}$, $j = 2, \ldots, k$. Choose $t_0 > 0$ so small such that $t_0\mu_j \in (-\pi, \pi) \setminus \{0\}$, $j = 2, \ldots, k$. Then in particular $\sum_{j=2}^{k} \lambda_j \exp(i[\mu_j t_0]n) \to -\lambda_1$, $n \to \infty$, and using (4) once more we conclude $\lambda_1 = 0$. $\square$

Remark 1. Theorem 1 and its proof are inspired by a result of Teicher (1963, p. 1267) on general integral transforms. See also Chandra (1977) or Al-Hussaini and El-Dab Ahmad (1981) for extensions and applications. Our proof extends Teicher's argument to Fourier transforms of type (1), which have an additional factor arising from a location parameter. In this case identifiability relies on Lemma 1.

3 Examples

3.1 Identifiability on the real line and on the circle.

Example 1 (*Location-scale families***).** Consider a location-scale family of densities on the real line

$$f_\alpha(x) := f_{\mu,a}(x) = a^{-1} f([x - \mu]/a), \tag{6}$$

where $\mu \in \mathbb{R}$, $a > 0$. If only either the scale or the location parameter are allowed to vary in the mixture, identifiability (even for arbitrary mixtures) was established by Teicher (1961, p. 246) (see also Yakowitz and Spragins (1968), for the case of finite mixtures of location families). By means of

Theorem 1 identifiability can be proved when both location and scale parameter are allowed to vary simultaneously. To this end suppose that the characteristic function Φ of f satisfies

$$\lim_{t\to\infty} \Phi(a_2 t)/\Phi(a_1 t) = 0, \quad a_2 > a_1 > 0. \tag{7}$$

Then finite mixtures are identifiable. Indeed, the characteristic function of $f_{\mu,a}$ is given by $\Phi_{\mu,a}(t) = e^{i\mu t}\Phi(at),\ t \in \mathbb{R}$. Evidently condition (7) implies (2). Ishwaran (1996) considered identifiability and rates of convergence of a joint scale parameter in a location mixture.

Remark 2. Condition (7) does not have an immediate interpretation. A slightly stronger condition is the following. Suppose that the characteristic function of f satisfies $\Phi(t) \sim \exp(-\log(t)\cdot s(t)), t \to \infty$, where s is eventually non-decreasing and $s(t) \to \infty$. Then condition (4) is satisfied.

Example 2 (*Stable distributions*). Symmetric stable distributions (Jammalamadaka and SenGupta 2001, p. 46) have characteristic function $\Phi(t) = \exp(-|t|^\alpha),\ \alpha \in (0, 2]$ and therefore satisfy condition (7). It follows that finite mixtures of the corresponding location-scale family are identifiable. Particular cases are the normal distribution ($\alpha = 2$) and the Cauchy distribution ($\alpha = 1$) (Yakowitz and Spragins (1963), p. 212).

In the next example we consider families of densities on the circle obtained by wrapping location-scale families on the real line. Recall that for a real density f the corresponding wrapped density f^w is given by $f^w(\theta) = \sum_{n\in\mathbb{Z}} f(\theta + 2\pi n),\ \theta \in [0, 2\pi)$ (Mardia and Jupp (2000), pp. 47-49).

Example 3 (*Wrapped distributions*). We show that finite mixtures of the family of wrapped densities of a location scale family (6) on the real line, for which condition (7) is satisfied, are also identifiable. Note that for the wrapped densities, $f^w_{\mu,a} = f^w_{\mu+2\pi n,a},\ n \in \mathbb{Z}$. Therefore the wrapped family can be parametrized by $\{f^w_\alpha = f^w_{\mu,a},\ \mu \in [0, 2\pi),\ a > 0\}$. Since the characteristic function $\phi_{\mu,a}$ of $f^w_{\mu,a}$ is given by $\phi_{\mu,a}(n) = \Phi_{\mu,a}(n) = e^{i\mu n}\Phi(an),\ n \in \mathbb{Z}$, (Mardia and Jupp (2000), p. 48), condition (3) is satified.
In particular this applies to finite mixtures of wrapped normal and wrapped Cauchy distributions or general symmetric wrapped stable distributions (cf. Jammalamadaka and SenGupta 2001, p. 46). For applications of mixture models involving such distributions see SenGupta and Chandranath (2001).

Example 4 (*Von Mises distribution*). The von Mises distribution has density of the form

$$f_{\mu,\kappa}(\theta) = [2\pi I_0(\kappa)]^{-1}\exp(\kappa\cos[\theta-\mu]), \quad \mu,\theta \in [0, 2\pi), \quad \kappa > 0,$$

and characteristic function (see e.g. Mardia and Jupp (2000), p. 39)

$$\phi_{\mu,\kappa}(n) = \exp(i\mu n)\cdot I_n(\kappa)/I_0(\kappa), \quad n \in \mathbb{Z},$$

where $I_n(\kappa)$ denotes the modified Bessel function of order n. Letting $h(n,a) = I_n(1/a)/I_0(1/a)$ and using the series expansion of $I_n(\kappa)$ (Mardia and Jupp 2000, p. 40) shows that condition (3) of Theorem 1 is satisfied (see also Fraser et al. (1981)).

In the next example we show that identifiability of finite mixtures on the line does not imply in general identifiability of finite mixtures of the corresponding wrapped densities on the circle.

Example 5 (*Counter examples*)**.** (a) It is well known that mixtures of Poisson distributions are identifiable (see Teicher (1961), §3). We will show that finite mixtures of wrapped Poisson distributions are not in general identifiable. The wrapped Poisson distribution is obtained as follows. If X has Poisson distribution with mean λ and $m \in \mathbb{N}$ is a fixed integer, then $\Theta = 2\pi X/m \mod 2\pi$ is a random variable on the lattice $\{2\pi r/m, r = 0,\dots,m-1\}$ on the circle (see Mardia and Jupp (2000), p. 49) and is called wrapped Poisson with parameters λ and m. Denote $P(\Theta = 2\pi r/m) = p_\lambda^r$. The characteristic function of Θ is

$$\phi_\lambda(n) = \sum_{r=0}^{m-1} p_\lambda^r \exp(2\pi i n r/m) = \exp(-\lambda[1 - e^{2\pi i n/m}]).$$

For wrapped Poisson distributions with fixed m we will consider finite mixtures with at most m components. These are identifiable if and only if for m different choices of λ the vectors $(p_\lambda^0, \dots, p_\lambda^{m-1})$ are linearly independent. After applying the Fourier transform this is equivalent to the matrix $(\phi_{\lambda_j}(n))_{0\le j,n\le m-1}$ having full rank. This means that for any choice of coefficients a_n, $n = 0,\dots,m-1$ the function

$$h(\lambda) = \sum_{n=0}^{m-1} a_n \exp(\lambda e^{2\pi i n/m})$$

has at most $m-1$ real zeros. But for $m \ge 3$ and $a_1 = -a_{m-1}$, $a_j = 0$ otherwise, it is easy to see that $h(\lambda)$ has infinitely many real zeros. Hence wrapped Poisson distributions are not identifiable in general.
(b) Consider the location family of densities $f_\mu(x) = f(x-\mu)$ on the real line, where f is the triangular density

$$f(x) = 1/(8\pi)(4 - \pi^2\rho + 2\pi\rho|x-\pi|)\mathbb{1}_{[0,2\pi)}(x), \quad x \in \mathbb{R}$$

for a fixed $0 \le \rho \le 4/\pi^2$ (Jammalamadaka and SenGupta (2001), p. 34). Then finite mixtures are identifiable (this always holds for a location family, see Yakowitz and Spragins 1968, p. 213). For the wrapped densities, however, the mixture $1/2(f(\theta-\mu) + f(\theta-\mu-\pi))$ gives the uniform distribution on the circle for any μ. Note that the family of wrapped densities is a location family on the circle, hence Proposition 6 in Yakowitz and Spragins (1968) does not extend to distributions on the circle. Observe that finite mixtures remain identifiable if we restrict the location parameter to $0 \leq \mu < \pi$.

We have, however, the following analog of a result by Teicher (1961, §4) about mixtures of location families. The proof is similar to Teicher's (1961) proof and therefore omitted.

Theorem 2. *Let f be a continuous density on the circle. Assume that its characteristic function ϕ satisfies $\phi(n) \neq 0 \quad \forall\, n \in \mathbb{Z}$. If $\nu_1,\ \nu_2$ are two Borel probability measures on $[0, 2\pi)$ such that*

$$\int_0^{2\pi} f(x-\mu)\, d\nu_1(\mu) = \int_0^{2\pi} f(x-\mu)\, d\nu_2(\mu) \quad \forall\, x \in [0, 2\pi),$$

then $\nu_1 = \nu_2$.

Examples include the von Mises distribution with fixed κ and the wrapped Cauchy and normal distribution with fixed scaling parameter.

3.2 Identifiability of multivariate distributions. In this section we indicate how the presented methodology can be extended to multivariate distributions. First let us consider d-dimensional wrapped normal distributions (Jammalamadaka and SenGupta (2001), p. 53). If X is a random vector in $\mathbb{R}^d$, then $\Theta = X \bmod 2\pi$ is a random element taking values on the d-dimensional torus. The characteristic function of Θ is given by $\phi(n) = \Phi(n),\ n \in \mathbb{Z}^d$, where $\Phi(t),\ t \in \mathbb{R}^d$ is the characteristic function of X. For X having a d-dimensional normal distribution $N(\mu, \Sigma)$, Θ is called d-dimensional wrapped normal.

Theorem 3. *Finite mixtures of d-dimensional wrapped normal densities are identifiable.*

Proof. Suppose not, then for some $m \geq 1$ and $\lambda_j \neq 0,\ j = 1, \ldots, m$,

$$\sum_{j=1}^{m} \lambda_j \exp(in'\mu_j)\exp(-n'\Sigma_j n/2) = 0 \quad \forall\, n \in \mathbb{Z}^d. \tag{8}$$

Following Yakowitz and Spragins (1968, p. 211) we find $n_0 \in \mathbb{Z}^d$ such that the pairs $(n_0'\mu_1,\ n_0'\Sigma_1 n_0), \ldots, (n_0'\mu_m,\ n_0'\Sigma_m n_0)$ are all distinct. Hence taking $n = ln_0,\ l \in \mathbb{Z}$, in (5) gives

$$\sum_{j=1}^{m} \lambda_j \exp(iln_0'\mu_j)\exp(-l^2 n_0'\Sigma_j n_0/2) = 0, \quad \forall\, l \in \mathbb{Z},$$

with $\lambda_j \neq 0$, contradicting the identifiability of one dimensional wrapped normal densities proven above (cf. Example 2). $\square$

Similarly, one can obtain distributions on the cylinder as introduced by Johnson and Wehrly (1976). More specifically, let (Y_1, Y_2) be bivariate normally distributed with means $\mu_1,\ \mu_2$, variances $\sigma_1,\ \sigma_2$ and covariance $\sigma_{1,2}$. Set $\Theta = Y_1 \bmod 2\pi$ and $X = Y_2$. Then $(\Theta,\ X)$ takes values on the cylinder and has characteristic function

$$\phi(n,t) = \exp(-[n^2\sigma_1^2 + t^2\sigma_2^2 + 2nt\sigma_{1,2}]/2)\exp(i[n\mu_1 + t\mu_2]),$$

where $n \in \mathbb{Z}$ and $t \in \mathbb{R}$. The above argument (just restrict t to the integers) can be used to conclude that finite mixtures of these distributions are also identifiable.

3.3 Strong identifiability. Chen (1995) introduced the notion of strong identifiability (see also Chen, Chen and Kalbfleisch (2001)). He called a family $\{F_\alpha : \alpha \in \Omega\}$ of distribution functions on the real line *strongly identifiable* if for each α, F_α is two times differentiable and furthermore for any distinct set of parameters $\alpha_1, \ldots, \alpha_m,\ m \geq 1$, the functions $\{F_{\alpha_j}, F'_{\alpha_j}, F''_{\alpha_j}\ j = 1, \ldots, m\}$ are linearly independent. Under strong identifiability (and certain further assumptions), Chen's (1995) result on the speed of convergence in estimating the mixing distribution and the asymptotics for the modified likelihood ratio statistic of Chen et al. (2001) for testing homogeneity apply to finite mixtures from $\{F_\alpha : \alpha \in \Omega\}$. Chen (1995, pp. 226-228) proved strong identifiability of certain location and certain scale families and also of Poisson mixtures. Our method yields the following

Theorem 4. *Let $\{f_\alpha : \alpha \in \Omega\}$ be a two-parameter family of continuous densities on the real line such that for each $\alpha \in \Omega$, f_α is two times differentiable and f''_α is integrable. Suppose that the characteristic function of f_α is of the form (1), where $h(\cdot, a)$ is a function which satisfies*

$$h(t, a_2)/h(t, a_1) = o(t^{-2}) \text{ as } t \to \infty, \quad a_2 > a_1. \tag{9}$$

Then the associated family $\{F_\alpha : \alpha \in \Omega\}$ of distribution functions is strongly identifiable.

Sketch of Proof. The proof is a variation of the argument used in the proof of Theorem 1. Suppose that $\sum_{j=1}^m [\lambda_j F_{\alpha_j} + \lambda'_j F'_{\alpha_j} + \lambda''_j F''_{\alpha_j}] = 0$. Taking the derivative and applying the Fourier transform gives

$$\sum_{j=1}^m [\lambda_j + (it)\lambda'_j + (it)^2\lambda''_j]\exp(i\mu_j t) \cdot h(t, a_j) = 0 \quad \forall\ t \in \mathbb{R}, \tag{10}$$

where we again can assume that $a_1 = \ldots = a_k < a_{k+1}$. Dividing (10) by $t^2\, h(a_1, t)$ and repeating the argument from the proof of Theorem 1 k-times, we get that $\lambda''_1 = \ldots = \lambda''_k = 0$. Next we divide (10) by $t\, h(a_1, t)$. By assumption,

$$\lim_{t \to \infty} t\, \frac{h(a_j, t)}{h(a_1, t)} = 0, \quad j > k,$$

therefore the same argument applies again and yields $\lambda_1' = \ldots = \lambda_k' = 0$. Similarly $\lambda_1 = \ldots = \lambda_k = 0$, and an induction argument finishes the proof. □

Theorem 4 applies in particular to location-scale mixtures from the normal and the Cauchy distribution. The argument can also be easily extended to circular distributions as discussed in Theorem 1, under the assumption

$$h(n, a_2)/h(n, a_1) = o(n^{-2}) \text{ as } n \to \infty, \quad a_2 > a_1. \tag{11}$$

This applies in particular to wrapped distributions and to the von Mises distribution (see Examples 3 and 4). For the von Mises distribution the tail behaviour (11) follows as before from the expansion of the modified Bessel function of the first kind (cf. Mardia and Jupp 2000, p. 40).

Acknowledgement. Hajo Holzmann acknowledges financial support of the Graduiertenkolleg "Gruppen und Geometrie".

References

Al-Hussaini, E. K., El-Dab Ahmad, K. (1981). On the identifiability of finite mixtures of distributions. *IEEE Trans. Inform. Theory*, **27**, 664–668.

Bickel, P. J., Ritov, Y. and Rydén, T. (1998). Asymptotic normality of the maximum-likelihood estimator for general hidden Markov models. *Ann. Statist.*, **26**, 1614-1635.

Chandra, S. (1977). On the mixtures of probability distributions. *Scand. J. Statist.*, **3**, 105-112.

Chen, J. (1995). Optimal rate of convergence for finite mixture models. *Ann. Statist.*, **23**, 221-233.

Chen, H., Chen, J. and Kalbfleisch, J.D. (2001). A modified likelihood ratio test for homogeneity in finite mixture models. *J.R. Statist. Soc. B*, **63**, 19-29.

Fraser, M. D., Hsu, Y. S. and Walker, J. J. (1981). Identifiability of finite mixtures of von Mises distributions. *Ann. Statist.*, **9**, 1130-1131.

Ishwaran, H. (1996). Identifiability and rates of estimation for scale parameters in location mixture models. *Ann. Statist.*, **24**, 1560-1571.

Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics.* River Edge, NJ: World Scientific Publishing Co. Inc.

Johnson, R. A. and Wehrly, T. (1977). Measures and models for angular correlation and angular-linear correlation. *J. Roy. Statist. Soc. Ser. B*, **39**, 222-229.

Kent, J. T. (1983). Identifiability of finite mixtures for directional data. *Ann. Statist.*, **11**, 984-988.

Leroux, B. G. (1995). Maximum-likelihood estimation for hidden Markov models.

Stochastic Process. Appl., **12**, 127-143.

LINDSAY, B. (1995). *Mixture Models: Theory, Geometry, and Applications.* Hayward, CA: Institute for Mathematical Statistics.

MARDIA, K. V. and JUPP, P. E. (2000). *Directional Statistics.* Chichester: John Wiley and Sons Ltd.

SENGUPTA, A. and CHANDRANATH, P. (2001). On optimal tests for isotropy against the symmetric wrapped stable-circular uniform mixture family. *J. Appl. Stat.*, **28**, 129-143.

TEICHER, H. (1961). Identifiability of mixtures. *Ann. Math. Statist.*, **32**, 244-248.

TEICHER, H. (1963). Identifiability of finite mixtures. *Ann. Math. Statist.*, **34**, 1265-1269.

YAKOWITZ, S. J. and SPRAGINS, J. D. (1968). On the identifiability of finite mixtures. *Ann. Math. Statist.*, **39**, 209-214.

Adress for correspondence:

Prof. Dr. Axel Munk
Institute for Mathematical Stochastics
Georg-August-University Göttingen
Maschmühlenweg 8-10,
D-37073 Göttingen, Germany

email: munk@math.uni-goettingen.de

Fon: +49/551/39-13501
Fax: +49/551/39-13505