ON REGULARITY AND IDENTIFIABILITY OF BLIND SOURCE SEPARATION UNDER CONSTANT-MODULUS CONSTRAINTS

Yingwei Yao and Georgios B. Giannakis

Department of Electrical & Computer Engineering
University of Minnesota
200 Union St. SE, Minneapolis, MN 55455

ABSTRACT

We investigate the information regularity and identifiability of the blind source separation problem with constant modulus constraints on the sources. We demonstrate that the information regularity (existence of a finite Cramér-Rao bound) is closely related to local identifiability. Sufficient and necessary conditions for local identifiability are derived. We also study the conditions under which unique (global) identifiability is guaranteed within the inherently unresolvable ambiguities on phase rotation and source permutation. Both sufficient and necessary conditions are obtained.

1. INTRODUCTION

One of the most widely used blind source separation (BSS) algorithms is the constant modulus algorithm (CMA) [7, 1, 9]. While extensive literature exists on the CMA and various uses of the CM criterion (see e.g., [3] and the references therein), identifiability issues have not been resolved fully.

Closely related to identifiability is the regularity of the Fisher information matrix (FIM). Cramér-Rao bound (CRB) has been widely used to investigate the performance limit of unbiased parameter estimators. Since the existence of a useful CRB depends on the regularity of the FIM, it is also of interest to study the conditions of information regularity.

In this paper, we will investigate the existence of a finite CRB for the BSS problem under CM constraints. The link between information regularity (a.k.a. existence of CRB) and local identifiability will be established, and the conditions for local identifiability will be derived. The issue of global identifiability for the BSS problem under CM constraints will also be studied. We will derive a sufficient and necessary condition for this problem to be globally identifiable up to a phase rotation and source permutation ambiguity, both of which are inherently unresolvable by any blind algorithm. Due to the space constraint, all proofs in this paper have been omitted, but can be found in [11].

This work was supported by the ARL/CTA grant no. DAAD19-01-2-0011.

Notation: Upper and lower case bold symbols will be used to denote matrices and column vectors, respectively; $(\cdot)^*$ will denote conjugation; $(\cdot)^{\mathcal{H}}$ Hermitian transpose; $(\cdot)^T$ transpose; $\mathbf{I}_N$ will stand for the $N \times N$ identity matrix; $\bar{x}$ and $\tilde{x}$ will denote the real and imaginary parts of x, respectively; finally, diag(x) will denote a diagonal matrix whose diagonal elements are the entries of the vector $\boldsymbol{x}$.

2. PROBLEM FORMULATION

Consider the following input-output matrix-vector model:

$$\boldsymbol{x}(i) = \mathbf{H}\boldsymbol{b}(i) + \boldsymbol{w}(i), \quad (1)$$

where $\mathbf{H}$ is an $N \times K$ mixing matrix, $\boldsymbol{x}(i)$ is the received vector in the ith symbol interval, and $\boldsymbol{b}(i)$ and $\boldsymbol{w}(i)$ are the source symbol vector and the noise vector, respectively. Equation (1) appears in many problems related to wireless communications and signal processing, e.g., multi-input multi-output (MIMO) systems and direct-sequence code-division multiple access (DS-CDMA) systems.

In this paper, we will assume that the mixing matrix $\mathbf{H}$ is of full column rank, the additive noise $\boldsymbol{w}(i)$ is zero-mean, white Gaussian with covariance matrix $\sigma^2\mathbf{I}_N$, and the source symbols have constant modulus, namely, $\boldsymbol{b}(i) := [b_1(i), \ldots, b_K(i)]^T$ satisfies $|b_k(i)| = 1$ for $k = 1, \ldots, K$.

3. INFORMATION REGULARITY

The Cramér-Rao lower bound benchmarks the covariance of unbiased estimators of unknown parameters. The nonexistence of a finite CRB often signifies the lack of identifiability of a particular parameter estimation problem. Thus, it is of interest to investigate the existence of a finite CRB for the BSS problem.

3.1. CRB for BSS under CM Constraints

We will first compute the CRB of estimators $\hat{\mathbf{H}}$, $\hat{\boldsymbol{b}}(i)$ in the BSS model (1) under CM constraints. Since a small amount of training data is needed to resolve the inherent phase ambiguity of the CM constraint, we will assume that a sufficient number of (say the first T) symbols of every source $b_k(i)$ is known.

3.1.1. A Constrained CRB Formulation

To derive the CRB under CM constraints, we will rely on the constrained CRB expression of [8]. Compared to the traditional method using reparameterization, this approach is more convenient and provides more insight to the effects of the constraints.

Let $\boldsymbol{y}$ be the vector of observations and $\boldsymbol{\theta} \in \mathbb{R}^n$ be the $n \times 1$ real parameter vector to be estimated. For a set of k ($k < n$) constraints $\boldsymbol{f}(\boldsymbol{\theta}) = \boldsymbol{0}$, define the gradient matrix of the constraints as

$$\mathbf{F}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}. \tag{2}$$

Assume that $\mathbf{F}(\boldsymbol{\theta})$ has full row rank for $\boldsymbol{\theta}$ in the constrained parameter space (rank deficiency of $\mathbf{F}(\boldsymbol{\theta})$ usually indicates that certain constraints are redundant, and can be eliminated). Let $\mathbf{U}(\boldsymbol{\theta})$ be the $n \times (n-k)$ matrix whose columns form an orthonormal basis for the null space of $\mathbf{F}(\boldsymbol{\theta})$ and denote the unconstrained FIM as $\mathbf{J}_\theta$. If $\mathbf{U}^T\mathbf{J}_\theta\mathbf{U}$ is nonsingular, then any unbiased estimator $\hat{\boldsymbol{\theta}}$ must satisfy

$$E\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\} \geq \mathbf{U}(\mathbf{U}^T\mathbf{J}_\theta\mathbf{U})^{-1}\mathbf{U}^T. \tag{3}$$

A finite CRB for the constrained parameter estimation exists if and only if $|\mathbf{U}^T\mathbf{J}_\theta\mathbf{U}| \neq 0$ [8].

3.1.2. Constrained CRB for BSS

Given M observed vectors $\mathbf{X} = \{\boldsymbol{x}(0), \ldots, \boldsymbol{x}(M-1)\}$, the likelihood function is

$$p(\mathbf{X}; \boldsymbol{\theta}) = \prod_{i=0}^{M-1} \frac{1}{(\pi\sigma^2)^N} \exp\left[-\frac{1}{\sigma^2}|\boldsymbol{x}(i) - \mathbf{H}\boldsymbol{b}(i)|^2\right], \tag{4}$$

where $\boldsymbol{\theta} := [\bar{\boldsymbol{h}}_0^T, \tilde{\boldsymbol{h}}_0^T, \ldots, \bar{\boldsymbol{h}}_{N-1}^T, \tilde{\boldsymbol{h}}_{N-1}^T, \bar{\boldsymbol{b}}^T(0), \tilde{\boldsymbol{b}}^T(0), \ldots, \bar{\boldsymbol{b}}^T(M-1), \tilde{\boldsymbol{b}}^T(M-1)]^T$ is the set of parameters to be estimated (expressed in terms of their real and imaginary parts), and we have denoted the nth row of $\mathbf{H}$ as $\boldsymbol{h}_i^T$. The unconstrained FIM for this problem can be shown to be:

$$\mathbf{J} = \frac{2}{\sigma^2}\left(\begin{array}{c|c} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \hline \mathbf{J}_{12}^T & \mathbf{J}_{22} \end{array}\right), \tag{5}$$

where

$$\mathbf{J}_{12} := \begin{pmatrix} \mathcal{M}_{0,0} & \cdots & \mathcal{M}_{0,M-1} \\ \cdots & \cdots & \cdots \\ \mathcal{M}_{N-1,0} & \cdots & \mathcal{M}_{N-1,M-1} \end{pmatrix}, \tag{6}$$

with

$$\mathcal{M}_{n,i} := \begin{pmatrix} \mathrm{Re}\left[\boldsymbol{b}(i)\boldsymbol{h}_n^{\mathcal{H}}\right] & \mathrm{Im}\left[\boldsymbol{b}(i)\boldsymbol{h}_n^{\mathcal{H}}\right] \\ -\mathrm{Im}\left[\boldsymbol{b}(i)\boldsymbol{h}_n^{\mathcal{H}}\right] & \mathrm{Re}\left[\boldsymbol{b}(i)\boldsymbol{h}_n^{\mathcal{H}}\right] \end{pmatrix}; \tag{7}$$

$\mathbf{J}_{11}$ and $\mathbf{J}_{22}$ are block diagonal matrices, whose diagonals consist of N copies of

$$\mathcal{A} := \begin{pmatrix} \mathrm{Re}\left[\mathbf{B}\mathbf{B}^{\mathcal{H}}\right] & \mathrm{Im}\left[\mathbf{B}\mathbf{B}^{\mathcal{H}}\right] \\ -\mathrm{Im}\left[\mathbf{B}\mathbf{B}^{\mathcal{H}}\right] & \mathrm{Re}\left[\mathbf{B}\mathbf{B}^{\mathcal{H}}\right] \end{pmatrix}, \tag{8}$$

and M copies of

$$\mathcal{D} := \begin{pmatrix} \mathrm{Re}\left[\mathbf{H}^{\mathcal{H}}\mathbf{H}\right] & -\mathrm{Im}\left[\mathbf{H}^{\mathcal{H}}\mathbf{H}\right] \\ \mathrm{Im}\left[\mathbf{H}^{\mathcal{H}}\mathbf{H}\right] & \mathrm{Re}\left[\mathbf{H}^{\mathcal{H}}\mathbf{H}\right] \end{pmatrix}, \tag{9}$$

respectively, where

$$\mathbf{B} := \begin{pmatrix} b_1(0) & \cdots & b_1(M-1) \\ \vdots & \ddots & \vdots \\ b_K(0) & \cdots & b_K(M-1) \end{pmatrix}. \tag{10}$$

Under the assumption that the sources are CM with amplitude 1 and the first T symbols are known, the parameter set $\boldsymbol{\theta}$ should satisfy the following set of constraint equations:

$$\bar{b}_k(i) - \bar{b}_{ki} = 0,\ i = 0, \ldots, T-1, \tag{11}$$

$$\tilde{b}_k(i) - \tilde{b}_{ki} = 0,\ i = 0, \ldots, T-1, \tag{12}$$

$$\bar{b}_k^2(i) + \tilde{b}_k^2(i) = 1,\ i = T, \ldots, M-1, \tag{13}$$

where $k = 1, \ldots, K$. The gradient matrix for these constraints is an $(M+T)K \times (2N+2M)K$ matrix:

$$\mathbf{F}(\boldsymbol{\theta}) = \left(\begin{array}{c|ccccc} & \mathbf{I}_{2TK} & & & & \\ & & \underline{\mathbf{B}}_T & & & \\ \mathbf{F}_h & & & \underline{\mathbf{B}}_{T+1} & & \\ & & & & \ddots & \\ & & & & & \underline{\mathbf{B}}_{M-1} \end{array}\right), \tag{14}$$

where $\mathbf{F}_h := \mathbf{0}(M+T)K \times 2NK$, $\underline{\mathbf{B}}_i := 2[\mathrm{diag}(\bar{\boldsymbol{b}}(i)), \mathrm{diag}(\tilde{\boldsymbol{b}}(i))]$, and the empty spaces indicate that the corresponding matrix entries are zero. From (14), the matrix $\mathbf{U}(\boldsymbol{\theta})$, and subsequently the constrained CRB can then be obtained straightforwardly.

Next, we proceed to link local identifiability with information regularity using Equations (4)–(14).

3.2. Local Identifiability and Information Regularity

The regularity (invertibility) of the Fisher information matrix is related to parameter identifiability. To investigate the link between them in the BSS context, we first introduce several relevant definitions [5]:

Definition 1. *Suppose a family of probability measures $\{F(\cdot,\theta), \theta \in \Theta\}$ on a measurable space $(\Omega, \mathcal{F})$ is dominated by a σ-finite measure μ*[1]. *Two parameter values θ_0 and θ_1 are said to be* observationally equivalent *if*

$$\frac{dF(\mathbf{x},\theta_0)}{d\mu} = \frac{dF(\mathbf{x},\theta_1)}{d\mu} a.e.[\mu]. \tag{15}$$

Definition 1 basically means that two parameter values can be seen as equivalent if the probability densities indexed by them are the same except on a set of measure 0.

Definition 2. *Suppose the family $\{F(\cdot,\theta), \theta \in \Theta\}$ is dominated by a σ-finite measure μ. A parameter value $\theta_0 \in \Theta$ is said to be* locally identifiable *if there exists an open neighborhood of θ_0 containing no other $\theta \in \Theta$ that is observationally equivalent to θ_0; it is said to be* (globally) identifiable *if there exists no other $\theta \in \Theta$ that is observationally equivalent to θ_0.*

Due to the permutation ambiguity present in the BSS problem, $\{\mathbf{H}, \boldsymbol{b}(i),\ i = 1,\ldots,M\}$ is not globally identifiable, in general. In this subsection, we will focus on local identifiability. The link between the local identifiability and the information regularity is given by the following theorem [6]:

Theorem 1. *Assume that: i) the unconstrained parameter space A is an open set in $\mathbb{R}^n$; ii) the likelihood function p is nonnegative and the equation $\int p(\mathbf{y}|\boldsymbol{\theta})d\mathbf{y} = 1$ holds for all $\boldsymbol{\theta} \in A$; iii) the support of p is the same for all $\boldsymbol{\theta} \in A$; iv) p is smooth in $\boldsymbol{\theta}$; v) the elements of the unconstrained FIM $\boldsymbol{J}(\boldsymbol{\theta})$ exist and are continuous functions of $\boldsymbol{\theta}$ everywhere in A; and vi) $\boldsymbol{\theta}$ satisfies a set of constraint equations $f_i(\boldsymbol{\theta}) = 0, \quad i = 1,\ldots,k$, where each f_i possesses continuous partial derivatives. If A' denotes the constrained parameter space, and both $\boldsymbol{F}(\boldsymbol{\theta})$ and $\boldsymbol{V}(\boldsymbol{\theta}) = [\boldsymbol{J}^T(\boldsymbol{\theta}), \boldsymbol{F}^T(\boldsymbol{\theta})]$ do not change ranks in an open neighborhood of $\boldsymbol{\theta}_0 \in A'$, then $\boldsymbol{\theta}_0$ is locally identifiable if and only if $\boldsymbol{V}(\boldsymbol{\theta})$ has full rank at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.*

Note that, the constant rank assumption of $\mathbf{F}$ and $\mathbf{V}$ is needed for establishing the necessity, but not the sufficiency. Recall that a sufficient and necessary condition for the constrained CRB to exist is $|\mathbf{U}^T\mathbf{J}\mathbf{U}| \neq 0$. It can be shown that this condition is equivalent to the condition that $\mathbf{V}$ has full rank:

Lemma 1. *If $\boldsymbol{F}(\theta_0)$ has full column rank, then $\boldsymbol{V}(\theta_0)$ has full rank if and only if $\boldsymbol{U}^T(\theta_0)\boldsymbol{J}(\theta_0)\boldsymbol{U}(\theta_0)$ has full rank, where $\boldsymbol{U}(\theta_0)$ is a matrix whose columns form an orthonormal basis of the null space of $\boldsymbol{F}(\theta_0)$.*

[1] A measure μ_1 is said to be dominated by μ_0, if $\mu_0(F) = 0$ implies that $\mu_1(F) = 0$, where $F \in \mathcal{F}$.

Checking (4) –(14), we can verified that conditions (i)–(vi) of Theorem 1 are satisfied. So, the existence of a finite constrained CRB indicates local identifiability. It can also be shown that both $\mathbf{F}(\boldsymbol{\theta})$ and $\mathbf{J}(\boldsymbol{\theta})$ have constant ranks in a neighborhood of $\boldsymbol{\theta}$, if $\boldsymbol{\theta}$ is in the constrained parameter space, the entries of $\mathbf{B}$ are independently chosen, and M goes to ∞. So, we expect that, in most cases, $\mathbf{V}(\boldsymbol{\theta})$ will not change rank around $\boldsymbol{\theta}$ and local identifiability will lead to information regularity.

3.3. Conditions of Local Identifiability

From (4), we can see that a set of parameters $\boldsymbol{\theta}$ is not locally identifiable if and only if given any $\epsilon > 0$, there exists another set of parameters $\boldsymbol{\theta}'$ such that $\boldsymbol{\theta}'$ belongs to the constrained parameter space, $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\| < \epsilon$, and

$$\mathbf{HB} = \mathbf{H}'\mathbf{B}'. \tag{16}$$

With this fact, we can obtain the following result:

Theorem 2. *Assume that: i) the data symbols of different sources are independent; ii) the kth source's symbols $b_k(i),\ i = 0,\ldots,M-1$ are independent and identically distributed (*i.i.d.*) random variables drawn from the CM constellation $\mathbb{A}_k$ such that $|b_k(i)| = 1$ and $E[b_k(i)] = 0$; iii) the mixing matrix $\boldsymbol{H}$ has full column rank; and iv) at most one source has binary antipodal constellation. Then, the true parameter set $[\boldsymbol{H}, \boldsymbol{B}]$ is locally identifiable (with probability 1, as $M \to \infty$), if and only if at least one data symbol from each source is known (i.e., the only local ambiguity is the phase ambiguity).*

When all sources have binary antipodal modulations, more training symbols are needed to ensure local identifiability:

Theorem 3. *Assume that all sources use binary antipodal constellations, data symbols from different sources and in different time slots are independent from each other, the first T symbols of all sources, $[b_1(m),\ldots,b_K(m)], m = 0,\ldots,T-1$, are known, and $\boldsymbol{H}$ has full column rank. Then, the true parameter set $\{\boldsymbol{H}, \boldsymbol{B}\}$ is locally identifiable (with probability 1, as $M \to \infty$) if and only if no two rows of the matrix $\boldsymbol{B}_T$ are dependent, where*

$$\boldsymbol{B}_T = \begin{pmatrix} b_1(0) & \cdots & b_1(T-1) \\ \vdots & \ddots & \vdots \\ b_K(0) & \cdots & b_K(T-1) \end{pmatrix}. \tag{17}$$

Remark: From Theorem 3, the minimum number of training symbols needed for each source is $T_{\min} = \lceil \log_2 K \rceil + 1$. To see this, first note that we can assume without loss of generality that $b_1(0) = \cdots = b_K(0) = 1$. If $T \geq T_{\min}$, then we can let each of $\boldsymbol{b}'_k := [b_k(1),\ldots,b_k(T-1)]^T$

($k = 1, \ldots, K$) take one distinct value among all 2^{T-1} ($T-1$)-vectors with ± 1 entries. Apparently, no two rows of the resulting matrix are dependent. If, on the other hand, $T < T_{\min}$, then there exist k_1 and k_2, such that $\boldsymbol{b}'_{k_1} = \boldsymbol{b}'_{k_2}$ and row k_1 and row k_2 are dependent.

4. UNIQUE IDENTIFIABILITY

A definition of identifiability that is particularly relevant to the BSS problem with CM constraints is the global (unique) identifiability that is defined as follows [10]:

Definition 3. *BSS problem under CM constraints is uniquely identifiable, if any set of parameters $\boldsymbol{\theta}'$ that is observationally equivalent to the true parameter set $\boldsymbol{\theta}$ satisfies*

$$\boldsymbol{H}' = \boldsymbol{H}\boldsymbol{T}^{-1}, \tag{18}$$

and

$$\boldsymbol{B}' = \boldsymbol{T}\boldsymbol{B}, \tag{19}$$

where $\boldsymbol{T}$ is an admissible transformation, that is, $\boldsymbol{T} = diag(\alpha_1, \cdots, \alpha_K)\boldsymbol{P}$, in which $\boldsymbol{P}$ is a permutation matrix, and $|\alpha_k| = 1, \forall k = 1, \ldots, K$.

In [10], it is shown that for CM signals, if the number of observed samples is large enough and if the signals are rich in phase, then these signals are uniquely identifiable. More recently, necessary and sufficient conditions for unique identifiability were obtained in [2] using Kruskal's permutation lemma. Both the phase richness condition of [10], as well as the one in [2] are difficult to verify. In [4], persistently exciting sources were shown to guarantee unique identifiability, and a lower bound of the finite sample identifiability was given for *i.i.d.* circularly symmetric CM sources.

We investigate the unique identifiability of the blind source separation problem where the sources have arbitrary finite CM constellations, and obtain a necessary and sufficient condition for unique identifiability that is easy to verify. More specifically, we have proved that:

Theorem 4. *If the data symbols from the kth CM source belong to the alphabet set $\mathbb{A}_k$, and the probability assigned to each vector of $\mathbb{A}_1 \times \cdots \times \mathbb{A}_K$ is nonzero, then when the number of observed samples is large enough, the transmitted data and the mixing matrix can be uniquely identified if and only if at most one alphabet set is binary and antipodal.*

5. CONCLUSIONS

In this paper, we investigated the information regularity, local identifiability, and unique identifiability of the BSS problem under CM constraints. We established the link between the information regularity and the local identifiability in the BSS context. Sufficient and necessary conditions for local identifiability (and hence information regularity) were derived. We also studied the unique identifiability of blindly separating multiple CM sources with finite alphabet. Sufficient and necessary conditions for unique identifiability were obtained.

6. ACKNOWLEDGEMENT

The authors would like to thank Prof. Vincent Poor for helpful discussions.

References

[1] D. N. Godard. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Trans. Commun.*, 28(11):1867–1875, Nov. 1980.

[2] T. Jiang and N. D. Sidiropoulos. Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models with constant modulus constraints. *IEEE Transactions on Signal Processing*, 2003, submitted, downloadable from http://www.ece.umn.edu/users/jiang.

[3] R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas. Blind equalization using the constant modulus criterion: a review. *Proc. IEEE*, 86(10):1927–1950, Oct. 1998.

[4] A. Leshem, N. Petrochilos, and A. J. van der Veen. Finite sample identifiability of multiple constant modulus sources. *IEEE Trans. Inform. Theory*, 49(9):2314–2319, Sept. 2003.

[5] B. L. S. Prakasa Rao. *Identifiability in Stochastic Models*. Academic Press Inc., CA, USA, 1992.

[6] T. J. Rothenberg. Identification in parametric models. *Econometrica*, 39(3):577–591, May 1971.

[7] Y. Sato. A method of self-recovering equalization for multilevel amplitude-modulation systems. *IEEE Trans. Commun.*, 23(6):679–682, June 1975.

[8] P. Stoica and B. C. Ng. On the Cramér-Rao bound under parametric constraints. *IEEE Signal Processing Lett.*, 5(7):177–179, July 1998.

[9] J. R. Treichler and B. G. Agee. A new approach to multipath correction of constant modulus signals. *IEEE Trans. Acoust., Speech, Signal Processing*, 31(2):459–471, Apr. 1983.

[10] Y. Wang, Y. C. Pati, Y. M. Cho, A. Paulraj, and T. Kailath. A matrix factorization approach to signal copy of constant modulus signals arriving at an antenna array. In *Proc. 28th Conference on Information Sciences and Systems*, pages 178–183, Princeton, NJ, March 1994.

[11] Y. Yao and G. B. Giannakis. On regularity and identifiability of blind source separation under constant-modulus constraints. *IEEE Transactions on Signal Processing*, submitted 2003.