BLIND SOURCE SEPARATION OF INDEPENDENT/DEPENDENT SIGNALS USING A MEASURE ON COPULAS

A. Keziou(1,2), *H. Fenniri*(3), *K. Messou*(4) *and E. Moreau*(5,6)

[1]Université de Reims Champagne-Ardenne (URCA),
LMR EA 4535 and ARC-Mathématiques FR 3399 du CNRS, France
[2]LSTA, Université Pierre et Marie Curie, France
[3]Université de Reims Champagne-Ardenne, CReSTIC, France
[4]Université d'Ibn Zohr, Ecole Supérieure de Technologie, Maroc
[5] Université de Toulon, CNRS, LSIS, UMR 7296, 83957 La Garde, France
[6] Aix Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, 13397 Marseille, France

ABSTRACT

We introduce a new BSS approach, based on modified Kullback-Leibler divergence between copula densities, for both independent or dependent source component signals. In the standard case of independent source components, the proposed method improves the mutual information (between probability densities) procedure, and it has the advantage to be naturally generalized to separate mixtures of dependent source components. Simulation results are presented showing the convergence and the efficiency of the proposed algorithms.

***Index Terms*—** Blind source separation, Modified Kullback-Leibler divergence between copulas, Mutual information.

1. INTRODUCTION

Blind source separation (BSS) is an important problem in signal processing which has been addressed in the last three decades. BSS consists of recovering unobserved signals called sources from observed mixtures of them. We consider instantaneous linear mixtures

$$\mathbf{x}(t) := \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \in \mathbb{R}^p, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{p\times p}$ is an unknown non-singular mixing matrix, $\mathbf{s}(t) := (s_1(t), \dots, s_p(t))^\top$ is the unknown vector of source signals to be estimated, as precisely as possible, from $\mathbf{x}(t) := (x_1(t), \dots, x_p(t))^\top$, the vector of observed signals (the number of sources and the number of observations are assumed to be equal). The presence of additive noise $\mathbf{n}(t)$ within the mixing model complicates significantly the BSS problem. Typically, additive noise is either considered as an additional set of sources, or it is reduced by applying some form of preprocessing. The goal is to estimate the vector source signals $\mathbf{s}(t)$ using only the observed signals $\mathbf{x}(t)$. The estimate $\mathbf{y}(t)$ of the source signals $\mathbf{s}(t)$ can be written as

$$\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t), \tag{2}$$

where $\mathbf{B} \in \mathbb{R}^{p\times p}$ is the de-mixing matrix. The question is how to obtain a good de-mixing matrix $\widehat{\mathbf{B}}$ which has to be close to the ideal solution $\mathbf{A}^{-1}$, in using only the observed signals $\mathbf{x}(t)$? It is well known, by Darmois theorem, that if the sources components are mutually independent and at most one component is gaussian, a consistent estimate $\widehat{\mathbf{B}}$ of $\mathbf{A}^{-1}$ (up to scale and permutation indeterminacies) is the one that makes the components of the vector $\mathbf{y}(t)$ independent, see e.g. [1]. The corresponding signal $\widehat{\mathbf{y}}(t) := \widehat{\mathbf{B}}\mathbf{x}(t)$ is the estimate of the source signals $\mathbf{s}(t)$. Under the above hypotheses, many procedures have been proposed in the literature. Some of these procedures use second or higher order statistics, see [2], [3] and the references therein, other consist of optimizing (on the de-mixing matrix space) an estimate of a measure of dependence of the components of the vector $\mathbf{y}(t)$. As measures of dependence used in BSS, we find in the literature the criterion of mutual information [4] [5], the criteria of α, β and Renyi's-divergences [6] [7], and the criteria of ϕ-divergences [8]. The procedures based on minimizing estimates of mutual information are considered as the most efficient, since this criterion can be estimated efficiently, other procedures using other divergences lead to robust method for appropriate choice of divergence criterion [8]. In this paper, we will focus on the criterion of mutual information (called also modified Kullback-Leibler divergence), viewed as measure of difference between copulas, and we will use it to propose a new BSS approach that applies both in the standard case of independent source components, and in the non standard one of dependent source components. In the following, we will show that the mutual information of a random vector $\mathbf{Y} := (Y_1, \dots, Y_p)^\top$ can be written as the modified Kullback-Leibler divergence (KL_m-divergence) between the copula of

Thanks to Egide PHC-Volubilis program for funding.

independence and the copula of the vector. Then, we propose a separation procedure based on minimizing an appropriate estimate of KL_m-divergence between the copula of independence and the copula of the vector. This approach applies in the standard case, and we will show that the proposed criterion can be naturally extended to separate mixture of dependent source components. The proposed approach can be adapted also to separate complex-valued signals. In all the sequel, we assume that at most one source is gaussian, and we will treat separately the case of independent source components, and then the case of dependent source components. In [9], the authors proposed a BSS algorithm (for independent sources) based on minimizing a distance between the parameter of the copula of the estimated source and the value of the parameter corresponding to independence. [10] proposed a different criterion combining the mutual information between probability densities and Shannon entropy of parametric models of copulas.

2. AN INTRODUCTION ON COPULA

Consider a random vector $\mathbf{Y} := (Y_1, \dots, Y_p)^\top \in \mathbb{R}^p$, $p \geq 1$, with joint distribution function (d.f.) $\mathbb{F}_{\mathbf{Y}}(\mathbf{y}) := \mathbb{F}_{\mathbf{Y}}(y_1, \dots, y_p) := \mathbb{P}(Y_1 \leq y_1, \dots, Y_p \leq y_p)$, and continuous marginal d.f.'s $F_{Y_j}(y_j) := \mathbb{P}(Y_j \leq y_j)$, for all $j = 1, \dots, p$. The characterization theorem of Sklar [11] shows that there exists a unique p-variate function $\mathbb{C}_{\mathbf{Y}}(\cdot) : [0,1]^p \mapsto [0,1]$, such that, $\mathbb{F}_{\mathbf{Y}}(\mathbf{y}) = \mathbb{C}_{\mathbf{Y}}(F_{Y_1}(y_1), \dots, F_{Y_p}(y_p))$, for all $\mathbf{y} := (y_1, \dots, y_p)^\top \in \mathbb{R}^p$. The function $\mathbb{C}_{\mathbf{Y}}(\cdot)$ is called a copula and it is in itself a joint d.f. on $[0,1]^p$ with uniform marginals. We have for all $\mathbf{u} := (u_1, \dots, u_p)^\top \in [0,1]^p$, $\mathbb{C}_{\mathbf{Y}}(\mathbf{u}) = \mathbb{P}\left(F_{Y_1}(Y_1) \leq u_1, \dots, F_{Y_p}(Y_p) \leq u_p\right)$. Conversely, for any marginal d.f.'s $F_1(\cdot), \dots, F_p(\cdot)$, and any copula function $\mathbb{C}(\cdot)$, the function $\mathbb{C}(F_1(y_1), \dots, F_p(y_p))$ is a multivariate d.f. on $\mathbb{R}^p$. On the other hand, since the marginal d.f.'s $F_{Y_j}(\cdot)$, $j = 1, \dots, p$, are assumed to be continuous, then the random variables $F_{Y_1}(Y_1), \dots, F_{Y_p}(Y_p)$ are uniformly distributed on the interval $[0,1]$. So, if the components $Y_1, \dots, Y_p$ are statistically independent, then the corresponding copula writes $\mathbb{C}_0(\mathbf{u}) := \prod_{j=1}^p u_j$, $\forall \mathbf{u} \in [0,1]^p$. It is called the copula of independence. Define, when it exists, the density copula (of the random vector $\mathbf{Y}$) $\mathbf{c}_{\mathbf{Y}}(\mathbf{u}) := \frac{\partial^p}{\partial u_1 \cdots \partial u_p} \mathbb{C}_{\mathbf{Y}}(\mathbf{u})$, $\quad \forall \mathbf{u} \in [0,1]^p$. Hence, the density copula of independence $\mathbf{c}_0(\cdot)$ is the function taking the value 1 on $[0,1]^p$ and zero otherwise, namely,

$$\mathbf{c}_0(\mathbf{u}) := \mathbb{1}_{[0,1]^p}(\mathbf{u}), \quad \forall \mathbf{u} \in [0,1]^p. \tag{3}$$

Let $f_{\mathbf{Y}}(\cdot)$, if it exists, be the probability density of the random vector $\mathbf{Y}$, and, respectively, $f_{Y_1}(\cdot), \dots, f_{Y_p}(\cdot)$, the marginal probability densities of the components $Y_1, \dots, Y_p$. Then, a straightforward computation shows that, for all $\mathbf{y} \in \mathbb{R}^p$, we have

$$f_{\mathbf{Y}}(\mathbf{y}) = \prod_{j=1}^p f_{Y_j}(y_j) \mathbf{c}_{\mathbf{Y}}(\mathbf{u}), \tag{4}$$

where $\mathbf{u} := (u_1, \dots, u_p)^\top := (F_{Y_1}(y_1), \dots, F_{Y_p}(y_p))^\top$. In the monographs by [12] and [13], the reader may find detailed ingredients of the modeling theory as well as surveys of the commonly used semi-parametric copulas.

3. MUTUAL INFORMATION AND COPULAS

The mutual information of a random vector

$$\mathbf{Y} := (Y_1, \dots, Y_p)^\top$$

is defined by

$$I(\mathbf{Y}) := \int_{\mathbb{R}^p} -\log \frac{\prod_{j=1}^p f_{Y_j}(y_j)}{f_{\mathbf{Y}}(\mathbf{y})} f_{\mathbf{Y}}(\mathbf{y}) \, dy_1 \dots dy_p. \tag{5}$$

It is called also the modified Kullbak-Leibler divergence (KL_m-divergence) between the product of the marginal densities and the joint density of the vector. Note also that $I(\mathbf{Y}) =: KL_m(\prod_{j=1}^p f_{Y_j}, f_{\mathbf{Y}})$ is nonnegative and takes the value zero if and only if the components of the vector are independent. An equivalent formula of (5) is

$$I(\mathbf{Y}) := E\left(-\log \frac{\prod_{j=1}^p f_{Y_j}(Y_j)}{f_{\mathbf{Y}}(\mathbf{Y})}\right), \tag{6}$$

where $E(\cdot)$ is the mathematical expectation. Using the relation (4), and applying the change variable formula in multiple integrals, we can show that the mutual information $I(\mathbf{Y})$ can be written as

$$I(\mathbf{Y}) = \int_{[0,1]^p} -\log\left(\frac{1}{\mathbf{c}_{\mathbf{Y}}(\mathbf{u})}\right) \mathbf{c}_{\mathbf{Y}}(\mathbf{u}) \, d\mathbf{u} =: KL_m(\mathbf{c_0}, \mathbf{c}_{\mathbf{Y}})$$
$$= E\left(\log \mathbf{c}_{\mathbf{Y}}(F_{Y_1}(Y_1), \dots, F_{Y_p}(Y_p))\right) =: -H(\mathbf{c}_{\mathbf{Y}}),$$

where $H(\mathbf{c}_{\mathbf{Y}}) := \int_{[0,1]^p} -\log\left(\mathbf{c}_{\mathbf{Y}}(\mathbf{u})\right) \mathbf{c}_{\mathbf{Y}}(\mathbf{u}) \, d\mathbf{u}$ is the Shannon entropy of the copula density $\mathbf{c}_{\mathbf{Y}}(\cdot)$. The relation above means that the mutual information of the random vector $\mathbf{Y}$ can be seen as the KL_m-divergence between the independent density copula $\mathbf{c}_0(\cdot)$, see (3), and the density copula $\mathbf{c}_{\mathbf{Y}}(\cdot)$ of the random vector $\mathbf{Y}$. We summarize the above results in the following Proposition.

Proposition 1. *Let $\mathbf{Y} \in \mathbb{R}^p$ be any random vector with continuous marginal distribution functions. Then, the mutual information $I(\mathbf{Y})$ of $\mathbf{Y}$ can be written as the KL_m-divergence between the density copula $\mathbf{c}_0$ of independence and the copula density of the vector $\mathbf{Y}$:*

$$I(\mathbf{Y}) = \int_{[0,1]^p} -\log\left(\frac{1}{\mathbf{c}_{\mathbf{Y}}(\mathbf{u})}\right) \mathbf{c}_{\mathbf{Y}}(\mathbf{u}) \, d\mathbf{u} =: KL_m(\mathbf{c_0}, \mathbf{c}_{\mathbf{Y}})$$
$$= E\left(\log \mathbf{c}_{\mathbf{Y}}(F_{Y_1}(Y_1), \dots, F_{Y_p}(Y_p))\right). \tag{7}$$

Moreover, $KL_m(\mathbf{c_0}, \mathbf{c}_{\mathbf{Y}})$ is non negative and takes the value zero if and only if $\mathbf{c}_{\mathbf{Y}}(\mathbf{u}) = \mathbf{c_0}(\mathbf{u}) =: \mathbb{1}_{[0,1]^p}(\mathbf{u})$, $\forall \mathbf{u} \in [0,1]^p$, namely, iff the components of the vector $\mathbf{Y}$ are independent.

4. A SEPARATION PROCEDURE FOR INDEPENDENT SOURCES THROUGH COPULAS

In this section, we describe our approach based on minimizing a non parametric estimate of the KL_m-divergence $KL_m(\mathbf{c}_0, \mathbf{c_Y})$, assuming that the source components are independent. Denote by $\mathbf{S} := (S_1, \ldots, S_p)^\top$ the random source vector, $\mathbf{X} := \mathbf{AS}$ the observed random vector and $\mathbf{Y} := \mathbf{BX}$ the estimated random source vector. The discrete (noise free) version of the mixture model (1) writes $\mathbf{x}(n) := \mathbf{As}(n)$, $n = 1, \ldots, N$. The source signals $\mathbf{s}(n)$, $n = 1, \ldots, N$, will be considered as N realizations of the random source vector $\mathbf{S}$, and then $\mathbf{x}(n), \mathbf{y}(n) := \mathbf{Bx}(n)$, $n = 1, \ldots, N$, are, respectively, N realizations of the random vectors $\mathbf{X}$ and $\mathbf{Y} := \mathbf{BX}$. In view of Proposition 1, the function $\mathbf{B} \mapsto KL_m(\mathbf{c}_0, \mathbf{c_Y})$ is non negative and achieves its minimum value zero if and only if $\mathbf{B} = \mathbf{A}^{-1}$ (up to scale and permutation indeterminacies). In other words, we have $\mathbf{A}^{-1} = \arg\inf_{\mathbf{B}} KL_m(\mathbf{c}_0, \mathbf{c_Y})$. Hence, to achieve separation, the idea is to minimize on $\mathbf{B}$ some statistical estimate $\widehat{KL_m}(\mathbf{c}_0, \mathbf{c_Y})$, of $KL_m(\mathbf{c}_0, \mathbf{c_Y})$, constructed from the data $\mathbf{y}(1), \ldots, \mathbf{y}(n)$. The separation matrix is then estimated by $\widehat{\mathbf{B}} := \arg\inf_{\mathbf{B}} \widehat{KL_m}(\mathbf{c}_0, \mathbf{c_Y})$, leading to the estimated source signals $\widehat{\mathbf{y}}(n) = \widehat{\mathbf{B}}\mathbf{x}(n)$, $n = 1, \ldots, N$. Based on the equality (7), we propose to estimate the KL_m-divergence $KL_m(\mathbf{c}_0, \mathbf{c_Y})$ by a plug-in type procedure. We obtain then the estimate $\widehat{KL_m}(\mathbf{c}_0, \mathbf{c_Y}) := \frac{1}{N}\sum_{n=1}^{N} \log \widehat{c}_{\mathbf{Y}}(\widehat{F}_{Y_1}(y_1(n)), \ldots, \widehat{F}_{Y_p}(y_p(n)))$, where,

$$\widehat{c}_{\mathbf{Y}}(\mathbf{u}) := \frac{1}{Nh_1 \cdots h_p} \sum_{m=1}^{N} \prod_{j=1}^{p} k\left(\frac{\widehat{F}_{Y_j}(y_j(m)) - u_j}{h_j}\right), \tag{8}$$

for all $\mathbf{u} \in [0,1]^p$, is the kernel estimate of the copula density $\mathbf{c_Y}(\cdot)$, and $\widehat{F}_{Y_j}(x)$, $j = 1, \ldots, p$, is the estimate of the marginal distribution function $F_{Y_j}(x)$ of the random variable Y_j, at any real value x, defined by

$$\widehat{F}_{Y_j}(x) := \frac{1}{N}\sum_{n=1}^{N} K\left(\frac{y_j(n) - x}{h_j}\right), \tag{9}$$

where $K(\cdot)$ is the primitive of a kernel $k(\cdot)$, a symmetric centered probability density. In order to estimate the marginal distribution functions, we will take as kernel a standard gaussian density. A more appropriate choice of the kernel $k(\cdot)$, for estimating the copula density, can be done according to [18], which copes with the boundry effect. The bandwidth parameters $h_1, \ldots, h_p$ in (8) and in (9) will be chosen according to the Silverman's rule of thumb, see [14].

4.1. The direct approach

In order to compute the estimate of the de-mixing matrix $\widehat{\mathbf{B}}$, we can use a gradient descent algorithm taking as initial matrix $B_0 = I_p$, the $p \times p$ identity matrix. The gradient in ($\mathbf{B}$) of $\widehat{KL_m}(\mathbf{c}_0, \mathbf{c_Y})$ can be computed explicitly from the definitions of the estimates. We propose then the following algorithm.

Data : the observed signals $\mathbf{x}(n), n = 1, \ldots, N$
Result : the estimated sources $\widehat{\mathbf{y}}(n), n = 1, \ldots, N$
Initialization : $\mathbf{y}_0(n) = \mathbf{B}_0\mathbf{x}(n)$, $\mathbf{B}_0 = I_p$. Given $\varepsilon > 0$ and $\mu > 0$ suitably choosen
Do
• Update $\mathbf{B}$ and $\mathbf{y}$:

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \mu \left.\frac{d}{d\mathbf{B}}\widehat{KL_m}(\mathbf{c}_0, \mathbf{c_Y})\right|_{\mathbf{B}=\mathbf{B}_k}$$

$$\mathbf{y}_{k+1}(n) = \mathbf{B}_{k+1}\mathbf{x}(n), n = 1, \ldots, N$$

• **Until** $\|\mathbf{B}_{k+1} - \mathbf{B}_k\| < \varepsilon$
$\widehat{\mathbf{y}}(n) = \mathbf{y}_{k+1}(n), n = 1, \ldots, N$

Algorithm 1: A copula based BSS algorithm for independent source components.

4.2. The two steps approach

This method proceeds in two steps: the first one consists of spatial whitening and the second one consists to apply a series of Givens rotations, minimizing the estimate of the KL_m-divergence. The whitened mixture vector $\mathbf{z}$ can be written as $\mathbf{z}(n) = \mathbf{Wx}(n)$, $n = 1, \ldots, N$, where $\mathbf{W}$ is the whitening $p \times p$-matrix. Let $\mathbf{U}$ be a unitary $p \times p$- matrix, namely, the matrix $\mathbf{U}$ satisfying $\mathbf{UU}^\top = I_p$. It can be written as $\mathbf{U}(\theta) := \prod_{1 \le i < k \le p} G(i, k, \theta_m)$, where $G(i, k, \theta_m)$ is the $p \times p$-matrix with entries

$$G(i,k,\theta_m)_{j,l} := \begin{cases} \cos(\theta_m) & \text{if} \quad j = i, l = i \\ & \text{or} \quad j = k, l = k; \\ \sin(\theta_m) & \text{if} \quad j = i, l = k; \\ -\sin(\theta_m) & \text{if} \quad j = k, l = i; \\ 1 & \text{if} \quad j = l; \\ 0 & \text{else,} \end{cases}$$

for all $1 \le j, l \le p$, and $\theta_m \in]-\pi/2, \pi/2[$, $m = 1, \ldots, p(p-1)/2$, are the rotation angles (the components of the vector θ). The estimated source signals take then the form $\mathbf{y}(n) = \mathbf{U}(\theta)\mathbf{z}(n)$, $n = 1, \ldots, N$, and the separating matrix is $\mathbf{B} = \mathbf{U}(\theta)\mathbf{W}$. The estimate $\widehat{KL_m}(\mathbf{c_0}, \mathbf{c_Y})$, of $KL_m(\mathbf{c_0}, \mathbf{c_Y})$, can be seen as a function of the parameter vector θ. Let $\widehat{\theta} := \arg\inf_\theta \widehat{KL_m}(\mathbf{c_0}, \mathbf{c_Y})$, which can be computed by a descent gradient (in θ) algorithm. The de-mixing matrix is then estimated by $\widehat{\mathbf{B}} := \mathbf{U}(\widehat{\theta})\mathbf{W}$, leading to the estimated source signals

$$\widehat{\mathbf{y}}(n) = \widehat{\mathbf{B}}\mathbf{x}(n) = \mathbf{U}(\widehat{\theta})\mathbf{Wx}(n), \; n = 1, \ldots, N.$$

We summarize the above methodology in the following algorithm.

Data : the observed signals $\mathbf{x}(n), n = 1, \ldots, N$
Result : the estimated sources $\widehat{\mathbf{y}}(n), n = 1, \ldots, N$
Whitening and Initialization : $\mathbf{z}(n) := W\mathbf{x}(n)$, $\mathbf{y}_0(n) = U(\theta_0)\mathbf{z}(n)$. Given $\varepsilon > 0$ and $\mu > 0$
Do
• Update θ and $\mathbf{y}$:

$$\theta_{k+1} = \theta_k - \mu \left. \frac{d}{d\theta} \widehat{KL_m}(\mathbf{c}_0, \mathbf{c}_{\mathbf{Y}}) \right|_{\theta = \theta_k}$$

$$\mathbf{y}_{k+1}(n) = U(\theta_{k+1})\mathbf{z}(n), n = 1, \ldots, N$$

• **Until** $\|\theta_{k+1} - \theta_k\| < \varepsilon$
$\widehat{\mathbf{y}}(n) = \mathbf{y}_{k+1}(n), n = 1, \ldots, N$

Algorithm 2: A two steps BSS algorithm for independent source components.

Data : the observed signals $\mathbf{x}(n), n = 1, \ldots, N$
Result : the estimated sources $\widehat{\mathbf{y}}(n), n = 1, \ldots, N$
Initialization : $\mathbf{y}_0(n) = \mathbf{B}_0\mathbf{x}(n)$, $\mathbf{B}_0 = I_p$. Given $\varepsilon > 0$ and $\mu > 0$ suitably choosen
Do
• Update $\mathbf{B}$ and $\mathbf{y}$:

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \mu \left. \frac{d}{d\mathbf{B}} \widehat{KL_m}(\mathbf{c}_{\widehat{\alpha}}, \mathbf{c}_{\mathbf{Y}}) \right|_{\mathbf{B} = \mathbf{B}_k}$$

$$\mathbf{y}_{k+1}(n) = \mathbf{B}_{k+1}\mathbf{x}(n), n = 1, \ldots, N$$

• **Until** $\|\mathbf{B}_{k+1} - \mathbf{B}_k\| < \varepsilon$
$\widehat{\mathbf{y}}(n) = \mathbf{y}_{k+1}(n), n = 1, \ldots, N$

Algorithm 3: A copula based BSS algorithm for mixtures of dependent component sources.

5. A SOLUTION TO THE BSS PROBLEM FOR MIXTURES OF DEPENDENT SOURCE COMPONENTS

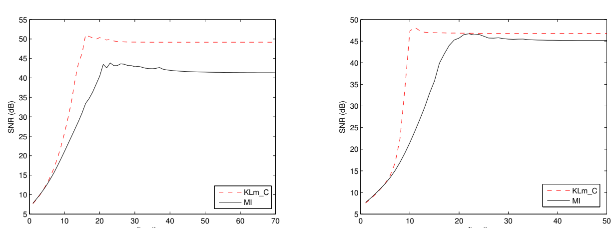
Assume that we dispose of some prior information about the density copula of the random source vector $\mathbf{S}$. Note that this is possible for many practical problems, it can be done, from realization of $\mathbf{S}$, by a model selection procedure in parametric copula density models $\left\{\mathbf{c}_\alpha(\cdot);\, \alpha \in \Theta \subset \mathbb{R}^d\right\}$ typically indexed by a multivariate parameter α, see [15]. The parameter α can be estimated using maximum semi-parametric likelihood, see [16] and [17]. Denote by $\widehat{\alpha}$ the obtained value of α and $\mathbf{c}_{\widehat{\alpha}}(\cdot)$ the copula density modeling the dependency structure of the source components. Obviously, since the source components are assumed to be dependent, $\mathbf{c}_{\widehat{\alpha}}(\cdot)$ is different from the density copula of independence $\mathbf{c}_0(\cdot)$. Hence, we naturally replace in the above criterion $KL_m(\mathbf{c}_0, \mathbf{c}_{\mathbf{Y}})$, $\mathbf{c}_0$ by $\mathbf{c}_{\widehat{\alpha}}$. Moreover, we can show that $\mathbf{B} \mapsto KL_m(\mathbf{c}_{\widehat{\alpha}}, \mathbf{c}_{\mathbf{Y}})$ is non negative and achieves its minimum value zero if $\mathbf{B} = \mathbf{A}^{-1}$ (up to scale and permutation indeterminacies), i.e., $\mathbf{A}^{-1} = \arg\inf_{\mathbf{B}} KL_m(\mathbf{c}_{\widehat{\alpha}}, \mathbf{c}_{\mathbf{Y}})$. Note also that the criterion $KL_m(\mathbf{c}_{\widehat{\alpha}}, \mathbf{c}_{\mathbf{Y}})$ can be written as $KL_m(\mathbf{c}_{\widehat{\alpha}}, \mathbf{c}_{\mathbf{Y}}) = E\left(\log\left(\frac{\mathbf{c}_{\mathbf{Y}}(F_{Y_1}(Y_1),\ldots,F_{Y_p}(Y_p))}{\mathbf{c}_{\widehat{\alpha}}(F_{Y_1}(Y_1),\ldots,F_{Y_p}(Y_p))}\right)\right)$. So as before, we propose to estimate the de-mixing matrix by $\widehat{\mathbf{B}} := \arg\inf_{\mathbf{B}} \widehat{KL_m}(\mathbf{c}_{\widehat{\alpha}}, \mathbf{c}_{\mathbf{Y}})$, where $\widehat{KL_m}(\mathbf{c}_{\widehat{\alpha}}, \mathbf{c}_{\mathbf{Y}}) := \frac{1}{N}\sum_{n=1}^{N} \log\left(\frac{\widehat{\mathbf{c}}_{\mathbf{Y}}(\widehat{F}_{Y_1}(y_1(n)),\ldots,\widehat{F}_{Y_p}(y_p(n)))}{\mathbf{c}_{\widehat{\alpha}}(\widehat{F}_{Y_1}(y_1(n)),\ldots,\widehat{F}_{Y_p}(y_p(n)))}\right)$. The estimates of copula density and the marginal distribution functions are defined as before. The solution $\widehat{\mathbf{B}}$ can be computed by a descent gradient algorithm. The estimated source signals are $\widehat{\mathbf{y}}(n) = \widehat{\mathbf{B}}\mathbf{x}(n)$, $n = 1, \ldots, N$; see Algorithm 3.

6. SIMULATION RESULTS

In this section, we present simulation results for the proposed method. We dealt with instantaneous mixtures of three kinds of sample sources, uniform i.i.d with independent components (Fig. 1.a and Fig. 2), i.i.d sources with independent components drawn from the 4-ASK (Amplitude Shift Keying) alphabet at which was added a centered Gaussian noise with variance equal to 0.25 (Fig. 1.b), and i.i.d vector sources with dependent components generated from Fairlie-Gumbel-Morgenstern (FGM)-copula with $\widehat{\alpha} = 0.8$ (Fig. 3). The accuracy of source estimation is evaluated through the SNR, defined by

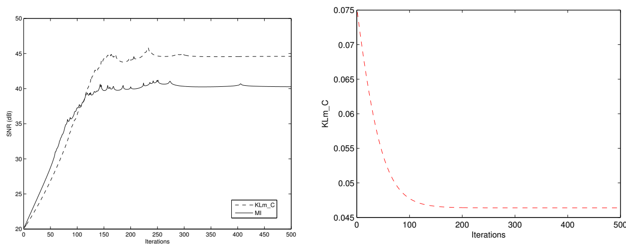
$$SNR_i := 10 \log_{10} \frac{\sum_{n=1}^{N} s_i(n)^2}{\sum_{n=1}^{N} (\widehat{y}_i(n) - s_i(n))^2}, i = 1, 2.$$

In the standard case of independent component sources, our algorithms (Algorithms 1 and 2) are compared with the MI method described in [4], see Fig. 1 and Fig. 2. In Fig. 1.b the used mixing matrix is $A := [1\ 0.85; 0.65\ 1]$, and in all the other cases $A := [1\ 0.5; 0.5\ 1]$. The number of samples is $N = 2000$, and all simulations are repeated 50 times. The gradient descent parameter is taken $\mu = 0.1$ in all cases. We observe from Fig. 1 and Fig. 2 that the proposed method gives good results for the standard case of independent component sources. Moreover, we see from Fig. 3 that our proposed Algorithm 3 is able to separate, with good performance, mixtures of dependent source components.



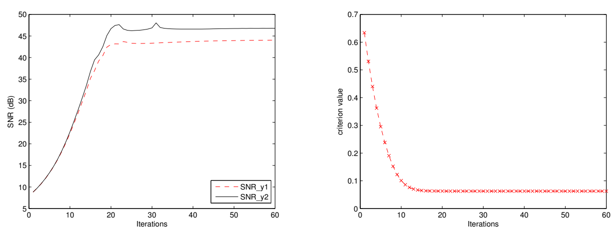
(a) Uniform i.i.d sources (b) ASK sources

Fig. 1. Average output SNRs versus iteration number



(a) SNRs vs iterations (b) Criterion value vs iterations

Fig. 2. The two steps copula based BSS for Uniform i.i.d with independent source components



(a) SNRs vs iterations (b) Criterion value vs iterations

Fig. 3. BSS for dependent source component signals

7. REFERENCES

[1] P. Comon, "Independent component analysis, a new concept?," *Signal Processing, IEEE*, vol. 36, no. 3, pp. 287–314, 1994.

[2] J.C. Pesquet and E. Moreau, "Cumulant-based independence measures for linear mixtures," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 1947–1956, 2001.

[3] M. Castella, S. Rhioui, E. Moreau and J.C. Pesquet, "Quadratic higher order criteria for iterative blind separation of a MIMO convolutive mixture of sources," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 218–232, 2007.

[4] D.T. Pham, "Mutual information approach to blind separation of stationary sources," *IEEE Trans. Information Theory*, vol. 48, no. 7, 2002.

[5] M. El Rhabi, G. Gelle, H. Fenniri and G. Delaunay, "A penalized mutual information criterion for blind separation of convolutive mixtures," *Signal Processing*, vol. 84, no. 10, 2004.

[6] A. Cichocki and S.I. Amari, "Families of alpha- beta- and gamma-divergences: flexible and robust measures of similarities," *Entropy*, vol. 12, no. 6, pp. 1532–1568, 2010.

[7] F. Vrins, D.T. Pham and M. Verleysen, *Is the general form of Renyi's entropy a contrast for source separation?*, vol. 4666/2007 of *Lecture Notes in Computer Science*, Springer, Berlin, 2007.

[8] M. El Rhabi, H. Fenniri, A. Keziou and E. Moreau, "A robust algorithm for convolutive blind source separation in presence of noise," *Signal Processing*, vol. 93, pp. 818–827, 2013.

[9] R.B. Chen, M. Guo, W. Hardle and S.F. Huang, "Independent component analysis via copula techniques," *SFB 649 Discussion paper 2008-004*, 2008.

[10] J. MA and J. Sun, "Copula component analysis," *ICA 2007*, pp. 73–80, 2007.

[11] A. Sklar, "Fonctions de répartition à n dimensions et leurs marges," *Publ. Inst. Statist. Univ. Paris*, vol. 8, pp. 229–231, 1959.

[12] R.B. Nelsen, *An introduction to copulas*, Springer Series in Statistics. Springer, New York, second edition, 2006.

[13] H. Joe, *Multivariate models and dependence concepts*, vol. 73 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, London, 1997.

[14] B.W. Silverman, *Density estimation for statistics and data analysis*, Monographs on Statistics and Applied Probability. Chapman & Hall, London, 1986.

[15] X. Chen and Y. Fan, "Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection," *Canad. J. Statist.*, vol. 33, no. 3, pp. 389–414, 2005.

[16] H. Tsukahara, "Semiparametric estimation in copula models," *Canad. J. Statist.*, vol. 33, no. 3, pp. 357–375, 2005.

[17] S. Bouzebda and A. Keziou, "New estimates and tests of independence in some copula models via ϕ-divergences," *Kybernetika*, vol. 46, no. 1, pp. 178–201, 2010.

[18] M. Omelka, I. Gijbels and N. Veraverbeke "Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing," *The Annals of Statistics*, vol. 37, no. 5B, pp. 3023–3058, 2009.