

ON MIMO INSTANTANEOUS BLIND IDENTIFICATION BASED ON THE EXPLOITATION OF THE TIME STRUCTURE OF SIGNALS USING ARBITRARY-ORDER CUMULANTS

J. van de Laar, Student member, IEEE

Technische Universiteit Eindhoven, Department of Electrical Engineering,
Signal Processing Systems Group, EH 3.27, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

ABSTRACT

This paper presents a new and unifying view at multiple-input multiple-output instantaneous blind identification based on the main assumptions that the cross-cumulant functions of the source signals of some arbitrary fixed order vanish for all time tuples and that the source auto-cumulant functions of the same order are linearly independent. Hence, the time structure of the signals is exploited. The main goal of the paper is to provide new insight into the algebraic and geometric structure of the problem, which is crucial for the development of solutions. The presented viewpoint is unifying in two senses. Firstly, the developed theory is general with respect to the considered order of the cumulants. Secondly, all types of statistical variability in the data, for example, the nonstationarity and the nonwhiteness, are incorporated into the problem in a unified manner.

1. INTRODUCTION

This paper is concerned with the Multiple-Input Multiple-Output (MIMO) Instantaneous Blind Identification (IBI) problem, shortly denoted by BI. In this problem, a number of mutually statistically independent sources are mixed by a MIMO instantaneous mixing system and only the mixed signals are available. The goal is to recover the mixing system from the observed mixtures of the sources only. It is widely recognized that many possible applications exist for BI. Common examples where BI is applied (in)directly are Instantaneous Blind Signal Separation (IBSS) and Direction Of Arrival (DOA) estimation. IBSS is slightly different from BI in the sense that the main goal is to recover the source signals instead of the mixing system. Several examples of IBSS can be found in the field of biomedical engineering, where the goal of several applications is to reveal independent sources in different kinds of signals like EEG's, ECG's, etc. DOA is in fact a parametrized version of BI. Several examples of DOA can be found in applications involving radar, sonar, etc. Many methods and algorithms for performing BI have been developed during the last decade. See, for example, [1], [2] and the references therein. Most of these methods can be classified into three distinctive approaches. The first approach exploits the *non-Gaussianity* of the sources and requires the use of higher order statistics (HOS) [2, 3]. The second approach assumes that the *sources are spatially uncorrelated and all have temporal correlation* on some domain of support, see, for example, [1], [2], [4] and the references therein. Finally, the third approach allows the exploitation of the *nonstationarity* of the sources under the assumption that the sources have different nonstationarity properties [5]. The latter two approaches are both subclasses of the same unifying principle, namely the exploitation of the second order time

structure of the signals. This paper is a generalization of a part of the work presented in [6]. It presents a new and unifying view at BI based on the main assumptions that the cross-cumulant functions of the source signals of some arbitrary fixed order are zero for all time tuples and that the source auto-cumulant functions of the same order are linearly independent, thereby exploiting the time structure of the signals for some arbitrary order. The first assumption is reasonable because many signals arising in practical situations are generated independently of each other by physically different sources at different positions, and the second because these sources have in general different temporal characteristics. Due to space limitations, the proofs for the results in the paper have been omitted and will be published in a subsequent paper. Based on the new insight, solutions that are general with respect to the order of the used cumulants can be developed. The paper also provides a basis for understanding and solving more advanced problems, such as the blind identification of convolutive MIMO systems.

The outline of the paper is as follows. Firstly, the used notation is introduced in Section 2, along with some definitions. Next, the considered BI model (including assumptions) is explained in detail in Section 3. Then, the algebraic and geometric structure of this model are discussed in Section 4, along with some examples. Finally, the conclusions are discussed in Section 5.

2. NOTATION AND DEFINITIONS

Both sub- and superscript indices are used to index quantities. Column vectors are denoted by lower case boldface letters, e.g. \mathbf{v} , and their elements are indexed by superscript indices, e.g. $\mathbf{v} = [v^i]$. The vector space of real column vectors of length N is denoted by \mathbb{R}^N . Matrices are denoted by upper case boldface letters, e.g. \mathbf{V} . The elements of matrices are denoted by lower case letters with both sub- and superscript indices, e.g. $\mathbf{V} = [v_j^i]$. The superscript indices correspond to row indices and the subscript indices correspond to column indices (see (4), for example). The space of matrices of size M by N with real-valued elements is denoted by \mathbb{R}_N^M . Usually, the dimensions of matrices and vectors are clear from the context, otherwise they are stated explicitly. The linear span of a set of vectors \mathcal{V} in a linear vector space is denoted by $\mathcal{L}(\mathcal{V})$. The cardinality of a set \mathcal{I} is denoted by $|\mathcal{I}|$.

Discrete-time functions and signals have their time index between square brackets. For example, $\mathbf{x}[n]$ represents a signal vector containing signal values at discrete time n and $x^i[n]$ denotes its i -th component. Finally, the operator that gives the expected value of a random variable will be denoted by $E\{\cdot\}$. For convenience, some additional notations and definitions are given and then explained (p and G are positive integers):

$$\bullet \tilde{v}^{i_p} \triangleq \prod_{m=1}^p v^{i_m} = v^{i_1} \dots v^{i_p} \quad (1a)$$

$$\bullet \mathbf{i}_p \triangleq (i_1, \dots, i_p) \in \mathbb{Z}_p \quad (1b)$$

$$\bullet (i)_p \triangleq (i, \dots, i) \in \mathbb{Z}_p \quad (1c)$$

$$\bullet \mathcal{I}_{t,p}^G \triangleq \{\mathbf{i}_p \mid 1 \leq i_1, \dots, i_p \leq G\}, |\mathcal{I}_{t,p}^G| = (G)^p \quad (1d)$$

$$\bullet \mathcal{I}_{a,p}^G \triangleq \{\mathbf{i}_p \mid 1 \leq i_1 \leq \dots \leq i_p \leq G\}, |\mathcal{I}_{a,p}^G| = \binom{G+p-1}{p} \quad (1e)$$

$$\bullet \mathcal{I}_{e,p}^G \triangleq \{(i)_p \mid 1 \leq i \leq G\} = \{(i)_p\}^{1 \leq i \leq G}, |\mathcal{I}_{e,p}^G| = G \quad (1f)$$

$$\bullet \mathcal{I}_{c,p}^G \triangleq \mathcal{I}_{t,p}^G \setminus \mathcal{I}_{e,p}^G, |\mathcal{I}_{c,p}^G| = (G)^p - G. \quad (1g)$$

The symbol \tilde{v}^{i_p} is merely a shorthand notation for the product $v^{i_1} \dots v^{i_p}$. An ordered tuple of integers like \mathbf{i}_p is used for three purposes. Firstly, the indices in a tuple can be used as running variables in summations or products. Secondly, the indices in a tuple can be used as free indices. Finally, a tuple can contain discrete time indices, in which case it is used as a p -dimensional argument to functions. Typically, the symbols \mathbf{i}_p and \mathbf{j}_p denote tuples used for the first two purposes, while the symbol \mathbf{n}_p denotes a tuple of discrete time indices. The tuple $(i)_p$ is a shorthand notation for \mathbf{i}_p that is used when all indices are equal. The subscript 't' in the set $\mathcal{I}_{t,p}^G$ stands for "total" because all indices in the tuple range from 1 to G . Likewise, the subscript 'a' in $\mathcal{I}_{a,p}^G$ stands for "ascending" because the indices in the tuple are ordered in ascending order. The subscript 'e' in $\mathcal{I}_{e,p}^G$ stands for "equal" because only those tuples are in the set whose indices are equal. Finally, the subscript 'c' in $\mathcal{I}_{c,p}^G$ stands for "cross" because only those tuples are in the set whose indices are unequal or "cross".

Given a length- G time dependent random vector $\mathbf{v}[n]$, the p -th order cross-cumulant function $\kappa_v^{i_p}[\mathbf{n}_p]$ of a subset $\{v^{i_1}[n_1], \dots, v^{i_p}[n_p]\}$ of p components indexed by i_1, \dots, i_p at possibly different times n_1, \dots, n_p is defined as:

$$\kappa_v^{i_p}[\mathbf{n}_p] \triangleq \text{cum}\left(v^{i_1}[n_1], \dots, v^{i_p}[n_p]\right) \quad \forall \mathbf{i}_p \in \mathcal{I}_{t,p}^G, \forall \mathbf{n}_p \in \mathbb{Z}_p,$$

where $\text{cum}(\cdot, \dots, \cdot)$ denotes the multi-argument cumulant function. See [7] for a detailed discussion about cumulants.

3. BI MODEL

An instantaneous mixing model with S source and D sensor signals is considered. Both the source and the sensor signals are assumed to be real-valued. The sensor signals $x^i[n]$ are given by:

$$x^i[n] = \sum_{j=1}^S a_j^i s^j[n] \quad \forall 1 \leq i \leq D, \forall n \in \mathbb{Z}, \quad (2)$$

where a_j^i is the real instantaneous transfer from the j -th source to the i -th sensor and $s^j[n]$ is the j -th source signal at time n . In matrix notation, (2) is written as:

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] = \sum_{j=1}^S \mathbf{a}_j s^j[n] \quad \forall n \in \mathbb{Z}, \quad (3)$$

where

$$\mathbf{x}[n] \triangleq \begin{bmatrix} x^1[n] \\ \vdots \\ x^D[n] \end{bmatrix}, \mathbf{s}[n] \triangleq \begin{bmatrix} s^1[n] \\ \vdots \\ s^S[n] \end{bmatrix} \text{ and } \mathbf{a}_j \triangleq \begin{bmatrix} a_j^1 \\ \vdots \\ a_j^D \end{bmatrix}$$

are vectors of sensor signals, source signals, and mixing elements respectively. The vectors $\mathbf{x}[n]$ and \mathbf{a}_j are elements of \mathbb{R}^D , while $\mathbf{s}[n]$ is an element of \mathbb{R}^S . The real-valued mixing matrix \mathbf{A} of size D by S can be written as:

$$\mathbf{A} = \begin{bmatrix} a_1^1 & \dots & a_S^1 \\ \vdots & \ddots & \vdots \\ a_1^D & \dots & a_S^D \end{bmatrix} \in \mathbb{R}_S^D. \quad (4)$$

The number of sources S may be smaller than, equal to, or larger than the number of sensors D . Considering l -th order cumulants, the main assumptions for the BI model, on which the results in this paper are based, are given in the following list:

- AS1: The mixing matrix \mathbf{A} is an element of \mathbb{R}_S^D and $\text{rank}(\mathbf{A}) = \min(D, S)$, i.e. \mathbf{A} is full rank
- AS2: If $D < S$, each tuple of D different columns is linearly independent
- AS3: The source signals are zero mean real-valued processes with zero l -th order cross-cumulant functions
- AS4: The l -th order source auto-cumulant functions are linearly independent.

Two indeterminacies are involved in blind identification that cannot be resolved without any prior knowledge. Firstly, from (3) it is clear that permuting the columns of \mathbf{A} together with an equal permutation of the source signals in $\mathbf{s}[n]$ still yields the same vector of sensor signals $\mathbf{x}[n]$. This means that the order of the columns, respectively sources, cannot be determined. Secondly, as is also clear from (3), scaling the columns of \mathbf{A} together with a corresponding inverse scaling of the source signals in $\mathbf{s}[n]$ also yields the same vector of sensor signals $\mathbf{x}[n]$. This means that the columns and sources can only be recovered up to a scale factor. Taking into account these two indeterminacies, the goal of BI is to recover the columns of the mixing system in arbitrary order and with arbitrary scaling, i.e. an estimate $\hat{\mathbf{A}}$ of the mixing system ideally satisfies $\hat{\mathbf{A}} = \mathbf{A}\mathbf{P}\mathbf{D}$, where \mathbf{P} is some permutation matrix and \mathbf{D} is some nonsingular diagonal matrix ($\mathbf{P}, \mathbf{D} \in \mathbb{R}_S^S$). For many applications in BI, most relevant information is in the "directions" of the columns rather than in their order or magnitudes.

4. ALGEBRAIC AND GEOMETRIC STRUCTURE

Intuitively it is clear that *all information about the mixing system that can be deduced from the set of l -th order sensor cumulant functions is present in the interrelationships between all these functions*. It will be shown in this section that when AS1-AS4 are satisfied, these interrelationships can be represented by a system of homogeneous polynomial equations satisfied by the elements of the columns of the mixing matrix. Several aspects of the algebraic and geometric structure of BI will be highlighted.

4.1. Linear span of sensor cumulant functions

AS3 and AS4 only need to hold over some finite region of support (ROS) $\Omega_s^{n_l} \subseteq \mathbb{Z}_l$ in the domain of length- l time tuples, i.e. the \mathbf{n}_l -domain. This ROS can be specified by a set of time tuples for which the condition holds. In practice, $\Omega_s^{n_l}$ has to be chosen in such a way that the conditions in AS1-AS4 hold. Furthermore, the tuples $\mathbf{n}_l \in \Omega_s^{n_l}$ can be chosen in such a way that different types of l -th order statistical information in the data are exploited. For example, if it is desired to exploit both the nonstationarity and the nonwhiteness of the source signals, the "time-distance" between different tuples must be large enough in order to exploit the nonstationarity, e.g. p must be large enough for two different tuples \mathbf{n}_l and $\mathbf{n}_l + p \cdot (1)_l$, and $n_1 \neq n_2 \neq \dots \neq n_l$ for some tuples in order to exploit the nonwhiteness. The values of a sensor cumulant

function for different time tuples are used in the same way. Hence, *all types of statistical variability in the data can be incorporated into the problem in a unified manner*. In the sequel, it is tacitly assumed that all real-valued functions depending on the integer-valued time tuple \mathbf{n}_l are defined only over $\Omega_s^{\mathbf{n}_l}$. For example, AS3 implies that the l -th order source cross-cumulant functions vanish on $\Omega_s^{\mathbf{n}_l}$:

$$\kappa_s^{\mathbf{i}_l}[\mathbf{n}_l] \triangleq \text{cum}(s^{\mathbf{i}_1}[n_1], \dots, s^{\mathbf{i}_l}[n_l]) = 0 \quad \forall \mathbf{i}_l \in \mathcal{I}_{c,l}^S, \quad \forall \mathbf{n}_l \in \Omega_s^{\mathbf{n}_l}.$$

Using AS3 and AS4, together with the properties of cumulants, all l -th order sensor cumulant functions $\{\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D}$ can be expressed in the l -th order source auto-cumulant functions as follows:

$$\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l] = \sum_{j=1}^S \tilde{a}_j^{\mathbf{i}_l} \kappa_s^{(j)\mathbf{i}_l}[\mathbf{n}_l] \quad \forall \mathbf{i}_l \in \mathcal{I}_{a,l}^D, \quad \forall \mathbf{n}_l \in \Omega_s^{\mathbf{n}_l}. \quad (5)$$

From this equation, it can easily be seen that $\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l] = \kappa_x^{\sigma(\mathbf{i}_l)}[\mathbf{n}_l]$ for all $\mathbf{i}_l \in \mathcal{I}_{a,l}^D$ and for all possible permutations σ of the indices in \mathbf{i}_l . Hence, the sensor cumulant functions are only essentially different for the set of index-tuples $\mathcal{I}_{a,l}^D$ with cardinality $|\mathcal{I}_{a,l}^D| = \binom{D+l-1}{l}$ (see (1e)). Therefore, in the sequel the set $\{\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D}$ is considered instead of $\{\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D}$. From AS1-AS4, it can be proven that the following proposition holds with probability 1 when the columns of \mathbf{A} are drawn independently from a continuous probability density distribution:

Proposition 1 Assume that AS1-AS4 are satisfied and $\binom{D+l-1}{l} \geq S$, then $\mathcal{L}(\{\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D}) = \mathcal{L}(\{\kappa_s^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{c,l}^S})$.

Hence, the linear vector space spanned by the sensor cumulant functions equals the space spanned by the source auto-cumulant functions. Note that this is intuitively clear from (5) when the transformation defined by the $\tilde{a}_j^{\mathbf{i}_l}$'s is full rank.

4.2. Derivation of system of homogeneous polynomial equations satisfied by columns of mixing matrix

Proposition 1 and AS4 imply that $\dim(\mathcal{L}(\{\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D})) = \dim(\mathcal{L}(\{\kappa_s^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{c,l}^S})) = S$. Therefore, when the number of functions $\binom{D+l-1}{l}$ in the set $\{\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D}$ is larger than the number of sources S , the sensor cumulant functions are linearly dependent. This implies that there exist nonzero and non-unique sets of coefficients $\{\varphi_{\mathbf{i}_l}^q\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D}$ indexed by an (arbitrarily integer-valued) index q such that:

$$\sum_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D} \varphi_{\mathbf{i}_l}^q \kappa_x^{\mathbf{i}_l}[\mathbf{n}_l] = 0 \quad \forall q \in \mathcal{Q}, \quad \forall \mathbf{n}_l \in \Omega_s^{\mathbf{n}_l}. \quad (6)$$

The index q is chosen to be an element of the set $\mathcal{Q} \triangleq \{1, \dots, Q\}$, where Q is the maximum number of linearly independent equations in (6). The information about the mixing system that is present in the sensor cumulant functions can be represented by the set $\Phi \triangleq \{\{\varphi_{\mathbf{i}_l}^q\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D}\}_{q \in \mathcal{Q}}$. The value of Q and a nontrivial set Φ can be computed from the available data using the Singular Value Decomposition (SVD) of the sensor cumulant functions, which is briefly discussed in Section 4.3. Note that all linear combinations of the sets in Φ also satisfy (6). Now, substituting (5) into (6) and using AS4, it can be shown that:

$$\sum_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D} \varphi_{\mathbf{i}_l}^q \tilde{a}_j^{\mathbf{i}_l} = 0 \quad \forall q \in \mathcal{Q}, \quad 1 \leq j \leq S. \quad (7)$$

This system of equations describes the relation between the coefficients of the mixing matrix \mathbf{A} and the set of coefficients Φ . It can easily be seen that the unknowns $\{a_j^{\mathbf{i}_l}\}_{1 \leq j \leq S}^{1 \leq l \leq D}$ in each individual equation are the elements of one column \mathbf{a}_j of \mathbf{A} . Hence, all columns $\mathbf{a}_1, \dots, \mathbf{a}_S$ satisfy the same system of polynomial equations. More specifically, let $\mathbf{z} = [z^1, \dots, z^D]^T$ be an arbitrary column of \mathbf{A} , then (7) states that each column \mathbf{z} of \mathbf{A} satisfies the following system of D -variate polynomial equations of degree l :

$$f^q(\mathbf{z}) \triangleq \sum_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D} \varphi_{\mathbf{i}_l}^q z^{\mathbf{i}_l} = 0 \quad \forall q \in \mathcal{Q}. \quad (8)$$

Each $f^q(\mathbf{z}) = f^q(z^1, \dots, z^D)$ is a D -variate l -homogeneous polynomial, meaning that (denoting η to the power l by $(\eta)^l$):

$$f^q(\eta \mathbf{z}) = (\eta)^l f^q(\mathbf{z}) \quad \forall q \in \mathcal{Q}, \quad \forall \eta \in \mathbb{R}, \quad \forall \mathbf{z} \in \mathbb{R}^D. \quad (9)$$

The following implication is immediate from (9):

$$f^q(\mathbf{z}) = 0 \quad \forall q \in \mathcal{Q} \Rightarrow f^q(\eta \mathbf{z}) = 0 \quad \forall q \in \mathcal{Q}, \quad \forall \eta \in \mathbb{R}. \quad (10)$$

Hence, if \mathbf{z} is a solution of the system $\{f^q(\mathbf{z}) = 0\}_{q \in \mathcal{Q}}$, then also $\eta \mathbf{z}$ is a solution. For this reason, a constraint has to be imposed on the norm of the solution vectors \mathbf{z} . This is a logical result of the *scaling indeterminacy* inherent to BI (see Section 3).

4.3. Singular value decomposition of cumulant functions

The (function-valued) vector $\underline{\kappa}_x[\mathbf{n}_l]$, defined by stacking all functions in the set $\{\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]\}_{\mathbf{i}_l \in \mathcal{I}_{a,l}^D}$ on top of each other, can be written in "SVD-form" as $\underline{\kappa}_x[\mathbf{n}_l] = \sum_{k=1}^S \sigma_k \underline{\mathbf{u}}_k v^k[\mathbf{n}_l]$ with the usual SVD-properties $\sigma_k \in \mathbb{R}$, $\langle \underline{\mathbf{u}}_m, \underline{\mathbf{u}}_n \rangle_1 = \delta_{mn} \quad \forall 1 \leq m, n \leq N$ and $\langle v^m[\mathbf{n}_l], v^n[\mathbf{n}_l] \rangle_2 = \delta^{mn} \quad \forall 1 \leq m, n \leq M$. Here, $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are properly defined inner products, $M \triangleq |\Omega_s^{\mathbf{n}_l}|$ and $N \triangleq |\mathcal{I}_{a,l}^D| = \binom{D+l-1}{l}$. Indexing the elements of the vectors $\underline{\mathbf{u}}_k$ by the tuples \mathbf{i}_l , in correspondence with the way the $\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]$'s are stacked in $\underline{\kappa}_x[\mathbf{n}_l]$, each function $\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]$ can be written as $\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l] = \sum_{k=1}^S \sigma_k u_k^{\mathbf{i}_l} v^k[\mathbf{n}_l]$. Now, from (6) and the properties of the SVD, it follows that all valid choices for the coefficients $\varphi_{\mathbf{i}_l}^q$ are: $\varphi_{\mathbf{i}_l}^q = \sum_{n=S+1}^N \alpha_n^q u_n^{\mathbf{i}_l} \quad \forall q \in \mathcal{Q}, \quad \forall \mathbf{i}_l \in \mathcal{I}_{a,l}^D, \quad \forall \alpha_n^q \in \mathbb{R}$. A natural choice for the coefficients $\varphi_{\mathbf{i}_l}^q$ in (8) is given by $\varphi_{\mathbf{i}_l}^q = u_{q+S}^{\mathbf{i}_l}$, i.e. $\alpha_n^q = \delta_n^{q+S}$. Since $\dim(\mathcal{L}(\{\{\underline{\mathbf{u}}_n\}_{S+1 \leq n \leq N}\})) = N - S$, the maximum number of linearly independent equations in system (8) equals: $Q = N - S = \binom{D+l-1}{l} - S$. In practice, a matrix $\mathbf{C}_x \in \mathbb{R}_M^N$ corresponding to $\underline{\kappa}_x[\mathbf{n}_l]$ is used as input to some SVD-algorithm. This matrix is constructed by putting the M estimated function values $\{\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l], \mathbf{n}_l \in \Omega_s^{\mathbf{n}_l}\}$ of each function $\kappa_x^{\mathbf{i}_l}[\mathbf{n}_l]$ in the \mathbf{i}_l -th row of $\mathbf{C}_x \in \mathbb{C}_M^N$.

4.4. Geometric structure

The purpose of this section is to clarify the geometric structure of the problem. Equation (10) means that the *zero contour levels of the l -homogeneous D -variate polynomials in the set $\{f^q(\mathbf{z})\}_{q \in \mathcal{Q}}$ define cones in the D -dimensional space \mathbb{R}^D . Therefore, geometrically, (7) and (8) signify that the S columns of the mixing matrix are determined by the intersections between $Q = \binom{D+l-1}{l} - S$ of these cones. This is illustrated in the following two examples. Each example deals with a two-sensor scenario because this easily allows visualization of the involved geometry. Because the purpose of this section is to demonstrate the geometric properties, the source auto-cumulant functions have been generated at random. Since the only fundamental assumption on these functions is that they are linearly independent, this is a completely valid approach.*

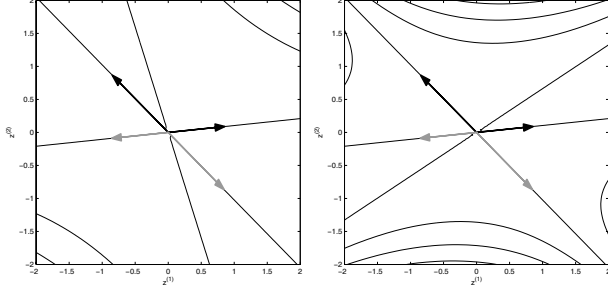


Fig. 1. Contour plots of functions in example 1.

It has been verified that simulations with real-world data yield similar functions with exactly the same zero contour levels.

Example 1: two sensors, two sources and third order cumulants

In this example with $D = 2$, $S = 2$ and $l = 3$, the following mixing matrix is used: $\mathbf{A} = \begin{bmatrix} 0.86 & -0.85 \\ 0.09 & 0.87 \end{bmatrix}$. The number of linearly independent equations in (8) equals $Q = \binom{2+3-1}{3} - 2 = 2$. Using the SVD method discussed in Section 4.3, (the coefficients of) two such linearly independent equations can be computed. The contour plots of each of the functions corresponding to the left hand sides of the two resulting linearly independent equations are plotted in Fig. 1. Each black arrow in a figure denotes a column of the mixing matrix and each grey arrow is the negative of a black arrow. The three crossing straight lines through the middle of each of the figures correspond to the zero contour level of the corresponding function (in two-dimensional space, cones are described by lines through the origin). It is evident from the contour plots that the columns of the mixing matrix are uniquely determined by the intersections between the zero contour levels of the two functions.

Example 2: two sensors, three sources and third order cumulants

In this example with $D = 2$, $S = 3$ and $l = 3$, the mixing matrix $\mathbf{A} = \begin{bmatrix} 0.86 & -0.85 & -0.44 \\ 0.09 & 0.87 & -0.43 \end{bmatrix}$ is used. Now there is only $Q = \binom{2+3-1}{3} - 3 = 1$ equation.

The contour plot of the corresponding function is plotted in Fig. 2. The three crossing straight lines in the middle of the figure correspond to the zero contour level. It can easily be seen that the arrows representing the (negative of the) columns of the mixing matrix are exactly in the directions of the lines describing the zero contour level of the function, thereby revealing that all three columns of the mixing matrix are determined uniquely by the zero contour level.

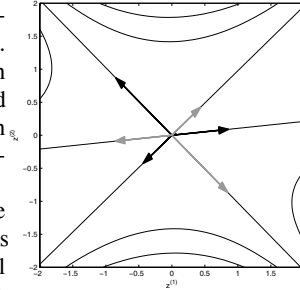


Fig. 2. Contour plot of function in example 2.

The zero contour level of a D -variate homogenous polynomial is in general, i.e. the non-degenerate case, a $(D - 1)$ -dimensional surface embedded in a D -dimensional Euclidian space. From geometric intuition, it is clear that in a D -dimensional space $D - 1$ surfaces of dimension $D - 1$ are required to define one-dimensional solution sets (lines through the origin), which in the current BI problem define (scalar multiples of) the columns of \mathbf{A} . Heuristically, from $Q = \binom{D+l-1}{l} - S \geq D - 1$ (each equation defines one surface), it follows that the maximum number S_{\max} of columns of \mathbf{A} that can be identified equals $S_{\max} = \binom{D+l-1}{l} - (D - 1)$.

5. CONCLUSIONS

This paper has presented a new and unifying view at Multiple-Input Multiple-Output Instantaneous Blind Identification based on the assumptions that the cross-cumulant functions of the source signals of some arbitrary fixed order l are zero and that the l -th order source auto-cumulant functions are linearly independent, thereby exploiting the time structure of the signals. Many real-world signals satisfy these assumptions. The presented viewpoint is unifying in two senses. Firstly, the developed theory is general with respect to the considered order of the cumulants. Secondly, all types of statistical variability in the data, for example, the nonstationarity and the nonwhiteness, are incorporated into the problem in a unified manner. Insight into the algebraic structure of the problem has been provided by showing that the columns of the mixing system satisfy a system of D -variate l -homogeneous polynomial equations, where D is the number of sensors and l is the considered cumulant order. Likewise, the geometric structure of the problem has been clarified by showing that the zero contour levels of the homogeneous polynomials are cones in D -dimensional space and that the columns of the mixing matrix are defined by the intersections between all these cones. Furthermore, it has been made plausible that, given D and l , the maximum number of columns of the mixing matrix that can be identified equals $\binom{D+l-1}{l} - (D - 1)$. The new viewpoint can be used to develop new solutions to the BI problem and also for understanding and tackling more advanced problems. A possible solution for the BI problem, based on the developed theory, will be presented in a subsequent paper.

Acknowledgements

The author would like to thank Ir. E. Habets, Dr. ir. E. Sarroukh and Dr. ir. D. Schobben very much for carefully proof reading the manuscript, and his supervisor Dr. ir. P. Sommen for his support.

6. REFERENCES

- [1] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, Wiley, 2002.
- [2] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley, New York, 2001.
- [3] J.-F. Cardoso, "Iterative techniques for blind source separation using only fourth-order cumulants," in *Proc. EUSIPCO*, Brussels, Belgium, 1992, pp. 739–742.
- [4] A. Cichocki and R. Thawonmas, "On-line algorithm for blind signal extraction of arbitrarily distributed, but temporally correlated sources using second order statistics," *Neural Processing Letters*, vol. 12, pp. 91–98, 2000.
- [5] S. Choi and A. Cichocki, "Blind separation of nonstationary sources in noisy mixtures," *Electronics Letters*, vol. 36, no. 9, pp. 848–849, apr 2000.
- [6] J. van de Laar, "Instantaneous blind source separation based on the exploitation of temporal correlations and nonstationarity," in *Proc. 14th Annual Workshop on Circuits, Systems and Signal Processing (ProRISC 2003)*, Veldhoven, The Netherlands, November 2003, pp. 391–398. Downloadable from: <http://www.sps.ele.tue.nl/members/J.Laar/index.html>.
- [7] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, no. 3, pp. 278–305, March 1991.