The Complex Version of the Minimum Support Criterion

Sergio Cruces, Auxiliadora Sarmiento, and Iván Durán

Dpto. de Teoría de la Señal y Comunicaciones*, Escuela de Ingenieros,
Universidad de Sevilla, Camino de los descubrimientos. s/n,
41092-Sevilla, Spain
sergio@us.es

Abstract. This paper addresses the problem of the blind signal extraction of sources by means of an information theoretic and geometric criterion. Our main result is the extension of the minimum support criterion to the case of mixtures of complex signals. This broadens the scope of its possible applications in several fields, such as communications.

1 Introduction

The paradigm of linear ICA consists in the decomposition of the observations into a linear combination of independent components (or sources), plus some added noise. The problem is named blind signal separation (BSS) when one tries to recover all the involved sources, whereas, it is named blind signal extraction (BSE) when one is interested in one or a subset of sources.

In the late 1970s, a powerful contrast function was proposed to solve the problem of blind deconvolution [1]. This contrast function, which minimizes the Shannon entropy of the output under a variance constraint on its signal component, was a direct consequence of the entropy power inequality [2]. A similar principle was much latter rediscovered in the field of ICA, where the minimization of the mutual information of the outputs, under a covariance constraint, was seen as a natural contrast function to solve the BSS problem [3]. Indeed, provided that the inverse system exists, there is a continuum of contrast functions based on marginal entropies which allows the simultaneous extraction of an arbitrary number of source signals [4].

Since them, the ICA literature explored the properties of other generalized entropy measures, like Renyi's entropies, to obtain novel information theoretic contrast functions [5,6]. A criterion, which involved the minimization of the sum of ranges of the outputs, was proposed in [7] for solving the BSS problem with order statistics. Some time latter, we independently proposed a similar criterion (the minimum support criterion) which minimizes zero order Renyi's entropy of the output for solving the problem of the blind extraction of one of the sources [8]. In [9] the minimum range criterion for extraction was rediscovered and proved

* Part of this research was supported by the MCYT Spanish project TEC2004-06451-C05-03.

M.E. Davies et al. (Eds.): ICA 2007, LNCS 4666, pp. 17–24, 2007.

to be free of erroneous minima, a very desirable property. The minimum support and the minimum range criteria coincide only when all the involved signals have convex support, otherwise they differ [10].

In this paper, we retake the minimum support criterion and extend its role as contrast function for mixtures of complex source signals.

The paper is organized as follows. In section 2 we present the signal model. Section 3 and section 4 detail some useful results and geometrical object definitions. Section 5 presents the complex version of the minimum support criterion and other extensions. Section 6 presents the simulations, and finally, section 7 discusses the conclusions.

2 Signal Model and Notation

We consider the standard linear mixing model of complex stationary processes in a noiseless situation. The observations random vector obeys the following equation

$$\boldsymbol{X} = \mathbf{A}\boldsymbol{S}, \tag{1}$$

where $\boldsymbol{S} = [S_1, \cdots, S_n]^T \in \mathbb{C}^{n\times 1}$ is a random vector with independent components, and $\mathbf{A} \in \mathbb{C}^{n\times n}$ is a mixing matrix of complex elements.

In order to extract one non-Gaussian source from the mixture, one can compute the inner product of the observations with the vector $\mathbf{u}$, to obtain the output random variable or estimated source

$$Y = \mathbf{u}^H \boldsymbol{X} = \mathbf{g}^H \boldsymbol{S}, \tag{2}$$

where $\mathbf{g}^H = \mathbf{u}^H \mathbf{A}$ denotes the vector with the coefficients of the mixture of the sources at the output.

The Darmois-Skitovitch theorem [3] guarantees the identifiability of non-Gaussian complex sources, up to a permutation, scaling and phase term. Let $\mathbf{e_i}$, $i = 1, \ldots, n$, denote the coordinate vectors; one source is extracted when

$$\mathbf{g} = \|\mathbf{g}\| e^{j\theta} \mathbf{e_i}, \quad i \in \{1, \ldots, n\}. \tag{3}$$

3 Support Sets and Geometric Inequalities

Consider two m-dimensional vectors of random variables $\boldsymbol{A}$ and $\boldsymbol{B}$, whose respective densities are $f_{\boldsymbol{A}}(\boldsymbol{a})$ and $f_{\boldsymbol{B}}(\boldsymbol{b})$.

Definition 1. *The support set of a random vector $\boldsymbol{A}$, which we denote by $\mathcal{S}_{\boldsymbol{A}} = supp\{\boldsymbol{A}\}$, is the set of points for which its probability density function is nonzero, i.e., $\mathcal{S}_{\boldsymbol{A}} = \{\boldsymbol{a} \in \mathbb{R}^m : f_{\boldsymbol{A}}(\boldsymbol{a}) > 0\}$.*

Definition 2. *The convex hull of the set $\mathcal{S}_{\boldsymbol{A}}$, which we denote by $\mathcal{S}_{\breve{A}} = conv\, \mathcal{S}_{\boldsymbol{A}}$, is the intersection of all convex sets in $\mathbb{R}^m$ which contain $\mathcal{S}_{\boldsymbol{A}}$.*

In this paper, we will consider that all the support sets of our interest are compact (bounded and closed), thus we will make no distinction between *convex hull* and the *convex closure*.

Definition 3. *The Minkowski sum of two given sets $\mathcal{S}_{\boldsymbol{A}}$ and $\mathcal{S}_{\boldsymbol{B}}$ is defined as the set $\mathcal{S}_{\boldsymbol{A}} \oplus \mathcal{S}_{\boldsymbol{B}} = \{\boldsymbol{a} + \boldsymbol{b} : \boldsymbol{a} \in \boldsymbol{A}, \boldsymbol{b} \in \boldsymbol{B}\}$ which contains all the possible sums of the elements of $\mathcal{S}_{\boldsymbol{A}}$ with the elements of $\mathcal{S}_{\boldsymbol{B}}$.*

In the case of two independent random vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, it is easy to observe that the support of their sum $\mathcal{S}_{\boldsymbol{A}+\boldsymbol{B}}$ is equal to the Minkowski sum of the original support sets $\mathcal{S}_{\boldsymbol{A}} \oplus \mathcal{S}_{\boldsymbol{B}}$.

The following famous theorem in geometry establishes the superadditivity of the n-th root of the volume of a Minkowsky sum of two sets.

Theorem 1 (Brunn-Minkowski inequality in $\mathbb{R}^m$). *Let $\mathcal{S}_{\boldsymbol{A}}$ and $\mathcal{S}_{\boldsymbol{B}}$ be nonempty bounded Lebesgue measurable sets in $\mathbb{R}^m$ such that $\mathcal{S}_{\boldsymbol{A}} \oplus \mathcal{S}_{\boldsymbol{B}}$ is also measurable. Then*

$$\mu_m(\mathcal{S}_{\boldsymbol{A}} \oplus \mathcal{S}_{\boldsymbol{B}})^{1/m} \geq \mu_m(\mathcal{S}_{\boldsymbol{A}})^{1/m} + \mu_m(\mathcal{S}_{\boldsymbol{B}})^{1/m} \tag{4}$$

The Brunn-Minkowski inequality is formulated for nonempty bounded measurable sets in $\mathbb{R}^m$. However, we want to apply it to obtain a criterion that works for complex data. The next section will help us in this task.

4 Isomorphisms Between Real and Complex Sets

The following bijective mapping

$$c = \Re\{c\} + j\Im\{c\} \mapsto T_1(c) = \begin{pmatrix} \Re\{c\} \\ \Im\{c\} \end{pmatrix}. \tag{5}$$

defines a well-known isomorphism between the space of complex scalar numbers $\mathbb{C}$ and the vector space $\mathbb{R}^2$ with the operation of addition and multiplication by a real number. However, the multiplication of two complex numbers is not naturally carried in $\mathbb{R}^2$. Hopefully, there is another isomorphism between the space of complex scalar numbers $c \in \mathbb{C}$ and the subfield of the M^2 vector space of real 2×2 which carries the operation of multiplication. It is defined by the following bijective mapping

$$c = \Re\{c\} + j\Im\{c\} \mapsto T_2(c) = \begin{pmatrix} \Re\{c\} & -\Im\{c\} \\ \Im\{c\} & \Re\{c\} \end{pmatrix}. \tag{6}$$

The two previously presented isomorphisms allow one to express the following operation of complex random variables

$$Y = \sum_{i=1}^{n} g_i^* S_i \tag{7}$$

as the equivalent real operation between real vectors of random variables

$$\begin{pmatrix} \Re\{Y\} \\ \Im\{Y\} \end{pmatrix} = \sum_{i=1}^{n} \begin{pmatrix} \Re\{g_i\} & \Im\{g_i\} \\ -\Im\{g_i\} & \Re\{g_i\} \end{pmatrix} \begin{pmatrix} \Re\{S_i\} \\ \Im\{S_i\} \end{pmatrix} . \tag{8}$$

Moreover, to any given set of complex numbers $S_{\boldsymbol{A}}$ we can associate an area $\mu_2(A)$ which represents the area of the equivalent set $T_1(S_{\boldsymbol{A}}) = \{T_1(a) : a \in S_{\boldsymbol{A}}\}$ of $\mathbb{R}^2$ defined by the real and imaginary pairs of coordinates. Thus, the measure of the support of a complex scalar random variable is defined as the measure of support of the random vector formed by its real and imaginary parts

$$\mu_1^c(\mathcal{S}_C) \equiv \mu_2 \left(\text{supp} \left\{ \begin{pmatrix} \Re\{C\} \\ \Im\{C\} \end{pmatrix} \right\} \right) . \tag{9}$$

Note that the measure of the support of the complex scalar multiplication $g_i^* S_i$ is invariant to the phase of the complex scalar g_i^*, because the phase term only implies a rotation of the space. This can be better seen from the fact that

$$\mu_1^c \left(\mathcal{S}_{(g_i^* S_i)} \right) = \begin{vmatrix} \Re\{g_i\} & \Im\{g_i\} \\ -\Im\{g_i\} & \Re\{g_i\} \end{vmatrix} \mu_2 \left(\text{supp} \left\{ \begin{pmatrix} \Re\{S_i\} \\ \Im\{S_i\} \end{pmatrix} \right\} \right) = |g_i|^2 \; \mu_1^c(\mathcal{S}_{S_i})$$

5 The Complex Version of the Minimum Support Criterion

Now we are ready to apply the Brunn-Minkowski theorem. We will implicitly assume complex sources whose densities have bounded Lebesgue measurable and non-empty supports. Under these conditions, we can exploit the previously defined isomorphisms, between real and complex sets, to rewrite the Brunn-Minkowski inequality in $\mathbb{R}^2$ (see equation (4)) as an inequality for the measure of the support of complex random variables

$$\left(\mu_1^c(\mathcal{S}_Y)\right)^{\frac{1}{2}} \geq \sum_{i=1}^{n} \left(\mu_1^c(\mathcal{S}_{g_i^* S_i})\right)^{\frac{1}{2}} = \sum_{i=1}^{n} |g_i| \left(\mu_1^c(\mathcal{S}_{S_i})\right)^{\frac{1}{2}} . \tag{10}$$

A theorem, originally formulated by Lusternik and whose proof was later corrected by Henstock and Macbeath [13], establishes the general conditions for the equality to hold in the Brunn-Minkowski theorem.

Theorem 2 (Conditions for equality). *Let $\mathcal{S}_{\boldsymbol{A}}$ and $\mathcal{S}_{\boldsymbol{B}}$ be nonempty bounded Lebesgue m-dimensional measurable sets, let $\mathcal{S}'_{\boldsymbol{A}}$ and $\check{\mathcal{S}}_{\boldsymbol{A}}$ denote, respectively, the complement and the convex closure of $\mathcal{S}_{\boldsymbol{A}}$.*

a) *If $\mu_m(\mathcal{S}_{\boldsymbol{A}}) = 0$ and $0 < \mu_m(\mathcal{S}_{\boldsymbol{B}}) < \infty$, then the necessary and sufficient condition for the equality in Brunn-Minkowski theorem is that $\mathcal{S}_{\boldsymbol{A}}$ should consist of one point only.*

b) *If* $0 < \mu_m(\mathcal{S}_{\boldsymbol{A}})\mu_m(\mathcal{S}_{\boldsymbol{B}}) < \infty$ *the equality in Brunn-Minkowski theorem holds if and only if*

$$\mu_m(\breve{\mathcal{S}}_{\boldsymbol{A}} \cap \mathcal{S}'_{\boldsymbol{A}}) = \mu_m(\breve{\mathcal{S}}_{\boldsymbol{B}} \cap \mathcal{S}'_{\boldsymbol{B}}) = 0,$$

and the convex closures $\breve{\mathcal{S}}_{\boldsymbol{A}}$ *and* $\breve{\mathcal{S}}_{\boldsymbol{B}}$ *are homothetic*[1]*.*

By the application of theorem 2, the equality in (10) is only obtained when one of the following conditions is true:

Case a) The mixture at the output is trivial, i.e.,

$$Y = g_i^* S_i, \quad i \in \{1, \ldots, n\}, \tag{11}$$

which happens when the output is an arbitrary scaled and rotated version of only one the sources.

Case b) When the sources whose contribution to the output does not vanish have support sets which are all convex and homothetic.

The connection between the zero order Rényi's entropy of a random vector in $\mathbb{R}^2$ and the volume of its support set (see [11]) leads us to identify the zero order entropy of a complex random variable with the joint zero order entropy of its real and imaginary parts,

$$h_0^c(Y) = \log \mu_1^c(\mathcal{S}_Y) \equiv h_0(\Re\{Y\}, \Im\{Y\}) \ . \tag{12}$$

Then, we can use equation (10) to obtain a different inequality which relates the zero order entropy of the output with those of the sources and which, at the same time, prevents the equality to hold true for the situations described in the case b). This new inequality is at the heart of the following result.

Theorem 3. *If the measure of the support set of the complex sources if finite and does not vanish for at least* $n-1$ *of them,*

$$\mu_1^c(\mathcal{S}_{S_{\pi_i}}) \neq 0, \quad i = 1, \ldots, n-1, \quad \pi \textit{ perm. of } \{1, \ldots, n\}, \tag{13}$$

the zero order entropy of the normalized output

$$\Psi(\boldsymbol{X}, \mathbf{u}) = h_0^c \left(\frac{\mathbf{u}^H}{\|\mathbf{u}\|_2} \boldsymbol{X} \right) = h_0^c \left(\frac{\boldsymbol{Y}}{\|\mathbf{u}\|_2} \right) \tag{14}$$

is a contrast function for the extraction of one of the sources. The global minimum of this contrast function is obtained for the source (or sources) with smallest scaled measure of support, i.e.,

$$\min_{\mathbf{u}} \Psi(\boldsymbol{X}, \mathbf{u}) = \min_i h_0^c \left(S_i / \|\mathbf{a}_i^-\|_2 \right), \tag{15}$$

where $\mathbf{a}_i^-$ *denotes the* i*th column of* $\mathbf{A}^{-H}$*, the inverse hermitian transpose of the mixing matrix.*

[1] They are equal sets up to translation and dilation.

Due to the lack of space, its proof is omitted. The result tells us that we can extract one of the sources by minimizing the area of the support set of the output.

Note that the theorem does not require the typical ICA assumption of the circularity of the complex sources nor the mutual independence between their real and imaginary parts.

The minimum support contrast function does not work for discrete sources (drawn from alphabets of finite cardinality) because they are of zero measure, a case not covered by the conditions of the theorem. Nevertheless, after replacing the support sets of the original random variables by its convex hull, we return to the conditions of the theorem, obtaining the well-behaved contrast function

$$\Psi(\breve{\boldsymbol{X}}, \mathbf{u}) = \log \mu_1^c(\breve{\mathcal{S}}_{Y/\|\mathbf{u}\|_2}) \equiv h_0^c\left(\frac{\breve{Y}}{\|\mathbf{u}\|_2}\right). \tag{16}$$

Indeed, in all of our experiments, and in similarity with the minimum range contrast for the case of real mixtures [9], this contrast function was apparently free of deceptive minima. Although we still don't know whether this property is true in general, we succeeded in proving the following result.

Theorem 4. *For a mixture of n complex sources with bounded circular convex hull, the minima of the contrast function $\Psi(\breve{\boldsymbol{X}}, \mathbf{u})$ can only be attained at the solutions of the extraction problem, i.e., there are no local deceptive minima.*

6 Simulations

In order to optimize the contrast function we first parametrized a complex unit norm vector $\mathbf{u}$ in terms of $2n-2$ angles (ignoring a common phase term). Let $\mathbf{R}(1, k+1, \alpha_k, \beta_k)$, for $k = 1, \ldots, n-1$, denote a class of planar rotation matrices, then

$$\mathbf{u} = \mathbf{e_1}^T \mathbf{R}(1, n, \alpha_{n-1}, \beta_{n-1}) \cdots \mathbf{R}(1, 2, \alpha_1, \beta_1).$$

Since the extraction solutions are non-differentiable points of the contrast function, we used the downhill simplex method of Nelder and Mead to optimize it in low dimensions [14]. In high dimensions, an improved convergence is obtained when combining the previous optimization technique with numerical gradient and line-search methods. Each function evaluation requires the computation of the planar convex hull of a set of T outputs. The optimal algorithms for this task, have, in the worst case, a computational complexity of $O(T \log V)$ where V is the number of vertices of the convex hull [15].

Consider the sample experiment of 200 observations of a complex mixture of two 16QAM sources (a typical constellation used in communications). The illustration of figure 1 presents the graph of the contrast function $\Psi(\breve{\boldsymbol{X}}, \mathbf{u})$ which periodically tessellates the (α_1, β_1)-plane. The figure shows a contrast function with no local deceptive minima, which is non-differentiable at those points where

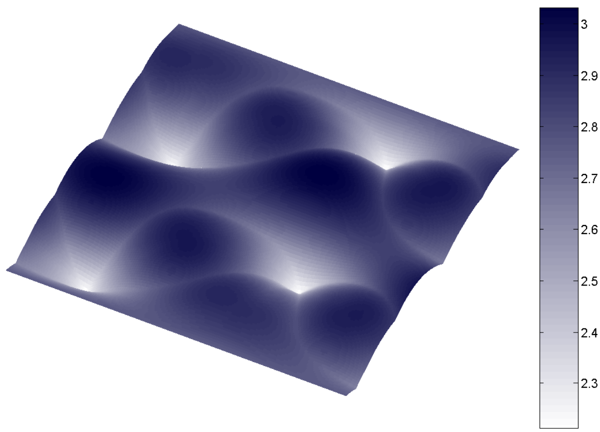

Fig. 1. Graph of the contrast function, with respect the parameters (α_1, β_1), for a mixture of two 16QAM sources. The solutions to the extraction problem are at the minima of the function.

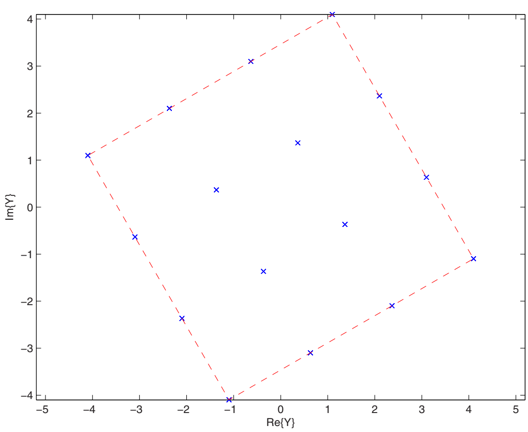

Fig. 2. The 16QAM source recovered by the extraction algorithm and the frontier of the convex hull of its support (dashed line)

the Brunn-Minkowski equality holds true. The illustration of figure 2 presents the 16QAM source extracted by the previously described algorithm and the frontier of the convex hull of its support.

7 Conclusions

We have presented a geometric criterion for the extraction of one independent component from of a linear mixture of complex and mutually independent signals. The criterion favors the extraction of the source signals with minimum scaled support and does not require the mutual independence between their real and imaginary parts. Under certain given conditions, the criterion is proved to be free of defective local minima, although, a general proof is still elusive.

References

1. Donoho, D.: On minimum entropy deconvolution. In: Findley, D.F. (ed.) Applied Time Series Analysis II, pp. 565–608. Academic Press, New York (1981)
2. Blachman, N.M.: The convolution inequality for entropy powers. IEEE Trans. on Information Theory IT-11, 267–271 (1965)
3. Comon, P.: Independent component analysis, a new concept? Signal Processing 3(36), 287–314 (1994)
4. Cruces, S., Cichocki, A., Amari, S-i.: From blind signal extraction to blind instantaneous signal separation: criteria, algorithms and stability. IEEE Trans. on Neural Networks 15(4), 859–873 (2004)
5. Bercher, J.-F., Vignat, C.: A Renyi entropy convolution inequality with application. In: Proc. of EUSIPCO, Toulouse, France (2002)
6. Erdogmus, D., Principe, J.C., Vielva, L.: Blind deconvolution with minimum Renyi's entropy. In: Proc. of EUSIPCO, Toulouse, France, vol. 2, pp. 71–74 (2002)
7. Pham, D.T.: Blind separation of instantaneous mixture of sources based on order statistics. IEEE Trans. on Signal Processing 48(2), 363–375 (2000)
8. Cruces, S., Durán, I.: The minimum support criterion for blind signal extraction. In: proc. of the int. conf. on Independent Component Analysis and Blind Signal Separation, Granada, Spain, pp. 57–64 (2004)
9. Vrins, F., Verleysen, M., Jutten, C.: SWM: A class of convex contrasts for source separation. In: proc. of the Int. Conf. on Acoustics, Speech and Signal Processing, Philadelphia (USA), vol. V, pp. 161–164 (2005)
10. Cruces, S., Sarmiento, A.: Criterion for blind simultaneous extraction of signals with clear boundaries. Electronics Letters 41(21), 1195–1196 (2005)
11. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley series in telecommunications. John Wiley, Chichester (1991)
12. Gardner, R.J.: The Brunn-Minkowski Inequality. Bulletin of the American Mathematical Society 39(3), 355–405 (2002)
13. Henstock, R., Macbeath, A.M.: On the measure of sum-sets. I. The theorems of Brunn, Minkowski, and Lusternik. In: Proceedings of the London Mathematical Society, third series, vol. 3, pp. 182–194 (1953)
14. Nelder, J.A., Mead, R.: A simplex method for function minimization. Computer Journal 7, 308–313 (1965)
15. Chan, T.M.: Optimal output-sensitive convex hull algorithms in two and three dimensions. Discrete and Computational Geometry 16, 361–368 (1996)