KASDI MERBAH UNIVERSITY
OUARGLA

Faculty of Mathematics and Matter Sciences

DEPARTEMENT OF MATHEMATICS

MASTER

specialty : Mathematics

Option : Probability and statistics

by : AMINA TRABELSI

Title

Higher Order Statistics And Source separation

date :30/09/2020

jury committee :

Mr.Akti Mohamed	KASDI Merbah University- Ouargla	Chairman
Mr.Mansoul Brahim	KASDI Merbah University- Ouargla	Examiner
Mr.Boussaad Abdalmalek	KASDI Merbah University- Ouargla	supervisor

DEDICATION

To my dear mother

God made the paradise under our feet, and a crown above our heads

To my father

The apple of my eyes that fought for us all the hardships of life. May God protect you

To my brother and sisters

you are my permanent bond, I hope you are bright stars in your world

To my loyal friend Sarah

To my family and friends

I wish to you the best

I Dedicate to you my thesis

ACKNOWLEDGEMENT

praise to allah ,the almighty ,with whose gracious help it made it easy to finish this project successfully.

Above all, I would like to express my great thanks and apperciation to my supervisor,"**Abdelmalek Boussaad**"whose support ,constant guidance and for all the advice he gave during this period that made my thesis work possible .

Many thanks to my parents for their supports during my whole life ,for their prayers and to all my teachers

Contents

List of Figures

Notations

➤ BSS : Blind Source Separation.

➤ HOS : Higher Order Statistics.

➤ TF : fourier transform.

➤ rv : random variable.

➤ H,G : mixing, and separating operators.

➤ R : Set of real numbers.

➤ E[.] : mathematical expectation.

➤ Var[.] : The variance.

➤ Det(C) : determinant of matrix C.

➤ z^*, x^* : Complex conjugation.

➤ $(.)^T$: Transposition (without conjugation).

➤ $(.)^\dagger$: Transposition and conjugation.

➤ $\mathfrak{Re}$: Real part.

➤ $\mathfrak{Im}$: Imaginary part.

➤ $Cum[.]$: Cumulant.

➤ $Cum[x,y,z]$: Cumulant des variables x,y,z.

➤ $\mu^{jk}_{x(i)}$: moments of random variables x_i, x^*_j, x^*_k.

➤ $\mu'_{x(p)}$: moments centered order p of x.

➤ $k^{(2)}_{x(2)}$: cumulant $cum[x,x,x^*,x^*]$.

➤ $k^{j}_{x_{il}}$: cumulants of random variables x_i, x^*_j, x_l.

➤ $k^{jk}_{x_{il}}$: standardized cumulants of random variable x_i, x^*_j, x_l.

➤ $I(p_X)$: mutual information of x.

➤ $k(p_V, p_W)$: Kullback divergence between p_V and p_W.

➤ $S(p_x)$: the entropy x

INTRODUCTION

Blind Source Separation (BSS) is an important signal processing problem, it was proposed in the late 1980s by [1].Due to its great practical value, many researchers in the area of signal processing and statistics have been focused on this problem in the aim to propose algorithms. BSS has become an essential tool of development in many engineering areas, mainly in the biomedical sciences, image processing, earth science, econometrics, text data mining, and speech signal communication. The BSS problem consists extracting and recovering a set of unknown source signals from a mixture of them, in general case the sources assumed that are non-Gaussian signals and statistically independent of one another. Mathematically, the BSS problem can be modeled as follows .Let $x(t) = (x_1(t), x_2(t), ..., x_p(t))^T$ the $p\times 1$ source vector, $y(t) = (y_1(t), y_2(t), ..., y_q(t))^T$ the $q\times 1$ observations vector, thus the BSS problem is defined by the following formula:

$$y(t) = H(x(t)) \tag{1}$$

where H is called the mixing operator which is unknown , the sources vector $x(t)$ is also unknown and the observations vector $y(t)$ is known. Indeed, the BSS problem estimates the inverse function G such that:

$$\widehat{x}(t) = G[H(x(t))] \tag{2}$$

where the vector $\widehat{x}(t)$ is the estimation of the sources vector $x(t)$.

In our dissertation, we focus on the linear BSS problems, where H is linear in eq (1) and methods which are based on the higher order statistics (HOS) to estimate the source signals. HOS have been proposed as a statistical tools to separate an i.i.d non-Gaussian signals with at most a single signal having a Gaussian distribution. Concerning this manuscript is organized as follows.

- **Chapter 01**: Blind Source Separation
  This chapter is devoted to the presentation of BSS problem. We explain the mathematical formulation of this problem, the different models of mixture, and BSS methods for instantaneous linear mixing systems.

- **Chapter 02**: Higher Order Statistics
  In this chapter, we outline the higher order statistics (HOS) (cumulants and moments of order greater than two) with their theoretical properties in order to separate systems of the class linear and instantaneous.

- **Chapter 03**:Blind source separation using an algebraic method
  In this final chapter, we try to detail an algebraic technique [16] for BSS problems which is based on the fourth-order cumulants. Experiments have been carried out in this chapter with different signals using this method in the aim to validate the effectiveness of HOS in the area of source separation.
  Finally, a conclusion is written to summarize this work.

Chapter 1

Blind source separation

1.1 Overview of Blind Source Separation

Blind Source Separation is a powerful signal processing tool that was proposed by [1] in the context of statistical signal processing, and information theory. BSS has become a very important topic in different areas of research and development, such as biomedical engineering, speech signal communication, image processing, earth science, artificial neural networks, econometrics, etc. In the broad sense, the problem of BSS consists extracting and recovering a set of unknown source signals from a set of observed signal that result from an unknown mixture model of these signal sources. Blinde means that the source signals are unknown (unobserved) [2]. As shown in [3] (BSS) is implicated in different areas such as, in [4, 5, 6] authors use (BSS) in various applications of biomedical engineering, N. Charkani in [7] is interested to apply the (BSS) in radio-communication field, especially for mobile-phones, some authors utilize (BSS) as a tool in the nuclear reactor monitoring [12]. A detailed overview concerning the different applications of (BSS) is shown in [2]. A famous example of (BSS) is the "cocktail party"problem. Suppose that you are in a room where there are a variety of sounds, for example a people who are talking. In this case, a person can distinguish between different voices getting mixed up in his ears, and so he can identify a particular speaker's voice and understand her. Its the same case for the problem of (BSS) which that offers a solution of this problem by separating the source signals from mixtures recorded by microphones placed in different locations. The Figure1.1 clarifies this example of blind source separation.

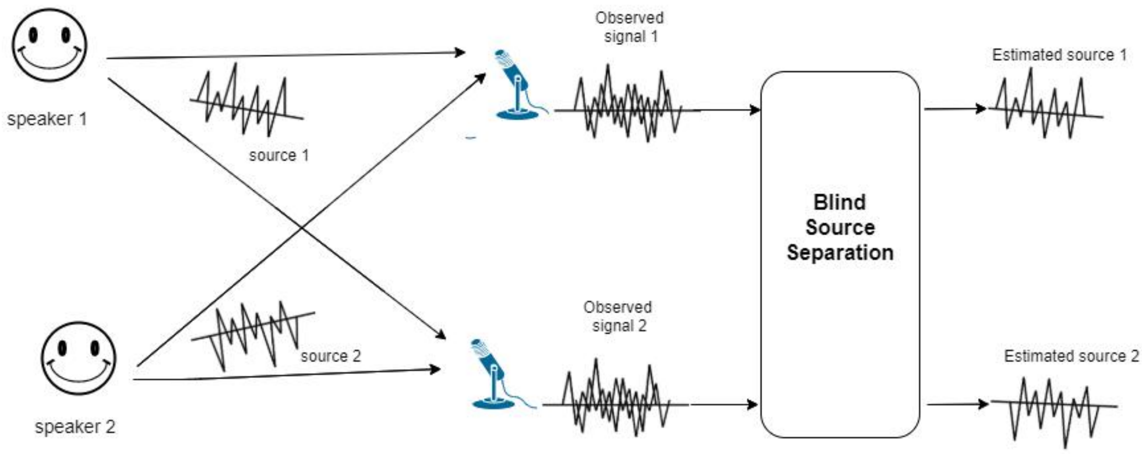

Figure 1.1: Blind source separation diagram

1.2 Mathematical modeling of the BSS problem:

As shown in [3],let $x(t) = (x_1(t), x_2(t), ..., x_p(t))^T$ denotes the $p \times 1$ source vector and $y(t) = (y_1(t), y_2(t), ..., y_q(t))^T$ the $q \times 1$ observation vector.

The problem of BSS can be formulated mathematically as follows:

$$y(t) = H(x(t)) \tag{1.1}$$

where H is called the mixing operator which is unknown , the source vector $x(t)$ is also unknown and the observation vector $y(t)$ is known. The aim of BSS consists on the estimation of an inverse function G such that:

$$\widehat{x}(t) = G[H(x(t))] \tag{1.2}$$

where the vector $\widehat{x}(t)$ is the estimation of the source vector $x(t)$. So we can represent the steps of BSS in the figure 1.2 below.

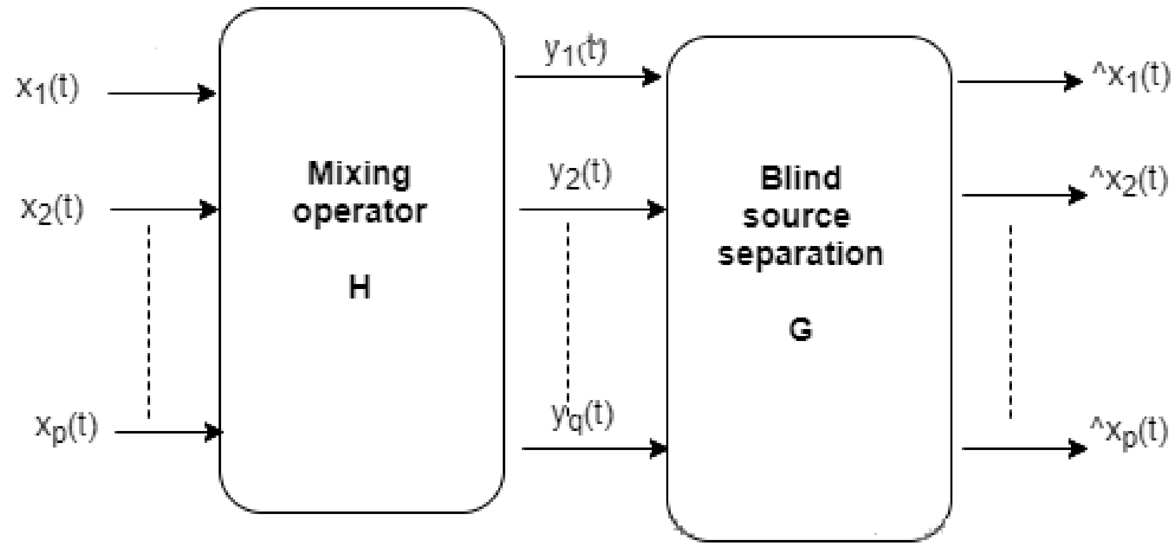

Figure 1.2: BSS problem

The majority of source separation algorithms can be classified according to the type of mixture, the assumptions assumed on the source signals and the number of observed signals compared to the number of source signals.

1.3 Models and assumptions:

In the general case of BSS problem,H in eq(1) is non-linear function which depends on the present and the past of the source signals, thus the BSS methods can be classified into two classes

In the first class, we find methods for BSS problems of linear mixture, where the function H in eq (1) is linear, that means the observations are linear mixtures of source signals. The BSS methods which treats this kind of problems are the most studied due to the simplicity of the linear model.

In the second class, we find methods for BSS problems of non-linear mixture, in this case H in eq(1) is non-linear. The extension of BSS methods to the non-linear mixture is still less studied due to their complexity. A non-linear models in this context have been proposed as shown in [2], for example, the model post-non-linear where the mixture is formed by a linear part and followed by a non-linear distortion. In our dissertation we focus on the BSS problems of linear mixture.

1.3.1 Linear mixtures:

In the subject of source separation, linear mixtures models which have been studied widely can be grouped in the following classes.

1.3.1.1 Instantaneous Linear Mixing:

This model assumes that, each observed signal from the L sensors is a linear combination of K statistically independent sources. Thus the following formula expresses the linear time-invariant instantaneous mixing model.

$$x_i(t) = \sum_{j=1}^{K} a_{ij} s_j(t) \qquad 1 \leq t \leq N \tag{1.3}$$

Where $x_i(t)/i = \overline{1,L}$ are the observed signals,$a_{ij}(t)/i = \overline{1,L}; j = \overline{1,K}$ are mixing parameters, and $s_j(t)/j = \overline{1,K}$ are the source signals.
This process of mixing can be written in the following matrix form:

$$X(t) = AS(t) \tag{1.4}$$

Where $X(t) = (x_1(t), x_2(t), ..., x_L(t)$ is the observed signals vector, $S(t) = (s_1(t), s_2(t), ..., s_K(t))^T$ is the source vector, and $A \in M_{L,K}$ is the mixing matrix.

1.3.1.2 Linear Convolution Mixing Model:

The convolution mixing model assumes that there are K statistically independent sources $s_j(t)/j = \overline{1,k}$ received by L sensors after the convolution mixing process. The observed signals are denoted by $x_i(t)/i = \overline{1,L}$ and the convolution mixing model can be expressed by the following formula

$$x_i(t) = \sum_{j=1}^{K} \sum_{\tau=-\infty}^{\infty} a_{ij}(\tau) s_j(t-\tau) \tag{1.5}$$

In this dissertation, we only concentrate on the instantaneous linear mixing models.

1.3.2 Assumptions:

Authors in [3] affirm that, in blind source separation it is widely used the following assumptions

- **Assumption 01:**The sources are statistically independent. This assumption is an important key for all blind separation algorithms.

- **Assumption 02:** The sources have a non-Gaussian distribution, or precisely, at most one of them can be Gaussian.

- **Assumption 03:**The channel can be instantaneous and the matrix H is assumed to be invertible.

1.4 BSS Methods for instantaneous linear mixing:

The instantaneous linear mixing style is the simplest model and forms the foundation for the other mixing styles. Different methods have been developed for solving this kind of BSS problem. As shown in [12], all types of BSS methods can be categorized as follows:

- Methods based on Independent Component Analysis (ICA)
- Methods which use Sparse Component Analysis (SCA) as an essential tool.
- Methods that involve Nonnegative Matrix Factorization (NMF) technique in the separation process.
- Methods based on Bayesian Approach (BA).

In our case, we try to focus on the first kind of methods which is based on Independent Component Analysis (ICA).

1.4.1 Independent Component Analysis (ICA):

ICA (independent component analysis) is among the most studied methods of Blind Source Separation. The aim of these methods is to apply some transformations on the observed signals in order to obtain a statistically independent signals output. From a mathematical point of view, the probability density function (PDF) is used to define the statistical independence concept of a random vector. If $s = (s_1, s_2, ..., s_L)$ is a random vector with a joint PDF $p(s_1, s_2, ..., s_L)$ then: The random variables $S_i / i = \overline{1, L}$ are statistically independent if and only if the probability $p(s_1, s_2, ..., s_L)$ can be decomposed into the product of the marginal densities, that means

$$p(s_1, s_2, ..., s_L) = \prod_{k=1}^{L} p_k(S_k) \tag{1.6}$$

The majority of ICA methods are based on the following assumptions, sources are statistically independent and have a non-Gaussian distribution, or precisely, at most one of them can be Gaussian. A several criteria have been used to measure the statistical independence between the source signals. In the next, we try to present the most well-known criteria [12].

1.4.1.1 Maximization of non-Gaussianity:

The central limit theorem is a classical result in probability theory, tells that the distribution of a sum of independent random variables tends toward a Gaussian distribution. Among the most used criteria in the BSS methods based on ICA is the maximization of non-Gaussianity. Indeed, the non-Gaussianity can be used as a necessary condition to separate the i.i.d. source signals. This condition is based on the central limit theorem (CLT) which is a classical result in probability theory, tells that the distribution of a sum of independent random variables tends toward a Gaussian distribution. Thus, in BSS problems, it suffices to increase the non-Gaussianity of the estimated sources to obtain independent components. The famous measure of non-Gaussianity is the kurtosis (fourth-order cumulant) [14] defined by the following formula.

$$Kurt(X) = \frac{E(X^4)}{E(X^2)^2} - 3 \tag{1.7}$$

Where x is a centered random variable and E(.) is the mathematical expectation operator. The kurtosis is a statistical measure that measures the distance between the probability density of a random variable and a Gaussian density. Therefore, maximizing the absolute value of the kurtosis equivalently to maximizing the non-Gaussianity of the estimated variable, which allows to extract the independent sources. Many works have been published in this context, among these works we cite the famous paper [13] where authors propose a fast ICA method.

1.4.1.2 Mutual Information Minimization:

Mutual information is another criterion which is involved in the ICA methods by Comon P in [15]. This criterion measures the statistical independence of random variables by evaluating the similarity between the joint probability density and the product of marginal densities. This similarity is calculated by the Kullback-Leibler divergence as follows

$$I(X_1, X_2, ..., X_n) = \int_{-\infty}^{\infty} p(x_1, x_2, ..., x_n) \log\left(\frac{p(x_1, x_2, ..., x_n)}{\prod_{i=1}^{n} p_i(x_i)}\right) \tag{1.8}$$

Where $X_1, X_2, ..., X_n$ are random variables, $p(x_1, x_2, ..., x_n)$ is the joint probability density of these random variables and $p_i(x_i)$ is the marginal density of x_i The mutual information criterion is always positive and is zero only when the variables $\{X_i\}_{i=1,n}$ are statistically independent.

1.4.1.3 Maximum Likelihood:

The idea of the ICA methods which are based on the maximization of likelihood is to find the mixing parameters by maximizing the probability density of the observed sources $p_X(x)$.

Assuming that the source signals are i.i.d, and each observed signal consists of N samples, the likelihood is calculated by the following formula

$$L = \prod_{j=1}^{N} \left[|Det(B)| \prod_{i=1}^{L} p_i(s_i(j)) \right] \tag{1.9}$$

Where B is the inverse of the mixing matrix A, and $p_i(s_i)$ is probability density of the source s_i. The limitation of this kind of ICA methods is that, when using maximum likelihood as the objective function, we first need to know the probability density function of the source signals to compute the likelihood function, otherwise we cannot.

1.4.1.4 Higher-Order-Statistics (HOS):

Higher-Order-Statistics (HOS) (cumulants and moments of order greater than two) have been proposed as a useful tool to separate an i.i.d non-Gaussian signals with at most a single signal having a Gaussian distribution. The basic idea of these methods is to separate signals using a function based on higher order cumulants. Indeed, the cross cumulants between signals are zero when these signals are independent. In practice , it is difficult to determine the cross cumulants at all orders, therefore, the majority of BSS methods using HOS are limited usually to order four.

Concerning our dissertation we try to detail the basic notions of HOS with their properties in chapter 02, and we present a BSS method based on HOS as an application in chapter 03.

Chapter 2

Higher Order Statistics

2.1 Introduction

Higher Order Statistics (HOS), cumulants and moments of order greater than two are extensions of second order statistics measures. They are well known tools that can be used widely in the description of data and their statistical properties. HOS are implied in BSS area due to its capability to separate mixing data, and allow for resolving insoluble problems in order 2. In this chapter we try to give the necessary definitions and properties for the introduction of HOS.

2.2 Real scalar random variables

Definition 2.2.1. *Let x be a real valueds scalar random variable.*
The distribution function of a continuous random variable X can be expressed as the integral of its probability density function $p_x(u)$ as follows :

$$F_x(u) = \int_{-\infty}^{u} p_x(t)dt \tag{2.1}$$

Definition 2.2.2. *The generalized moments of x are defined for any real application g by:*

$$E[g(x)] = \int_{-\infty}^{\infty} g(u)p_x(u)\mathrm{d}u \tag{2.2}$$

We use polynomial functions $g(u)$, leading to moments different orders, we associate with random variables characteristic functions.

Definition 2.2.3. *The first characteristic function of x is :*

$$\Phi_x(v) = E[e^{jvx}] = \int_{-\infty}^{\infty} e^{jvu} p_x(u)\mathrm{d}u \tag{2.3}$$

where j denote the root of -1. When the random variable x admits a density of probability $p_x(u)$.

We find the density of probability from the first characteristic function by inverse Fourier transformation :

$$p_x(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-jvu}\Phi_x(v)\mathrm{d}v \tag{2.4}$$

Definition 2.2.4. *we define the second characteristic function as follows:*

$$\Psi_x(v) = \log\Big[\Phi_x(v)\Big] \tag{2.5}$$

2.2.1 Moments

Definition 2.2.5. *The rth-order moment $\mu_{x(r)}$ of random variable x is defined by:*

$$\mu_{x(r)} = E[x^r] = (-j)^r \frac{d^r\Phi_x(v)}{dv^r}\bigg|_{v=0} \tag{2.6}$$

and $\mu'_{x(r)}$ the ***centered moments****:*

$$\mu'_{x(r)} = E\Big[(x-\mu_{x_{(1)}})^r\Big] \tag{2.7}$$

2.2.2 Cumulants

Definition 2.2.6. *The derivatives of the second characteristic function define* ***cumulants****.*

$$Cum[x, x...x] = k_{x(r)} = (-j)^r \frac{d^r\Psi_x(v)}{dv^r}\bigg|_{v=0} \tag{2.8}$$

Which are the Taylor series development coefficients of the second characteristic function
We can calculate the cumulants of order r from moments of order lower or equal to r, for example (for r = 4):

$$
\begin{aligned}
k_{x(1)} &= \mu_{x(1)} \\
k_{x(2)} &= \mu'_{x(2)} = \mu_{x(2)} - \mu^2_{x(1)} \\
k_{x(3)} &= \mu'_{x(3)} = E[(x-\mu_{x(1)})^3] = \mu_{x(3)} - 3\mu_{x(2)}\mu_{x(1)} + 2\mu^3_{x(1)} \\
k_{x(4)} &= \mu'_{x(4)} = \mu_{x(4)} - 4\mu_{x(3)}\mu_{x(1)} - 3\mu^2_{x(2)} + 12\mu_{x(2)}\mu^2_{x(1)} - 6\mu^4_{x(1)}
\end{aligned}
$$

In the case of centered random variables($\mu_{x(1)} = 0$), the expressions of cumulants are simplified by:

$$
\begin{aligned}
k_{x(1)} &= 0 \\
k_{x(2)} &= \mu_{x(2)} = E[x^2] \\
k_{x(3)} &= \mu_{x(3)} = E[x^3] \\
k_{x(4)} &= E[x^4] - 3E[x^2]^2
\end{aligned}
$$

When the variable x is Gaussian, the second characteristic function is:

$$\Psi_x(v) = j\mu_{x(1)}v - \frac{1}{2}\mu_{x(2)}v^2 \tag{2.9}$$

Proof.

$$\Phi_x(v) = E[e^{jvu}] = \int_{-\infty}^{\infty} e^{jvu} p_x(u)\mathrm{d}u$$

$$\Phi_x(v) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{jvu}\exp\left(\frac{-1}{2\mu_{x(2)}}(u-\mu_{x(1)})^2\right)\mathrm{d}u$$

$$\Phi_x(v) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(jvu - \frac{1}{2\mu_{x(2)}}\left(u^2 + \mu^2_{x(1)} - 2u\mu_{x(1)}\right)\right)\mathrm{d}u$$

$$\Phi_x(v) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(jvu - \frac{u^2}{2\mu_{x(2)}} - \frac{\mu^2_{x(1)}}{2\mu_{x(2)}} + \frac{u\mu_{x(1)}}{\mu_{x(2)}} + \frac{\mu_{x(2)}v^2}{2} - \frac{\mu_{x(2)}v^2}{2} + j\mu_{x(1)}v - j\mu_{x(1)}v\right)\mathrm{d}u$$

$$\Phi_x(v) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(\frac{-1}{2\mu_{x(2)}}(-2\mu_{x(2)}jvu + u^2 + \mu^2_{x(1)} - 2u\mu_{x(1)} - \mu^2_{x(2)}v^2 + 2jv\mu_{x(2)}\mu_{x(1)}\right)$$

$$\exp\left(j\mu_{x(1)}v - \frac{\mu_{x(2)}v^2}{2}\right)\mathrm{d}u$$

$$\Phi_x(v) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(j\mu_{x(1)}v - \frac{\mu_{x(2)}v^2}{2}\right)\int_{-\infty}^{\infty}\frac{-1}{2\mu_{x(2)}}\left(u - (\mu_{x(1)} + jv\mu_{x(2)})\right)^2\mathrm{d}u$$

So

$$\Phi_x(v) = \exp\left(j\mu_{x(1)}v - \frac{\mu_{x(2)}v^2}{2}\right)$$

Hence

$$\Psi_x = j\mu_{x(1)}v - \frac{\mu_{x(2)}v^2}{2}$$

■

Remark. *In the Gaussian distribution case, the cumulants of order greater than 2 are therefore all zero [8].*

Definition 2.2.7. *The standardized random variable defined by:*

$$x^* = \frac{x - \mu_{x(1)}}{\sqrt{k_{x(2)}}} \tag{2.10}$$

Theorem 2.2.8. *The skewness of a random variable X is the third standardized cumulant, defined as:*

$$M_{x(3)} = \frac{E(x^3)}{E(x^2)^{\frac{3}{2}}}$$

Theorem 2.2.9. *The kurtosis(the factor flattening) is the fourth standardized cumulant,defined as :*

$$M_{x(4)} = \frac{E[x^4]}{E[x^2]^2} - 3$$

2.2.3 Examples of random variables

Let's take a look at some examples of random variables:

Uniform variable :

The distribution of a random variable in the part [-a, a] has the following characteristic function :

$$\begin{aligned}
\Phi_x(v) &= E[e^{jvx}] \\
&= \int_{-a}^{a} e^{jvx} p_x(u)\mathrm{d}x \\
&= \int_{-a}^{a} e^{jvx} \frac{1}{2a}\mathrm{d}x \\
&= \frac{1}{2a}\frac{1}{jv} e^{jvx}\Big|_{-a}^{a} \\
&= \frac{1}{2ajv}(e^{jva} - e^{-jva}) \\
&= \frac{1}{2ajv}(cos(va) + jsin(va) - cos(-va) - jsin(-va)) \\
&= \frac{1}{2ajv}(cos(va) + jsin(va) - cos(va) + jsin(va))
\end{aligned}$$

Hence

$$\Phi_x(v) = \frac{sin(va)}{av}$$

the rth-order moments is $\mu_{x(r)} = \frac{a^r}{r+1}$, and for kurtosis $k_{x(4)} = \frac{\mu_{x(4)}}{(\mu_{x(2)})^2} - 3 = \frac{-6}{5}$.

Gaussian variable :

The density of the centered Gaussian variable is :

$$p_x(u) = \tfrac{1}{\sigma\sqrt{2\pi}} exp(-\tfrac{u^2}{2\sigma^2})$$

The characteristic function is:

$$\begin{aligned}
\Phi_x(v) &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{jvu}\exp(\frac{-1}{2}\frac{u^2}{\sigma^2})\mathrm{d}u \\
&= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(jvu - \frac{1}{2}\frac{u^2}{\sigma^2})\mathrm{d}u \\
&= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(jvu - \frac{1}{2}\frac{u^2}{\sigma^2} + \frac{\sigma^2 v^2}{2} - \frac{\sigma^2 v^2}{2})\mathrm{d}u \\
&= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(\frac{-1}{2\sigma^2}(u - jv\sigma^2)^2)\exp(\frac{-\sigma^2 v^2}{2})\mathrm{d}u \\
&= \exp(\frac{-\sigma^2 v^2}{2})
\end{aligned}$$

Because

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(\frac{-1}{2\sigma^2}(u - jv\sigma^2)^2) = 1$$

The odd moments are zero and even moments are equal:

$$\mu_{x(2r)} = \sigma^{2r}\frac{(2r)!}{r!2^r}$$

generalized Gaussian variable :

the random variable x is called Gaussian generalized if its probability density is written as follows :

$$p_x(u) = B\exp(-A|u|^g)$$

g is a positive real number, and The coefficients A and B are introduced in order to normalize the sum of $p_x(.)$ to 1, and the variance to 1. These coefficients are :

$$A = \frac{\Gamma(3/g)}{\Gamma(1/g)} \qquad \text{and} \qquad B = g\frac{\Gamma(3/g)^{1/2}}{\Gamma(1/g)^{3/2}}$$

where $\Gamma(.)$ is the gamma function, $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$. The moments in the odd order are zero and the moments in the even order are :

$$\mu_{x(r)} = \frac{\Gamma(1/g)^{r-1}}{\Gamma(3/g)^r}\Gamma\left(\frac{2r+1}{g}\right) \qquad \forall g \geq 1$$

We find $\mu_{x(r)} = 1$, if $g = 2$, we are in the Gaussian case, and we find the results given above using the relationships $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(n+1) = n\Gamma(n)$. By tending g towards infinity, we find the uniform case mentioned above. if g = 1 we obtain a bilateral Laplace variable [9].

2.3 Multidimensional random variables

Multivariate random variables can be represented by the column vector

$$X^T = (x_1, x_2 ... x_N)$$

Definition 2.3.1. *We define the first characteristic function of N random variables x_n by the relation :*

$$\Phi_X(V) = E\left[e^{j\sum_n v_n x_n}\right] = E\left[e^{jV^T X}\right] \tag{2.11}$$

Where

$$V^T = (v_1, v_2 ..., v_N)$$

If the components x_n of the random vector x admit a joint density $p_x(u)$. The first characteristic function of x is given by the Fourier transform of this density :

$$\Phi_X(V) = \int_{R^N} e^{jV^T U} p_X(U) \mathrm{d}U \tag{2.12}$$

Definition 2.3.2. *the second characteristic function is :*

$$\Psi_x(V) = \log \Phi_x(V) \tag{2.13}$$

The characteristic functions can be used to generate moments and cumulants.

In the multidimensional random variables case, we take examples of cumulants and write them as follows :

Second order cumulant

the second order cumulants, can be stored in a matrix (covariance matrix)

$$k_{X(2)ij} = Cum[x_i, x_j]$$

Third order cumulant

$$k_{X_{ijk}} = Cum[x_i, x_j, x_k] = k_{X(3)}$$
$$k_{X_{iii}} = Cum[x_i, x_i, x_i] = k_{x_i(3)}$$

Fourth order cumulant

$$k_{X_{hijk}} = Cum[x_h, x_i, x_j, x_k] = k_{X(4)}$$
$$k_{X_{iiii}} = Cum[x_i, x_i, x_i, x_i] = k_{x_i(4)}$$

Definition 2.3.3. *The cross moments define as follows :*

$$\mu_{X i_1 i_2 \ldots i_r} = E[x_{i_1} \ldots x_{i_r}] = (-j)^r \frac{\partial^r \Phi_x(V)}{\partial v_{i_1} \partial v_{i_2} \ldots \partial v_{i_r}}\bigg|_{V=0} \tag{2.14}$$

With

$$r = i_1 + i_2 + \ldots i_r$$

Remark. *By developing the exponential function $[e^{jV^T X}]$ into a series around $V = 0$. The cross moments is the coefficients terms of degree r : $[j^r \mu_{x^{ij\ldots k}}/r!]$, As shown in [17].*

Definition 2.3.4. *The cross cumulants define as follows :*

$$k_{X_{i_1 i_2 \ldots i_r}} = (-j)^r \frac{\partial^r \Psi_x(V)}{\partial v_{i_1} \partial v_{i_2} \ldots \partial v_{i_r}}\bigg|_{V=0} \tag{2.15}$$

As in the scalar case, and by developing the logarithm function in series, we can write a relation between moments and cumulants.
For example the second-order cumulant is :

$$k_{X_{ij}} = \mu_{X_{ij}} - \mu_{X_i} \mu_{X_j}$$

Notation 2.3.5. *The integer number that appears in bracket [m] is the number of characteristic monomers that can be obtained by permutation. In short, we take the examples as follows:*

$$[3]\delta_{ij}\delta_{kl} = \delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}$$
$$[3]a_{ij} b_k c_{ijk} = a_{ik} b_j c_{ijk} + a_{ij} b_k c_{ijk} + a_{jk} b_i c_{ijk}$$
$$[3]x_i \delta_{jk} = x_i \delta_{jk} + x_j \delta_{ik} + x_k \delta_{ij}$$
$$[6]x_i x_j \delta_{kl} = x_i x_j \delta_{kl} + x_k x_j \delta_{il} + x_l x_j \delta_{ki} + x_i x_k \delta_{jl} + x_i x_l \delta_{jl} + x_k x_l \delta_{ij}$$

In the non-centered case, the third-order and fourth-order cumulants by the moments function are given as follows:

$$k_{X_{ijk}} = \mu_{X_{ijk}} - [3]\mu_{X_i} \mu_{X_{jk}} + 2\mu_{X_i} \mu_{X_j} \mu_{X_k} \tag{2.16}$$

$$k_{X_{ijkl}} = \mu_{X_{ijkl}} - [4]\mu_{X_i} \mu_{X_{jkl}} - [3]\mu_{X_{ij}} \mu_{X_{kl}} + 2[6]\mu_{X_i} \mu_{X_j} \mu_{X_{kl}} - 6\mu_{X_i} \mu_{X_j} \mu_{X_k} \mu_{X_l} \tag{2.17}$$

In the centered case, these expressions are simplified by:

$$k_{x_{ij}} = \mu_{x_{ij}}$$
$$k_{x_{ijk}} = \mu_{x_{ijk}}$$
$$k_{x_{ijkl}} = \mu_{x_{ijkl}} - [3]\mu_{x_{ij}}\mu_{x_{kl}}$$

In the scalar case, we replace the above notation $n[m]$ by nm.

Definition 2.3.6. *[Leonov and Shiryayev]*
From the definition of the second characteristic function, the cumulants are related to moments by the formula called ***Leonov and Shiryayev****, and it writes as follows:*

$$Cum[x_1...x_r] = \sum(-1)^{k-1}(k-1)!E\left[\prod_{i\in v_1} x_i\right]E\left[\prod_{j\in v_2} x_j\right]...E\left[\prod_{k\in v_p} x_k\right] \tag{2.18}$$

In the second-order , the possible partitions are (1,2) and (1)(2), so we find:

$$Cum[x_1,x_2] = (-1)^0 0!E[x_1x_2] + (-1)^{2-1}1!E[x_1]E[x_2]$$

In the third-order, the possible partitions are (1, 2, 3), (1)(2, 3), et (1)(2)(3).

Remark. *There are* 3 *partitions of type* $(1)(2,3)$*: Note that there are three partitions of type (1) (2 , 3): We write the number of this partitions in bracket as we indicated in the above notation:*

(1) (2, 3); (2) (1, 3); (3) (1, 2)

The third-order cross cumulant is :

$$Cum[x_1,x_2,x_3]=(-1)^0 0!E[x_1x_2x_3] + (-1)^{2-1}1![3]E[x_1]E[x_2x_3] + (-1)^{3-1}2!E[x_1]E[x_2]E[x_3]$$

In the fourth-order , the partitions are:

- $(1,2,3,4)$ au nombre de 1,$k-1$=0($k = 1 partition$).

- $(1)(2,3,4)$ number of 4,$k-1$=1($k=2partitions$).

- $(1)(2)(3,4)$ number of 6,$k-1$=2($k=3partitions$).

- $(1)(2)(3)(4)$ number of 1,$k-1$=3($k=4partitions$).

- $(1,2)(3,4)$ number of 3,$k-1$=1($k=2partitions$).

The fourth-order cross cumulant is:

$$Cum[x_1,x_2,x_3,x_4] = (-1)^0 0!E[x_1x_2x_3,x_4] + (-1)^1 1![4]E[x_1]E[x_2x_3x_4] +$$
$$(-1)^2 2![6]E[x_1]E[x_2]E[x_3x_4] + (-1)^1 1![3]E[x_1x_2]E[x_3x_4] + (-1)^3 3!E[x_1]E[x_2]E[x_3]E[x_4]$$

Remark. *We can find the moments by the inverse formula of* ***Leonov et Shiryayev*** *:*

$$E[x_1...x_r]{=}\sum cum[x_i, i \in v_1].cum[x_j, j \in v_2]....cum[x_k, k \in v_p]$$

We find The fourth-order moment as follows :

$$\mu_{X_{ijkl}} = k_{X_{ijkl}} + [4]k_{X_i}k_{X_{jkl}} + [3]k_{X_{ij}}k_{x_{kl}} + [6]k_{X_i}k_{X_j}k_{X_{kl}} + k_{X_i}k_{X_j}k_{X_k}k_{X_l}$$

2.4 Random variables with complex values

Definition 2.4.1. *z is the random variable with complex values, and it represente by a real random variable of* 2*-dimension as :*

$$z = x + iy$$

With

$$x, y \in \boldsymbol{R}^N \qquad and \qquad j^2 = -1$$

This complex random variable z has a density if and only if its real and imaginary parts admit a joint density .

Definition 2.4.2. *The characteristic function of the complex vector variable z is :*

$$\Phi_z(u) = E\left[e^{j[X^T V + y^T w]}\right] = E\left[e^{j\Re[z^\dagger u]}\right] \tag{2.19}$$

With

$$u = v + jw$$

Definition 2.4.3. *We note that any function in x and y can be represented as a function From z and z^*, the first characteristic function of z is:*

$$\Phi_{z,z^*}(u,u^*) = E\left[e^{\dfrac{j[z^\dagger u + u^\dagger z]}{2}}\right]$$

Definition 2.4.4. *The second characteristic function of z is :*

$$\Psi_{z,z*}(u,u*) = \log(\Phi_{z,z^*}(u,u^*)) \tag{2.20}$$

We can find the relation of moments and cross cumulants between a variable z and its conjugate z^*.

Definition 2.4.5. *The moments for the complex random variable z are written as follows:*

$$\mu_{z(p)}^{(q)} = E[z^p z^{*q}] \tag{2.21}$$

Definition 2.4.6. *The cumulants for the complex random variable z define as :*

$$k_{z(p)}^{q} = Cum[z,...,z;z^*,...,z^*] = (-2j)^r \left.\frac{\partial^r \Psi_{z,z*}(u,u^*)}{\partial u^q \partial u^{*p}}\right|_{u=0} \tag{2.22}$$

Definition 2.4.7. *The complex random vector z and his components z_i,the moments are:*

$$\mu_{z_{i_1 \dots i_p}}^{j_1 \dots j_q} = E[z_{i_1} ... z_{i_p}, z_{j_1}^* ... z_{j_q}^*] \tag{2.23}$$

And his cumulants as follows :

$$k_{z_{i_1 \dots i_p}}^{j_1 \dots j_q} = Cum[z_{i_1},...,z_{i_p},z_{j_1}^*,...,z_{j_q}^*] \tag{2.24}$$

2.4.1 Standardization

Definition 2.4.8. *The Standardization is the affine transformation that is associated with x a centered random vector, its covariance matrix is identity I. The standardization random vector is:*

$$x^* = W(x - \mu_{x(1)}) \tag{2.25}$$

Where W is the matrix verifying $WCW^{\dagger}=I$, $\mu_{x(1)}$ *is the random vector mean,* $C = k_{X^{ij}}$ *for real value and* $C = k_{x_i}^{j}$ *for complex values.*
we take :

$$W{=}\Lambda^{\frac{-1}{2}}U^{\dagger}$$

where Λ *is the diagonal matrix of eigenvalues of C and U is the unit matrix.*

2.4.2 Circularity

Definition 2.4.9. *A complex random vector Z of dimension N is circular, if and only if:*

$$\forall\theta : \qquad Z \text{ and } Ze^{j\theta} \text{ have the same statistical properties}$$

We will take the Necessary proprietes of circularity, as shown in [9]

1. If a complex random variable z is circular, then :

$$\Phi_Z(e^{j\theta}u) = \Phi_Z(u), \forall\theta \tag{2.26}$$

2. Z complex random vector, Z is circular if and only if all its moments of the form:

$$\mu_{Z(p)}^{(q)} = E\left[\prod_{\sum a_i=p} z_i^{a_i} \prod_{\sum b_j=q} z_j^{*b_j}\right] \tag{2.27}$$

It Equals zero when $p \neq q$ and $p + q \leq r$, and :

$$\mu_{Z(p)}^{(q)} = \mu_{Ze^{j\alpha}(p)}^{(q)} = \mu_{Z(p)}^{(q)} e^{j\alpha(p-q)} \tag{2.28}$$

The above properties of circular complex random variable prove that :

$$E[Z] = 0; E[Z^2] = 0; E[Z^2 Z^*] = 0$$

3. Z is a circular variable with order r if its density of probability is invariant by rotation of angle $2\pi/r+1$, and if Z and $Ze^{2\pi j/r+1}$ have the same statistical properties.

In Gaussian case, the second order circularity leads to circularity of all orders. if z is circular, then : $E[ZZ^T] = 0$ implies that $[xx^T - yy^T] = 0$ and $E[xy^T + yx^T] = 0$.

2.5 Properties of moments and cumulants

2.5.1 Multilinearity

Moments and cumulants satisfy the property of multilinearity.

Definition 2.5.1. *If $y = Ax$ and $A = \{A_{ij}\} \in R^{mn}$, the property of miltilinearity implies that the moments of y are function linear moments of x and the same for cumulants.*

For example, we have:

$$Cum[y_i, y_j, y_k^*] = \sum_{a,b,c} A_{ia} A_{jb} A_{kc}^* Cum[x_a, x_b, x_c]$$

In the scalar case, we write :

$$k_{\lambda x(r)} = \lambda^r k_{x(r)} \tag{2.29}$$

The general case of multilinearity is :

$$\begin{gathered} Cum[x+y, z_1, ..., z_p] = Cum[x, z_1, ..., z_p] + Cum[y, z_1, ..., z_p] \quad \forall x, y, z_1, ..., z_p \\ Cum[\lambda x, z_1, ..., z_p] = \lambda Cum[x, z_1, ..., z_p] \end{gathered}$$

2.5.2 Translation invariance

The Cumulants are deterministic translation invariant. if y=$x+t$ (t is deterministic), then the order cumulants greater or equal than 2 of y are identical with cumulants of x. The second characteristic function of y and x are linked by:

$$\Psi_y(v) = j\Re\mathfrak{e}[t^\dagger v] + \Psi_x(v) \tag{2.30}$$

2.5.3 Independent random variables

Let x and y be two independent random vectors, with reals values or complex, of respective dimensions n and p, and either $Z^T = x^T y^T$
Hence

$$\Phi_Z = \Phi_X \Phi_Y$$
$$\Psi_z(u,v) = \Psi_x(u) + \Psi_y(v)$$

Remark. *The cross cumulants between x and y are :*

$$k_{Z_{i_1 \dots i_n, j_1 \dots j_p}} = Cum[x_{i_1} \dots x_{i_n}, y_{j_1} \dots y_{j_p}] = (-j)^r \frac{\partial^r \Psi_Z(U,V)}{\partial u_{i_1} \dots \partial u_{i_n} \partial v_{j_1} \dots \partial v_{j_p}} \Bigg|_{V=0} \tag{2.31}$$

Where

$$r = n + p$$

The cross cumulants of x and y are zero if one of the i_k and one of the j_k are non-zero simultaneously.
If x and y a independents vectors, for all z, we find :

$$Cum[x,y,z] = 0 \tag{2.32}$$

2.6 Higher order statistics and probability density.

In this paragraph, we examine the link between higher order statistics and probability density.

2.6.1 Tendency towards gaussianity

We consider N independent random variables, a bounded cumulant of order r is denoted by $k_{(r)}(n)$. we put :

$$\overline{k_{(r)}} = \frac{1}{N} \sum_{n=1}^{N} k_{x(r)}(n) \qquad \text{and} \qquad y = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} (x(n) - \overline{k_{(1)}})$$

The random variable y is the normalized sum of N random variables independent. When $N \longrightarrow \infty$ the random variable y tends in law towards a Gaussian random variable y.

$\lambda_{(r)}$ is the cumulants of the random variable y, and are define by:

$$\lambda_{(r)} = k_{y(r)} = \frac{1}{N^{r/2-1}}\overline{k_{(r)}} \quad \forall r \geq 2 \tag{2.33}$$

2.6.2 Gaussianity and independence criteria

This part is based on the concept of independent statistics, Entropy, Kullback divergence and mutual information allow the introduction of criteria of gaussianity and independence.

2.6.2.1 Entropy and gaussianity

The observation of a random density vector x of probability $p_x(u)$ provides a quantity of information quantified by entropy:

$$S(p_x) = -\int p_x(u) \log p_x(u) du \tag{2.34}$$

To encrypt the Entropy deficit, compared to the random vector Gaussian x_g, of a random vector x belonging to the set we introduce the negoentropy.

$$J(p_X) = S(p_{X_g}) - S(p_X) \tag{2.35}$$

2.6.2.2 Kullback divergence

The Kullback divergence measures the distance between two densities of probability, $p_V(u)$ and $p_W(u)$, by:

$$k(p_V, p_W) = \int p_V(u) \log \frac{p_V(u)}{p_W}(u) du$$

The divergence of Kullback $k(p_V, p_W)$ is negative if $p_V \neq p_W$, it is zero if $p_V = p_W$.

Proof. (demonstration of Neguentropy)

we have :

$$\log\ w \leq w - 1$$

If $(p_V = p_W)$

$$k(p_V, p_W) = \int p_V(u)\log\frac{p_W(u)}{p_V(u)}du \leq \int p_V(u)(\frac{p_W(u)}{p_V(u)} - 1)du = 0$$

And we have

$$S(p_{X_g}) = -\int p_{X_g}(u)\log p_{X_g}(u)du = -\int p_X(u)\log p_{X_g}(u)du$$

We find that the divergence of Kullback

$$\begin{aligned}
J(p_X) &= S(p_{X_g}) - S(p_X)\\
&= -\int p_X(u)\log p_{X_g}(u)du + \int p_X(u)\log p_X(u)du\\
&= \int p_X(u)\log\frac{p_X(u)}{p_{X_g}(u)}du\\
&= k(p_X, p_{X_g})
\end{aligned}$$

■

Let be a random vector x of probability density $p_X(u)$. If this vector is made up of independent variables its density of probability is the product of the marginal probability densities of each of its components $\prod_i p_{x_i}(u_i)$.
The Kullback divergence between $p_X(u)$ and $\prod_i p_{x_i}(u_i)$ gives a measure of the statistical independence of the components of x that are call: mutual information.

$$I(p_X) = k(p_X, \prod_i p_{x_i}(u_i)) = \int p_X(u)\log\frac{p_X(u)}{\prod_i p_{x_i}(u_i)}du \tag{2.36}$$

The random vector x and the Gaussian random vector x_g have the same first and second order moments, we can write :

$$J(p_X) - \sum_i J(p_{x_i}) = I(p_X) - I(p_{X_g})$$

Proof.

$$J(p_X) - \sum_i J(p_{x_i}) = \int p_X(\mathrm{u})\log\frac{p_X(u)}{p_{X_g}(u)}du - \sum_i \int p_{x_i}(u_i)\log\frac{p_{x_i}(u_i)}{p_{x_{g_i}}(u_i)}du_i$$

We have

$$\sum_i \int p_{x_i}(u_i) \log p_{x_i}(u_i) du_i = \int p_X(u) \log \prod_i p_{x_i}(u_i) du$$

$$\begin{aligned}
J(p_X) - \sum_i J(p_{x_i}) &= \int p_X(u) \log \frac{p_X(u)}{p_{X_g}(u)} du - \int p_X(u) \log \frac{\prod_i p_{x_i}(u_i)}{\prod_i p_{x_{g_i}}(u_i)} du \\
&= \int p_X(u) \left[\log \frac{p_X(u)}{p_{X_g}(u)} - \log \frac{\prod_i p_{x_i}(u_i)}{\prod_i p_{x_{g_i}}(u_i)} \right] du \\
&= \int p_X(u) \left[\log p_X(u) - \log p_{X_g}(u) - \log \prod_i p_{x_i}(u_i) + \prod_i p_{x_{g_i}}(u_i) \right] du \\
&= \int p_X(u) \left[\log p_X(u) - \log \prod_i p_{x_i}(u_i) \right] - \left[\log p_{X_g}(u) - \log \prod_i p_{x_{g_i}}(u_i) \right] du \\
&= \int p_X(u) \left[\log \frac{p_X(u)}{\prod_i p_{x_i}(u_i)} - \log \frac{p_{X_g}(u)}{\prod_i p_{x_{g_i}}(u_i)} \right] du \\
&= I(p_X) - I(p_{X_g})
\end{aligned}$$

■

So :

$$I(p_X) = J(p_X) - \sum_i J(p_{x_i}) + I(p_{X_g})$$

2.7 Estimation of moments and cumulants

If you use higher order statistic, you must pass by their estimate. This paragraph presents some elements on estimation, in the scalar case.

2.7.1 moment estimators

Let x be a centered scalar random variable, and $x_n, 1 \leq n \leq N, N$ realisations of x.
The classic estimator of the rth order moment of X is given by:

$$\widehat{\mu_{(r)}} = \frac{1}{N} \sum_{n=1}^{N} x_n^r \qquad (2.37)$$

This estimator is unbiased [i.e $E[\widehat{\mu_{(r)}}] = \mu_{(r)}$], In addition, if x_n is realizations independent of x, the variance estimate is:

$$Var[\widehat{\mu_{(r)}}] = \frac{1}{N}\text{Var}[x^r]$$

Note that the moment estimator is a consistent estimator when:

$$\lim_{N \longrightarrow \infty} Var[\widehat{\mu_{(r)}}] = 0$$

2.7.2 Cumulants estimators

A cumulants estimator is obtained by substituting the moments in Leonov and Shiryayev formula by their estimators.

Example of fourth-order cumulants

The fourth order cumulants are written by moments, as follows :

$$k_{x(4)} = \mu_{x(4)} - 3\mu_{x(2)}^2 \tag{2.38}$$

So, the fourth order cumulant estimator is:

$$\widehat{k_{x(4)}} = \widehat{\mu_{x(4)}} - 3\widehat{\mu_{x(2)}}^2 \tag{2.39}$$

$$= \frac{1}{N}\sum_{i=1}^{N} x_i^4 - 3\left(\frac{1}{N^2}\sum_{i,j=1}^{N} x_i^2 x_j^2\right) \tag{2.40}$$

$$= \frac{1}{N}\sum_{i=1}^{N} x_i^4 - \frac{3}{N^2}\left(\sum_{i=j=1}^{N} x_i^2 x_j^2 + \sum_{i=1}^{N} x_i^2 \sum_{j=1,i\neq j}^{N} x_j^2\right) \tag{2.41}$$

Study of bias :

From the above Eq(2.40), we put :

$$f(N)=\frac{1}{N} \quad \text{and} \quad g(N)=\frac{-3}{N^2}$$

We study the bias of fourth order cumulants:

$$E[\widehat{k_{x(4)}}] = E\left[\frac{1}{N}\sum_{i=1}^{N} x_i^4 - \frac{3}{N^2}\left(\sum_{i=j=1}^{N} x_i^2 x_j^2 + \sum_{i=1}^{N} x_i^2 \sum_{j=1, i\neq j}^{N} x_j^2\right)\right] \tag{2.42}$$

$$= \mu_{x(4)} - \frac{3}{N}\mu_{x(4)} - \frac{3}{N^2}\left(N(N-1)(\mu_{x(2)})^2\right) \tag{2.43}$$

$$= \mu_{x(4)} - \frac{3}{N}\mu_{x(4)} - 3\mu_{x(2)}^2 + \frac{3}{N}\mu_{x(2)}^2 \tag{2.44}$$

$$= (\mu_{x(4)} - 3\mu_{x(2)}^2) - \frac{3}{N}(\mu_{x(4)} - \mu_{x(2)}^2) \tag{2.45}$$

$$= k_{x(4)} - \frac{3}{N}(k_{x(4)} + 3\mu_{x(2)}^2 - \mu_{x(2)}^2) \tag{2.46}$$

$$= k_{x(4)} - \frac{3}{N}(k_{x(4)} + 2\mu_{x(2)}^2) \tag{2.47}$$

From the Eq(2.43), we can write the bias of this form :

$$E[\widehat{k_{x(4)}}] = (Nf(N) + Ng(N))\mu_{x(4)} + N(N-1)g(N)\mu_{x(2)}^2 \tag{2.48}$$

The bias is zero if :

$$\begin{cases} Nf(N) + Ng(N) = 1 \\ N(N-1)g(N) = -3 \end{cases}$$

Then :

$$\begin{cases} f(N) = \dfrac{N+2}{N(N-1)} \\ g(N) = \dfrac{-3}{N(N-1)} \end{cases}$$

The fourth k-statistic is :

$$\widehat{k_{x(4)}} = \frac{N+2}{N(N-1)}\sum_{i=1}^{N} x_i^4 - \frac{3}{N(N-1)}\sum_{i,j=1}^{N} x_i^2 x_j^2$$

Variance of estimator[17]

The k-statistic variance to $1/N$ order is:

$$Var[\widehat{k_{x(4)}}] = \frac{1}{N}(k_{x(8)} + 16k_{x(6)}k_{x(2)} + 48k_{x(5)}k_{(3)} + 34k_{x(4)}^2 + 72k_{x(4)}k_{(2)}^2 + 144k_{x(3)}^2 k_{x(2)} + 24k_{x(2)}^4)$$

The cumulants estimators are consistants since the estimator variance tends to 0 when the number of N tends to infinity.

Tendency towards Gaussianity [18]

The K-order cumulant ($k_{x(4)}$) is asymptotically normal. Then, the cumulants of order greater than or equal to 3 is zero.

2.7.3 Estimation of asymmetry and kurtosis

For the standardized random variable define :

- The skewness $M_{x(3)}$ is estimated by the following quantity:

$$\widehat{M_{x(3)}} = \frac{\widehat{k}_{x(3)}}{\widehat{k}_{x(2)}^{3/2}} \tag{2.49}$$

- The kurtosis $M_{x(4)}$ is estimated by the following quantity:

$$\widehat{M_{x(4)}} = \frac{\widehat{k}_{x(4)}}{(\widehat{k}_{x(2)})^2} \tag{2.50}$$

- In general, the standardized estimators defined as:

$$\widehat{M}_{x(r)} = \frac{\widehat{k}_{x(r)}}{\widehat{k}_{x(2)}^{r/2}} \tag{2.51}$$

There are exact results in the Gaussian case, shown in [19]:

$$\begin{aligned}
E[\widehat{k_{x(3)}}] &= 0 \\
E[\widehat{k_{x(4)}}] &= 0 \\
Var[\widehat{k_{x(3)}}] &= \frac{6N(N-1)}{(N-2)(N+1)(N+3)} \approx \frac{6}{N} \\
Var[\widehat{k_{x(4)}}] &= \frac{24N(N-1)^2}{(N-3)(N-2)(N+3)(N+5)} \approx \frac{24}{N}
\end{aligned}$$

Chapter 3

Blind source separation using an algebraic method

3.1 Blind separation of sources using a new polynomial equation

In this chapter, we present a simple algebraic method for estimating the mixing matrix in the two source separation problem between two sources separation, this problem proposed in [16].By using fourth-degree cumulants, then equating them with zero, and we solve a second-degree polynomial equation.
We assume that the sources are a zero mean, non-Gaussian, and statistically independent.

Model of mixtures:

Let $x_1(n)$ and $x_2(n)$ be the unknown sources. with help of two sensors, we observe two instantaneous linear mixtures $y_1(n)$ and $y_2(n)$ of the two zero-mean sources $x_1(n)$ and $x_2(n)$.
By defining $H = (h_{ij})$ the mixture matrix, we have :

$$\begin{bmatrix} y_1(n) \\ y_2(n) \end{bmatrix} = \begin{bmatrix} 1 & h_{12} \\ h_{21} & 1 \end{bmatrix} \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix}$$

Where h_{12} and h_{21} are unknown. The diagonal of the matrix $\mathbf{H}$ has been set to one.

Our objective is to finding the sources by a linear combination of the signals $y_1(n)$ and $y_2(n)$. We define $G = (g_{ij})$ as the separation matrix, we obtain the outputs $\widehat{x_1}(n)$ and $\widehat{x_2}(n)$:

$$\begin{aligned}
\begin{bmatrix} \widehat{x_1}(n) \\ \widehat{x_2}(n) \end{bmatrix} &= \begin{bmatrix} 1 & g_{12} \\ g_{21} & 1 \end{bmatrix} \begin{bmatrix} y_1(n) \\ y_2(n) \end{bmatrix} \\
&= \begin{bmatrix} 1 & g_{12} \\ g_{21} & 1 \end{bmatrix} \begin{bmatrix} 1 & h_{12} \\ h_{21} & 1 \end{bmatrix} \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} \\
&= \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix}
\end{aligned}$$

If H is regular, that is: $1 - h_{12}h_{21} \neq 0$, the separation is achieved by estimatiing a matrix G such that $GH = PD$, where P is a permutation matrix and D a diagonal matrix.
The above form of **G** leads to the two separation solutions :

$$\begin{aligned}
& g_{12} = -h_{12} \quad g_{21} = -h_{21} \\
or \quad & g_{12} = -1/h_{21} \quad g_{21} = -1/h_{12}
\end{aligned}$$

Now, we begin to compute the fourth order cross cumulants. And as explained earlier, the sources must be zero-mean stationary, non-Gaussian and statistically independent
First, let us denote :

$$Mom_{kl}(y_1, y_2) = E[y_1^k(n) y_2^l(n)] \tag{3.1}$$

$$cum_{kl}(y_1(n), y_2(n)) = cum(y_1^k(n) y_2^l(n)) = c_{kl} \tag{3.2}$$

$$p_i = E[x_i^2(n)] \tag{3.3}$$

$$\gamma_i = E[x_i^4(n)] \tag{3.4}$$

$$\beta_i = cum(x_i^4(n)) = E[(x_i^4(n)] - 3[E(x_i^2(n)]^2 \tag{3.5}$$

From [14], we've got:

$$cum_{13}(y_1(n), y_2(n)) = Mom_{13}(y_1, y_2) - 3Mom_{20}(y_1(n), y_2(n)) Mom_{11}(y_1(n), y_2(n)) \tag{3.6}$$

$$cum_{31}(y_1(n),y_2(n)) = Mom_{31}(y_1,y_2) - 3Mom_{02}(y_1(n),y_2(n))Mom_{11}(y_1(n),y_2(n)) \quad (3.7)$$

$$cum_{22}(y_1(n),y_2(n)) = Mom_{22}(y_1,y_2) - Mom_{02}(y_1(n),y_2(n))Mom_{20}(y_1(n),y_2(n)) - 2Mom_{11}^2(y_1(n),y_2(n)) \quad (3.8)$$

We calculate and substitute Eqns [(3.1), (3.2), (3.3) ,(3.4), (3.5)] in Eqns [(3.6) , (3.7), (3.8)], we find this equations [10] :

$$c_{31} = h_{21}\beta_1 + h_{12}^3\beta_2 \quad (3.9)$$

$$c_{13} = h_{21}^3\beta_1 + h_{12}\beta_2 \quad (3.10)$$

$$c_{22} = h_{21}^2\beta_1 + h_{12}^2\beta_2 \quad (3.11)$$

$$c_{40} = \beta_1 + h_{12}^4\beta_2 \quad (3.12)$$

$$c_{04} = h_{21}^4\beta_1 + \beta_2 \quad (3.13)$$

New separation solution:

We obtain the cross cumulants of the outputs $\widehat{x_1}(n)$ and $\widehat{x_2}(n)$ in the same method, shown in [11].

$$cum_{31}(\widehat{x_1}(n),\widehat{x_2}(n)) = k_{11}^3k_{21}\beta_1 + k_{12}^3k_{22}\beta_2 \quad (3.14)$$

$$cum_{13}(\widehat{x_1}(n),\widehat{x_2}(n)) = k_{11}k_{21}^3\beta_1 + k_{12}k_{22}^3\beta_2 \quad (3.15)$$

$$cum_{22}(\widehat{x_1}(n),\widehat{x_2}(n)) = k_{11}^2k_{21}^2\beta_1 + k_{12}^2k_{22}^2\beta_2 \quad (3.16)$$

We have:

$$k_{ii} = 1 + h_{ji}g_{ij} \quad (3.17)$$

$$k_{ij} = h_{ij} + g_{ij} \quad (3.18)$$

We substitute Eqns[(3.17)-(3.18)] in Eqns [(3.14)-(3.15)], we find :

$$cum(\widehat{x_1}^3(n),\widehat{x_2}(n)) = c_{31} + g_{21}c_{40} + 3g_{12}[c_{22} + g_{21}c_{31}] + 3g_{12}^2[c_{13} + g_{21}c_{22}] + g_{12}^3[c_{04} + g_{21}c_{13}] \quad (3.19)$$

and

$$cum(\widehat{x_1}(n),\widehat{x_2}^3(n)) = c_{13} + g_{12}c_{04} + 3g_{21}[c_{22} + g_{12}c_{13}] + 3g_{21}^2[c_{31} + g_{12}c_{22}] + g_{21}^3[c_{40} + g_{12}c_{31}] \quad (3.20)$$

We calculate the second partial derivatives of equations (3.19) and (3.20) and set it equal to zero :

$$\frac{\partial^2 cum(\widehat{x}_1^3(n), \widehat{x}_2(n))}{\partial^2 g_{12}} = 6[c_{13} + g_{21}c_{22}] + 6g_{12}[c_{04} + g_{21}c_{13}] = 0 \tag{3.21}$$

$$\frac{\partial^2 cum(\widehat{x}_1(n), \widehat{x}_2^3(n))}{\partial^2 g_{12}} = 6[c_{31} + g_{12}c_{22}] + 6g_{21}[c_{40} + g_{12}c_{31}] = 0 \tag{3.22}$$

From Eqn (3.22), we find :

$$g_{21} = -\frac{c_{31} + c_{22}g_{12}}{c_{40} + g_{12}c_{31}} \tag{3.23}$$

We substitute the Eqn (3.23) into Eqn (3.21), we obtain the following second-degree polynomial equation of the variable g_{12}:

$$c_{13}c_{40} - c_{22}c_{31} + [c_{40}c_{04} - c_{22}^2]g_{12} + [c_{31}c_{04} - c_{13}c_{22}]g_{12}^2 = 0 \tag{3.24}$$

We replace c_{ij} by their estimates, and we calculate the discriminant of this second-degree polynomial equation $\Delta \geq 0$, we get the roots g_{12a} and g_{12b} as follows:

$$\begin{cases} g_{12a} = -h_{12} \\ g_{12b} = -1/h_{21} \end{cases}$$

3.2 Experiments and results

First experiment :

In this experiment we try to mixing two audio signals and we apply this method to separate.The mixing matrix in this case is $H = \begin{bmatrix} 1 & 0.7 \\ 0.9 & 1 \end{bmatrix}$. we obtain the following visual result :

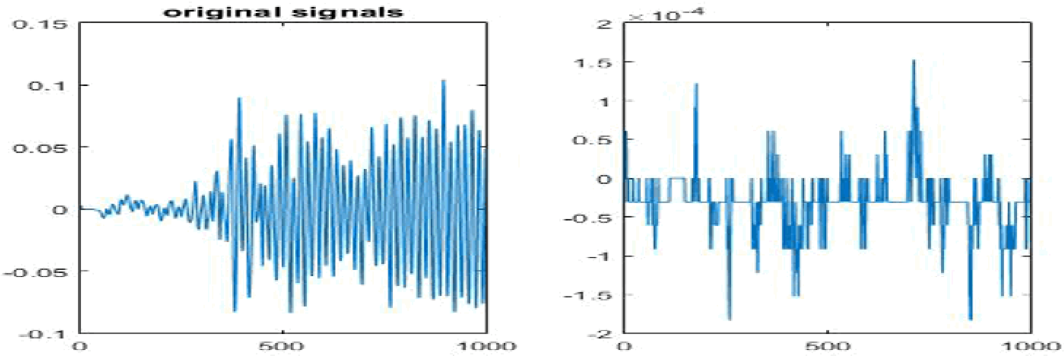

Figure 3.1: Original signals

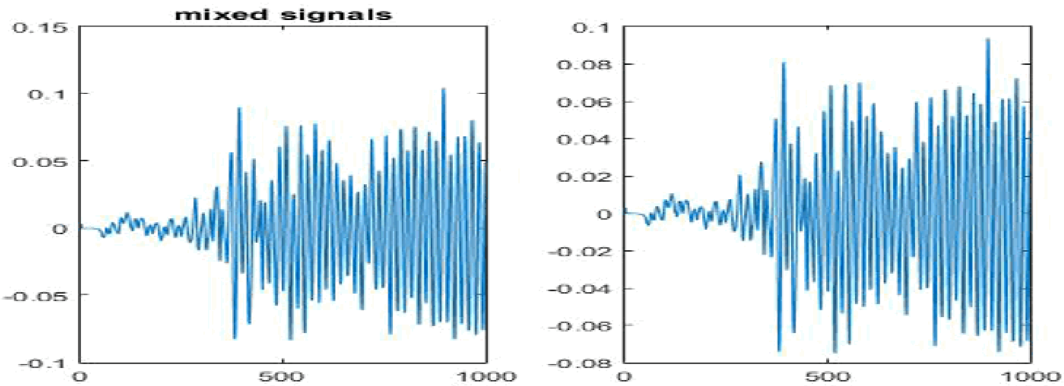

Figure 3.2: Mixed signals

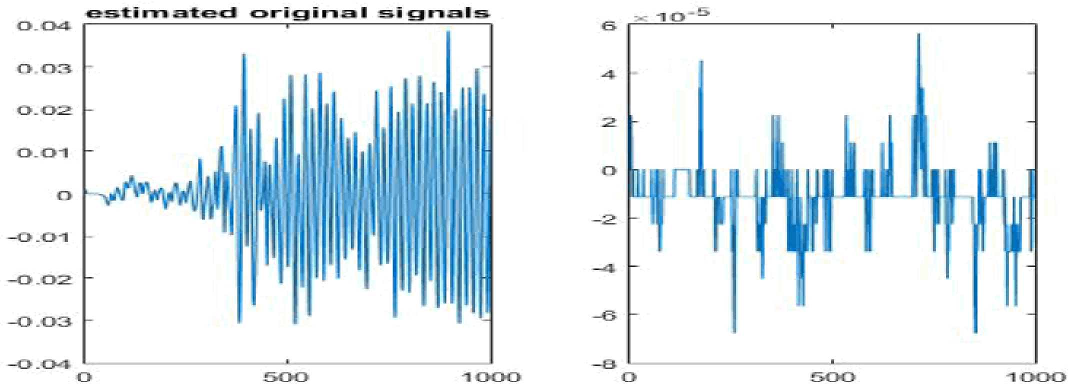

Figure 3.3: Estimated original signals

Second experiment :

In this experiment we mix an ECG signal (100.dat from MIT-BIH) of 1000 samples with a white Gaussian noise. In this case we choose the variance of the noise = 0.5, and the mixing matrix is $H = \begin{bmatrix} 1 & 10 \\ 0.02 & 1 \end{bmatrix}$.

We obtain the following result :

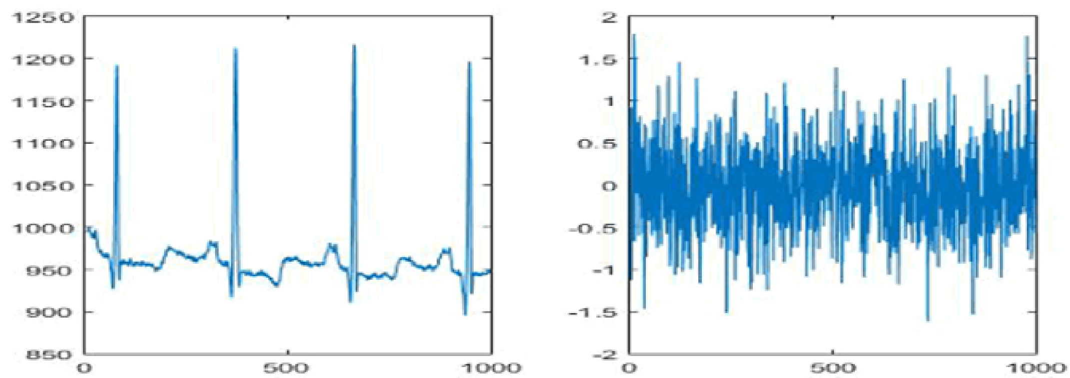

Figure 3.4: Original signals

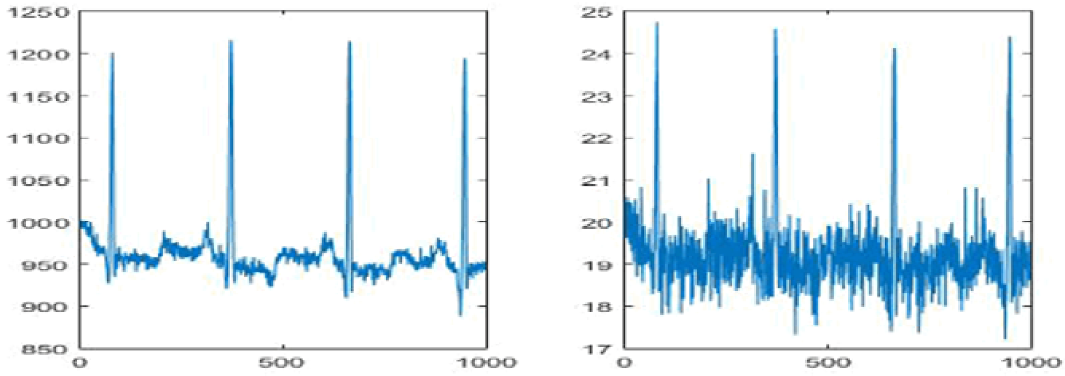

Figure 3.5: Mixed signals

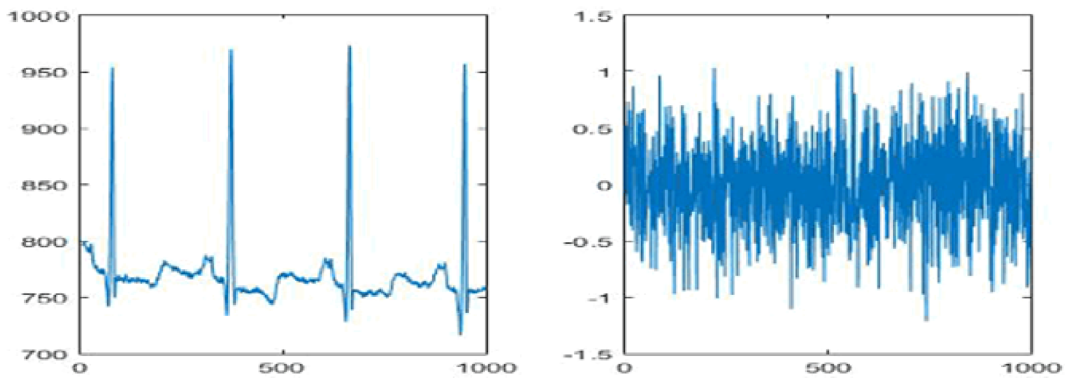

Figure 3.6: Estimated original signals

Third experiment :

In this final experiment, we mix two images with mixing matrix $H = \begin{bmatrix} 1 & 0.7 \\ 0.9 & 1 \end{bmatrix}$.and we obtain the following result :

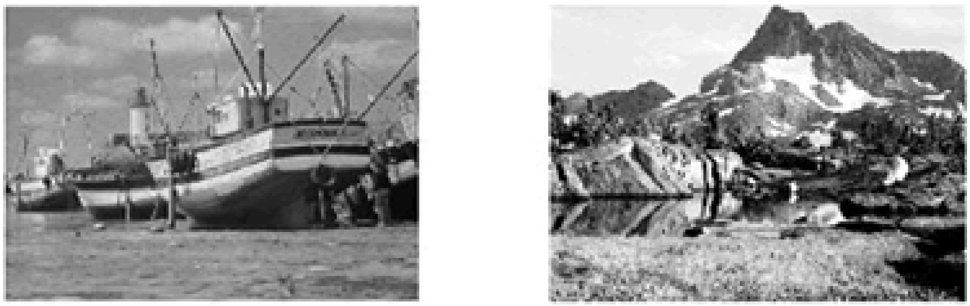

Figure 3.7: Original images

Figure 3.8: Mixed images

Figure 3.9: Estimated original images

Conclusion

During the preparation of this modest work, we have tried in the first part to present the Blind Source Separation (BSS) problem with mathematical formulation, and the different techniques of separation concerning the instantaneous linear mixing systems, then we outline the higher order statistics (HOS) (cumulants and moments of order greater than two) with their theoretical properties and their effectiveness in the area of source separation. At the end of this dissertation, we detail an algebraic technique for separation which is based on the fourth-order cumulants with experiments have been carried out on different signals using this method.
Finally, we hope to have the ability to explore this vast field of signal processing.

Bibliography

[1] Hérault J, Jutten C, Ans B (98) Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In: GRETSI, Paris, France, pp 1014-1022 .

[2] Xianchuan Yu, Dan Hu and Jindong Xu, Blind source separations theory and applications, (2014) Wieley.

[3] Mansour A, Barros AK, Ohnishi N (2000) Blind separation of sources: Methods, assumptions and applications. EICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E83-A(8):1498-1512 .

[4] A. Kardec Barros, A. Mansour, and N. Ohnishi, Adaptive blind elimination of artifacts in ecg signals, in International ICSC Workshop o Independence & Artificial Neural Networks 98, Tenerife-Spain, 11-13 February 1998.

[5] A. Kardec Barros, A. Mansour, and N. Ohnishi, Removing artifacts from ecg signals using independent componentsanalysis,NeuroComputing, vol. 22, pp. 173-186,1999.

[6] N. Charkani, Séparation auto-adaptative de sources pour les mélanges convolutifs. Application à la téléphonie mains-libres dans les voitures, Ph.D. thesis, INP Grenoble, Novembre 1996.

[7] G. D'urso and L. Cai, Sources separation method applied to reactor monitoring, in Proc. Workshop Athos working group, Girona, Spain, June 1995.

[8] A. M. KAGAN, Y. V. LINNIK, C. R. RAO, Characterization Problems in Mathematical Statistics, Wiley, 1973.

[9] Jean-Louis Lacoume, Pierre-Olivier Amblard, Pierre Comon. Statistiques d'ordre supérieur pour le traitement du signal. MASSON, pp.290, 1997.

[10] MANSOUR, A., and JUTTEN, c.: 'A direct solution for blind separation of sources', IEEE Trans. Sig. Process., 1996.

[11] CRUCES, s., MARTIN, R., and ACHA, J.I.: 'Blind separation of sources: stability results and comparisons'. Fourth Bayona Workshop on Intelligent Methods for Signal Processing and Communications,Bayona,24-26June 1996 .

[12] . L.JARBOUI, Méthodes avancées de séparation de sources applicables aux mélanges linéaires-quadratiques. Ph.D thesis , École Nationale d'Ingénieurs de Sfax, Novembre 2017.

[13] Hyvarinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks 10(3):626 -634 .

[14] Mansour A, Jutten C (1999) What should we say about the kurtosis? IEEE Signal Processing Letters 6(12):321-322 .

[15] Comon P (1994) Independent component analysis, a new concept? Signal Process 36(3):287-314 .

[16] R.M. Clemente and J.I. Acha : ' Blind separation of sources using a new polynomial equation ' , electronics letters , 30 January 1997 , Vol: 33 , pp :176 - 177 .

[17] P. McCULLAGH, Tensor Methods in Statistics, Monographs on Statistics and Applied Probability. Chapman and Hall, 1987 .

[18] Vrabie , V.D, statistiques d'ordre supérieur : applications en géophisique et électrotechnique ,Ph.D.thesis ,I.N.P.of Grenoble,France,2003 .

[19] K. V.MARDIA, "Tests of univariate and multivariate normality", in Handbook of Statistics, Vol.1, P. R. Krishnaiah, Ed., pp. 279–320. North-Holland, 1980 .

Résumé

Notre objectif dans ce mémoire est d'englober les différents types de problèmes concernant la séparation aveugle des sources (SAS) , définir les statistiques d'ordre supérieures (SOS) comme un outil de base dans les différentes techniques de séparation, et effectuer des expériences sur des mélanges de signaux utilisant une méthode de séparation basée sur les (SOS).

Les mots clés : Séparation aveugle des sources (SAS), Analyse en composantes indépendantes (ACI), Divergence de Kullback, Information mutuelle, Entropie, Statistiques d'ordre supérieures .

ملخص

هدفنا في هذه الرسالة هو تغطية مختلف أنواع المشاكل المتعلقة بالفصل الأعمى للمصادر ، (ف أ م) وتحديد إحصائيات الترتيب الاعلى (إ ت أ) كأداة أساسية في تقنيات الفصل المختلفة ، وإجراء تجارب على خليط الإشارات باستخدام طريقة فصل تعتمد على (إ ت أ) .

الكلمات المفتاحية : الفصل الأعمى للمصادر، تحليل المكونات المستقلة ، تباعد كيلباك ، المعلومات المتبادلة ، الانتروبيا ، إحصائيات الترتيب الاعلى .

Abstract

Our aim in this dissertation is to cover the different kind of problems concerning blind source separation (BSS), to define higher order statistics (HOS) as a basic tool in different separation techniques, and to perform experiments on signal mixtures using a separation method based on (HOS) .

Key-words: Blind source separation (BSS), Independent component analysis (ICA), Kullback divergence, Mutual Information, Entropy, Higher order statistics .