**RAIRO Operations Research**
Will be set by the publisher

# BLIND SOURCE SEPARATION USING HELLINGER DIVERGENCE AND COPULAS

AMAL OURDOU[1], ABDELGHANI GHAZDALI[1] AND ABDELMOUTALIB METRANE[1]

**Abstract**. Whenever there is a mixture of signals of any type, e.g. sounds, images or any other form of source signals, Blind Source Separation (BSS) is the method utilized to separate these signals from the observations. The separation is done without any prior knowledge about the mixing process nor the source signals. In literature multiple algorithms have been deployed for this particular problem, however most of them depends on Independent Component Analysis (ICA) and its variations assuming the statistical independence of the sources. In this paper, we develop a new algorithm improving the separation quality for both independent and dependent sources. Our algorithm used copulas to accurately model the dependency structure and the Hellinger divergence as a distance measure since it can convergence faster and it is robust against noisy source signals. Many simulations were conducted for various samples of sources to illustrate the superiority of our approach compared to other methods.

**Keywords:** Blind Source Separation, Hellinger divergence, Copulas, Dependent sources, Noise-contaminated sources.

## INTRODUCTION

Blind Source Separation is the technique used to extract sources, from observations of their mixtures without the knowledge of the original signals or the mixing process. Signal processing and Machine learning communities have widely explored the challenges in BSS during the last three decades. It was first used for the cocktail party problem, where the aim was to separate the sound signals of each person's speech. Then exploited in other scientific fields such as signal processing, image processing, medical signal processing,

...

[1] LIPIM, ENSA Khouribga, Sultan Moulay Slimane University, Khouribga, Morocco

artificial neural networks, statistics, and information theory, speech recognition systems, telecommunications.
BSS is an ill posed problem, therefore, in literature, various assumptions on the sources has been made to enable the separation of the observed mixtures. For the linear and static mixing environment, Independent Component Analysis (ICA) [1] is used. It considers the sources to be mutually independent and non-Gaussian. Under these assumptions, the source signals can be estimated by optimizing a cost function. Numerous variations of ICA were introduced to literature citing for examples, maximizing likelihood [2], minimizing the mutual information [3,4], minimizing the criteria of $\phi$-divergences [5,6], the second or higher order statistics [7,8], etc. A good overview on the problem can be found in [9].
In [6, 10, 11] a new BSS algorithm was proposed to overcome the drawbacks of ICA techniques. This algorithm uses Copula to accurately model the dependency structure between the source components, hence omitting the mutual independence assumption. In this paper we make use of this copula and focus on the Hellinger divergence between the copula densities as our cost function to minimize, due to its efficiency and robustness in improving the results even for noisy data [12,13], moreover one of its main characteristics is its rapid convergence compared to any other divergences.
This paper is organized as follows, section "Blind Source Separation" gives an overview on the BSS principle and model, then a review on copula in section "What are copulas?". After that in section "Hellinger divergence and copula" we present our cost function as the Hellinger divergence between the copula densities. We then introduce our new approach detailing separately the independent and the dependent cases, in section "The proposed approach". Then the comparison between our approach and various other methods is made, illustrating that the superiority of our approach in section "Simulation results". Finally, we conclude the paper and give some further research directions.

## 1. Blind Source Separation

The linear BSS problem states that the $n$ unknown source components $\boldsymbol{s}(t) \in \mathbb{R}^n$, are blindly mixed together through a matrix $\boldsymbol{M}$ containing the mixing coefficients. In vector notations

$$\boldsymbol{x}(t) := \boldsymbol{M}\boldsymbol{s}(t) + \boldsymbol{b}(t),\ t \in \mathbb{R}, \tag{1}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^p$ are the observations, and $\boldsymbol{b}(t)$ is an additive noise. In our work we consider the determine case where the number of sources is the same as the number of the observations e.g. $p = n$ and that the additive noise is omitted using a pre-processing technique [5]. The new BSS model is as follows:

$$\boldsymbol{x}(t) := \boldsymbol{M}\ \boldsymbol{s}(t),\ \forall t \in \mathbb{R}, \tag{2}$$

having only the observed signals $\boldsymbol{x}(t)$, and no prior knowledge about the mixing process, the BSS solution searches for the optimum $p \times p$ un-mixing matrix $B$, which gives the recovered sources

$$\boldsymbol{y}(t) := \boldsymbol{B}\,\boldsymbol{x}(t),\ \forall t \in \mathbb{R}. \tag{3}$$

Where $\boldsymbol{y}(t)$ is the estimated sources that would be similar to the wanted sources $\boldsymbol{s}(t)$ if the un-mixing matrix $B$ is as close as possible to $M^{-1}$.

## 2. What are copulas?

Copulas has become very popular recently as a method to model the dependency structure of random variables. We can define copula as the function that helps us to connect univariate marginal distributions to a joint multivariate distribution function with a specific form of dependency. The Sklar's theorem [14], which is the fundamental theorem for copulas, affirm the copula function's existence, which it is of the form:

$$F(\mathbf{Z}) = \mathbb{C}_{\mathbf{Z}}(F_1(Z_1), \ldots, F_p(Z_p)), \forall \mathbf{Z} := (Z_1, \ldots, Z_p)^\top \in \mathbb{R}^p. \tag{4}$$

Where $F$ is an p-dimensional distribution function with marginals $F_1, ..., F_p$. $\mathbb{C}_{\mathbf{Z}}(\cdot)$ is the copula function which is also a joint distribution function on $[0,1]^p$ in itself, with uniform margins. We have the following:

$$\forall \mathbf{u} := (u_1, \ldots, u_p)^\top \in [0,1]^p,\ \mathbb{C}_{\mathbf{Z}}(\mathbf{u}) = \mathbb{P}\left(F_1(Z_1) \le u_1, \ldots, F_p(Z_p) \le u_p\right).$$

If $F_1, ..., F_p$ are all continuous, then $\mathbb{C}_{\mathbf{Z}}(\cdot)$ is unique. In the opposite direction, consider a copula, $\mathbb{C}_{\mathbf{Z}}(\cdot)$, and univariate distribution functions, $F_1, ..., F_n$. Then $F$ as defined in (4) is a joint multivariate distribution function with marginals $F_1, ..., F_n$.
For the case, where the components of a random vector variable, $\mathbf{Z} := (Z_1, \ldots, Z_p)^\top \in \mathbb{R}^p$ are statistically independent, we have the copula of independence denoted $\mathbb{C}_{\Pi}(.)$ of the form:

$$\mathbb{C}_{\Pi}(\mathbf{u}) := \mathbb{C}_{\mathbf{Z}}(\mathbf{u}) = \prod_{j=1}^{p} u_j,\ \forall \mathbf{u} \in [0,1]^p,$$

If the copula has a density, then it is obtained in the following manner as

$$c_{\mathbf{Z}}(\mathbf{u}) := \frac{\partial^p \mathbb{C}_{\mathbf{Z}}(\mathbf{u})}{\partial u_1 \cdots \partial u_p},\ \forall \mathbf{u} \in [0,1]^p.$$

Using the last formula we can obtain the density of the copula of independence as follows:

$$c_{\Pi}(\mathbf{u}) := 1_{[0,1]^p}(\mathbf{u}), \quad \forall \mathbf{u} \in [0,1]^p. \tag{5}$$

For the the random vector $\mathbf{Z} := (Z_1, \ldots, Z_p)^\top$, let $f_{\mathbf{Z}}(\cdot)$ be its probability density if it exists, and $f_1(\cdot), \ldots, f_p(\cdot) \in \mathbb{R}^p$ the marginal probability densities of $Z_1, \ldots, Z_p$ respectively. We can obtain the following relation after some uncomplicated computations

$$f_{\mathbf{Z}}(\mathbf{Z}) = \left( \prod_{j=1}^{p} f_j(Z_j) \right) \; c_{\mathbf{Z}} \left( F_1(Z_1), \ldots, F_p(Z_p) \right). \tag{6}$$

Numerous models for copulas have been proposed in the literature. Semi-parametric copula models class are the most popular for modeling and estimating the structure of dependency. For this class the parametric copulas $\mathbb{C}(\cdot, \boldsymbol{\theta})$ is indexed by a parameter $\theta \in \Theta \subset \mathbb{R}^d$, with a non-parametric margins.

In Table 1 we recall a description of three models of copula : Clayton [15], Ali-Mikhail-Haq (AMH) [16] and Frank [17], which were used in Section "Simulation results" in our simulation study. We provide the respective parameter space $\Theta$ for each model and the parameter $\theta$ corresponding to the independence hypothesis of margins denoted $\theta_0$, in other words

$$\begin{aligned} \mathbb{C}\left(u_1, \ldots, u_p; \boldsymbol{\theta}_0\right) &= \mathbb{C}_{\Pi}\left(u_1, \ldots, u_p\right) \\ &:= \prod_{i=1}^{p} u_i, \; \forall \left(u_1, \ldots, u_p\right)^{\top} \in [0,1]^p. \end{aligned} \tag{7}$$

For a better understanding on the widely used semi-parametric copulas, one may refer to [18, 19].

In the following lines, we outline briefly one of the copula model selection procedures and the method of estimating the parameter $\theta$ from the data. For a random vector $\boldsymbol{Z} \in \mathbb{R}^p$, let's assume that a training sample of $Z$ is available, that is, we dispose of i.i.d. realizations $\boldsymbol{z}(1), \ldots, \boldsymbol{z}(N)$ of $\boldsymbol{Z}$.

The objective is to select the "best" copula model from the data, among a list of candidate models, that models the dependence structure of the components $\boldsymbol{Z}$, and to estimate the parameter $\theta$ of the model selected. Let $\left\{C_1(\cdot, \theta_1); \theta_1 \in \Theta_1 \subset \mathbb{R}^{d_1}\right\}, \ldots, \left\{C_K(\cdot, \theta_K); \theta_K \in \Theta_K \subset \mathbb{R}^{d_K}\right\}$ be a list of candidate copula models. The selection of

TABLE 1. Examples of semiparametric copulas

| Family | $\mathbb{C}(u_1, u_2, \ldots, u_p, \theta)$ | $\Theta$ | $\theta_0$ |
|---|---|---|---|
| **AMH** | $\dfrac{\prod_{i=1}^{p} u_i}{1 - \theta \left( \prod_{i=1}^{p} (1 - u_i) \right)}$ | $[-1, 1]$ | 0 |
| **Clayton** | $\max \left[ \left( \sum_{i=1}^{p} u_i^{-\theta} - p + 1 \right), 0 \right]^{-\frac{1}{\theta}}$ | $[-1, +\infty[ \setminus \{0\}$ | 0 |
| **Frank** | $-\dfrac{1}{\theta} \ln \left( 1 + \dfrac{\prod_{i=1}^{p} \left(e^{-\theta u_i} - 1\right)}{\left(e^{-\theta} - 1\right)^{p-1}} \right)$ | $\mathbb{R} \setminus \{0\}$ | 0 |

models can be done using the Bayesian information criterion (BIC) [20], resulting from the semi-parametric log-likelihood, see e.g. [21] and [22]. Denote by $c_k(\cdot, \theta_k)$ the density of the copula $\mathbb{C}_k(\cdot, \theta_k)$, for all $k$.
The BIC, of a given model $k$, is defined by

$$BIC(k) = -2 \sup_{\theta_k \in \Theta_k} \sum_{n=1}^{N} \log c_k \left( \widehat{F}_{Z_1}(\boldsymbol{z}_1(n)), \ldots, \widehat{F}_{Z_p}(\boldsymbol{z}_p(n)), \theta_k \right) \\ + d_k / \log(N),$$

The ideal model is the one which minimizes the BIC values, namely, the density copula model $\left\{ c_{k^*}(\cdot, \theta_{k^*});\ \theta_{k^*} \in \Theta_{k^*} \subset \mathbb{R}^{d_{k^*}} \right\}$ where

$$k^* = \arg \min_{k \in \{1, \ldots, K\}} BIC(k).$$

Denote, simply, $\left\{ c(\cdot, \theta); \theta \in \Theta \subset \mathbb{R}^d \right\}$ a selected model according to the above procedure. The parameter $\theta$ of the copula model in question can be estimated by maximizing the semi-parametric log-likelihood

$$\widehat{\theta} = \arg \sup_{\theta \in \Theta} \sum_{n=1}^{N} \log\ c \left( \widehat{F}_{Z_1}(\boldsymbol{z}_1(n)), \ldots, \widehat{F}_{Z_p}(\boldsymbol{z}_p(n)), \theta \right).$$

## 3. Hellinger divergence and copula

One of the important issues in many applications of probability theory is finding an appropriate measure of distance between two probability distributions. A number of divergence measures for this purpose have been studied. In this paper, we singled out Hellinger divergence [23] as the measure of instantaneous information because it improves the maximum likelihood in terms of efficiency-robustness for noisy data. It also converges faster than other divergences, see, e.g., [24, 25].
The Hellinger distance denoted $H$ between two probability density functions is defined through

$$H(q,p) := \int_{\mathbb{R}^M} 2 \left( \sqrt{\frac{q(\boldsymbol{t})}{p(\boldsymbol{t})}} - 1 \right)^2 p(\boldsymbol{t}) dt. \tag{8}$$

where $p$ and $q$ are two probabilities on $\mathbb{R}^M$ and $q$ is absolutely continuous with respect to $p$. Note that the function $q \rightarrow H(q,p)$ is convex and non-negative, for any given probability $p$. Furthermore, we have the following fundamental property which was proved in [26]:

$$H(q,p) = 0 \text{ iff } q = p.$$

The Hellinger distance $H$ between the joint density $f_{\boldsymbol{Y}}(\cdot)$ of the random vector $\boldsymbol{Y} := (Y_1, \ldots, Y_p)^\top \in \mathbb{R}^p$, $p \geq 1$, and the product of the the marginal densities $f_{Y_i}$ of the components $Y_i$, $i \in \{1, \ldots, p\}$, is given by

$$H\left(\prod_{i=1}^{p} f_{Y_i}(y_i), f_{\boldsymbol{Y}}(\boldsymbol{y})\right) := \int_{\mathbb{R}^M} 2\left(\sqrt{\frac{\prod_{i=1}^{p} f_{Y_i}(y_i)}{f_{\boldsymbol{Y}}(\boldsymbol{y})}} - 1\right)^2 f_{\boldsymbol{Y}}(\boldsymbol{y}) dy_1, \ldots, dy_p, \\ := \mathbb{E}\left[2\left(\sqrt{\frac{\prod_{i=1}^{p} f_{Y_i}(y_i)}{f_{\boldsymbol{Y}}(\boldsymbol{y})}} - 1\right)^2\right]. \tag{9}$$

where $\mathbb{E}$ is the mathematical expectation.
Note that $H\left(\prod_{i=1}^{p} f_{Y_i}, f_{\boldsymbol{Y}}\right)$ is non-negative and reaches its minimum value zero only when the components of the random vector $\boldsymbol{Y}$ are statistically independent, in other words: $\prod_{i=1}^{p} f_{Y_i}(\cdot) = f_{\boldsymbol{Y}}(\cdot)$.
From eq. (9) and using formula (6), the hellinger distance $H\left(\prod_{i=1}^{p} f_{Y_i}, f_{\boldsymbol{Y}}\right)$ can be reformulated as follows

$$H\left(\prod_{i=1}^{p} f_{Y_i}, f_{\boldsymbol{Y}}\right) := \int_{[0,1]^p} 2\left(\sqrt{\frac{1}{\boldsymbol{c}_{\boldsymbol{Y}}(\boldsymbol{u})}} - 1\right)^2 \boldsymbol{c}_{\boldsymbol{Y}}(\boldsymbol{u}) d\boldsymbol{u}, \\ := \mathbb{E}\left[2\left(\sqrt{\frac{1}{\boldsymbol{c}_{\boldsymbol{Y}}(\boldsymbol{u})}} - 1\right)^2\right]. \tag{10}$$

This last equation implies that the Hellinger distance between the product of the marginal densities and the joint density of the random vector $\boldsymbol{Y}$ can be also defined as the Hellinger distance between the copula density of independence $\boldsymbol{c}_{\prod}$, and copula density $\boldsymbol{c}_{\boldsymbol{Y}}$ of the random vector $\boldsymbol{Y}$

$$H\left(\prod_{i=1}^{p} f_{Y_i}, f_{\boldsymbol{Y}}\right) := H\left(\boldsymbol{c}_{\prod}, \boldsymbol{c}_{\boldsymbol{Y}}\right). \tag{11}$$

## 4. The proposed approach

Before going into details let us first present the discrete version of the BSS problem. Considering the source signals $\boldsymbol{s}(n)$, $n = 1, \ldots, N$ as $N$ copies of the random source vector $\boldsymbol{S}$ the eq. (2) will take the following form:

$$\boldsymbol{X} := \boldsymbol{M}\boldsymbol{S},\ n = 1, \ldots, N. \tag{12}$$

Hence, $\boldsymbol{y}(n) := \boldsymbol{B}\boldsymbol{x}(n)$, $n = 1, \ldots, N$ is, $N$ copies of the random source vector $\boldsymbol{Y} := \boldsymbol{B}\boldsymbol{X}$.

### 4.1. A SEPARATION PROCEDURE FOR INDEPENDENT SOURCES.

As shown in the previous section the Hellinger distance $H\left(\boldsymbol{c}_{\prod}, \boldsymbol{c}_{\boldsymbol{Y}}\right)$ between the copula density of independence and the copula density of the random variable $\boldsymbol{Y}$ is always positive and only achieve its minimum zero if the components of $\boldsymbol{Y}$ are statistically independent and the un-mixing matrix $\boldsymbol{B} = \boldsymbol{D}\boldsymbol{P}\boldsymbol{M}^{-1}$, where $\boldsymbol{D}$ and $\boldsymbol{P}$ are, a diagonal and permutation matrix respectively.
For a successful separation, the idea is to minimize an estimate $\widehat{H}\left(\boldsymbol{c}_{\prod}, \boldsymbol{c}_{\boldsymbol{Y}}\right)$ constructed from the data $\boldsymbol{y}(1), \ldots, \boldsymbol{y}(n)$. Therefore, the separation matrix is calculated in this fashion

$$\widehat{\boldsymbol{B}} = \arg \min_{\boldsymbol{B}} \widehat{H}\left(\boldsymbol{c}_{\prod}, \boldsymbol{c}_{\boldsymbol{Y}}\right). \tag{13}$$

That results in approximating the components $\widehat{\boldsymbol{y}}(n) = \widehat{\boldsymbol{B}}\,\boldsymbol{x}(n), \quad n = 1, \ldots, N$. Considering eq. (10), we introduce the following estimate of the distance $H\left(\boldsymbol{c}_{\prod}, \boldsymbol{c}_{\boldsymbol{Y}}\right)$ as

$$\widehat{H}\left(\boldsymbol{c}_{\prod}, \boldsymbol{c}_{\boldsymbol{Y}}\right) := \frac{2}{N}\sum_{n=1}^{N}\left(\sqrt{\frac{1}{\widehat{c}_{\boldsymbol{Y}}(\widehat{F}_{Y_1}(y_1(n)), \ldots, \widehat{F}_{Y_p}(y_p(n)))}} - 1\right)^2 \tag{14}$$

where the kernel estimate of the copula density $c_{\boldsymbol{Y}}(.)$ is of the form

$$\widehat{c}_{\boldsymbol{Y}}(\boldsymbol{u}) := \frac{1}{NH_1 \cdots H_p}\sum_{m=1}^{N}\prod_{j=1}^{p} k\left(\frac{\widehat{F}_{Y_j}(y_j(m)) - u_j}{H_j}\right), \forall \boldsymbol{u} \in [0,1]^p, \tag{15}$$

with $\widehat{F}_{Y_j}(x)$, $j = 1, \ldots, p$ the smoothed estimate of the marginal distribution functions $F_{Y_j}(x)$ for the random variable $Y_j$. For any real value $x \in \mathbb{R}$, $\widehat{F}_{Y_j}(x)$ is defined by

$$\widehat{F}_{Y_j}(x) := \frac{1}{N}\sum_{m=1}^{N} K\left(\frac{y_j(m) - x}{h_j}\right), \forall j = 1, \ldots, p \tag{16}$$

$K(.)$ is a symmetric and centered probability density and the primitive of a kernel $k(.)$ which we chose it to be the regular Gaussian density in this study. A more acceptable kernel choice $k(.)$ that copes with the boundary effect can be rendered according to [27] to approximate the copula density.
$H_1, \ldots, H_p$ and $h_1, \ldots, h_p$ which are the bandwidth parameters seen in eqs. (15) and (16) are chosen with conform to the Silverman's thumb rule [28]. Hence, for all $j = 1, \ldots, p$, we have:

$$\begin{cases} H_j = \left(\dfrac{4}{p+2}\right)^{\frac{1}{p+4}} N^{\frac{-1}{p+4}} \widehat{\Sigma}_j, \\ h_j = \left(\dfrac{4}{3}\right)^{\frac{1}{5}} N^{\frac{-1}{5}} \widehat{\sigma}_j, \end{cases} \tag{17}$$

where $\widehat{\sigma}_j$ and $\widehat{\Sigma}_j$ are, respectively, the empirical standard deviation of $y_j(1), \dots, y_j(N)$ and $\widehat{F}_{Y_j}(y_j(1)), \dots, \widehat{F}_{Y_j}(y_j(N))$.

We present a two-steps approach to estimate the separation matrix $\widehat{\boldsymbol{B}}$. First, it is normal to start with a normalization stage, namely the spacial whitening, where we transform the data $\boldsymbol{x}$ by a $p \times p$-matrix $\boldsymbol{W}$ such that

$$\boldsymbol{z}(n) = \boldsymbol{W}\boldsymbol{x}(n),\ n = 1, \dots, N, \tag{18}$$

with

$$\mathcal{R}_{\boldsymbol{Z}} = \boldsymbol{W}\mathcal{R}_{\boldsymbol{X}}\boldsymbol{W}^T = \mathcal{I}_d \tag{19}$$

$\mathcal{R}_{\boldsymbol{Z}}$ and $\mathcal{R}_{\boldsymbol{X}}$ are the auto-covariances of $\boldsymbol{Z}$ and $\boldsymbol{X}$ successively. The spatial whitening of the observations consists in un-correlating the signals paired with a unit power constraint [1,29].

The second step consists in applying series of Givens rotations, minimizing the estimate of the Hellinger distance. Let $\boldsymbol{U} \in \mathbb{R}^{p\times p}$ be a unitary matrix, satisfying $\boldsymbol{U}\boldsymbol{U}^\top = \boldsymbol{I}_p$. This matrix can be written as $\boldsymbol{U}(\alpha) := \prod_{1\le i<\le p} G(i,k,\alpha_m)$, where $G(i,k,\alpha_m)$ is the $p \times p$-matrix with the following inputs, for all $1 \le j, l \le p$,

$$G(i,k,\alpha_m)_{j,l} := \begin{cases} cos(\alpha_m) & \text{if} \quad j=i, l=i \quad \text{or} \quad j=k, l=k; \\ sin(\alpha_m) & \text{if} \quad j=i, l=k; \\ -sin(\alpha_m) & \text{if} \quad j=k, l=i; \\ 1 & \text{if} \quad j=l; \\ 0 & \text{else}, \end{cases} \tag{20}$$

the rotation angles $\alpha_m \in ]-\pi/2, \pi/2[$, $m = 1, \dots, p(p-1)/2$, are the elements of the vector $\alpha$.

The un-mixing matrix is written as follow: $\boldsymbol{B} = \boldsymbol{U}(\alpha)\boldsymbol{W}$ hence, the estimated sources take the upcoming form: $\boldsymbol{y}(n) = \boldsymbol{U}(\alpha)\boldsymbol{z}(n),\ n = 1, \dots, N$. Accordingly, the estimate $\widehat{H}\left(\boldsymbol{c}_\Pi, \boldsymbol{c}_{\boldsymbol{Y}}\right)$ is a function of the parameter vector $\alpha$ which can be computed using a gradient descent algorithm by minimizing $\widehat{\alpha} := \arg\min_{\boldsymbol{\alpha}} \widehat{H}\left(\boldsymbol{c}_\Pi, \boldsymbol{c}_{\boldsymbol{Y}}\right)$ with respect to $\alpha$. The un-mixing matrix is then estimated by

$$\widehat{\boldsymbol{B}} = \boldsymbol{U}(\widehat{\alpha})\boldsymbol{W}, \tag{21}$$

which results in approximating the source signals:

$$\widehat{\boldsymbol{y}}(n) = \widehat{\boldsymbol{B}}\,\boldsymbol{x}(n) = \boldsymbol{U}(\widehat{\alpha})\boldsymbol{W}\boldsymbol{x}(n), \quad n = 1, \dots, N. \tag{22}$$

The gradient in $\alpha$ of $H\left(\boldsymbol{c}_{\Pi}, \boldsymbol{c}_{\boldsymbol{Y}}\right)$ can be calculated from the proper definitions of the estimates as follows:

$$\begin{aligned}
\frac{d\widehat{H}\left(\boldsymbol{c}_{\Pi}, \boldsymbol{c}_{\boldsymbol{Y}}\right)}{d\boldsymbol{\alpha}} &:= -\frac{2}{N}\sum_{n=1}^{N}\frac{d}{d\boldsymbol{\alpha}}\left(\sqrt{\frac{1}{\boldsymbol{c}_{\boldsymbol{Y}}(\widehat{F}_{Y_1}(y(n)), \ldots, \widehat{F}_{Y_p}(y_p(n)))}} - 1\right)^2 \\
&:= -\frac{2}{N}\sum_{n=1}^{N}\left(\frac{1}{\boldsymbol{c}_{\boldsymbol{Y}}(\widehat{F}_{Y_1}(y(n)), \ldots, \widehat{F}_{Y_p}(y_p(n)))} - \sqrt{\boldsymbol{c}_{\boldsymbol{Y}}(\widehat{F}_{Y_1}(y(n)), \ldots, \widehat{F}_{Y_p}(y_p(n)))}\right) \\
&\times \frac{\frac{d}{d\boldsymbol{\alpha}}\boldsymbol{c}_{\boldsymbol{Y}}(\widehat{F}_{Y_1}(y(n)), \ldots, \widehat{F}_{Y_p}(y_p(n)))}{\boldsymbol{c}_{\boldsymbol{Y}}(\widehat{F}_{Y_1}(y(n)), \ldots, \widehat{F}_{Y_p}(y_p(n)))^2}
\end{aligned} \tag{23}$$

where,

$$\begin{aligned}
\frac{d\widehat{c}_{\boldsymbol{Y}}(\widehat{F}_{Y_1}(y(n)), \ldots, \widehat{F}_{Y_p}(y_p(n)))}{d\boldsymbol{\alpha}} &= \frac{1}{NH_1\cdots H_p}\sum_{m=1}^{N}\prod_{j=1, j\neq i}^{p} k\left(\frac{\widehat{F}_{Y_j}(y_j(m)) - \widehat{F}_{Y_j}(y_j(n))}{H_j}\right) \\
&\times k'\left(\frac{\widehat{F}_{Y_i}(y_i(m)) - \widehat{F}_{Y_i}(y_i(n))}{H_i}\right)\frac{1}{H_i}\frac{d(\widehat{F}_{Y_i}(y_i(m)) - \widehat{F}_{Y_i}(y_i(n)))}{d\boldsymbol{\alpha}},
\end{aligned} \tag{24}$$

with

$$\frac{d(\widehat{F}_{Y_i}(y_i(m))}{d\boldsymbol{\alpha}} = \frac{1}{Nh_i}\sum_{n=1}^{N} k\left(\frac{y_i(n) - y_i(m)}{h_i}\right)\frac{d}{d\boldsymbol{\alpha}}\left(\frac{y_i(n) - y_i(m)}{h_i}\right), \tag{25}$$

with $\boldsymbol{y}(n) = \boldsymbol{U}(\alpha)\boldsymbol{z}(n),\ n = 1, \ldots, N.$
The following algorithm sums up the proposed approach for the separation of independent components:

**Algorithm 1** The separation algorithm for independent source components.

**Data:** the observations $\boldsymbol{x}(n),\ n = 1, \ldots, N.$
**Result:** the estimated sources $\widehat{\boldsymbol{y}}(n),\ n = 1, \ldots, N.$
**Whitening and Initialization:** $\boldsymbol{z}(n) := \boldsymbol{W}\boldsymbol{x}(n), \widehat{\boldsymbol{y}}_0(n) = \boldsymbol{U}(\widehat{\alpha_0})\boldsymbol{z}(n)$. Given $\varepsilon > 0$ and $\mu > 0$.
**Do:** • **Update** $\boldsymbol{\alpha}$ **and** $\boldsymbol{y}$

$$\boldsymbol{\alpha}_{k+1} = \boldsymbol{\alpha}_k - \mu\frac{d\widehat{H}\left(\boldsymbol{c}_{\Pi}, \boldsymbol{c}_{\boldsymbol{Y}}\right)}{d\boldsymbol{\alpha}}.$$
$$\boldsymbol{y}_{k+1}(n) = \boldsymbol{U}(\alpha_{k+1})\boldsymbol{z}(n),\ n = 1, \ldots, N.$$

• **Until** $||\alpha_{k+1} - \alpha_k|| < \varepsilon$
$\widehat{\boldsymbol{y}}(n) = \boldsymbol{y}_{k+1}(n),\ n = 1, \ldots, N.$

### 4.2. A SEPARATION PROCEDURE FOR DEPENDENT SOURCES.

In this section we tackle the case of dependent source components, hence, we can't use the independent copula density as in the previous section. Denote by $c_{\boldsymbol{S}}(\cdot)$ the unknown semi-parametric copula density of $\boldsymbol{S}$. We assume that it belongs to a set of $L$ candidate semi-parametric models, say,

$$M_l := \{c_l(\cdot;\theta_l);\, \theta_l \in \Theta_l \subset \mathbb{R}\},\, l = 1, \ldots, L. \tag{26}$$

Table 1 gives some examples of semi-parametric copula density models. Each "semiparametric" model $M_l$ for $l = 1, \ldots, L$, satisfies the following identifiability condition : for any regular matrix $\boldsymbol{Q}$, if the copula density, of $\boldsymbol{QS}$, belongs to $\{c_{\theta_l}(\cdot);\, \theta_l \in \Theta_l \subset \mathbb{R}\}$, then $\boldsymbol{Q} = \boldsymbol{DP}$, where $\boldsymbol{D}$ is diagonal and $\boldsymbol{P}$ is a permutation. To get the objective function for dependent sources all we have to do is to replace the copula density of the independence sources in 14 by the semi-parametric copula density $\{c_{\theta_l}(\cdot)$ [10]. The new objective function will be of the following form:

$$H\left(\boldsymbol{c}_{\theta_l}, \boldsymbol{c}_{\boldsymbol{Y}}\right) := \mathbb{E}\left[2\left(\sqrt{\frac{\boldsymbol{c}_{\theta_l}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p)))}} - 1\right)^2\right], \tag{27}$$

this term is non-negative and achieves its minimum value zero iff $\boldsymbol{B} = \boldsymbol{M}^{-1}$ (up to scale and permutation indeterminacies). Therefore, we estimate the demixing matrix by

$$\widehat{\boldsymbol{B}} := \arg\inf_{\alpha} \inf_{\theta_{l*} \in \Theta_{l*}} \widehat{H}\left(\boldsymbol{c}_{\theta_l}, \boldsymbol{c}_{\boldsymbol{Y}}\right), \tag{28}$$

where

$$l^* = \arg\min_{l=1,\ldots,L} \inf_{\alpha} \inf_{\theta_l \in \Theta_l} \widehat{H}\left(c_{\theta_l}, c_{\boldsymbol{Y}}\right), \tag{29}$$

The copula density as well as the marginal distribution functions estimates are defined as before. The solution $\widehat{\boldsymbol{B}}$ can be computed by a gradient descent algorithm with respect to both $\alpha$ and $\theta$ of the criterion function $(\alpha, \theta_l) \mapsto \widehat{H}\left(c_\theta, c_{\boldsymbol{Y}}\right)$ for each model and then choose the solution minimizing the criterion over all considered models. The calculations for the gradient of the Hellinger divergence is the same as the one stated in the case of independence (23). The algorithm (2) summarizes the presented method.

## 5. SIMULATION RESULTS

In the following, we present the results of various simulations that were conducted to test our proposed approach and to better illustrate its performance. Our results will be compared with those obtained by [30](Copula-Alpha), [10](Copula-MI), [3](MI), [31] (JADE), [32] (FastICA), [33] (RADICAL) and [34] (InfoMax) under the same conditions. In all the instances of our experiments the number of samples is $N = 3000$. The matrix used to mix the source components is $\boldsymbol{A} := [1\ 0.7\ 0.7\ ;\ 0.7\ 1\ 0.7\ ;\ 0.7\ 0.7\ 1]$, and $\mu = 0.1$ is the chosen gradient descent parameter. All simulations are iterated 80 times,

**Algorithm 2** The separation algorithm for dependent source components.

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

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \nu \frac{d\widehat{H}\left(\boldsymbol{c}_{\theta_k}, \boldsymbol{c}_{\boldsymbol{Y}}\right)}{d\boldsymbol{\theta}}.$$

$$\boldsymbol{\alpha}_{k+1} = \boldsymbol{\alpha}_k - \mu \frac{d\widehat{H}\left(\boldsymbol{c}_{\theta_{k+1}}, \boldsymbol{c}_{\boldsymbol{Y}}\right)}{d\boldsymbol{\alpha}}.$$

$$\boldsymbol{y}_{k+1}(n) = \boldsymbol{U}(\alpha_{k+1})\boldsymbol{z}(n),\ n = 1, \ldots, N.$$

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

and the accuracy of the estimated sources is calculated using the signal-to-noise-ratio criterion, which is defined by

$$SNR_i := 10 \log_{10} \frac{\sum_{n=1}^{N} s_i(n)^2}{\sum_{n=1}^{N} (\widehat{s}_i(n) - s_i(n))^2},\ i = 1, 2, 3. \tag{30}$$

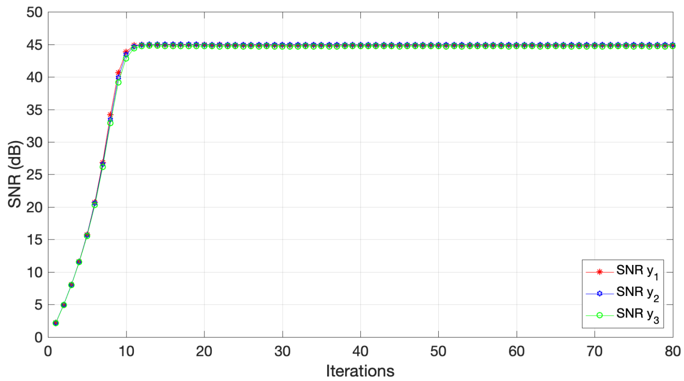
### 5.1. INDEPENDENT SOURCE COMPONENTS

We consider in this experiment three mixed signals from two types of sample sources:
- uniform i.i.d with independent components (see Fig. 1a).
- i.i.d sources with independent components drawn from the 4-ASK (Amplitude Shift Keying) alphabet (see Fig. 1b).

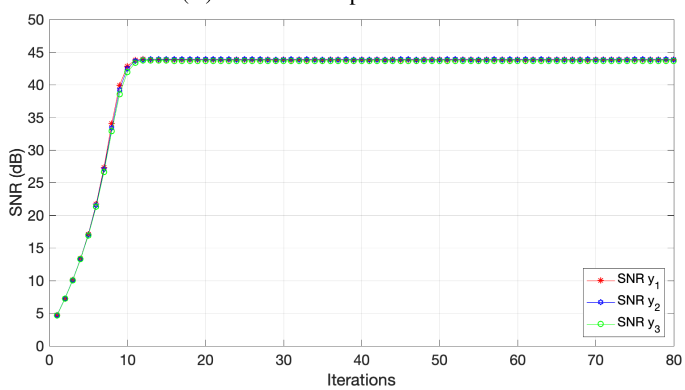
From Figs. 1a- 1b, we observe that for both independent samples the SNR is close to 45dB which is considered highly satisfying for this classical case.In the other hand, figs. 2a- 2b present the criterion value vs iterations. We can see that the separation is achieved when our criterion converges to its minimum value 0.
Table 2 illustrate the different SNR values of the sources, for our approach and other methods, we can see that the method proposed achieves the separation with similar accuracy with a slight improvement for the independent source components case.

### 5.2. DEPENDENT SOURCE COMPONENTS

Within this subsection we demonstrate the ability of the proposed approach (algorithm 2 for dependent sources) to successfully separate mixtures of three dependent signals, we dealt with instantaneous mixtures of tree kinds of sample sources:
- i.i.d (with uniform marginals) sources with dependent components generated from AMH copula with $\theta = 0.75$.
- i.i.d (binary phase-shift keying(BPSK)) sources with dependent components generated from Clayton copula, with $\theta = 1.5$




(A) Uniform independent sources.

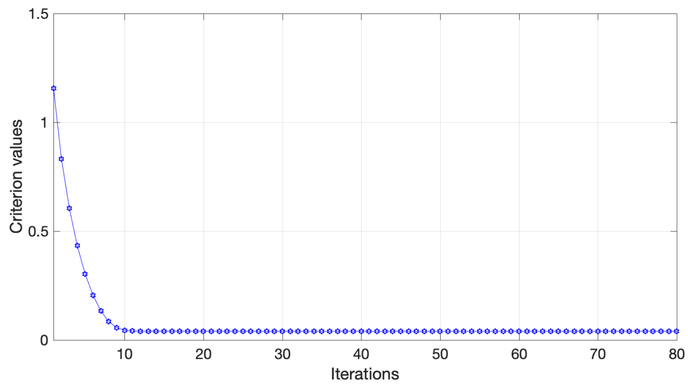



(B) ASK independent sources.

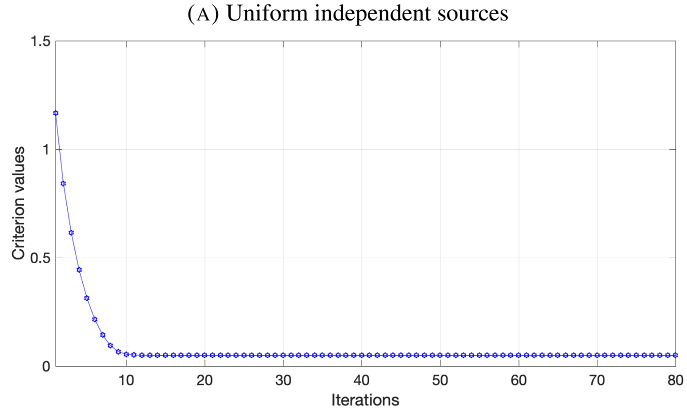
FIGURE 1. Average output SNRs versus iteration number

TABLE 2. Output SNR's for independent source components

| **Type** | **Uniform** | | | **ASK** | | |
|---|---|---|---|---|---|---|
| **Sources** | **S1** | **S2** | **S3** | **S1** | **S2** | **S3** |
| **Our method** | **46.7660** | **46.9336** | **46.8821** | **44.9767** | **44.7364** | **44.8914** |
| **Copula-Alpha** | 46.0311 | 45.3267 | 45.8112 | 43.9051 | 44.1076 | 44.3188 |
| **Copula-MI** | 45.8914 | 45.9560 | 45.7158 | 43.8195 | 43.8786 | 43.6588 |
| **MI** | 45.6757 | 45.6216 | 45.6142 | 43.1125 | 43.0654 | 43.0589 |
| **FastICA** | 43.8497 | 44.4492 | 40.1810 | 42.7949 | 42.0838 | 40.2107 |
| **JADE** | 45.5146 | 44.9522 | 44.9522 | 43.1671 | 43.1508 | 43.1889 |
| **RADICAL** | 44.9902 | 43.9986 | 44.7361 | 42.9023 | 42.4095 | 42.6272 |
| **InfoMax** | 45.3687 | 45.6243 | 44.8909 | 43.8846 | 43.3195 | 43.1528 |




(A) Uniform independent sources




(B) ASK independent sources

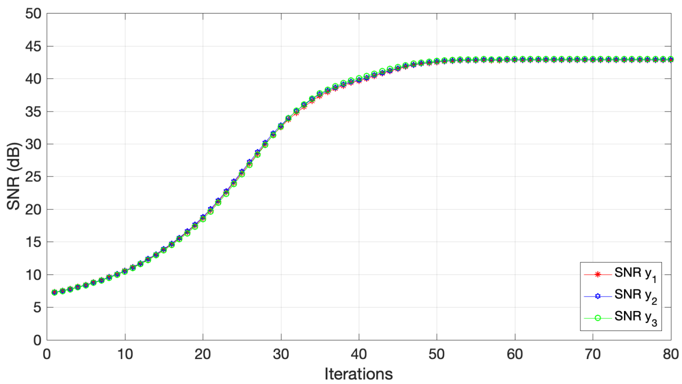
FIGURE 2. The criterion value vs iterations

- i.i.d (with uniform marginals) sources with dependent components generated from Frank copula, with $\theta = 2$.

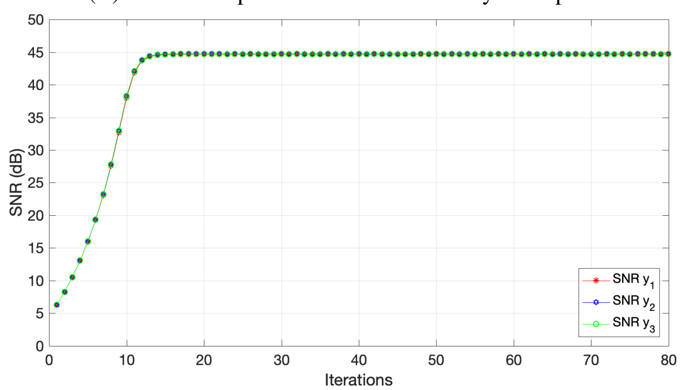
In Figs. 3a- 3b, we have shown the SNRs for dependent sources from Clayton and Frank copulas. From the simulation results it is noticeable that the proposed approach can separate the mixtures of dependent source components, with good performance.

Moreover, Figs. 4a- 4b show the criterion value versus iterations for Clayton and Frank copulas. We can see that the separation is achieved when our criterion converges to its minimum value 0.

Table 3 exhibits the superiority of our proposed approach compared to [30](Copula-Alpha), [10](Copula-MI), [3](MI), [31] (JADE), [32] (FastICA), [33] (RADICAL) and [34] (InfoMax).




(A) Uniform dependent sources from Clayton-copula

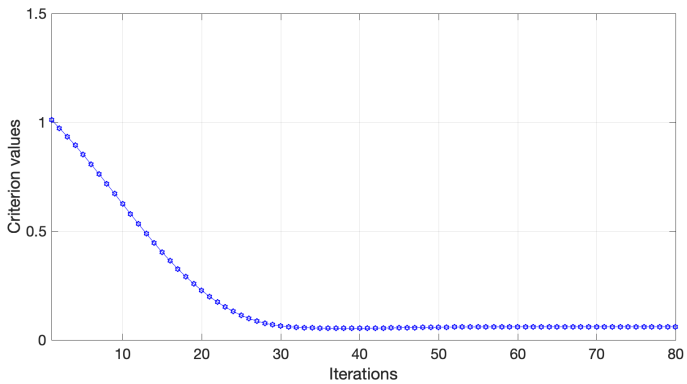



(B) Uniform dependent sources from Frank-copula

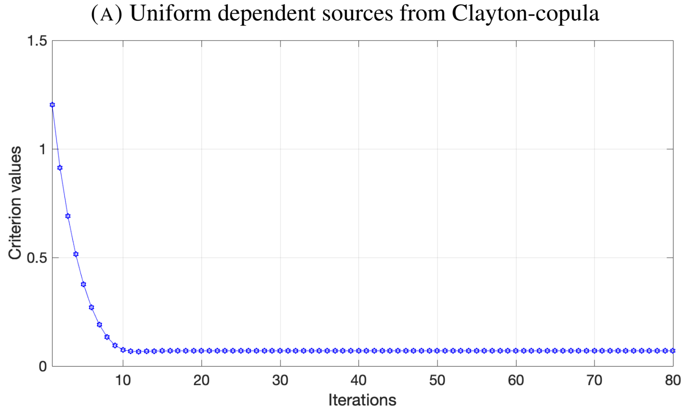
FIGURE 3. Average output SNRs versus iteration number

### 5.3. NOISY SOURCE COMPONENTS

In this subsection we test the accuracy of our approach for noisy data. We work with the same source signals as above and the same conditions with an added white gaussian noise to the observed signals. We take $SNR = -25dB$.
Fig. 5a illustrate the SNR of the independent sources, it can be seen that the proposed approach is able to separate noisy independent sources with good performance, and Fig. 5b shows that when the separation is achieved our criterion converges to its minimum 0. Fig. 6a showcase the SNR of the dependent sources from Clayton copula, the proposed approach is able to separate even noisy dependent sources. Moreover Fig. 6b shows that the criterion in this case also converges to its minimum 0.
Table 4 present the output SNR values of the estimated sources using our approach and the other methods, we can see that the approaches are equivalent, with superiority of our method, in case of noise-contaminated independent source components. On the other




(A) Uniform dependent sources from Clayton-copula




(B) Uniform dependent sources from Frank-copula

FIGURE 4. The criterion value vs iterations

hand, our approach is apt to separate even noisy mixtures of dependent source components with higher accuracy.

TABLE 3. Output SNR's for dependent source components

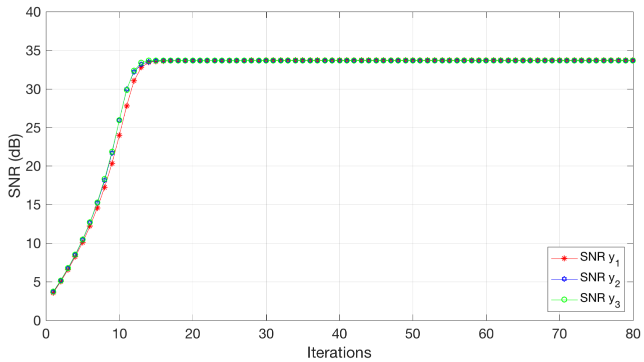
| Copulas | AMH | | | Clayton | | | Frank | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sources** | **S1** | **S2** | **S3** | **S1** | **S2** | **S3** | **S1** | **S2** | **S3** |
| **Our method** | **45.224** | **45.468** | **45.330** | **43.784** | **43.839** | **44.511** | **46.017** | **46.262** | **46.1243** |
| **Copula-Alpha** | 43.144 | 45.093 | 44.350 | 43.569 | 41.934 | 42.885 | 44.947 | 45.841 | 45.0578 |
| **Copula-MI** | 44.230 | 44.277 | 44.251 | 42.815 | 42.825 | 42.953 | 44.678 | 44.724 | 44.6984 |
| **MI** | 15.301 | 15.322 | 15.308 | 10.130 | 10.327 | 10.275 | 18.958 | 18.987 | 18.9681 |
| **FastICA** | 41.160 | 8.0131 | 8.0765 | 38.090 | 5.9141 | 2.4933 | 42.943 | 11.786 | 8.7365 |
| **JADE** | 14.994 | 14.134 | 14.314 | 9.7169 | 9.6885 | 10.710 | 17.063 | 17.225 | 17.5806 |
| **RADICAL** | 9.3148 | 8.7314 | 6.7465 | 9.4081 | 8.7376 | 10.993 | 17.804 | 17.165 | 17.4945 |
| **InfoMax** | 9.2981 | 9.2577 | 9.2871 | 8.2981 | 8.2577 | 8.2871 | 10.513 | 10.398 | 10.4529 |

TABLE 4. Output SNR's for independent and dependent noisy source components

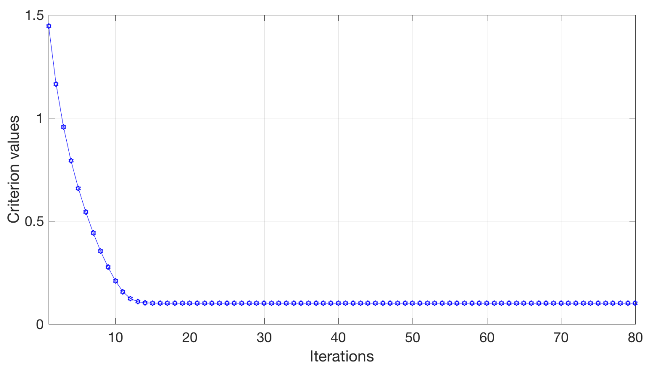
| Copulas | Independence | | | Clayton | | | Frank | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sources** | **S1** | **S2** | **S3** | **S1** | **S2** | **S3** | **S1** | **S2** | **S3** |
| **Our method** | **36.188** | **36.238** | **36.051** | **34.965** | **34.817** | **34.661** | **36.277** | **36.533** | **36.3887** |
| **Copula-Alpha** | 31.021 | 31.065 | 30.901 | 29.137 | 29.185 | 29.767 | 30.444 | 30.427 | 30.4370 |
| **Copula-MI** | 31.021 | 31.065 | 30.901 | 29.137 | 29.185 | 29.767 | 30.444 | 30.427 | 30.4370 |
| **MI** | 30.453 | 30.417 | 30.413 | 7.1605 | 7.0444 | 6.6586 | 12.285 | 12.378 | 12.1843 |
| **FastICA** | 30.211 | 29.679 | 28.277 | 25.576 | 3.0302 | 3.7841 | 28.547 | 8.4895 | 6.6148 |
| **JADE** | 30.233 | 30.220 | 30.249 | 6.9056 | 6.8649 | 7.1214 | 13.281 | 13.371 | 12.8820 |
| **RADICAL** | 30.316 | 29.947 | 30.110 | 6.5064 | 7.4440 | 6.9503 | 12.428 | 14.750 | 12.5084 |
| **InfoMax** | 30.088 | 30.665 | 30.540 | 5.3751 | 5.3627 | 5.3168 | 8.3977 | 8.2703 | 8.3137 |

## 6. CONCLUSION

We have presented a new BSS algorithm, that is able to separate instantaneous linear mixtures of both independent and dependent source components. Our approach proceeds in two steps: First a normalization stage with spatial whitening and the then the application of Givens rotations, minimizing the estimate of the Hellinger distance. This divergence works better in presence of noise and it also converge faster than the usual Kullback-Leibler divergence as illustrated in section "Simulation results" for $3 \times 3$ mixture-sources, where the efficiency and the accuracy of the proposed algorithms is evaluated through the signal-to-noise-ratio criterion. It should be noted that our proposed algorithms are more time-consuming compared to the classic ones, considering that we estimate both copulas




(A) Average output SNRs versus iterations




(B) The criterion value vs iterations

FIGURE 5. Uniform noisy independent sources

density of the vector and the marginal distribution function of each component, however, ours gives better results especially for the dependent noisy source components case.

## REFERENCES

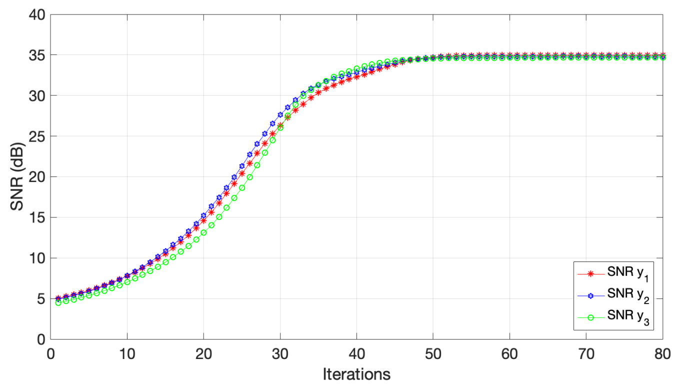

[1] Comon, P.: 'Independent component analysis, a new concept?', *Signal Processing*, 1994, **36**, (3), pp. 287–314

[2] Cardoso, J.F.: 'Blind signal separation: statistical principles', *Proceedings of the IEEE*, 1998, **86**, (10), pp. 2009–2025

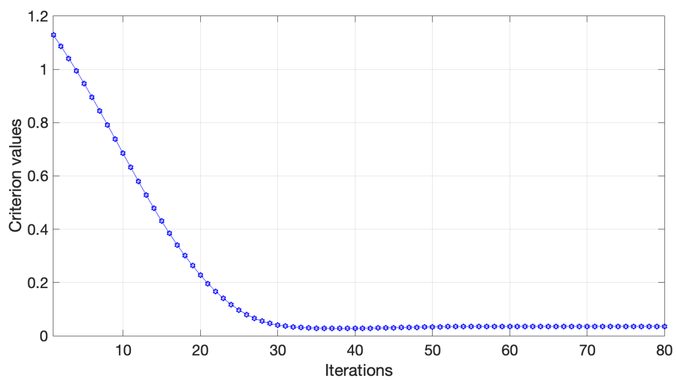
[3] Pham, D.T.: 'Mutual information approach to blind separation of stationary sources', *IEEE Transactions on Information Theory*, 2002, **48**, (7), pp. 1935–1946

[4] El.Rhabi, M., Gelle, G., Fenniri, H., Delaunay, G.: 'A penalized mutual information criterion for blind separation of convolutive mixtures', *Signal Processing*, 2004, **84**, (10), pp. 1979–1984

[5] El.Rhabi, M., Fenniri, H., Keziou, A., Moreau, E.: 'A robust algorithm for convolutive blind source separation in presence of noise', *Signal Processing*, 2013, **93**, (4), pp. 818–827




(A) Average output SNRs versus iterations




(B) The criterion value vs iterations

FIGURE 6. Uniform dependent noisy sources from Clayton-copula


[6] Ghazdali, A., El.Rhabi, M., Fenniri, H., Hakim, A., Keziou, A.: 'Blind noisy mixture separation for independent/dependent sources through a regularized criterion on copulas', *Signal Processing*, 2017, **131**, pp. 502–513

[7] Pham, D.T.: 'Blind separation of instantaneous mixture of sources based on order statistics', *IEEE Trans on Signal Processing*, 2000, **2**, pp. 363–375

[8] M..Castella, E.M. S. Rhioui, Pesquet, J.: 'Quadratic higher order criteria for iterative blind separation of a mimo convolutive mixture of sources', *IEEE Trans on Signal Processing*, 2007, **55**, (1), pp. 218–232

[9] Comon, P., Jutten, C.: 'Handbook of Blind Source Separation: Independent component analysis and applications'. (Oxford: Academic press, 2010)

[10] Keziou, A., Fenniri, H., Ghazdali, A., Moreau, E.: 'New blind source separation method of independent/dependent sources', *Signal Processing*, 2014, **104**, pp. 319–324

[11] Ghazdali, A., Hakim, A., Laghrib, A., Mamouni, N., Raghay, S.: 'A new method for the extraction of fetal ecg from the dependent abdominal signals using blind source separation and adaptive noise cancellation techniques', *Theoretical Biology and Medical Modelling*, 2015, **12**, (1), pp. 25

[12] Beran, R.: 'Minimum hellinger distance estimates for parametric models', *The annals of Statistics*, 1977, **5**, (3), pp. 445–463

[13] Lindsay, B.G.: 'Efficiency versus robustness: the case for minimum hellinger distance and related methods', *The annals of statistics*, 1994, **22**, (2), pp. 1081–1114

[14] Sklar, M.: 'Fonctions de répartition à $n$ dimensions et leurs marges', *Publ Inst Statist Univ Paris*, 1959, **8**, pp. 229–231
[15] Clayton, D.G.: 'A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence', *Biometrika*, 1978, **65**, (1), pp. 141–151
[16] M..M..Ali, M.S.H. N. N. Mikhail: 'A class of bivariate distributions including the bivariate logistic', *J Multivariate Anal*, 1978, **8**, (3), pp. 405–412
[17] Frank, M.J.: 'On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$', *Aequationes Math*, 1979, **19**, (2-3), pp. 194–226
[18] Nelsen, R.B.: 'An introduction to copulas'. (New Haven: Springer, 2007)
[19] Joe, H.: 'Multivariate models and dependence concepts'. (Chapman & Hall London: Boca Raton, 2001)
[20] Schwarz, G.: 'Estimating the dimension of a model', *The annals of statistics*, 1978, **6**, (2), pp. 461–464
[21] C..Genest, L.P.R. K. Ghoudi: 'A semiparametric estimation procedure of dependence parameters in multivariate families of distributions', *Biometrika*, 1995, **82**, (3), pp. 543–552
[22] Tsukahara, H.: 'Semiparametric estimation in copula models', *Canad J Statist*, 2005, **33**, (3), pp. 357–375
[23] Csiszár, I.: 'Eine informationstheoretische ungleichung und ihre anwendung auf beweis der ergodizitaet von markoffschen ketten', *Magyer Tud Akad Mat Kutato Int Koezl*, 1964, **8**, pp. 85–108
[24] Beran, R., et al.: 'Minimum hellinger distance estimates for parametric models', *The annals of Statistics*, 1977, **5**, (3), pp. 445–463
[25] Jiménz, R., Shao, Y.: 'On robustness and efficiency of minimum divergence estimators', *Test*, 2001, **10**, (2), pp. 241–248
[26] Csiszár, I.: 'Information-type measures of difference of probability distributions and indirect observation', *studia scientiarum Mathematicarum Hungarica*, 1967, **2**, pp. 229–318
[27] Omelka, M., Gijbels, I., Veraverbeke, N., et al.: 'Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing', *The Annals of Statistics*, 2009, **37**, (5B), pp. 3023–3058
[28] Silverman, B.W.: 'Density estimation for statistics and data analysis'. (New Haven: CRC press, 1986)
[29] Belouchrani, A., Abed.Meraim, K., Cardoso, J.F., Moulines, E.: 'A blind source separation technique using second-order statistics', *IEEE Transactions on signal processing*, 1997, **45**, (2), pp. 434–444
[30] Ourdou, A., Ghazdali, A., Laghrib, A., Abdelmoutalib, M.: 'Blind separation of instantaneous mixtures of independent/dependent sources', *Circuits, Systems, and Signal Processing*, 2021, **40**, pp. 1–24
[31] Cardoso, J.F.: 'A. souloumiac, blind signal beamforming for non gaussian signals', *Proceedings of the IEE*, 1993, **140**, (6), pp. 362–370
[32] Hyvärinen, A., Oja, E.: 'A fast fixed-point algorithm for independent component analysis', *Neural computation*, 1997, **9**, (7), pp. 1483–1492
[33] Miller, E.G., Fisher.III, J.W.: 'Independent components analysis by direct entropy minimization', *Computer Science*, 2003,
[34] Bell, A.J., Sejnowski, T.J.: 'An information-maximization approach to blind separation and blind deconvolution', *Neural computation*, 1995, **7**, (6), pp. 1129–1159