FINITE SAMPLE IDENTIFIABILITY OF MULTIPLE CONSTANT MODULUS SOURCES

Amir Leshem[1,2] *, *Nicolas Petrochilos* [1,3] *and Alle-Jan van der Veen* [1]

[1]Delft University of Technology, Dept. ITS/Electrical Engineering, 2628 CD Delft, The Netherlands
[2]Metalink Broadband Access, Yakum Business Park, 60972, Yakum, Israel
[3] I3S–CNRS, Algorithmes, 2000 route des Lucioles, BP.121, 06903 Sophia-Antipolis, France
Email: petro@ieee.org, leshem@cas.et.tudelft.nl, allejan@cas.et.tudelft.nl.

ABSTRACT

We prove that mixtures of continuous constant modulus sources can be identified with probability 1 with a finite number of samples (under noise-free conditions). This strengthens earlier results which only considered an infinite number of samples. The proof is based on the linearization technique of the Analytical Constant Modulus Algorithm, together with a simple inductive argument. We then study the finite alphabet case. In this case we provide an upper bound on the probability of non-identifiability for finite sample of sources. We show that under practical assumptions, this upper bound is tighter than the currently known bound.

Key words: Constant modulus signals, blind source separation, identifiability, finite sample analysis, PSK, Chernoff bound, large deviations.

I. INTRODUCTION

The constant modulus algorithm (CMA) is very popular for blind equalization [1], [2]. Similarly the separation of constant modulus (CM) signals has attracted much attention in the signal processing literature, e.g., [3], [4] and [5].It was also recognized that the underlying CM cost function can be used also for the separation of non-Gaussian signals, and more specifically finite alphabet signals [6]. While practical algorithms do exist the issue of identifiability is not well treated. Identifiability is an important issue, establishing that the only existing solutions are the original source signals up to inherent indeterminacies. Identifiability analysis has been mostly based on the expected value of the CM cost function, so that the results are only valid for infinitely many samples and ergodic scenarios. Not much is known about identifiability based on a *finite* number of samples. For the separation of a linear mixture of d continuous CM sources, [5] conjectured that about $2d$ samples should be sufficient. The provided argument was unsatisfying and based on counting the number of equations and unknowns, ignoring possible indeterminacies. For binary signals (BPSK), a sufficient condition for identifiability in [6] was based on the premise that all 2^{d-1} combinations of constellation points (up to sign) have been received. This means that an average of approximately $(d-1)2^{(d-1)}$ many samples is needed for BPSK signals and much more for higher constellations. Moreover there is always a nonzero probability that any finite number of samples does not provide identifiability (e.g., if all inputs are identical). The proof in [6] does not generalize to continuous CM sources.

In this paper we give a rigorous proof of identifiability of a mixture of d continuous or discrete complex CM sources, with finitely many samples. We use the linearization technique of [5], together with a simple inductive argument, to show that for continuous CM sources, $d(d-1)+1$ many samples suffice with probability 1. The analysis of the finite alphabet case is harder because there is a nonzero probability that sample vectors are repeated. For sufficiently large N, we specify an upper bound on the probability that a data set with N samples is not yet identifiable. The probability decays exponentiallyas a function of the number of samples and as L^{-N-1} as a function of alphabet size L.

II. THE IDENTIFICATION PROBLEM

Consider an array with p sensors receiving d narrowband constant modulus signals. Under standard assumptions for the array manifold, we can describe the received signal as an instantaneous linear combination of the source signals,

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n) \tag{1}$$

where

$\mathbf{x}(n) = [x_1(n), \cdots, x_p(n)]^T$ is a $p \times 1$ vector of received signals at discrete time n (^{T} denotes matrix transposition),

$\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_d]$, where $\mathbf{a}_i$ is the array response vector towards the i-th signal,

*Amir Leshem was partially supported by the NOEMI project of the STW under contract no. DEL77-4476.

$\mathbf{s}(n) = [s_1(n), \cdots, s_d(n)]^T$ is a $d \times 1$ vector of source signals at time n.
We further assume that all sources have constant modulus, i.e. for all n, $|s_i(n)| = 1$ ($i = 1, \cdots, d$), and that $\mathbf{A}$ has full column rank (this implies $p \geq d$).

In our problem, the array is assumed to be uncalibrated so that the array response vectors $\mathbf{a}_i$ are unknown. Unequal source powers are absorbed in the mixing matrix. Phase offsets of the sources after demodulation are part of the s_i. Thus we can write $s_i(n) = e^{j\phi_i(n)}$, where $\phi_i(n)$ is the unknown phase modulation for source i, and we define $\boldsymbol{\phi}(n) = [\phi_1(n), \cdots, \phi_d(n)]^T$ as the phase vector for all sources at time n. Note that this leads to the fundamental indeterminacy of phase exchange between a source and the corresponding column in the mixing matrix. Furthermore we can permute the sources and simultaneously permute the columns of $\mathbf{A}$. Thus, $\mathbf{A}$ is determined only up to a permutation of its columns and a complex unit-modulus scaling of each column.

The identifiability problem asks for the number of samples needed in order to ensure (with probability 1) that in the *noiseless* case we have a unique solution up to the above indeterminacies.

III. IDENTIFIABILITY WITH INFINITELY MANY SAMPLES

Let $\mathbb{T} = \{z : |z| = 1\}$ be the complex unit circle, and let $\mathbb{T}^d$ be the Cartesian product of d copies of $\mathbb{T}$, representing the collection of d-dimensional CM source vectors.

We first characterize linear transformations $\mathbf{G}$ mapping $\mathbb{T}^d$ into itself. Consider the set $\mathbb{G}$,

$$\mathbb{G} = \left\{ \mathbf{G} \in \mathbb{C}^{d \times d} \mid \mathbf{G} \text{ invertible; } \ \mathbf{s} \in \mathbb{T}^d \Rightarrow \mathbf{G}\mathbf{s} \in \mathbb{T}^d \right\} .$$

Lemma 1: Let $\mathbf{G} \in \mathbb{G}$. Then $\mathbf{G} = \mathbf{P}\mathbf{\Lambda}$, where $\mathbf{P}$ is a permutation matrix and $\mathbf{\Lambda}$ a diagonal matrix with diagonal elements on the unit circle.
Proof We will prove that each row of $\mathbf{G}$ contains at most one non-zero element with magnitude 1. Let

$$\mathbf{g} = [g_1, \ldots, g_d] = [r_1 e^{j\phi_1}, \ldots, r_d e^{j\phi_d}]$$

be a row of $\mathbf{G}$ where r_i is the magnitude of g_i. For each $\mathbf{s} \in \mathbb{T}^d$, we know that $|\mathbf{g}\mathbf{s}| = 1$. Choose $\mathbf{s}_1$ such that $\mathbf{s}_1 = [e^{-j\phi_1}, \ldots, e^{-j\phi_d}]^T$. We obtain

$$\mathbf{g}\mathbf{s}_1 = r_1 + \sum_{i>1}^{d} r_i = 1 \qquad (2)$$

since all r_i are non-negative real numbers. Similarly define $\mathbf{s}_2$ by $(\mathbf{s}_2)_1 = e^{-j\phi_1}$ and $(\mathbf{s}_2)_i = -e^{-j\phi_i}$ for $2 \leq i \leq d$. Then

$$\mathbf{g}\mathbf{s}_2 = r_1 - \sum_{i>1}^{d} r_i$$

Since $|\mathbf{g}\mathbf{s}_2| = 1$ we have either

$$r_1 - \sum_{i>1}^{d} r_i = 1 \qquad (3)$$

or

$$r_1 - \sum_{i>1}^{d} r_i = -1.$$

From (2) and (3), we obtain in the first case that $r_1 = 1$ and $\sum_{i>1} r_i = 0$, whereas in the second case $r_1 = 0$ and $\sum_{i>1} r_i = 1$. Proceeding inductively we obtain that exactly one element of $\mathbf{g}$ is non-zero with magnitude 1. Since all the rows of $\mathbf{G}$ have this property and $\mathbf{G}$ is invertible, it must be a permutation of a diagonal matrix with unit-modulus diagonal entries. □

The identifiability theorem for infinite samples follows directly from the preceding lemma:

Theorem 2: Consider an infinite collection of vectors $\mathbf{s}(n) \in \mathbb{T}^d$, $n = 1, \cdots, \infty$, and suppose that the collection is dense in $\mathbb{T}^d$. Suppose that we have available the observations $\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n)$, where $\mathbf{A} \in \mathbb{C}^{p \times d}$ is full column rank d. Then $\mathbf{A}$ is uniquely determined by the observations, up to a permutation and a unit-modulus complex scaling of the columns.
Proof Suppose that there is another matrix $\mathbf{B} \in \mathbb{C}^{p \times d}$ and collection of source signals $\mathbf{z}(n) \in \mathbb{T}^d$ which generate the same data $\{\mathbf{x}(n)\}$.

The linear span of the collection $\{\mathbf{s}(n)\}$ is $\mathbb{C}^d$, so that the linear span of $\{\mathbf{x}(n)\}$ is a d-dimensional subspace in $\mathbb{C}^p$. Hence $\mathbf{B}$ is full column rank d. Since its column span must be the same as that of $\mathbf{A}$, $\mathbf{G} := \mathbf{B}^{\dagger}\mathbf{A} \in \mathbb{C}^{d \times d}$ is full rank. Moreover, since $\mathbf{s}(n)$ is dense in $\mathbb{T}^d$, it follows that for any $\mathbf{s} \in \mathbb{T}^d$, $\mathbf{G}\mathbf{s} = \mathbf{z} \in \mathbb{T}^d$. Hence $\mathbf{G} \in \mathbb{G}$, and lemma 1 claims that $\mathbf{G} = \mathbf{P}\mathbf{\Lambda}$, so that $\mathbf{A} = \mathbf{B}\mathbf{P}\mathbf{\Lambda}$. □

The proof of the theorem shows that the infinite collection of vectors $\{\mathbf{s}(n)\}$ is only used to quickly deduce that $\mathbf{G} \in \mathbb{G}$. The question is whether this can be done using a finite set of vectors.

IV. IDENTIFIABILITY WITH FINITELY MANY SAMPLES

In this section, we derive a sufficient condition on the number of samples needed to guarantee identifiability with probability 1, for the case of constant modulus signals with continuous alphabet. Based on the discussion of the previous section we restrict ourselves to invertible linear transformations from $\mathbb{T}^d$ to $\mathbb{T}^d$.

Consider a collection of N vectors $\mathcal{S} = \{\mathbf{s}(n) \in \mathbb{T}^d,$

$n = 1, \ldots, N\}$, and let

$$\Psi = \begin{bmatrix} 1 & s_1(1)s_2^*(1) & s_1^*(1)s_2(1) & \cdots & s_d^*(1)s_{d-1}(1) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & s_1(N)s_2^*(N) & s_1^*(N)s_2(N) & \cdots & s_d^*(N)s_{d-1}(N) \end{bmatrix} \quad (4)$$

where $*$ denotes complex conjugate and Ψ has size $N \times d(d-1)+1$. We call $\mathcal{S}$ *"persistently exciting"* if Ψ has full column rank. Note that this implies that $N \geq d(d-1)+1$. It also implies that the constellation is complex (for BPSK constellations, columns of Ψ are repeated and a modified definition can be introduced).

Lemma 3: Let $N \geq d(d-1)+1$, and let $\mathbf{s}(n) \in \mathbb{T}^d$, $n = 1, \ldots, N$, be a persistently exciting collection. Consider an invertible linear transformation $\mathbf{G} \in \mathbb{C}^{d \times d}$ such that $\mathbf{G}\mathbf{s}(n) \in \mathbb{T}^d$, for $n = 1, \ldots, N$. Then $\mathbf{G} = \mathbf{P}\mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is a diagonal matrix with unit norm diagonal entries $\mathbf{\Lambda}$ and $\mathbf{P}$ is a permutation matrix.

Proof Since $\mathbf{G}$ is invertible, it is sufficient to prove that each row $\mathbf{g}$ of $\mathbf{G}$ contains exactly one non-zero element which is unit norm. Let $\mathbf{g} = [g_1, \ldots, g_d]$, and let $y(n) = \mathbf{g}\mathbf{s}(n)$ be the corresponding entry of $\mathbf{G}\mathbf{s}(n)$.

Then for each $n \in \{1, \ldots, N\}$, we have:

$$\begin{aligned} y(n) &= \sum_{i=1}^{d} g_i s_i(n) \\ \Rightarrow \quad |y(n)|^2 &= |\sum_{i=1}^{d} g_i s_i(n)|^2 \\ \Rightarrow \quad 1 &= \sum_{1 \leq i,j \leq d} g_i g_j^* s_i(n) s_j^*(n) \end{aligned} \quad (5)$$

Denote $P_{ij} = g_i g_j^*$ and $P_T = \sum_{i=1}^{d} P_{ii}$. By linearizing (5) and considering all n, we obtain (as in ACMA [5])

$$\Psi \mathbf{p} = \mathbf{1} \quad (6)$$

where $\mathbf{p} = [P_T, P_{12}, P_{21}, \ldots, P_{d,d-1}]^T$, $\mathbf{1} = [1, \ldots, 1]^T$, and Ψ is as in (4). A particular solution of (6) for $\mathbf{p}$ is given by

$$\begin{cases} P_T = 1 \\ P_{ij} = 0, \quad \forall i \neq j \end{cases} \quad (7)$$

Suppose that $g_j \neq 0$ for some j, then since $P_{ij} = g_i g_j^*$ we immediately obtain that $g_i = 0$ for all $i \neq j$. Since $P_T = 1$, $|g_j|^2 = 1$. Hence each row $\mathbf{g}$ of $\mathbf{G}$ has precisely one non-null entry, which is unit-modulus. It follows that $\mathbf{G} = \mathbf{P}\mathbf{\Lambda}$.

Since Ψ is full column rank, this is the only solution to (7). □

Combining with theorem 2 we obtain

Theorem 4: Identifiability as in theorem 2 already holds for a finite collection of source signals $\mathbf{s}(n)$, $n = 1, \cdots, N$, where $N \geq d(d-1)+1$, if this collection is persistently exciting.

V. PERSISTENCE OF EXCITATION

The remaining issue is to establish when a collection of vectors in $\mathbb{T}^d$ is persistently exciting. As usual, this is hard to characterize in a deterministic setting. In a stochastic sense, any "sufficiently random" collection of $N \geq d(d-1)+1$ complex vectors in $\mathbb{T}^d$ is expected to be persistently exciting. Although this appears a reasonable argument, the inter-relations of the elements of Ψ make it not completely evident that this is the case. Moreover, in the case of discrete alphabet CM sources, e.g. QPSK, proofs are harder because the randomness is much less and perhaps not sufficient. We first make a more explicit statement for continuous CM sources, and then consider the discrete alphabet CM case.

V-A. The continuous alphabet case

Lemma 5: Let $\mathbf{s}(n)$, for $n = 1, \ldots, N$, be a collection of continuous-alphabet independent identically distributed complex vectors in $\mathbb{T}^d$ with stochastically independent components. If $N \geq d(d-1)+1$ then the matrix Ψ has full column rank with probability 1.

Proof Given $N \geq d(d-1)+1$ samples of $\mathbf{s}(n)$, assume towards contradiction that there exists a vector $\boldsymbol{\alpha} \neq \mathbf{0}$ such that $\Psi\boldsymbol{\alpha} = 0$, or equivalently $\exists\{\alpha_0, \alpha_{ij}, 1 \leq i \neq j \leq d\}$, not all zeros such that for every $n = 1, \cdots, N$, the next equation holds:

$$\alpha_0 + \sum_{i \neq j} \alpha_{ij} s_i(n) s_j^*(n) = 0 \quad (8)$$

After multiplying every equation by $s_1^*(n)$, this becomes, for all $n = 1, \cdots, N$:

$$\sum_{j>1} \alpha_{1j} s_j^*(n) \quad + s_1^*(n) \left(\alpha_0 + \sum_{i \neq j > 1} \alpha_{ij} s_i(n) s_j^*(n) \right) + (s_1^*(n))^2 \sum_{i>1} \alpha_{i1} s_i(n) = 0$$

After taking the conjugate of this expression, we see that it is a set of N independent quadratic equations in $s_1(n)$: $a(n) + b(n)s_1(n) + c(n)s_1^2(n) = 0$. Hence one of the following holds: (a) $s_1(n)$ is a function of $(s_2(n), \cdots, s_d(n))$ which contradicts the independence assumption, and for $s_1(n)$ in a continuous alphabet is a zero-measure event, or (b) the coefficients satisfy: $a(n) = b(n) = c(n) = 0$, $\forall n \in \{1, \ldots, N\}$, hence:

(1) $a(n) = 0 \Rightarrow \sum_{j>1} \alpha_{1j} s_j^*(n) = 0$. For each n, this is a linear condition on the $d-1$ coefficients $\{\alpha_{1j}\}$. Using $d-1$ independent samples suffices to derive that $\forall i$, $\alpha_{1i} = 0$.

(2) Similarly, from the condition $c(n) = 0$ and using $d-1$ other independent samples, we obtain that $\forall i$, $\alpha_{i1} = 0$.

(3) $b(n) = 0 \Rightarrow \alpha_0 + \sum_{i \neq j > 1} \alpha_{ij} s_i(n) s_j^*(n) = 0$. This condition is verified by applying inductively the same argument on $b(n)$ as we did on equation (8).
We thus obtain that all α_{ij} are equal to zero, and $\alpha_0 = 0$. Therefore $\mathbf{\Psi}$ is full rank with probability one.

If $d = 1$, then we trivially need 1 sample to conclude that $\mathbf{\Psi}$ is full rank. Hence the recursive application of the argument needs $2(d-1) + \cdots + 2 + 1 = d(d-1) + 1$ independent samples, and this number is sufficient with probability 1. □

V-B. The finite alphabet case—large deviations bound

For discrete-alphabet sources, we have to make a different approach since the independence of the conditions is not evident. Let $\mathbf{s}(n)$, for $n = 1, \ldots, N$, be a collection of zero mean independent identically distributed complex vectors in $\mathbb{T}^d$ with stochastically independent and circularly symmetric components, or more in particular,

$$\begin{array}{lcll} E(|s_i|^2) & = & 1, & \\ E(s_i^2) & = & 0, & \\ E(s_i s_j^*) & = & 0, & i \neq j, \\ E(s_i^2 s_j^{*2}) & = & 0, & i \neq j, \\ E(s_i s_j s_k^{*2}) & = & 0, & i \neq j \neq k, \\ E(s_i s_j s_k^* s_l^*) & = & 0, & i \neq j \neq k \neq l. \end{array} \tag{9}$$

Denote a generic n'th row of $\mathbf{\Psi}$ by

$$\mathbf{v}(n) = [1, s_1(n) s_2^*(n), s_2(n) s_1^*(n), \cdots]. \tag{10}$$

Then (ommitting the index n) we have

$$\mathbf{v}^H(n)\mathbf{v}(n) = \begin{bmatrix} 1 & s_1 s_2^* & s_2 s_1^* & \cdots \\ s_2 s_1^* & 1 & s_2^2 s_1^{*2} & \cdots \\ s_1 s_2^* & s_1^2 s_2^{*2} & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

With the assumptions (9), it follows that $E(\mathbf{v}^H \mathbf{v}) = \mathbf{I}$. Note that $\frac{1}{N} \mathbf{\Psi}^H \mathbf{\Psi} \to E(\mathbf{v}^H \mathbf{v})$ as $N \to \infty$. Hence for sufficiently large N, $\mathbf{\Psi}$ must have full column rank. For continuous CM sources we already proved that $N \geq d(d-1) + 1$ is sufficient w.p. 1. For discrete-alphabet sources it can happen that the same constellation vector is received multiple times and hence N might have to be larger.

We next quantify the probability that N samples of the array output are sufficient. We first provide a simple proof which gives subexponentially decreasing probability of non-identifiability. Subsequently, in the next subsection we provide a more accurate analysis providing an exponentially decreasing upper bound on the probability of non-identifiability fopr the case of $L-$PSK i.i.d. signals. Let

$$\hat{\mathbf{R}}_N = \frac{1}{N} \mathbf{\Psi}^H \mathbf{\Psi} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{v}(n)^H \mathbf{v}(n)$$

As we have shown $E(\hat{\mathbf{R}}_N) = \mathbf{I}$. We now analyze the rate of convergence of $\hat{\mathbf{R}}_N$ to $\mathbf{I}$ and provide an upper bound on the probability that $\hat{\mathbf{R}}_N$ is singular. By the argument of section To that end we use the following consequence of Gershgorin's theorem.

Theorem 6 ([7, p.349]) Let $\mathbf{A} = [a_{ij}]$ be a Hermitian matrix. Assume that for all i, $|a_{ii}| > 0$, and that $\mathbf{A}$ is diagonally dominated, i.e., for all i

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|,$$

then $\mathbf{A}$ is strictly positive definite.

Assume that all off-diagonal elements of $\hat{\mathbf{R}}_N$ have magnitude less than $\frac{1}{d(d-1)+1}$. Then for all i

$$\sum_{j \neq i} (\hat{\mathbf{R}}_N)_{ij} \; < \; \frac{d(d-1)}{d(d-1)+1} \; < \; (\hat{\mathbf{R}}_N)_{ii} \; = \; 1 \tag{11}$$

and by theorem 6 we can conclude that $\hat{\mathbf{R}}_N$ is strictly positive definite. It remains to compute a bound on the probability that all off-diagonal elements of $\hat{\mathbf{R}}_N$ have magnitude less than $\frac{1}{d(d-1)+1}$. This will provide a lower bound on the probability of persistence of excitation since as discussed above, if $\hat{\mathbf{R}}_N$ is non-singular then $\mathbf{\Psi}$ is full rank. To obtain the bound we use large deviation theory. The proof of the following lemma will be omitted.

Lemma 7: For every $i \neq j$, for all $\varepsilon > 0$ and N sufficiently large

$$P\left((\hat{\mathbf{R}}_N)_{ij} > \frac{1}{d(d-1)+1}\right) \; < \; 2e^{-\frac{1}{2}(1-\varepsilon)\frac{N}{(d(d-1)+1)\log\log N}}$$

We now use the lemma above to bound the probability that $\hat{\mathbf{R}}_N$ is non-singular. Note that since $\hat{\mathbf{R}}_N$ is Hermitian it is sufficient to obtain that all entries above the diagonal are sufficiently small. There are $\frac{1}{2}[d(d-1)+1][d(d-1)]$ entries, and since most entries are uncorrelated (although not independent) we obtain that the probability that all entries are smaller than $1/d(d-1)$, for N sufficiently large, is bounded by (but not equal to)

$$\left(1 - 2e^{-\frac{1}{2}(1-\varepsilon)\frac{N}{(d(d-1)+1)^2 \log\log N}}\right)^{\frac{1}{2}(d(d-1)+1)(d(d-1))} \tag{12}$$

Since for any x such that $0 < x < 1$ we have $(1-x)^n > 1 - nx$, we can bound (12) by

$$1 - d^4 e^{-\frac{1}{2}(1-\varepsilon)\frac{N}{(d(d-1)+1)^2 \log\log N}}. \tag{13}$$

In summary, the probability of having a data set that is not persistently exciting is asymptotically less than $d^4 e^{-\frac{1}{2}(1-\varepsilon)\frac{N}{(d(d-1)+1)\log\log N}}$ (for any $\epsilon > 0$).

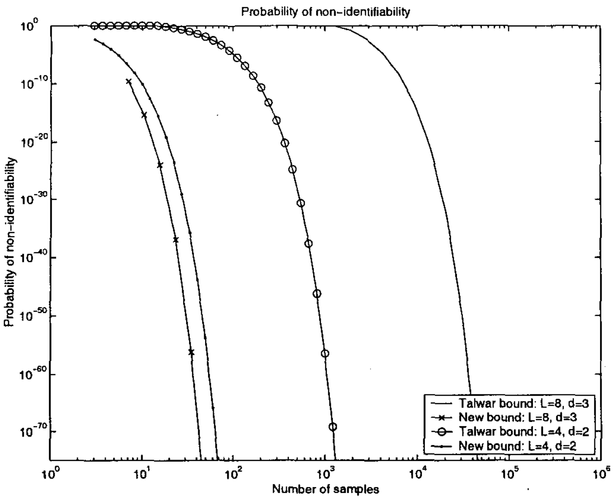

Fig. 1. Finite alphabet CM sources: Upper bound on the probability that N source samples are not persistently exciting. d sources and L-PSK constellations.

V-C. Chernoff bound and finite alphabet CM signals

We now provide a more accurate bounding. The proof uses Chernoff's bound on finite alphabet L-PSK signals. This bound holds for all values of N. Furthermore it also shows that for any fixed $N > d(d-1)$ increasing the alphabet size L decreases the probability of non-identifiability at least as $\frac{1}{L^{N-1}}$. Our goal is to bound

$$P\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{v}_i(n)^H\mathbf{v}(n) < \frac{1}{d(d+1)}\right) \qquad (14)$$

where $\mathbf{v}_i(n), \mathbf{v}_j(n)$ are the i'th and j'th rows of $\boldsymbol{\Psi}$ as defined as in (10). Let the probability of identifiability of d sources using N vector samples taken from L-PSK i.i.d source (L even) be denoted by $P_{id}(L,d,N)$. Then the following inequality holds:

$$P_{id}(L,N,d) \geq$$
$$1 - 2d^4\left(\tfrac{2}{L}\right)^{N-1} e^{-\frac{N}{d(d+1)}\tanh^{-1}\left(\frac{1}{d(d+1)}\right)}\left(\frac{1}{\sqrt{d^2(d+1)^2-1}}\right)^N \qquad (15)$$

This is better than the large deviation bound (13) since the dependence on N is exponential and not sub-exponential and is also valid for all values of N. Moreover we can see that as the alphabet size is increased the probability of non-identifiability approaches 0 as $L^{-(N-1)}$. The proof will be provided elsewhere.

VI. SIMULATIONS

We now illustrate a comparison of the new upper bound on failure of identifiability (13) to the bound by Talwar [6], see figure 1. We can clearly see that the new bound is much better with orders of magnitudes less samples necessary for a given probability of identifiability.

VII. CONCLUSIONS

We presented a rigorous proof of a sufficient condition for the identifiability of mixtures of CM signals, based on finitely many samples. For continuous-CM sources, $N = d(d-1)+1$ samples are sufficient with probability 1. For finite-alphabet cases, only an upper bound on the probability of non-identifiability given alphabet size, number of sources and number of samples could be derived. However the new bound is much tighter than previously known bounds.

VIII. ACKNOWLEDGEMENT

We would like to thank S. Litsyn for helpfull discussion on Chernoff's bound.

IX. REFERENCES

[1] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Communications*, vol. 28, pp. 1867–1875, Nov. 1980.

[2] J. Treichler and B. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 459–471, Apr. 1983.

[3] J. Shynk and R. Gooch, "The constant modulus array for cochannel signal copy and direction finding," *IEEE Trans. Signal processing*, vol. 44, pp. 652–660, Mar. 1996.

[4] A. Leshem, "Maximum likelihood separation of constant modulus signals," *IEEE trans. on Signal Processing*, vol. 48, pp. 2948 –2952, October 2000.

[5] A. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 1136–1155, May 1996.

[6] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array. I. algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1184–1197, May 1996.

[7] R. Horn and C. Johnson, *Matrix Analysis.* Cambridge, 1994.