

UNIVERSITÉ SAAD DAHLEB DE BLIDA

Faculté des Sciences de l'Ingénieur

Département D'Electronique

MÉMOIRE DE MAGISTER

Spécialité : Communication

**REPRÉSENTATIONS GRAPHIQUES EN TROIS
DIMENSIONS DES SIGNAUX EEG SÉPARES PAR LA
MÉTHODE D'ANALYSE EN COMPOSANTES
INDÉPENDANTES**

Par

Mme KAOULA. YKRELEF. Ikram

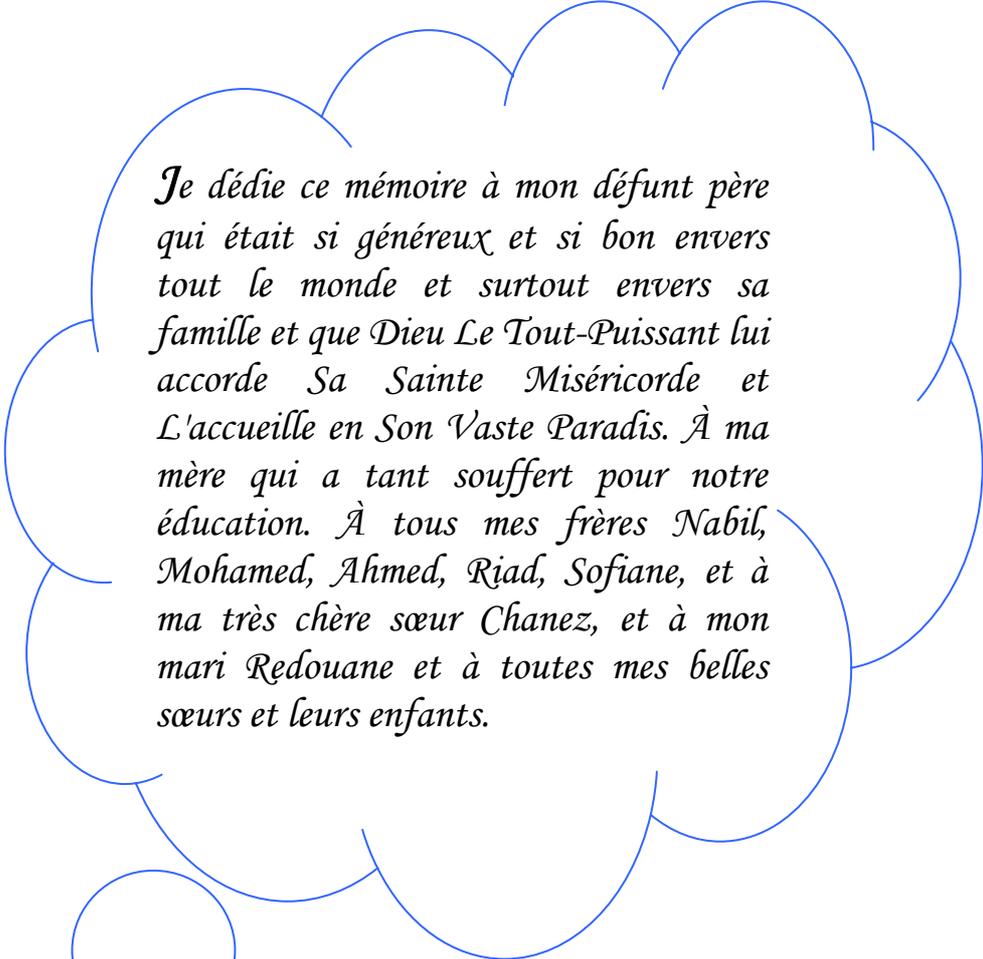
Devant le jury composé de :

Mr A. GUESSOUM	Professeur. U de Blida	Président.
Mr M. BENSEBTI	MC, U de Blida	Examineur
Mme H. BOUGHRIRA	CC, U de Blida	Examinatrice
Mr M. OULD-ZMIRLI	MC, CU de Médéa	Examineur
Mr M. ATTARI	Professeur, USTHB	Invité
Mr B. KAZED	CC, U de Blida	Rapporteur

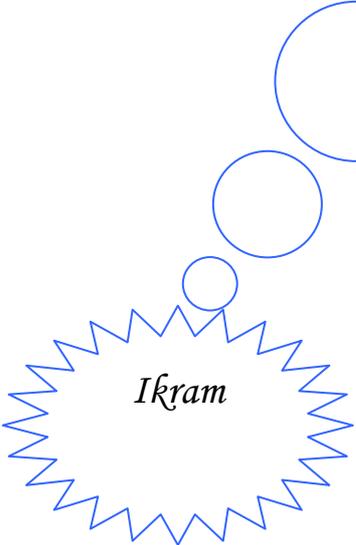
Blida, juillet 2005.



DÉDICACE



Je dédie ce mémoire à mon défunt père qui était si généreux et si bon envers tout le monde et surtout envers sa famille et que Dieu Le Tout-Puissant lui accorde Sa Sainte Miséricorde et L'accueille en Son Vaste Paradis. À ma mère qui a tant souffert pour notre éducation. À tous mes frères Nabil, Mohamed, Ahmed, Riad, Sofiane, et à ma très chère sœur Chanez, et à mon mari Redouane et à toutes mes belles sœurs et leurs enfants.



Ikram

REMERCIEMENTS

Je tiens à remercier vivement mon promoteur M^{er} KAZED BOUALLEM, qui n'a pas ménagé son temps et ses efforts afin de m'aider à réaliser ce modeste travail avec beaucoup de dévouement et de patience.

Je remercie tous mes professeurs durant tout mon cycle d'étude.

Je remercie également toute ma famille et surtout ma mère pour sa compréhension ainsi que pour son dévouement total durant toute sa vie pour sa famille, et pour son soutien moral et financier durant mes études, ainsi que tous mes frères et ma sœur ainsi que mon mari pour son soutien et son aide et sa compréhension.

Mon remerciement se retourne de ma part et de la part de mon promoteur M^{er} B. KAZED, en premier lieu envers le Professeur AREZKI le chef de service de NEUROLOGIE de l'Hôpital CHU de Blida secteur Frantz Fanon, pour son autorisation d'accéder au matériel en toute liberté pour pouvoir effectuer nos tests, ainsi qu'au maître assistant D^r DRAI, et aux techniciens du EEG M^{er} et M^{me} BENAROUS. Notre remerciement se retourne aussi envers M^{er} A. Belouchrani pour ses conseils ainsi que son orientation lors de mes premiers pas dans ce projet sur l'Analyse en Composantes Indépendantes.

Et enfin je remercie tous ceux qui m'ont aidée de près ou de loin pour achever ce travail si intéressant. Je cite parmi eux M^{er} NEHAL Abdelkader bibliothécaire à l'institut d'électronique pour son aide au niveau de la documentation.

IKRAM

RESUME

Ce travail a pour but d'acquérir et séparer les signaux EEG accompagnés d'artéfacts par la méthode d'Analyse en Composantes Indépendantes. Une représentation cartographique en trois dimensions de l'activité cérébrale sera aussi implémentée.

Mots clés : ACI, Séparation de sources, EEG.

ABSTRACT

The purpose of this work is to acquire and separate EEG signals with artifacts by the method of the Independent Component Analysis. A three dimensions cartographic representation of the cerebral activity will also be carried out.

Key words: ICA, Source Separation, EEG.

ملخص

، إشارات الدماغ.

كلمات مفتاحية: التحليل المركبات المستقلة، فصل

LISTE DES SYMBOLES ET DES ABREVIATIONS

- w : vecteur colonne (minuscule).
- w^T : vecteur ligne (minuscule).
- W** : représente une matrice (majuscule en gras).
- ACI* : **A**nalyse en **C**omposantes **I**ndépendantes.
- ACP* : **A**nalyse en **C**omposantes **P**incipales.
- $\langle x \rangle$: La moyenne de la variable x .
- $\langle xx^T \rangle$: La matrice de covariance de la donnée x .
- $p_1(y_1)$: La fonction de densité de probabilité marginale de la variable, y_1
- $p(y_1, y_2)$: La fonction de densité de probabilité joint de y_1 et y_2
- ERP* : **E**vent-**R**elated **P**otential.

INTRODUCTION

L'étude des signaux issus du cerveau humain, ne cesse de progresser depuis **1929** lors de la découverte du premier appareil *EEG – l'Electroencéphalographie* - par **Hans Berger**, jusqu'à ses développements actuels. Cependant, les origines dynamiques exactes de ces signaux *EEG*, et la relation avec la fonction cérébrale sont difficiles à interpréter, car les signaux enregistrés au niveau du cuir chevelu sont en fait un mélange de signaux de multiples générateurs cérébraux.

La séparation aveugle de sources par l'Analyse en Composantes Indépendantes *ACI* a reçue une attention particulière à cause de ses applications potentielles dans le traitement de signal comme pour la reconnaissance des systèmes de la parole, télécommunications et le traitement des signaux médicaux. Le but *d'ACI* est de retrouver les sources indépendantes à partir des signaux enregistrés par des capteurs; ces enregistrements sont le mélange de sources de signaux indépendants. Cette méthode ne décorrèle pas seulement les signaux (statistiques du second ordre) mais aussi réduit la dépendance statistique d'ordre supérieur, afin de rendre les signaux aussi indépendants que possible.

Le problème de la séparation de sources été étudié par différents chercheurs, Comon [28] a développé le concept de la méthode *d'Analyse en Composantes Indépendantes* en supprimant les statistiques d'ordre supérieur aux cumulants d'ordre quatre, Bell et Sejnowski [10] ont dérivé un nouvel algorithme d'apprentissage d'organisation automatique "learning self-organising algorithm" qui maximise l'information transférée dans un réseau d'unités non linéaires, Bell et Sejnowski [12] ont proposé une solution pour la séparation aveugle en utilisant la théorie d'information basée sur les sigmoïdes; Amari et al [36] ont dérivé un nouvel algorithme en ligne qui minimise la dépendance statistique entre les sorties pour la séparation aveugle des signaux mélangés, Makeig et al [33] ont appliqué l'Analyse en Composantes Indépendantes pour décomposer un ou plusieurs potentiels évoqués (*ERP*) enregistrés à plusieurs capteurs au niveau du cuir

chevelu, Hyvärinen [6] a présenté un nouvel algorithme rapide pour l'Analyse en Composantes Indépendantes, où il montre comment une règle d'apprentissage d'un système neuronal peut être transformée en itérations de point fixe, Amari [31] présente explicitement l'efficacité d'utiliser le gradient naturel au lieu du gradient ordinaire en apprentissage, Hyvärinen [3] a présenté la combinaison de deux approches différentes de l'Analyse en Composantes Indépendantes linéaire; l'approche de la théorie d'information de P. Comon dans [28] et de la poursuite de projection "projection pursuit". T.- P. Jung et al [41] ont présenté les suppositions d'Analyse en Composantes Indépendantes et démontré son application aux enregistrements électriques et hémodynamiques du cerveau humain, T. W. Lee et al [42] ont proposé des résultats pratiques obtenus pour la séparation aveugle de sources où le nombre de sources est supérieur aux nombre de mélanges en utilisant les représentations d'apprentissage "learning overcomplete representation".

Les données analysées par l'*Analyse en Composantes Indépendantes* peuvent être originaire de différentes sortes de domaine d'application, y compris la séparation d'artéfacts dans les données *EMG* et *EEG*, indicateurs économiques, images digitales, télécommunications.

Dans notre travail on essayera d'utiliser cette technique d'Analyse en Composantes Indépendantes pour retrouver plusieurs composantes élémentaires (les sources *EEG*) à partir de combinaisons linéaires de ces sources (les signaux *EEG* mélangés) enregistrés par plusieurs capteurs. Cette technique sera utilisée pour retrouver les signaux sources émis par le crâne qui sont mélangés, puis captés par les électrodes. Cette séparation sera exploitée pour extraire les artéfacts, très souvent indésirables pour l'analyse et le diagnostic.

Après la séparation des signaux *EEG*, nous allons procéder à une présentation en trois dimensions avant et après la séparation, l'activité cérébrale sera schématisée sur un crâne avec des nuances de couleurs indiquant l'intensité du signal, ceci permettra une meilleure visibilité pour l'utilisateur, ce traitement sera effectué sous *MATLAB*.

L'acquisition des signaux *EEG* passe par l'utilisation d'une carte de conversion analogique numérique, réalisés par mes collègues ingénieurs [27]. Nous développerons un programme en langage *BUILDER C++* pour l'affichage et le stockage des données numérisées et ce afin de les appliquer dans les algorithmes de séparation cités précédemment.

Le contenu de ce mémoire est présenté sous forme de quatre chapitres :

Le premier chapitre résumera les principes fondamentaux de la méthode d'*Analyse en Composantes Indépendantes* et la validation des algorithmes utilisés, le second chapitre contiendra des généralités sur l'*EEG*, le troisième résumera la partie matérielle de ce projet pour la réalisation des cartes d'acquisition et de conversion, et enfin le dernier chapitre comportera les différents résultats obtenus lors de la séparation des signaux *EEG* et les différentes représentations graphiques des signaux *EEG* ainsi que celles après la séparation des artéfacts.

CHAPITRE 1

ANALYSE EN COMPOSANTES INDEPENDANTES

1.1 Introduction

L'Analyse en Composantes Indépendantes *ACI* est utilisée pour la séparation aveugle de sources, qui concerne les techniques de traitement du signal, mis en œuvre pour retrouver plusieurs composantes élémentaires (les sources) à partir de combinaisons linéaires de ces sources sur plusieurs capteurs [20].

ACI ([28]), est une technique de traitement de signal dont le but est d'écrire un ensemble de variables aléatoires sous forme de combinaisons linéaires de composantes statiquement indépendantes. Une des applications d'*ACI* est la séparation aveugle de Sources.

► Motivation

Pour illustrer la méthode d'Analyse en Composantes Indépendantes, on a opté pour un exemple connu sous l'appellation " Le problème Cocktail Party ". Lorsque deux personnes parlent simultanément dans une chambre dans des microphones. Ces derniers nous donnent deux signaux enregistrés, qu'on dénotera par $x_1(t)$ et $x_2(t)$. Chacun de ces enregistrements est une somme de chaque signal émis par les deux orateurs, quand dénotera par $s_1(t)$ et $s_2(t)$.

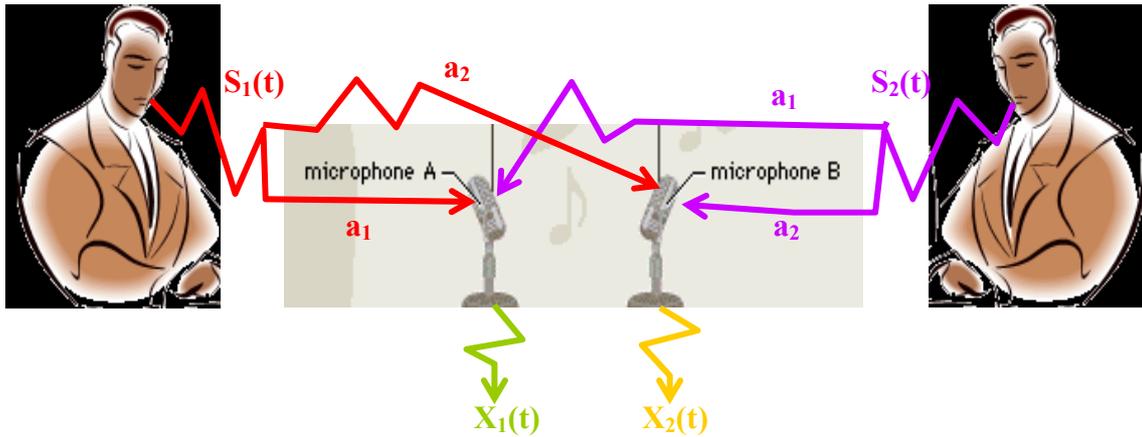


Figure 1 -1. Le problème Cocktail Party.

On peut écrire ceci comme une équation linéaire :

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t)$$

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t)$$

Où a_{11} , a_{12} , a_{21} , a_{22} sont des paramètres qui dépendent des distances des microphones des orateurs.

► Formules générales

Dans la forme la plus simple d'ACI ([28]), et si on observe m variables aléatoires scalaires x_1, x_2, \dots, x_m qui sont supposées être les combinaisons linéaires de n composantes indépendantes inconnues s_1, s_2, \dots, s_n mutuellement indépendantes et centrées. On suppose que $n \leq m$. On peut écrire ceci comme une équation linéaire :

$$x_1 = a_{11}s_1 + a_{12}s_2 + \dots + a_{1n}s_n.$$

$$x_2 = a_{21}s_1 + a_{22}s_2 + \dots + a_{2n}s_n.$$

$$\vdots$$

$$x_m = a_{m1}s_1 + a_{m2}s_2 + \dots + a_{mn}s_n$$

Si on range les variables observées x_i dans un vecteur ligne $x = (x_1, x_2, \dots, x_m)^T$ et les composantes s_i dans le vecteur s respectivement, alors l'expression est donnée par :

$$x = As \tag{1.1}$$

Où : A est une matrice $m \times n$ inconnue, nommée *Matrice Mélange*.

Le problème de base d'*ACI* est d'estimer les sources s_i à partir des mélanges x_j , ou similairement estimer la matrice mélange A [6].

Après l'estimation de A , on peut calculer son inverse, notée W et obtenir les composantes indépendantes simplement par [5] :

$$s = Wx \quad (1.2)$$

La condition fondamentale du modèle est qu'on peut estimer seulement les composantes indépendantes non-gaussiennes (excepté si juste une composante indépendante est gaussienne). Cependant, ni les énergies ni les signes des composantes indépendantes ne peuvent être estimés, car n'importe quelle constante multipliant une composante indépendante dans l'Eq. (1.1) peut être éliminée en divisant la colonne correspondante de la matrice mélange A par la même constante. Il faut noter aussi que l'ordre entre les composantes indépendantes n'est pas défini.

Le problème d'estimation de la matrice A dans Eq. (1.1) peut être simplifié en appliquant un prétraitement sur les données, par un blanchiment de la donnée x . [28]. Le blanchiment transforme la matrice mélange en une nouvelle matrice notée \tilde{A} ; tel que le vecteur observé x est linéairement transformé en un vecteur $\tilde{x} = Mx$ dont les éléments \tilde{x}_i sont mutuellement décorrélés et ayant tous une variance unitaire. Donc la matrice de corrélation de \tilde{x} est égale à la matrice unité : $E\{\tilde{x}\tilde{x}^T\} = I$. Cette transformation peut être réalisée par l'Analyse en Composantes Principales *PCA* classique [6]. En même temps, la dimensionnalité de la donnée peut être réduite de sorte que la dimension du vecteur de donnée transformée \tilde{x} est n , ce dernier représente le nombre des composantes indépendantes [6], la taille de la matrice M est $n \times m$.

Les algorithmes courants pour *ACI* peuvent être divisés en deux catégories ; les algorithmes de la première catégorie ([26]), c'est une approche en bloc 'batch mode', ce sont les méthodes dites directes consistant en la maximisation ou la minimisation des critères. Les algorithmes de la deuxième catégorie sont les algorithmes adaptatifs, ce sont les méthodes en boucle souvent, basés sur les méthodes du gradient stochastique. [10], [12], [6].

1.2 Indépendance

1.2.1 Définition et principes fondamentaux

Pour définir le concept de l'**indépendance**, on considère deux variables aléatoires scalaires y_1 et y_2 . Fondamentalement, les variables y_1 et y_2 sont considérés indépendants si l'information dans la valeur de y_1 ne rapporte aucune information sur la valeur de y_2 , et vice versa.

Techniquement, l'indépendance peut être définie par la densité de probabilité. Nous dénotons par $p(y_1, y_2)$ la fonction de densité de probabilité jointe de y_1 et y_2 . Ainsi on dénote la fonction de densité de probabilité marginale par $p_1(y_1)$, alors, la fonction de densité de probabilité de y_1 quand on la considère seule est :

$$p_1(y_1) = \int p(y_1, y_2) dy_2 \quad (1.3)$$

De même pour $p_2(y_2)$. On considère alors que y_1 et y_2 sont indépendants si et seulement si la fonction de densité de probabilité jointe est factorisable de la manière suivante :

$$p(y_1, y_2) = p_1(y_1) p_2(y_2) \quad (1.4)$$

Cette définition s'étend naturellement pour tout nombre n de variables aléatoires, dans ce cas la fonction de densité de probabilité jointe doit être le produit de n termes. [5]

La définition peut être utilisée pour dériver l'une des plus importantes propriétés des variables aléatoires indépendantes; Soit deux fonctions, h_1 et h_2 données, on a :

$$E \{h_1(y_1)h_2(y_2)\} = E \{h_1(y_1)\} E\{h_2(y_2)\} \quad (1.5)$$

Ceci peut être prouvé comme suit :

$$\begin{aligned}
 E\{h_1(y_1)h_2(y_2)\} &= \iint h_1(y_1)h_2(y_2)p(y_1,y_2)dy_1dy_2 \\
 &= \iint h_1(y_1)p_1(y_1)h_2(y_2)p_2(y_2)dy_1dy_2 = \int h_1(y_1)p_1(y_1)dy_1 \int h_2(y_2)p_2(y_2)dy_2 \\
 &= E\{h_1(y_1)\}E\{h_2(y_2)\}.
 \end{aligned} \tag{1.6}$$

1. 2. 2 Variables incorréllées et variables indépendantes

Une forme plus faible de l'indépendance est l'incorrélation [16]. Deux variables aléatoires y_1 et y_2 sont dites incorréllées, si leurs covariance est nulle :

$$E\{y_1y_2\} - E\{y_1\}E\{y_2\} = 0 \tag{1.7}$$

Si les variables sont indépendantes, alors elles sont incorréllées, ce qui suit directement de l'Eq. (1.5), en prenant $h_1(y_1) = y_1$ et $h_2(y_2) = y_2$. [5]

D'autres parts, l'incorrélation n'implique pas l'indépendance. Par exemple, supposons que (y_1, y_2) sont des valeurs discrètes qui suivent une distribution telles que leurs probabilités est égale à 1/4 pour toutes les valeurs suivantes : (0, 1), (0, -1), (1, 0), (-1, 0). Alors y_1 et y_2 sont incorréllées, comme il peut être simplement calculé.

D'autre part,

$$E\{y_1^2y_2^2\} = 0; E\{y_1^2\}E\{y_2^2\} = 1/4 \tag{1.8}$$

$$\text{donc : } E\{y_1^2y_2^2\} \neq E\{y_1^2\}E\{y_2^2\}$$

tel que $(y_1 * y_2 = 0)$; et puisque $E(0) = 0$, donc $E\{y_1^2y_2^2\} = 0$

Alors la condition dans l'Eq. (1.5) n'est pas vérifiée, et les variables ne peuvent pas être indépendantes bien qu'elles soient incorréllées.

1. 2. 3 Cas de variables gaussiennes

La restriction fondamentale dans *ACI* est que les composantes indépendantes doivent être non gaussiennes.

Pour voir pourquoi les variables gaussiennes rendent *ACI* impossible, supposons que la matrice mélange est orthogonale et les variables sources indépendantes sont gaussiennes. Alors x_1 et x_2 -variables mélangés- sont gaussiennes, incorréllées, et d'une variance unitaire.

Leurs densité jointe est donnée par :

$$p(x_1, x_2) = \frac{1}{2} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right) \quad (1.9)$$

Cette distribution est illustrée dans la figure 1. 2 ci-dessous. La figure montre que la densité est complètement symétrique. Par conséquent elle ne contient aucune information sur les directions des colonnes de la matrice mélange. Ceci est la raison pour laquelle la matrice mélange ne peut pas être estimée. – En réalité, si juste une des composantes indépendantes est gaussienne, la matrice mélange peut être estimée. [5].

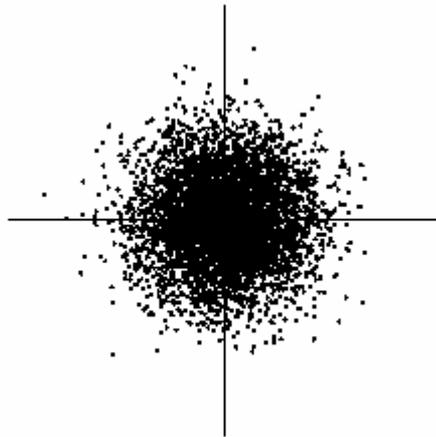


Figure 1- 2. Distribution multivariable de deux variables gaussiennes indépendantes.

Par contre si deux variables aléatoires indépendantes ne sont pas gaussiennes, leur distribution jointe est illustrée dans la figure 1 -3.

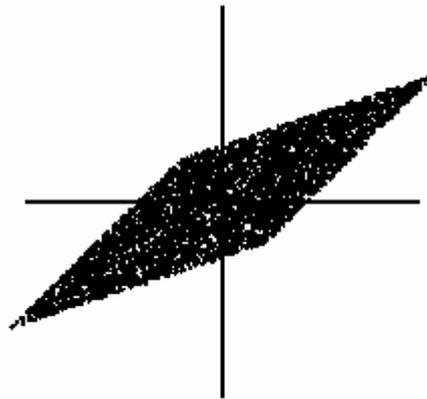


Figure 1 -3. La distribution jointe des mélanges observés x_1 et x_2 . Axe horizontale : x_1 ,
axe vertical : x_2 .

En réalité, on peut voir de la figure ci-dessus un moyen intuitif pour l'estimation de la matrice mélange : les arrêtes du parallélogramme sont dans les directions des colonnes de la matrice mélange [16].

1.3 Principe d'Analyse en Composantes Indépendantes (ACI)

La séparation à l'ordre deux est essentiellement fondée sur la décorrélation des signaux mélangés. Cette décorrélation est une conséquence de l'indépendance des sources. L'utilisation de la décorrélation n'explique pas complètement l'indépendance qui est une propriété bien plus forte que la décorrélation. C'est ici qu'apparaissent les statistiques d'ordre supérieur en séparation de sources. En utilisant des statistiques d'ordre supérieur à deux, on exploite plus complètement l'indépendance des sources. On obtient ainsi des critères permettant de réaliser la séparation sans informations supplémentaires. Ceci conduit à un nouveau concept : l'Analyse en Composantes Indépendantes (ACI), introduit par C. Jutten et J. Héroult dans [13], et développé dans [28]. Cette nouvelle approche vient compléter l'Analyse en Composantes Principales (ACP), fondée sur la seule décorrélation, cette dernière est une méthode très utilisée en statistiques [26].

1.3.1 Les principes d'estimation d'ACI

La séparation de sources, est fondée sur la notion d'indépendance statistique. L'entropie, la divergence de Kullback et l'information mutuelle permettent d'introduire des critères de gaussianité et d'indépendance.

Le critère d'indépendance, est basé sur le fait que, si un vecteur aléatoire x est formé de variables indépendantes, alors sa densité de probabilité est le produit des densités de probabilité marginales de chacune de ses composantes.

Les variables aléatoires dans le modèle *ACI* sont supposées non gaussiennes [9, 7]. En réalité, si juste une des composantes indépendantes est gaussienne, les composantes indépendantes peuvent être encore estimées.

1.3.1.1 Mesure de non-gaussianité

Pour utiliser la non-gaussianité dans une estimation *ACI*, on doit avoir une mesure de non-gaussianité des variables aléatoires, y . Pour illustrer cela, supposons que y est centrée (de moyenne nulle) et ayant une variance unitaire. [5].

a) Kurtosis

La mesure classique de la non-gaussianité est le *Kurtosis*. Classiquement le *Kurtosis* de y est défini par :

$$kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (1.10)$$

En fait, puisqu'on a supposé que y est de variance unitaire, alors le terme à droite de l'Eq. (1.10) est simplifié à $E\{y^4\} - 3$. Ceci montre que le *Kurtosis* est simplement une version normalisée du moment d'ordre quatre $E\{y^4\}$. Pour les y gaussiens, le moment d'ordre quatre est à $3(E\{y^2\})^2$. Donc, le *Kurtosis* des variables gaussiennes est nul [5]. Les variables aléatoires qui ont un *Kurtosis* négatif sont appelées sous-gaussiennes et celles avec le *Kurtosis* positif sont appelées sur-gaussiennes. Cependant il existe des variables aléatoires non gaussiennes qui ont un *Kurtosis* nul, mais on peut considérer ce cas comme très rare [5].

Les composantes indépendantes peuvent être trouvées par la maximisation ou la minimisation du *Kurtosis* dont les propriétés de linéarité sont :

$$kurt(x_1 + x_2) = kurt(x_1) + kurt(x_2) \quad (1.11)$$

et

$$kurt(\alpha x_1) = \alpha^4 kurt(x_1) \quad (1.12)$$

Où α est un scalaire.

b) Néguentropie

La deuxième mesure très importante de la non-gaussianité est donnée par la *Néguentropie*, elle est basée sur la quantité d'information de (la différentielle) l'*Entropie*.

L'*entropie* est un concept de base de la théorie de l'information ; l'entropie d'une variable aléatoire peut être interprétée comme le degré d'information qu'apporterait l'observation d'une variable. Plus la variable est aléatoire, c'est-à-dire la variable est imprévisible et non structurée, plus l'entropie est grande.

Considérons l'ensemble des vecteurs aléatoires dont la moyenne et la matrice de covariance sont fixées. Dans cet ensemble, le vecteur aléatoire ayant une entropie maximale et apportant donc une information maximale, et le vecteur aléatoire x_{gauss} .

L'entropie H est définie pour une variable aléatoire discrète Y comme :

$$H(Y) = -\sum_i P(Y = a_i) \log P(Y = a_i) \quad (1.13)$$

Où, les a_i sont les valeurs possibles de y , et $P(y)$ est la probabilité de y .

Cette définition très connue peut être généralisée pour le cas continu, où, dans ce cas, il est souvent appelé *différentielle d'entropie*. La différentielle d'entropie H d'un vecteur aléatoire y avec la densité $f(y)$ est définie comme suit :

$$H(y) = -\int f(y) \log f(y) dy \quad (1.14)$$

Un résultat fondamental de la théorie d'information est qu'une variable gaussienne a la plus grande entropie parmi toutes les variables aléatoires de variances égales. Cela veut dire que l'entropie peut être utilisée comme une mesure de non-gaussianité.

Pour obtenir une mesure de non-gaussianité nulle. On utilise souvent une version légèrement modifiée de la définition de la différentielle d'entropie, nommée *Néguentropie* notée J et définie comme suit :

$$J(y) = H(y_{gauss}) - H(y) \quad (1.15)$$

Où y_{gauss} est une variable aléatoire gaussienne de même matrice de covariance de celle de y .

Du point de vue théorique, la théorie d'information justifie bien cette notion de *Néguentropie* comme mesure de non-gaussianité. Par contre l'estimation de cette quantité peut s'avérer difficile en estimant la densité. Des approximations simples de *Néguentropie* sont très utiles.

✓ Approximations de la Néguentropie

L'estimation de la *Néguentropie* est difficile, comme il a été mentionné ci-dessus. La méthode classique d'estimation de la *Néguentropie* est l'utilisation des moments d'ordre supérieur [5], par exemple comme suit :

$$J(y) \approx \frac{1}{12} E\{y^3\}^2 + \frac{1}{48} kurt(y)^2 \quad (1.16)$$

La variable aléatoire y est supposée être de moyenne nulle et de variance unitaire. Cependant, la validité de telles approximations peut être plutôt limitée. En particulier, cette approximation souffre de la non robustesse produite par le kurtosis.

Pour éviter les problèmes rencontrés avec l'approximation précédente de néguentropie, une nouvelle approximation a été développée dans [8]. Cette approximation était basée sur le principe du maximum d'entropie. En général on obtient l'approximation suivante :

$$J(y) \approx \sum_{i=1}^p k_i [E\{G_i(y)\} - E\{G_i(v)\}]^2 \quad (1.17)$$

Où les k_i sont des constantes positives, et v une variable gaussienne de moyenne nulle et variance unitaire. La variable y est supposée centrée et de variance unitaire, et les fonctions G_i sont des fonctions non quadratiques.

Dans le cas où on utilise seulement une fonction non quadratique G , l'approximation devient [5] :

$$J(y) \propto [E\{G(y)\} - E\{G(v)\}]^2 \quad (1.18)$$

Dans [3], trois fonctions non quadratiques ont été proposées :

$$G_1(y) = \frac{1}{a_1} \log \cos(a_1 y) \quad (1.19)$$

$$G_2(y) = -\frac{1}{a_2} \exp(-a_2 y^2 / 2) \quad (1.20)$$

$$G_3(y) = \frac{1}{4} y^4 \quad (1.21)$$

Par la suite on aura besoin de leurs dérivées :

$$g_1(y) = \tanh(a_1 y) \quad (1.22)$$

$$g_2(y) = y \exp(-a_2 y^2 / 2) \quad (1.23)$$

$$g_3(y) = y^3 \quad (1.24)$$

Où $1 \leq a_1 \leq 2$ souvent on prend a_1 égale à 1 et $a_2 \approx 1$.

Par la suite la fonction $g_1(y)$ sera appelée fonction *tght*, la fonction $g_2(y)$ sera appelée fonction *gauss* et enfin la fonction $g_3(y)$ sera appelée fonction *puissance de trois*.

1.3.2 Prétraitement pour ACI

Avant l'application d'un algorithme *ACI* sur les données mélangées, il est habituellement très utile de faire des prétraitements. Dans cette section, on discutera sur des techniques de prétraitement qui mettraient le problème d'estimation d'*ACI* plus simple et mieux conditionné et surtout avec une convergence plus rapide.

La vitesse de convergence en utilisant juste les algorithmes de séparation qui seront discutés par la suite peut être très lente [12]. Cependant, si on prétraite notre donnée en enlevant les statistiques du premier et second ordre, l'accélération de la vitesse de convergence est ainsi considérablement élevée. Ce processus de prétraitement, se résume dans le fait de soustraire les moyennes des signaux d'entrées –mélangés- c'est le **centrage**. Par la suite décorrèle ces entrées centrées, tout en rendant la variance de chaque entrée unitaire, ce procédé est le **Blanchiment**. [12].

1. 3. 2. 1 Centrage

Le centrage est un prétraitement de base et nécessaire. Le centrage de la donnée captée est effectué en soustrayant son vecteur moyenne $m = E\{x\}$, en la rendant une variable de moyenne nulle. Tel que la nouvelle donnée sera donnée par : $x \leftarrow (x - \langle x \rangle)$. Tel que $\langle x \rangle$ marquerait la moyenne de x .

Après estimation de la matrice mélange A avec les données centrées, on peut compléter l'estimation en ajoutant le vecteur moyenne de la variable estimée s de nouveau à la variable centrée estimée s . Le vecteur moyenne de s est donnée par $A^{-1}m$, tel que m est l'espérance mathématique de x [5].

1. 3. 2. 2 Blanchiment

Le problème d'estimation de la matrice A , peut être simplifié par le blanchiment des données observées. Cette technique nous permet de transformer le vecteur observé x linéairement à un vecteur $\tilde{x} = Mx$.

Ce nouveau vecteur \tilde{x} sera blanc, c'est-à-dire que ces composantes sont décorréelées, et leurs variances sont égales à l'unité $E\{\tilde{x}\tilde{x}^T\} = I$, on dit que ce vecteur est blanc. [16], [5].

$$\tilde{x} = M x \quad (1. 25)$$

La matrice, M , est la matrice de décorrélation lorsque la matrice de covariance du vecteur de sortie, \tilde{x} , satisfait :

$$\langle \tilde{x}\tilde{x}^T \rangle = D \quad (1.26)$$

Où D est une matrice diagonale.

En général, il y a plusieurs M qui décorréllent. Par exemple, dans le cas de l'Eq. (1.25) lorsque $\langle \tilde{x}\tilde{x}^T \rangle = I$, alors :

$$M^T M = \langle xx^T \rangle^{-1} \quad (1.27)$$

Qui nous laisse clairement la liberté dans le choix de M . il y a, cependant, plusieurs solutions spéciales de l'Eq. (1.27). [15]

La matrice mélange A sera transformée en une nouvelle matrice, \tilde{A} , et on aura

$$\tilde{x} = Mx = MA s = \tilde{A} s \quad (1.28)$$

L'utilité du blanchiment réside dans le fait que la nouvelle matrice sera orthogonale ceci peut être vu par [5] :

$$E\{\tilde{x}\tilde{x}^T\} = \tilde{A} E\{s s^T\} \tilde{A}^T = \tilde{A} \tilde{A}^T = I \quad (1.29)$$

a) Solution Orthogonale ACP

La méthode d'Analyse en Composantes Principales est la solution orthogonale de l'Eq. (1.26) [15], c'est une méthode classique des statistiques du second ordre dont elle est basée sur la décomposition de la matrice de covariance de la donnée d'entrée, avec $\text{cov}\{x\} = E\{xx^T\}$. Notre but est de trouver M (Eq. (1.25)), tel que $\text{cov}\{\tilde{x}\} = I$. [15].

Ce blanchiment peut être effectué via la méthode *EVD* la décomposition en vecteurs propres (*Eigenvectors decomposition*) prenant $R_{xx} = E\{xx^T\}$ et trouvant les valeurs propres ainsi que les vecteurs propres :

$$R_{xx} = EDE^T \quad (1.30)$$

Où, E est la matrice des vecteurs propres et $D = \text{diag}(d_1, d_2, \dots, d_n)$ est la matrice diagonale des valeurs propres.

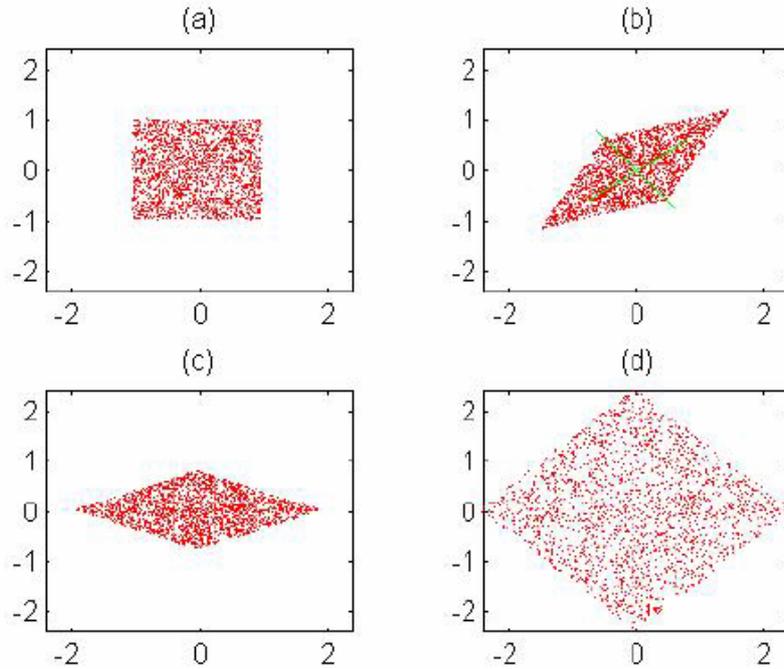


Figure 1 -4. ACP à deux dimensions. (a) Donnée Originale (b) Donnée Mélangée (c) Rotation E^T (d) Graduation par $D^{-1/2}$

On peut réécrire l' Eq. (1.30) de la manière suivante :

$$\begin{aligned}
 R_{xx} &= EDE^T = ED^{1/2}D^{1/2}E^T = E\{xx^T\} \\
 &= E \left\{ M^{-1} \tilde{x} \tilde{x}^T M^{-T} \right\} = M^{-1}E\{\tilde{x}\tilde{x}^T\}M^{-T} = M^{-1}M^{-T}
 \end{aligned}
 \tag{1.31}$$

Ce qui donnera :

$$M = (ED^{1/2})^{-1} = D^{-1/2}E^{-1} = D^{-1/2}E^T
 \tag{1.32}$$

Où, $D^{-1/2} = \text{diag}(d_1^{-1/2}, d_2^{-1/2}, \dots, d_n^{-1/2})$ [15].

b) Solution Symétrique ACZ

On force M à être symétrique, telle que $M^T = M$, alors la solution. M de l'Eq. (1.27) est

$$M = \langle xx^T \rangle^{-1/2} \quad (1.33)$$

On pourrait appeler cette solution ACZ puisque les filtres qu'elle produit sont de phases nulles (symétriques). [15]

Dans [12], cette méthode a été proposée pour blanchir les données. Tout en multipliant la matrice de décorrélation par deux. Alors la matrice M sera donnée par l'expression suivante :

$$M = 2 \sqrt{\langle xx^T \rangle^{-1}} \quad (1.34)$$

1. 3. 3 Algorithmes pour ACI

Le concept $d'ACI$ a été introduit par C. Jutten et J. Héroult dans [13] et développé dans [28], cette nouvelle approche vient compléter l'Analyse en Composantes Principales, fondée sur la seule décorrélation [28]. Plusieurs algorithmes ont été proposés pour la mise en œuvre $d'ACI$. On peut citer La Minimisation de moments d'ordre supérieur [20], minimisation de l'information mutuelle de signaux de sorties ou Maximisation de l'entropie de signaux de sorties [10], Minimisation de la divergence de Kullback-Leibler entre la densité conjointe du vecteur de sortie et le produit de distributions marginaux des signaux de sorties [36], [37].

Les algorithmes développés pour la séparation de sources ont pour objectif de construire un système linéaire donnant en sortie des signaux proportionnels aux signaux de

sources. Deux familles de méthodes ont été utilisées pour construire le séparateur comme la montre la figure 1. 5, **les méthodes dites directes** et **les méthodes en boucles**.

Dans les méthodes directes, un système annexe détermine les paramètres du séparateur à partir des signaux observés; mais les systèmes en boucle, utilisent la rétroaction. Un critère issu des signaux de sortie et fournit la règle d'adaptation des paramètres du séparateur.

Une question importante se pose d'entrée de jeu aux concepteurs de systèmes. Nous avons vu qu'une partie de la séparation peut-être faite à l'ordre 2 par des techniques classiques déjà développées. On peut donc construire des systèmes à deux étages enchaînant le blanchiment, réalisé à l'ordre 2, et la séparation finale, utilisant des statistiques d'ordre supérieur à 2. Ces systèmes découplant les deux étapes ont déjà fait leurs preuves. [26]

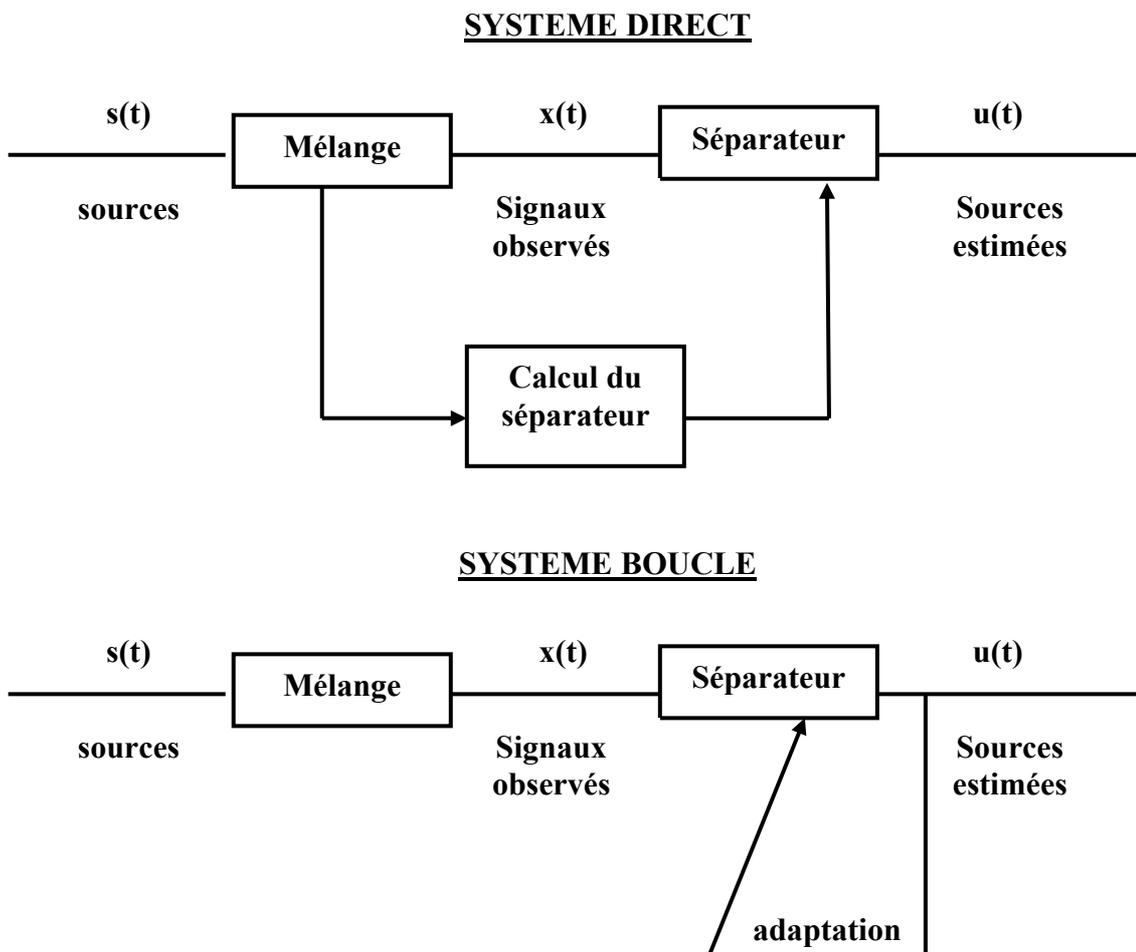


Figure 1. 5. Les systèmes directs et les systèmes en boucle pour la séparation de sources.

L'approche standard pour le système direct, est qu'il procède en deux étapes [24]. La première étant le prétraitement, et suivi d'un traitement utilisant des statistiques d'ordre supérieur pour compléter la séparation, cette dernière est fondée sur la minimisation ou la maximisation d'un critère[26]. Parmi les algorithmes directs pour *ACI*, *FastICA* de Hyvärinen [2,7], l'algorithme *JADE* pour (Joint Approximate Diagonalization of Eigenmatrices) de Cardoso qui a utilisé largement les cumulants d'ordre quatre [15,1].

La forme générale du deuxième système, dit en boucles, repose sur un critère d'adaptation qui fait varier les coefficients w_{ij} de la matrice séparatrice W tel que $u(t) = W x(t)$, et transforme ainsi le mélange observé $x(t)$, pour donner une sortie $u(t)$ dont les composants sont en relation proportionnelle aux sources $s(t)$. Alors la matrice W est obtenue récursivement en optimisant un critère $v(r)$ calculé à partir de la sortie. [26]. On cite quelques algorithmes ; méthode *EASI* pour (*Equivariant Adaptive Source Separation*) [1], l'algorithme de maximum de vraisemblance (Maximum likelihood, *ML*) de Pham et al [18] qui a formé ultérieurement la base de la méthode d' *EASI* de Cardoso et de Laheld [25].

Ces méthodes se rapprochent de l'algorithme "*Infomax*" [10], dont on va esquisser la dérivation et le développement dans la section suivante.

L'algorithme "*Infomax*" pourrait être appelé aussi *Infomax/ML ICA*. Auparavant, Cichocki et al [1] ont proposé un algorithme qui a motivé Amari [31] et ses collègues à montrer que son succès était dû à sa relation d'une modification du « Gradient Naturel » du gradient d' *Infomax/ML ICA*. Cette modification a considérablement simplifié l'algorithme, et rendu la convergence plus rapide et plus stable [41].

1.3.3.1 Algorithme Infomax

L'Analyse en Composantes Indépendantes *ACI* [4, 9] est le nom donné aux techniques pour trouver la matrice A tel que :

$$u = W\tilde{x} = WMx = WMAs \quad (1.35)$$

Où, u est le vecteur de données estimées.

Dans cette section, on supposera que les données sont prétraitées par le centrage et le blanchiment.

Les techniques de décorrélation telles que l'Analyse en Composantes Principales *ACP*, et qui assure $\langle u_i u_j \rangle = 0 \quad \forall ij$, *ACI* impose pour les critères les plus robustes, que la fonction de densité du vecteur de u soit factorisable tel que : $f_u(u) = \prod_{i=1}^N f_{u_i}(u_i)$. La recherche d'une telle factorisation implique que l'information mutuelle entre les u_i tendra vers zéro : $I(u_i, u_j) = 0 \quad \forall ij$. L'information mutuelle est une mesure qui dépend de toutes les statistiques d'ordre supérieur des u_i tandis que la décorrélation prend en compte les statistiques du second ordre seulement.

En utilisant le concept de la *différentielle d'entropie* défini dans Eq. (1.14), on définit l'*information mutuelle I* entre m variables aléatoires (scalaire) $y_i, i=1 \dots m$, comme suit :

$$I(y_1, y_2, \dots, y_m) = \sum_i H(y_i) - H(y) \quad (1.36)$$

Dans [5], un nouvel algorithme a été proposé pour exécuter *ACI*. Les seules suppositions antérieures sont que les composantes indépendantes inconnues, u_i , ont la même forme de la fonction de répartition notée : $F_u(u)$ [32].

L'algorithme *infomax* de base change les poids (matrice W) selon le gradient de l'entropie. Définissant $y_i = g(u_i)$ à être les variables transformées sigmoïdalement ([10]), la règle d'apprentissage est alors :

$$\Delta W \propto \frac{\partial H(y)}{\partial W} = E \left[\frac{\partial \ln |J|}{\partial W} \right] \quad (1.37)$$

Dans ceci on dénote, E comme l'espérance, $y = [g(u_1) \dots g(u_N)]^T$ et $|J|$ est la valeur absolue du *Jacobien* donné par la relation suivante :

$$J = \det \left[\frac{\partial y_i}{\partial x_j} \right]_{ij} \quad (1.38)$$

Dans le gradient stochastique ascendant on omet l'opération de l'espérance dans l' Eq. (1.37), ce qui donne [12] :,

$$\Delta W \propto [W^T]^{-1} + \hat{y} \tilde{x}^T \quad (1.39)$$

Où $\hat{y} = [\hat{y}_1 \cdots \hat{y}_N]^T$, dont les éléments sont :

$$\hat{y}_i = \frac{\partial}{\partial y_i} \frac{\partial y_i}{\partial u_i} = \frac{\partial}{\partial u_i} \ln \frac{\partial y_i}{\partial u_i} \quad (1.40)$$

Cependant, on utilisera souvent des fonctions sigmoïdales standard. Par exemple, pour la "fonction logistique", $y = (1 + \exp(-u))^{-1}$, et on dérive $\hat{y} = 1 - 2y$, et pour la fonction "fonction tangente hyperbolique", $y = \tanh(u)$, et on dérive $\hat{y} = 1 - y^2$ qui sont proposées dans [12].

Une modification de la règle d'apprentissage Eq. (1.39) a été proposée dans [36], qui utilise le gradient naturel plutôt que le gradient absolu de $H(y)$. Il revient à la multiplication du gradient absolu par $W^T W$, donnant dans notre cas la version modifiée suivante de l' Eq. (1.39). [15] :

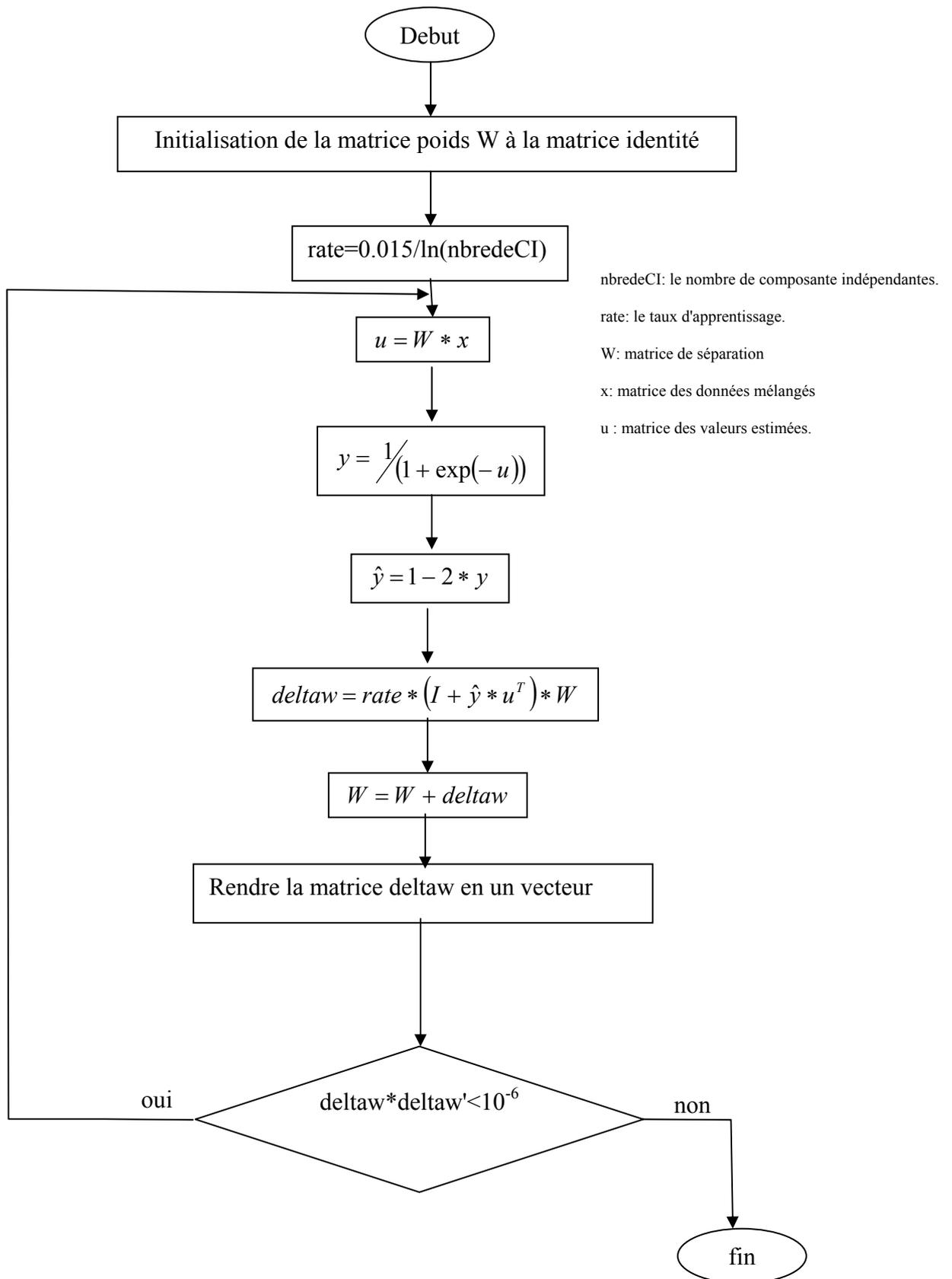
$$\Delta W \propto \frac{\partial H(y)}{\partial W} W^T W = (I + \hat{y} u^T) W \quad (1.41)$$

L'avantage de la loi présentée dans l'Eq. (1.41) sur celle présentée dans l'Eq. (1.39), est qu'on évite le calcul de la matrice inverse, ce qui veut dire que la loi d'apprentissage est simplifiée dans cette nouvelle formule par rapport à celle présentée dans l'Eq. (1.39), et que la vitesse de convergence est considérablement accélérée. [38].

La loi d'apprentissage donnée dans l'Eq. (1.41) converge, c'est-à-dire que $\langle \Delta W \rangle = 0$, lorsque $I = -\langle \hat{y} u^T \rangle$.

a) Organigramme de *Infomax*

La figure 1. 6 présente l'organigramme d'*Infomax*, cet algorithme s'applique à l'estimation des composantes indépendantes.

Figure I. 6. Algorithme *Infomax*

1.3.3.2 Algorithme point fixé ou "FastICA"

Dans la section § 1.3.1.1, nous avons introduit différentes mesures de non-gaussianité. En pratique, nous avons encore besoin d'un algorithme pour la maximisation de fonction contraste, par exemple celle présentée dans l'Eq. (1.18). Dans cette section, on introduit une méthode très efficace adaptée à cette tâche. On supposera que les données sont prétraitées par le centrage et le blanchiment discutés dans la section § 1.3.2.

a) Estimation d'une unité par FastICA

Pour commencer, on montrera tout d'abord la version d'une unité de *FastICA*. Par "unité" nous nous référons à un calcul d'une unité, enfin un neurone artificiel, ayant un vecteur poids w que le neurone est capable de se mettre à jour par une règle d'apprentissage.

La règle d'apprentissage *FastICA* trouve une direction, c'est-à-dire le vecteur unité w tel que la projection $w^T \tilde{x}$ maximise la non-gaussianité. Non-gaussianité est mesurée ici par l'approximation de la néguentropie $J(w^T \tilde{x})$ donnée dans l' Eq. (1.18).

FastICA est basé sur un schéma d'itération de point fixé (*a fixed-point iteration scheme*) pour trouver un maximum de non-gaussianité de $w^T \tilde{x}$, comme elle a été mesurée dans Eq. (1.18) [21, 22]. Il peut être encore dérivé comme une itération de Newton approximative [22, 29]. Dénoteons par g la dérivée de la fonction non quadratique G utilisée dans l'Eq. (1.18).

La forme de base de l'algorithme *FastICA* est comme suit :

1. choisir un vecteur poids w initial aléatoire.
2. $w^+ = E\{\tilde{x}g(w^T \tilde{x})\} - E\{g'(w^T \tilde{x})\}w$. L'espérance peut être estimée en utilisant un grand échantillon des vecteurs \tilde{x} (dire, 1000 points).
3. $w = w^+ / \|w^+\|$
4. Si pas de convergence, aller à 2.

Noter que la convergence signifie que l'ancienne et la nouvelle valeur de w pointe dans la même direction, c'est-à-dire que leur produit scalaire est presque égale à 1 [5].

La dérivation de *FastICA* est comme suit. Premièrement, il faut noter que les maxima de l'approximation de la néguentropie de $w^T \tilde{x}$ sont obtenus à certains optimum de $E\{G(w^T \tilde{x})\}$ [5]. Les optimums de $E\{G(w^T \tilde{x})\}$ selon les conditions de *Kuhn-Tucker* sous la contrainte $E\{(w^T \tilde{x})^2\} = \|w\|^2 = 1$ sont obtenus aux points où :

$$F(w) = E\{xg(w^T \tilde{x})\} - \beta w = 0 \quad (1.42)$$

Essayons de résoudre cette équation par la méthode de *Newton* ([40]). C'est-à-dire :

$$w^+ = w - (JF(w))^{-1} f(w) \quad (1.43)$$

Où $JF(w)$ est le jacobien de $F(w)$ de l'Eq. (1.42) [40].

En dérivant le jacobien on aura :

$$JF(w) = \frac{\partial}{\partial w} [E\{\tilde{x}g(w^T \tilde{x})\} - \beta w] = E\{xg'(w^T \tilde{x})\tilde{x}^T\} - \beta I = E\{\tilde{x}\tilde{x}^T g'(w^T \tilde{x})\} - \beta I \quad (1.44)$$

Pour simplifier l'inversion de cette matrice, on approxime le premier terme dans l'éq.(40). Puisque la donnée est blanchie, une approximation raisonnable donnera :

$$E\{\tilde{x}\tilde{x}^T g'(w^T \tilde{x})\} \approx E\{\tilde{x}\tilde{x}^T\} E\{g'(w^T \tilde{x})\} = E\{g'(w^T \tilde{x})\} I \quad (1.45)$$

Donc la matrice jacobienne devient diagonale, et peut être facilement inversée ; et alors on obtient l'itération de *Newton* approximative suivante :

$$w^+ = w - [E\{\tilde{x}g(w^T \tilde{x})\} - \beta w] / [E\{g'(w^T \tilde{x})\} - \beta] \quad (1.46)$$

Cet algorithme peut être plus simplifié en multipliant les deux côtés de l'Eq. (1.46) par $\beta - E\{g'(w^T \tilde{x})\}$. Ceci donne, après des simplifications algébriques, l'itération *FastICA*.

En pratique, l'espérance dans *FastICA* doit être remplacée par leurs estimations. Les estimations naturelles sont bien sûr les moyennes des échantillons correspondants [5].

b) Estimation de plusieurs unités par *FastICA*

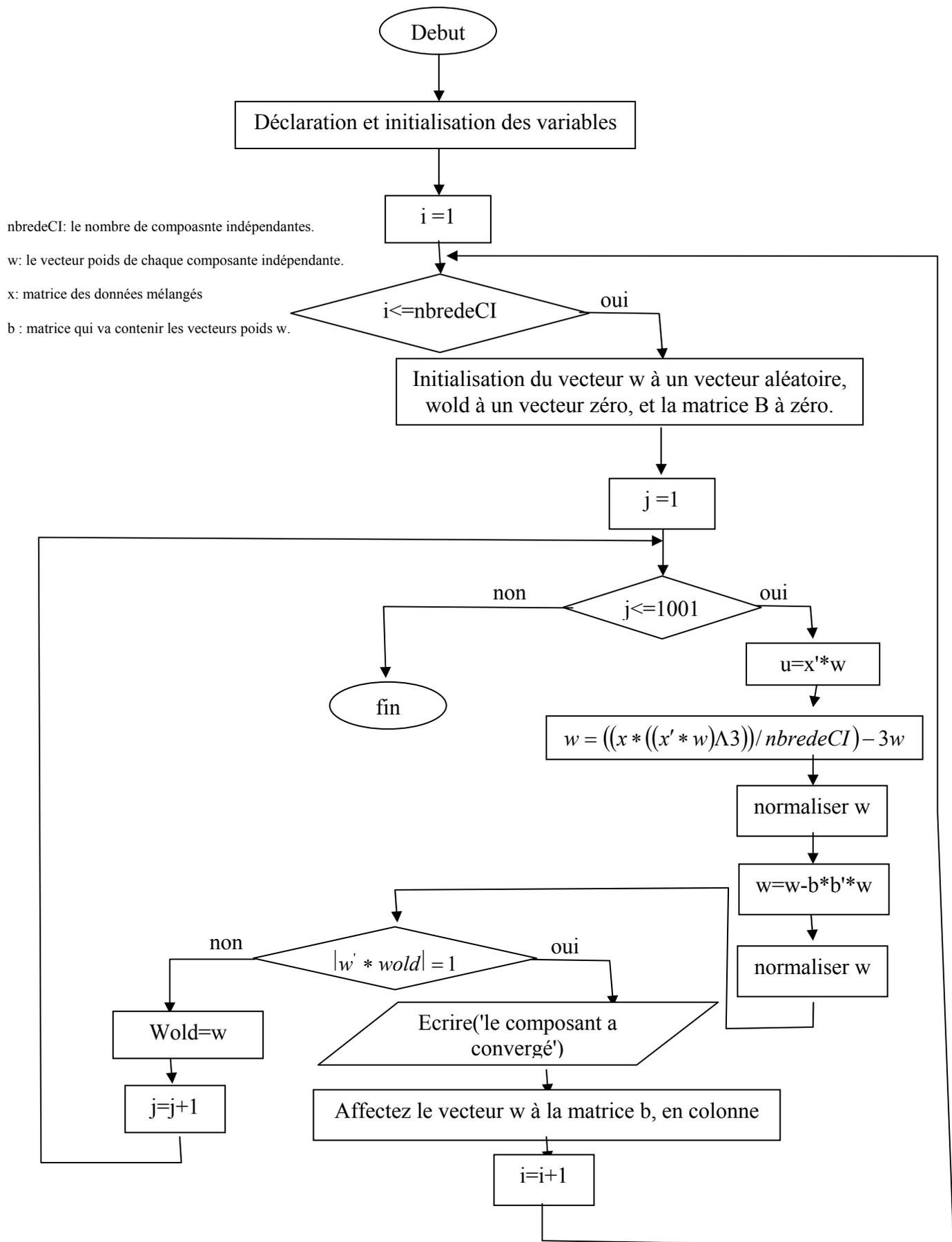
L'algorithme d'une unité présenté dans la sous section précédente, estime juste une des composantes indépendantes. Pour estimer plusieurs composantes indépendantes, on a besoin d'exécuter l'algorithme *FastICA* pour une unité en utilisant plusieurs unités (exp. neurones) avec les vecteurs poids w_1, \dots, w_n .

Pour empêcher les différents vecteurs de converger au même maxima on doit décorréler les sorties $w_1^T \tilde{x}, \dots, w_n^T \tilde{x}$ après chaque itération. Un moyen simple pour réaliser la décorrélation est le plan de *Déflation (a deflation scheme)* basé sur la décorrélation de *Gram-Schmit-Like*; ceci signifie qu'on estime les composantes indépendantes une par une. Lorsqu'on estime p composantes indépendantes, ou p vecteurs w_1, \dots, w_p , on exécute l'algorithme du point fixé (d'un seul élément) pour w_{p+1} et après chaque étape d'itération soustraire de w_{p+1} les projections $w_{p+1}^T w_j w_j$, avec $j=1, \dots, p$ des p vecteurs estimés précédemment, et puis normalisé w_{p+1} :

$$\begin{aligned}
 1. \quad w_{p+1} &= w_{p+1} - \sum_{j=1}^p w_{p+1}^T w_j w_j \\
 2. \quad w_{p+1} &= w_{p+1} / \sqrt{w_{p+1}^T w_{p+1}}
 \end{aligned}
 \tag{1.47}$$

c) Organigramme de *FastICA*

La figure 1. 7 présente l'organigramme de *FastICA*, cet algorithme s'applique à l'estimation des composantes indépendantes composante par composante.

Figure I. 7. Algorithme *FastICA*

1.4 Domaines d'applications de l'Analyse en Composantes Indépendantes

Des algorithmes de séparation de sources par des statistiques d'ordre supérieur ont été proposés dans divers domaines d'applications. Parmi ces applications qui ont utilisées cette analyse pour résoudre des problèmes on cite :

- ✓ Les problèmes des finances des entreprises.
- ✓ Le domaine de télécommunications.
- ✓ Le domaine de l'aéronautique, on cite la résolution de l'encombrement du trafic aérien.
- ✓ Le domaine médical, on cite comme exemple le traitement de l'électrocardiogramme, ainsi que le traitement de l'électroencéphalogramme...

1.5 Validation des algorithmes

Cette partie fait l'objet des résultats obtenus lors d'exécution des algorithmes *FastICA* ainsi que *Infomax*. Ces deux derniers ont été développés sous *MATLAB*.

L'application de ces algorithmes se résume dont le fait de séparer un mélange de signaux périodiques dans un premier temps, puis de séparer un mélange de signaux audio pour pouvoir certifier les résultats.

1.5.1 Séparation des sources

Les signaux utilisés dans notre validation, ont été générés sous *MATLAB*. Ces sources ont été multipliées par une matrice via le *script Rand* dont ses éléments sont compris entre 0 et 1, quand peut appelées matrice mélange, ainsi on obtient le mélange de nos sources, qu'on va séparer par nos algorithmes.

1.5.1.1 Séparation par *FastICA*

Pour vérifier la validité et la performance de cet algorithme, les mêmes mélanges pour l'authentification de l'algorithme *Infomax*.

a) Séparation des signaux périodiques

La première application, pour cette authentification est une application de quatre signaux périodiques (sinusoïdale, triangulaire,..) de mille échantillons chacun.

La figure 1. 8 nous montre le résultat de la séparation par l'algorithme *FastICA* des signaux périodiques mélangés, avec l'observation on constate que les quatre signaux périodiques sources ont été retrouvées avec une petite déformation dans la quatrième composante indépendante. Il faut noter que la vitesse de convergence de cet algorithme est considérablement élevée où cette séparation a été effectuée par trois itérations pour séparer la première composante, cinq itérations pour la deuxième, six itérations pour la troisième, et enfin deux itérations pour la dernière composante indépendante.

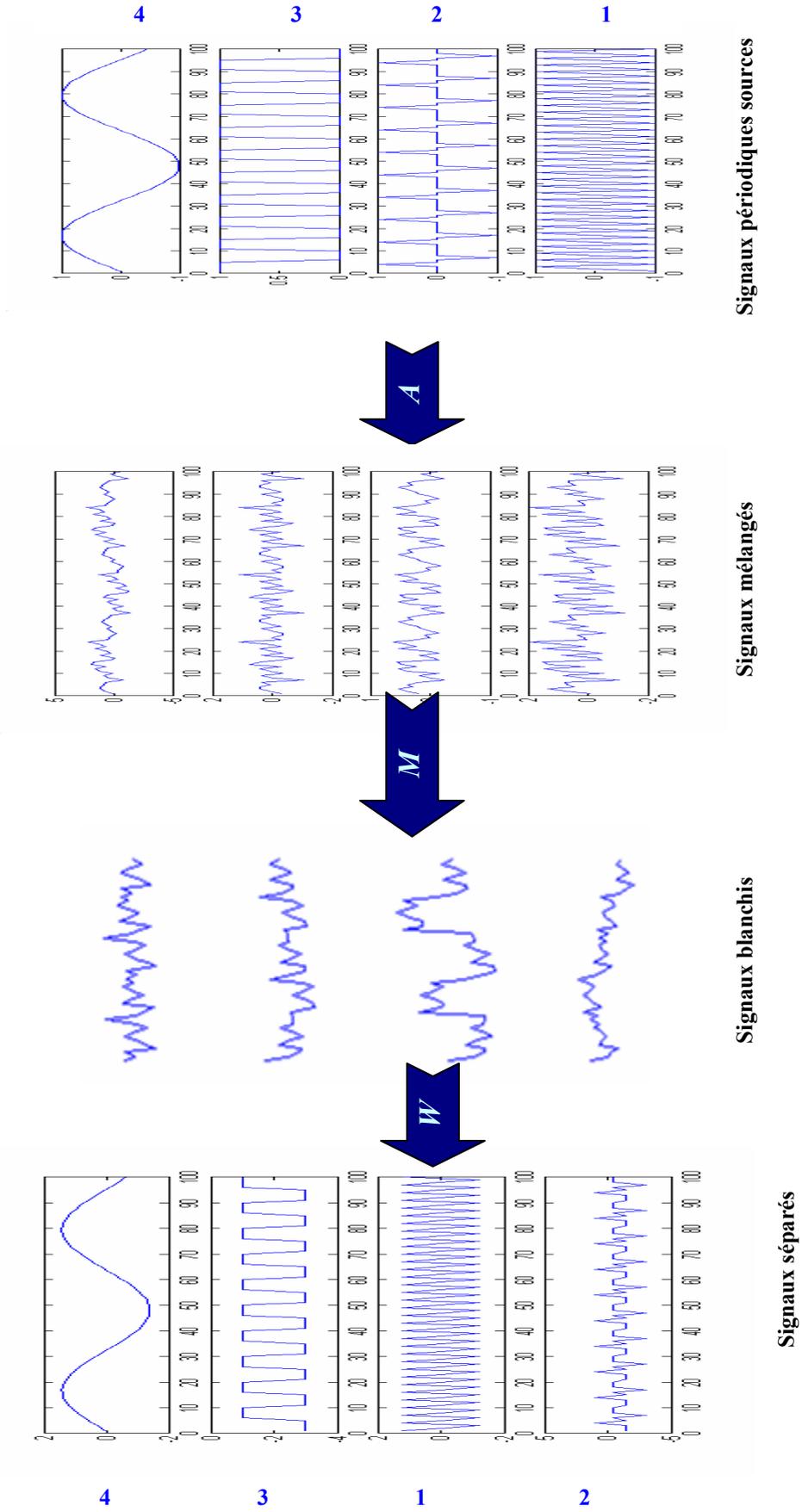


Figure 1. 8. Séparation de signaux périodiques par FastICA via ACP et la fonction Puiss3

b) Séparation des signaux audio

La deuxième application, est un mélange de quatre sons de trois secondes chacun. Les signaux sources, 'signaux audio', ont été convertis à des fichiers de données au format MATLAB ".MAT" avant de les mélanger.

La figure 1. 9, montre le résultat de la séparation par l'algorithme *FastICA* des signaux audio mélangés. À première vue de cette figure on peut distinguer les signaux sources. Mais le résultat doit être certifié par la conversion des fichiers de données au format MATLAB ".MAT" à des fichiers audio au format ".WAV".

En entendant ces sons, on constate que les sons ont été partiellement séparés.

Nous avons obtenu deux sons parmi les quatre sons clairement séparés, un troisième son est clairement séparé sauf qu'on constate une légère déformation vers la fin de cet échantillon et le dernier est mélangé.

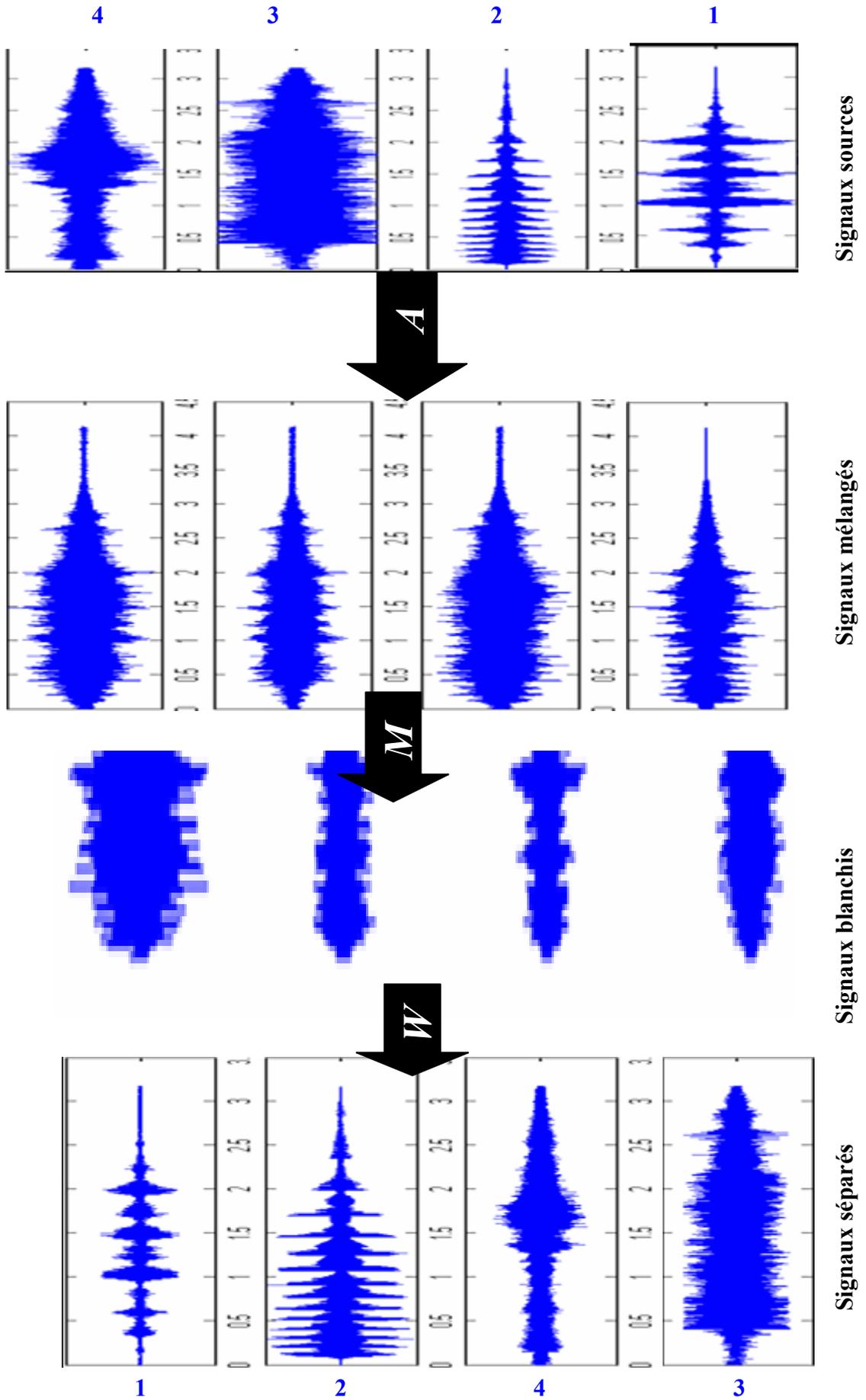


Figure. 1. 9 Séparation de signaux audio par l'algorithme *FastICA* via *ACZ* et la fonction gauss

1. 5. 1. 2 Séparation par *Infomax*

Pour vérifier la validité et la performance de cet algorithme, la simulation a été faite pour un mélange de signaux périodiques et un mélange de signaux audio, le taux d'apprentissage est égale à : $0.015/\log(\text{nombre CI})$.

a) Séparation des signaux périodiques

Comme pour l'algorithme *FastICA*, on utilisera les mêmes applications. Tout d'abord, la séparation du mélange des quatre signaux périodiques de milles échantillons chacun par l'algorithme *Infomax*.

La figure 1. 8 nous montre le résultat de la séparation par l'algorithme *Infomax*, on constate que les quatre signaux ont été séparés avec une légère déformation pour deux signaux.

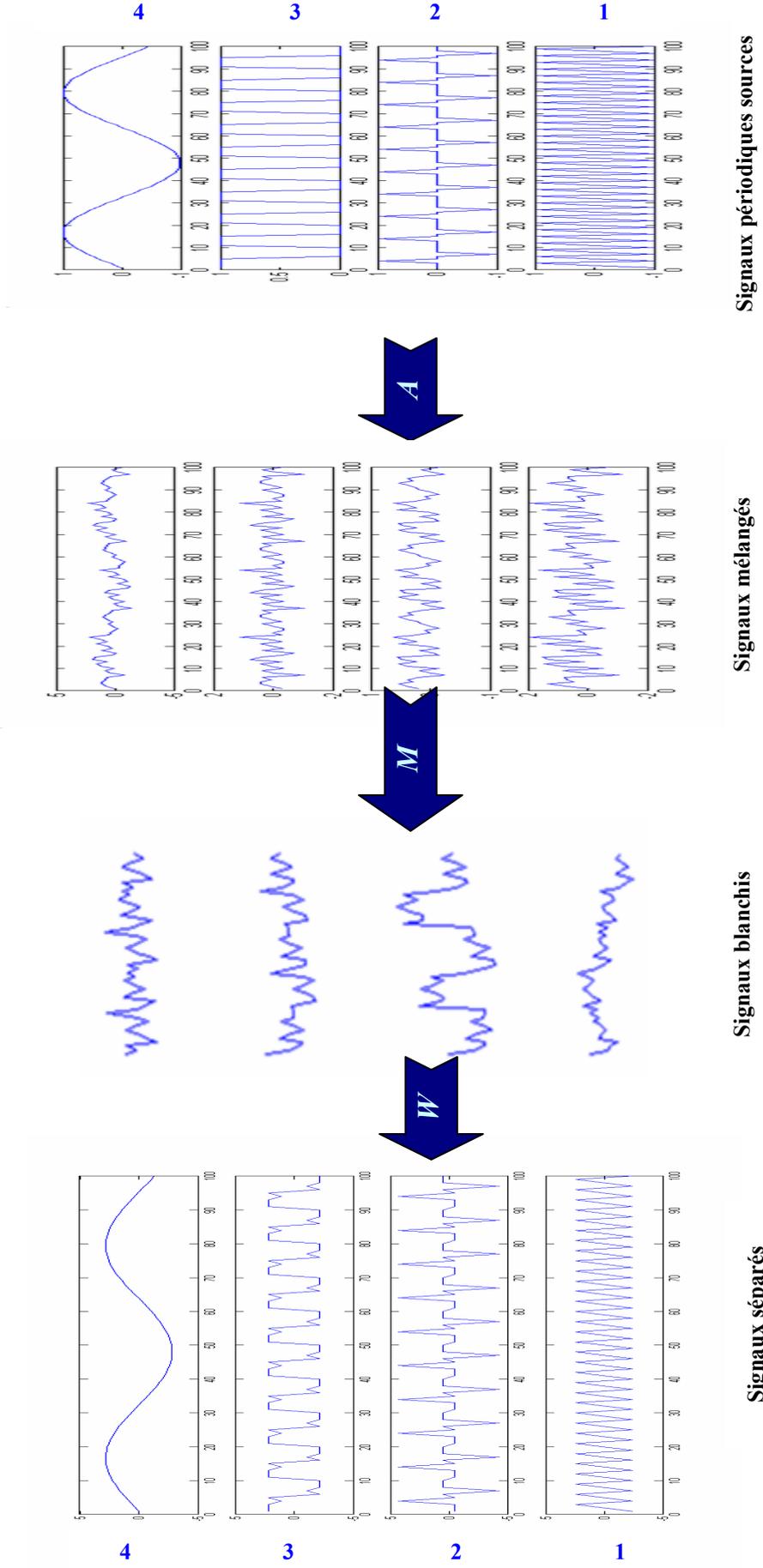


Figure. 1. 10. Séparation de signaux périodiques par Infomax

b) Séparation des signaux audio

La figure 1. 11, montre le résultat de la séparation du mélange des sons par l'algorithme *Infomax*. Le résultat doit être certifié par la conversion des fichiers de données au format MATLAB ".MAT " à des fichiers audio au format ".WAV ".

En entendant ces signaux audio séparés, on constate que les sons ont été partiellement séparés.

Nous avons obtenu l'un des sons clairement séparé sauf qu'on constate une légère déformation vers la fin de cet échantillon, pour les trois autres signaux il y a apparition d'un nouveau son avec une intensité plus élevée par rapport aux signaux testés.

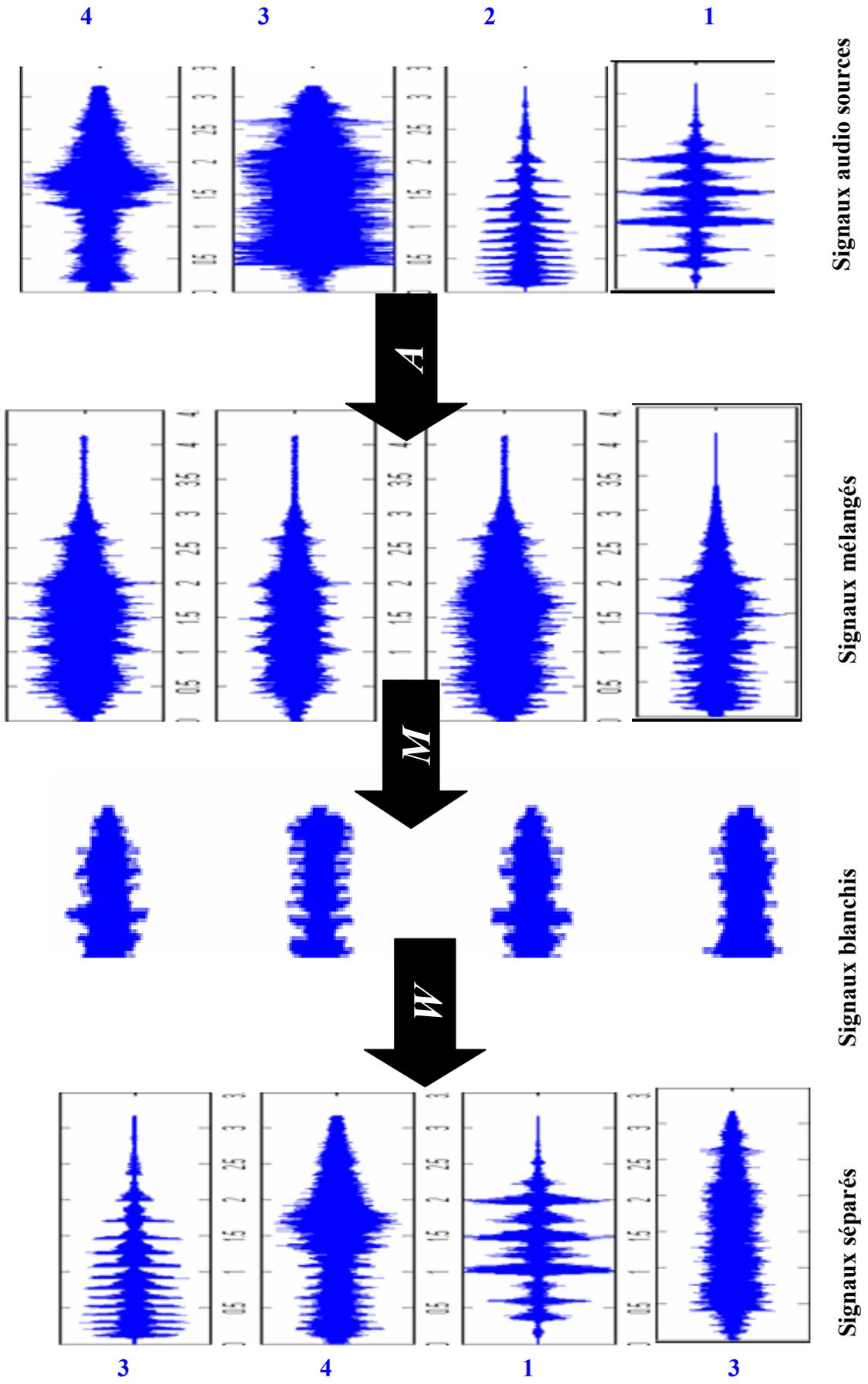


Figure. 1. 11. Séparation de signaux audio par l'algorithme *Infomax*

1.6 Conclusion

La technique qu'on a illustré dans ce chapitre, est une technique de traitement de signal, avec laquelle on a séparé les signaux captés mélangés par des capteurs pour retrouver les signaux originaux sources.

CHAPITRE 2

GENERALITES SUR L'EEG

2.1 Electroencéphalographie

2.1.1 Présentation

L'électroencéphalographie, est un procédé permettant d'obtenir un enregistrement de l'activité électrique (bioélectrique) spontanée du cerveau au moyen d'électrodes posées sur le cuir chevelu. Ce qui est enregistré est une différence de potentiel entre les deux électrodes, due à des mouvements d'ions électriquement chargés qui rentrent et sortent des neurones constituant le cortex cérébral.

Le résultat, l'électroencéphalogramme (EEG) représenté dans la figure 2. 1, se présente comme une série de tracés (courbes, dérivations), les uns au-dessous des autres, correspondant chacun à certaines des électrodes, et donc à la région du cerveau. Chaque tracé est à peu près sinusoïdal et comprend une succession d'ondes de forme variable (pointue, arrondie, large, étroite, etc.) et plus ou moins complexe (par exemple, une pointe vers le haut suivie d'une pointe vers le bas, dans certaines maladies) [17].

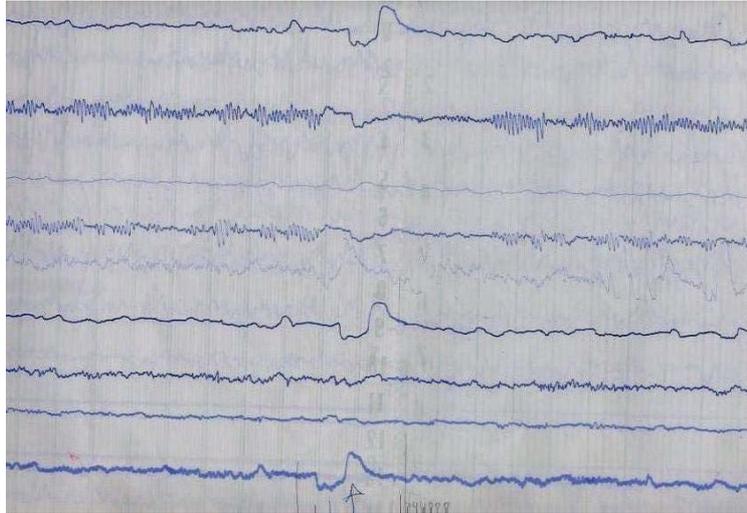


Figure 2. 1. L'électroencéphalogramme.

2. 1. 2 Électroencéphalogramme normal

Chez l'adulte en état de veille mais au repos (allongé, les yeux fermés), le profil normal d'un électroencéphalogramme est en rythme alpha. Si les ondes alpha sont nettement dominantes, bien visibles, et réparties sur presque toutes les dérivation (sur la majeure partie du cortex), on dit que le tracé est synchronisé. Sinon, il est désynchronisé, ce qui peut être pathologique. Quand le sujet ouvre les yeux, il se produit une réaction d'arrêt : le rythme alpha est plus ou moins remplacé par le rythme bêta.

Les rythmes delta et thêta s'observent surtout chez l'enfant. Certains pensent que l'onde thêta, inconstante chez les adultes, reflète chez eux l'activité créatrice. Le tracé se modifie considérablement au cours du sommeil.

2. 1. 3 les ondes cérébrales

Au niveau du cerveau humain, l'activité électrique des millions de neurones localisés principalement dans le cortex, peuvent être observés avec l'électroencéphalographie "EEG". Ce dernier est un dispositif qui enregistre l'activité des cellules du cerveau pour plusieurs états d'une personne, de l'état éveillé à l'état du sommeil profond.

Les cellules nerveuses présentent des différences de potentiels électriques en relation avec le liquide dans lequel elles sont submergées. L'action potentielle se rapporte à une fluctuation brève de la charge électrique dans la membrane du neurone, causée par

l'ouverture et fermeture rapide des canaux ioniques, qui dépend de la tension (flux ionique).

L'influx nerveux se propage comme une onde à travers les axones des neurones, transférant l'information d'une place à une autre dans le système nerveux. Une onde peut être d'amplitude et de fréquence variable selon l'état du patient, en tenant compte que chaque amplitude des signaux *EEG* change approximativement entre -100 et +100 microvolts.

Les signaux électrochimiques dans le cerveau humain sont la base de transmission entre les cellules de nerf, ces signaux peuvent être enregistrés au niveau du cuir chevelu. Les premiers enregistrements de ces signaux ont été édités en 1924 par *Hans Berger* lors de la découverte du premier *Electroencéphalogramme*.

L'électroencéphalogramme enregistré au niveau du cuir chevelu représente seulement une mesure brute de l'activité d'un grand nombre de neurones, ces enregistrements ont été divisés en plusieurs rythmes, et cette configuration est basée sur la bande de fréquences pour pouvoir distinguer les différents signaux *EEG*. Ces rythmes sont cités ci-dessous avec leurs spécificités :

- ✘ **Le rythme Alpha** à 8-13 hertz, a l'amplitude relativement grande et les fréquences modérées.
- ✘ **Le rythme Bêta** à 14-30 hertz, c'est le résultat de l'activité mentale intensive, montre typiquement des oscillations rapides avec de petites amplitudes.
- ✘ **Le rythme Thêta** à 4-7 hertz, qui peut accompagner des sentiments d'émotions fortes, est caractérisé par des fréquences modérément basses.
- ✘ **Le rythme Delta** à 1-5 hertz, c'est une oscillation à très basse fréquence qui se produit pendant des périodes de sommeil profond.

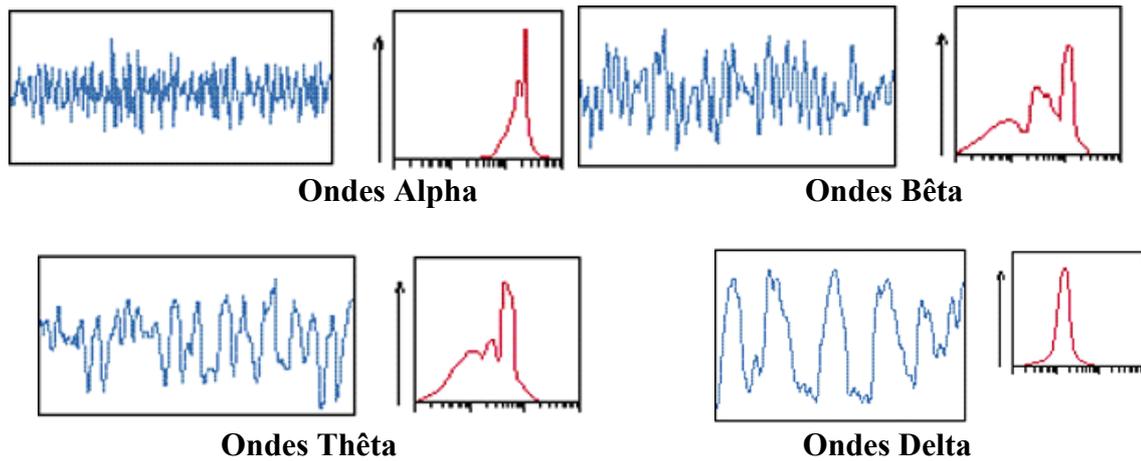


Figure 2. 2. Echantillons des ondes cérébrales.

2. 2. Utilisation de l'Électroencéphalographie

2. 2. 1 Diagnostique

L'utilisation médicale principale de l'électroencéphalographie a pour but d'aider au diagnostic et au suivi de l'épilepsie. Le tracé de l'EEG permet d'identifier le site spécifique où d'éventuelles lésions cérébrales épileptogènes ont eu lieu. Il peut aussi servir de méthode diagnostique d'appoint lors d'affections diffuses de l'encéphale (maladie de Creutzfeldt-Jakob). Une autre indication rare est l'enregistrement de troubles du sommeil ou d'anomalies survenant au cours du sommeil, telles que le syndrome des apnées du sommeil. En fait, on parle de polysomnographie, car on enregistre en même temps d'autres données, par exemple sur la respiration.

En revanche, toutes les indications qui faisaient partie naguère de la routine, comme le diagnostic des tumeurs du cerveau, sont désuètes du fait des progrès des techniques d'imagerie médicale (scanographie, par exemple).

L'obtention d'un EEG plat, dépourvu d'ondes, chez une personne plongée dans le coma, est considérée comme une absence de fonctionnement cérébral. C'est l'une des preuves légales de la mort, nécessaire, par exemple, pour effectuer un prélèvement d'organe en vue d'une greffe.

2. 2. 2 Montages et placements des électrodes

Pour pouvoir enregistré les ondes illustrés dans la figure 2. 1, on a besoin d'utiliser les électrodes qui seront fixés au cuir chevelu à l'aide d'un gel, l'emplacement de ces électrodes se fait en général selon le système international '10/20' (*Annexe. A*) développé par H. Jasper (1958). Dans ce système chaque électrode est placée en terme de sa proximité aux régions particulières du cerveau - Frontales, Centrales, Temporales, Pariétales et à l'Occipital -; le problème principal qui se pose dans le placement des électrodes, est les résistances qui se forment entre ces derniers et le cuir chevelu, elles doivent être les plus petites possible, de l'ordre de quelque $K\Omega$.

La figure 2. 3 montre ce montage pour 19 électrodes, chaque électrode à un descripteur alphanumérique qui lui est assigné. Chaque descripteur se compose d'une lettre qui se rapporte à la zone du cerveau : le "O" pour Occipital, "C" pour la bande Centrale, "T" pour Temporale et "F" pour Frontale ; alors que le composant numérique indique l'orientation du devant à l'arrière et la médiane indiquant à l'extérieur une valeur croissante, comme avec le T3 et le T5 plus postérieur ou le F3 et le F7 plus transversal. Tous les chiffres pairs sont du côté droit et les numéros impairs sont du côté gauche. La médiane est indiquée avec une deuxième lettre plus petite "z".

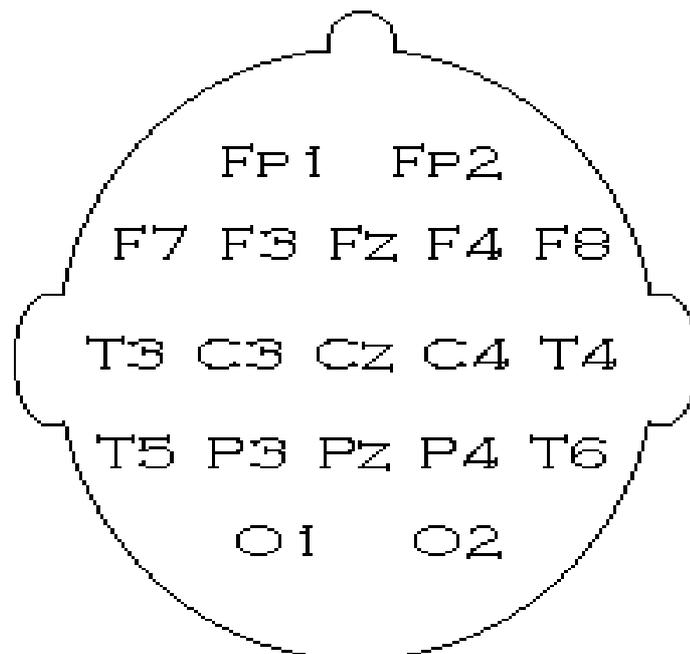


Figure 2. 3. Placements des électrodes selon le système international '10/20'.

2.3 Recherche médicale

L'étude de l'EEG a contribué aux recherches sur la nature et sur la physiologie du sommeil en permettant d'en identifier les différentes étapes. Les spécialistes ont pu aussi analyser les ondes cérébrales engendrées par des stimulations sensorielles comme un éclair lumineux ou un son ; c'est l'une des méthodes grâce auxquelles ils ont pu déterminer les fonctions de chaque partie du cerveau.

En pathologie, l'électroencéphalogramme permet de faire des recherches sur les troubles du sommeil. On peut non seulement mieux comprendre et mieux classer ces affections, mais aussi observer les effets sur le sommeil de chaque type de médicament.

2.3.1 Techniques voisines et perspectives

L'électroencéphalographie quantitative fait analyser le résultat par un ordinateur, ce qui augmente considérablement l'efficacité de la méthode.

La technique des potentiels évoqués consiste à provoquer une brève stimulation sensorielle (un éclair lumineux, un son), puis à enregistrer une courbe selon le même principe que l'électroencéphalographie, mais seulement en regard de la petite région du cerveau correspondante. On obtient ainsi le tracé d'un potentiel évoqué (c'est-à-dire provoqué) visuel, auditif, ou sensoriel. L'apport diagnostique est très intéressant dans certaines affections sensorielles ou sensibles, et chez l'enfant trop jeune pour coopérer dans le cadre d'autres techniques (il ne peut pas dire, par exemple, s'il entend ou non un son).

La magnétoencéphalographie est une technique encore à l'étude pour enregistrer l'activité magnétique du cerveau (liée à son activité électrique). Elle permettrait notamment de saisir des phénomènes plus profondément situés qu'avec l'électroencéphalogramme [17].

2.3.2 Représentation cartographique

Il y a eu dans ces dernières années un très grand développement en ce qui concerne la représentation graphique de l'électroencéphalogramme, puisque il était auparavant limité aux traçage de ce dernier sur du papier, ce qui est le cas dans la plupart de nos hôpitaux. Les nouvelles techniques de visualisations de ces signaux ont vu l'introduction de

l'ordinateur, comme un moyen d'ouverture des possibilités infinies pour extraire des informations utiles et pour travailler plus aisément.

Le signal enregistré par l'EEG doit subir un traitement adéquat pour pouvoir le visualiser sur ordinateur, les éléments essentiels d'un système de cartographie d'EEG ont les rôles suivants :

1. Acquisition des signaux EEG ;
2. Traitement des données ;
3. Stockage des informations ;
4. représentation des résultats.

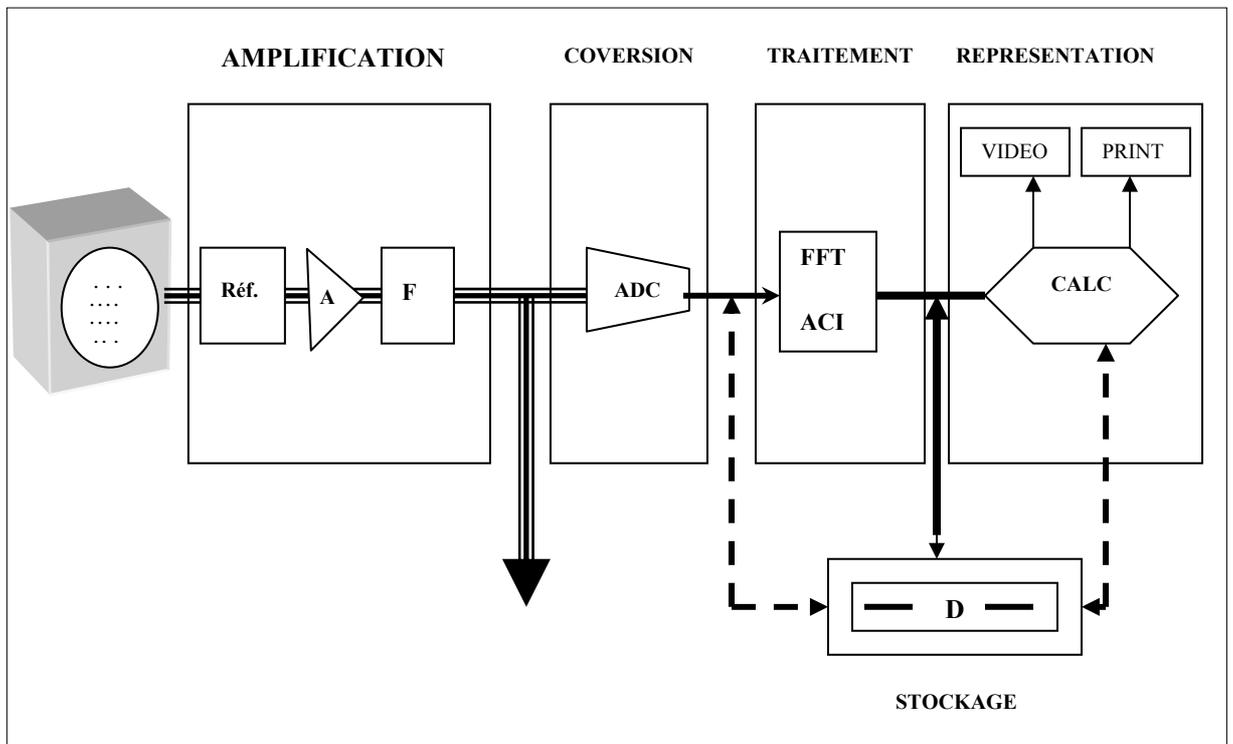


Figure 2. 4. - Structure générale d'un équipement de cartographie.

La figure 2. 4 décrit les différentes étapes conduisant à la représentation cartographique.

2.3.2.1 Amplification

Les signaux recueillis au niveau des électrodes sont tellement faible qu'il leur faut une amplification, cette opération est suivi d'un filtrage qui permet d'adapter la bande passante du signal à la fréquence d'échantillonnage du convertisseur. A l'entrée du convertisseur, le signal ne doit pas présenter de composantes fréquentielles supérieures à la moitié de la fréquence d'échantillonnage, sinon celles-ci seraient « repliées » et parasiteraient le spectre final.

2.3.2.2 Conversion analogique – numérique

La fréquence d'échantillonnage doit être au moins égale au double de la fréquence maximale analysée.

2.3.2.3 Traitement des données

Il existe plusieurs traitements pour les données enregistrées par exemple la *FFT*, ou bien l'Analyse en Composantes Indépendantes *ACI* connue par son abréviation en anglais « *ICA* », cette dernière a été introduite par C. Jutten et J. Héroult et développé par P. Comon.

2.3.2.4 Représentation des résultats

Une fois l'acquisition faite, les données seront stockées dans un fichier. En utilisant des algorithmes adéquats on pourrait visualiser les signaux sur écran ou les imprimés [29].

Le développement de la représentation des signaux EEG ne cesse d'évoluer. Dans ce qui suit on mentionnera quelques exemples de cette évolution :

Analyse spectrale

Cette méthode consiste à visualiser les signaux *EEG* en trois dimensions, c'est-à-dire les pics de tensions vont être plus visibles, donc un meilleur diagnostique pour les maladies telle que l'Epilepsie et bien d'autres maladies.

Dans cette représentation, on peut visualiser les pics de tension, on peut ajouter une barre pour différencier les amplitudes qui seront en couleurs différentes chacune selon son amplitude. La figure 2.5 représente cette technique :

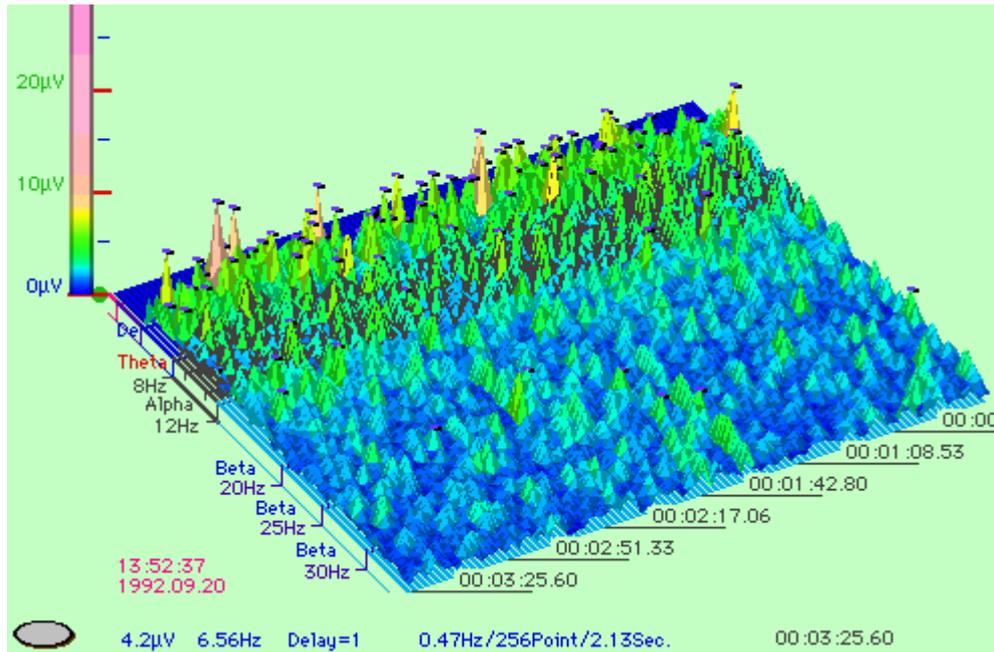


Figure 2. 5. Analyse spectrale

Topographie du cerveau

Cette technique a été développée vers la fin des années 80, en utilisant plusieurs électrodes placées sur la tête. A l'aide d'un logiciel spécial à l'intérieur de l'ordinateur, qui permet de tracer l'activité cérébrale sur un écran, en utilisant les couleurs pour spécifier l'amplitude des ondes enregistrées (par exemple, le noir et le bleu représenteraient les basses amplitudes du signal *EEG*, alors que le jaune et le rouge représenteraient les amplitudes élevées).

Les points spatiaux se trouvant entre les électrodes sont calculés par des techniques mathématiques d'interpolation entre les électrodes voisines, et une gradation de couleurs douce est ainsi réalisée.

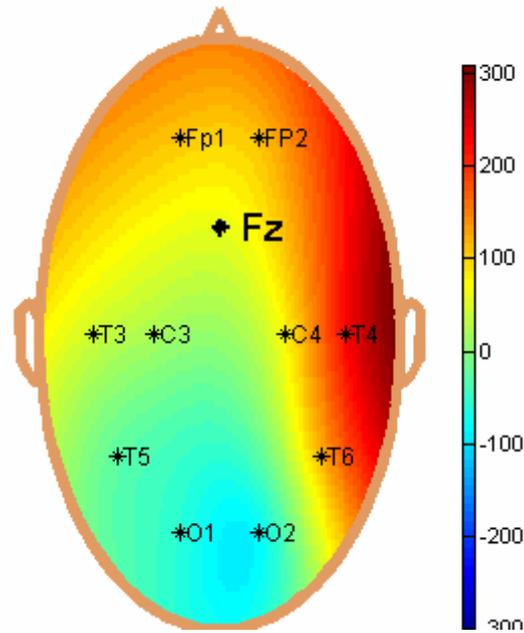


Figure 2. 6. Carte topographique du *EEG*

Une amélioration a été faite sur la représentation topographique du cerveau, c'est la reconstruction tridimensionnelle de la tête et du cerveau, en plus des animations, montrant les changements d'activité électrique en fonction du temps.

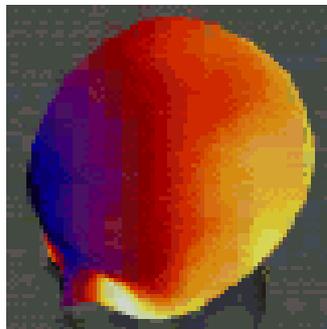
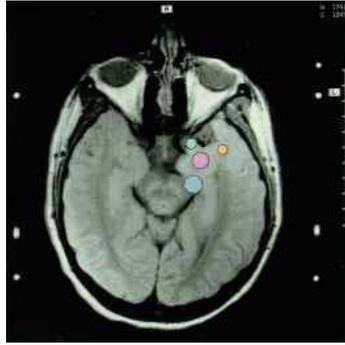


Figure 2. 7. Représentation en trois dimensions.

La représentation des données ne cesse de se développer, le futur de ces représentation est une méthode basée sur le couplage entre le traitement d'image et le traitement du signal numérique; la combinaison de deux procédés le scanner *MRI* qui donne une image du cerveau avec un grand détail et une clarté remarquable et le Magnétoencéphalographe *MEG*. Ce qui donne le résultat représenté dans la figure suivante :



Alors les deux images celles du MRI et du MEG sont superposées. Les points d'activité électrique altérée (Zones épileptique) sont schématisés en cercle, en plus de l'image du *MRI* on aura l'emplacement exact dans le cerveau, comme le montre la figure ci-dessus [30].

CHAPITRE 3

CARTES D'ACQUISITION ET DE CONVERSION

3.1 Introduction

L'étude du cerveau humain ne cesse de se développer, vu l'importance de cet organe dans notre vie ainsi que notre quotidien. Ce développement a débuté en 1929 lors de la découverte du premier appareil *EEG* (*Electroencéphalophe*) par *Hans Berger*, jusqu'au développement actuel de cet appareil ainsi que les nouvelles techniques d'étude du cerveau humain.

L' Electroencéphalophe n'a pas échappé à l'informatisation bien sûre; d'où la possibilité de traiter ces signaux *EEG* aisément et de plusieurs façons.

Pour pouvoir visualiser les signaux *EEG* captés via l'électroencéphalophe au niveau de l'hôpital, on a du utiliser une carte de conversion analogique /numérique de seize signaux réalisée par des étudiantes en *PFE* [27], et pour gérer ces données on a utilisé une carte d'interface utilisant le *PPI 8255*; cette dernière sera branchée sur le bus *ISA* du *PC*. Les deux cartes nous permettront de communiquer avec l'*EEG* avec un transfert parallèle des données numériques.

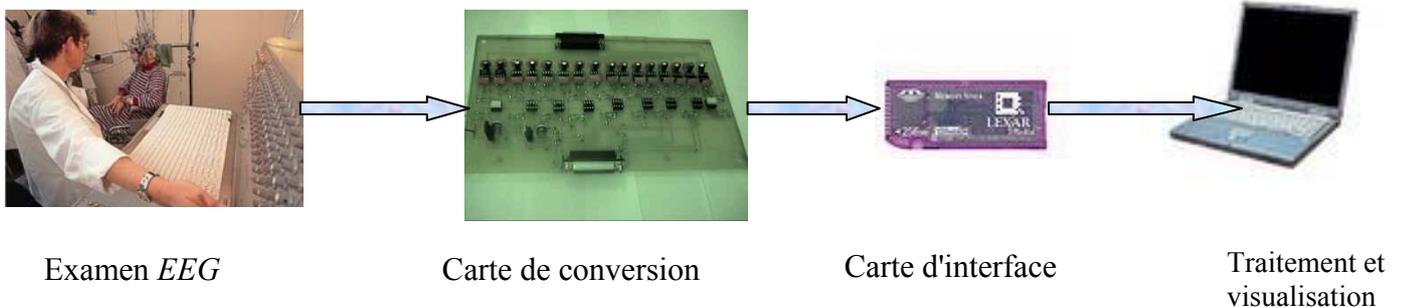


Figure 3. 1. Synoptique du système.

3.2 Carte de Conversion

Cette carte permet de convertir seize signaux à la fois en utilisant les convertisseurs *MCP 3301*, c'est des convertisseurs analogiques numériques bipolaires à treize bits.

Ce convertisseur opère sur une seule alimentation de 5 *Volts*, possède une entrée différentielle $V_{in (+)}$ et $V_{in (-)}$ dont l'amplitude maximale est de 5 *Volts* et utilise une tension de référence de 2.5 *Volts*.

Le procédé de conversion est conduit par une horloge externe et doit inclure treize cycles d'horloge un pour chaque bit. La conversion commence si le \overline{CS} est au niveau bas. Les treize cycles d'horloge font ressortir le résultat de la conversion avec le bit signe en premier puis les 12 bits de données restants. [27].

Après le bloc de conversion une isolation galvanique a été mise en œuvre par des optocoupleurs ou photocoupleurs (*Annexe. C*) pour assurer l'isolement électrique donc de prévenir des perturbations de fonctionnement des équipements et matériel, vue que cet optocoupleur est un dispositif qui transmet des informations logiques ou analogiques sous formes de signal électrique via une voie optique qui isole électriquement l'entrée de la sortie.

La logique de contrôle des 16 convertisseurs sera gérée par le microprocesseur via le port *C* du *PPI* de la carte d'interface; les deux bits du quarté inférieur du port *C* sont utilisés pour commander respectivement l'horloge et le chip select (\overline{CS}). Tandis que l'acheminement des données de la conversion vers le *PPI* se fera à travers le port *A* et *B* de la carte d'interface. [27]

Les échanges d'informations entre le microprocesseur et la carte de conversion se feront à l'aide d'un programme d'acquisition qui sera détaillé par la suite.

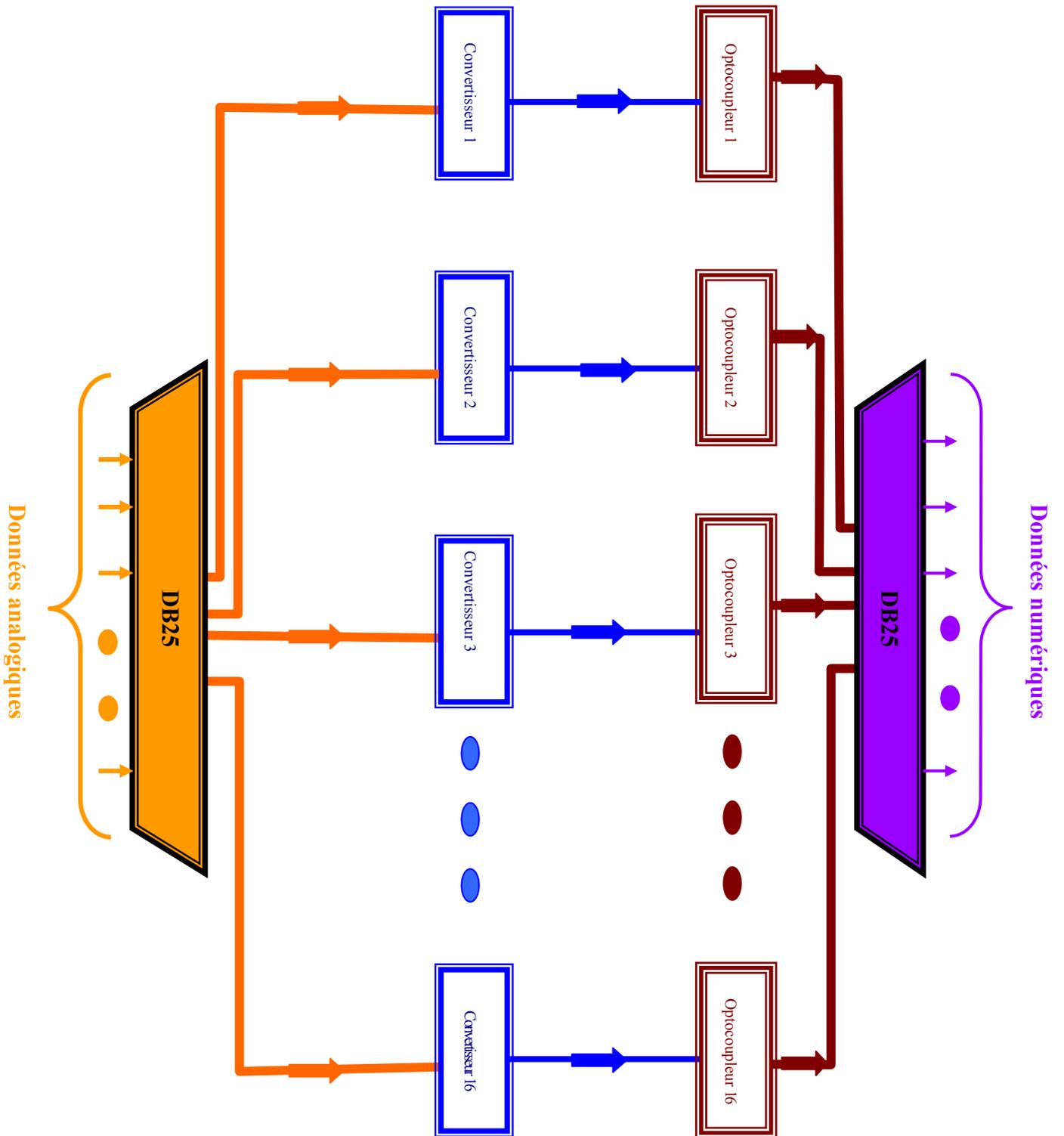


Figure 3. 2. Schéma synoptique de la Carte de conversion

3.3 Carte d'Interface

Cette carte est utilisée pour permettre l'échange des informations entre le microprocesseur et la carte de conversion décrite dans le paragraphe précédent. Cette carte a été conçue et réalisée sur le bus *ISA* du *PC*. Cette dernière intégrera le *PPI 8255A* d'*Intel* pour assurer la gestion de données entre le *PC* et la carte de conversion analogique numérique.

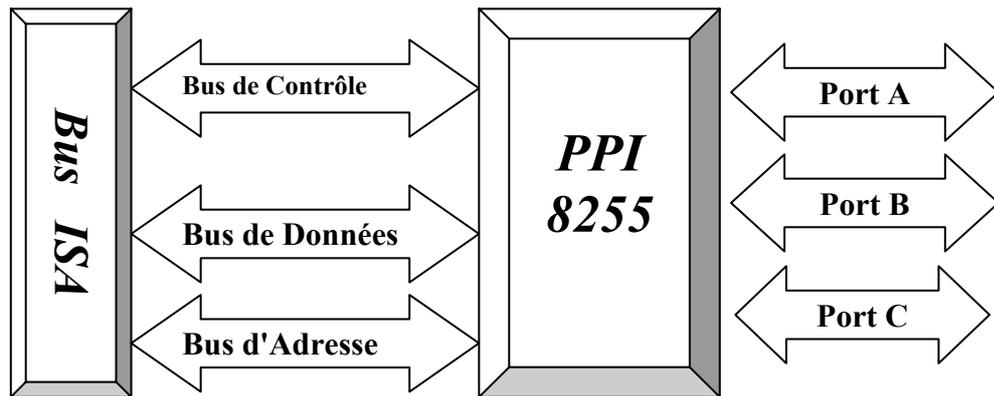


Figure 3. 3. Schéma synoptique de la carte d'interface

Le *PPI8255* (l'interface programmable parallèle) est un circuit d'interface parallèle qui assure l'interface entre des périphériques et des microprocesseurs de la famille *Intel* (8080, 8085, 8086). Il communique avec le périphérique à travers 24 bits :

- ✘ Deux groupes de 8 lignes (port *A*, port *B*) utilisés seulement pour le transfert d'information.
- ✘ Deux groupes de 4 lignes (port *C* utilisé pour le transfert d'information ainsi que le contrôle de ce transfert).

Le *PPI8255* dispose de 3 modes de fonctionnement sélectionnés par le contenu du registre de contrôle, le **Mode 0** le **mode1** et le **Mode 2**. Pour la programmation elle s'effectue en envoyant dans le registre correspondant le mot de commande approprié. La composition de ce mot, placé par le microprocesseur sur le bus de données (*D0 – D7*). L'adressage du registre de commande est assuré sur 2 bits *A0* et *A1* qui combinés avec les signaux \overline{RD} , \overline{WR} et \overline{CS} , en assurant différentes opérations.

3. 4 Acquisition des signaux EEG

Les échanges d'informations entre le microprocesseur et la carte de conversion se feront à l'aide d'un programme d'acquisition.

Ce programme a été développé sous *C++ BUILDER*, permettant une acquisition de seize signaux simultanément via les cartes de conversion et d'interface, la fréquence d'échantillonnage pour l'acquisition des données est égale à 1,042 KHz (*Annexe. E*). Cette acquisition nous a permis de visualiser des signaux à des fréquences arrivant jusqu'à 130 Hz.

3. 4. 1 Représentation de l'organigramme principal de l'acquisition

L'algorithme présenté dans la figure 3. 4. représente la schématisation du programme principal d'acquisition. Après la conversion et le transfert des données vers le PC via le PPI à travers les deux ports *A* et *B* ayant huit lignes chacun, permettant ainsi d'acquérir nos seize données simultanément.

L'algorithme dans la figure 3. 4. s'applique à l'acquisition du port A et le port B du PPI. Chaque ligne du port A et du port B, nous fournit les bits d'une donnée qui est constitué de treize bits dont un est le bit signe. Le déroulement de cette acquisition se déroule en effectuant tout d'abord la lecture des bits signes de toutes les données, ensuite faire la lecture des douze bits de chaque donnée en faisant un décalage à gauche à chaque lecture de nouveau bit, à la fin de la lecture des seize données, on teste le signe de celles-ci, par la suite on les traces, ce qui nous permet d'avoir un affichage en temps réel, et enfin on procède à l'écriture de ces données numériques dans un fichier de données.

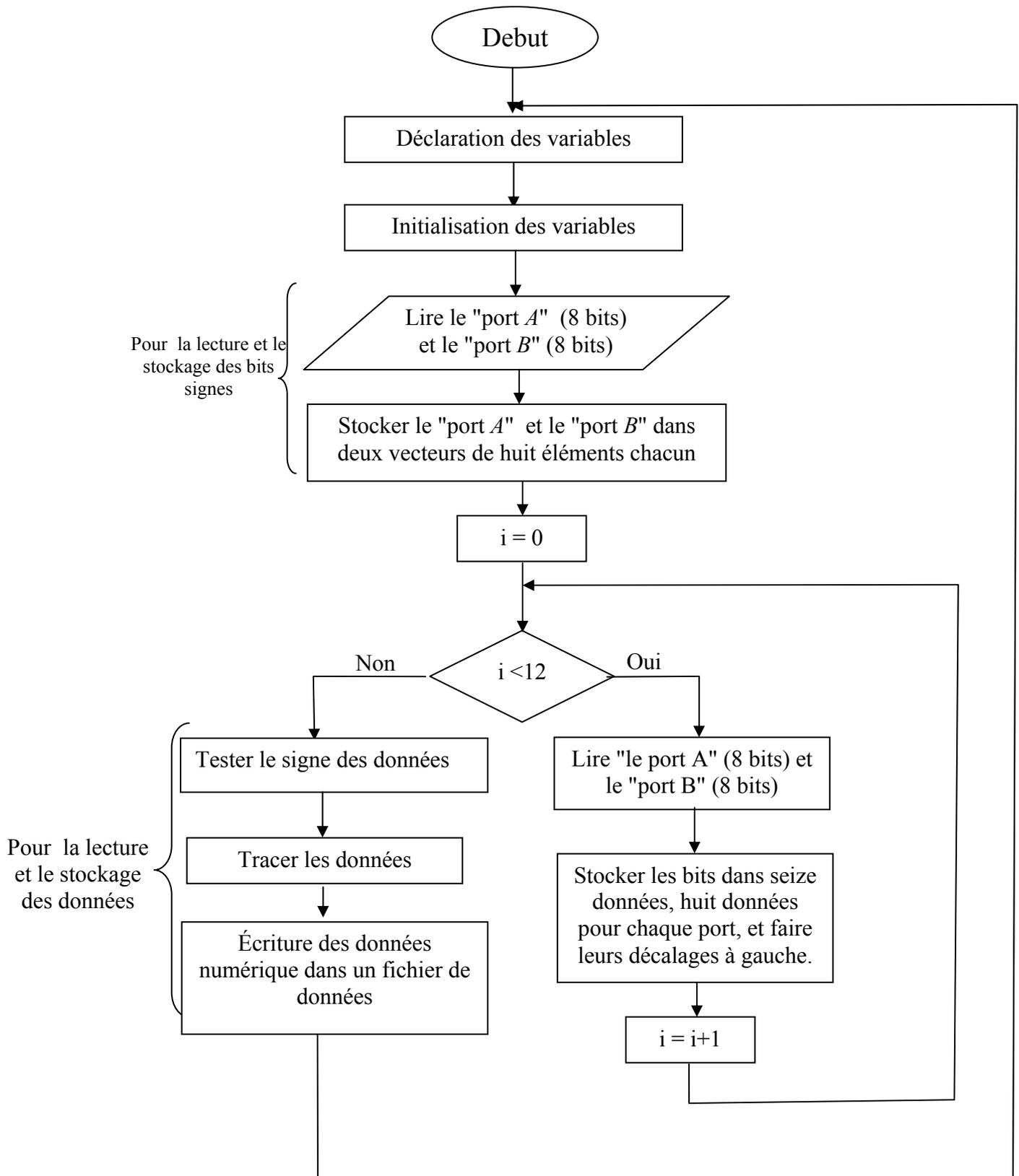


Figure 3. 4. Organigramme principal d'acquisition

3.5 Conclusion

Actuellement, la simulation a gagné beaucoup de place dans le terrain des recherches effectués, d'où l'intérêt d'utiliser la partie matérielle dans notre projet, où l'utilisation des résultats obtenus par cette acquisition sont plus crédibles que d'utiliser dans le cas de notre application des signaux apportés directement *d'Internet*. Cette acquisition nous permettra de stocker les signaux *EEG* numérisés sur le disque dur.

CHAPITRE 4

TESTS ET RESULTATS

4.1 Introduction

Dans ce chapitre nous présenterons les résultats obtenus des expériences réalisées au niveau de l'hôpital "*CHU de Blida secteur Frantz Fanon*". En premier lieu, on a dû faire une première approche pour un examen *EEG* sur un homme âgé de trente six ans, et ceci pour confirmer le bon fonctionnement de notre matériel et logiciels, en comparant les signaux enregistrés à l'ordinateur avec ceux tracés sur du papier. Ces signaux ont été séparés et l'activité de ces signaux séparés a été représenté dans le plan et dans l'espace; par la suite nous présenterons le logiciel multi-tâches *SEPAREEG version 1.0* réalisé pour avoir une manipulation plus facile de nos programmes, et pour visualiser les différentes formes d'applications.

En deuxième lieu, le deuxième examen *EEG* a été fait sur un homme âgé de trente trois ans, cette fois l'expérience a été réalisée afin de séparer les artéfacts, tout en représentant leurs activités dans le plan et dans l'espace.

4.2 Domaine d'application et terminologie

Le domaine d'application de nos algorithmes est le traitement de l'électroencéphalogramme. Il s'agit de séparer les signaux sources émis par le cerveau humain, après avoir enregistré leur mélange au niveau du scalp, capté par des électrodes.

Après le recueillement des données numériques *EEG*, on a appliqué la technique *ACI* pour séparer ces signaux mais tout en débutant avec un blanchiment pour un prétraitement des données.

Après la séparation des signaux *EEG*, on obtiendra la matrice de blanchiment M ainsi que la matrice de séparation notée W , alors la matrice de séparation totale est le produit de la matrice de séparation W par la matrice de blanchiment M , pour plus de convivialité nous allons appelé la matrice de séparation totale, la matrice de séparation.

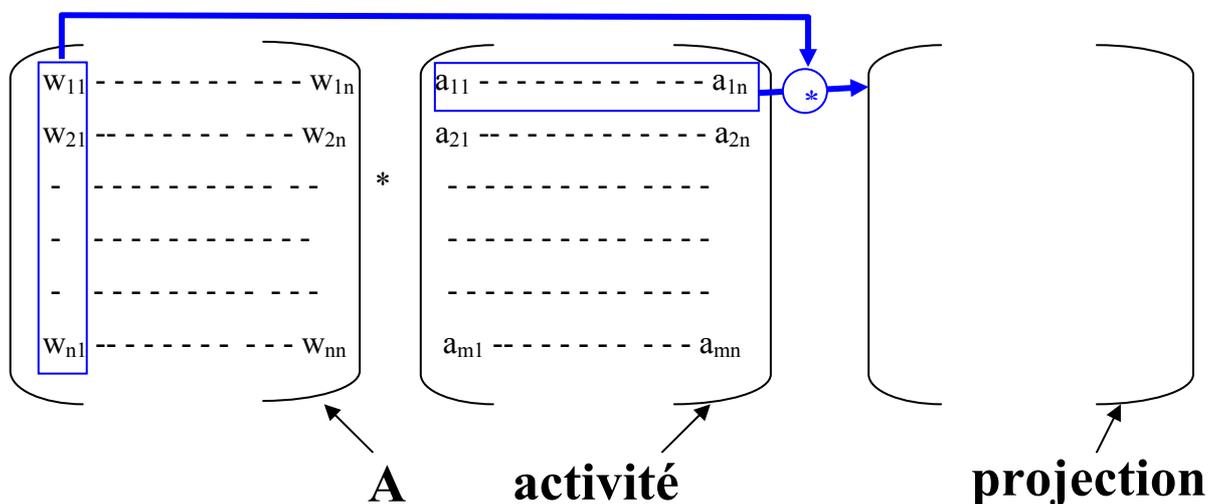
Pour obtenir l'activité des composantes indépendantes, on multiplie la matrice de séparation par les données mélangées, puisque les lignes de la matrice de séparation sont des filtres linéaires décomposant les données mélangées en activités des composantes indépendantes. Les lignes du produit de la matrice de séparation par les données représenteront les ondes d'activations des composantes *ACI*, en terme d'équation ceci veut dire [33] :

$$\text{activité} = W * \text{donnée} \quad (4.1)$$

Cependant les composantes *cartes du scalp* sont les colonnes de la matrice Mélange A . Donc chaque composante indépendante (composante *Infomax* ou *FastICA*) est constituée d'une activation des composantes *ACI* (qui représente une ligne de la matrice *activité* Eq. (4.1)) et la carte du scalp (colonne de la matrice mélange).

Finalement, la projection de la $i^{\text{ème}}$ composante indépendante est donnée par :

$$\text{projection} = \text{winv}(:,i) * \text{activité}(i,:) \quad (4.2)$$



La projection de la composante est dans l'espace de la donnée originale, c'est-à-dire dans les mêmes canaux (exemple. Fz, Cz, etc.) et les unités (μV) comme la donnée originale. Et puisque *ACI* est une décomposition linéaire, la projection de la somme des composantes est la même que la somme des projections de chaque composante.

Alors *ACI* distribue les informations entre les matrices poids (w_{inv}) et *activité*. [35].

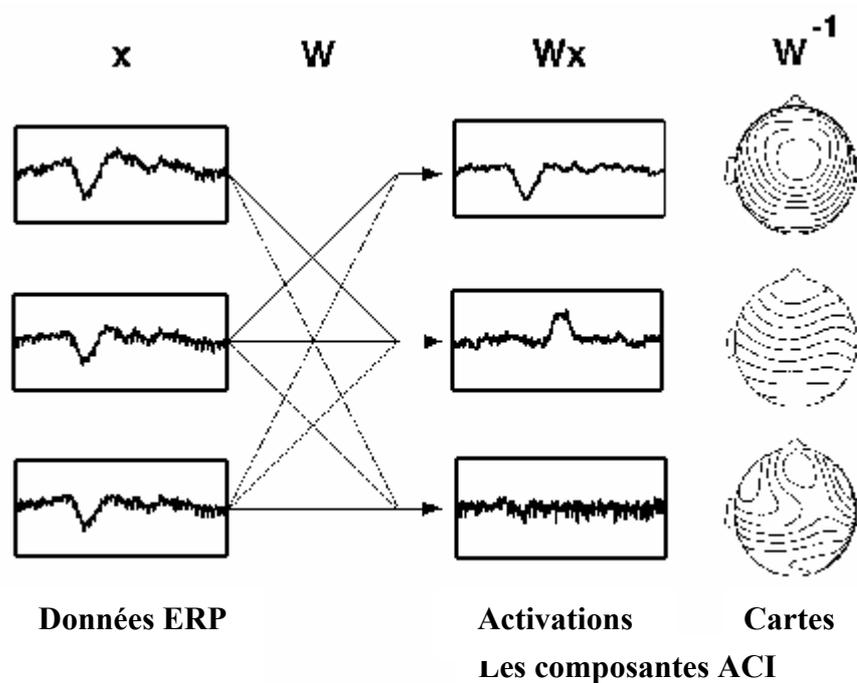


Figure 4. 1. Schéma de décomposition d'Analyse en Composantes Indépendantes de l'électroencéphalogramme.

4.3 Acquisition des données et premier examen *EEG*

Cette acquisition a été réalisée dans l'hôpital "*CHU de Blida secteur Frantz Fanon*" sur un homme âgé de 36 ans. La machine *EEG* disponible à l'hôpital, dispose de quatorze plumes de traçage dont seulement dix plumes marchent normalement ce qui nous mène à utiliser dix électrodes, et qui nous a permis d'acquérir dix signaux *EEG*. Le montage de ces électrodes sera expliqué dans le paragraphe suivant.

Le montage utilisé dans ce test est le montage référentiel. La figure 4. 2 montre ce montage proposé à l'hôpital. En utilisant dix électrodes (O1, O2, CZ, PZ, T3, T4, T5, T6, C3, C4) avec l'électrode FZ comme électrode de référence.

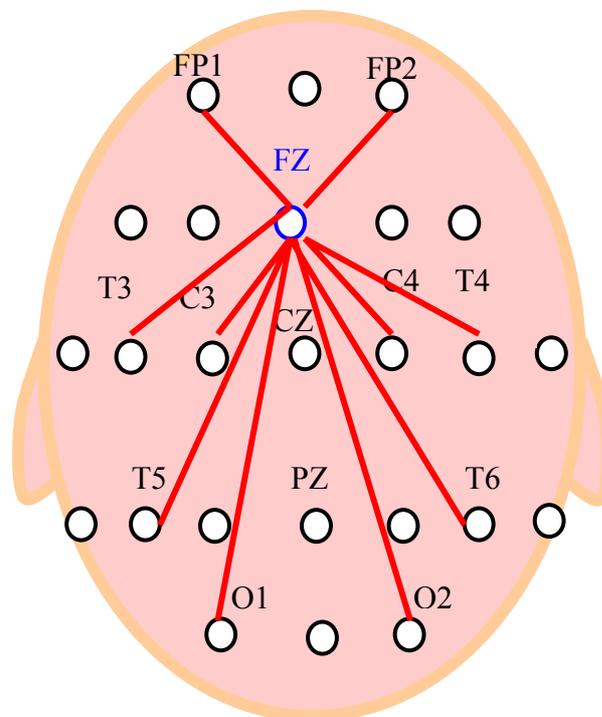


Figure 4. 2. Montage Référentiel

Pour réussir cette étape il a fallu mettre en œuvre une partie électronique (§ Chap. 3), et à l'aide d'un logiciel d'acquisition présenté dans la figure 4. 3. Ce logiciel nous a permis l'acquisition, l'affichage et l'enregistrement de données via des boutons bien appropriés avec une possibilité de changer l'échelle du plan du tracé, avec la possibilité de visualiser l'acquisition des seize signaux en temps réel, tout en enregistrant en même temps ces données visualisées.

La figure 4. 3 représente les signaux de calibrage de la machine *EEG* lors du réglage de cette dernière par le technicien du laboratoire de l'hôpital avant le commencement de nos tests *EEG*.

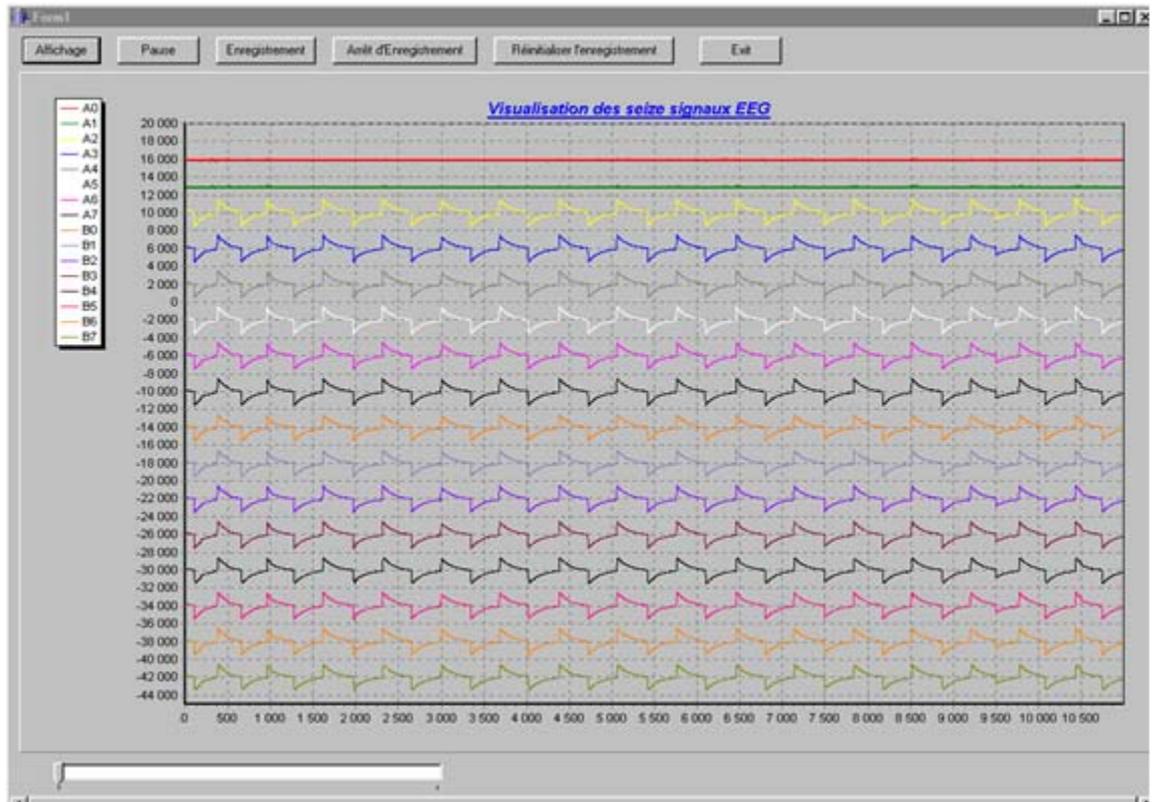


Figure 4. 3. Représentation du logiciel d'acquisition

Pour authentifier nos résultats obtenus par l'acquisition, nous avons prévu le traçage des signaux enregistrés sur du papier lors de notre test à l'hôpital. Une partie de ce tracé est représenté dans la figure 4. 4. B, cependant la figure 4. 4. A montre les signaux acquis sur l'écran via le logiciel d'acquisition. Cette partie du tracé choisie représente un clignement des yeux du sujet. Si on regarde la figure 4. 4. B on remarque qu'il n'existe que neuf signaux et un signal manque sur ce tracé du papier, ceci est dû à une défaillance mécanique au niveau d'une plume de traçage de l'appareil *EEG*. Par contre sur la figure 4. 4. A, on présente seize signaux dont les le onzième et le douzième sont non utilisés, et le reste du tracé sont ceux connectés à l'*EEG*. Notons que le signal manquant sur le tracé du papier est le signal tracé en blanc (la flèche), dans la figure 4. 4. A.

En comparant les deux figures présentées dans 4. 4. A et 4. 4. B, on constate que les signaux enregistrés sur le papier de la machine *EEG* sont identiques à ceux présentés dans le plan de notre présentation sur l'écran.

Cette comparaison nous permet de constater l'avantage de l'acquisition, car faute de plume, de papier ou même d'encre, les tests *EEG* ne se font pas, par contre avec notre application ceci est toujours possible bien sûr si l'appareil *EEG* est toujours en marche.

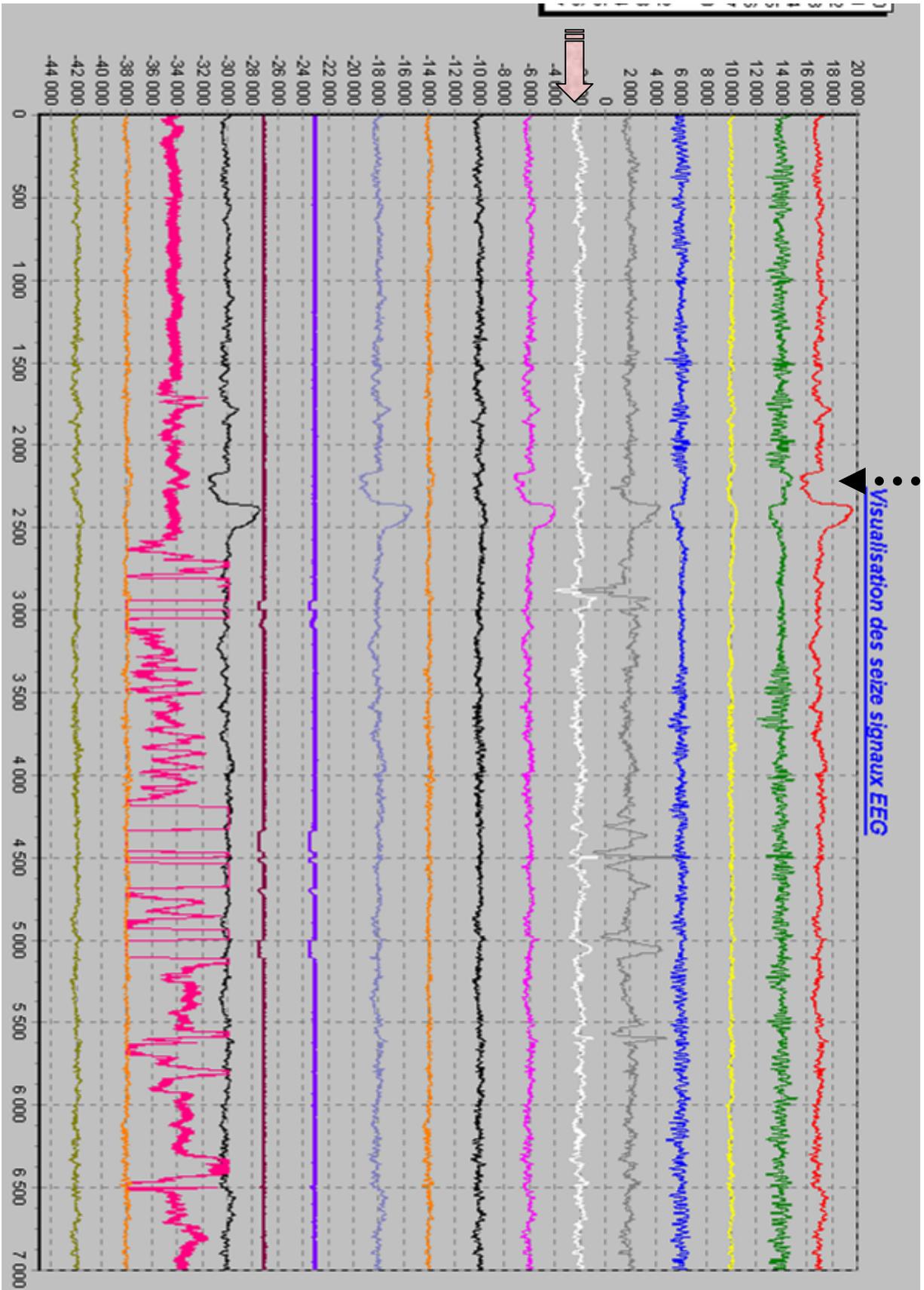


Figure 4. 4. A Représentations des signaux acquis sur l'écran

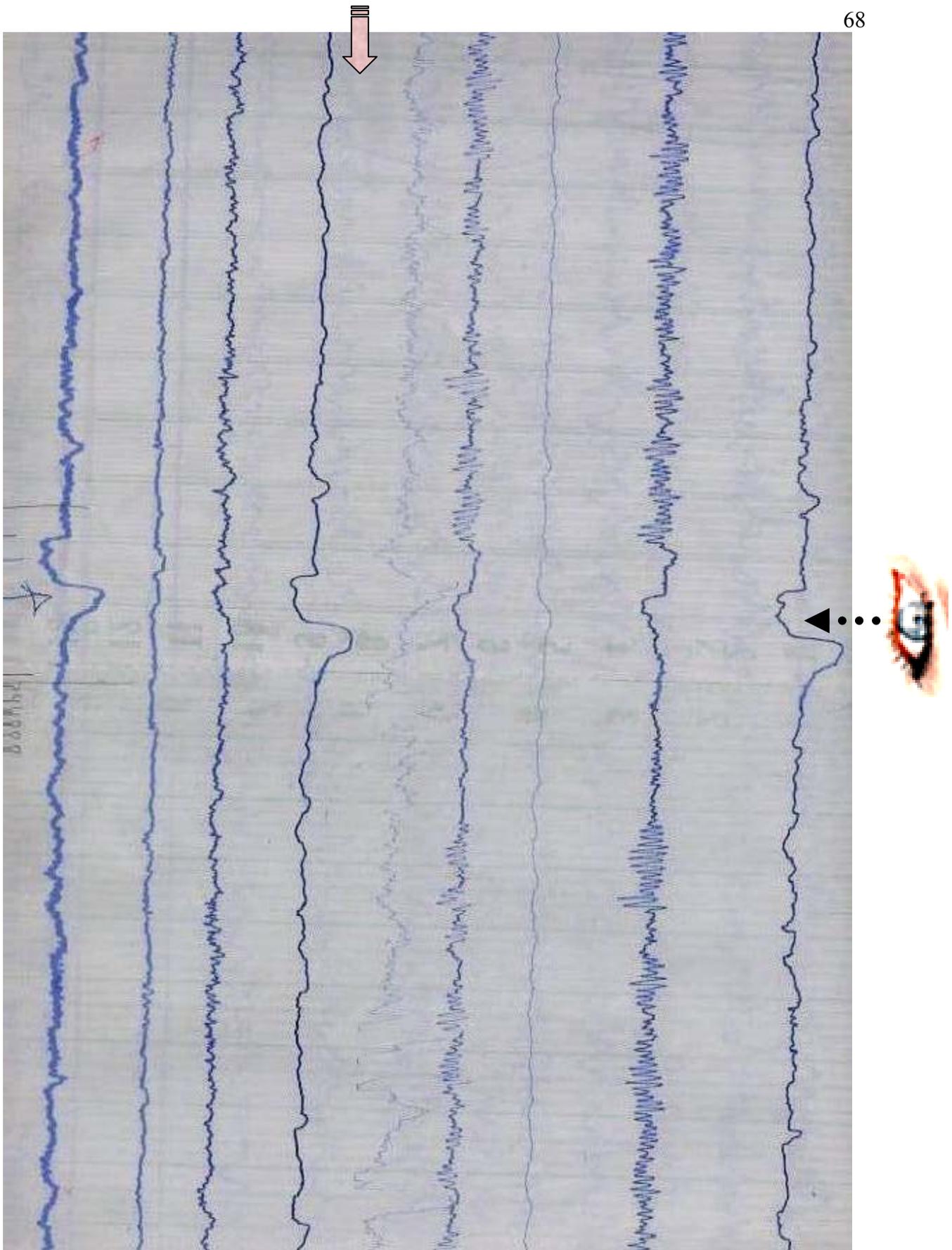


Figure 4. 4. B Représentations des signaux acquis sur le papier

4.3.1 Traitement des signaux *EEG* réels

4.3.1.1 Séparation des signaux *EEG* réels enregistrés

Dans cette partie, on présente nos résultats obtenus avec la séparation des signaux *EEG*, acquis lors du premier test *EEG*.

La figure 4.5 représente les signaux numérisés acquis; le segment choisi est un **clignotement du patient**. Les dix premiers signaux représentent les signaux *EEG* sources avec les descripteurs des électrodes respectives à gauche et le onzième signal représente la moyenne de ces dix signaux. Les signaux présentés dans la figure 4.5, ont été traités par la méthode d'Analyse en Composantes Indépendantes. Les résultats de la séparation de ces signaux sont présentés dans la figure 4.6, ces résultats sont obtenus en utilisant l'algorithme *FastICA* avec un blanchiment par l'algorithme *ACZ* et la fonction contraste *tght*.

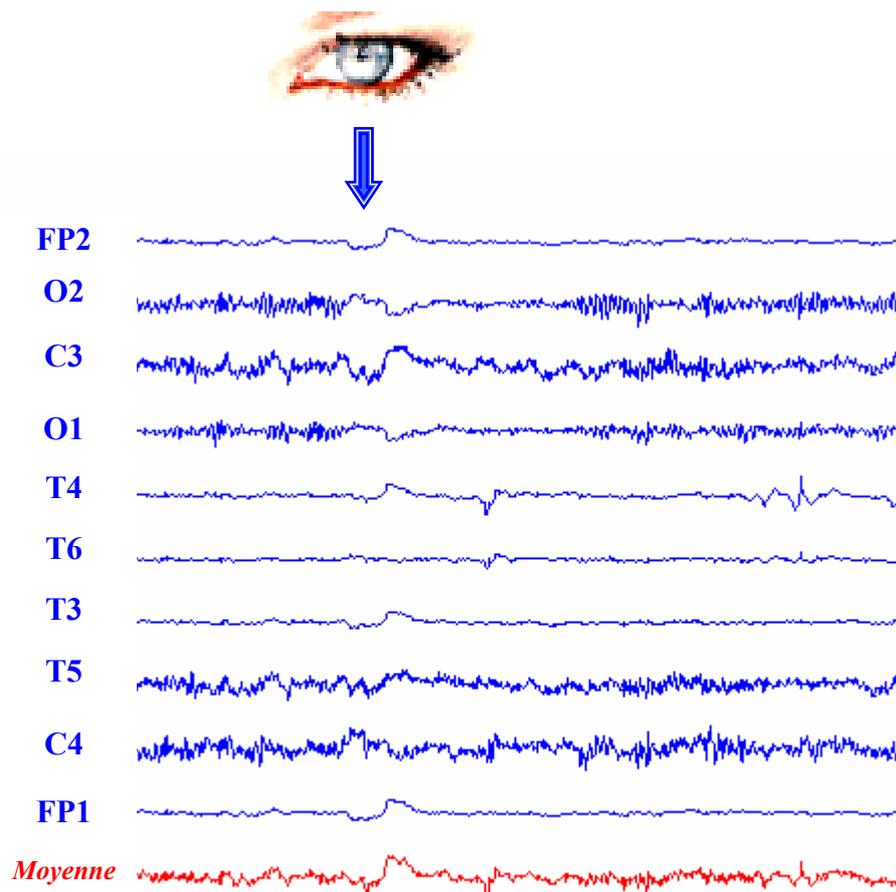


Figure 4.5. Représentations des signaux *EEG* réels.

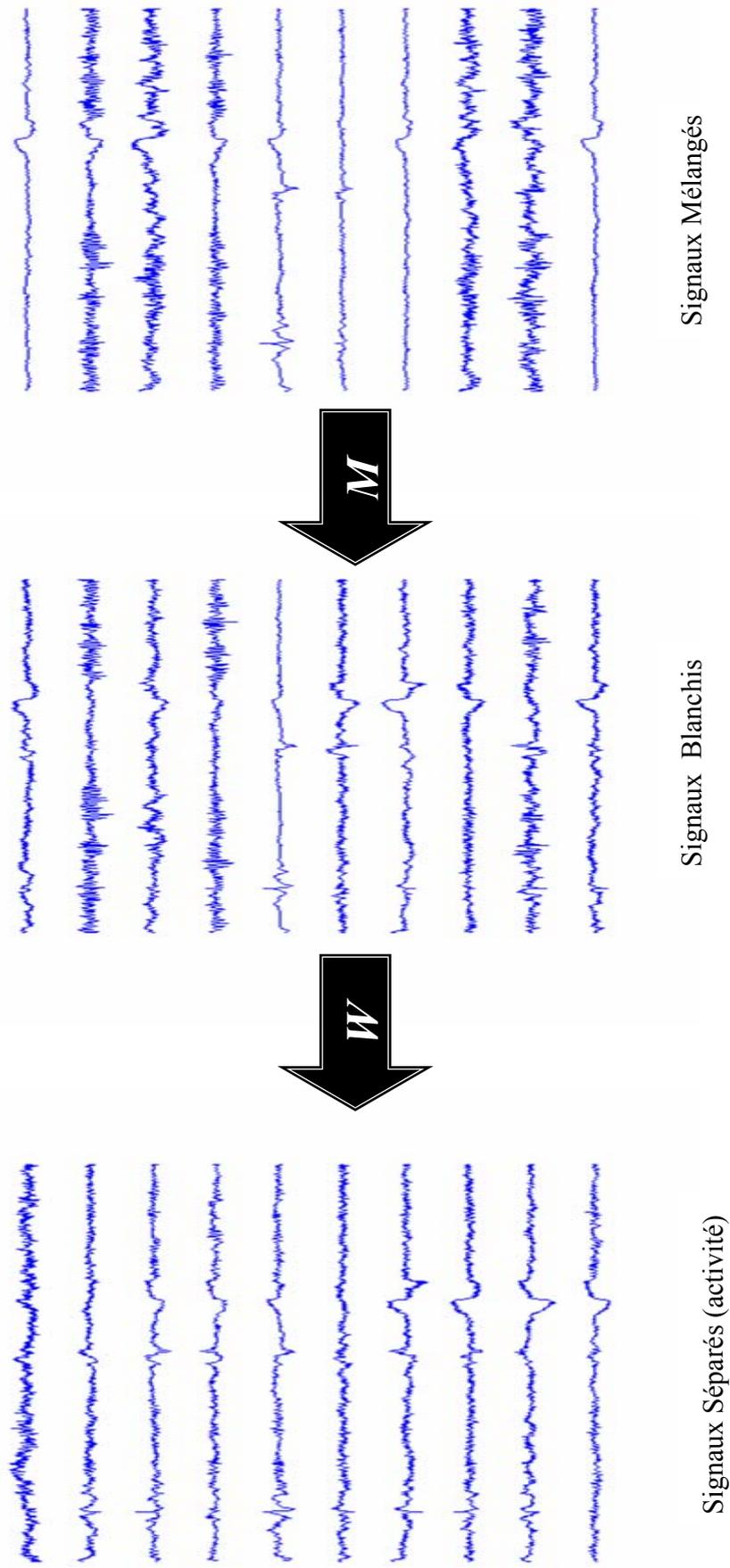


Figure 4. 6. Séparation des dix signaux réels EEG

La figure 4. 7 représente ce qui a été expliqué dans la § 4. 2. Alors chaque composante indépendante (composante Infomax ou FastICA) est constituée d'une activation des composantes ACI et qui représentent une ligne de la matrice W x et la carte du scalp qui représente une colonne de la matrice mélange.

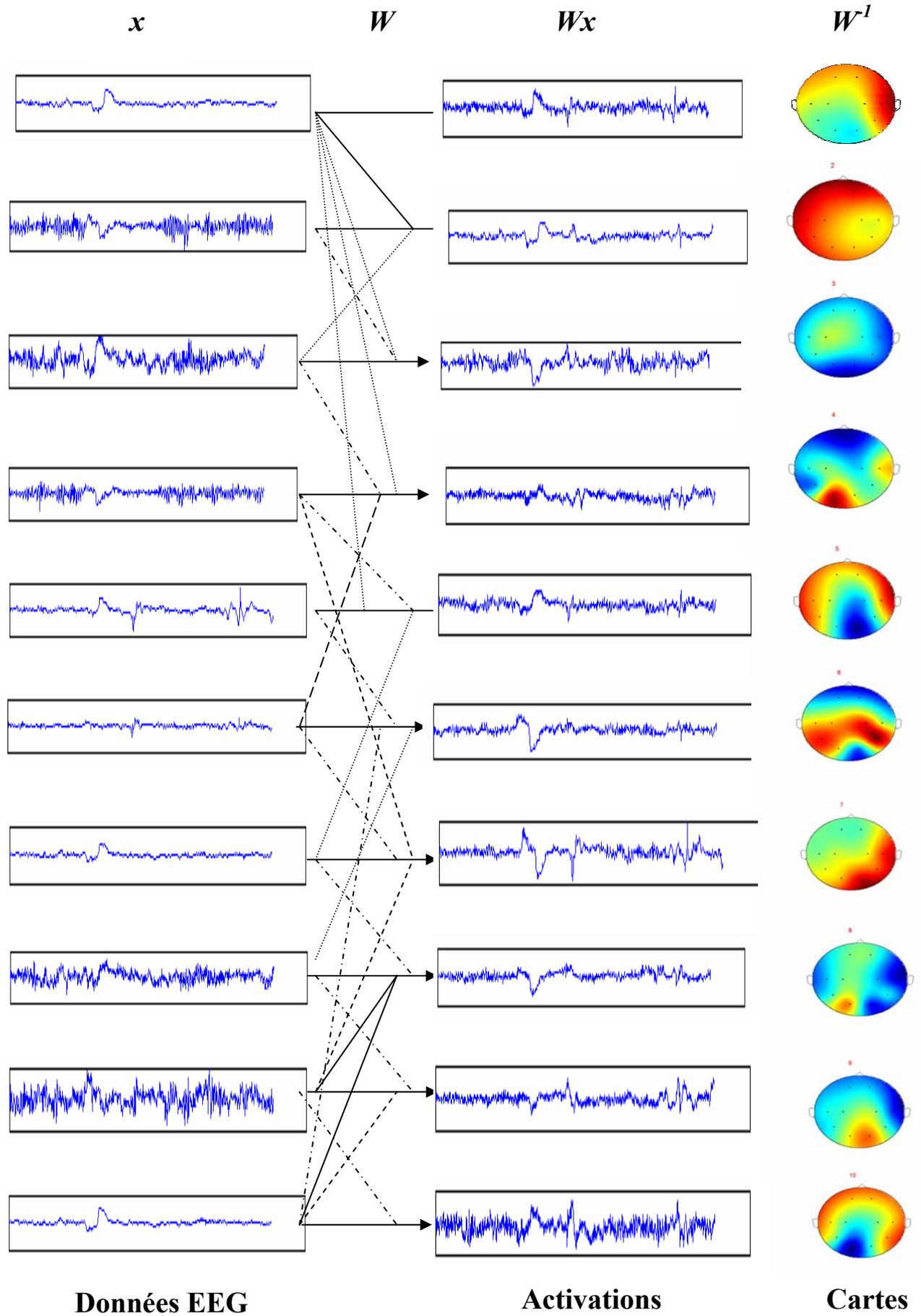


Figure 4. 7. Schéma de décomposition d'Analyse en Composantes Indépendantes de l'électroencéphalogramme.

IV. 3. 1. 2 Représentations des signaux EEG dans le plan et l'espace par le logiciel

a) Présentation du logiciel *SEPAREEG*

Pour permettre une manipulation conviviale pour cette séparation de sources ainsi que pour les différentes représentations des différents résultats, on a développé un logiciel en utilisant une programmation multi-tâches basée sur des interfaces graphiques développée sous *MATLAB*.

L'exécution de notre logiciel nous permet d'avoir en premier lieu le menu général présenté dans la figure 4. 8.

Ce menu est divisé en trois champs, le premier champ est destiné pour les signaux mélangés, le second pour la séparation de sources tandis que le troisième est destiné pour la sauvegarde des résultats ainsi que les différentes représentations graphiques et enfin l'aide et le bouton pour fermer toutes les interfaces graphiques. (*Annexe. B*).



Figure 4. 8. Menu général de *SEPAREEG* vers. 1. 0

b) Représentations des différents résultats de séparation des signaux EEG

Pour visualiser les représentations graphiques nous avons utilisé les données réelles acquises; on a dix signaux de soixante seize milles échantillons chacun, notons qu'à cause de cette grande taille de nos données que la représentations des signaux paraissent un peu serrés. Les résultats qui seront présentés par la suite sont dus à la séparation de ces données via *FastICA* en utilisant comme algorithme de blanchiment *ACZ* et la fonction contraste *tght*.

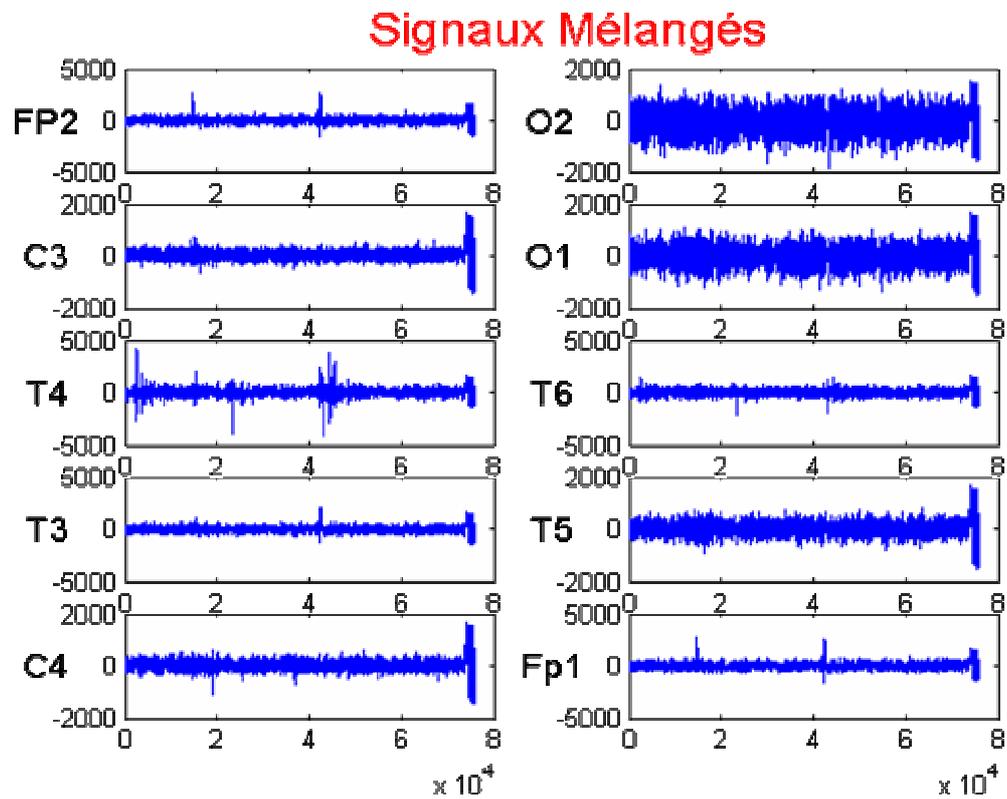


Figure 4. 9. Signaux Mélangés

Cette figure présente les signaux mélangés captés sur chaque électrode respectivement à l'examen *EEG*.

Les deux figures 4. 10 et 4. 11, représentent respectivement les signaux *EEG* blanchis après un prétraitement de ces données, par la suite la représentation des signaux séparés.

Signaux Blanchis

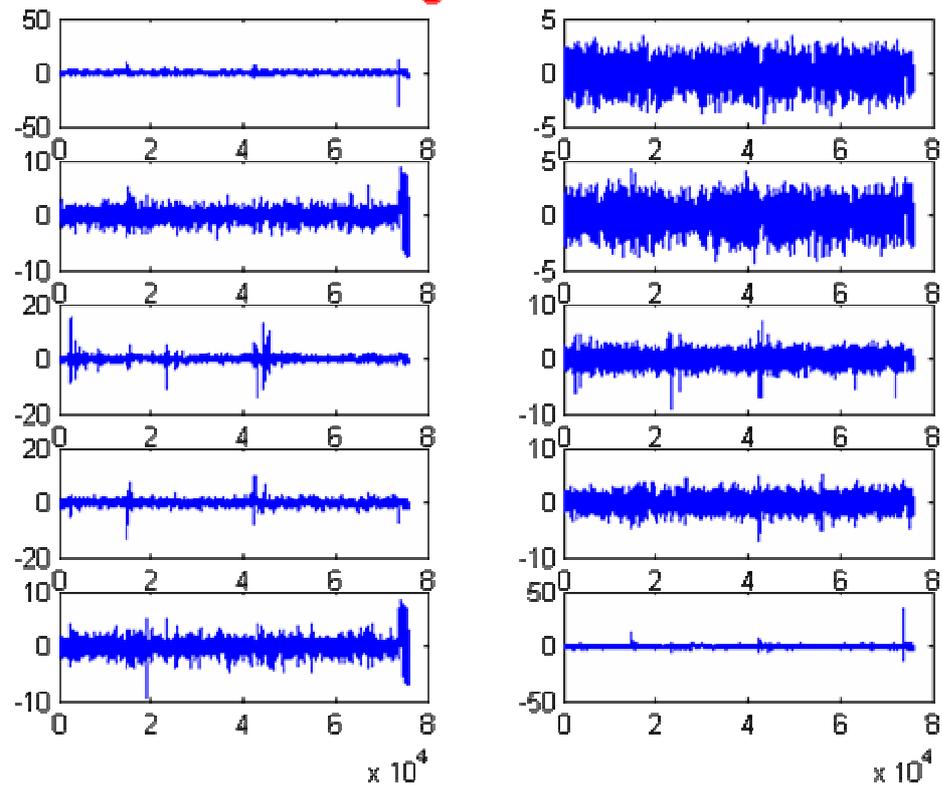


Figure 4.10. Signaux Blanchis

Signaux Séparés

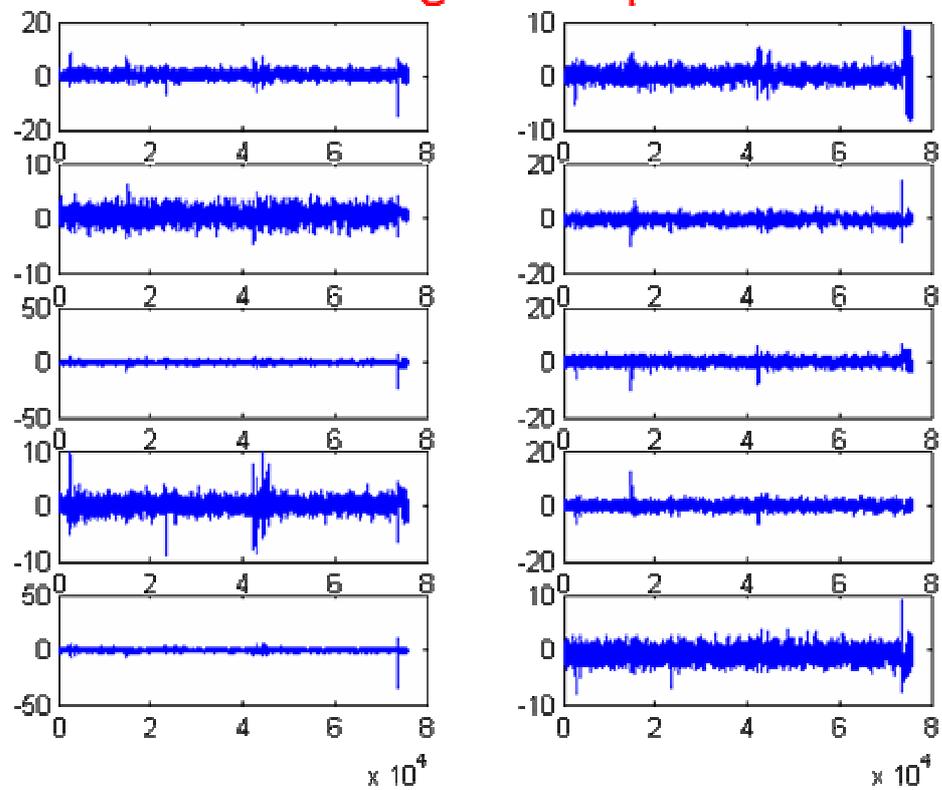


Figure 4.11. Signaux Séparés

Une représentation de l'activité cérébrale sur un crâne permet d'avoir une meilleure visibilité pour un meilleur diagnostique. Cette représentation est très utile si celle-ci est représentée sur le plan, mais elle peut rapporter encore plus de détail si cette représentation est faite dans l'espace avec la possibilité de faire tourner le crâne en trois dimensions.

Cette activité sera représentée sous forme de nuances de couleurs, par exemple la couleur rouge signifie que l'amplitude dans cette région du crâne est élevée, par contre le bleu marque une faible amplitude. Pour faciliter cette lecture de couleurs les représentations dispose d'une barre de couleurs qui montre les couleurs avec leurs intensités respectives.

La représentation de l'activité cérébrale se représente en premier lieu, par une fenêtre comme le montre la figure 4. 13, et qui s'ouvre avec un tableau dont tous les tracés des signaux sont représentés avec une barre glissante placée au milieu des échantillons.

Cette barre affichée en pointillé bleu sur le tableau des tracés est glissante changeante ainsi le temps selon sa position sur l'axe du temps, et l'affichage à gauche du tableau de l'intensité de chaque signal à cette position du temps choisi. Le crâne sera affiché à gauche de la fenêtre; si la barre est à nouveau glissée alors l'activité cérébrale sera remplacée par la nouvelle activité correspondante au nouveau temps choisi. En plus de la possibilité d'avoir cette activité en deux dimensions on peut l'avoir en trois dimensions. Le crâne en trois dimensions peut être pivoté à l'aide de la souris.

L'organigramme de cette représentation est présenté ci-dessous :

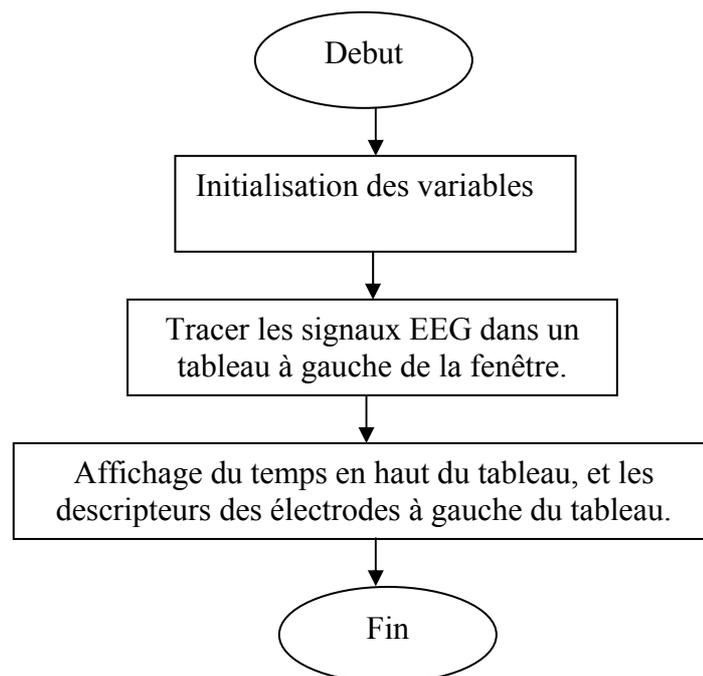


Figure IV. 12. Organigramme de la représentation de l'activité cérébrale

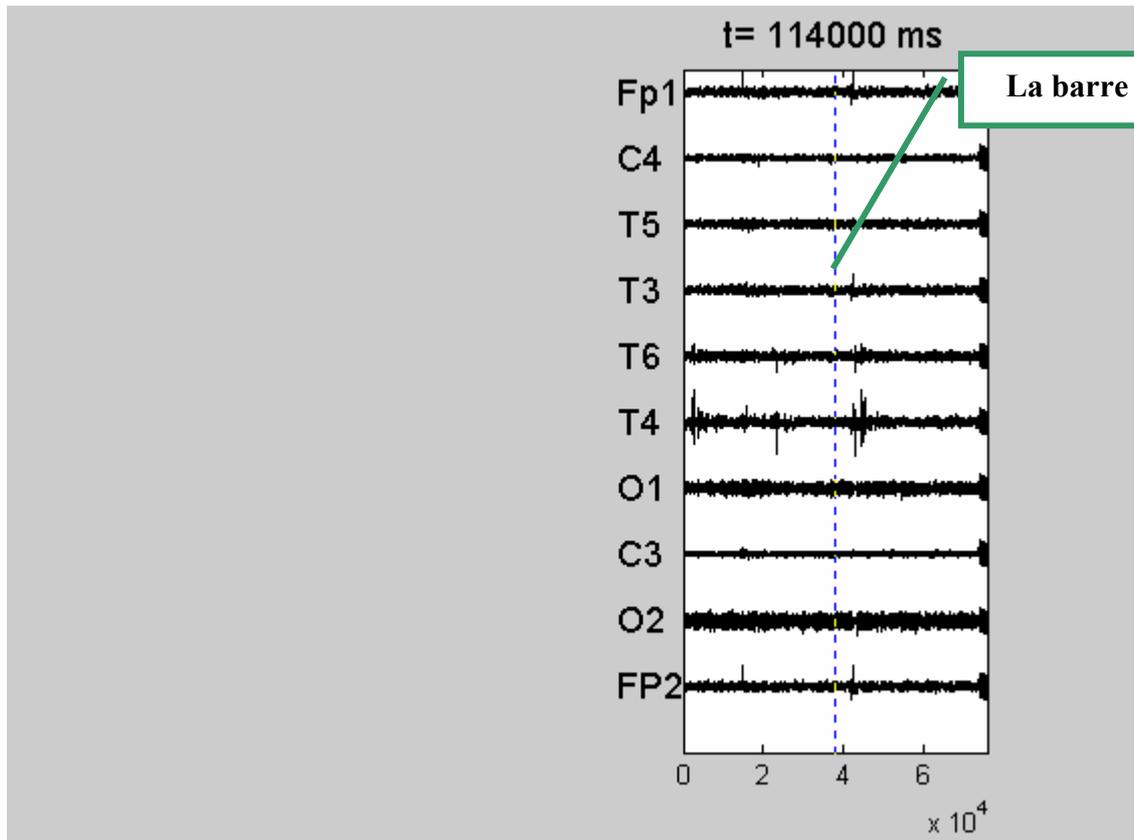


Figure 4. 13. Activité cérébrale

La figure 4. 15 montre l'activité cérébrale des signaux *EEG* dans le plan à l'instant choisi, on remarque que l'activité est grande au niveau de la partie frontale et temporale. Pour la représentation sur la figure 4. 16. c'est l'activité de la projection de la première composante indépendante dans le plan à l'instant choisi.

L'organigramme de l'affichage de l'activité cérébrale pour les signaux EEG ainsi pour l'activité de la projection des composantes indépendantes est présenté Figure IV. 14:

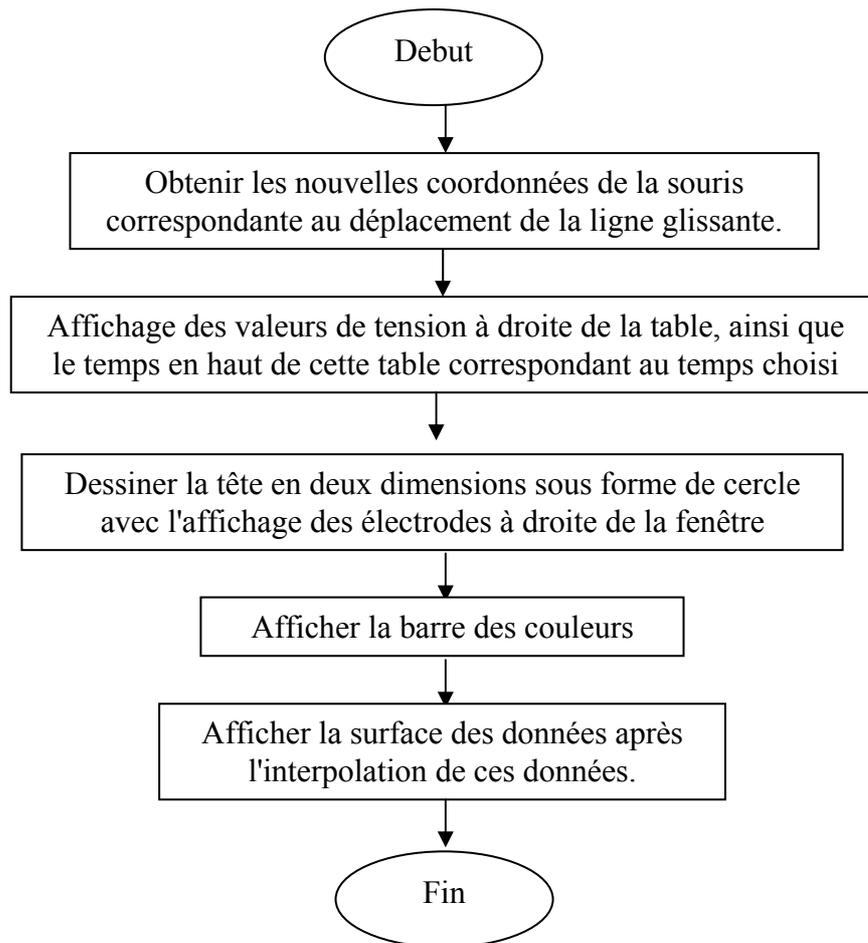


Figure IV. 14. Organigramme de la représentation de l'activité cérébrale en deux dimensions

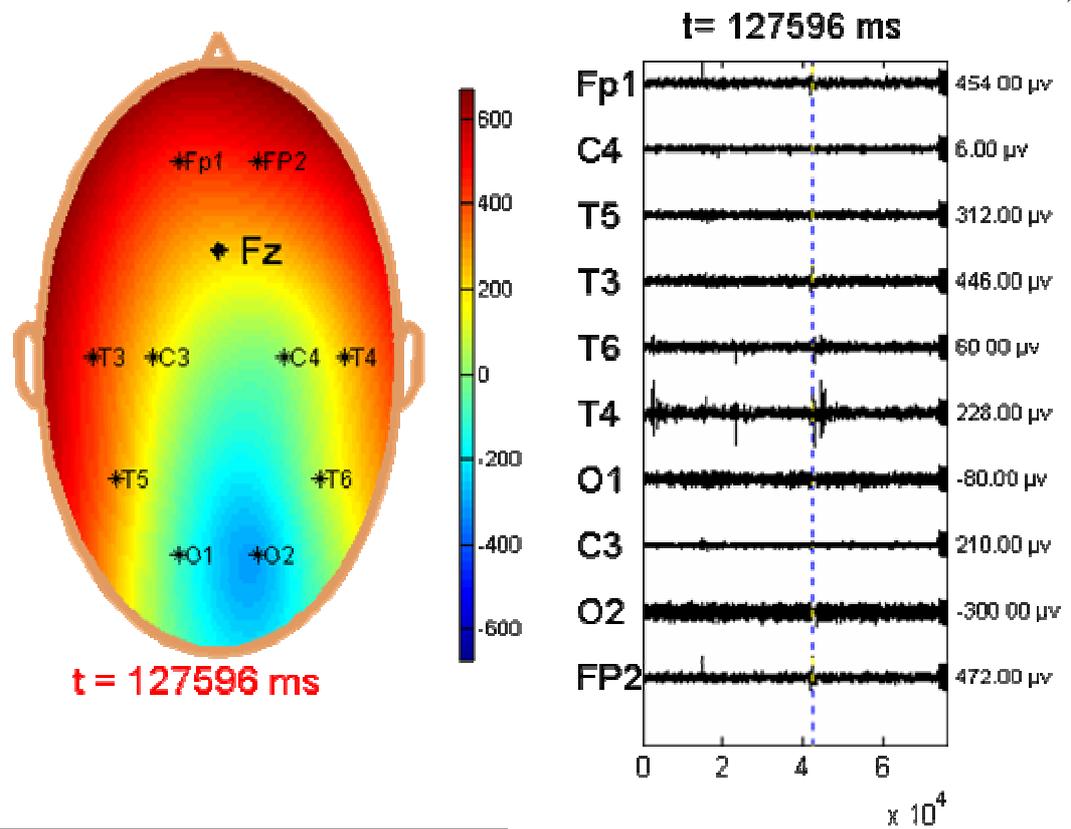


Figure 4. 15. Activité cérébrale des signaux *EEG* mélangés dans le plan à l'instant 127596

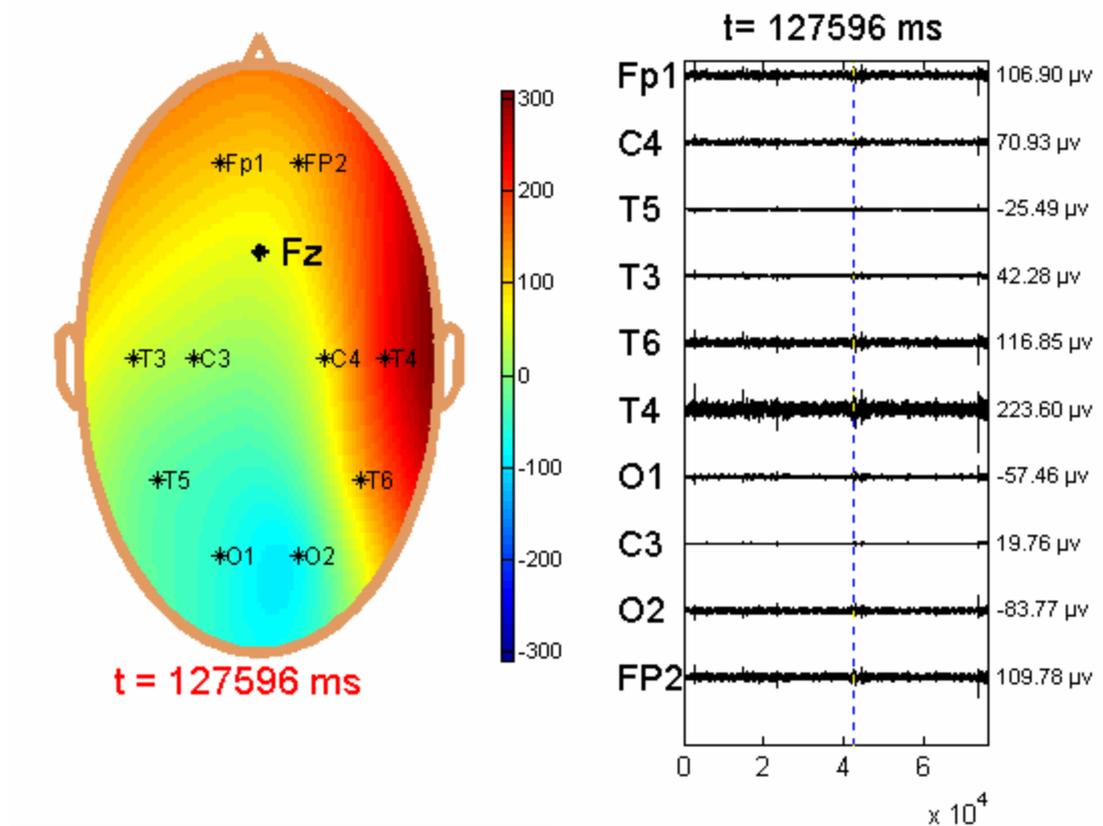


Figure 4. 16. Activité cérébrale de la projection de la première composante indépendante dans le plan.

Les figures 4. 17 et 4. 18, montrent respectivement les cartes du scalp dans le plan et les cartes du scalp dans l'espace. Ces cartes ne sont que les colonnes de la matrice de mélange. Les étoiles représentent l'emplacement des électrodes placé sur le crâne du sujet. Chaque numéro de carte représente le numéro de la colonne de la matrice mélange correspondante.

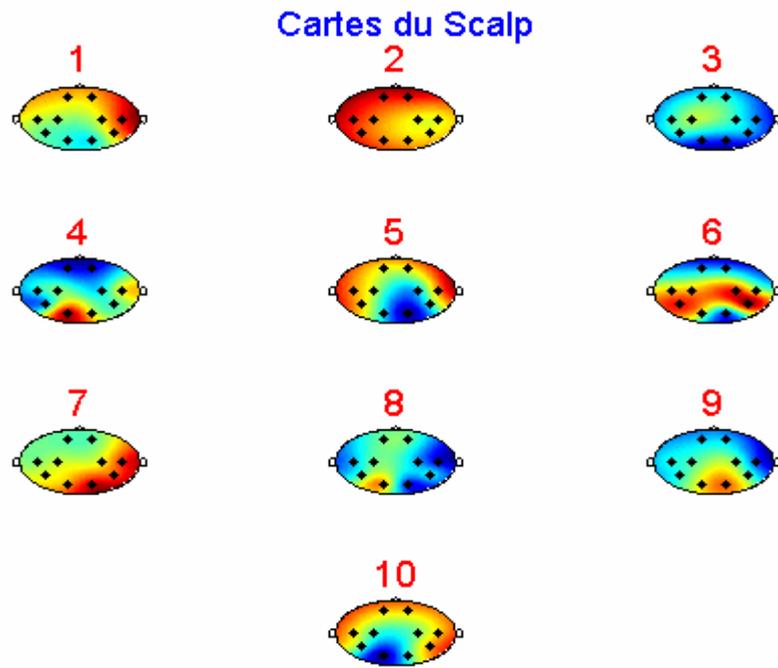


Figure 4. 17. Cartes du scalp dans le plan

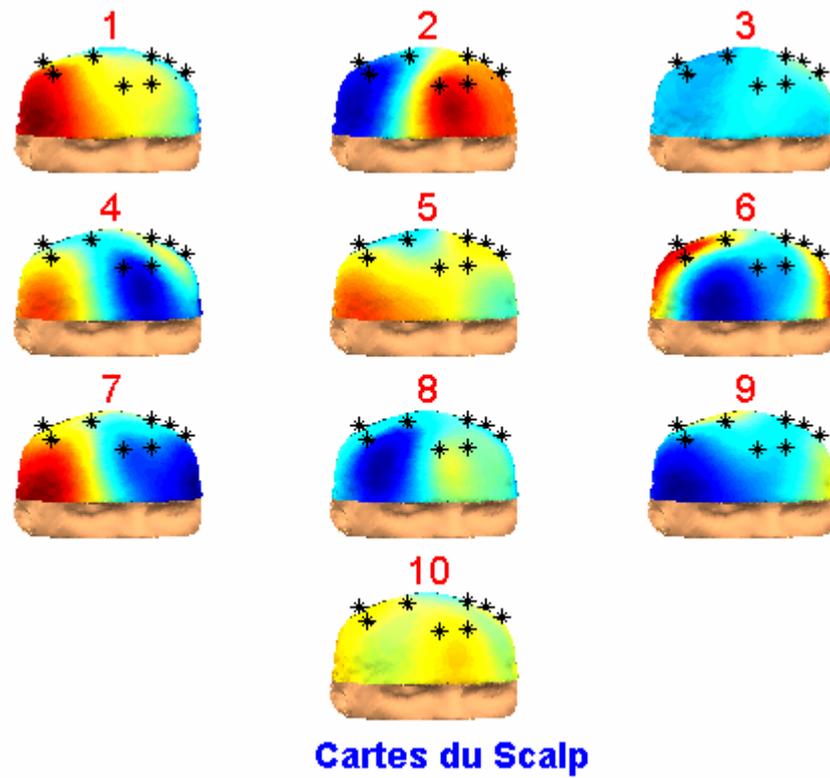


Figure 4. 18. Cartes du scalp dans l'espace.

Les figures 4. 20 et 4. 21, présentent l'activité cérébrale des signaux *EEG* et l'activité cérébrale de la projection de la deuxième composante indépendante respectivement, cette représentation est réalisée dans le plan ainsi que dans l'espace.

En changeant la position de la barre, on obtient l'activité du cerveau du sujet à l'instant choisi ; cette représentation en trois dimensions est avantageuse, vue l'exploitation de l'option de pivoter le crâne par la souris, ceci nous permet d'avoir une meilleure perspective du crâne d'où un diagnostic plus aisé.

L'organigramme de l'affichage de l'activité cérébrale pour les signaux EEG ainsi pour l'activité de la projection des composantes indépendantes en trois dimensions est présenté ci-dessous :

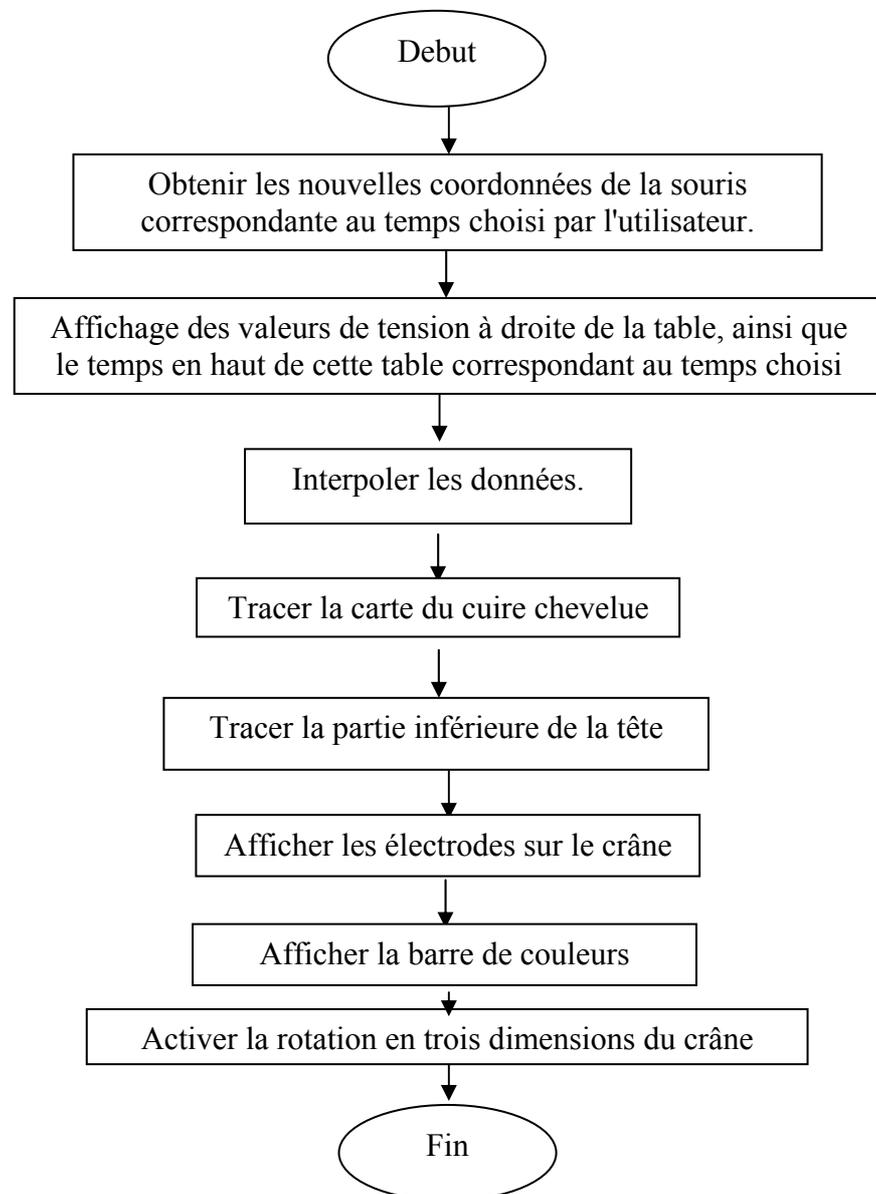


Figure 4. 19. Organigramme de la représentation de l'activité cérébrale en trois dimensions

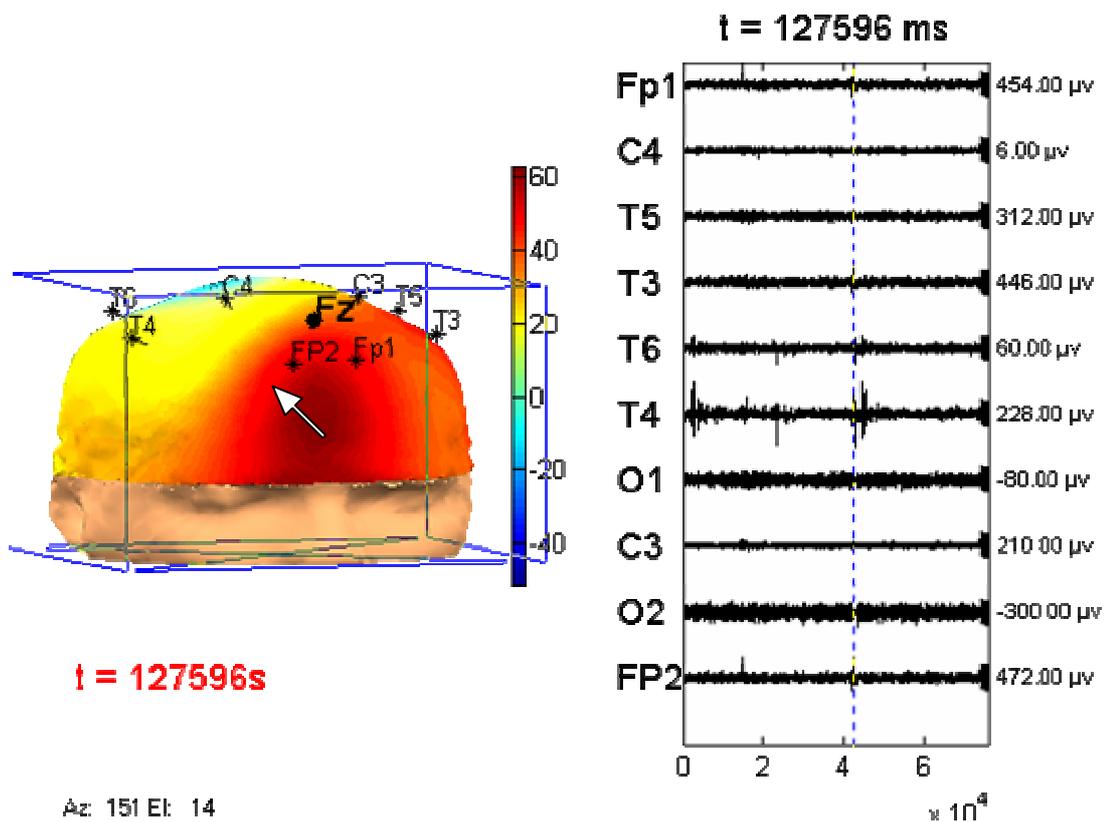


Figure 4. 20. Activité cérébrale des signaux *EEG* dans l'espace

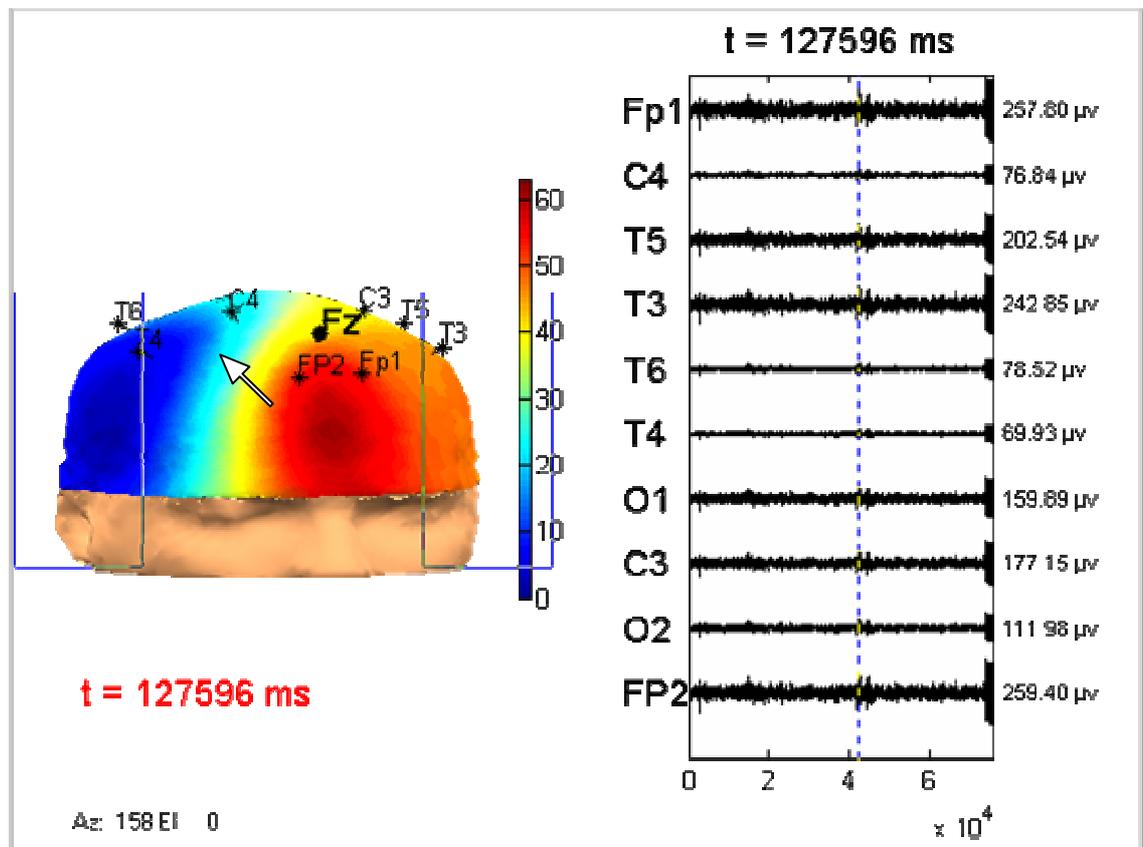


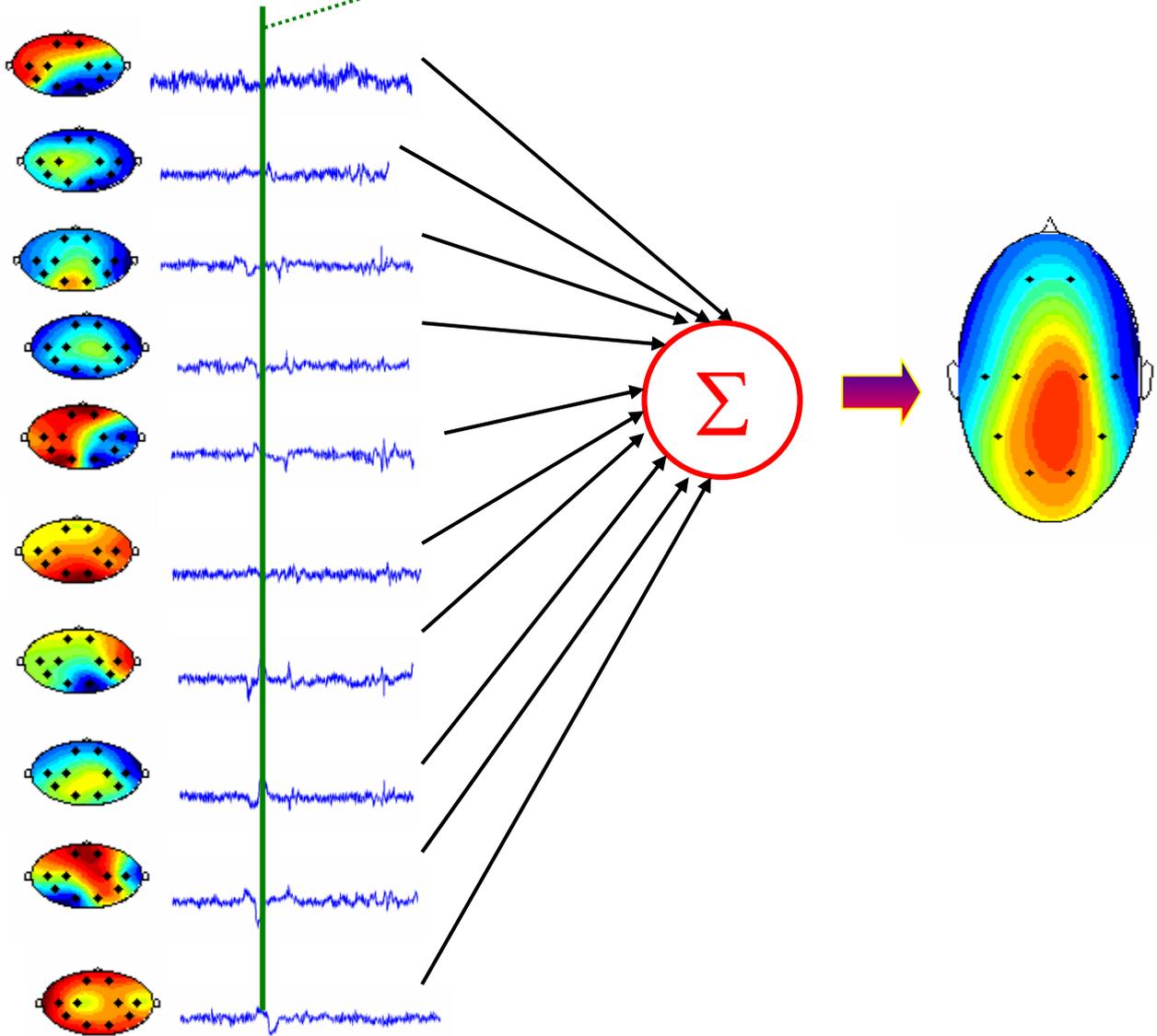
Figure 4. 21. Activité cérébrale de la projection de la deuxième composante indépendante dans l'espace

✘ Comparaison des signaux réels et la somme des projections

Dans ce qui suit nous présenterons l'activité cérébrale des signaux *EEG* à l'instant *127797 ms*, qui correspond aux clignotement des yeux du patient. Ce clignotement nous permettra de visualiser une grande activité au niveau de la partie Occipitale du crâne, qui se traduise par une nuance rouge sur le crâne au niveau des électrodes *O1* et *O2*.

Dans la figure 4. 22, nous présenterons tout d'abord en haut l'activité cérébrale résultante de la somme de toutes les projections de l'activité des composantes indépendantes (§ 4. 2), celle d'en bas représente l'activité cérébrale due aux signaux *EEG* réels enregistrés. On remarque que les représentations du crâne sont presque identique, dont la zone de la grande activité cérébrale des signaux *EEG* marqué par des contours rouges au niveau de la partie occipitale est plus large que celle représentée dans la somme des projection.

Projections *Composantes ACI* t=127797 ms



t=127797 ms *Signaux EEG Réels*

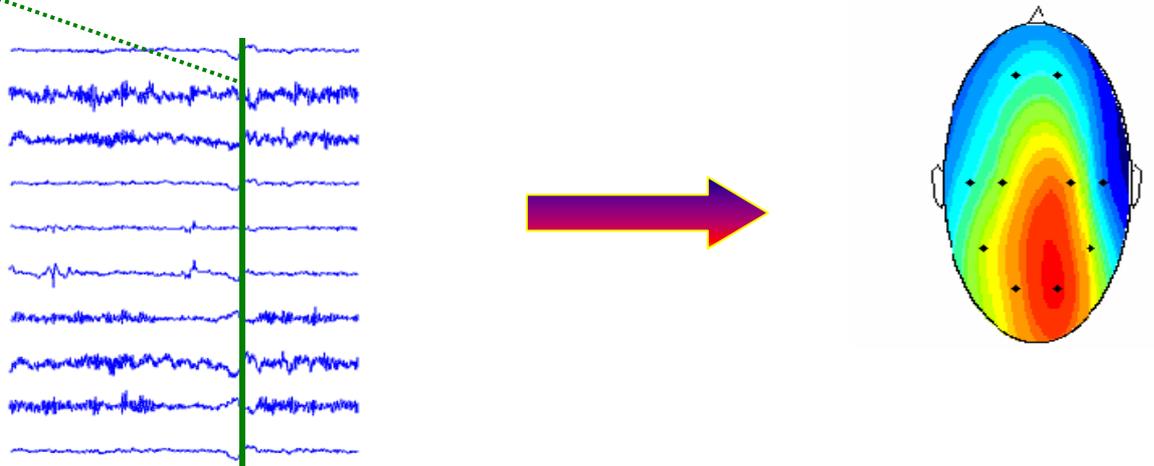


Figure 4. 22. Activités Cérébrales.

4.4 Séparation des artefacts

Ce deuxième examen, a été fait toujours au niveau de l'hôpital "CHU de Blida secteur Frantz Fanon". Les signaux EEG ont été enregistrés dans la salle d'examen avec huit électrodes, -cette diminution est due à une défaillance du matériel de l'hôpital au niveau de la fiche de sortie des signaux numériques- avec un montage référentiel illustré dans la figure 4. 23, sur un homme âgé de 33 ans.

Pour produire les artefacts musculaires, nous avons demandé au sujet de bien vouloir mordre ses dents durant 38 secondes, puis de cligner ses yeux pendant 37 secondes, et finalement pour créer un autre artefacts le sujet devait subir des flash de lumières à fréquences variantes 3 HZ, 6 HZ et 9 HZ c'est ce qu'on appelle la *Simulation Lumineuse Intermédiaire SLI* durant 5 minutes.

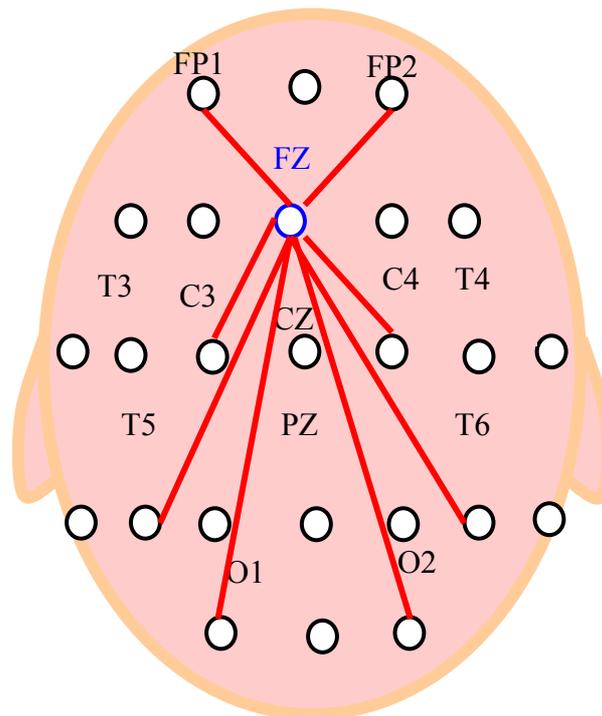


Figure 4. 23. Montage Référentiel.

La figure 4. 25 montre les huit composantes indépendantes retrouvés à partir des signaux enregistrés et illustrés dans la figure 4. 24. le résultat de la séparation des artéfacts présentés dans la figure 4. 25, montre les composantes indépendantes six et sept représentent les artéfacts originaires du serrement des dents, on a retrouvé deux composantes ce qui correspond au deux muscles responsables du serrement des dents, cette activité correspond au temps de la contraction musculaire pendant l'examen *EEG*, malgré que la composante indépendante sept ne devient pas aussi fine que la composante indépendante six après le temps de la contraction musculaire. Le deuxième artéfact qu'on a pu extraire est l'artéfact originaire du clignotement des yeux du patient, cet artéfact a été séparé dans la composante indépendante quatre, où on remarque que l'activité du signal devient très rapide pendant le clignotement du patient.

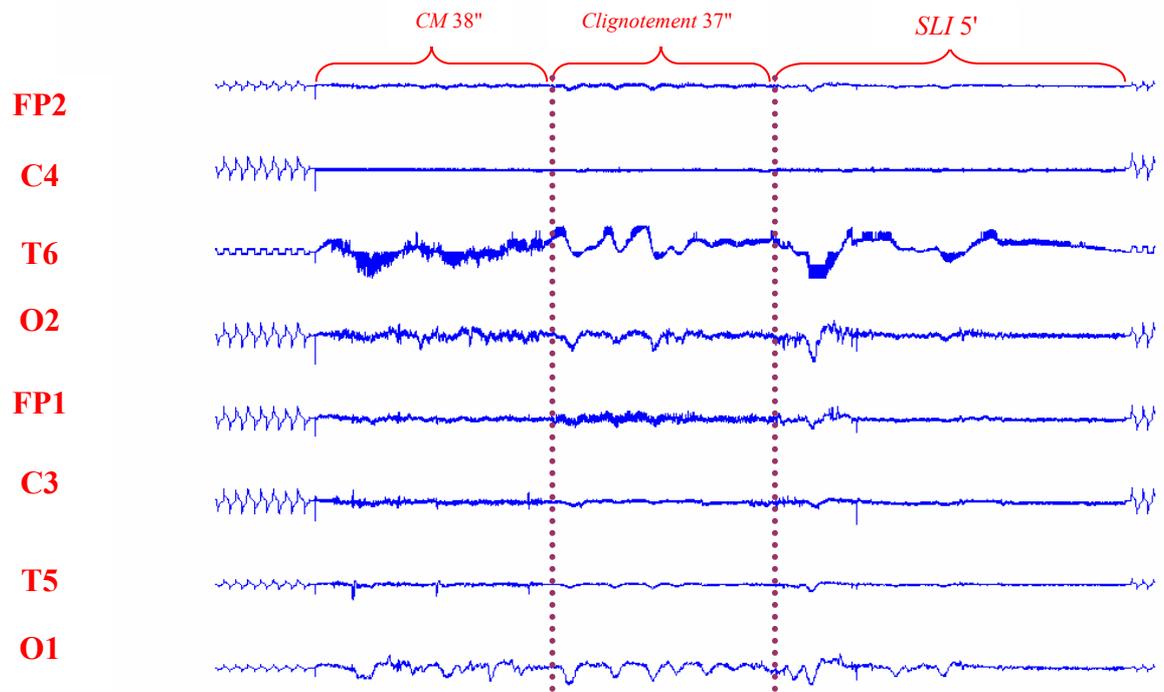


Figure 4. 24. Échantillons des signaux *EEG*, montrant les artéfacts produits par la contraction musculaire (*CM*), clignotement et la simulation limoneuse intermittente (*SLI*) avec les yeux fermés.

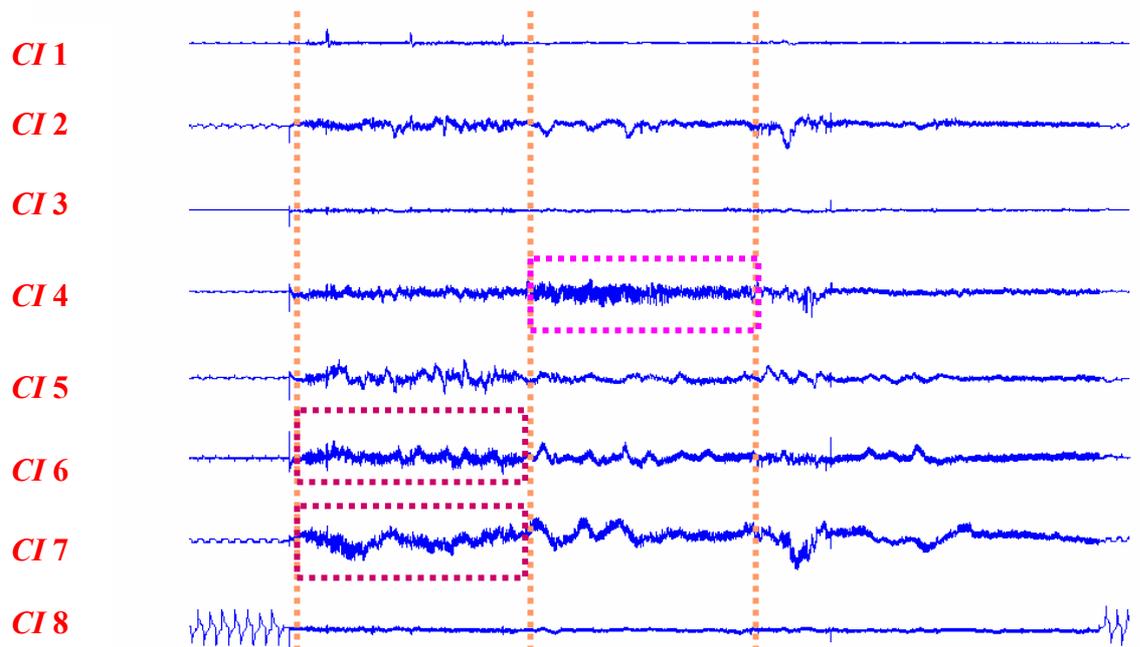


Figure 4. 25. Huit composantes indépendantes trouvées via *FASTICA-ACP* à partir des données *EEG*.

4.5 Conclusion

Ce chapitre a résumé la majorité des résultats obtenus par la séparation des signaux *EEG* par la méthode *d'ACI* acquis au sein d'un hôpital, ainsi que les différentes représentations des signaux dans le plan et l'espace, finalement la séparation des artefacts des signaux *EEG* qui sont indésirables au diagnostic. Ces résultats peuvent être améliorés en procédant à plusieurs expériences ainsi avec une augmentation considérable des signaux acquis donc le nombre d'électrodes utilisées pendant le test *EEG*.

CONCLUSION

La réalisation de ce mémoire nous a permis de faire des recherches approfondies dans le domaine du traitement de signal, en explorant une nouvelle technique d'analyse; l'Analyse en Composantes Indépendantes, cette dernière est utilisée pour séparer les signaux sources mélangés.

Nous avons utilisé cette technique pour séparer les artéfacts des signaux *EEG*, ces artéfacts sont très souvent indésirables pour l'analyse et le diagnostic. Pour cela nous avons tout d'abord testé nos algorithmes (*FastICA* et *Infomax*) sur des signaux connus, un mélange de signaux périodiques qui ont été mélangés puis séparés avec succès par la méthode *FastICA*. Nous avons aussi testé ces algorithmes sur des signaux audio, où les résultats étaient satisfaisants avec l'algorithme *FastICA* pour deux des quatre signaux utilisés, par contre avec l'algorithme *Infomax* les résultats n'étaient pas aussi performants.

Les signaux *EEG* utilisés pour la séparation ont été enregistrés à partir d'un appareil analogique classique au niveau du *CHU- Blida* secteur *Frantz Fanon*. Pour réaliser cette étape il a fallu effectué plusieurs expériences. Cette acquisition nous a permis d'avoir une disponibilité totale des données numérisées, puisqu'elles sont enregistrées sur le disque dur.

Les signaux obtenus ont été séparés puis projeté sur le crâne ensuite représentés sous forme topographique en trois dimensions. Cette représentation a été effectuée sur un crâne dans le plan et l'espace avec des nuances de couleurs indiquant l'activité des différentes zones du cerveau ce qui a amélioré la visibilité de l'activité cérébrale.

Le but de la création d'artéfacts artificielles était de tester les algorithmes de séparation, puisque nous connaissons la nature de ces artéfacts, les résultats obtenus peuvent être améliorés avec l'utilisation de plus d'électrodes et en effectuant plusieurs tests, comme il a été le cas dans [5].

REFERENCES

1. A. Cichocki, R. Unbehauen, E. Rummert, "Robust learning algorithm for blind separation of signals. *Electron. Lett.*, vol. 30, pp. 1386–1387, 1994.
2. A. Dale & M. Sereno. "Improved Localization of Cortical Activity By Combining EEG and MEG with MRI Cortical Surface Reconstruction". *J.Cogn. Neurosci*, vol. 5, pp. 162-176, 1993.
3. A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis". *IEEE Tran. On Neural Network*, vol. 10, pp. 626-634, 1999.
4. A. Hyvärinen, "Survey on Independent Component Analysis ", *Neural Computation. Sur.*, vol. 2, pp. 94-128, 1999.
5. A. Hyvärinen, E Oja. "Independent Component Analysis: A Tutorial". April 1999. téléchargeable sur :  @ <http://www.cis.hut.fi/projects/ica/>.
6. A. Hyvärinen, E. Oja, "A Fast Fixed-Point Algorithm for Independent component analysis". *Appear in Neural Computation*, vol. 9, pp. 1483-1492, 1997.
7. A. Hyvärinen, E. Oja. "Independent component analysis by general nonlinear Hebbian-like learning rules", *signal processing*, vol. 64, pp. 301-313, 1998.
8. A. Hyvärinen. "New Approximations of differential entropy for Independent Component Analysis and Projection Pursuit ". In *Neural Information Processing Systems*. vol. 10, pp. 273-279. MIT Press, 1998.
9. A. j. Bell & T. J. Sejnowski. "Learning the higher-order structure of a natural sound". *Neural Systems*, vol 7, pp. 261-266, 1996.
10. A.-J. Bell, T.-J. Sejnowski, "An Information-maximisation approach to blind separation and blind deconvolution ". *Neural Computation*, vol. 7, pp. 1129-1159, 1995(a).
11. A.-J. Bell, T.-J. Sejnowski, "The Independent Components' of Natural Scenes are Edge Filters". *Vision Research*, vol. 37, pp. 3327-3338, 1997.
12. A.-J. Bell, T.-J. Sejnowski, "Fast blind separation based on information theory", *Proc. Intern. Symp. Nonlinear Theory and Applications*, Las Vegas. Dec. 1995(b).

13. C. Jutten, J. Héroult, "Independent Components versus PCA ". *In Proc. EUSIPCO*, Grenoble, France, pp. 643-646, 1998.
14. C. Jutten, J. Héroult, "Independent Component Analysis versus PCA", *proc. EUSIPCO*, Grenoble, France, pp. 643-646, 1988.
15. C. Jutten, J. Héroult, " Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture ", *Signal Processing*, vol. 24, n° 1, pp. 1-10, 1991.
16. *Chapter I. Introduction*. Paris. Téléchargeable sur l'adresse:  @. <http://web.media.mit.edu/~Paris/phd/Paris-intro.pdf>.
17. Collection Microsoft ® Encarta ® 2005. © 1993-2004 Microsoft Corporation.
18. D.-T. Pham, P. Garat, C. Jutten, "Separation of a mixture of independent sources through a maximum Likelihood approach. ". *Dans Proc. EUSIPCO*, pp. 771-774, 1992.
19. G. Dara, S. Makeig, T-P. Jung, A. j. Bell, T. J. Sejnowski. "Independent Component Analysis of Simulated EEG Using a three-Shell Spherical Head Model". *Tech. Rep. INC-9601*, Institute for Neural Computation, San Diego, CA, 1996.
20. J. F. Cardoso and P. Comon. "Independent Component Analysis, a survey of some algebraic methods". In *Proc. ISCAS'96*, vol.2, pp. 93-96, 1996.
21. J. Max et collaborateurs, *Méthodes et techniques de traitement du signal et applications aux mesures physiques*, Tome 1 et Tome 2, 4^e édition. Ed. Masson, 1985.
22. J. Max et collaborateurs. *Méthodes et techniques de traitement du signal et applications aux mesures physiques*. Tome. 1. 4^{ème} édition, ed. Masson.
23. J.-F. Cardoso, "Source separation using higher order moments", *Proc. ICASSP* Glasgow, pp. 2109-2112, 1989.
24. J.-F. Cardoso, A. Souloumiak, "Blind beamforming for non gaussian signals ". *IEEE Proceedings-F*, 140, vol. 6, pp. 362-370, Dec. 1993.
25. J.-F. Cardoso, B.-H. Laheld, "Equivariant adaptive source separation ", *IEEE Tran. Signal Processing*, vol. 44, pp. 3017-3030, Dec. 1996.
26. J.-L. Lacoume, P.-O. Amblard, P. Comon, *Statistiques d'ordre supérieur pour le traitement du signal*, oct 1997, ed. Masson.
27. O. Belkina, N. Boussaid. *Acquisition de 16 signaux simultanément d'un EEG*. Thèse PFE ingénieur d'état en électronique, option contrôle, département d'électronique. USTB session Novembre 2002.
28. P. Comon, "Independent Component Analysis, a new concept? ", *Signal Processing*, vol. 36, n° 3, pp. 287-314, Apr. 1994.

29. P. Rondot, J. Gaches, C. Sebban, *CARTOGRAPHIE E.E.G. METHODOLOGIE ET APPLICATIONS CLINIQUES*, nov. 1987, Ed. Médicales Frison-Roche.
30. R. M. E. Sabbatini. *The Future of EEG Brain Mapping*. Téléchargeable sur l'adresse :  @ <http://www.epub.org.br/cm/n03/tecnologia/futuro.htm>.
31. S. Amari, "Natural gradient works efficiently in learning". *Neural Comput.*, vol. 10, pp. 251–276, 1998.
32. S. Makeig, A.J. Bell, T.-P. Jung, T.-J. Sejnowski, "Independent Components Analysis of electroencephalographic data". *In advances in Neural Information Processing Systems*, vol. 8, pp. 145- 151. MIT Press, 1996.
33. S. Makeig, T. -P. Jung, A. J. Bell, D. Ghahremani, T. J. Sejnowski. "Blind Separation of Auditory Event-related Brain Responses into Independent Components ".*Proc. Natl. Acad.Sci. USA*, vol 94, pp. 10979-10984, 1997.
34. S. Makeig, T-P. Jung, D. Ghahremani, A. J. Bell, T. J. Sejnowski. "What (Not Where) are the Sources of the EEG?". *18th Annual Meeting of the Cognitive Science Society* 802, 1996.
35. S. Makeig, with M. Westerfield, J. –R. Duann, T. –P. Jung. *ICA Toolbox Tutorial*. Téléchargeable sur l'adresse :  @ <http://www.cnl.salk.edu/~scott/tutorial.tar.gz>
36. S.-I. Amari, A. Cichocki, H.H. Yang, "A New Learning Algorithm for Blind signal Separation", *In advances in Neural Information Processing Systems*, vol. 8, pp. 757-763. MIT Press, Cambridge, MA, 1996.
37. T. -W. Lee, "Introduction to Independent Component Analysis", Mars 1998. Téléchargeable sur l'adresse :  @ <http://www.cnl.salk.edu/~tewon/iICA/intro.html>.
38. T. –W. Lee, M. Girolami, and T. J. Sejnowski, "Independent Component Analysis using an Extended Infomax Algorithm for Mixed sub-gaussian and super-gaussian sources ". *Neural Computation*, vol. 11, pp. 417-441, 1999.
39. T. Zeman, "BSS – Preprocessing Steps for Separation Improvement ". May 2000.
40. T. Zeman. "Newton's' method for FastICA algorithm", May 2000.
41. T.-P. Jung, S. Makeig, M.-j. Mckeown, A.-J. Bell, T.-W. Lee, J. Sejnowski, "Imaging Brain Dynamics Using Independent Component Analysis". *IEEE*. Vol. 89, N° 7, July 2001.
42. T.-W. Lee, M.-S. Lewicki, M. Girolami, T.-J. Sejnowski "Blind Source Separation of More Sources Than Mixtures Using Overcomplete Representations ". *IEEE signal processing letters*, vol. 6, N° 4, April 1999.

43. T-p. Jung, "What is Independent Component Analysis (ICA)? ", disponible sur l'adresse:  @ <http://www.cnl.salk.edu/~jung/ica.html>.

ANNEXE. A

ANNEXE. B

Présentation du Logiciel SEPAREEG 1.0

Pour permettre une manipulation conviviale et pratique, pour notre séparation de sources ainsi que pour les représentations des différents résultats, on a développé un logiciel en utilisant une programmation multi-tâches basée sur des interfaces graphiques développé sous *MATLAB 6.5*.

L'exécution de notre logiciel nous permet d'avoir en premier lieu le menu général présenté en dessous.



Figure 1. Menu général de *SeparEEG ver.1.0*

Ce menu est divisé en trois champs, le premier champ est destiné pour les signaux mélangés, le second pour la séparation de sources tandis que le troisième pour la sauvegarde des résultats ainsi que les différentes représentations graphiques et enfin l'aide et le bouton pour fermer toutes les interfaces graphiques.

Pour pouvoir séparer des sources, faudrait bien charger le fichier contenant les données qui seront les signaux *EEG* mélangés. Pour cela il faudrait cliquer sur le bouton **Chargement** dans le champ *Signaux Mélangés* du menu général; alors le nom du fichier complet doit être écrit dans le champ approprié, puis cliquer sur le bouton *Charger* pour que le fichier soit charger.

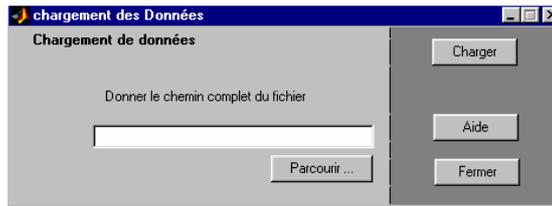


Figure 2. Chargement de données

A la fin du chargement du fichier, le nombre des signaux ainsi que le nombre d'échantillons seront affichés dans le premier champ du menu général, en plus on a la possibilité de tracer les signaux *EEG* mélangés avant de procéder à la séparation en cliquant sur le bouton **Tracer des Comp. Indép.** du premier champ *Signaux Mélangés* du menu.

Pour séparer les données chargées préalablement, on a le choix dans le champ *Séparation de Sources* entre les deux algorithmes de séparation *Infomax* et *FastICA*. Si le choix fait est l'algorithme *Infomax* il suffit de cliquer sur le bouton **Infomax**, mais si le choix fait est sur l'algorithme **FastICA** il faut choisir une fonction contraste ainsi que la fonction de blanchiment selon le choix l'algorithme sera exécuté, sinon le menu dispose des choix fait par défaut pour exécuter *FastICA* directement, puis cliquer sur le bouton *FastICA*.

Dans le troisième champ du menu général on dispose de l'aide ainsi que la sauvegarde des résultats; les données pouvant être sauvegardées toutes dans un même fichier ou bien dans l'espace de travail *MATLAB* en donnant le même suffixe pour toutes les données.

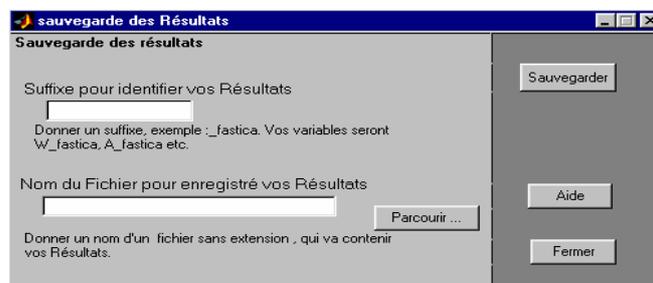


Figure 3. Sauvegarde de données

Le dernier bouton dans le troisième champ du menu général est le bouton **Sauvegarde des résultats**. Qui nous permet d'avoir plusieurs formes de représentations.

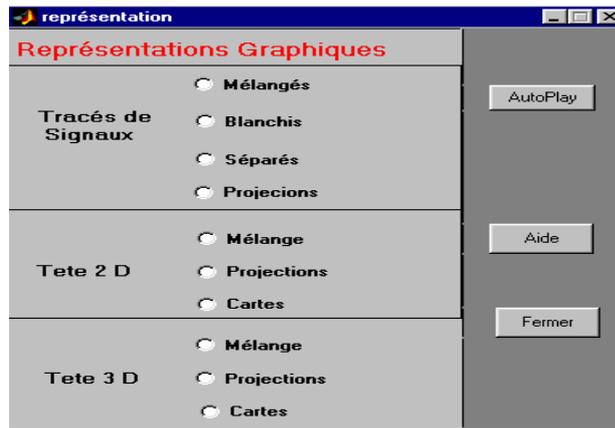
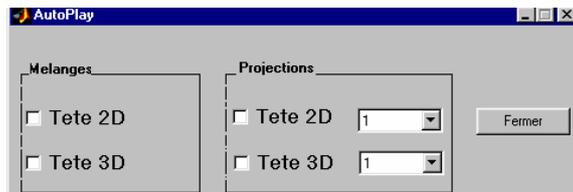
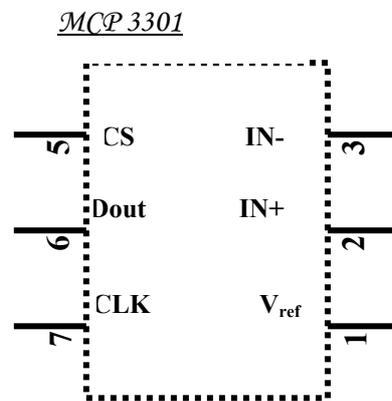
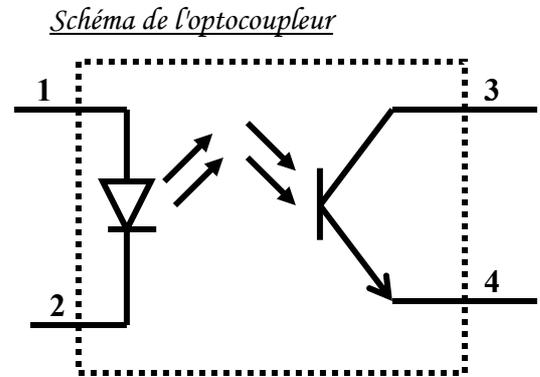
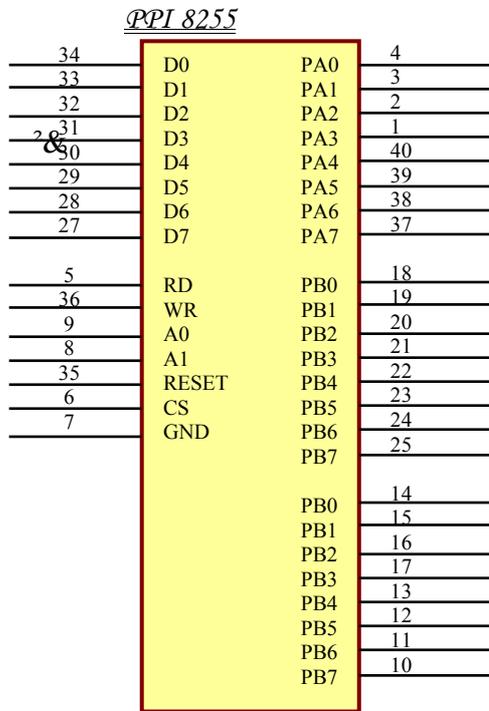


Figure 4. Représentations Graphiques

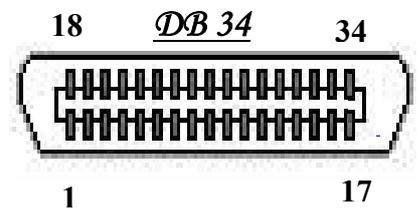
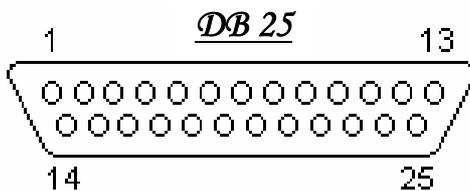
Comme nous la montre la figure précédente on a la possibilité de tracer les signaux blanchis mélangés, séparés ainsi que les projections dans le plan. On nous permet également d'avoir l'activité cérébrale dans le plan dû aux signaux EEG mélangés et dû aux projections avec la possibilité de choisir les projections voulus à l'aide d'un menu déroulant ainsi que toutes les cartes du scalp ainsi que l'activité cérébrale dans l'espace des signaux mélangés et séparés. Et finalement une exécution automatique à l'aide du bouton  nous permet d'avoir les représentations automatiquement des possibilités figurantes dans la figure suivante.



ANNEXE. C

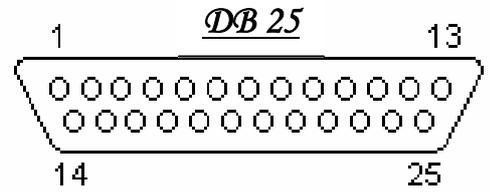
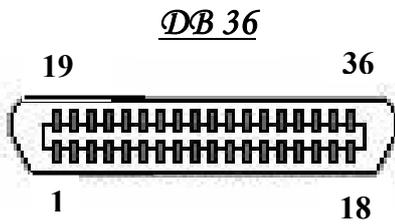


▶ La liaison : Pc → Carte de conversion:



<i>PPI</i>	4	3	2	1	40	39	38	37	18	19	20	21	22	23	24	25	14	15	7	
<i>255A</i>																				
<i>DB 25</i>	4	5	6	7	8	9	10	11	12	25	24	23	22	19	20	21	2	3	13	
<i>de PC</i>																				
<i>DB 34</i>	4	2	20	22	21	19	3	5	9	26	6	23	7	24	8	25	13	12	34	
	PA0	PA1	PA2	PA3	PA4	PA5	PA6	PA7	PB0	PB1	PB2	PB3	PB4	PB5	PB6	PB7	PC0	PC1	GND	

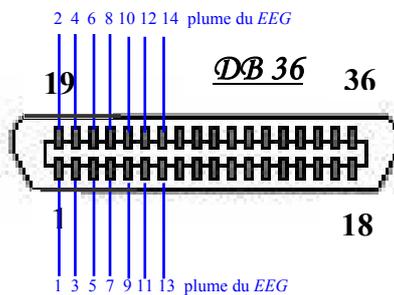
► La liaison : Carte de conversion: —► EEG:



DB 25 (côté EEG) 25 13 12 11 10 23 24 7 6 5 4 3 2 1
 22
V_{cc}=2,5 v

DB 36 19 GND 20 21 22 23 24 25 7 6 5 4 3 2 1

► Connexion du DB 36 à l'appareil EEG :



lum 1^{ère} 2° 3° 4° 5° 6° 7° 8° 9° 10° 11° 12° 13° 14°
 s du
 EEG PB PB PB PA PA PB PB PA PA PB PB PA PB PA
 mné 0 5 1 2 7 6 2 3 6 7 3 0 4 1
 es

ANNEXE. D

Validation des Algorithmes

1 Introduction

Dans cette partie nous allons voir les résultats donnés par l'authentification des algorithmes proposés dans le chapitre un de la méthode de l'Analyse en Composantes Indépendantes, les deux algorithmes *FastICA* ainsi que *Infomax* ont été développés sous *MATLABR 6.5*. Les application de cette authentification et qui fait l'objet de cette partie, représentent deux applications, la première étant la séparation de quatre signaux périodiques, la deuxième la séparation de quatre signaux audio de trois secondes chacun.

2 Séparation des signaux périodiques

La première application, pour cette authentification est une application de quatre signaux périodiques de mille échantillons chacun créer sous *MATLAB 6.5*. Le mélange de ces quatre signaux périodiques a été réalisé par une matrice dont ces éléments sont compris entre 0 et 1 générées par *MATLABR 6.5* par la fonction *Rand*.

Tous les résultats obtenus pour cette application seront présentés ci-après. La séparation de cette application a été faite via les deux algorithmes *FastICA* et *Infomax*, en utilisant les deux méthodes de blanchiment *ACZ* et *ACP* pour l'algorithme *FastICA* ainsi que toutes les fonctions contrastes proposées dans le chapitre un, en ce qui concerne l'algorithme *Infomax* le blanchiment a été effectué via l'algorithme *ACZ*, et la fonction sigmoïdale utilisée et la fonction *logistique*.

Les figures suivantes montrent que la séparation de sources a été complète via l'algorithme *FastICA* en utilisant toutes les fonctions contrastes et les deux méthodes de blanchiment, la séparation via l'algorithme *Infomax* a séparé les quatre signaux sources. Notons que l'ordre des composantes Indépendantes n'est pas respecté par la séparation.

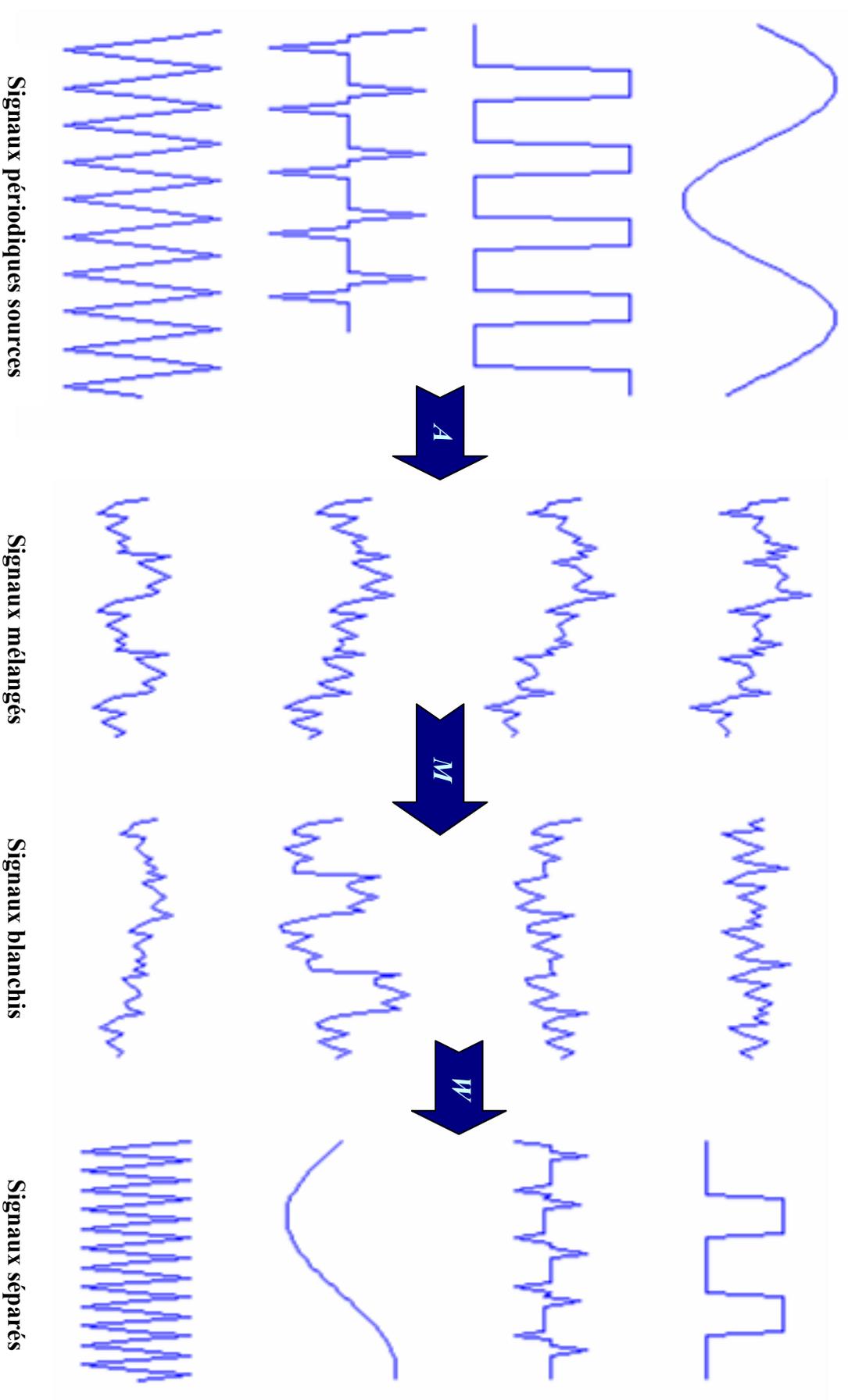


Figure. 1. Séparation de signaux périodiques par FastICA via ACZ et la fonction gauss

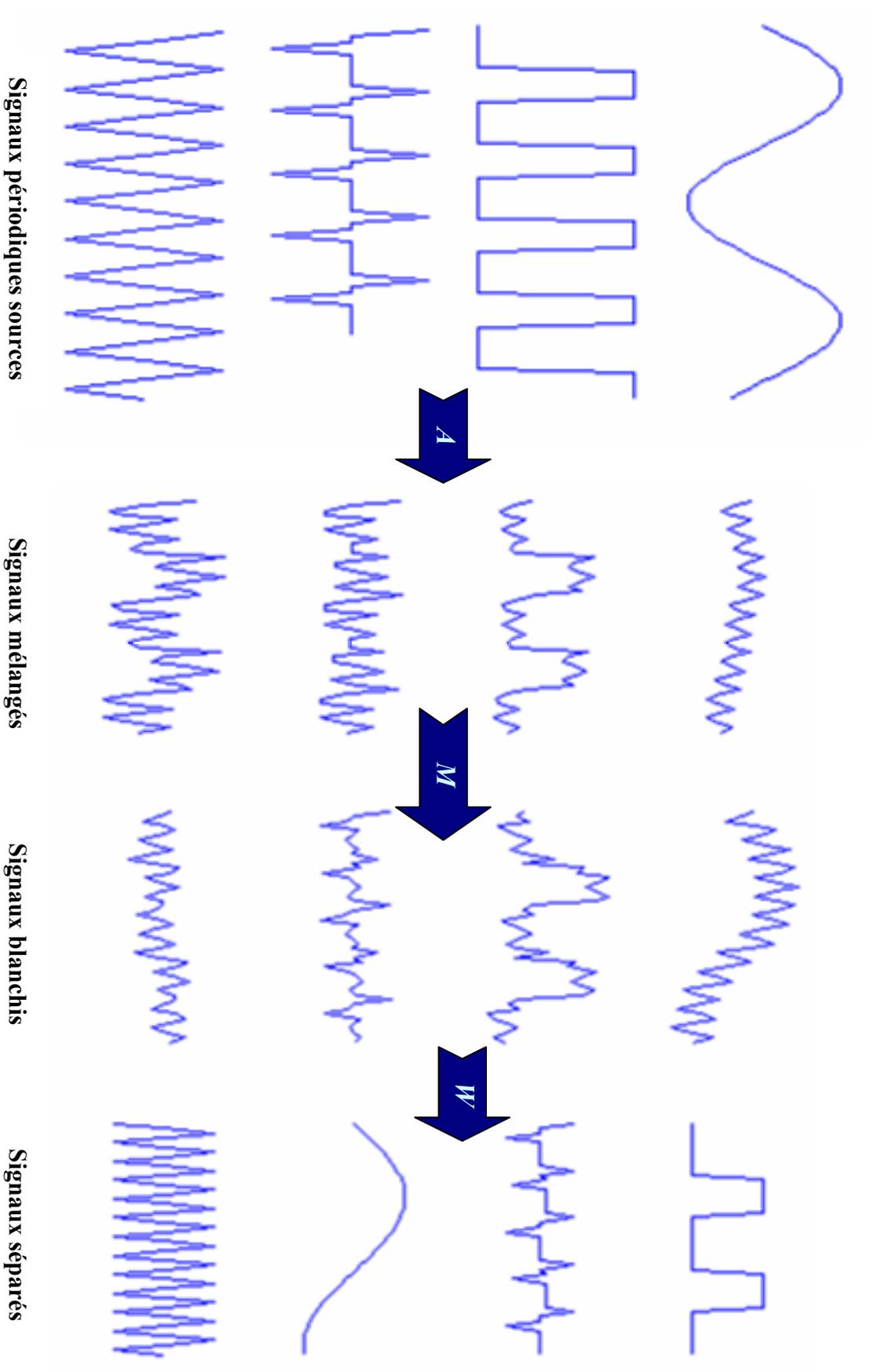


Figure 2. Séparation de signaux périodiques par FastICA via ACZ et la fonction $tanh$

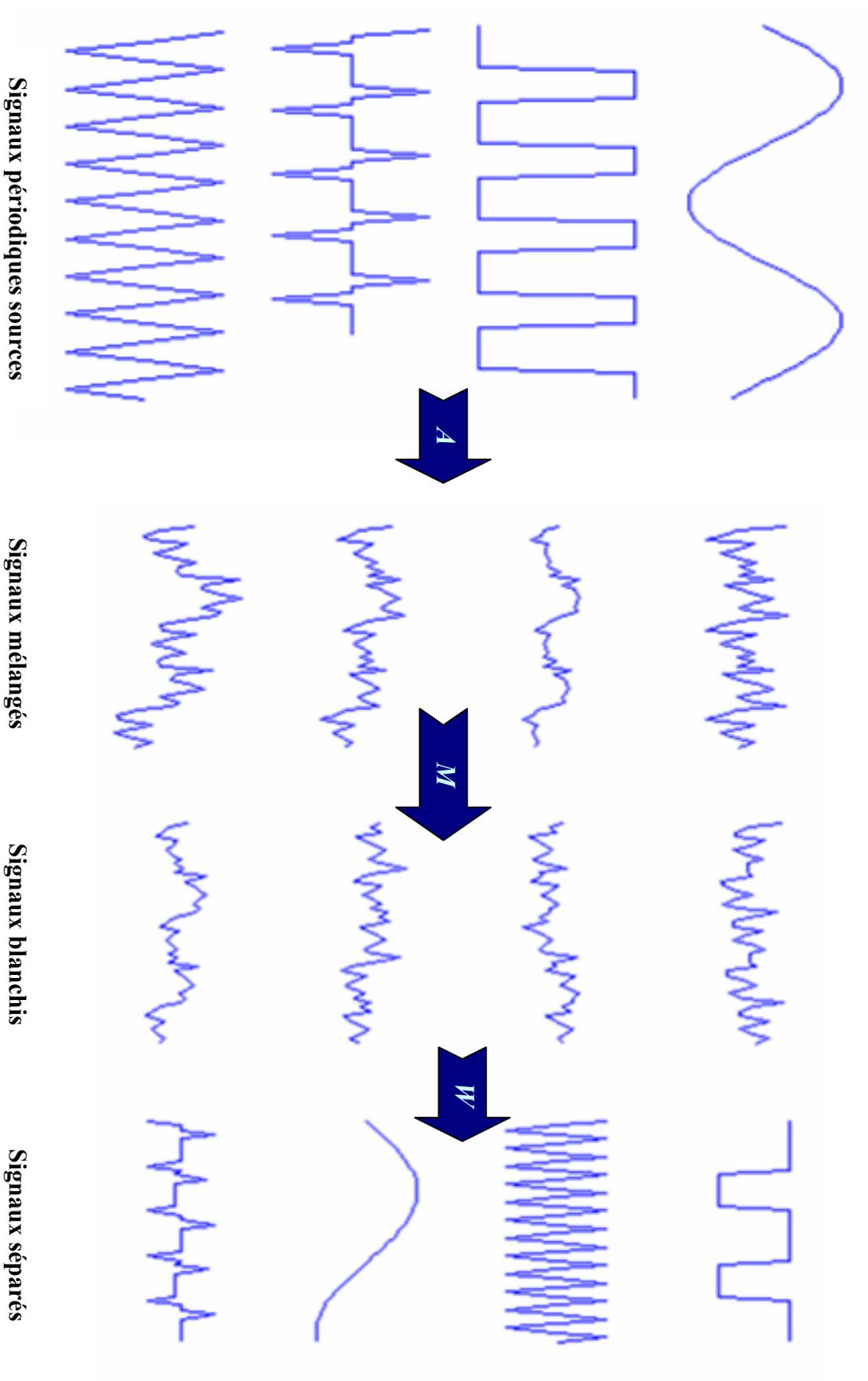


Figure.3. Séparation de signaux périodiques par FastICA via ACZ et la fonction puissance de trois

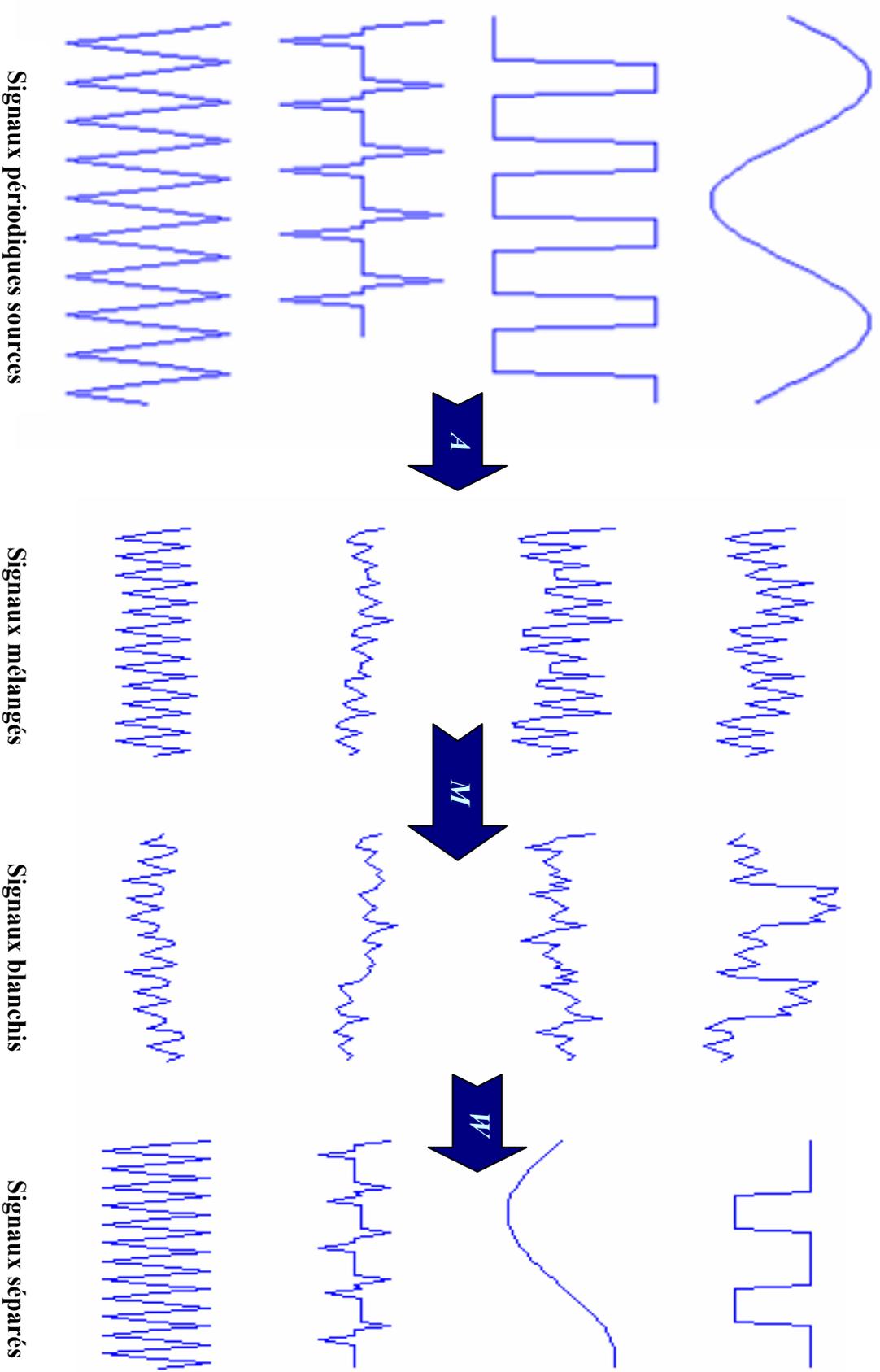


Figure.4. Séparation de signaux périodiques par FastICA via ACP et la fonction puissance de trois

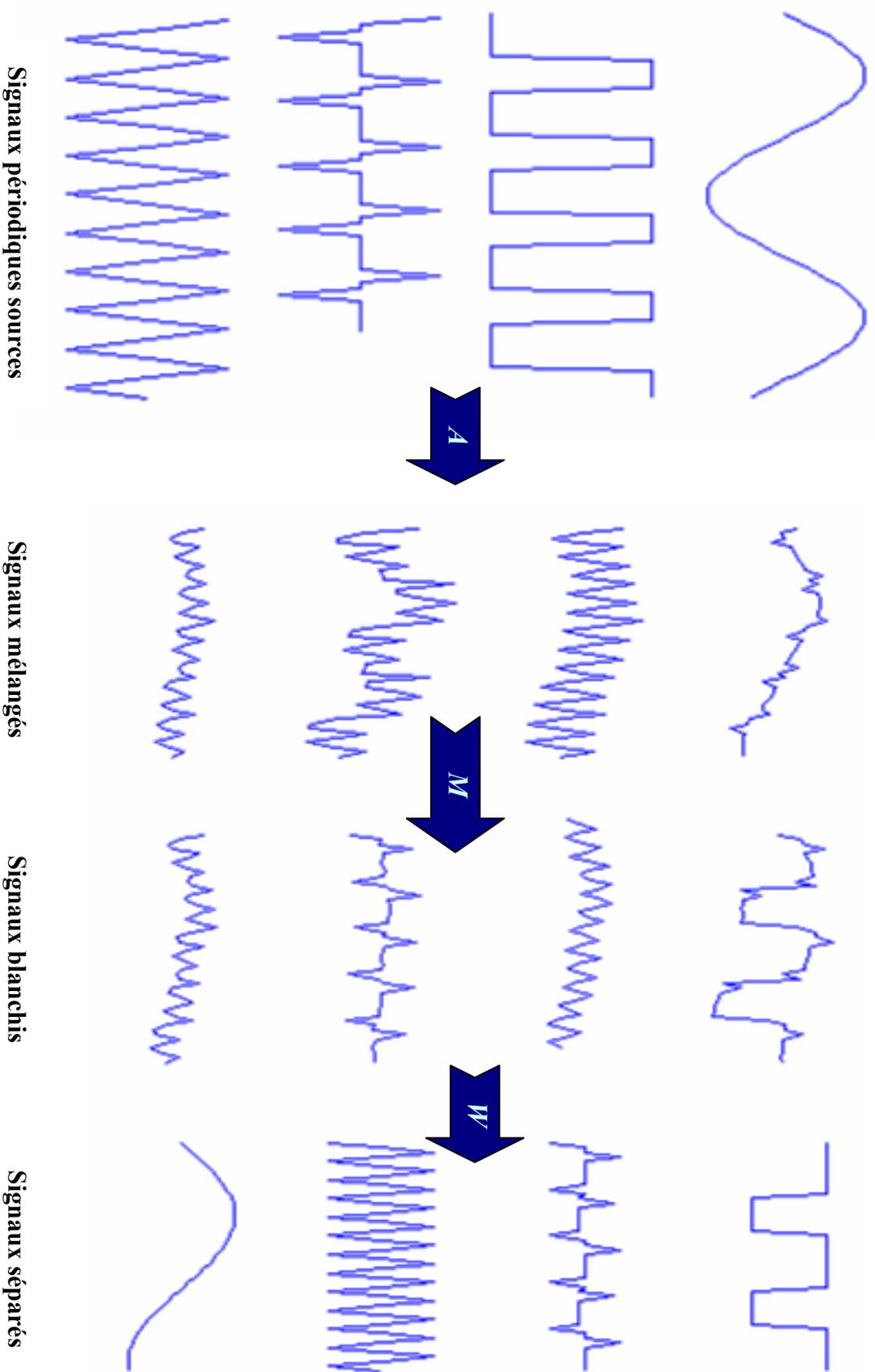


Figure.5. Séparation de signaux périodiques par FastICA via ACP et la fonction *tgh*.

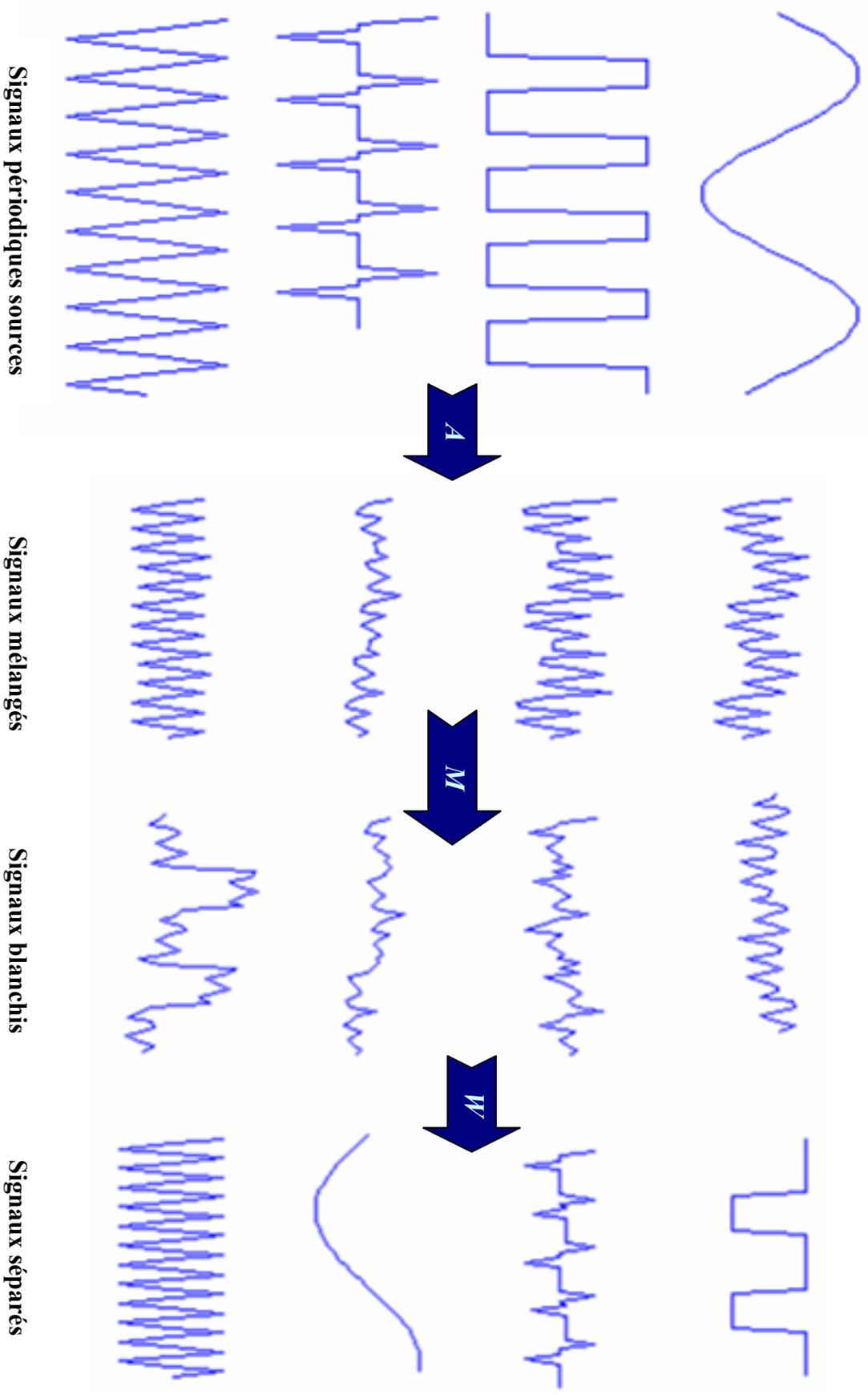


Figure. 6. Séparation de signaux périodiques par *FastICA* via *ACP* et la fonction gauss

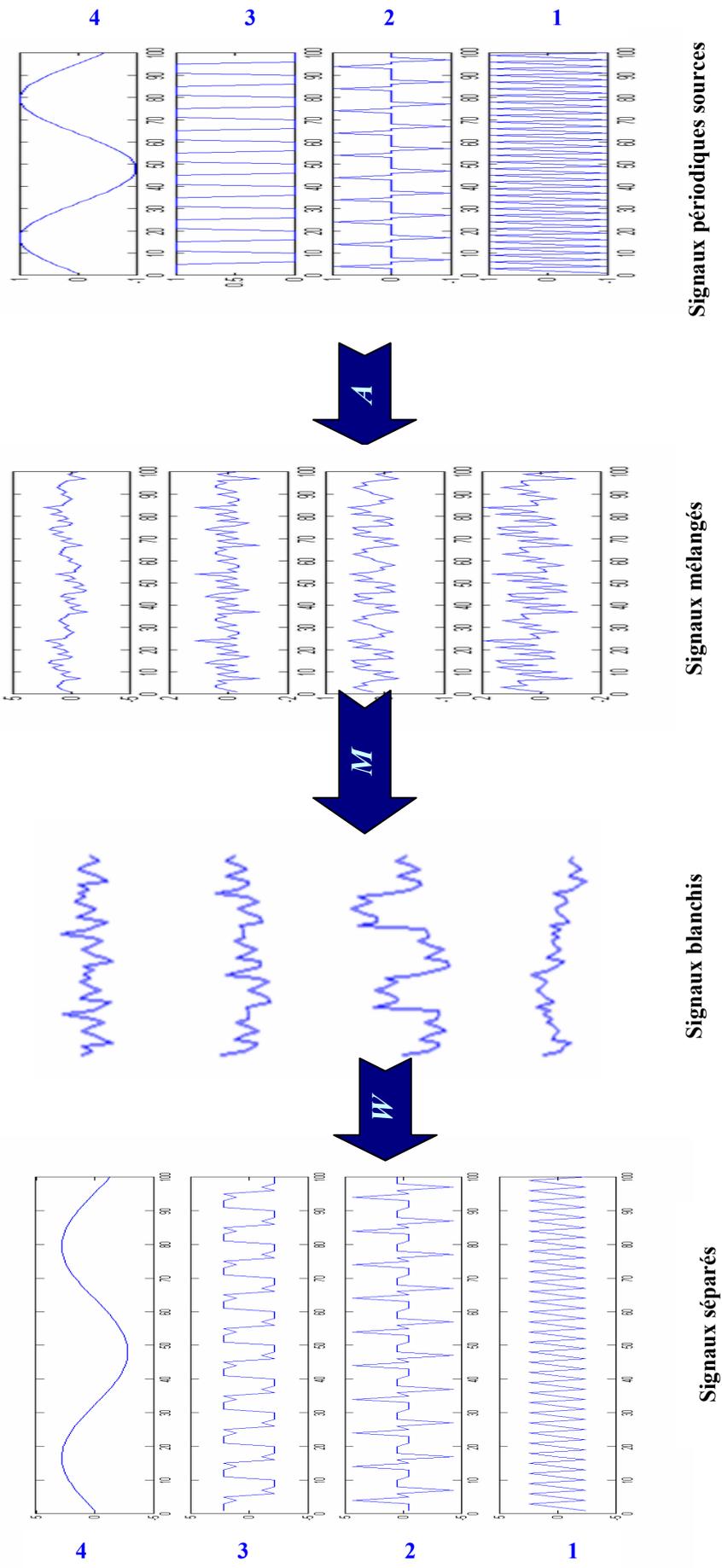


Figure. 7. Séparation de signaux périodiques par Infomax

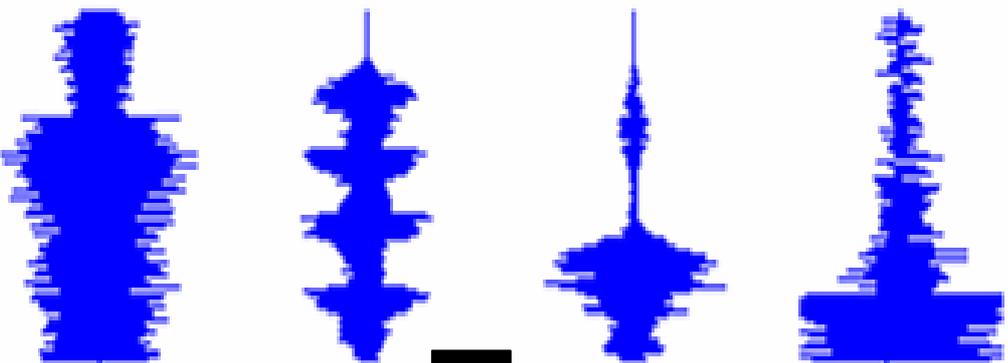
3 Séparation des signaux audio

La deuxième application mis en œuvre pour authentifier les deux algorithmes est un mélange de quatre signaux audio de trois secondes chacun. Cette application se présente comme cohérente par rapport à la précédente. Ces quatre sons ont été mélangés par une matrice mélange aléatoire dont ses éléments sont compris entre 0 et 1 et qui est générée sous *MATLABR 6.5* par le *Script "Rand"*.

Comme la première application, l'algorithme *Infomax* a été testé pour séparer ce mélange avec un blanchiment utilisant l'algorithme *ACZ*, et la fonction sigmoïdale utilisée et la fonction *logistique*, ainsi que toutes les fonctions contrastes et les deux méthodes de blanchiment qui ont été testées pour la séparation via l'algorithme *FastICA* des signaux audio sources.

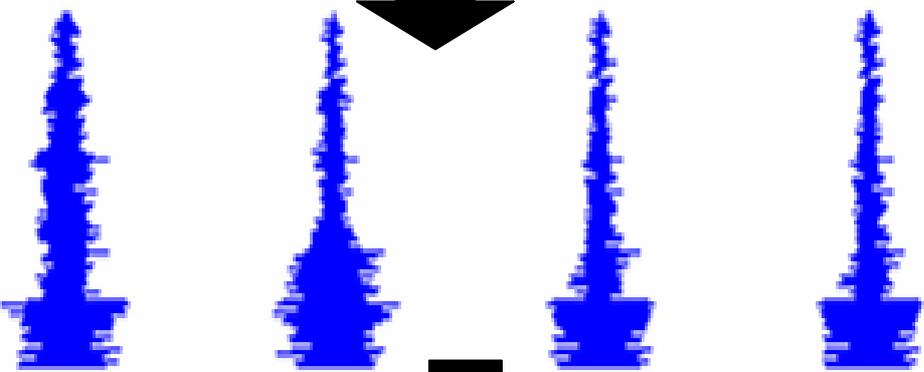
Les résultats obtenus pour cette deuxième séparation sont représentés dans les figures suivantes. Ces dernières montrent que toutes les sources audio ont été totalement séparées par tous les algorithmes, et ceci en comparant les signaux audio séparés avec les signaux audio sources on remarque qu'il sont identiques. Ce résultat peut se certifier tout en écoutant les sons après séparation avec les sons sources. On note que l'ordre des signaux sources n'est pas retenu pour les composantes indépendantes après la séparation.

Signaux audio sources



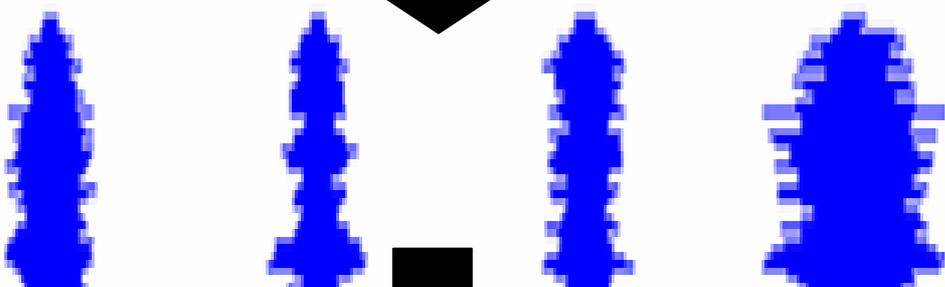
A

Signaux mélangés



M

Signaux blanchis



W

Signaux séparés

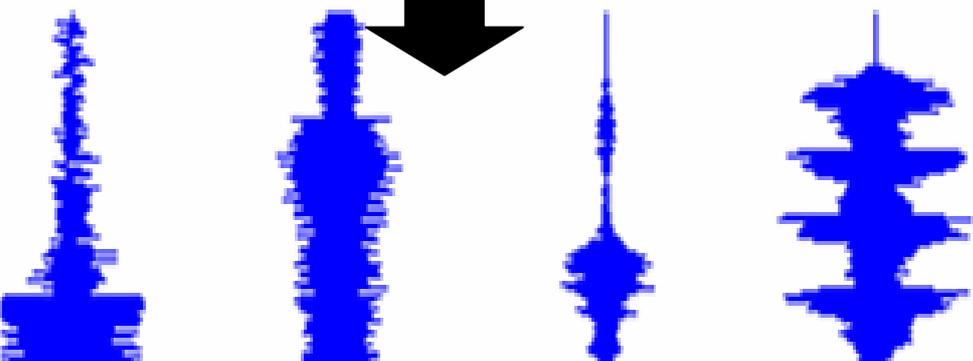


Figure 8. Séparation de signaux audio par l'algorithme *FastICA* via *ACZ* et la fonction gauss

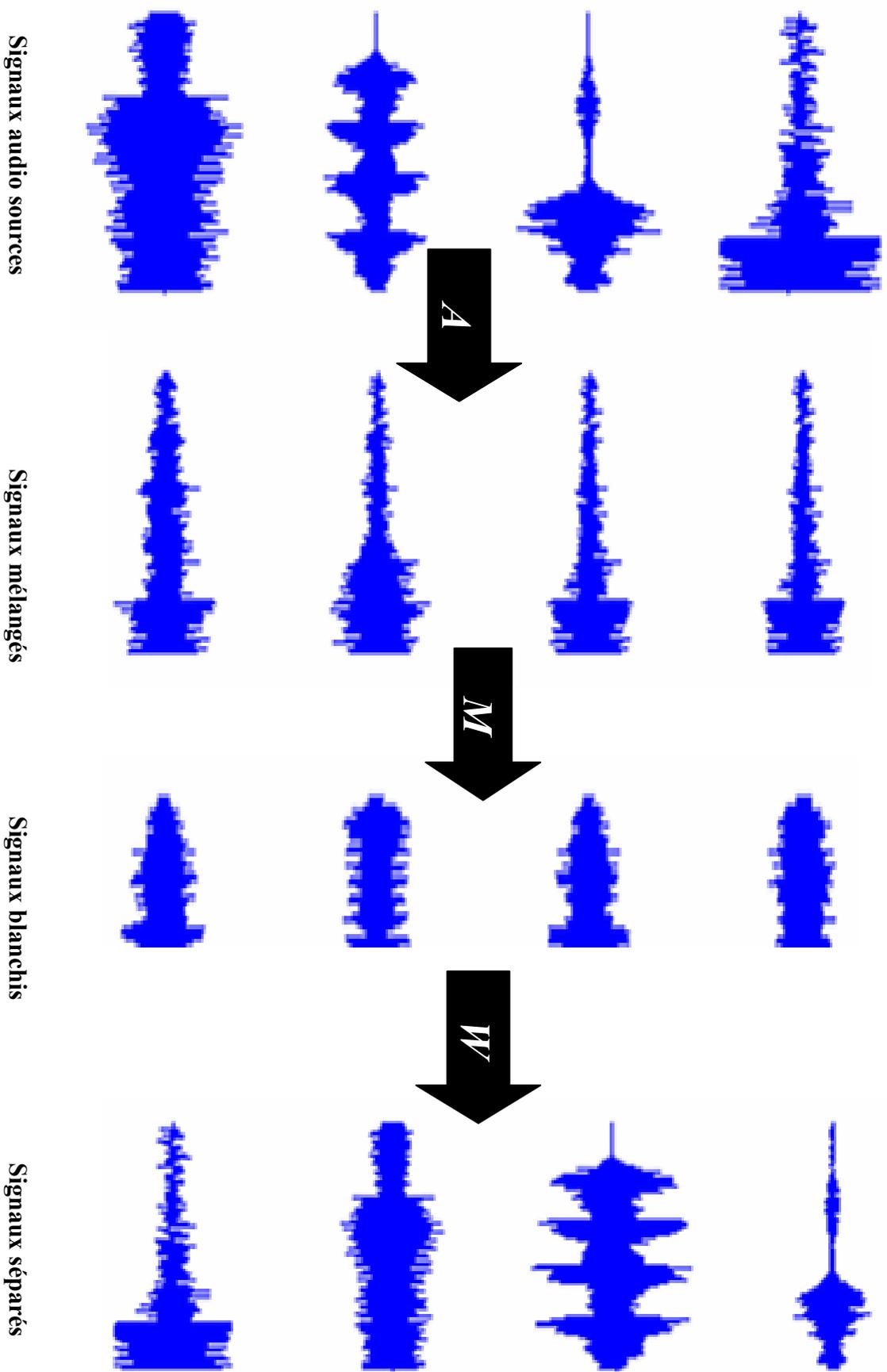


Figure 9. Séparation de signaux audio par l'algorithme *FastICA* via *ACZ* et la fonction *Puissance de trois*

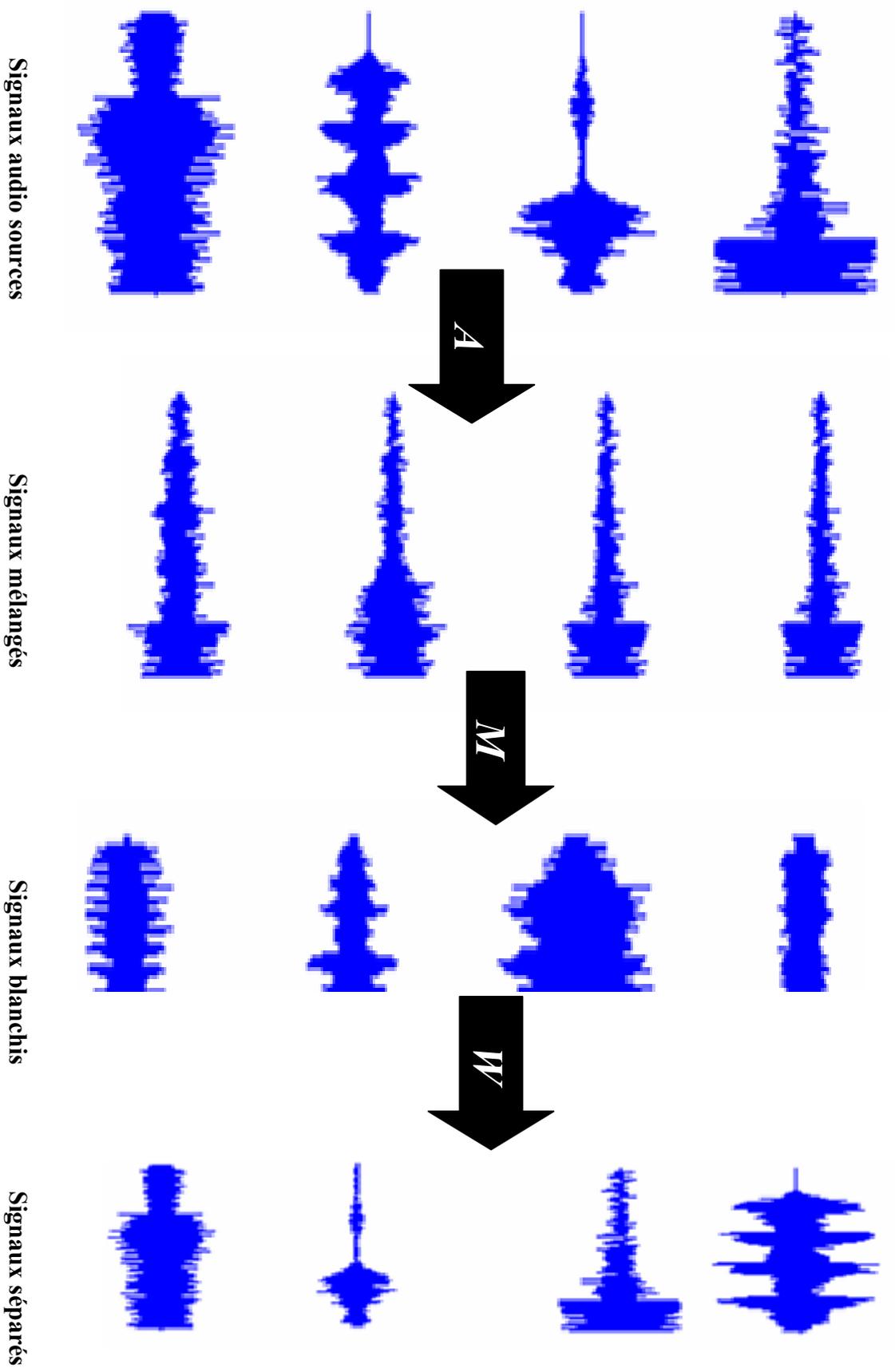


Figure.10. Séparation de signaux audio par l'algorithme *FastICA* via *ACP* et la fonction *igt*

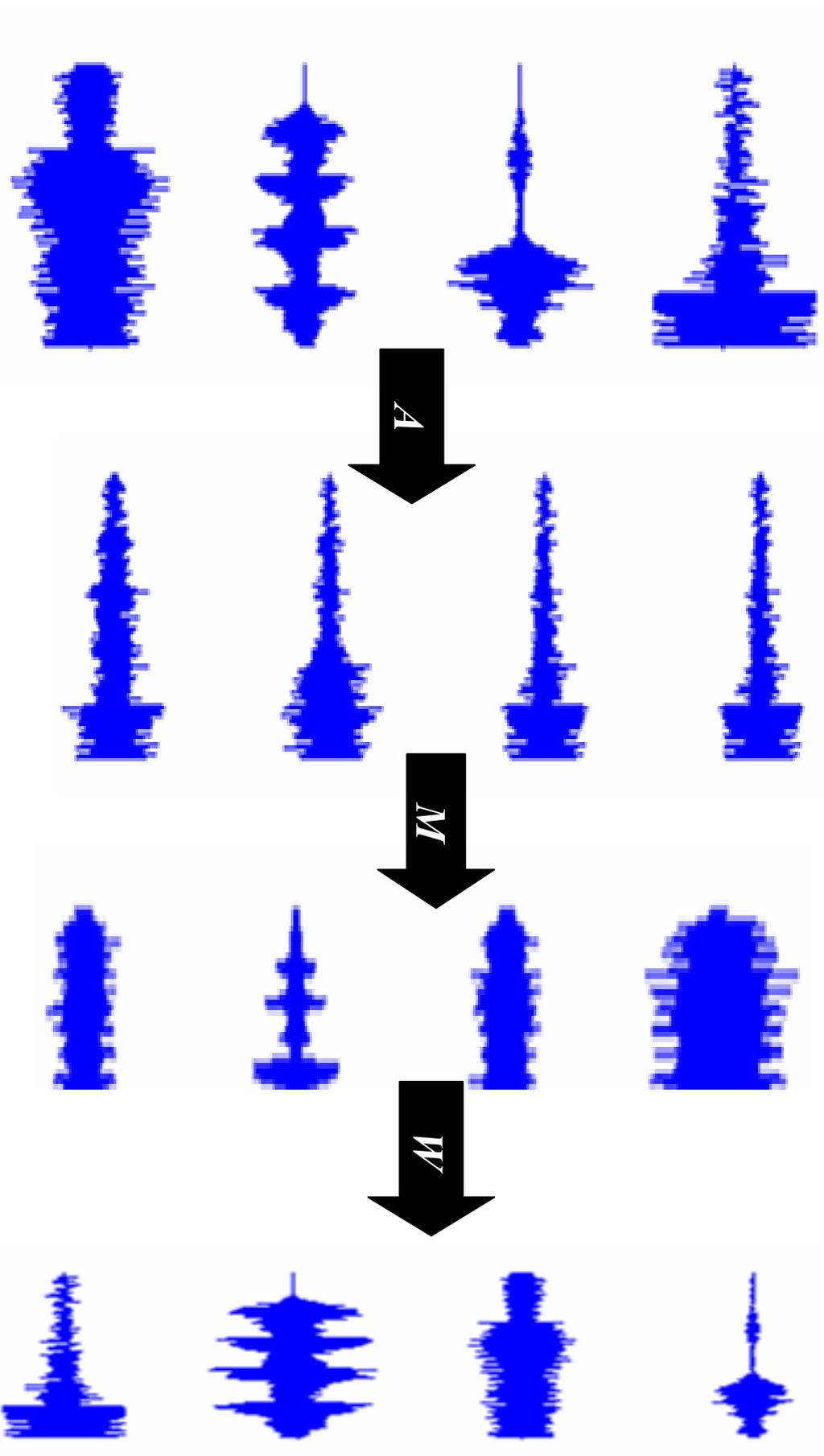


Figure. 11. Séparation de signaux audio par l'algorithme *FastICA* via *ACZ* et la fonction *igh*

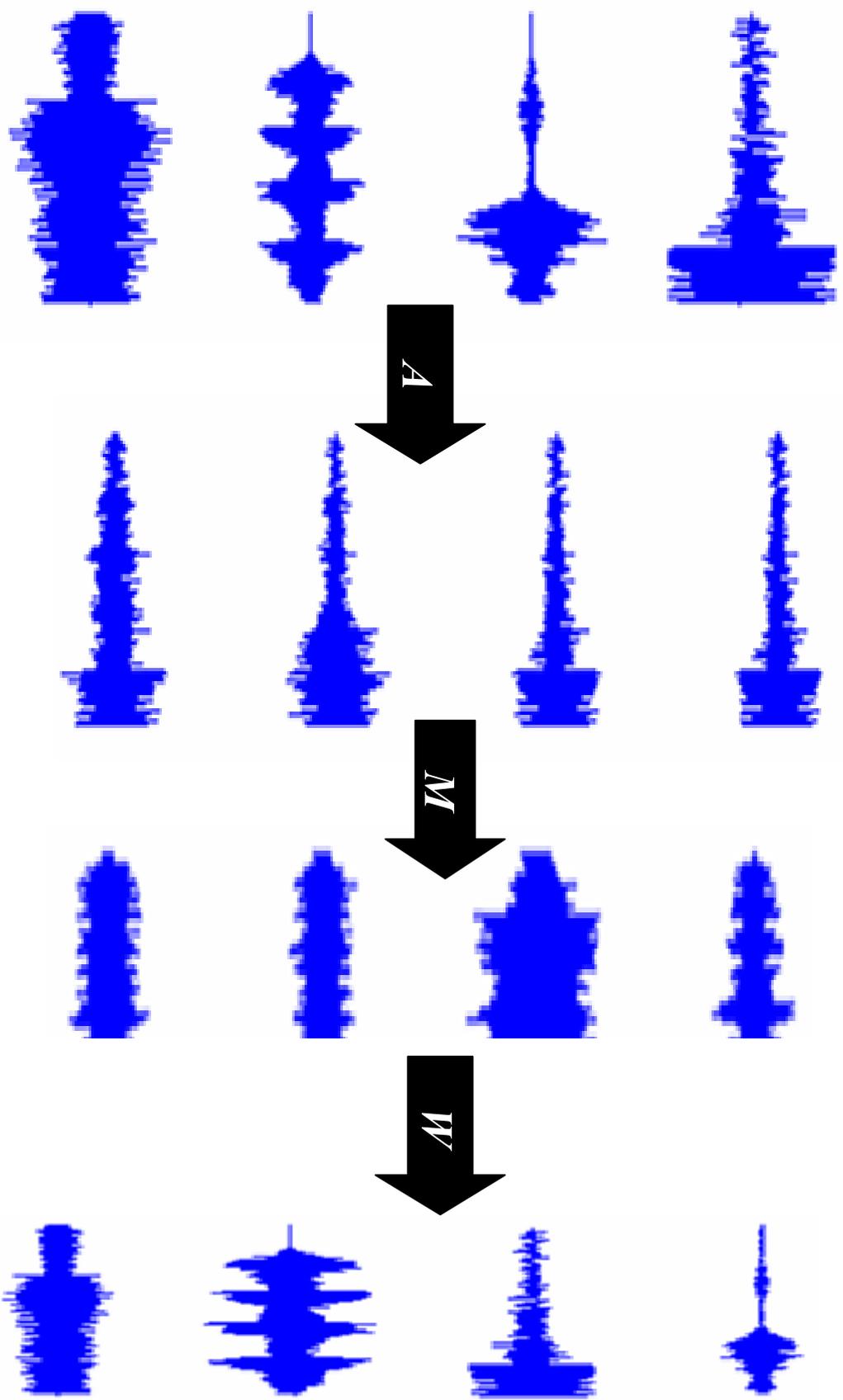


Figure. 12. Séparation de signaux audio par l'algorithme *FastICA* via *ACP* et la fonction *Puissance de trois*

aux séparés

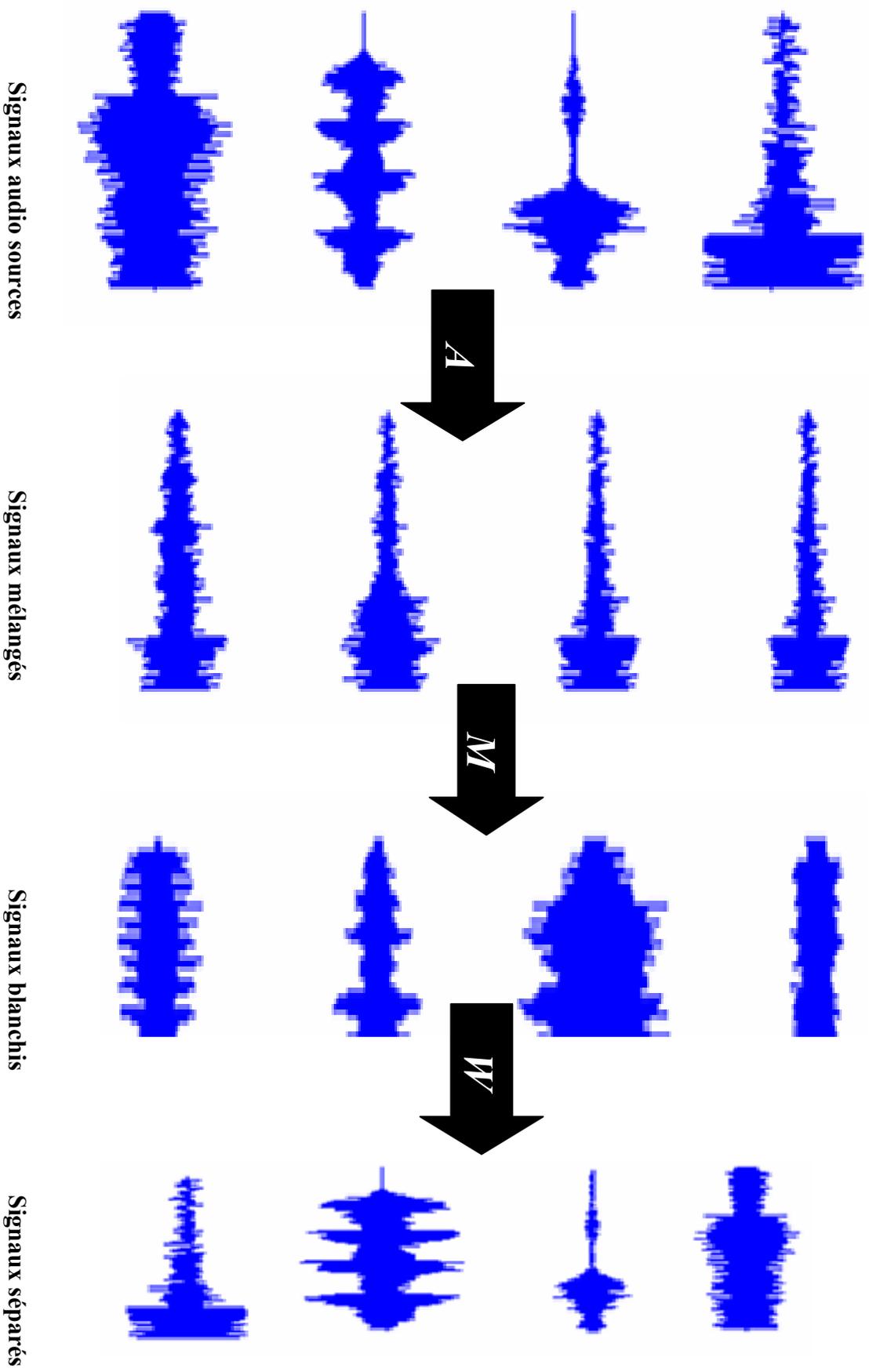
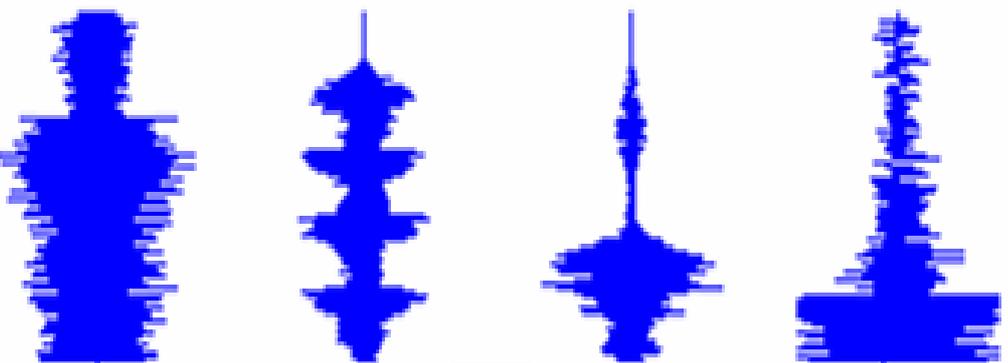
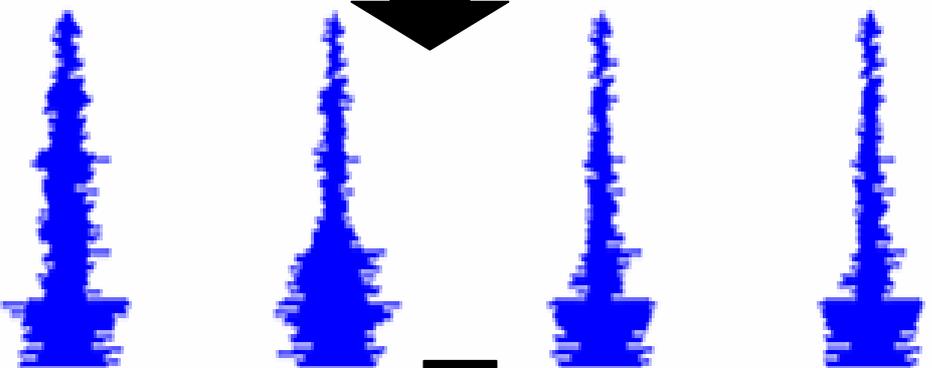


Figure.13. Séparation de signaux audio par l'algorithme *FastICA* via *ACP* et la fonction *ighi*

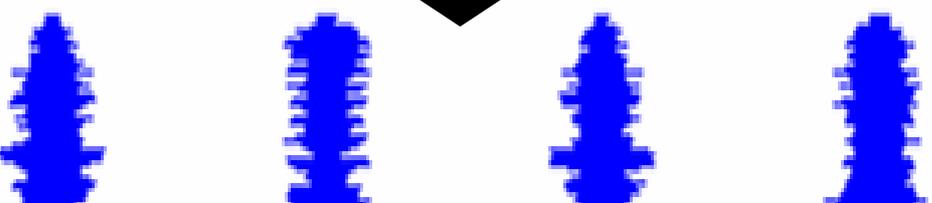
Signaux audio sources



Signaux mélangés



Signaux blanchis



Signaux séparés



Figure.14. Séparation de signaux audio par l'algorithme *Infomax*.

ANNEXE. E

INDEX

A

Activité	31, 49, 63
Cérébrale	39, 63, 71
des Composantes Indépendantes	49, 71
ACZ	15, 55
Analyse en composantes	
Indépendantes	1, 7, 17, 22, 38, 55
Principales	3, 7, 15, 17

C

Corrélation	3
-------------	---

D

Décorrélation	7, 12, 15, 17, 21
---------------	-------------------

E

Entropie	7, 9, 15, 18
Electroencéphalographie	
34	

F

fonctions	
Non quadratiques	11
Sigmoïdales	19

I

Incorrélation	5
---------------	---

M

Matrice

de Blanchiment	49
de covariance	9, 12
Diagonale	12, 13
Inverse	19
Jacobienne	21
Mélange	2, 5, 13, 23
de Séparation	49
de Corrélacion	3

Moyenne	9, 10, 12
---------	-----------

N

Néguentropie	8, 9, 10, 20
Approximation	10, 20

S

Statistiques

d'ordre supérieur	7, 16, 18
du second ordre	13, 18

Systèmes

Boucle	15, 16
Directe	3, 15, 16

V

Variance	3, 5, 8, 12
----------	-------------

Vecteur

aléatoire	7, 9
moyenne	12
observé	3, 12
poids	19, 20
propres	13

TABLE DES MATIERES

RESUME.....	
REMERCIEMENTS.....	
TABLE DES MATIERES.....	
LISTE DES SYMBOLES ET DES ABREVIATIONS.....	
LISTE DES ILLUSTRATIONS, GRAPHIQUES ET TABLEAUX.....	
INTRODUCTION.....	
1. ANALYSE EN COMPOSANTES INDEPENDANTES.....	9
1.1 Introduction.....	9
1.2 Indépendance.....	12
1.3 Principe d'Analyse en Composantes Indépendantes (<i>ACI</i>).....	15
1.4 Domaines d'applications de l'Analyse en Composantes Indépendantes.....	33
1.5 Validation des algorithmes.....	33
1.6 Conclusion.....	42
2. GENERALITES SUR L'EEG.....	43
2.1 Electroencéphalographie.....	43
2.2 Utilisation de l'Électroencéphalographie.....	46
2.3 Recherche médicale.....	48
3. CARTES D'ACQUISITION ET DE CONVERSION.....	54
3.1 Introduction.....	54
3.2 Carte de Conversion.....	55
3.3 Carte d'Interface.....	57
3.4 Acquisition des signaux EEG.....	58
3.5 Conclusion.....	60
4. TESTS ET RESULTATS.....	61
4.1 Introduction.....	61

4.2	Domaine d'application et terminologie.....	61
4.3	Acquisition des données et premier examen EEG.....	64
4.4	Séparation des artefacts.....	86
4.5	Conclusion.....	89
CONCLUSION.....		
REFERENCES.....		
ANNEXE. A.....		
ANNEXE. B.....		
ANNEXE. C.....		
ANNEXE. D.....		
ANNEXE. E.....		
INDEX.....		

LISTE DES ILLUSTRATIONS, GRAPHIQUES ET TABLEAUX

Figure 1. 1	Le problème Cocktail-Party	10
Figure 1. 2	Distribution multivariable de deux variables gaussiennes indépendantes	14
Figure 1. 3	La Distribution jointe des mélanges observés x_1 et x_2 . Axe horizontale: x_1 , axe vertical: x_2	15
Figure 1. 4	ACP à deux dimensions (a) Données Originales (b) Donnée Mélangée (c) Rotation E^T (d) Gradation par $D^{-1/2}$	22
Figure 1. 5	Les systèmes directs et les systèmes en boucle pour la séparation de sources	24
Figure 1. 6	Algorithme Infomax	28
Figure 1. 7	Algorithme FastICA	32
Figure 1. 8	Séparation de signaux périodiques par FastICA	35
Figure 1. 9	Séparation de signaux audio par FastICA	37
Figure 1. 10	Séparation de signaux périodiques par Infomax	39
Figure 1. 11	Séparation de signaux audio par Infomax	41
Figure 2. 1	L'électroencéphalogramme	44
Figure 2. 2	Echantillons des ondes cérébrales	46
Figure 2. 3	Placements des électrodes selon le système international '10/20'	47
Figure 2. 4	Structure générale d'un équipement de cartographie	49
Figure 2. 5	Analyse spectrale	51
Figure 2. 6	Carte Topographique du EEG	52
Figure 2. 7	Représentation en trois dimensions	52
Figure 3. 1	Synoptique du système	54
Figure 3. 2	Schéma synoptique de la carte de conversion	56
Figure 3. 3	Schéma synoptique de la carte d'interface	57
Figure 3. 4	Organigramme principal d'acquisition	59
Figure 4. 1	Schéma de décomposition d'analyse en composantes de l'électroencéphalogramme	63
Figure 4. 2	Montage référentiel	64

Figure 4. 3	Représentation du logiciel d'acquisition	65
Figure 4. 4. A	Représentations des signaux acquis sur l'écran	67
Figure 4. 4. B	Représentations des signaux acquis sur le papier	68
Figure 4. 5	Représentations des signaux EEG réels.	69
Figure 4. 6	Séparation des dix signaux réels EEG	70
Figure 4. 7	Schéma de décomposition d'analyse en composantes indépendantes de l'électroencéphalogramme	72
Figure 4. 8	Menu général de SeparEEG vers. 1.0.	73
Figure 4. 9	Signaux Mélangés	74
Figure 4. 10	Signaux Blanchis	75
Figure 4. 11	Signaux Séparés	75
Figure 4. 12	Organigramme de la représentation de l'activité cérébrale	76
Figure 4. 13	Activité cérébrale	77
Figure 4. 14	Organigramme de la représentation de l'activité cérébrale en deux dimensions	78
Figure 4. 15	Activité cérébrale des signaux EEG mélangés dans le plan	79
Figure 4. 16	Activité cérébrale de la projection de la première composante indépendante dans le plan	79
Figure 4. 17	Cartes du scalp dans le plan	81
Figure 4. 18	Cartes du scalp dans l'espace	81
Figure 4. 19	Organigramme de la représentation de l'activité cérébrale en trois dimensions	82
Figure 4. 20	Activité cérébrale des signaux <i>EEG</i> dans l'espace	83
Figure 4. 21	Activité cérébrale de la projection de la deuxième composante indépendante dans l'espace	83
Figure 4. 22	Activités cérébrales	85
Figure 4. 23	Montage référentiel	86
Figure 4. 24	Échantillons des signaux <i>EEG</i> , montrant les artéfacts produites par la contraction musculaire (<i>CM</i>), clignotement et la simulation limuneuse intermittente (<i>SLI</i>) avec les yeux fermés.	88
Figure 4. 25	composantes indépendantes trouvées via <i>FASTICA-ACP</i> - à partir des données <i>EEG</i> .	88