

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



**Université Hadj Lakhdar Batna**  
**Faculté de Technologie**  
**Département d'Électronique**



## **Mémoire**

**Présenté pour l'obtention du diplôme de  
MAGISTER en Électronique**

### **OPTION**

**Traitement du signal**

**Par**

**SMATTI El mouataz billah**

### **Thème**

---

**Etude et mise en œuvre des techniques pour  
la séparation aveugle des signaux**

---

Soutenu devant le jury composé de :

Pr. BENOUDJIT Nabil	Prof.	Université de Batna	Président
Pr. ARRAR Djemai	Prof.	Université de Batna	Rapporteur
Pr. BENZID Redha	Prof.	Université de Batna	Co-Rapporteur
Pr. BENYOUCEF Moussa	Prof.	Université de Batna	Examineur
Pr. SOLTANI Fouzi	Prof.	Université de Constantine	Examineur

## *Remerciements*

- ❖ Je tiens à exprimer mes vifs remerciements à mon encadreur Monsieur: **ARRAR Djemai** professeur à l'université de Batna, pour avoir dirigé mon travail, son aide m'a toujours été précieuse.
- ❖ Je voudrais exprimer, ma gratitude et ma considération à Monsieur **BENZID Redha**, Professeur à l'université de Batna, pour avoir suivi mon travail avec une attention soutenue.
- ❖ Je remercie très sincèrement Monsieur le Professeur **BENOUDJIT Nabil**, de l'université de Batna, pour l'intérêt qu'il a manifesté pour ce travail et d'avoir accepté de présider le jury de ce travail.
- ❖ J'adresse mes vifs remerciements à **BENYOUCEF Moussa**, Professeur au département d'électronique de l'université de Batna, pour avoir accepté d'examiner ce travail
- ❖ Je voudrais également exprimer mes vifs remerciements à monsieur **SOLTANI Fouzi** Professeur à l'université de Constantine pour l'honneur d'assister à ce jury et d'enrichir le manuscrit.
- ❖ Enfin, j'exprime mes remerciements à tous ceux qui ont contribué de près ou de loin à l'élaboration de ce travail.

## Sommaire

Remerciements

Sommaire.....i-1

Liste des figures.....ii-1

Liste des tableaux.....iii-1

Liste des symboles.....iv-1

Introduction générale.....1

**Chapitre I: Concepts généraux sur les probabilités et aux statistiques d'ordre supérieur.....4**

1.1: Introduction.....5

I.2: Probabilités.....5

    I.2.1: Ensemble fondamental et événements.....5

    I.2.2: Diagramme de Venn.....6

    I.2.3: La notion de la probabilité et ses axiomes .....7

    I.2.4: Probabilité conditionnelle et la règle de Bayes.....7

    I.2.5: Événements indépendants.....8

I.3: Variables aléatoires.....8

    I.3.1: Événements définis par les variables aléatoires.....9

    I.3.2: Fonction de distribution (répartition) .....10

    I.3.3: Fonction de densité de probabilité (PDF).....11

    I.3.4: Espérance et variance d'une variable aléatoire.....11

I.4: Variables aléatoires multiples.....13

    I.4.1 Fonctions de distribution (répartition) conjointe .....14

    I.4.2 Fonctions de densité de probabilité conjointe la PDF conjointe.....15

I.4.3:Distribution conditionnelle .....	15
I.4.4:Variables aléatoires indépendantes.....	16
I.4.5: Covariance et le coefficient de corrélation .....	16
I.5: Fonctions caractéristiques et Cumulants.....	17
I.5.1:Fonction d'une variable aléatoire.....	17
I.5.2:Fonction de deux variables aléatoires.....	18
I.5.3:Fonction caractéristique d'une seule variable aléatoire.....	19
I.5.4: Fonctions caractéristique conjointe.....	21
I.5.5 Moments et cumulants.....	22
I.5.5.1: Moment et cumulants d'une seule variable aléatoire.....	22
I.5.5.2: Relation entre moments et cumulants d'une seule variable aléatoire.....	23
I.5.5.3: Moment et cumulants de plusieurs variables aléatoires.....	23
I.5.5.4: Relation entre moments et cumulants d'un vecteur aléatoire .....	23
I.6:Processus stochastique (aléatoire).....	24
I.6.1:Caractérisation d'un processus stochastique (aléatoire).....	25
I.6.1.1:Description probabiliste.....	25
I.6.1.2:Espérance, corrélation et fonction de covariance.....	25
I.6.2: Processus stochastique stationnaire.....	26
I.6.3:Processus stochastiques independants.....	26
I.6.4:Histogramme.....	26
I.7:Théorème de la limite centrale et le Kurtosis.....	28
I.7.1:Kurtosis.....	28
I.7.2: Théorème de la limite centrale.....	29
I.8:Conclusion .....	30

<b>Chapitre II Concepts généraux sur la SAS.....</b>	<b>31</b>
II. 1:Introduction.....	32
II.2:Vue historique.....	33
II.3: Le problème de la séparation de sources .....	34
II.4.Caractéristiques du système de mélange.....	35
II.4.1:Linéaire ou non linéaire.....	35
II.4.2: Instantané ou convolutif.....	35
II.4.3 : Selon le nombre de sources et le nombre d'observations.....	35
II.5:Principe de la solution du problème de la SAS.....	36
II.6:L'analyse en composantes indépendantes (ICA).....	37
II.6.1:Théorème de (Darmois-Skitovich).....	37
II.6.2:Définition de la (ICA).....	38
II.6.3: Théorème de séparabilité.....	38
II.7:Le prétraitement (WHITENING-le blanchiment) .....	38
II.8: Critères pour l'analyse en composantes indépendantes.....	40
II.8.1:Critère basé sur l'information mutuelle.....	40
II.8.1.1:Entropie.....	40
II.8.1.2:Entropie conditionnelle .....	41
II.8.1.3:Information mutuelle.....	41
II.8.2:Critères basé sur les statistiques d'ordre supérieur .....	41
II.8.3:Critère basé sur la maximisation de la non gaussienne .....	42
II.9: Systèmes de séparation.....	43
II.10:Applications de la séparation de sources.....	44

II.10.1:Le problème du Cocktail Party .....	44
II.10.2 : Application de la SAS pour les systèmes de Communications numériques.....	45
II.11:Conclusion.....	46
<b>Chapitre III: Algorithmes directs et adaptatifs pour les mélanges linéaires instantanés.....</b>	<b>47</b>
III.1:Introduction.....	48
III.2:Méthodes directes (hors ligne).....	48
III.2.1:Approche géométrique de (Ali Mansour).....	49
III.2.1.1:Transformation.....	49
III.2.1.2:Rotation.....	50
III.2.2: Méthode basée sur les cumulants croisés (Ali Mansour).....	52
III.2.2.1:Modèle du mélange.....	52
III.2.2.2:Polynôme du quatrième degré pour le cas deux sources.....	53
III.2.3 Méthode basée sur les cumulants croisés (Taro Yamaguchi).....	53
III.2.3.1: Modèle du séparateur.....	53
III.2.3.2: Polynôme du quatrième degré et les expressions de ses coefficients .....	54
III.3:Méthodes adaptatives (en ligne).....	55
III.3.1:Principe des méthodes en ligne.....	55
III.3.2:Approche neuronale de (Herault et Jutten).....	56
III.3.2.1:Réseau direct.....	56
III.3.2.1.1: Solutions pour le cas de deux sources.....	57
III.3.2.1.2: Loi d'adaptation .....	57
III.3.2.2:Réseau récurrent.....	58
III.3.2.2.1: Solutions pour le cas de deux sources.....	58
III.3.3:Algorithme de (A-Cichocki et R Unbehauen).....	59

III.3.3.1: Problèmes des réseaux de (Herault et Jutten).....	59
III.3.3.2: Principe de l'algorithme de (A-Cichocki et R Unbehauen).....	59
III.3.3.3: Loi d'adaptation .....	60
III.3.4: Algorithme de (A-Cichoki et S-Amari).....	61
III.3.4.1: Principe de l'algorithme de (A-Cichoki et S-Amari).....	61
III.3.4.2: Loi d'adaptation .....	61
<b>Chapitre IV Implémentation des algorithmes et comparaison des résultats.....</b>	<b>63</b>
IV. 1 : Introduction.....	64
I.V.2: Hypothèse sur le système et les signaux sources et le critère de comparaison.....	65
I.V.3: Implémentation des Méthodes directes.....	66
I.V.3.1: Méthode géométrique (Ali Mansour).....	66
IV.3.1.1: Deux sources (directe).....	66
IV.3.1.2: Influence de l'angle de rotation sur la valeur du Kurtosis.....	68
IV.3.1.3: Deux sources:(max Kurtosis).....	69
IV.3.1.4: Trois sources:(max Kurtosis).....	72
IV.3.2: Méthode basée sur les cumulants croisés (Ali Mansour).....	73
IV.3.2.1: Résultats de l'application de l'algorithme .....	73
IV.3.2.2: Décision .....	74
IV.3.3: Méthode basée sur les cumulants croisés (Taro Yamaguchy).....	75
IV.3.3.1: Résultats de l'application de l'algorithme.....	75
IV.3.3.2: Décision.....	76
IV.4: Comparaison des résultats des méthodes directes.....	77
IV.5: Implémentation des Méthodes Adaptatives.....	78

IV.5.1:Algorithme de (Herault et Jutten) réseau direct.....	78
IV.5.1.1:(Deux sources).....	78
IV.5.1.1.1:Convergence et influence du gain d'apprentissage.....	78
IV.5.1.1.2:Valeur du Kurtosis pour différentes matrices.....	81
IV.5.1.1.3:Problème des échelles mal conditionnées .....	82
IV.5.1.2:(Trois sources).....	82
IV.5.2:Algorithme de (Herault et Jutten ) réseau récurrent.....	83
IV.5.2.1:(Deux sources).....	83
IV.5.2.2:Remarques concernant les résultats des deux réseaux.....	83
IV.5.3:Algorithme de (A-Cichocki et R- Unbehauen).....	84
IV.5.3.1:(Deux sources).....	84
IV.5.3.1.1:Valeur du Kurtosis pour différentes matrices.....	85
IV.5.3.2:(Trois sources).....	86
IV.5.3.3:(Quatre sources).....	87
IV.5.4:Algorithme de (A-Cichoki et S-Amari).....	88
IV.5.4.1:(Deux sources).....	88
IV.5.4.2:(Trois sources).....	89
IV.5.4.3:(Quatre sources).....	90
IV.6:Comparaison des méthodes adaptatives.....	90
IV.7:Comparaison des méthodes adaptatives et directes .....	92
IV.8:Présence d'une source gaussienne .....	93
IV.8.1:Présence d'une source gaussienne (méthodes directes).....	93
IV.8.1.1:Solutions.....	93



IV.8.1.2:Décision.....	94
IV.8.2:Présence d'une source gaussienne (méthodes adaptatives).....	94
IV.8.2.1:Deux sources.....	94
IV.8.2.2:Trois sources.....	95
IV.9:Influence du gain d'apprentissage sur le nombre de sources.....	96
IV.9.1:Cinq sources.....	96
IV.10:Conclusion.....	97
Conclusion générale.....	98
Bibliographie .....	100
Appendice A.....	103
Appendice B.....	105
Résumé	

## **Liste des figures**

### **Chapitre: I**

Figure 1.1:Diagramme de Venn.....	6
Figure I.2:Représentation des variables aléatoires d'une seule dimension.....	9
Figure I.3:Fonction de répartition de l'expérience du lancement d'une pièce 3 fois.....	10
Figure I.4:Distributions gaussiennes.....	12
Figure I.5: Représentation des variables aléatoires multiples.....	13
Figure I.6:Exemple d'une variable aléatoire générée à partir d'une seule variable aléatoire....	18
Figure I.7: Exemple d'une variable aléatoire générée à partir de deux variables aléatoires....	19
Figure I.8:Exemple d'une fonction caractéristique d'une seule variable aléatoire.....	21
Figure I.9: Schéma simplifié de la représentation d'un processus aléatoire .....	24
Figure I.10:Exemple de la construction d'un histogramme.....	27
Figure I.11: Représentation de plusieurs distributions gaussiennes et le calcul du Kurtosis....	28
Figure I.12:Exemple du théorème de la limite centrale pour des signaux uniformes.....	29
Figure I.13:Exemple du théorème de la limite centrale pour des signaux sinusoïdaux.....	30

### **Chapitre: II**

Figure II.1: Le problème général de la séparation de sources.....	35
Figure II.2- Exemple démonstratif sur le but de la séparation de sources.....	36
Figure II.3: La procédure du blanchiment.....	40
Figure II.4: Mélange entre la distribution uniforme et la distribution Laplacienne.....	42
Figure II.5: Les systèmes pour la séparation de sources.....	43
Figure II.6: Le problème du Cocktail party.....	44
Figure II.7: Le problème de la séparation de sources pour les réseaux de communication ....	45

**Chapitre: III**

Figure III.1: Transformation des signaux mélanges en des signaux non corrélés.....50

Figure III.2: La procedure de rotation (deux sources).....51

Figure III.3: La procedure de rotation (trois sources).....52

Figure III.4: Principe des méthodes en ligne pour la SAS.....55

Figure III.5: Réseau neuronal direct de (Herault et Jutten).....56

Figure III.6: Réseau neuronal direct pour deux sources de (Herault et Jutten).....57

Figure III.7: Réseau neuronal récurrent de (Herault et Jutten).....58

Figure III.8: Réseau neuronal récurrent pour deux sources de (Herault et Jutten).....58

Figure III.9: Schéma de principe du réseau de (A-Cichocki et R Unbehauen).....60

Figure III.10: Réseau neuronal pour la SAS de (A-Cichocki et RUnbehauen).....60

Figure III.11: Réseau neuronal pour la SAS de (A-Cichocki et S-Amari).....61

**Chapitre IV**

Figure IV.1: Différents signaux utilisés pour la simulation.....64

**IV.I:Méthodes directes:**

**IV.I.1:Méthode géométrique (directe)**

Figure IV.2: Distributions des amplitudes (signaux sinusoïdaux).....66

Figure IV.3:Résultats de l'application (signaux sinusoïdaux).....67

Figure IV.4: Distributions des amplitudes (signal sinusoïdal et signal carré).....67

Figure IV.5: Résultats de l'application (signal sinusoïdal et signal carré).....68

Figure IV.6: Influence de ( $\alpha$ ) sur ( $K$ ) (signaux sinusoïdaux).....69

**IV.I.2:Méthode géométrique deux sources (max Kurtosis)**

Figure IV.7: Distributions des amplitudes (signaux sinusoïdaux).....69

Figure IV.8: Résultats de l'application (deux signaux sinusoïdaux) .....	70
Figure IV.9: Influence de $(\alpha)$ sur $(K)$ (signaux non sinusoïdaux).....	70
Figure IV.10: Distributions des amplitudes (signaux non sinusoïdaux).....	71
Figure IV.11: Résultats de l'application (deux signaux non sinusoïdaux).....	71
 <b>IV.I.3: Méthode géométrique trois sources (max Kurtosis)</b>	
Figure IV.12: Distributions des amplitudes.....	72
Figure IV.13: Résultats de l'application.....	72
 <b>IV.I.4: Méthode basée sur les cumulants (Ali mansour)</b>	
Figure IV.14: Résultats de l'application (les 4 solutions).....	73
Figure IV.15: Résultats de l'application (décision).....	74
Figure IV.16: Valeur finale du Kurtosis (décision).....	74
 <b>IV.I.5: Méthode basée sur les cumulants (Taro Yamaguchy)</b>	
Figure IV.17: Résultats de l'application (les 4 solutions).....	75
Figure IV.18: Résultats de l'application (décision).....	76
Figure IV.19: Valeur finale du Kurtosis (décision).....	76
 <b>IV.II: Méthodes adaptatives:</b>	
 <b>IV.II.1: Algorithme d'Hérault et Jutten réseau direct</b>	
 <b>Deux sources</b>	
Figure IV.20: Valeurs finales du Kurtosis pour chaque valeur du gain d'apprentissage(A1)...	78
Figure IV.21: Valeurs des poids du réseau(A1).....	79
Figure IV.22: Résultat de séparation(A1).....	79
Figure IV.23: Valeurs finales du Kurtosis pour chaque valeur du gain d'apprentissage(A2)...	80
Figure IV.24: Valeurs des poids du réseau(A2).....	80

Figure IV.25: Résultat de séparation(A2).....81

Figure IV.26: Résultat pour différentes matrices de mélanges.....81

Figure IV.27: Problème des échelles mal conditionnées .....82

**Trois sources**

Figure IV.28: Valeurs finales du Kurtosis pour chaque valeur du gain d'apprentissage .....82

Figure IV.29: Résultat de séparation.....82

**IV.II.2: Algorithme de Héroult et Jutten réseau récurrent**

**Deux sources**

Figure IV.30: Valeurs finales du Kurtosis pour chaque valeur du gain d'apprentissage.....83

Figure IV.31: Résultat de séparation .....83

**IV.II.3: Algorithme de A-Cichocki et R -Unbehauen**

**Deux sources:**

Figure IV.32: Valeurs du Kurtosis en fonction du nombre d'itérations.....84

Figure IV.33: Résultat de séparation .....84

Figure IV.34: Valeurs du Kurtosis pour différentes matrices .....85

**Trois sources:**

Figure IV.35: Valeurs du Kurtosis en fonction du nombre d'itérations.....86

Figure IV.36: Résultat de séparation .....86

**Quatre sources**

Figure IV.37: Valeurs du Kurtosis en fonction du nombre d'itérations.....87

Figure IV.38: Résultat de séparation .....87

**IV.II.4: Algorithme de A-Cichoki et S-Amari**

**Deux sources**

Figure IV.39: Valeurs du Kurtosis en fonction du nombre d'itérations.....88

Figure IV.40: Résultat de séparation .....88

**Trois sources**

Figure IV.41: Valeurs du Kurtosis en fonction du nombre d'itérations.....89

Figure IV.42: Résultat de séparation .....89

**Quatre sources:**

Figure IV.43: Valeurs du Kurtosis en fonction du nombre d'itérations.....90

Figure IV.44: Résultat de séparation .....90

**IV.III: Présence d'une source gaussienne**

**IV.III.1: Méthodes directes**

Figure IV.45: Résultats de l'application (les 4 solutions).....93

Figure IV.46: Résultats de l'application (décision).....94

**IV.III.1: Méthodes adaptatives**

**Deux sources**

Figure IV.47: Valeurs du Kurtosis en fonction du nombre d'itérations.....94

Figure IV.48: Résultat de séparation .....95

**Trois sources:**

Figure IV.49: Valeurs du Kurtosis en fonction du nombre d'itérations.....95

Figure IV.50: Résultat de séparation .....95

**IV.IV: Influence du gain d'apprentissage (cinq sources)**

Figure IV.51: Résultat de séparation de cinq sources.....96

*Liste des tableaux*

**TABLEAU 1:** Comparaison des méthodes directes.....77

**TABLEAU 2:** Comparaison des méthodes adaptatives.....91

**TABLEAU 3:** Comparaison des méthodes directes et adaptatives.....92

**Liste des Symboles**

$P(A)$	probabilité de l'événement A
$P(A/B)$	probabilité conditionnelle de l'événement A
$P(A \cap B)$	probabilité conjointe des événements A et B
$n(A)$	nombre d'éléments élémentaires de l'événement A
$X(\zeta)$	variable aléatoire (image de l'élément $\zeta$ )
$(x_1 \leq X \leq x_2)$	l'événement entre les deux valeurs $x_1$ et $x_2$
$P(X \leq x)$	probabilité de l'événement $X \leq x$
$F_x(X)$	fonction de répartition de la variable aléatoire X
$f_x(x)$	fonction de densité de probabilité de la variable aléatoire X
PDF	Probability density function
$\mu_x, E\{X\}$	Esperance mathématique de la variable aléatoire X
$\sigma^2, VAR(X)$	Variance de la variable aléatoire X
$F_{XY}(x, y)$	Fonction de répartition conjointe des variable aléatoires X et Y
$P(X \leq x, Y \leq y)$	Probabilité de l'événement $X \leq x$ et $Y \leq y$
$f_{xy}(x, y)$	fonction de densité de probabilité conjointe des variables aléatoires X et Y
$f_{y/x}(y/x)$	fonction de densité de probabilité conditionnelle de la variable Y
$m_{kn}, E\{X^k Y^n\}$	le (k,n) nième moment des variables aléatoires X et Y
$Cov(X, Y)$	la covariance des deux variables aléatoires X et Y
$\rho(X, Y)$	facteur de corrélation des deux variables aléatoires X et Y
$\psi_x(\omega)$	la première fonction caractéristique de la variable aléatoire X
$\psi_{xy}(\omega_1, \omega_2)$	la première fonction caractéristique conjointe des variables aléatoire X et Y
$Y_x(\omega)$	la deuxième fonction caractéristique de la variable aléatoire X
$C_k$	cumulant d'ordre k
$C(x_1^{k1}, \dots, x_p^{kp})$	cumulant croisé des variables aléatoires $x_1, \dots, x_p$
$C(x_1, x_2, x_3, x_4)$	cumulant croisé du quatrième ordre des variables aléatoires $x_1, x_2, x_3, x_4$
$F_x(x_1; t_1)$	fonction de répartition du premier ordre du processus stochastique X(t)
$R_x(t, s)$	fonction d'auto corrélation du processus stochastique X(t)
$K_x(t, s)$	fonction d'auto covariance du processus stochastique X(t)
K	Kurtosis
SAS	Séparation aveugle de sources
ICA	Independent component analysis
MIMO	Multi input Multi output



$A, H$	matrice de mélange
$W$	matrice de séparation
$S(n)$	vecteur source
$Y(n)$	vecteur mélange
$E(n)$	vecteur estimation
$H(\alpha)$	entropie
$H(\alpha \beta)$	entropie conditionnelle
$I(\alpha, \beta)$	information mutuelle
$R(\alpha)$	matrice de rotation (deux dimensions)
$R_x(\alpha)$	matrice de rotation (trois dimensions) selon l'axe x
$R_y(\alpha)$	matrice de rotation (trois dimensions) selon l'axe y
$R_z(\alpha)$	matrice de rotation (trois dimensions) selon l'axe z

## *Introduction générale*

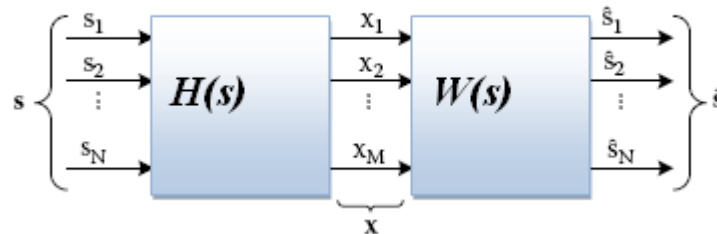
L'échange d'informations dans les systèmes de communications n'arrête pas de s'accroître dans la vie quotidienne, et le traitement du signal s'est introduit et joue un rôle très important dans le domaine de l'extraction de l'information, et il est impliqué dans ce domaine d'une façon considérable en raison de l'augmentation des demandes de systèmes de visualisation et d'enregistrement de haute qualité et fiabilité.

le traitement du signal a fondé initialement sa problématique sur des théories mathématiques en se basant sur les probabilités et l'analyse fonctionnelle [6], la théorie des probabilités a été mise en contribution dans le domaine du traitement du signal d'une façon indispensable et spécifiquement pour le traitement des signaux aléatoires (non déterministes), cette branche est riche en mathématiques depuis son origine dans la théorie des jeux en 17<sup>ième</sup> siècle par les fondateurs qui furent Bernoulli, Fermat, Pascal et d'autres [6]. Les concepts de base pour la théorie de la probabilité se reposent sur la notion de la variable aléatoire, où cette dernière est une représentation mathématique ou une projection des événements de la probabilité dans un autre domaine, cette représentation a pour but de simplifier et de ramener le problème à un traitement mathématique [6] [10].

Le traitement du signal aveugle est maintenant l'un des axes les plus importants, il est abordé avec des fondations théoriques solides, et appliqué dans beaucoup d'applications. En effet le traitement aveugle des signaux est devenu un sujet de recherche et de développement très important dans plusieurs domaines spécialement en ingénierie biomédicale, amélioration de la qualité pour la parole, la télédétection, exploration sismologique et de données. etc. Principalement le traitement aveugle du signal n'utilise pas les techniques de l'apprentissage supervisé [2], ni les connaissances préalables sur les signaux qui doivent être traités, ce type de du traitement inclue trois domaines majeurs (BSS, blind signal separation) la séparation aveugle des signaux, (ICA, independent component analysis) l'analyse en composantes indépendantes et la déconvolution aveugle des canaux. Les deux premiers domaines sont le sujet discutés dans ce travail.

En traitement du signal, le principal problème est de trouver une relation liant les signaux sources et leurs transformations avant d'être captés, ces signaux transformés on les appelle les mélanges des signaux sources, et ces mélanges sont principalement la conséquence de la

détérioration du signal utile par le canal de transmission et le mélange des signaux entre eux, donc le but de ce travail peut être résumé dans cette figure simplifiée:



À partir de cette figure, on a un ensemble de données originales ( $s$ ) qu'on veut les transmettre ou les extraire et on n'a pas un accès direct pour les obtenir, car elles sont transformées par un système ( $H$ ) en d'autres données ( $x$ ) que on les nomme les observations [7] [8].

$$x = H(s)$$

Notre tâche est de trouver un système inverse ( $W$ ) qui doit être appliqué aux observations afin d'avoir l'estimation la plus proche des données originales.

$$\hat{s} = W(x)$$

Pour cela il est nécessaire de se baser sur certains critères pour recouvrer les données originales.

Après cette brève introduction sur le problème, ce mémoire est organisé et devisé en quatre chapitres, dans le but de présenter les explications et les solutions proposées pour le problème de la séparation de sources.

On commence par le premier chapitre qui est consacré aux notions de bases sur la théorie des probabilités et les statistiques d'ordre supérieur. On va introduire dans ce chapitre le concept de la variable aléatoire ainsi que le concept de la fonction caractéristique, qui est un outil essentiel pour les mesures statistiques de la variable aléatoire, tel que la mesure des moments et les cumulants. Une brève introduction sur les processus stochastiques a été introduite dans ce chapitre avec quelques descriptions probabilistes sur ces derniers. Et le chapitre se termine par énoncer le théorème de la limite centrale, et aussi par la mesure du Kurtosis qui est le critère qu'on a choisi pour la décision et la comparaison dans ce travail.

Le deuxième chapitre est une introduction générale sur le problème de la séparation de sources où on va le présenter par sa forme la plus simple, qui est la séparation des mélanges linéaires instantanés et qui sont un cas particulier du mélange convolutif [8]. On introduit aussi quelques critères pour l'achèvement de la (SAS) tel que l'information mutuelle et le critère basé sur les statistiques d'ordre supérieur. À la fin de ce chapitre nous donnerons quelques domaines d'applications qui utilisent les techniques de la séparation de sources.

Le troisième chapitre inclut les célèbres algorithmes proposés pour la séparation aveugle de sources, ces algorithmes sont divisés en deux catégories, les algorithmes directs et les algorithmes adaptatifs. Pour les algorithmes directs on va discuter deux approches basées sur l'estimation de la matrice de mélange en utilisant la notion des cumulants croisés pour l'indépendance statistique [23], ces algorithmes sont l'algorithme de Ali mansour et l'algorithme de Taro Yamaguchi, le troisième algorithme direct est l'algorithme Ali mansour qui se base sur les techniques géométriques pour la séparation après l'obtention de la non corrélation des signaux mélanges [22]. Pour les algorithmes adaptatifs on commencera par l'algorithme de (J.Herault et C.Jutten) qui est le premier algorithme proposé pour le problème de la séparation des signaux [8], ensuite on conclura par deux algorithmes célèbres qui sont plus performants, ces algorithmes sont l'algorithme de (A-Cichocki et R-Unbehauen) et l'algorithme de (A-Cichoki et S-Amari).

Le dernier chapitre est consacré pour la mise en œuvre des algorithmes par simulation, où on a utilisé le langage de programmation matlab, et on va faire dans cette partie une étude comparative des différents algorithmes en y'ajoutant quelques remarques.

Finalement on terminera ce travail par une conclusion générale avec des perspectives envisagées.

# ***Chapitre I***

***Concepts généraux sur  
les probabilités  
et aux  
statistiques d'ordre supérieure***

### **I.1: Introduction:**

Dans le but de présenter l'intérêt de la théorie des probabilités et les statistiques d'ordre supérieur pour la séparation aveugle de sources (SAS), il est nécessaire de définir quelques notions de base sur ce domaine des mathématiques. On va focaliser ce chapitre sur la notion de l'indépendance statistique des variables aléatoires, et ce chapitre est restreint que sur les variables aléatoires réelles, afin de présenter les notions d'une façon simple. On définira aussi le Kurtosis qui est une mesure de la distance d'un processus aléatoire par rapport à la gaussienne [18], et qui est aussi utilisé comme un critère de décision pour atteindre la séparation dans ce travail.

### **I.2: Probabilités**

L'étude de la probabilité provient de l'analyse de certains jeux de chance [6], et elle a trouvé l'application dans beaucoup de domaines scientifiques et de l'ingénierie, pour cela on doit définir les concepts fondamentaux de cette branche des mathématiques.

#### **I.2.1: Ensemble fondamental et événements**

L'ensemble  $S$  de tous les résultats possibles d'une expérience donnée est appelé ensemble fondamental. Chaque point élémentaire de  $S$  s'appelle un résultat. Un événement  $A$  est un ensemble de résultats, ou en d'autres termes, un sous-ensemble de l'ensemble fondamental  $S$ . Pour les sous ensembles qui contiennent un seul élément on les appelle les événements élémentaires, on note aussi que l'ensemble vide  $\emptyset$  est appelé l'événement impossible et l'ensemble universel  $S$  est appelé l'événement certain. On peut utiliser l'algèbre des ensembles pour former de nouveaux événements en utilisant les opérations élémentaires tel que l'intersection et l'union [4] [6], et pour clarifier ce concepts, considérons le lancement d'une pièce de monnaie une seule fois. Toutes les possibilités sont l'ensemble fondamentale  $S = \{P, F\}$ , où  $P$  est l'événement de l'apparition du coté pile, et  $F$  est l'apparition du coté face. Si par exemple on lance cette pièce deux fois alors l'ensemble fondamental  $S$  est différent,  $S = \{PP, PF, FP, FF\}$  et comme exemple  $PF$  c'est l'événement de l'apparition du coté pile dans le premier lancement et du coté face dans le deuxième.

### I.2.2: Diagramme de Venn

La représentation graphique est très utile pour l'illustration des différentes opérations des ensembles, tel que l'appartenance et le complément, cette représentation est appelée le diagramme de Venn [4] [9] [12], et on illustre quelques exemples sur la figure ci dessous.

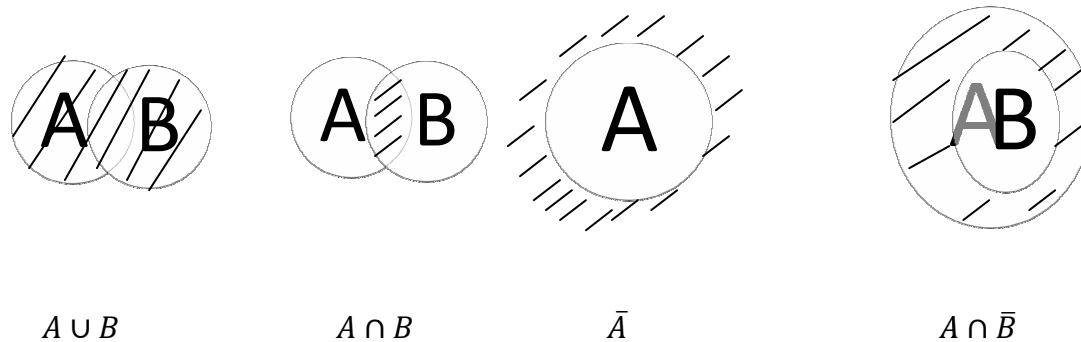


Figure 1.1:Diagramme de Venn

*Quelques exemples pour la représentation des opérations de l'algèbre des ensembles par le diagramme de Venn*

On dit que deux ensembles sont disjoints si leur intersection est l'ensemble vide, et l'exemple le plus évident est celui de l'intersection d'un ensemble et de son complément. En se référant au diagramme on peut déduire facilement que l'intersection de l'ensemble fondamental et d'un sous-ensemble A est l'ensemble A lui même, par conséquent l'union d'un sous ensemble et de son complément est l'ensemble universel (fondamental). Il est nécessaire de présenter quelques propriétés concernant l'algèbre des ensembles sous forme d'équations [4] [9] [12].

**Commutativité :**  $A \cup B = B \cup A$  (1.1)

$$A \cap B = B \cap A \quad (1.2)$$

**Associativité:**  $A \cup (B \cup C) = (A \cup B) \cup C$  (1.3)

$$A \cap (B \cap C) = (A \cap B) \cap C \quad (1.4)$$

**Distributivité:**  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$  (1.5)

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad (1.6)$$

**La loi de De Morgan:**  $\overline{A \cup B} = \bar{A} \cap \bar{B}$  (1.7)

$$\overline{A \cap B} = \bar{A} \cup \bar{B} \quad (1.8)$$

**I.2.3: La notion de la probabilité et ses axiomes**

Si on attribue à chaque événement défini dans l'ensemble universel un nombre réel on dit que on a mesuré la probabilité de cet événement. Considérons une expérience dans l'ensemble fondamental  $S$  et posons  $A$  un événement particulier, si on répète cette expérience avec les mêmes conditions  $n$  fois la probabilité d'occurrence de l'événement  $A$  est :

$$P(A) = \lim_{n \rightarrow \infty} \frac{n(A)}{n} \tag{1.9}$$

Cette relation définit la fréquence relative de l'événement  $A$  où  $n(A)$  est le nombre d'apparitions de l'événement  $A$  [4] [9] [12]. Après la définition de la fréquence d'occurrence, on définit la probabilité de l'événement  $A$  qui est un nombre réel positif assigné à  $A$  et qui doit satisfaire ces trois axiomes:

$$\text{Axiome1: } P(A) \geq 0 \tag{1.10}$$

$$\text{Axiome2: } P(S) = 1 \tag{1.11}$$

$$\text{Axiome3: } P(A \cup B) = P(A) + P(B) \text{ si } A \cap B = \emptyset \tag{1.12}$$

En utilisant ces axiomes on obtient les propriétés suivantes:

$$1. \ P(\bar{A}) = 1 - P(A) \tag{1.13}$$

$$2. \ P(\emptyset) = 0 \tag{1.14}$$

$$3. \ P(A) \leq P(B) \text{ si } A \subset B \tag{1.15}$$

$$4. \ P(A) \leq 1 \tag{1.16}$$

$$5. \ P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1.17}$$

6. si  $A_1, \dots, A_n$  sont des événements dans  $S$

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) + \sum_{i \neq j} P(A_i \cap A_j) + \sum_{i \neq j \neq k} P(A_i \cap A_j \cap A_k) - \dots (-1)^{n-1} P(A_1 \cap A_2 \dots \dots \dots \cap A_n) \tag{1.18}$$

**I.2.4: Probabilité conditionnelle et la règle de Bayes**

Soit un événement  $B$  dans un ensemble universel  $S$  tel que  $P(B) > 0$ , on appelle la probabilité pour que l'évènement  $A$  se produise, et l'évènement  $B$  s'étant produit auparavant la probabilité conditionnelle de  $A$  sachant que  $B$  se réalise, et on écrit  $P(A/B)$  [4] [9] [12].

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad P(B) > 0 \tag{1.19}$$



De la même façon:

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad P(A) > 0 \quad (1.20)$$

À partir de ces deux équations on peut trouver la probabilité conjointe  $P(A \cap B)$ :

$$P(A \cap B) = P(A/B) P(B) = P(B/A) P(A) \quad (1.21)$$

Et on définit aussi la règle de Bayes :

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)} \quad (1.22)$$

### **I.2.5: Événements indépendants**

On dit que les deux événements A et B sont statistiquement indépendants si et seulement si leur probabilité conjointe égale au produit de leurs probabilités.

$$P(A \cap B) = P(A) P(B) \quad (1.23)$$

Et immédiatement on peut déduire que si A et B sont indépendants alors:

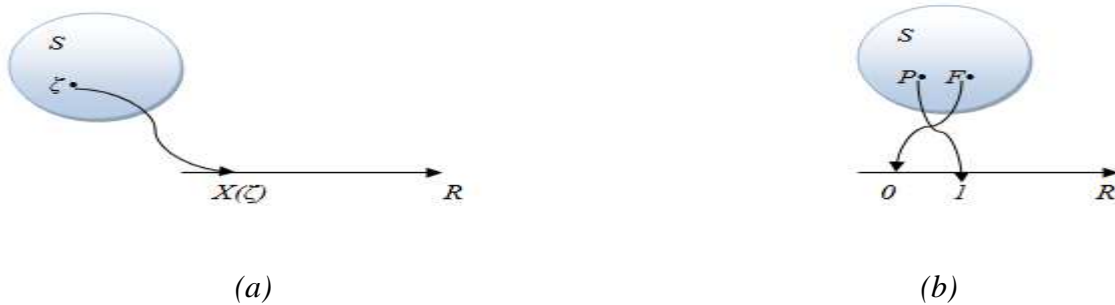
$$P(A/B) = P(A) \quad \text{et} \quad P(B/A) = P(B)$$

Il a été prouvé que si A et B sont indépendants alors A et  $\bar{B}$  sont aussi indépendants [4].

### **I.3: Variables aléatoires**

On a introduit le concept de la probabilité pour la description des résultats d'une expérience donnée et ces caractéristiques. Un événement nous permet de déterminer les propriétés d'une expérience seulement en se référant aux résultats de l'expérience eux mêmes [4] [9]. Pour cela on va introduire un nouveau concept pour représenter les événements d'une expérience sous forme numérique en projetant les sorties (événements) sous un autre espace qu'on l'appelle La variable aléatoire. Si on considère une expérience aléatoire dans l'ensemble fondamental S, la variable aléatoire  $X(\zeta)$  est une fonction réel d'une seule valeur et qui représente un nombre réel qui assigne chaque point de l'ensemble fondamental S. on appelle S le domaine de la variable aléatoire et toutes les valeurs que  $X(\zeta)$  puisse prendre sont l'image de ce domaine, par exemple le lancement d'une pièce de monnaie de l'expérience précédente, on peut définir la variable aléatoire comme suit:  $X(P)=1, X(F)=0$ .

Et la figure ci dessous nous montre la représentation:



**Figure I.2:** Représentation des variables aléatoires d'une seule dimension.

(a): Représentation de la variable aléatoire  $X$  en fonction de  $\zeta$ , (b): représentation des événements de l'exemple du lancement d'une pièce

### I.3.1: Événements définis par les variables aléatoires

Puisque la variable aléatoire est une fonction pour représenter les événements de l'ensemble fondamental  $S$ , il faut que cette fonction représente chaque point de l'ensemble  $S$  par une seule valeur et vice versa [4] [9], par conséquent chaque intervalle de la variable aléatoire représente un événement dans l'ensemble fondamental  $S$ .

Si  $X$  est une variable aléatoire et  $x$  un nombre réel fixe, on peut définir l'événement  $(X=x)$  par :

$$(X=x) = \{\zeta: X(\zeta)=x\} \quad (1.24)$$

De la même façon pour les nombres  $x$ ,  $x_1$  et  $x_2$ , nous définissons les événements suivants:

$$(X \leq x) = \{\zeta: X(\zeta) \leq x\} \quad (1.25)$$

$$(X > x) = \{\zeta: X(\zeta) > x\} \quad (1.26)$$

$$(x_1 \leq X \leq x_2) = \{\zeta: x_1 \leq X(\zeta) \leq x_2\} \quad (1.27)$$

Une condition est nécessaire pour que  $X$  soit une variable aléatoire, cette condition stipule que les probabilités de  $(X=\infty)$  et  $(X=-\infty)$  égales à zéro même si la variable  $X$  prend ces valeurs. Il est important aussi de mentionner que la variable aléatoire est divisée en trois types. Le premier type est celui d'une variable aléatoire discrète, l'ensemble fondamental de ce type peut être continu ou discret ou les deux à la fois. Le deuxième type c'est la variable aléatoire continue, et son ensemble fondamental ne peut être que du type continu.

Le troisième type est le type d'une variable aléatoire mixte qui est le type le moins important sauf pour quelques problèmes de signification pratique [4] [9].

**I.3.2: Fonction de distribution (répartition)**

Si  $P(X \leq x)$  est la probabilité de l'événement  $(X \leq x)$ , alors c'est une fonction de  $x$ . On appelle cette fonction, la fonction de distribution de probabilité ou cumulative:  $F_x(X) = P(X \leq x)$  [4] [9].

La fonction de distribution a les propriétés suivantes:

$$1. 0 \leq F_x(x) \leq 1 \tag{1.28}$$

$$2. F_x(x_1) \leq F_x(x_2) \quad \text{si } x_1 \leq x_2 \tag{1.29}$$

$$3. \lim_{x \rightarrow +\infty} F_x(x) = F_x(+\infty) = 1 \tag{1.30}$$

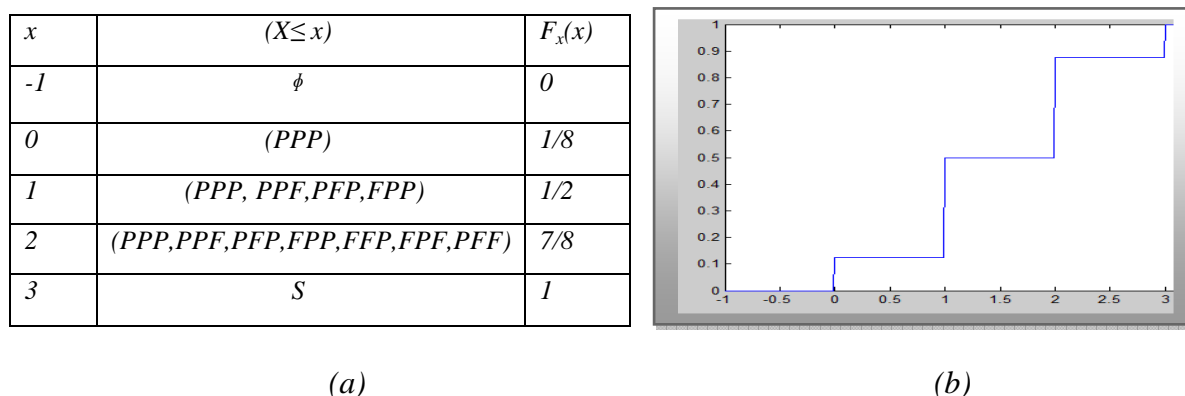
$$4. \lim_{x \rightarrow -\infty} F_x(x) = F_x(-\infty) = 0 \tag{1.31}$$

$$5. \lim_{x \rightarrow a^+} F_x(x) = F_x(a^+) = F_x(a) \quad a^+ = \lim_{0 < \varepsilon \rightarrow 0} a + \varepsilon \tag{1.32}$$

$$\lim_{x \rightarrow \infty} P(X \leq x) = P(X \leq \infty) = P(S) = 1$$

$$\lim_{x \rightarrow -\infty} P(X \leq x) = P(X \leq -\infty) = 0$$

Considérons le lancement d'une pièce trois fois respectivement, donc l'ensemble fondamental  $S = \{PPP, \dots, FFF\}$ , si  $X$  est la variable aléatoire qui représente le nombre de piles, et on veut trouver la probabilité  $P(X=2)$  alors,  $P(X=2) = P(A) = 3/8$ , où  $A = \{PPF, PFP, FPP\}$ . On peut aussi représenter cette variable aléatoire dans un tableau comme le montre la figure suivante:



**Figure I.3:** Fonction de répartition de l'expérience du lancement d'une pièce 3 fois.

(a): tableau pour la représentation de la fonction de répartition  $F_x(x)$ , (b): représentation graphique de la fonction de répartition  $F_x(x)$

On peut aussi déterminer la probabilité  $P(A)$  en utilisant ces deux équations:

$$P(a < X \leq b) = F_x(b) - F_x(a) \quad (1.33)$$

$$P(X > a) = 1 - F_x(a) \quad (1.34)$$

Alors pour l'exemple précédent la probabilité de l'événement  $P(X=2)$  est:

$$P(A) = P(1 < X \leq 2) = F_x(2) - F_x(1) = (7/8) - (1/2) = 3/8.$$

### I.3.3: Fonction de densité de probabilité (PDF)

Soit: 
$$f_x(x) = \frac{dF_x(x)}{dx} \quad (1.34)$$

La fonction  $f_x(x)$  est appelée fonction de densité de probabilité et elle a les propriétés suivantes:

$$1. f_x(x) \geq 0 \quad (1.35)$$

$$2. \int_{-\infty}^{+\infty} f_x(x) dx = 1 \quad (1.36)$$

$$3. f_x(x) \text{ est continue par segment} \quad (1.37)$$

$$4. P(a \leq x \leq b) = \int_a^b f_x(x) dx \quad (1.38)$$

Si  $X$  est une variable aléatoire continue, on peut trouver la fonction de distribution et même les probabilités en utilisant ces équations [4] [9][12].

$$F_x(x) = P(X \leq x) = \int_{-\infty}^x f_x(\xi) d\xi \quad (1.39)$$

$$P(a < X \leq b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X < b) \quad (1.40)$$

$$= \int_a^b f_x(x) dx = F_x(b) - F_x(a) \quad (1.41)$$

### I.3.4: Espérance et variance d'une variable aléatoire

L'espérance est le nom de la procédure par laquelle on mesure la moyenne d'une variable aléatoire et son symbole est  $E\{X\}$ . On l'appelle aussi la moyenne statistique de la variable aléatoire [4] [9] [8], et nous la calculons par ces équations:

$$\mu_x = E\{X\} = \begin{cases} \sum_k X_k P_x(x_k) & \text{cas discret} \\ \int_{-\infty}^{\infty} x f_x(x) dx & \text{cas continu} \end{cases} \quad (1.42)$$

La variance d'une variable aléatoire  $X$  notée  $VAR(X)$  est définie par:

$$\sigma^2 = VAR(X) = \begin{cases} \sum_k (X_k - \mu x)^2 P x(x_k) & \text{cas discret} \\ \int_{-\infty}^{\infty} (x - \mu x)^2 f x(x) dx & \text{cas continu} \end{cases} \quad (1.43)$$

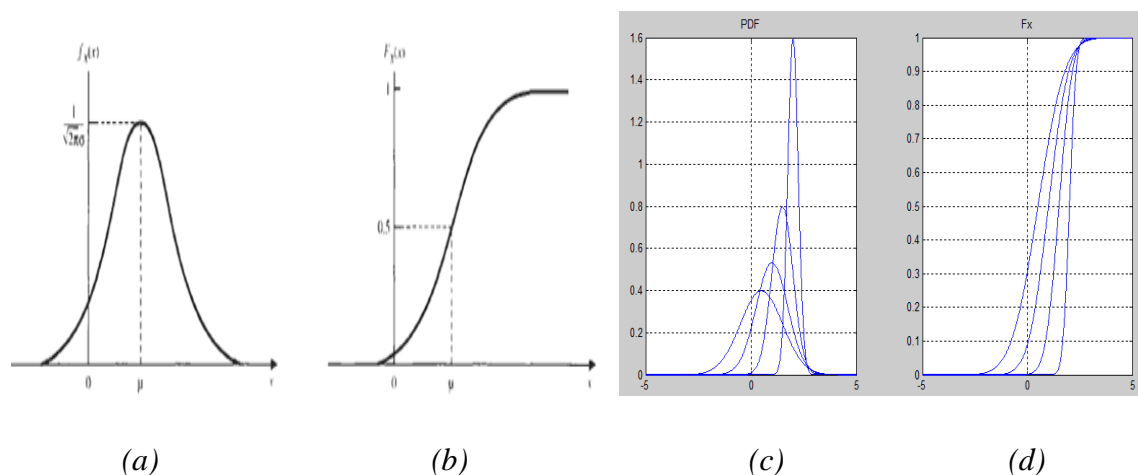
La racine carrée de la variance est ce qu'on appelle l'écart-type [4] [12], et c'est une mesure de la répartition des probabilités de la variable aléatoire par rapport à la moyenne, et pour mieux comprendre ces mesures on va utiliser une distribution spéciale, c'est la distribution gaussienne définie comme suit:

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\pi(x-\mu)^2}{2\sigma^2}} \quad (1.44)$$

$$F_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{\pi(\xi-\mu)^2}{2\sigma^2}} d\xi = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{\xi^2}{2}} d\xi \quad (1.45)$$

$$\mu_x = E\{X\} = \mu$$

$$\sigma_x^2 = VAR(X) = \sigma^2$$



**Figure I.4:**Distribution gaussienne.

(a) : fonction de la densité de probabilité gaussienne cas général, (b): fonction de répartition gaussienne cas général, (c): représentations des fonctions de densité de probabilité de différentes gaussiennes, (d) représentations des fonctions de répartition de probabilité de différentes gaussiennes.

### I.4: Variables aléatoires multiples

Dans beaucoup d'applications il est important de présenter les événements et les probabilités de l'ensemble fondamental par deux ou plusieurs variables aléatoires [4] [6], Dans ce qui suit, on va introduire la représentation d'un ensemble par deux variables aléatoires  $X$  et  $Y$ , et ce concept peut être généralisé pour plusieurs variables aléatoire  $X_1, X_2, \dots, X_n$ . Considérons un ensemble fondamental  $S$  d'une expérience aléatoire et posons  $X$  et  $Y$  deux variables aléatoires, dans ce cas  $(X, Y)$  est appelé vecteur aléatoire à deux dimensions [6]. Si chaque paire de  $(X, Y)$  est un nombre réel positif et qui est assigné par un seul élément de l'ensemble fondamental  $S$ , on dit qu'on a une fonction à deux variables qui représente l'image de chaque élément  $\zeta$  de l'ensemble  $S$  sur le plan  $(XOY)$ .

$$R_{xy} = \{(x, y); \xi \in S \text{ et } X(\xi) = x, Y(\xi) = y\} \quad (1.46)$$

Par exemple le lancement d'une pièce deux fois successives où  $X$  représente le nombre de piles et  $Y$  le nombre de faces, alors on a  $R_x = R_y = \{0, 1, 2\}$ . L'image de la variable aléatoire multiple sur le plan  $(XOY)$  est  $R_{xy} = \{(2,0), (1,1), (0,2)\}$ , et la figure ci dessous nous montre la représentation:

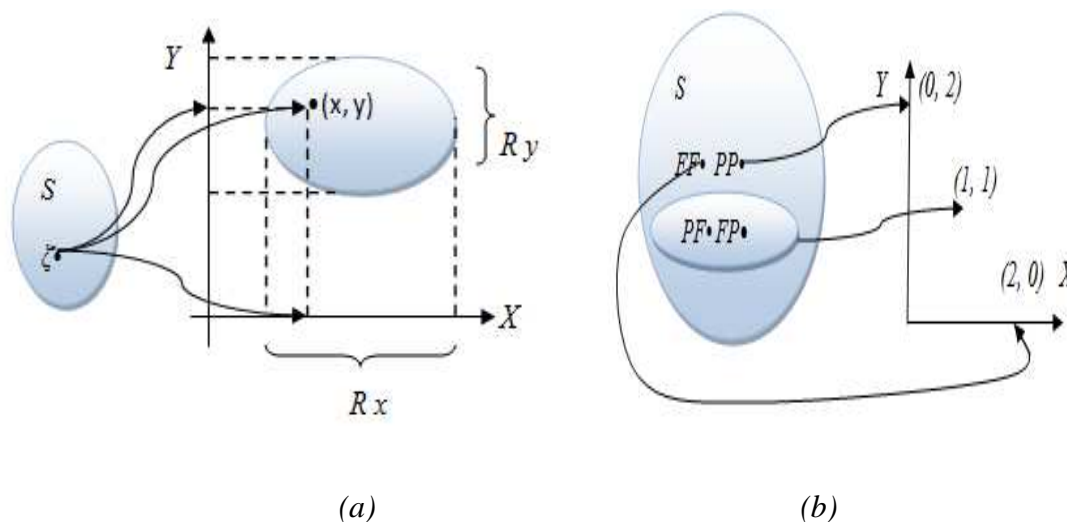


Figure I.5: Représentation des variables aléatoires multiples.

(a): représentation des variables aléatoire  $X$  et  $Y$  en fonction de  $\zeta$ , (b): représentation des événements de l'exemple du lancement d'une pièce deux fois par deux variables aléatoires.

**I.4.1: Fonctions de distribution (répartition) conjointe**

De la même façon pour la définition de la fonction de répartition d'une seule variable aléatoire, on définit la fonction de répartition conjointe des deux variables aléatoires X et Y comme suit:  $F_{XY}(x, y) = P(X \leq x, Y \leq y)$  (1.47)

Et l'événement  $(X \leq x, Y \leq y)$  est équivalent à l'intersection des deux événements A et B de l'ensemble fondamental S [4].

$$P(A) = F_x(x) \quad \text{et} \quad P(B) = F_y(y)$$

$$F_{XY}(x, y) = P(A \cap B) \tag{1.48}$$

La fonction de répartition conjointe a des propriétés analogues à celles d'une fonction de distribution d'une seule variable aléatoire:

$$1. \quad 0 \leq F_{XY}(x, y) \leq 1 \tag{1.48}$$

$$2. \quad \text{si } x_1 \leq x_2 \text{ et } y_1 \leq y_2$$

$$F_{XY}(x_1, y_1) \leq F_{XY}(x_2, y_1) \leq F_{XY}(x_2, y_2) \tag{1.49}$$

$$F_{XY}(x_1, y_1) \leq F_{XY}(x_1, y_2) \leq F_{XY}(x_2, y_2) \tag{1.50}$$

$$3. \quad \lim_{\substack{x \rightarrow \infty \\ y \rightarrow \infty}} F_{XY}(x, y) = F_{XY}(\infty, \infty) = 1 \tag{1.51}$$

$$4. \quad \lim_{x \rightarrow -\infty} F_{XY}(x, y) = F_{XY}(-\infty, y) = 0 \tag{1.52}$$

$$\lim_{y \rightarrow -\infty} F_{XY}(x, y) = F_{XY}(x, -\infty) = 0 \tag{1.53}$$

$$5. \quad \lim_{x \rightarrow -a^+} F_{XY}(x, y) = F_{XY}(a^+, y) = F_{XY}(a, y) \tag{1.54}$$

$$\lim_{y \rightarrow -b^+} F_{XY}(x, y) = F_{XY}(x, b^+) = F_{XY}(x, b) \tag{1.56}$$

$$6. \quad P(x_1 < X \leq x_2, Y \leq y) = F_{XY}(x_2, y) - F_{XY}(x_1, y) \tag{1.57}$$

$$P(X \leq x, y_1 < Y \leq y_2) = F_{XY}(x, y_2) - F_{XY}(x, y_1) \tag{1.58}$$

Et comme les conditions  $(y \leq \infty)$  et  $(x \leq \infty)$  sont toujours satisfaites on définit aussi les fonctions de répartition marginales de  $F_{XY}(x, y)$  et qui sont définies comme suit:

$$F_x(x) = \lim_{y \rightarrow \infty} F_{XY}(x, y) = F_{XY}(x, \infty) \tag{1.59}$$

$$F_y(y) = \lim_{x \rightarrow \infty} F_{XY}(x, y) = F_{XY}(\infty, y) \tag{1.60}$$

**I.4.2: Fonctions de densité de probabilité conjointe (la PDF conjointe)**

Soient (X, Y) deux variables aléatoires définies par la fonction de répartition  $F_{XY}(x, y)$ , et soit:

$$f_{XY}(x, y) = \frac{\partial^2 F_{xy}(x,y)}{\partial x \partial y} \tag{1.61}$$

La fonction  $f_{xy}(x, y)$  est appelée la fonction de densité de probabilité conjointe des deux variables aléatoires X et Y [4] [9], et qui a les propriétés suivantes:

$$1. f_{XY}(x, y) \geq 0 \tag{1.62}$$

$$2. \iint_{-\infty}^{+\infty} f_{XY}(x, y) dx dy = 1 \tag{1.63}$$

$$3. f_{XY}(x, y) \text{ est continue} \tag{1.64}$$

$$4. P [(X, Y) \in A] = \iint_{R_A} f_{XY}(x, y) dx dy \tag{1.65}$$

$$5. P (a < X \leq b, c < Y \leq d) = \int_a^b \int_c^d f_{XY}(x, y) dx dy \tag{1.66}$$

De la même manière on définit les fonctions de densité de probabilité marginale  $f_x(x)$  et  $f_y(y)$  comme suit:

$$f_x(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy \tag{1.67}$$

$$f_y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx \tag{1.68}$$

**I.4.3: Distribution conditionnelle**

Si X et Y sont deux variables aléatoires avec la fonction de densité de probabilité conjointe  $f_{XY}(x,y)$ , donc la fonction de densité de probabilité de Y à condition que X=x, est définie:

$$f_{Y/X}(y/x) = \frac{f_{XY}(x,y)}{f_X(x)} \quad f_X(x) > 0 \tag{1.69}$$

De la même façon, la fonction de densité de probabilité de X à condition que Y=y, est définie:

$$f_{X/Y}(x/y) = \frac{f_{XY}(x,y)}{f_Y(y)} \quad f_Y(y) > 0 \tag{1.70}$$



**I.4.4: Variables aléatoires indépendantes**

On dit que les deux variables aléatoires X et Y sont indépendantes si:

$$F_{XY}(x, y) = P(A \cap B) = P(A)P(B) = F_x(x)F_y(y) \tag{1.71}$$

$$\frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} = \frac{\partial F_x(x)}{\partial x} \frac{\partial F_y(y)}{\partial y} \tag{1.72}$$

$$f_{XY}(x, y) = f_x(x)f_y(y) \tag{1.73}$$

$$f_{Y/X}(y/x) = f_y(y) \quad \text{et} \quad f_{X/Y}(x/y) = f_x(x)$$

Prenons comme exemple les deux fonctions de densité de probabilités conjointes suivantes:

$$f_{1XY}(x, y) = \begin{cases} 0.125(x + y) & \text{si } (0 < x < 2) \text{ et } (0 < y < 2) \\ 0 & \text{ailleurs} \end{cases}$$

$$f_{2XY}(x, y) = \begin{cases} 4xy & \text{si } (0 < x < 1) \text{ et } (0 < y < 1) \\ 0 & \text{ailleurs} \end{cases}$$

\* Pour la première fonction les PDF marginales sont,  $f_{1X}(x) = 0.25(x+1)$  et  $f_{1Y}(y) = 0.25(y+1)$

il est clair que ces deux variables sont dépendantes car  $f_{1X}(x) \times f_{1Y}(y) \neq f_{1XY}(x, y)$ .

\* Pour la deuxième fonction les PDF marginales sont  $f_{2X}(x) = 2x$  et  $f_{2Y}(y) = 2y$ , ces deux

variables sont indépendantes, puisque  $f_{2X}(x) \times f_{2Y}(y) = f_{2XY}(x, y)$ .

**I.4.5: Covariance et le coefficient de corrélation**

Avant d'introduire la notion de la covariance des variables aléatoires X et Y on doit définir le (k, n) ième moment qui est défini par :

$$m_{kn} = E\{X^k Y^n\} = \begin{cases} \sum_{y_i} \sum_{x_i} x_i^k y_i^n P_{xy}(x_i, y_i) & \text{cas discret} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^n f_{xy}(x, y) dx dy & \text{cas continu} \end{cases} \tag{1.74}$$

Si  $n = 0$ , ce moment est l'espérance mathématique de la variable aléatoire X,  $m_{10} = E\{X\}$ , de même si  $k = 0$ , alors  $m_{01} = E\{Y\}$ . De la même façon pour la variance d'une seule variable aléatoire on définit la variance des variables aléatoires X et Y à partir de la fonction de densité de probabilité conjointe en appliquant l'équation (1.74), donc la variance de X et Y est le (1,1) ième moment de (X,Y) =  $m_{11} = E\{XY\}$  [4] [9].

La covariance de X et Y est définie comme suit :

$$Cov(X, Y) = E\{XY\} - E\{X\}E\{Y\} \quad (1.75)$$

Si  $Cov(X, Y) = 0$ , on dit que les deux variables aléatoires sont non corrélées. Il est important de mentionner que si deux variables aléatoires sont indépendantes alors elles sont aussi non corrélées mais pas le contraire.

Le facteur de corrélation des variables aléatoires X et Y est une mesure de l'indépendance entre X et Y et qui est défini :

$$\rho(X, Y) = \rho_{xy} = \frac{Cov(X, Y)}{\sigma_x \sigma_y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (1.76)$$

Où  $|\rho_{xy}| \leq 1$

### **I.5: Fonctions caractéristiques et Cumulants**

La fonction de densité de probabilité joue un rôle crucial pour la caractérisation d'une variable aléatoire, et aussi pour la détermination des informations stochastiques concernant la variable aléatoire, mais dans certains cas avoir les informations à partir d'une PDF peut être une tâche difficile, donc il faut avoir des représentations alternatives pour la représentation, comme la fonction caractéristique, que cette dernière projette la PDF sur un autre plan et qui nous facilite la tâche pour extraire les informations nécessaires d'une variable aléatoire [8].

#### **I.5.1: Fonction d'une variable aléatoire**

Considérons la variable aléatoire X qui est caractérisée par sa fonction de distribution  $F_x(x)$  et sa PDF  $f_x(x)$ . La transformation ou la fonction  $Y = g(X)$  est aussi une variable aléatoire générée par la variable aléatoire X, où les probabilités de la variable aléatoire Y peuvent être déterminées par cette équation [4] [9].

$$(Y \leq y) = [g(X) \leq y] = (X \in D_y) \quad (1.77)$$

De même pour la fonction de répartition et la fonction de densité de probabilité, on peut les obtenir à partir de la fonction de répartition de la variable aléatoire X en utilisant ces équations:

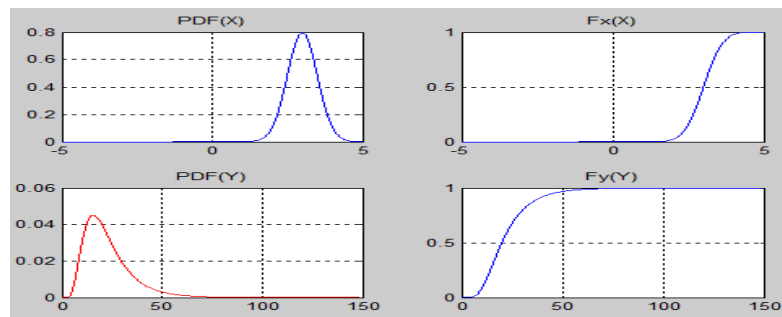
$$F_y(y) = P(Y \leq y) = P[g(X) \leq y] = P(X \in D_y) \quad (1.78)$$

$$F_y(y) = \int_{D_y} f_x(x) dx \tag{1.79}$$

Si par exemple la variable aléatoire  $X$  qui suit la loi gaussienne  $X=N(\mu,\sigma^2)$ , la PDF de la variable aléatoire  $Y$  générée par  $X$  par la fonction  $Y=e^X$  est une variable aléatoire, et sa PDF peut être obtenue après intégration ensuite après dérivation de la PDF de  $X$ . Et après le calcul analytique la PDF de  $Y$  a l'expression suivante.

$$f_y(y) = \frac{f_x(\ln y)}{y} = \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}} \quad 0 < y < \infty$$

Et la figure ci dessous, représente les deux variables:



**Figure I.6:**Exemple d'une variable aléatoire générée à partir d'une seule variable aléatoire.

*Représentation graphique des fonctions de distributions de la variable aléatoire Y générée par la variable aléatoire X gaussienne (3,0.25) par la fonction  $Y=e^X$*

**I.5.2:Fonction de deux variables aléatoires**

De la même façon pour générer une variable aléatoire à partir d'une seule variable aléatoire, on peut aussi générer une variable à partir de deux variables aléatoires.

Soit  $Z=g(X, Y)$  une fonction de deux variables aléatoires  $X$  et  $Y$ .  $Z$  définit la nouvelle variable aléatoire suivante:

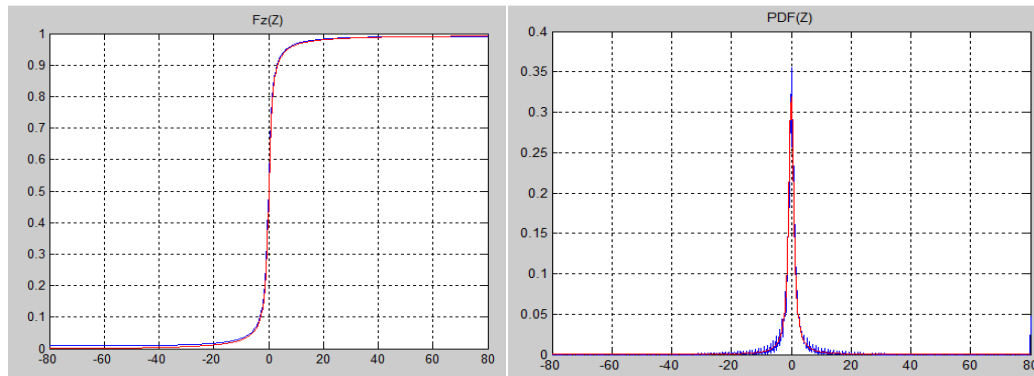
$$(Z \leq z) = [g(X, Y) \leq z] = \{(X, Y) \in D_z\} \tag{1.80}$$

La fonction de répartition et la PDF peuvent être obtenues à partir de de la PDF conjointe en utilisant ces deux equations:

$$F_z(Z) = P(Z \leq z) = P[g(X, Y) \leq z] = P\{(X, Y) \in D_z\} \tag{1.81}$$

$$F_z(Z) = \int \int_{D_z} f_{XY}(x, y) dx dy \tag{1.82}$$

Si par exemple X et Y sont des variables aléatoires gaussiennes et indépendantes, et Z une variables aléatoire définie par:  $Z=X/Y$  alors la PDF de Z est obtenue en appliquant numériquement la relation (1.82), et les distributions de cette nouvelle variables sont représentées sur la figure suivante:



**Figure I.7:** Exemple d'une variable aléatoire générée à partir de deux variables aléatoires. Représentation graphique des fonctions de distributions de la variable aléatoire Z générée par les variables aléatoires gaussiennes X (0, 1) et Y (0, 1) par la fonction  $Z=X/Y$

Les relations (1.81) et (1.82) peuvent être généralisées pour plusieurs variables aléatoires, par exemple la détermination de la fonction de répartitions conjointes de deux variables Z et W à partir de la fonction de densité de probabilité conjointe des variables aléatoires X et Y est obtenue en appliquant ces équations [4][9].

Si  $Z=g(X, Y)$  et  $W=h(X, Y)$

$$F_{ZW}(z,w) = P(Z \leq z, W \leq w) = P\{g(X,Y) \leq z, h(X, Y) \leq w\} = P\{(X,Y) \in D_{ZW}\} \quad (1.83)$$

$$F_{ZW}(z,w) = \int \int_{D_{ZW}} f_{xy}(x, y) dx dy \quad (1.84)$$

### I.5.3:Fonction caractéristique d'une seule variable aléatoire

Avant d'introduire la notion de la fonction caractéristique d'une variable aléatoire, il est important de définir l'espérance mathématique d'une variable aléatoire générée par une autre variable aléatoire.

Comme dans le cas pour mesurer la moyenne ou l'espérance mathématique de X on peut mesurer l'espérance de la variable  $Y=g(X)$  à partir de ces équations:

$$E\{Y\} = E\{g(X)\} = \begin{cases} \sum_i g(x_i)P_x(x_i) & \text{cas discret} \\ \int_{-\infty}^{\infty} g(x)f_x(x)dx & \text{cas continu} \end{cases} \quad (1.85)$$

Et la mesure peut être généralisée pour plusieurs variables aléatoires  $X_1, \dots, X_n$ .

Soit  $Y=g(X_1, \dots, X_n)$ , alors:

$$E\{Y\} = E[g(X)] = \begin{cases} \sum_{x_1} \dots \sum_{x_n} g(x_1 \dots x_n)P_{x_1, \dots, x_n}(x_1 \dots x_n) & \text{cas discret} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1 \dots x_n)f_{x_1, \dots, x_n}(x_1 \dots x_n) & \text{cas continu} \end{cases} \quad (1.86)$$

La fonction caractéristique  $\psi_x(\omega)$  de la variable aléatoire X est une transformation de la fonction de densité de probabilité de X en une autre fonction dans le plan  $\omega$  [4][9], où  $\omega$  est une variable réelle, et le calcul de cette fonction se fait de la même façon pour la mesure de l'espérance mathématique, sauf qu'on doit mesurer l'espérance de  $e^{j\omega X}$  pour chaque valeur de  $\omega$ [8].

$$\psi_X(\omega) = E\{e^{j\omega x}\} = \begin{cases} \sum_i e^{j\omega x_i}P_X(x_i) & \text{cas discret} \\ \int_{-\infty}^{\infty} e^{j\omega x}f_X(x)dx & \text{cas continu} \end{cases} \quad (1.87)$$

Pour le cas inverse, la détermination de la PDF de la variable aléatoire X est obtenue à partir de la fonction caractéristique en appliquant cette équation

$$f_x(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi_X(\omega)e^{-\omega x j} d\omega \quad (1.88)$$

On remarque que ces deux dernières équations sont de la même forme pour la détermination de la transformée de Fourier [4], donc on peut utiliser le tableau de la TF pour la détermination de la fonction caractéristique d'une fonction de densité de probabilité. Si on prend comme exemple la fonction de densité de probabilité de Cauchy avec le paramètre (a) et qui a l'expression suivante:

$$f_x(x) = \frac{1}{\pi(a^2+x^2)} \quad -\infty < x < +\infty$$

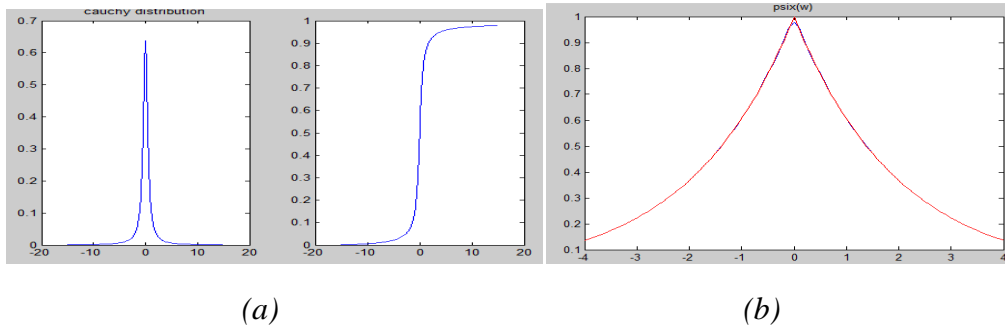
En utilisant la transformée de Fourier on peut avoir la fonction caractéristique  $\psi_x(\omega)$  de cette façon:

$$e^{-a/|x|} \leftrightarrow \frac{2a}{(a^2+\omega^2)}$$

$$\frac{2a}{(a^2+x^2)} \leftrightarrow 2\pi e^{-a/|\omega|}$$

$$\frac{2a}{\pi(a^2+x^2)} \leftrightarrow e^{-a/|\omega|}, \quad \text{donc: } \psi_x(\omega) = e^{-a/|\omega|}$$

Et la représentation des distributions de cette variable et sa fonction caractéristique son sur la figure ci dessous:



**Figure I.8:**Exemple d'une fonction caractéristique d'une seule variable aléatoire

(a): représentation des fonctions de distribution de la variable de Cauchyle avec le paramètre a (dans cet exemple a =0.5). (b): la fonction caractéristique de cette distribution

**I.5.4: Fonctions caractéristique conjointe**

On définit la fonction caractéristique conjointes des deux variables aléatoire X et Y par:

$$\psi_{xy}(\omega_1, \omega_2) = E\{ e^{j(\omega_1x_i + \omega_2y_k)} \} \tag{1.89}$$

$$\text{Alors: } \psi_{xy}(\omega_1, \omega_2) = \begin{cases} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j(\omega_1x + \omega_2y)} f_{xy}(x, y) dx dy & \text{cas continu} \\ \sum_i \sum_k e^{j(\omega_1x_i + \omega_2y_k)} P_{xy}(x_i, y_k) & \text{cas discret} \end{cases} \tag{1.90}$$

cette fonction a la même expression de la transformée de fourier à deux dimension [4], et la fonction inverse de la fonction caractéristique est définie par:

$$f_{xy}(x, y) = \frac{1}{2\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_{xy}(\omega_1, \omega_2) e^{-j(\omega_1x + \omega_2y)} d\omega_1 d\omega_2 \tag{1.91}$$

La fonction caractéristique conjointe a ces deux propriétés :

$$\psi_x(\omega) = \psi_{xy}(\omega, 0) \quad \text{et} \quad \psi_y(\omega) = \psi_{xy}(0, \omega)$$

Si les variables aléatoires sont indépendantes alors:

$$\psi_{x1...xn}(\omega_1, \dots, \omega_n) = \psi_{x1}(\omega_1) \times \dots \times \psi_{xn}(\omega_n)$$

**I.5.5 : Moments et cumulants**

On a introduit au paragraphe (I.4.5) que les moments d'une variable aléatoire peuvent être calculés à partir de la fonction de densité de probabilité conjointe, dans cette partie on introduit le concept des cumulants qui sont des mesures pour la caractérisation des propriétés stochastiques des variables aléatoires [8] [18], et pour cela on va commencer par définir le calcul des moments à partir de la première fonction caractéristique.

**I.5.5.1: Moment et cumulants d'une seule variable aléatoire**

On peut mesurer les moments d'ordre n d'une variable aléatoire X à partir de la première fonction caractéristique en utilisant cette expression:

$$m_k = (-j)^k \frac{\partial^k \psi_x(\omega)}{\partial \omega^k} \Big|_{\omega=0} \tag{1.92}$$

Où  $m_k$  est le k nième moment. Une autre mesure statistique très importante est appelée cumulants de la variable aléatoire X, et qui nous révèlent certains aspects de la fonction de densité de probabilité de la variable aléatoire. Les cumulants sont générés par la deuxième fonction caractéristique qui est définie par cette expression [4] [8] [18].

$$Y_x(\omega) = \ln [\psi_x(\omega)] \tag{1.93}$$

Et les cumulants sont obtenus comme suit:

$$C_k = (-j)^k \frac{\partial^k Y_x(\omega)}{\partial \omega^k} \Big|_{\omega=0} \tag{1.94}$$

Où k est l'ordre du cumulant.

Les cumulants ont les propriétés suivantes:

$$* C_1 (X+\alpha) = C_1 (X) + \alpha \tag{1.95}$$

$$* C_k (X+\alpha) = C_k (X) \tag{1.96}$$

$$* C_k (\alpha X) = \alpha^k C_k (X) \tag{1.97}$$

Si X et Y sont deux variables aléatoires statistiquement indépendantes alors

$$* C_k (X+Y) = C_k (X) + C_k (Y) \tag{1.98}$$

**I.5.5.2: Relation entre moments et cumulants d'une seule variable aléatoire**

On peut mesurer les moments et les cumulants en utilisant la règle récursive suivante:

$$C_k = m_k - \sum_{i=1}^{k-1} \binom{k-1}{i-1} C_i m_{k-i} \tag{1.99}$$

À partir de cette dernière expression les moments peuvent être calculés en utilisant les cumulants d'ordres inférieurs [8] [18]. Et voici les quatre premiers moments comme exemple:

$$m_1 = C_1$$

$$m_2 = C_2 + C_1^2$$

$$m_3 = C_3 + 3C_2C_1 + C_1^3$$

$$m_4 = C_4 + 4C_3C_1 + 3C_2^2 + 6C_2C_1^2 + C_1^4$$

**I.5.5.3: Moment et cumulants de plusieurs variables aléatoires**

Le calcul des moments des variables aléatoires à plusieurs dimensions sont obtenus en dérivant la première fonction caractéristique conjointe partiellement, et de la même manière nous calculons les cumulants par la dérivation partielle de la deuxième fonction caractéristique conjointe [8] [18], en utilisant cette relation:

$$C(x_1^{k_1}, \dots, x_p^{k_p}) = (-j)^r \frac{\partial^r \gamma(\omega_1, \dots, \omega_p)}{\partial \omega_1^{k_1} \dots \partial \omega_p^{k_p}} \Big|_{\omega_1 = \dots = \omega_p = 0} \tag{1.100}$$

**I.5.5.4: Relation entre moments et cumulants d'un vecteur aléatoire**

Les moments et les cumulants d'un vecteur aléatoire de n dimensions sont liés par cette relation:

$$C(x_1, x_2, x_3, x_4) = \sum (-1)^{q-1} (q-1)! E \left[ \prod_{i \in v_1} x_i \right] \left[ \prod_{j \in v_2} x_j \right] \dots \left[ \prod_{k \in v_p} x_k \right] \tag{1.101}$$

où  $v_i$  est la répartition  $i$  des éléments du vecteur aléatoire

Par exemple les moments conjoints (croisés) et les cumulants croisés des variables aléatoires  $x_1, x_2, x_3, x_4$  sont liés par ces relations:



$$C(x_1) = E\{x_1\} \tag{1.102}$$

$$C(x_1, x_2) = E\{x_1 x_2\} - E\{x_1\} E\{x_2\} \tag{1.103}$$

$$C(x_1, x_2, x_3) = E\{x_1 x_2 x_3\} - E\{x_1\} E\{x_2 x_3\} - E\{x_2\} E\{x_1 x_3\} - E\{x_3\} E\{x_1 x_2\} + 2 E\{x_1\} E\{x_2\} E\{x_3\} \tag{1.104}$$

$$C(x_1, x_2, x_3, x_4) = E\{x_1 x_2 x_3 x_4\} - E\{x_1\} E\{x_2 x_3 x_4\} - E\{x_2\} E\{x_1 x_3 x_4\} - E\{x_3\} E\{x_1 x_2 x_4\} - E\{x_4\} E\{x_1 x_3 x_2\} - E\{x_1 x_2\} E\{x_3 x_4\} - E\{x_1 x_3\} E\{x_2 x_4\} - E\{x_1 x_4\} E\{x_2 x_3\} + 2 E\{x_1 x_2\} E\{x_3\} E\{x_4\} + 2 E\{x_1 x_3\} E\{x_2\} E\{x_4\} + 2 E\{x_1 x_4\} E\{x_2\} E\{x_3\} + 2 E\{x_2 x_3\} E\{x_1\} E\{x_4\} + 2 E\{x_2 x_4\} E\{x_1\} E\{x_3\} + 2 E\{x_3 x_4\} E\{x_1\} E\{x_2\} - 6 E\{x_1\} E\{x_2\} E\{x_3\} E\{x_4\} \tag{1.105}$$

Il faut mentionner que, si les variables aléatoires sont statistiquement indépendantes, alors tous les cumulants croisés sont nuls [8].

### I.6: Processus stochastique (aléatoire)

Un processus aléatoire est une famille de variables aléatoires avec le paramètre (t), {X(t), t ∈ T} donc c'est une fonction de deux paramètres, ζ pour l'ensemble fondamental S des variables aléatoires, et le second (t) pour l'ensemble T. Si (t) est fixe alors c'est une simple variable aléatoire, et si (t) varie on appelle cette fonction la réalisation du processus aléatoire des éléments ζ<sub>i</sub> de l'ensemble (A) à partir de l'ensemble fondamental S, et la valeur de cette réalisation pour chaque index (t) est appelée état. si le paramètre (t) est continu on dit que le processus est un processus stochastique continu, et si (t) est discret on dit que le processus est un processus stochastique discret [4] [8].

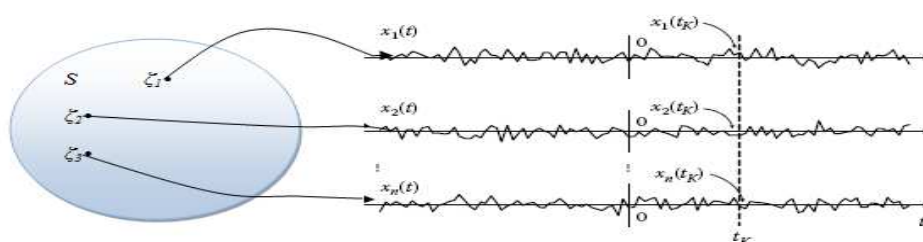


Figure I.9: Schéma simplifié pour la représentation d'un processus aléatoire

**I.6.1:Caractérisation d'un processus stochastique (aléatoire)**

**I.6.1.1:Description probabiliste**

Considérons un processus stochastique pour  $t=t_1$ , alors  $X(t_1)=X_1$  est une variable aléatoire et sa fonction de répartition est donnée comme suit:

$$F_x(x_1;t_1)=P\{X(t_1) \leq x_1\} \tag{1.106}$$

cette dernière fonction est appelée la fonction de répartition de  $X(t)$  du premier ordre, et pour deux paramètres on aura deux variables aléatoires  $X(t_1)=X_1$  et  $X(t_2)=X_2$ , et on dit que c'est une fonction de répartition du second ordre [4] [8].

$$F_x(x_1, x_2; t_1, t_2)=P\{X(t_1) \leq x_1, X(t_2) \leq x_2\} \tag{1.107}$$

Le concept s'étend pour tous les paramètres:

$$F_x(x_1, \dots, x_n; t_1, \dots, t_n)=P\{X(t_1) \leq x_1, \dots, X(t_n) \leq x_n\} \tag{1.108}$$

**I.6.1.2:Espérance, corrélation et fonction de covariance**

Comme dans le cas de la variable aléatoire on peut mesurer l'espérance mathématique d'un processus aléatoire par la relation suivante:

$$\mu_x(t)=E\{X(t)\} \tag{1.109}$$

On définit aussi l'auto corrélation du processus stochastique, qui est l'esperence mathématique du processus pour chaque couple des variables aléatoires  $X(t_i)$  en appliquant cette relation:

$$R_x(t,s)=E\{X(t)X(s)\} \tag{1.110}$$

Et on a aussi les propriétés suivantes:

$$R_x(t,s)= R_x(s,t) \tag{1.111}$$

$$R_x(t,t)=E\{X(t)^2\} \tag{1.112}$$

l'auto covariance d'un processus stochastique est définie par :

$$K_x(t,s)=Cov[X(t),X(s)]=E\{X(t)-\mu_x(t)\}[E\{X(s)-\mu_x(s)\}] \tag{1.113}$$

$$=R_x(t,s)-\mu_x(t) \mu_x(s) \tag{1.114}$$

**I.6.2: Processus stochastique stationnaire**

On dit que le processus stochastique X(t) est stationnaire [4] [8], si pour tous les (n) et pour tous les instants (t):

$$F_x(x_1, \dots, x_n; t_1, \dots, t_n) = F_x(x_1, \dots, x_n; t_1 + \tau, \dots, t_n + \tau) \tag{1.115}$$

pour la distribution du premier ordre on a:

$$F_x(x; t) = F_x(x; t + \tau) = F_x(x) \tag{1.116}$$

$$f_x(x; t) = f_x(x) \tag{1.117}$$

$$\mu_x(t) = E\{X(t)\} = \mu \tag{1.118}$$

$$Var(X(t)) = \sigma^2 \tag{1.119}$$

**I.6.3: Processus stochastiques indépendants**

Si pour tous les (n) d'un processus stochastique X(t) la fonction de répartition

$$F_x(x_1, \dots, x_n; t_1, \dots, t_n) = \prod_{i=1}^n F_x(x_i; t_i) \tag{1.120}$$

X(t) est un processus stochastique indépendant, et la fonction de répartition du premier ordre suffit pour le caractériser [4] [9].

**I.6.4: Histogramme**

L'histogramme est une représentation graphique du nombre d'apparition de chaque état d'une réalisation d'une variable aléatoire, si par exemple un processus stochastique est strictement stationnaire et on a une réalisation prise dans un intervalle de temps suffisamment grand, l'histogramme tend vers la fonction de densité de probabilité de la variable aléatoire X prise à partir d'un ensemble fondamental S où tous les états sont les éléments d'un événement A [7]. Et pour clarifier la notion supposons qu'on a dans un intervalle T une réalisation d'un processus stochastique, les états sont  $[x_1, x_2, \dots, x_M]$ , ces états appartiennent à un intervalle que ce dernier est le domaine des images d'une variable aléatoire, si on divise cet intervalle pour obtenir des intervalles élémentaires  $\Delta x$  en assignant pour chaque intervalle le nombre d'éléments ( $m_i$ ) qu'il contient, alors l'intervalle total contient la somme de tous les  $m_i$ , et la somme de toutes les régions en fonction des intervalles élémentaires est:

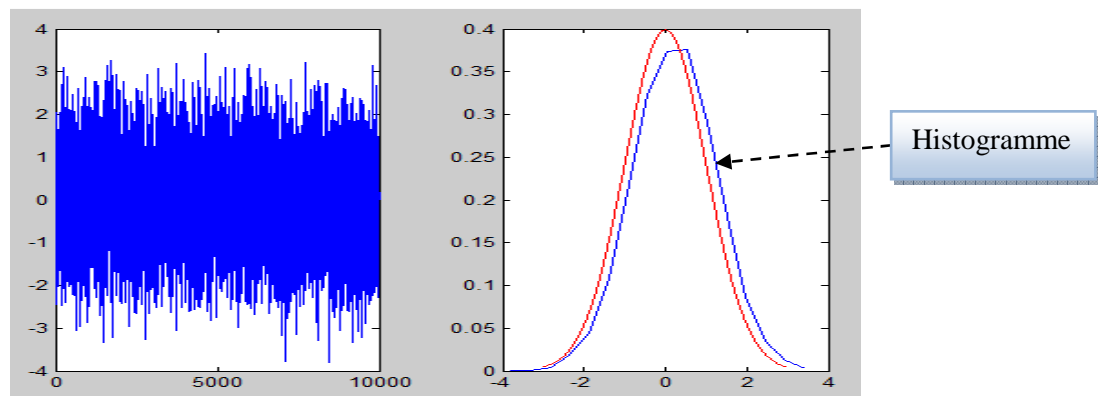
$$m_1 \Delta x + m_2 \Delta x + \dots + m_M \Delta x$$

$$\begin{aligned}
 &= (m_1+m_2+\dots+m_M) \Delta x \\
 &= (\sum_{i=1}^M m_i) \Delta x
 \end{aligned}
 \tag{1.121}$$

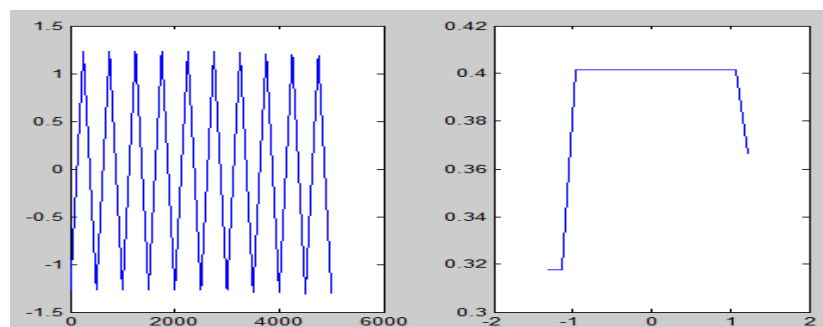
La probabilité qu'un état se présente dans le nième intervalle est:

$$P(x_i < x \leq (x_i + \Delta x)) = \frac{m_i \Delta x}{\sum_j m_j \Delta x} = \frac{m_i}{N}
 \tag{1.122}$$

Si on applique cette dernière relation pour tous les intervalles élémentaires on dit qu'on a construit l'histogramme du processus stochastique. À partir de l'histogramme on peut déterminer d'une façon approximative les caractéristiques d'un processus aléatoire et appliquer toutes les formules concernant la variable aléatoires comme le montre la figure suivante:



(a)



(b)

**Figure I.10:**Exemple de la construction d'un histogramme.

(a): à gauche un processus stochastique qui suit la loi gaussienne normale à droite en bleu l'histogramme du processus, en rouge la densité de probabilité gaussienne normale. (b):à gauche signal triangulaire, à droite l'histogramme du signal.

## I.7: Théorème de la limite centrale et le Kurtosis

### I.7.1: Kurtosis

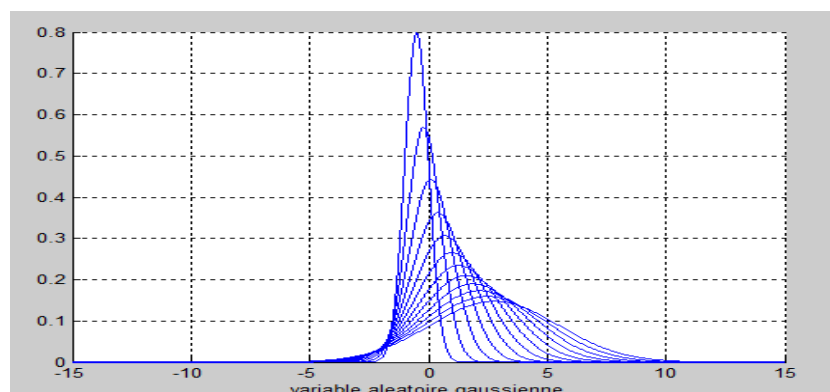
Le Kurtosis est une mesure pour les variables aléatoires [6], on le calcul à partir des moments du second ordre et d'ordre quatre par cette expression:

$$K = \frac{E(x^4)}{E(x^2)^2} \quad (1.123)$$

Le Kurtosis est la mesure du degré d'aplatissement d'une distribution [6] [7]. Si les densités de probabilités tendent vers 0 à l'infini moins vite que la gaussienne alors le kurtosis est positive, et si ces distribution tendent vers 0 plus rapidement que la gaussienne alors le Kurtosis est négative [6]. L'expression du Kurtosis pour une variable aléatoire standardisée est la suivante:

$$K = \frac{E\{(\bar{x}-x)^4\}}{E\{(\bar{x}-x)^2\}^2} - 3 \quad (1.124)$$

la constante 3 assure que la distribution gaussienne a le Kurtosis nul.



kurtosis =

0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0002 0.0012

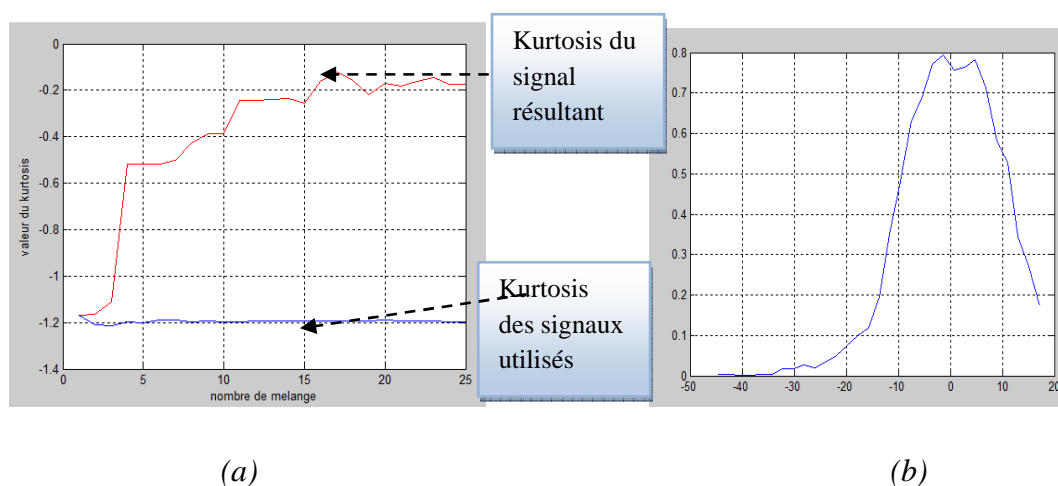
**Figure I.11:** Représentation de plusieurs distributions gaussiennes et le calcul du Kurtosis

En haut plusieurs PDF de distribution gaussienne, en bas le calcul numérique du Kurtosis de chacune d'elles en utilisant l'équation (1.124)

**I.7.2: Théorème de la limite centrale:**

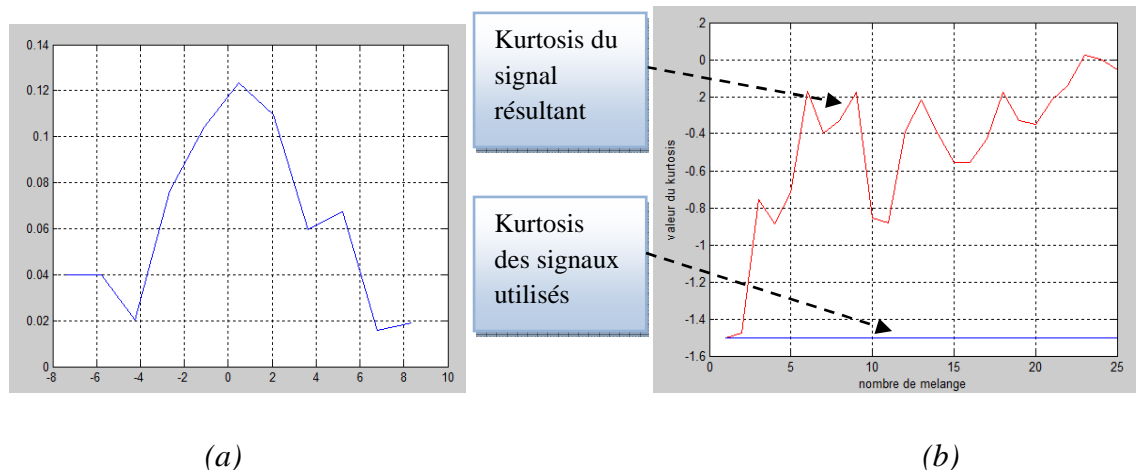
Si on a (n) variables aléatoires  $X_1, X_2, \dots, X_n$  statistiquement indépendantes, avec les moyennes  $(\mu_1, \dots, \mu_n)$  et les variances  $(\sigma_1^2, \dots, \sigma_n^2)$ , la variable aléatoire générée par la somme de ces n variables aléatoires, tend à avoir une densité de probabilité gaussienne [4][7], dont la moyenne égale à la somme des moyennes des variables aléatoires et une variance égale à la somme des variances des variables aléatoires génératrices. et le théorème est valable pour les processus stochastique. Considérons par exemple (n) processus stochastiques  $(S_1, \dots, S_n)$ , la somme de toutes ces réalisation sur un intervalle de temps T tend vers une réalisation x et que cette dernière possède après la construction de son histogramme une distribution gaussienne avec une moyenne égale à  $\sum_j \mu_j$ , et une variance égale à  $\sum_j \sigma_j^2$ .

Et pour clarifier ce théorème, on a construit l'histogramme de la somme de plusieurs signaux non gaussiens, et qui sont du même type, et on a appliqué cette procédure pour deux types de signaux, et les figures suivantes nous montrent le résultat:



**Figure I.12:**Exemple du théorème de la limite centrale pour des signaux uniformes

(a): en rouge la valeur du Kurtosis du siganal résultant de la somme de plusieurs signaux de de distribution uniforme en fonction du nombre des signaux, en bleu la valeur moyenne du Kurtosis de tous les signaux utilisés pour chaque nombre. (b): l'histogramme de la somme de 20 signaux différents de distribution uniforme



**Figure I.13:**Exemple du théorème de la limite centrale pour des signaux sinusoïdaux

(a): l'histogramme de la somme de 25 signaux sinusoïdaux de différentes fréquences, (b) : en rouge la valeur du Kurtosis du signal résultant de la somme de plusieurs signaux sinusoïdaux de différentes fréquences. En bleu la valeur moyenne du Kurtosis pour chaque nombre de signaux utilisés.

### I.8:Conclusion

Nous avons présenté dans ce chapitre les notions de base sur la théorie de probabilité et aux statistiques d'ordre supérieur, qui sont un outil indispensable pour le traitement aveugle des signaux. Nous avons commencé par les concepts fondamentaux sur les probabilités afin de présenter comment ce domaine de mathématique s'est introduit dans le domaine du traitement de signal par la variable aléatoire, qui est une projection des différents événements de la probabilité. Ensuite nous avons présenté les outils essentiels pour caractériser les variables aléatoires, après cela une petite introduction a été présentée sur les processus stochastiques et leurs traitement par la construction de leurs histogrammes, que ce dernier est une approximation de la fonction de densité de probabilité si ces processus sont stationnaires. À la fin de ce chapitre nous avons présenté le théorème de la limite centrale qui est l'un des concepts que le traitement aveugle des signaux s'est basé pour la résolution du problème de la séparation aveugle de sources.

# ***Chapitre II***

## ***Concepts généraux sur la Séparation de sources***



## II. 1: Introduction

Lorsqu'on traite des signaux provenant d'un canal de communication numérique, deux hypothèses sont présentées implicitement.

1- Les signaux transmis ont été formés d'une façon qu'ils soient des séquences de symboles à partir d'un ensemble fini de symboles [8].

2- Ces séquences sont supposées être des variables aléatoires indépendantes et identiquement distribuées [8].

Le scénario le plus général survient quand on éloigne les hypothèses précédentes, et en remplaçant le système (MIMO) par un système (MIMO) générique qui engendre le mélange et la distorsion de l'ensemble des signaux d'entrées, la reconstitution de ces signaux d'entrées après la procédure du mélange c'est ce qu'on appelle le problème de la séparation de sources ou la séparation de signaux (SS). On dit la séparation aveugle de sources (SAS) si la séparation est accomplie par les méthodes non supervisées [2] [8] [10].

Les techniques de la séparation de sources n'arrêtent pas de s'accroître intensivement depuis leurs apparitions dans les débuts des années 1980 jusqu'à nos jours. Du point de vue théorique le problème de la (SAS) prend remarquablement la notion de l'extraction de l'information [8], par conséquent et à partir du point de vue pratique, les techniques de séparation de sources sont pertinentes en de grands nombres d'applications, et pour citer quelques unes, la (SAS) est préoccupée par la compréhension et l'extraction d'informations à partir des données aussi diverses tel que les activités neuronales, audio, vidéo,....etc. [2] [8].

L'outil principal utilisé pour effectuer la séparation de sources est ce qu'on appelle l'(ICA), l'analyse en composantes indépendantes [7] [8], et qui peut être confirmé par le fait que l'(ICA) a pris le nom de la plupart des conférences concernant la séparation de sources et ces applications, néanmoins la séparation de sources n'utilise pas seulement l'ICA ou les techniques aveugles, puisque la séparation de sources semi aveugle prend en considération les informations préalables sur le problème posé, et qui se traduit par le nouveau nom (LVA latent variable analysis) [8], l'analyse en variables latentes et qui accentue le problème général en traitement du signal et l'extraction de l'information.

La structure de ce chapitre est caractérisée par les concepts généraux sur le sujet. Notre but est de présenter la fondation théorique de la (SAS), et nous avons pas l'intention de présenter

tous les aspects pour un sujet qui a reçu l'attention et la contribution par de nombreuses disciplines comme le traitement du signal, l'intelligence artificielle...etc. Pour cela notre travail est organisé d'une façon à présenter le problème de la (SAS) sous une forme générale et discuter quelques éléments de modélisation qui sont importants pour définir un modèle spécifique.

On va introduire les concepts pertinents qui relient entre la SAS avec l'ICA, en énonçant les définitions et les théorèmes fondamentaux, ensuite on énonce l'opération du blanchiment qui est une opération de prétraitement [8], après cela, on présente quelques critères pour effectuer l'opération de séparation.

Et on terminera le chapitre par donner quelques applications pratiques qui utilisent les techniques de la séparation aveugle de sources pour accomplir leurs tâches techniques.

## **II.2:Vue historique**

L'origine de la séparation de sources commence à être tracée dans les débuts des années 1980 particulièrement par les efforts de Héroult et Jutten, et Ans dans l'étude du problème du codage du mouvement, et durant cette décennie le problème de la séparation de sources attire graduellement l'attention de la communauté du traitement du signal, principalement en France et après cela en Europe. Depuis 1990, la séparation de sources et ses applications ont pris largement de la considération par de nombreux journaux et de conférences, et elle est devenue un sujet de recherche très important [8].

Un point de repère très important est celui du travail de P.Comon, qui relie la (SAS) avec l'(ICA), ses travaux ont ouvert de nouvelles perspectives et ont donné de nouveaux critères pour accomplir la séparation [8].

Les premiers travaux concernant l'utilisation des statistiques d'ordre supérieur, sont associés à J.Lacoum et de J-F Cardoso, et qui ont utilisé les approches de la (PCA non linéaire), et que les concepts de cette dernière sont très proches aux concepts des réseaux de neurones artificiels. Ces techniques ont aussi pris naissance par les travaux de E.Oja. Une autre approche classique qui a une relation forte avec le domaine des réseaux de neurones artificiels est celle qui utilise la méthode pour la maximisation de l'information (Infomax) et qui a débuté par les travaux de Bell et Sejnowski. Notre liste se termine par les travaux de A.Cichocki et R-Unbenhauen, qui sont les premiers à utiliser l'idée du gradient naturel pour la (SAS) et qui sont aussi les premiers à appliquer la méthode de la (Fast ICA). D'autres

approches ont été développées en utilisant d'autres suppositions concernant le signal à extraire, par exemple les approches qui se basent sur l'idée d'explorer la structure de corrélation des sources. Une autre approche est celle de M.D Plumbly qui suppose que les signaux sources sont toujours non négatifs, et c'est ce qui a donné naissance à ce qu'on appelle la (non négative ICA) [8].

Le problème de la (SAS) non linéaire a reçu une très grande importance depuis 1990, et les efforts les plus importants sont ceux qui ont été fondés en utilisant les réseaux de neurones artificiels par les travaux de Burel, et les clés de référence pour ce sujet qu'on l'appelle (PNL Post non linéaire mixtures) sont les travaux de Jutten et de Taleb, et l'étude sur ce domaine reste un sujet de recherche prolifique [8].

#### Remarque:

Après cette brève vue historique il est important de mentionner que notre travail est focalisé et restreint que pour l'étude de la séparation de sources des mélanges linéaires instantanés, afin de mieux comprendre le sujet et pour avoir une bonne initiation sur ce domaine.

### **II.3: Le problème de la séparation de sources**

La figure II.1 représente le problème général de la séparation de sources, où on a un ensemble de signaux d'informations qui ont subi une transformation par un système de mélange, et aussi une distorsion éventuelle du système (MIMO). Les signaux résultants sont captés par un ensemble de capteurs, le but de la séparation de sources est de récupérer les signaux originaux à partir des sorties des capteurs en utilisant un système de séparation adéquat [2] [6] [8].

On parle d'une séparation aveugle, si le système de mélange est inconnu et que les signaux désirés ne sont pas disponibles pour n'importe quel procédure d'apprentissage [6] [7] [8], et le problème peut être formulé comme suit:

Considérons un ensemble de N signaux désignés par sources, dont les échantillons forment le vecteur source  $S(n)$ , et on a aussi un ensemble de M signaux d'observations qui sont organisés dans le vecteur  $Y(n)$ . En général les observations se présentent comme suit:

$$Y(n)=F(S(n),\dots, S(n-L),n(n),n) \quad (2.1)$$

Où  $F(\cdot)$  est la fonction du mélange, il est claire que si  $F(\cdot)$  est connue, il suffit d'estimer le système inverse pour récupérer les signaux sources .en plus si on a la possibilité d'utiliser les

techniques d'apprentissage, il est nécessaire d'identifier la fonction du mélange par une manière non supervisée afin d'obtenir le système inverse. Le défi à relever est comment obtenir ce système inverse où les estimés doivent être des versions approximatives des signaux sources, en se basant seulement sur les observations et avec un minimum d'informations sur les signaux sources?. Même si on atteint la solution pour le problème mentionné, cette solution n'est pas une solution générale pour le problème de la séparation de sources, pour cela il faut traiter chaque problème en se basant sur ces caractéristiques spécifiques [8].

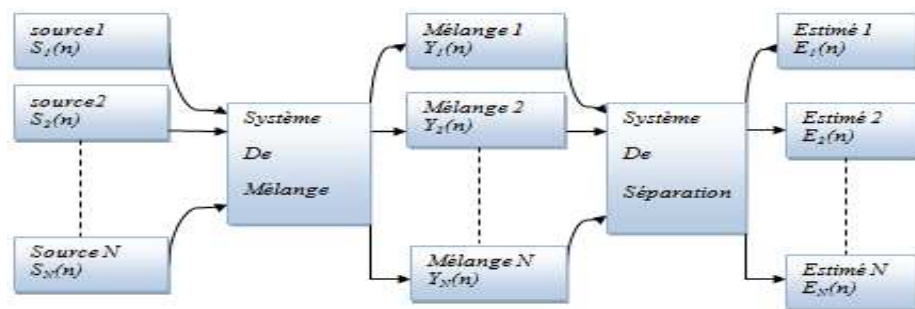


Figure II.1: Le problème général de la séparation de sources

## II.4:Caractéristiques du système de mélange

### II.4.1:Linéaire ou non linéaire:

On dit que le système de mélange  $F(\cdot)$  est linéaire si il obéit au principe de superposition :

$$F(\alpha_1 S_1(n) + \alpha_2 S_2(n)) = \alpha_1 F(S_1(n)) + \alpha_2 F(S_2(n)) \quad (2.2)$$

Pour toutes les constantes  $\alpha_i$  et tous les vecteurs sources  $S_1(n)$  et  $S_2(n)$ , autrement on dit que le système de mélange est non linéaire.

### II.4.2: Instantané ou convolutif

Si les observations représentent un mélange des sources avec des échantillons de différents instant  $L > 0$ , alors le système est dit convolutif, contrairement à cela si on a affaire avec un système où  $L = 0$ , le système de mélange est dit instantané.

### II.4.3 : Selon le nombre de sources et le nombre d'observations

Dans le cas où le nombre de sources est inférieur au nombre de capteurs (observations)  $M > N$ , on dit que le système est surdéterminé, dans le cas contraire  $M < N$  on dit que le système de mélange est indéterminé.

**II.5:Principe de la solution du problème de la SAS**

Considérons le cas le plus simple où le système de mélange est linéaire instantané et sans bruit, donc ce système est caractérisé par sa matrice de mélange  $A_{M \times N}$  et le vecteur d'observations est donné par:

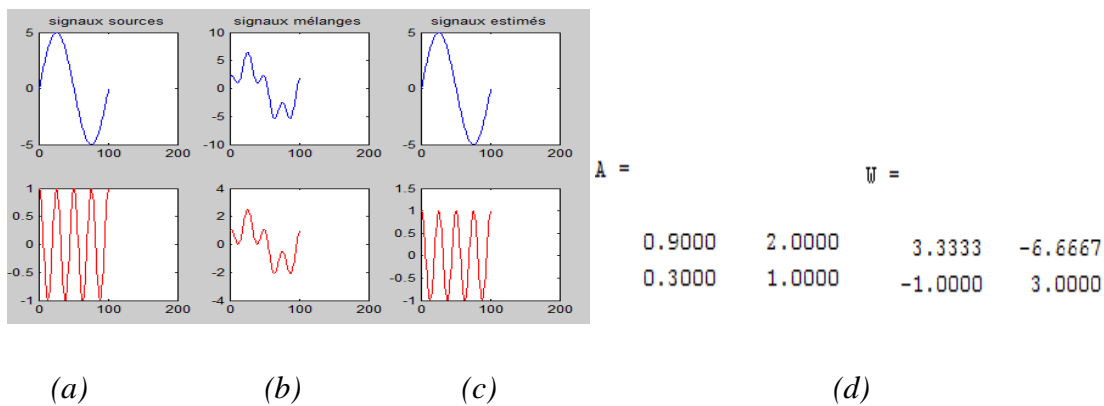
$$Y(n) = AS(n) \tag{2.3}$$

Ce qui veut dire que les observations sont des combinaisons linéaires des sources.

Et afin de présenter le principe de base pour la résolution du problème de la séparation aveugle de sources dans le cas où le mélange est sans bruit, on doit supposer que le nombre de sources égale au nombre de capteurs (observations)  $A_{N \times N}$ , et on suppose aussi que les sources sont des processus stochastiques stationnaires[8]. En se basant sur ces conditions et par la supposition que la matrice de mélange est inversible, la séparation est achevée par l'obtention du système inverse  $W$  tel que

$$E(n) = AY(n) = WAS(n) = IS(n) = S(n) \tag{2.4}$$

c.à.d.  $W=A^{-1}$  comme le montre la figure ci dessous:



**Figure II.2-** Exemple démonstratif sur le but de la séparation de sources

(a): signaux sources, (b): observations, (c): signaux estimés

(d): A: système de mélange, W : le système de séparation

Donc le but de la (SAS) est de trouver le système inverse  $W$  sans aucune information sur le système de mélange  $A$  ni sur les signaux sources  $S$ , pour cela il faut se baser sur l'hypothèse que les signaux sources sont statistiquement indépendants afin que la séparation puisse être atteinte.

## II.6:L'analyse de la composante indépendante (ICA)

Mathématiquement les éléments d'un vecteur aléatoire  $S$  sont statistiquement indépendants

Si:

$$P_s(S) = P_{s1}(S_1) \times P_{s2}(S_2) \times \dots \times P_{sN}(S_N) \quad (2.5)$$

où  $P_s(S)$ , est la densité de probabilité conjointe (PDF) des variables aléatoires  $S_1, S_2, \dots, S_N$  et  $P_{Si}(S_i)$  est la densité de probabilité marginale de la source  $S_i$ . Il est démontré que si les sources sont mutuellement indépendantes, alors il est possible d'obtenir la matrice de séparation d'une manière non supervisée [8]. L'idée principale est de trouver une matrice  $W$  d'une façon où les signaux estimés auraient cette expression:

$$E=WY \quad (2.6)$$

Et qui sont aussi mutuellement indépendants. La démonstration de cette propriété est basée sur le théorème de (Darmois-Skitovich)

### II.6.1:Théorème de (Darmois-Skitovich[8])

Soient  $S_1, S_2, \dots, S_N$  un ensemble de variables aléatoires de valeurs moyennes nulles et statistiquement indépendantes, et soient  $Y_1$  et  $Y_2$  définies comme suit:

$$Y_1 = a_1 S_1 + a_2 S_2 + \dots + a_N S_N$$

$$Y_2 = b_1 S_1 + b_2 S_2 + \dots + b_N S_N$$

Si  $Y_1$  et  $Y_2$  sont mutuellement indépendantes, alors toutes les variables aléatoires  $S_i$  pour  $a_i, b_i \neq 0$  sont des variables aléatoires de distribution gaussienne. Selon ce théorème il n'est pas possible d'obtenir des variables aléatoires indépendantes à partir d'un mélange linéaire de sources non gaussiennes, par conséquent si on considère que les coefficients  $a_i$  et  $b_i$  sont les paramètres de la matrice  $WA$ , alors on ne peut avoir des estimés indépendants à partir des signaux sources non gaussiens, sauf si les signaux sources ne sont pas mélangés, pour cela la séparation de sources est achevée via l'ICA

### II.6.2: Définition de la (ICA)

la ICA d'un vecteur aléatoire  $Y = [Y_1, Y_2, \dots, Y_N]^T$  consiste à déterminer la matrice  $W$  d'une manière où les éléments du vecteur  $E = WY$  soient statistiquement indépendants que possible [2] [7] [8], pour cela la récupération des sources originales se base sur deux conditions:

a) L'indépendance statistique des sources.

b) La non gaussienne des sources.

Mais il est important de mentionner que les estimés seront des versions des sources permutées et multipliées par des facteurs comme le montre cette expression [2][7][8].

$$E = \rho PS \quad (2.7)$$

$\rho$ : une matrice diagonale.

$P$ : matrice de permutation.

Et les conditions pour que les sources soient reconstituées en utilisant l'ICA peuvent être résumées par ce théorème.

### II.6.3: Théorème de séparabilité

Le système présenté sur la figure II.1 est séparable en utilisant l'ICA, c.à.d. il est possible d'obtenir  $W$  de telle façon que  $E = WY$  représente les sources permutées et multipliées par des facteurs, si la matrice  $A$  est inversible et il y'a tout au plus une source gaussienne, pour la première condition il est évident puisque on cherche la matrice de séparation qui doit dans un sens inverser la matrice de mélange, pour la deuxième condition c'est pour des raisons des limitations des statistiques d'ordre supérieur [2] [8].

### II.7: Le prétraitement (WHITENING- le blanchiment)

Si deux variables aléatoires sont statistiquement indépendantes alors elles sont aussi non corrélées, considérons le cas où les deux sources ont les valeurs moyennes nulles et les variances égales à l'unité donc

$$R_s = E \{SS^T\} = I \quad (2.8)$$

En se basant sur cette condition la matrice d'auto corrélation des observations est donnée par:

$$R_y = E \{YY^T\} = A R_s A^T = AA^T \quad (2.9)$$

Par cette équation la matrice de mélange peut être exprimée par la décomposition des valeurs singulières (SVD) [8], en :

$$A=U \rho^{1/2} V^T \quad (2.10)$$

Et puisque V est une matrice unitaire ( $V^{-1}=V^H=conjV^T$ ) alors:

$$R_y=U \rho U^T \quad (2.11)$$

Considérons maintenant cette transformation :

$$Z=TY \quad (2.12)$$

Où:  $T= \rho^{-1/2} U^T \quad (2.13)$

Pour cette transformation la matrice d'auto corrélation est donnée par :

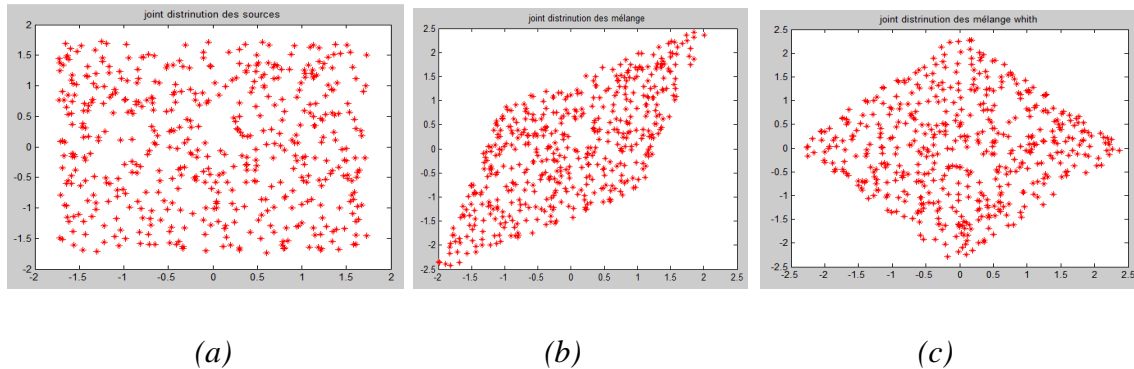
$$\begin{aligned} R_z &= E \{ZZ^T\} \\ &= \rho^{1/2} U^T R_y U \rho^{-1/2} \\ &= \rho^{1/2} U^T U \rho U^T U \rho^{-1/2} = I \end{aligned}$$

c.à.d. les nouvelles observations sont non corrélées, dans ce cas on dit que l'observation a été blanchie. il est important de mentionner que la transformation T n'est pas unique depuis que  $T=Q \rho^{-1/2} V^T$  pour n'importe quelle matrice orthogonale Q produit aussi des observations non corrélées [8], donc il est claire que ce prétraitement ne produit pas les sources originales, c'est pour cela il n'est pas possible de séparer des sources gaussiennes. Par exemple deux sources gaussiennes et de valeurs moyenne nulles, pour avoir des estimées non corrélées implique aussi qu'elles soient aussi indépendantes, et les signaux obtenus ne sont pas nécessairement les sources originales.

Prenons comme exemple deux sources indépendantes et uniformément distribuées dans l'intervalle  $[-\sqrt{3}, \sqrt{3}]$ , et avec la distribution conjoint  $P(x, y)=1/12$  si  $-\sqrt{3} \leq x, y \leq \sqrt{3}$ , comme l'illustre la figure II.3.a, la matrice de mélange  $A = \begin{bmatrix} 1 & 0.5 \\ 0.3 & 0.9 \end{bmatrix}$  [8], après le mélange des signaux la distribution du mélange deviendra comme l'illustre la figure II.3.b. On remarque que la matrice de mélange déforme et fait tourner la distribution originale, et après la procédure du prétraitement, la distribution obtenue ressemble à la distribution originale sauf



pour le facteur de rotation comme le montre la figure II.3.c, donc l'étape qui reste, est de déterminer la matrice de rotation pour recouvrer les sources originales.



**Figure II.3:** La procédure du blanchiment

(a) *distribution des signaux sources* (b) *distribution des signaux mélanges*, (c) *distribution des signaux après blanchiment (sans rotation)*

## II.8: Critères pour l'analyse en composantes indépendantes

Comme on a mentionné au début de ce chapitre, on a pour but de trouver une structure de séparation qui nous fournisse des signaux estimés indépendants que possible. Différents critères ont été proposés afin d'implémenter l'idée de l'ICA, on va discuter dans ce qui va suivre quelques critères.

### II.8.1: Critère basé sur l'information mutuelle

Les concepts de la théorie de l'information et de l'information mutuelle peuvent être utilisés pour quantifier l'indépendance entre les variables aléatoires [2] [7] [8]. Dans le but de clarifier cette idée il est nécessaire de définir quelques concepts concernant la théorie de l'information.

#### II.8.1.1: Entropie

L'entropie est la mesure du degré de désordre d'un système par rapport à son état initial probable. En communication c'est la mesure de l'incertitude par rapport à la nature du message [5][7][8]. Soit  $\alpha$  une variable aléatoire qui est caractérisée par sa fonction de la densité de probabilité  $p_{\alpha}(\alpha)$ , l'entropie de la variable aléatoire est définie par:

$$H(\alpha) = -E[\log p_{\alpha}(\alpha)] = -\int_{-\infty}^{\infty} p_{\alpha}(\zeta) \log p_{\alpha}(\zeta) d\zeta \quad (2.14)$$

Et pour le cas discret :

$$H(\alpha) = -\sum_i P_i \log P_i \quad (2.15)$$

### II.8.1.2: Entropie conditionnelle

Soit  $P_{\alpha|\beta}(\alpha|\beta)$  la densité de distribution de probabilité conditionnelle de  $\alpha$  selon  $\beta$ , l'entropie conditionnelle est donnée par :

$$H(\alpha|\beta) = -E[\log P_{\alpha|\beta}(\alpha|\beta)] = -\iint_{-\infty}^{\infty} P_{\alpha,\beta}(\zeta, \nu) \log P_{\alpha|\beta}(\zeta|\nu) d\zeta d\nu \quad (2.16)$$

Cette équation représente l'incertitude d'une variable aléatoire par rapport à l'observation d'une autre variable aléatoire [5][8].

### II.8.1.3: Information mutuelle

L'information mutuelle entre deux variables aléatoires  $\alpha$  et  $\beta$  est définie comme suit

$$I(\alpha, \beta) = H(\alpha) - H(\alpha|\beta) = H(\beta) - H(\beta|\alpha) \quad (2.17)$$

Le concept de l'information mutuelle peut être étendu pour le cas d'un vecteur de  $N$  variables aléatoires  $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_N$  [5] [8], et peut être exprimé comme suit:

$$I(\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_N) = \sum_i^N H(\alpha_i) - H(\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_N) \quad (2.18)$$

Notons que la minimisation de l'information mutuelle est équivalente à rendre l'entropie conjoint  $H(\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_N)$  proche à la somme des entropies marginales [8]. Dans le cas d'un système de mélange linéaire, et en évaluant l'information mutuelle des sorties nous aurons:

$$\begin{aligned} I(E) &= \sum_i^N H(E_i) - H(E) \\ &= \sum_i^N H(E_i) - H(Y) - \log |\det W| \end{aligned} \quad (2.19)$$

Il est important de mentionner que  $H(Y)$  ne dépend pas de la matrice  $W$ , et ce qui mène au critère de séparation suivant :

$$\text{MIN} \sum_i^N H(E_i) - \log |\det W| \quad (2.20)$$

### II.8.2: Critères basé sur les statistiques d'ordre supérieur

Un autre critère pour que l'ICA accomplisse la tâche de séparation, est le critère basé sur l'exploitation des moments est cumulants croisés des observations. L'idée principale se base sur le principe de l'indépendance statistique. Si on a plusieurs signaux ou un ensemble de processus stochastiques  $(X_1, \dots, X_k)$  qui sont physiquement et statistiquement indépendants alors tous les cumulants croisés de ces variables aléatoires sont toujours nuls [8] [10][13]. si on considère par exemple les deux variables aléatoires  $X$  et  $Y$ , et afin d'assurer qu'elles soient indépendantes la relation suivante doit être vérifiée:

$$C(X^s, Y^q) = 0 \quad (2.21)$$

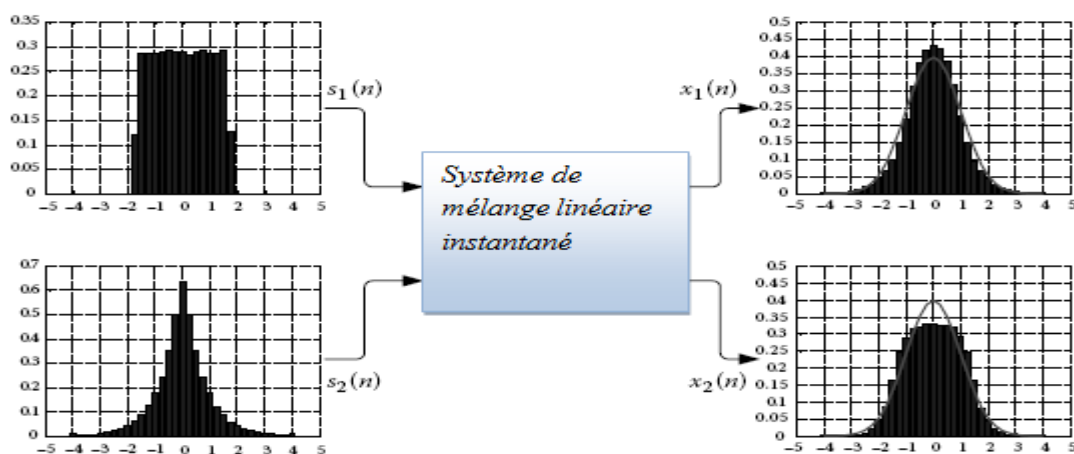
Pour n'importe q, s = 1, ..., ∞. Donc l'indépendance entre les signaux est liée par tous les cumulants croisés. En séparation de sources il a été prouvé que la séparation peut être achevée en utilisant les cumulants d'ordre quatre [10] [13] [18], à condition que les signaux ont le Kurtosis non nul ou tout au plus un signal a le Kurtosis nul. Donc ce critère est basé sur la minimisation des cumulants croisés du quatrième ordre en utilisant cette expression:

$$\min \sum |C(E_i, E_j, E_k, E_l)|^2 \quad (2.22)$$

À l'exception de i=j=k=l,

### II.8.3: Critère basé sur la maximisation de la non gaussienne

Selon le théorème de la limite centrale la fonction de densité de probabilité de la somme des variables statistiquement indépendantes tend vers une distribution gaussienne [7] [8], de ce fait la somme de deux variables aléatoires où l'une d'elles a une distribution uniforme et l'autre une distribution laplacienne, génère une nouvelle variable que cette dernière tend approximativement à avoir une distribution gaussienne comme le montre la figure ci dessous:

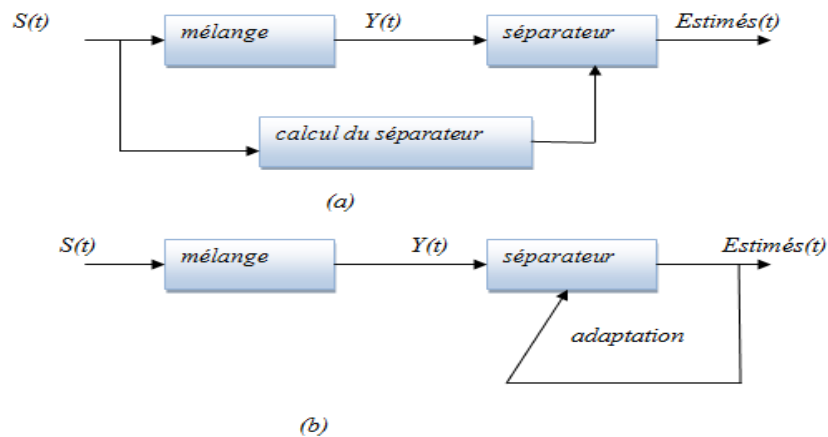


**Figure II.4:** Mélange entre la distribution uniforme et la distribution Laplacienne

On a introduit la mesure du Kurtosis au premier chapitre, et une de ces propriétés est qu'il a la valeur nulle pour les distributions gaussiennes, alors en se référant au modèle de mélange de la figure (Figure II.1) on déduit facilement que les observations doivent avoir un Kurtosis qui tend vers la valeur nulle, alors pour que la séparation soit atteinte on doit appliquer au système de séparation une fonction coût que cette dernière doit maximiser le Kurtosis de chaque observation si tous les signaux originaux ont les Kurtosis non nuls.

## II.9: Systèmes de séparation

Deux méthodes ont été utilisées pour construire le système de séparation, ces méthodes sont les méthodes directes appelées aussi hors ligne, et les méthodes adaptatives (en ligne) [6], Comme le montre la figure suivante:



**Figure II.5:** Les systèmes pour la séparation de sources

(a): les systèmes directs (b): les systèmes en boucle

Pour les méthodes directes, elles se basent principalement sur l'application des techniques et les théorèmes des méthodes numériques et du calcul matriciel pour la résolution des systèmes d'équations. Ces techniques exploitent les caractéristiques stochastiques des observations en utilisant des calculs analytiques et algébriques par un développement mathématique pour avoir des expressions bien formées, par exemple avoir les expressions des moments et cumulants croisés des observations ou les signaux originaux en fonctions des paramètres du système de mélange, ensuite estimer les paramètres du séparateur en se basant sur une fonction coût que cette dernière doit minimiser toutes les valeurs des cumulants croisés des solutions trouvées ou estimées, ces méthodes donnent de meilleurs résultats comparant aux méthodes adaptatives, mais leur problème est comment trouver l'expression mathématique la plus pertinente pour des observations de dimension supérieure à deux ?

les méthodes adaptatives et comme dans le cas des filtres adaptatifs en traitement du signal qui se basent essentiellement sur la minimisation de l'erreur quadratique en utilisant la technique du gradient naturel [1][3], on doit trouver les paramètres les plus pertinents pour avoir la meilleure approximation. Les méthodes adaptatives en séparation de sources ont pour but de trouver les paramètres optimaux du séparateur pour atteindre la séparation. L'idée principale

de ces techniques est d'utiliser le gradient d'une fonction cout pour minimiser ou maximiser un critère [8]. Par exemple si le système de séparation est  $W$  et que ce système doit maximiser le Kurtosis de chaque observation, alors les paramètres de  $W$  doivent passer par des transformations de tel façon que le Kurtosis de chaque estimation soit le plus grand que possible.

## II.10: Applications de la séparation de sources

On a mentionné à l'introduction de ce chapitre que les algorithmes et les techniques de la SAS sont appliqués dans plusieurs champs de recherches pour accomplir leurs taches techniques, et pour cela on va donner deux exemples d'applications et qui sont les domaines les plus connus en traitement du signal numérique.

On commence par le premier, qui est le traitement de la parole où la SAS est introduite dans ce domaine par ce que on appelle le problème du Cocktail Party.

### II.10.1: Le problème du Cocktail Party

Le problème du Cocktail Party peut être décrit par l'habilité d'une personne à se concentrer sur un seul locuteur parmi plusieurs conversations et avec le bruit de l'entourage, ce problème a pris une très grande attention et considéré comme un problème de challenge. Techniquement ce problème est appelé la déconvolution aveugle des canaux multiples. Le but est de séparer plusieurs signaux (paroles) captés à partir de plusieurs microphones [2].

Le cerveau humain peut accomplir cette tache facilement c'est pour cela les premiers algorithmes de la séparation aveugle de sources sont basés sur l'application des techniques des réseaux de neurones artificiels, [2]. Le cerveau humain peut focaliser le sens de l'ouïe sur une source parmi plusieurs mais la procédure n'est pas complètement définie pour la simuler artificiellement, pour ce problème la SAS est une approximation de cette procédure.

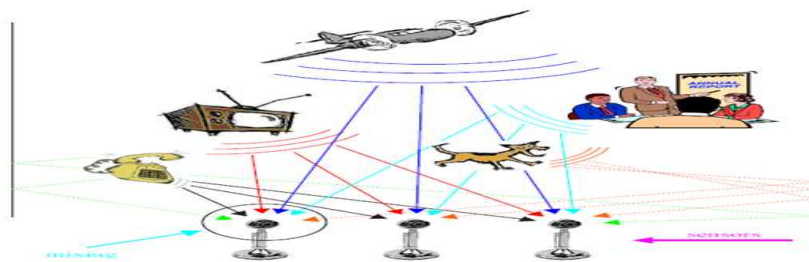
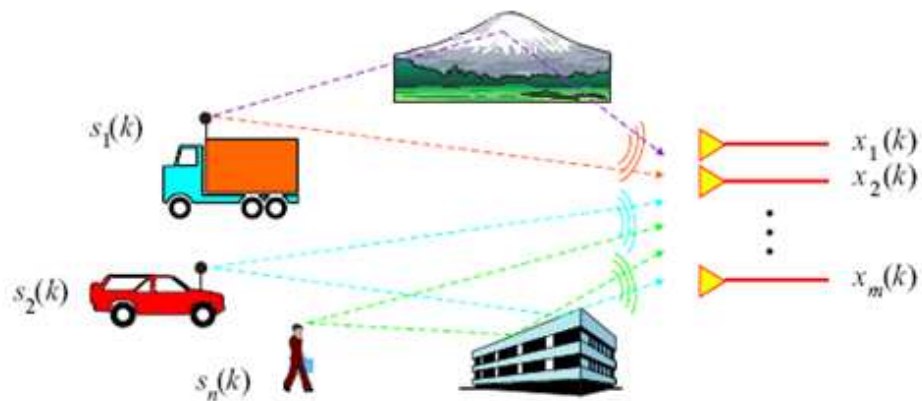


Figure II.6: Le problème du Cocktail party

### II.10.2: Application de la SAS pour les systèmes de Communications numériques

Les techniques du traitement du signal aveugle surviennent aussi dans une large variété d'applications des communications numériques, par exemple en Radio numérique, les lignes d'abonnés numériques, enregistrement magnétique numérique multi piste, les systèmes de communications multi utilisateurs et a accès multiple, les capteurs sonores des radars.Etc. [2].

La SAS est un outil prometteur pour ces système qu'on les symbolise par MIMO, par exemple on peut effectuer les techniques de la séparation de sources pour la suppression de l'interférence inter symboles (ISI), L'interférence des canaux adjacents (ACI) et l'interférence des accès. L'état de l'art de ces champs de recherche nécessite des connaissances préalables pour estimer les systèmes optimaux qui accomplissent chaque tâche concernant la séparation aveugle dans les systèmes MIMO [2].



**Figure II.7:** Le problème de la séparation de sources pour les réseaux de communication

**II.11: Conclusion**

Ce chapitre avait pour but de présenter une vue générale sur le problème de la séparation de sources, où on a commencé par présenter la nature du problème et dans quel domaine il appartient, ensuite on a présenté une petite vue historique pour le traitement de ce problème, après ça on l'a exposé d'une façon générale par la représentation de différents cas possibles de son apparition, mais pour sa solution on a pris que le cas simple qui est le cas du mélange linéaire instantané des signaux, où on a donné un exemple sur le but de la séparation des signaux mélangés par un système inversible, et dans cet exemple la solution est idéale et ce n'est pas le cas en pratique, car comme tout domaine il y'a des indéterminations concernant la résolution du problème et qui ont été mentionnées dans ce chapitre. L'idée de base pour la résolution du problème de la séparation de sources est ce qu'on appelle l'indépendance statistique des sources, si les sources sont physiquement indépendantes alors elles sont aussi statistiquement indépendantes, donc on doit extraire des signaux statistiquement indépendants que possible. Beaucoup de critères ont été utilisés pour atteindre la séparation mais on a présenté que trois. Même les techniques utilisées pour extraire les signaux sources se différencient selon le critère et la procédure utilisés, dans ce chapitre on a mentionné les techniques directes et les techniques adaptatives pour résoudre le problème et on va présenter quelques algorithmes de chacune d'elles dans le chapitre suivant. Bien que ce problème se présente dans beaucoup de domaines, on a terminé ce chapitre par donner que deux exemples sur l'utilisation des techniques de la séparation de sources et comment ce problème se présente dans ces deux domaines.

# ***Chapitre III***

***Algorithmes directs et adaptifs pour  
les mélanges linéaires instantanés***



### III.1:Introduction

Ce chapitre a pour but de présenter quelques algorithmes célèbres qui ont été mis en œuvre pour la résolution du problème de la séparation aveugle de sources. On a mentionné dans le chapitre précédent que aveugle signifie que la résolution est atteinte d'une façon non contrôlée, et dans le domaine de la SAS signifie que on n'a pas des connaissances préalables ni sur les signaux sources ni sur le système de mélange. On a aussi discuté les différentes techniques pour achever la séparation en utilisant l'ICA, ces techniques se divisent en deux catégories, les techniques directes (hors ligne) et les techniques adaptatives (en ligne), et chacune de ces méthodes possède ses avantages et ses inconvénients.

Ce chapitre est organisé d'une façon à donner une vue globale sur ces techniques, où on commence par les techniques algébriques et on débute cette catégorie par l'algorithme de Ali Mansour qui se base sur la non corrélation des sources [22]. Les deux autres techniques algébriques que on va les présenter dans ce chapitre sont l'algorithme de Ali Mansour et l'algorithme de Taro Yamaguchi, ces deux derniers algorithmes ont été fondés sur l'indépendance statistique des sources par l'estimation des paramètres du système de mélange en utilisant les cumulants croisés des observations [19][28].

Pour les algorithmes adaptatifs on commence par l'algorithme de (Hérault et Jutten), qui se présente en deux formes, le réseau direct et le réseau récurrent, ces deux réseaux se basent sur les méthodes d'apprentissages des réseaux de neurones artificiels [26]. Il a été prouvé en pratique que ces algorithmes sont fiables que dans des cas spécifiques et sous des conditions nécessaires, hors ces conditions l'application de ces algorithmes ne donne pas des solutions acceptables [20].

On terminera ce chapitre par deux algorithmes qui se basent aussi sur les techniques neuronales, ces techniques sont l'algorithme de (A-Cichocki et R Unbehauen) et l'algorithme (A-Cichoki et S-Amari). Ces techniques sont l'extension du réseau récurrent de (Hérault et Jutten), mais elles sont plus stables et donnent de meilleurs résultats. [20]

### III.2:Méthode directes (hors ligne)

Les techniques algébriques pour la SAS sont des algorithmes qui s'exécutent algébriquement et elles sont basées sur l'estimation des paramètres du système de mélange. Même ces techniques se différencient selon les suppositions posées sur le système de mélange. On va introduire dans ce chapitre trois catégories de ces techniques hors ligne, où on débute par

une technique qui se base sur une approche géométrique pour extraire les signaux sources, ensuite on introduira deux techniques qui se basent sur l'exploitation des cumulants croisés des observations pour l'estimation des paramètres de la matrice de séparation.

### **III.2.1: Approche géométrique de (Ali Mansour [22])**

Cet algorithme utilise les techniques du blanchiment (Sphering ou whitening) des observations, bien que ces techniques de sphering nous donnent des signaux non corrélés, mais les signaux estimés ne sont pas nécessairement une version des signaux sources originaux. Donc il faut avoir une dernière procédure pour recouvrer les signaux originaux, cette procédure est la rotation des signaux estimés par un angle de rotation qu'on doit le déterminer dans le cas de deux sources, afin que les signaux soient indépendants que possible.

Alors l'idée de base de cet algorithme est de trouver un séparateur de telle façon que ce système mène à la solution suivante :

$$G=WA=PD \tag{3.1}$$

Où P est une matrice de permutation et D une matrice inversible.

L'algorithme proposé par Ali mansour s'exécute en deux étapes, étape de transformation et l'étape de rotation.

#### **III.2.1.1: Transformation**

Cette étape est l'étape de la transformation des signaux de l'observation en des signaux orthogonaux en utilisant la décomposition de Cholesky [22] [1], de la matrice d'auto covariance des signaux mélangés:

$$R=LL^T \tag{3.2}$$

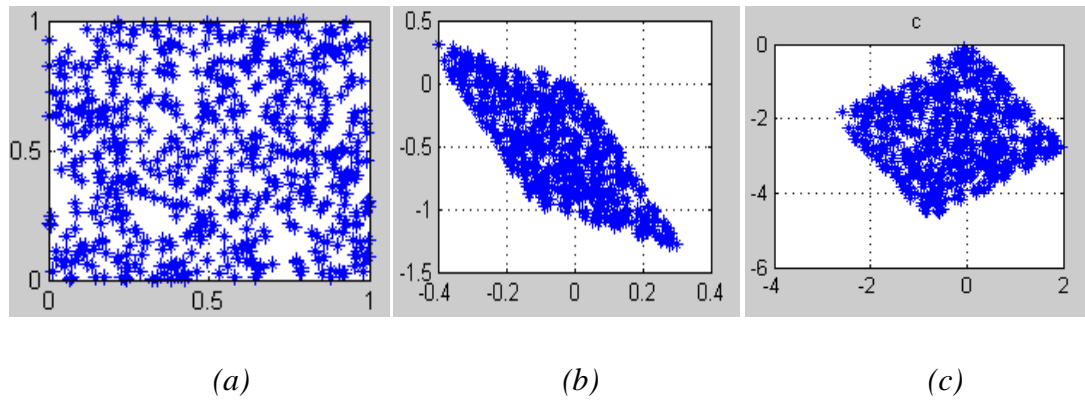
Si on pose  $K = L^{-1}$

l'expression des signaux blanchis est la suivante:

$$Z(t)=KY(t) \tag{3.3}$$

Et les signaux obtenus sont non corrélés comme le montre la figure suivante pour le cas de deux signaux qui ont les fonctions de densité de probabilités uniformes et la matrice de mélange :

$$A = \begin{pmatrix} 0.3188 & -0.4336 \\ -1.3077 & 0.3426 \end{pmatrix}$$



**Figure III.1:** Transformation des signaux mélanges en des signaux non corrélés

(a) : distribution des amplitudes des signaux sources, (b): distribution des signaux mélange,  
(c) : distribution des signaux non corrélés.

### III.2.1.2:Rotation

Le but de cette étape est de rendre les signaux non corrélés en des signaux indépendants. Pour le cas de deux sources, la matrice de rotation a l'expression suivante:

$$R(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \quad (3.4)$$

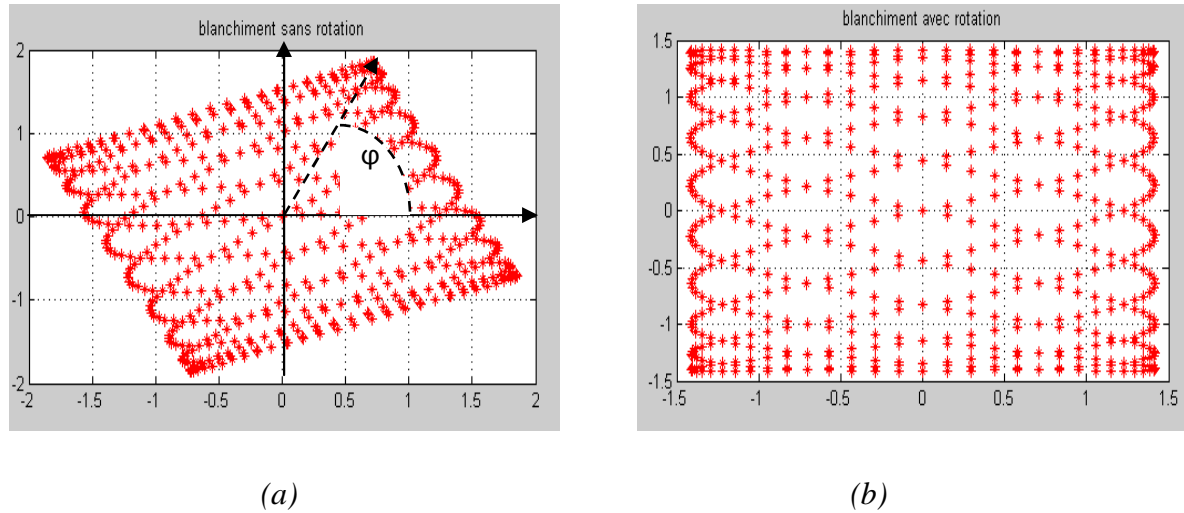
Alors les signaux estimés sont obtenus en appliquant cette équation:

$$E(t) = R(\alpha) Z(t) \quad (3.5)$$

Il a été prouvé que pour les signaux qui ont une fonction de densité de probabilité uniforme ou presque, l'angle de rotation des signaux non corrélés est le suivant:

$$\alpha = \pi/4 - \varphi \quad (3.6)$$

où  $\varphi$  est l'angle qui doit être ramené à la valeur  $\pi/4$ . Si cet angle est égal à cette valeur, on n'a pas besoin de faire la rotation pour le cas des sources uniformes [22]. Pour le cas des signaux sinusoïdaux et les signaux qui ont la densité de probabilité uniforme, l'angle  $\varphi$  peut être déterminé par localiser le point le plus loin de la distribution des amplitudes des signaux non corrélés, comme le montre la figure suivante:



**Figure III.2:** La procédure de rotation (deux sources)

(a) détermination de l'angle  $\alpha = \pi/4 - \varphi$ , (b) résultat de la rotation pour le cas des signaux sinusoidaux

Pour le cas de trois sources, la rotation doit être faite en trois étapes et par rapport aux trois axes par les relations suivantes:

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix} \quad (3.7)$$

$$R_y(\alpha) = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} \quad (3.8)$$

$$R_z(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.9)$$

Et les signaux estimés sont obtenus comme suit:

$$E(t) = R_z(\alpha) Z_2(t) \quad (3.10)$$

Où  $Z_2(t) = R_y(\alpha) Z_1(t)$  et  $Z_1(t) = R_x(\alpha) Z(t)$

Un exemple démonstratif se présente sur la figure suivante pour le cas des signaux uniformes qui sont mélangés par la matrice:

$$A = \begin{bmatrix} -0.0396 & 0.0306 & 0.4593 \\ 0.0945 & 1.0699 & -0.1500 \\ 0.6210 & -0.6220 & 0.4346 \end{bmatrix}$$

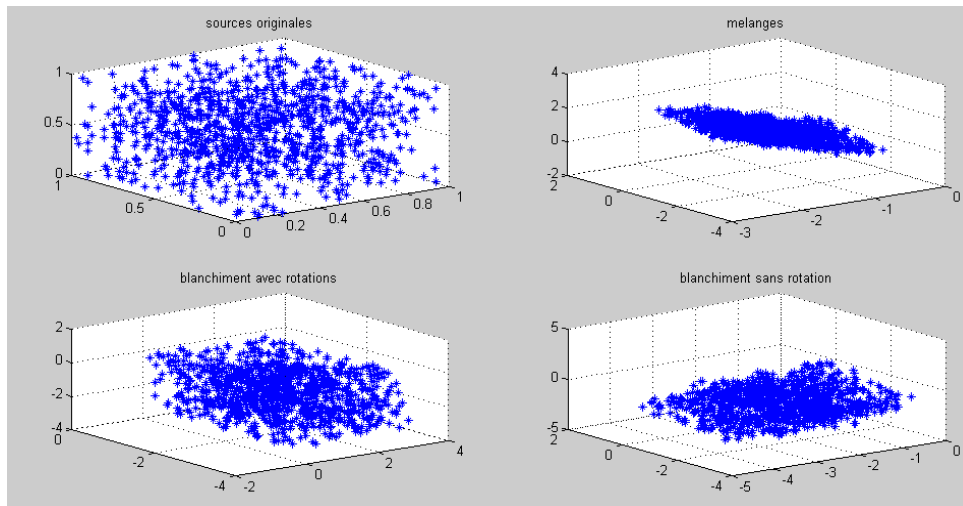


Figure III.3: La procedure de rotation(trois sources)

*Exemple des résultats de l'application de l'approche géométrique sur trois signaux de distributions uniformes.*

### III.2.2 Méthode basée sur les cumulants croisés (Ali Mansour [19])

Cet algorithme est basé sur l'estimation de la matrice de mélange  $A$  pour le cas des mélanges linéaire instantané, en se basant sur l'hypothèse que les signaux sources sont statistiquement indépendants et non gaussiens, la matrice de mélange est estimée par la résolution du polynôme du quatrième degré dans le cas de deux sources en impliquant les cumulants croisés des observations pour la détermination des coefficients du polynôme.

#### III.2.2.1:Modèle du mélange

L'algorithme est proposé pour le cas de deux sources linéairement mélangées par la matrice  $A$  non singulière. il est connu que les signaux sources peuvent être récupérer mais pas avec les mêmes positions ni les même amplitudes [8]. de ce fait l'algorithme suppose que la matrice de mélange  $A$  est de la forme suivante:

$$A = \begin{bmatrix} 1 & \alpha \\ \beta & 1 \end{bmatrix} \quad (3.10)$$

Alors les signaux mélanges ont l'expression suivante :  $X=AS$

Pour que la séparation soit atteinte il faut que les coefficients  $\beta$  et  $\alpha$  satisfassent la condition suivante :

$$1 - \alpha\beta \neq 0 \quad (3.11)$$

En appliquant la relation (1.105) du premier chapitre et en développant les cumulants du quatrième ordre des observations. Les cumulants croisés en fonction des coefficients de la matrice de mélange sont :

$$\begin{aligned}
 C_{11} &= \beta P_1 + \alpha P_2 & C_{31} &= \beta B_1 + \alpha^3 B_2 & C_{40} &= B_1 + \alpha^4 B_2 \\
 C_{20} &= P_1 + \alpha^2 P_2 & C_{13} &= \beta^3 B_1 + \alpha B_2 & C_{04} &= \beta^4 B_1 + B_2 \\
 C_{02} &= \beta^2 P_1 + P_2 & C_{22} &= \beta^2 B_1 + \alpha^2 B_2 & &
 \end{aligned}$$

où 
$$P_i = E \{X_i^2(t)\} \quad \text{et} \quad B_i = C(X_i^2)$$

### III.2.2.2: Polynôme du quatrième degré pour le cas deux sources

L'équation du polynôme du quatrième degré en fonction du paramètre  $\beta$  de la matrice de mélange est:

$$\begin{aligned}
 & (C_{40}C_{13}^2 - C_{22}^3)\beta^4 + 2C_{13}(C_{22}^2 - C_{40}C_{04})\beta^3 \\
 & + (C_{40}C_{04}^2 + C_{04}C_{22}^2 - 2C_{22}C_{13}^2)\beta^2 + C_{04}(C_{13}^2 - C_{22}C_{04}) = 0 \quad (3.12)
 \end{aligned}$$

Et pour l'obtention du coefficient  $\alpha$  on doit appliquer cette relation:

$$\alpha = \frac{C_{13} - C_{22}\beta}{C_{04} - C_{13}\beta} \quad (3.13)$$

Puisque on doit avoir quatre solutions, donc il est nécessaire d'avoir un critère de décision pour extraire les signaux sources parmi toutes les solutions estimées.

### III.2.3 Méthode basée sur les cumulants croisés (Taro Yamaguchi [28])

#### III.2.3.1: Modèle du séparateur

Pour le cas de deux sources, le séparateur proposé est donné par l'inverse de la matrice de mélange, en se basant sur les mêmes conditions de la matrice de mélange supposée par Ali mansour donc la solution de cet algorithme mène directement à l'estimation de la matrice de séparation et les signaux estimés ont la forme matricielle suivante:

$$E = a WX \quad (3.14)$$

Où 
$$a = (1 - \alpha \beta)^{-1} \quad (3.15)$$

$$W = \begin{bmatrix} 1 & -\alpha \\ -\beta & 1 \end{bmatrix} \quad (3.16)$$

et l'indépendance entre les signaux mène à ce système d'équations [28].

$$E \{S_1 S_2\} = E \{S_1\} E \{S_2\} \quad \text{et} \quad E \{S_1^3 S_2\} = E \{S_1^3\} E \{S_2\} \quad (3.17)$$

### III.2.3.2: Polynôme du quatrième degré et les expressions de ses coefficients

Le polynôme du quatrième degré en fonction du paramètre  $\alpha$  proposé par Taro Yamaguchi est le suivant:

$$\begin{aligned} & (C_2 C_{10} - C_{11} C_3) \alpha^4 + (3C_9 C_3 - 3C_8 C_2 - C_3 C_{10} + C_1 C_{11}) \alpha^3 \\ & + (3C_6 C_2 + 3C_8 C_3 - 3C_9 C_1 - 3C_7 C_3) \alpha^2 \\ & + (C_5 C_3 + 3C_7 C_1 - 3C_6 C_3 - C_2 C_4) \alpha + C_3 C_4 - C_1 C_5 = 0 \end{aligned} \quad (3.18)$$

Et les coefficients peuvent être obtenus en utilisant que l'expression pour le calcul des moments, en appliquant ces expressions pour chaque coefficient [28].

$$\begin{aligned} C_1 &= E \{X_1^2\} - E\{X_1\}^2 & C_6 &= E \{X_1^3 X_2\} - E \{X_1^2 X_2\} E\{X_1\} \\ C_2 &= E \{X_2^2\} - E\{X_2\}^2 & C_7 &= E \{X_1^2 X_2^2\} - E \{X_1^2 X_2\} E\{X_2\} \\ C_3 &= E\{X_1 X_2\} - E\{X_1\} E\{X_2\} & C_8 &= E \{X_1^2 X_2^2\} - E \{X_1^2 X_2\} E\{X_1\} \\ C_4 &= E \{X_1^4\} - E \{X_1^3\} E\{X_1\} & C_9 &= E \{X_1 X_2^3\} - E \{X_1 X_2^2\} E \{X_2^3\} \\ C_5 &= E \{X_1^3 X_2\} - E \{X_1^3\} E\{X_2\} & C_{10} &= E \{X_1 X_2^3\} - E\{X_1\} E \{X_2^3\} \\ & & C_{11} &= E \{X_2^4\} - E \{X_2^3\} E\{X_2\} \end{aligned}$$

et pour l'obtention du coefficient  $\beta$  on doit appliquer cette relation:

$$\beta = \frac{\alpha C_2 - C_3}{\alpha C_3 - C_2} \quad (3.19)$$

Cet algorithme est plus performant du point de vue de la vitesse d'exécution que celui de Ali Mansour, et c'est évident car le nombre d'opérations pour le calcul des coefficients du polynôme est inférieur par rapport au nombre d'opérations pour le calcul des coefficients du polynôme de Ali Mansour même si on ne prend pas en considération les termes répétés.

### III.3: Méthodes adaptatives (en ligne)

#### III.3.1: Principe des méthodes en ligne

Du point de vue mathématique la solution du problème de la séparation de sources consiste à trouver le séparateur de telle façon que les signaux estimés aient le degré d'indépendance le plus grand que possible, on a discuté à la section précédente que le problème peut être achevé d'une façon directe. Dans cette partie on va présenter les méthodes alternatives pour la séparation en utilisant les techniques adaptatives. Ces méthodes se basent généralement sur l'utilisation des techniques neuronales, qui sont une approximation mathématique du comportement des réseaux de neurones du cerveau humain. Comme toutes les techniques adaptatives les réseaux de neurones artificiels ont aussi des lois d'adaptations afin d'adapter les paramètres du réseau pour avoir la meilleure approximation avec l'erreur la plus petite. Les réseaux de neurone artificiels adaptent les paramètres par ce qu'on appelle les fonctions d'activations qui ont pour but de contrôler le seuil de la décision pour que chaque cellule neuronale soit active ou non. Ces fonctions doivent être non linéaires et impaires [20] [21], leur utilisation dépend du type du réseau utilisé dans l'application. Le schéma de principe pour l'utilisation des méthodes neuronales pour la séparation de source est le suivant:

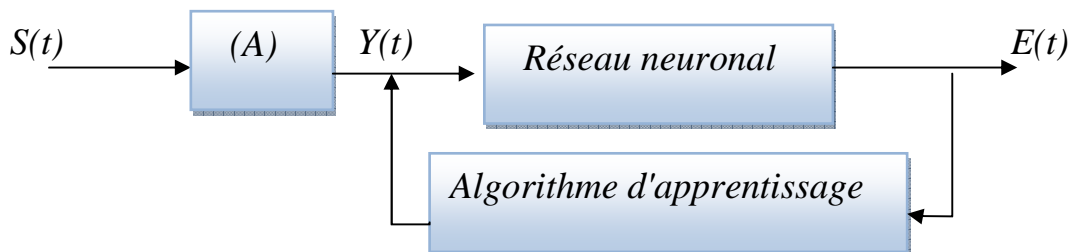


Figure III.4: Principe des méthodes en ligne pour la SAS

Pour cette partie on présentera trois algorithmes pour la séparation de sources par les techniques neuronales, où on débutera par l'algorithme de (Herault et Jutten) puis deux extensions de cet algorithme qui sont l'algorithme de (A-Cichocki et R-Unbehauen) et l'algorithme (A-Cichoki et S-Amari).



### III.3.2: Approche neuronale de (Herault et Jutten [26][15])

Les réseaux de (Herault et Jutten) se présentent en deux formes, réseau direct et réseau récurrent, et qui sont deux techniques qui mènent à une solution pour la séparation de sources. On débute par le réseau direct ensuite le réseau récurrent en donnant quelques remarques sur ces deux techniques.

#### III.3.2.1: Réseau direct

Le schéma de principe du réseau direct pour le cas général est sur la figure suivante:

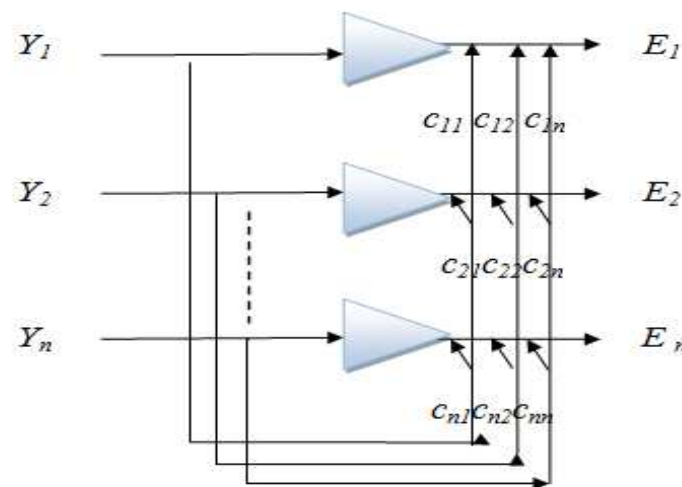


Figure III.5: Réseau neuronal direct de Herault et Jutten

La sortie du réseau est donnée sous la forme matricielle suivante:

$$E = CY \tag{3.20}$$

Où (C) est la matrice des poids du réseau où les éléments de la diagonale sont tous nuls. La sortie du réseau en fonction de la matrice de mélange a l'expression suivante:

$$E = CAS \tag{3.21}$$

Donc les signaux des sorties en fonction des signaux mélanges sont donnés par:

$$E(t) = Y(t) - CY(t) \tag{3.22}$$

Et l'expression des signaux de sorties en fonction de la matrice de mélange et les signaux sources est donnée par:

$$E(t) = (I - C) AS(t) \tag{3.23}$$

### III.3.2.1.1: Solutions pour le cas de deux sources

Pour le cas de deux sources le réseau est le suivant [26] [15].

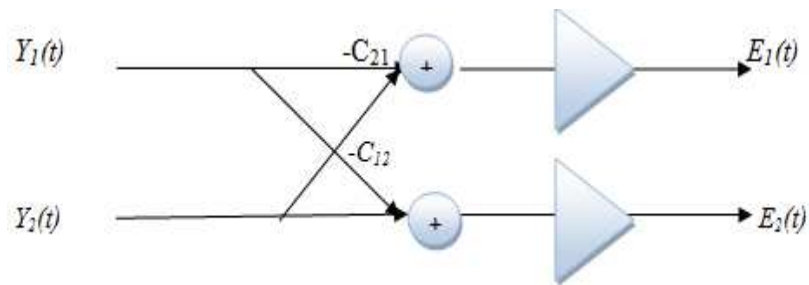


Figure III.6: Réseau neuronal direct pour deux sources de Herault et Jutten

En se référant à la figure, les expressions des signaux de sortie sont les suivantes:

$$E_1(t) = (a_{11}-c_{12}a_{21}) S_1 + (a_{12}-c_{12}a_{22}) S_2 \quad (3.24)$$

$$E_2(t) = (a_{21}-c_{21}a_{11}) S_1 + (a_{22}-c_{21}a_{12}) S_2 \quad (3.25)$$

Si  $C_{12} = a_{12}/a_{22}$  et  $C_{21} = a_{21}/a_{11}$  les sorties du réseau sont:

$$E_1(t) = \alpha S_1 \quad \text{et} \quad E_2(t) = \beta S_2$$

Où  $\alpha$  et  $\beta$  sont des constantes réelles

si  $C_{12} = a_{11}/a_{21}$  et  $C_{21} = a_{22}/a_{12}$  les sorties du réseau sont:

$$E_1(t) = \alpha S_2 \quad \text{et} \quad E_2(t) = \beta S_1$$

### III.3.2.1.2: Loi d'adaptation

Donc pour deux sources on a deux solutions, alors le but de ce réseau adaptatif est d'adapter les poids du réseau pour avoir l'une de ces deux solutions, et la loi d'adaptation des poids du réseau directe pour le cas discret est la suivante [26] [15].

$$C_{ij}(k+1) = C_{ij}(k) - \alpha f(Y_i(k))g(Y_j(k)) \quad (3.26)$$

\*f et g: les fonctions d'activations du réseau.

\* $\alpha$  : le gain d'apprentissage.

III.3.2.2: Réseau récurrent

Pour le réseau récurrent les poids du réseau sont réinjectés dans les cellules neuronales et le schéma de principe pour le réseau récurrent est le suivante:

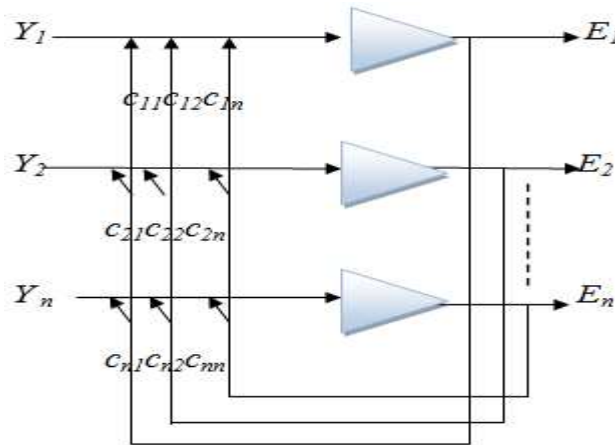


Figure III.7: Réseau neuronal récurrent de Herault et Jutten

L'équation du réseau est la suivante:

$$E(t) = (I+C)^{-1} Y(t) \tag{3.27}$$

Et l'équation du réseau en fonction des signaux sources est la suivante:

$$E(t) = (I+C)^{-1} A S(t) \tag{3.28}$$

III.3.2.2.1: Solutions pour le cas de deux sources

Pour le cas de deux sources le réseau a la forme suivante:

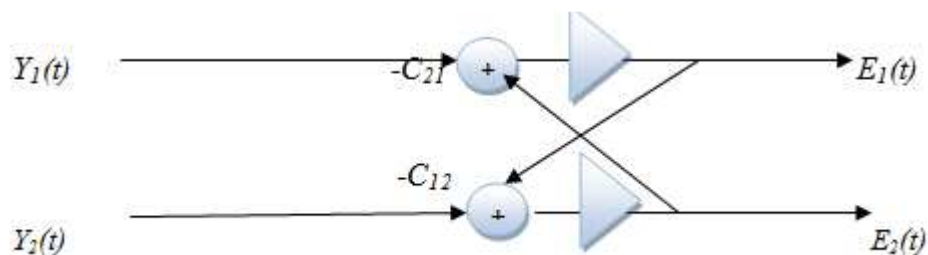


Figure III.8: Réseau neuronal récurrent pour deux sources de Herault et Jutten

Les sorties du réseau en fonction des paramètres de la matrice de mélange et les signaux sources sont traduites par les expressions suivantes:

$$E_1(t) = ((a_{11}-C_{12} a_{21})S_1 + (a_{12}-C_{12}a_{22})S_2) / (1-C_{12}C_{21}) \quad (3.29)$$

$$E_2(t) = ((a_{21}-C_{21} a_{11})S_1 + (a_{22}-C_{21}a_{12})S_2) / (1-C_{12}C_{21}) \quad (3.30)$$

Si  $C_{12} = a_{12}/a_{22}$  et  $C_{21} = a_{21}/a_{11}$  les sorties du réseau sont:

$$E_1(t) = a_{11} S_1 \quad \text{et} \quad E_2(t) = a_{22} S_2$$

Si  $C_{12} = a_{11}/a_{21}$  et  $C_{21} = a_{22}/a_{12}$  les sorties du réseau sont :

$$E_1(t) = a_{12} S_2 \quad \text{et} \quad E_2(t) = a_{21} S_1$$

Pour ce réseau on a aussi deux solutions, et qu'on doit atteindre l'une d'elles, et la loi d'adaptation utilisée pour ce réseau est la même loi utilisée pour le réseau directe.

### III.3.3: Algorithme de (A-Cichocki et R -Unbehauen [20] [24])

#### III.3.3.1: Problèmes des réseaux de (Herault et Jutten)

Les problèmes pour les réseaux proposés par Herault et Jutten sont:

- 1- ils divergent si la matrice A de mélange est mal conditionnée.
- 2- divergent aussi ces les signaux ont des amplitudes très petites (les variables aléatoire ont un domaine très étroit autour de leurs moyennes).
- 3- un autre problème qui mène à la divergence de l'algorithme est le rapport d'amplitudes, si par exemple le rapport de la longueur du domaine le plus grand des variables aléatoires des signaux sources sur la longueur du domaine le plus petit est grand, alors l'algorithme diverge pour ce problème on dit aussi qu'on a le problème des échelles mal conditionnées. Donc la convergence de l'algorithme dépend de plusieurs facteurs. Les algorithmes de (A-Cichocki et R Unbehauen) et (A-Cichoki et S-Amari) traitent ces problème et ils sont plus stables pour la convergence vers la solution.

#### III.3.3.2: Principe de l'algorithme de (A-Cichocki et R- Unbehauen[20] [24])

L'algorithme de (A-Cichocki et R Unbehauen) est une extension de l'algorithme récurrent de Herault et Jutten, mais plus performant en vitesse et en résultats. et puisque il est aussi un réseau neuronal, ce réseau est composé d'une seule couche comme le réseau de (Herault et Jutten), mais le nombre des poids du réseaux est différents. La sortie du réseau est donnée par cette équation:

$$E(t) = (I+W)^{-1}Y(t) \tag{3.31}$$

I : c'est une matrice diagonale et W : la matrice des poids

Et le schéma de principe de l'algorithme est le suivant:

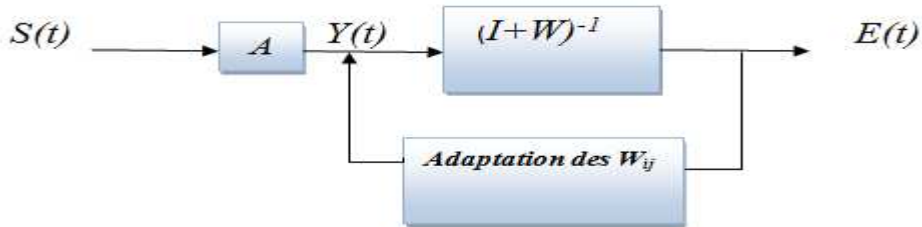


Figure III.9: Schéma de principe du réseau de (A-Cichocki et R Unbehauen)

### III.3.3.3: Loi d'adaptation

La loi d'adaptations des poids du réseau est donnée pour le cas discret comme suit:

$$W(k+1) = W(k) - \alpha(k)[I+W] \times [I - f(E(t)) \times g(E(t)^T)] \tag{3.32}$$

\*  $f(\cdot)$  et  $g(\cdot)$  : Les fonctions d'activations du réseau.

\*  $\alpha$  : le gain d'apprentissage.

Donc le schéma complet de l'algorithme est le suivant:

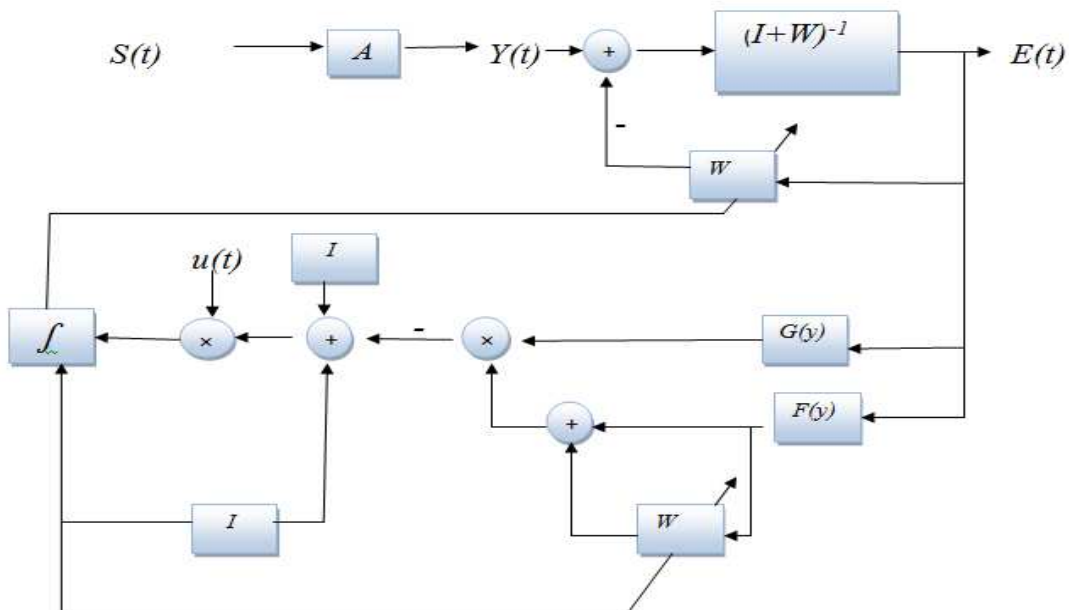


Figure III.10: Réseau neuronal pour la SAS de (A-Cichocki et R Unbehauen)

III.3.4:Algorithme de (A-Cichoki et S-Amari [21] [25])

III.3.4.1: Principe de l'algorithme de (A-Cichoki et S-Amari)

Cet algorithme est plus rapide et plus performant que celui de (A-Cichocki) et (Runbehauen), il a aussi l'avantage de ne pas calculer la matrice inverse comme le cas du réseau discuté précédemment. la motivation principale de cet algorithme est d'améliorer les résultats de séparation et en moins de temps, il corrige aussi le problème de la matrice mal conditionnée et donne de bons résultats même si on a la présence du problème des échelles mal conditionnées. La sortie du réseau se comporte comme les filtres prédictifs, et elle a l'expression suivante:

$$E(t)=Y(t)-WY(t- \tau) \tag{3.33}$$

Et le schéma de principe de l'algorithme est le suivant:

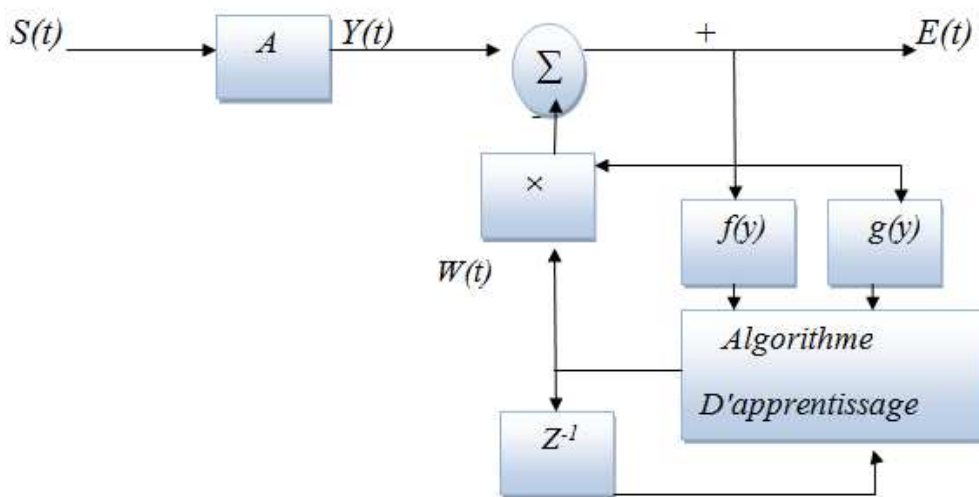


Figure III.11:Réseau neuronal pour la SAS de (A-Cichocki) et (S-Amari)

III.3.4.2: Loi d'adaptation

la loi d'adaptation des poids proposée par (A-Cichoki et S-Amari) a l'expression suivante:

$$W(k+1)=W(k)- \alpha[W+(W \times I)-Z(t).g^T(E(t))] \tag{3.34}$$

\*  $Z(t)=f(E(t)+W \times (E(t)))$

\*  $f(.)$  et  $g(.)$  les fonctions d'activations du réseau.

\* $\alpha$  : le gain d'apprentissage.

Tout comme le réseau de (A-Cichocki et RUnbehauen), les fonctions d'activation du réseau sont des fonctions impaires et non linéaires. Ces fonctions d'activations jouent un rôle très important pour la séparation, car elles distinguent les moments d'ordre supérieur et assurent l'indépendance mutuelle des signaux estimés [21] [25].

# ***Chapitre IV***

***Implémentation des algorithmes***

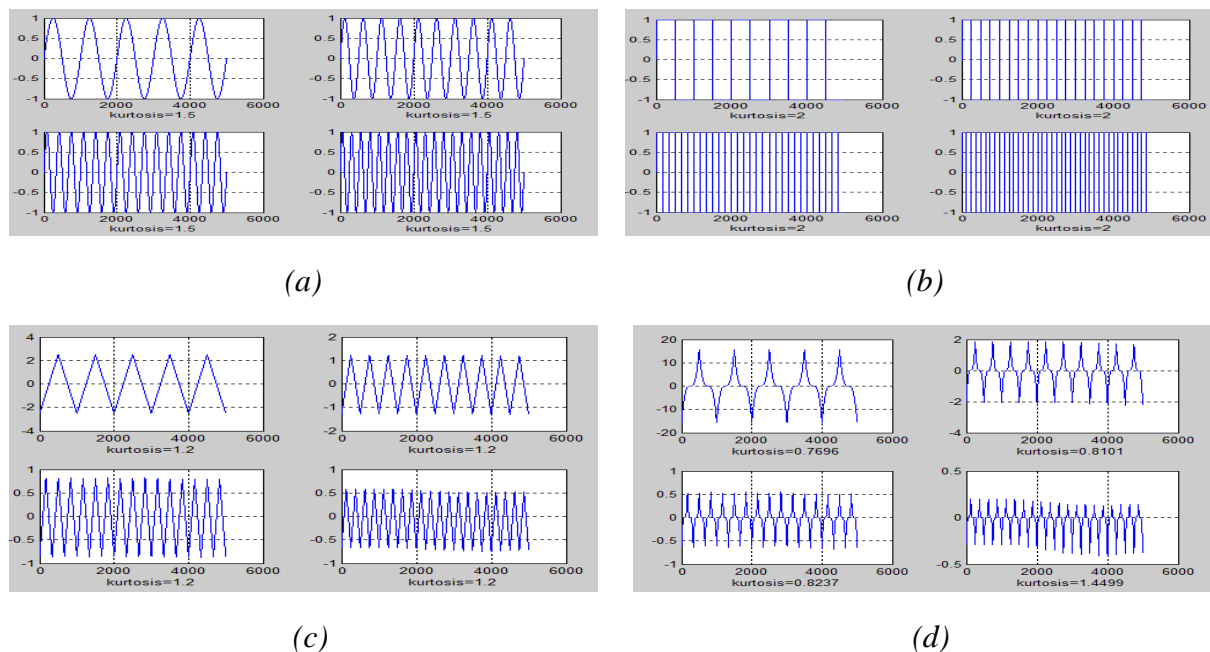
***et***

***comparaison des résultats***



## IV. 1 : Introduction

Après avoir donné une idée générale sur les théories des probabilités au premier chapitre, et donner les concepts généraux sur le problème de la séparation de sources et de sa résolution, ce chapitre est consacré à l'application des techniques discutées dans le troisième chapitre, et comme on a mentionné en deuxième chapitre, les solutions ne sont pas toutes des solutions générales pour le problème de la SAS, chaque solution est fiable selon le type du système de mélange et la nature des signaux afin d'extraire les signaux sources. L'application de ces techniques est faite dans notre travail par simulation numérique en utilisant le langage de programmation (Matlab). Et pour donner une vision claire sur le problème de la SAS et comment on le résout, on a utilisé des signaux de différents types (sinusoïdaux, carrés, triangulaires) et avec différentes fréquences, et on a ajouté un autre signal qui est le cube du signal triangulaire, et les signaux sources sont présentés dans la figure ci dessous et avec le calcul de leurs Kurtosis en valeur absolue.



**Figure IV.1:** Différents signaux utilisés pour la simulation

(a) signaux sinusoïdaux, (b) signaux carrés, (c) signaux triangulaires, (d) le cube du signal triangulaire.

**Remarque:** la fréquence d'échantillonnage utilisée pour l'implémentation, égale à 1000 HZ, et le décalage du signal complexe pour la dernière fréquence est causé par l'approximation du signal à partir du signal triangulaire, mais ce n'est pas un problème puisque dans nos applications le but est d'avoir le Kurtosis non nul et c'est le cas pour ce signal.

**I.V.2: Hypothèse sur le système et les signaux sources et le critère de comparaison**

Afin d'accomplir le but de ce travail, on suppose tout d'abord que le système de mélange est linéaire instantané, et puisqu'on parle d'une séparation aveugle il faut se baser sur des hypothèses concernant l'application des algorithmes pour le but de comparer les résultats trouvés avec les signaux originaux. Pour la comparaison des résultats ou pour la décision finale des résultats on a appliqué le concepts du Kurtosis discuté au premier et au deuxième chapitre, mais en ce qui concerne le calcul du Kurtosis pour la comparaison des différentes méthodes on a calculé la moyenne des Kurtosis en valeurs absolues de tous les signaux, car les signaux utilisés ont le Kurtosis non nul et positif, si par exemple on a (n) signal, les valeurs des Kurtosis pour chaque signal sont:  $K_1, K_2, \dots, K_n$ , alors la valeur moyenne du kurtosis de tous les signaux est:

$$K = \frac{\sum_{i=1}^n |K_i|}{n} \quad (4.1)$$

Et la comparaison est faite en calculant le pourcentage d'erreur du Kurtosis des signaux sources utilisés dans l'application par rapport au Kurtosis des signaux estimés après l'application de chaque algorithme par l'application de cette relation:

$$\text{Pourcentage d'erreur} = \left( \frac{|\Delta K|}{K} \times 100 \right) \% \quad (4.2)$$

Et pour les hypothèses de l'application des algorithmes, elles sont les suivantes:

**Hypothèse 1:** Le nombre de sources égale au nombre d'observations.

**Hypothèse 2:** Les sources sont statistiquement indépendantes.

**Hypothèse 3:** Le bruit est considéré nul.

**Hypothèse 4:** Les signaux sont considérés ergodiques.

**Hypothèse 5:** Le système de mélange est stable.

On a discuté dans le deuxième chapitre que, si la séparation est atteinte, les signaux estimés sont des versions des signaux sources multipliés par des facteurs, pour cela dans ce travail, la représentation de chaque signal estimé ( $E_i$ ) est assignée par le signal source approprié ( $S_j$ ) multiplié par une constante réelle positives  $\alpha_i$ .

### IV.3: Implémentation des Méthodes directes

Du même ordre suivi pour le troisième chapitre on commence par l'application des méthodes directes.

#### IV.3.1: Méthode géométrique (Ali Mansour [22])

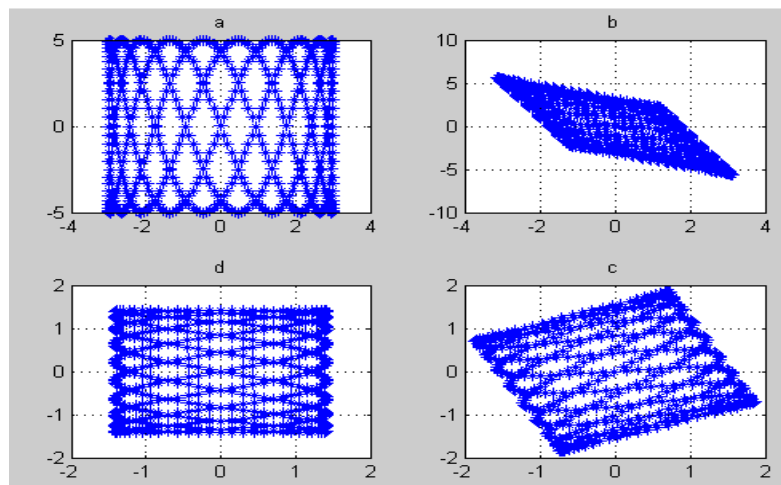
On a discuté au troisième chapitre que pour les techniques géométriques l'angle de rotation de la distribution des deux signaux orthogonaux peut être trouvé par la détermination du point le plus loin, si les fonctions de densité de probabilité des deux sources sont uniformes, ou les signaux sources sont du type sinusoïdal, en appliquant la relation (3.6), et on commence par deux sources sinusoïdales.

##### IV.3.1.1: Deux sources (directe)

Pour cette application la matrice de mélange utilisée est:

$$A = \begin{bmatrix} 0.3188 & -0.4336 \\ -1.3077 & 0.3426 \end{bmatrix}$$

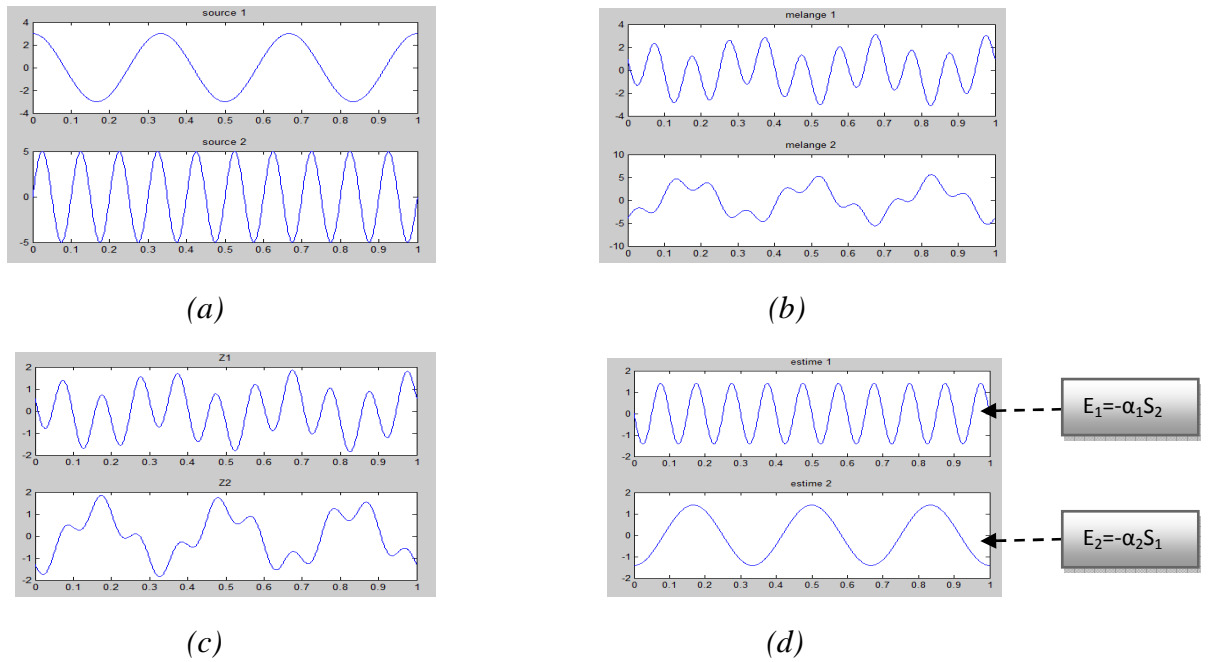
Pour des signaux sinusoïdaux les étapes et les résultats sont sur les figures suivantes:



**Figure IV.2:** Distributions des amplitudes (signaux sinusoïdaux).

*Application de la méthode géométrique directe d'Ali Mansour (directe) sur des signaux sinusoïdaux cas de deux sources, (a): distribution des amplitudes des signaux sources, (b): distribution des amplitudes des signaux mélanges, (c): distribution des amplitudes des signaux non corrélés, (d): distribution des amplitudes des signaux estimés.*

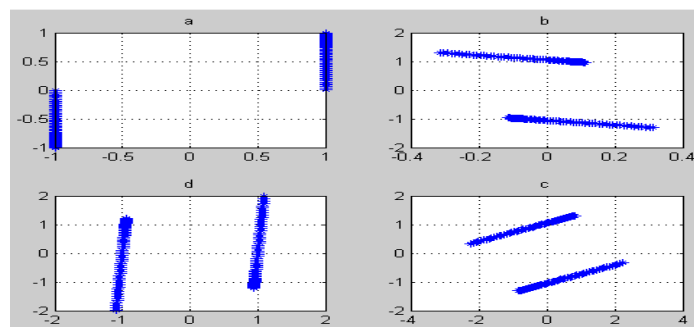
On remarque à partir de la figure qu'on a pu ajuster la distribution des signaux non corrélés, mais avec une rotation de  $\Pi/2(2n+1)$ , ce qui engendre une permutation des positions. Et le résultat de la séparation est le suivant:



**Figure IV.3:** Résultats de l'application (signaux sinusoidaux)

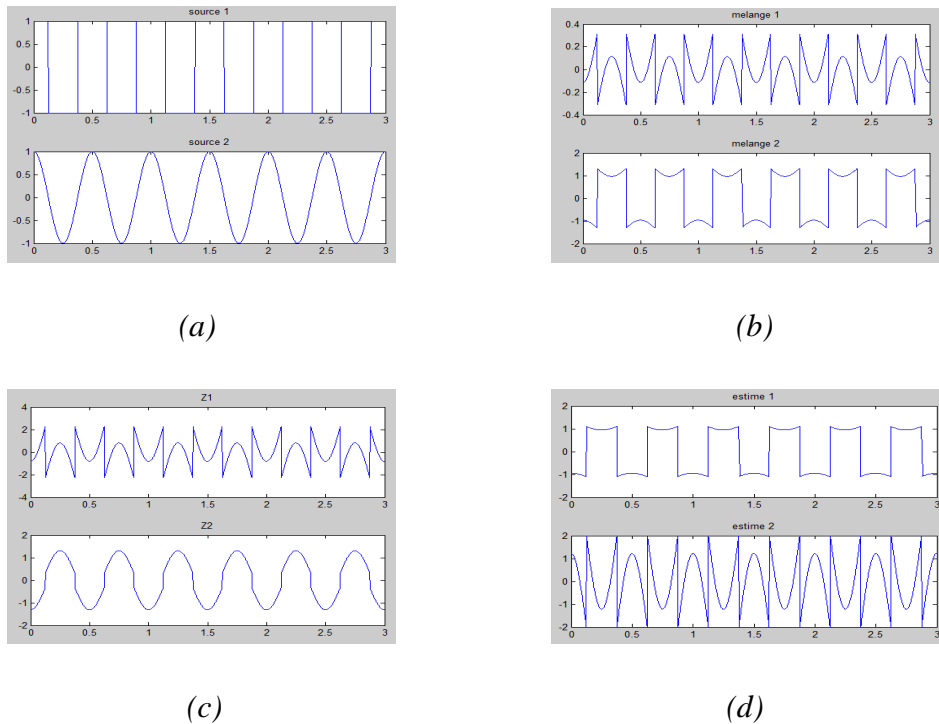
Résultat de séparation des signaux sinusoidaux (de deux sources), (a): signaux sources, (b): signaux mélanges, (c): signaux non corrélés, (d): signaux estimés, ( $\alpha_i > 0$ )

Pour ce cas on n'a pas besoin d'avoir un critère pour atteindre la séparation, mais si on applique deux signaux sources qui ne sont pas tous des signaux sinusoidaux, on n'aura pas des résultats satisfaisants. Exemple pour le cas d'un signal carré avec un autre sinusoidal, les résultats sont les suivants:



**Figure IV.4:** Distributions des amplitudes (signal sinusoidal et signal carré)

Application de la méthode géométrique directe de Ali Mansour sur le mélange d'un signal sinusoidal et un signal carré, (a) distribution des amplitudes des signaux sources, (b) distribution des amplitudes des signaux mélanges, (c) distribution des amplitudes des signaux non corrélés, (d) distribution des amplitudes des signaux estimés.



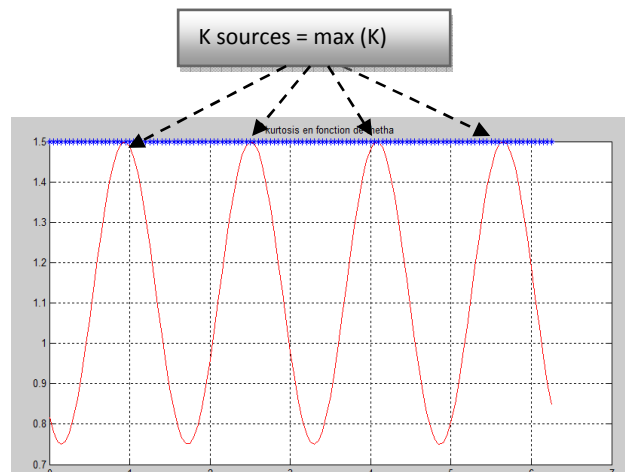
**Figure IV.5:** Résultats de l'application (signal sinusoidal et signal carré)

*Application de la méthode géométrique directe de Ali Mansour pour le mélange d'un signal sinusoidal et un signal carré, (a): signaux sources, (b): signaux mélanges, (c): signaux non corrélés, (d): signaux estimés.*

On remarque que les signaux estimés ne sont pas une version des sources originales, en utilisant cette technique de rotation.

#### IV.3.1.2: Influence de l'angle de rotation sur la valeur du Kurtosis

Il faut savoir le comportement de la valeur du Kurtosis finale des signaux estimés en fonction de l'angle de rotation appliqué sur les signaux non corrélés, et pour cela on a fait plusieurs rotations des signaux blanchis dans un intervalle de  $[0, 2\pi]$ , et pour avoir une idée sur la maximisation du Kurtosis, on a commencé par deux sources sinusoidales ensuite on a appliqué cette procédure sur d'autres signaux qui ne sont pas sinusoidaux, et les résultats sont les suivants pour deux signaux sources sinusoidaux :



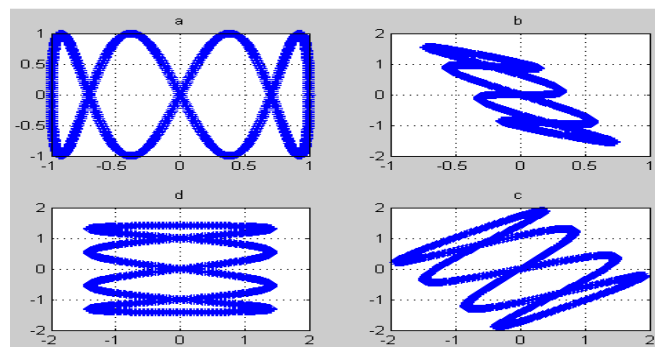
**Figure IV.6:** Influence de  $(\alpha)$  sur  $(K)$  (signaux sinusoïdaux)

(a): le graphe continu montre la valeur du Kurtosis des signaux estimés en fonction de l'angle de rotation appliqué sur les signaux non corrélés et qui est pris dans l'intervalle de  $[0, 2\pi]$ ,

(b): le graphe discontinu le kurtosis des signaux sources.

#### IV.3.1.3: Deux sources (max Kurtosis)

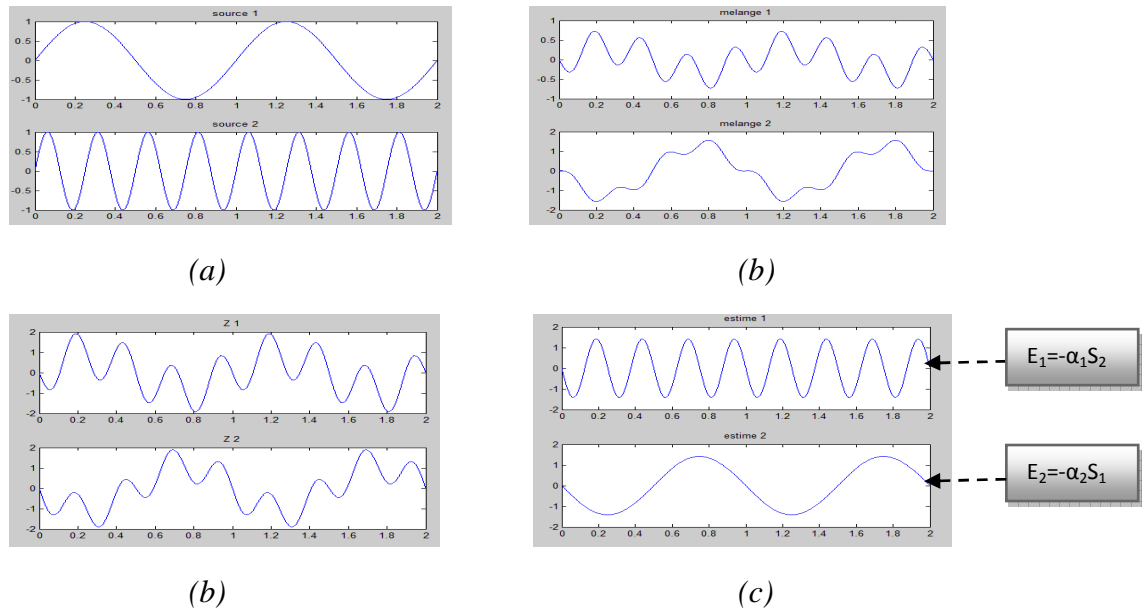
à partir de la figure précédente, on remarque qu'on peut avoir les signaux sources si on choisit la valeur du Kurtosis maximale. et on remarque aussi qu'on a quatre valeurs maximales alors on doit choisir qu'une, et le résultat pour des signaux sinusoïdaux est le suivant:



**Figure IV.7:** Distributions des amplitudes (signaux sinusoïdaux)

Application de la méthode géométrique de Ali Mansour sur des signaux sinusoïdaux par la détermination de l'angle de rotation en se basant sur l'angle qui mène au Kurtosis maximal pour le cas de deux sources sinusoïdales, (a): distribution des amplitudes des signaux sources, (b): distribution des amplitudes des signaux mélanges, (c): distribution des amplitudes des signaux non corrélés, (d): distribution des amplitudes des signaux estimés.

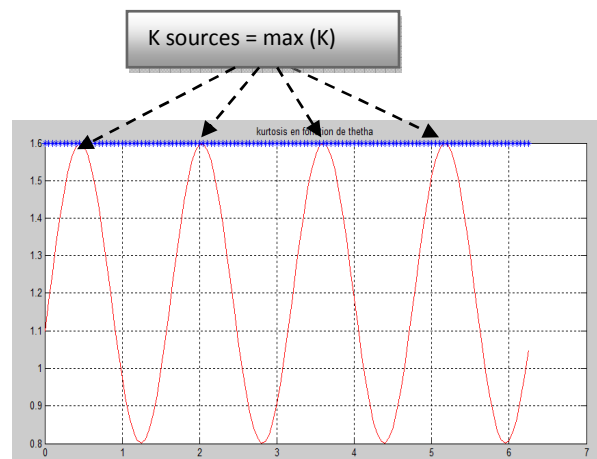
On remarque la même chose, la distribution finale est une version de la distribution originale avec une rotation de  $\pi/2(2n+1)$ , et le résultat de la séparation est le suivant:



**Figure IV.8:** Résultats de l'application (deux signaux sinusoidaux)

*Application de la méthode géométrique de Ali Mansour sur des signaux sinusoidaux en se basant sur l'angle qui mène au Kurtosis maximal (a): signaux sources, (b): signaux mélanges, (c): signaux non corrélés, (d): signaux estimés ( $\alpha_i > 0$ )*

Revenons maintenant dans le cas où les sources ne sont pas sinusoidales, les resultats sont les suivants où la matrice de mélange A est la même, et les signaux sources sont du type carré et du type triangulaire. l'influence de l'angle de rotation sur la valeur du kurtosis est la suivante:



**Figure IV.9:** Influence de  $(\alpha)$  sur  $(K)$  (signaux non sinusoidaux)

*(a) le graphe continu montre la valeur du Kurtosis des signaux estimés en fonction de l'angle de rotation appliqué sur les signaux non corrélés et qui est pris dans l'intervalle de  $[0 2\pi]$ , (b) le graphe discontinu représente le kurtosis des signaux sources.*

Si on choisit une des quatre valeurs maximales du Kurtosis les résultats de la séparation sont les suivants:

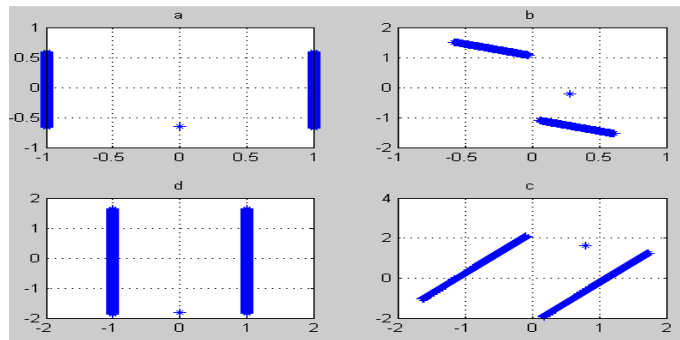


Figure IV.10: Distributions des amplitudes (signaux non sinusoïdaux)

Application de la méthode géométrique de Ali Mansour sur le mélange d'un signal triangulaire et un signal carré par la détermination de l'angle de rotation qui mène au Kurtosis maximal , (a) distribution des amplitudes des signaux sources, (b) distribution des amplitudes des signaux mélanges, (c) distribution des amplitudes des signaux non corrélés, (d) distribution des amplitudes des signaux estimés.

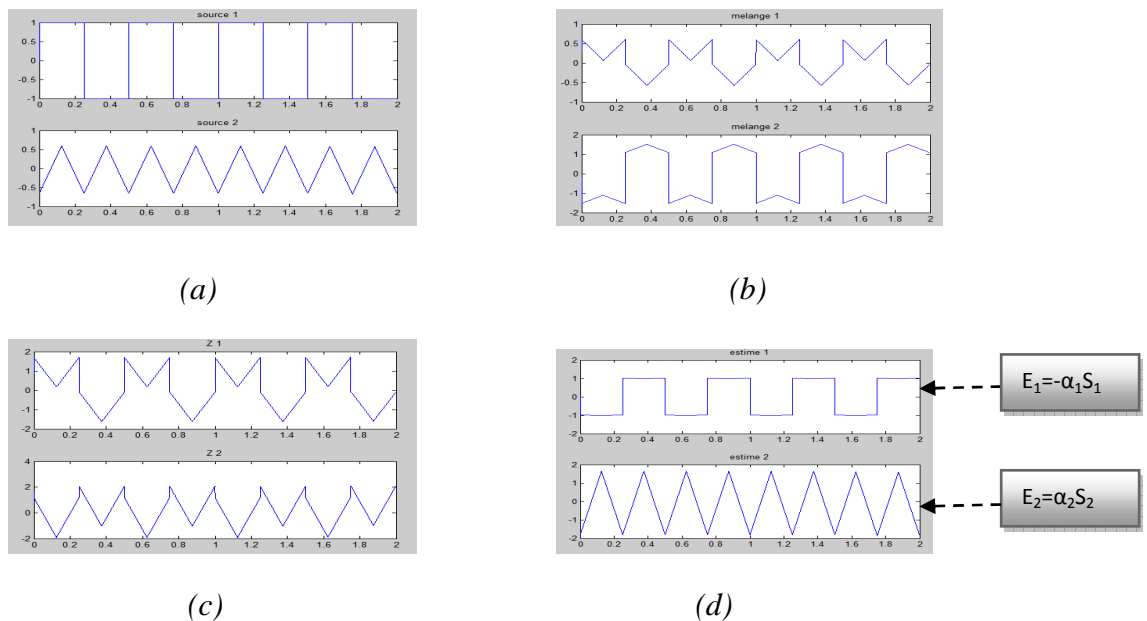


Figure IV.11: Résultats de l'application (deux signaux non sinusoïdaux)

Application de la méthode géométrique de Ali Mansour sur le mélange d'un signal triangulaire et un signal carré en se basant sur l'angle qui mène au Kurtosis maximal (a): signaux sources, (b): signaux mélanges, (c): signaux non corrélés, (d): signaux estimés ( $\alpha_i > 0$ )



IV.3.1.4: Trois sources (max Kurtosis)

Après avoir transformé les observations en des signaux orthogonaux la rotation doit être faite pour les trois axes en appliquant les relations ((3.7), (3.8), et (3.9)) du chapitre3, est les résultats son les suivants pour le mélange de trois sources du type différents et la matrice du mélange utilisée est:

$$A = \begin{bmatrix} 0.5377 & 0.8622 & -0.4336 \\ 1.8339 & 0.3188 & 0.3426 \\ -2.2588 & -1.3077 & 3.5784 \end{bmatrix}$$

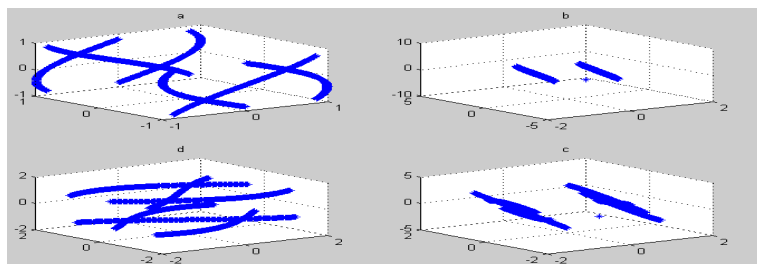


Figure IV.12: Distributions des amplitudes

Application de la méthode géométrique d'Ali Mansour sur le mélange trois signaux du type différent par la détermination de l'angle de rotation qui mènent au Kurtosis maximal, (a):distribution des amplitudes des signaux sources, (b):distribution des amplitudes des signaux mélanges, (c) : distribution des amplitudes des signaux non corrélés, (d): distribution des amplitudes des signaux estimés.

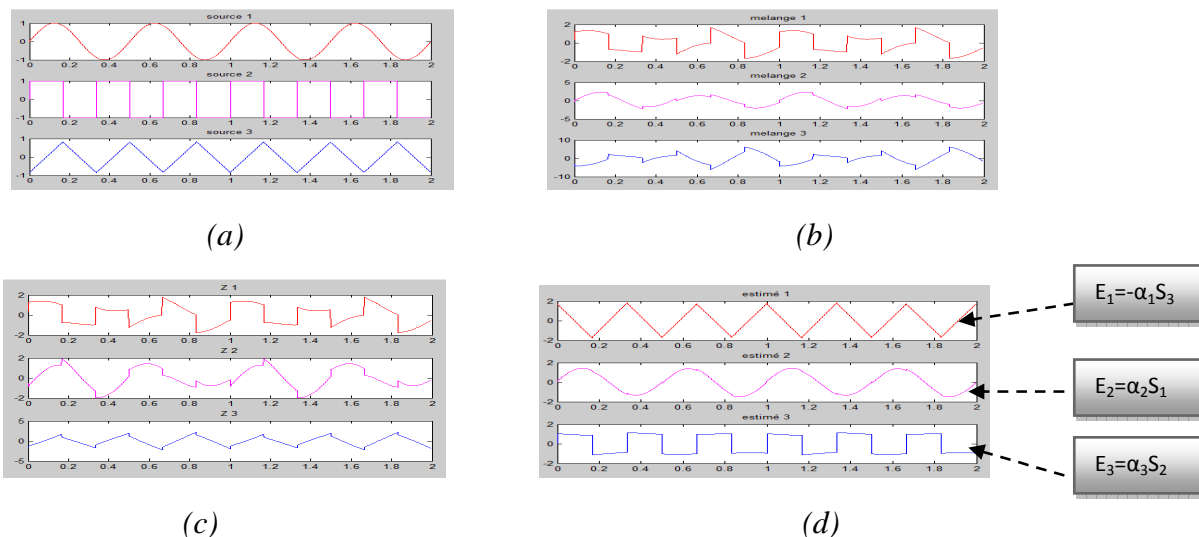


Figure IV.13: Résultats de l'application

Application de la méthode géométrique de Ali Mansour sur le mélange de trois signaux différents en se basant sur les angles qui mènent au Kurtosis maximal (a):signaux sources, (b): signaux mélanges, (c): signaux non corrélés, (d) signaux estimés ( $\alpha_i > 0$ )

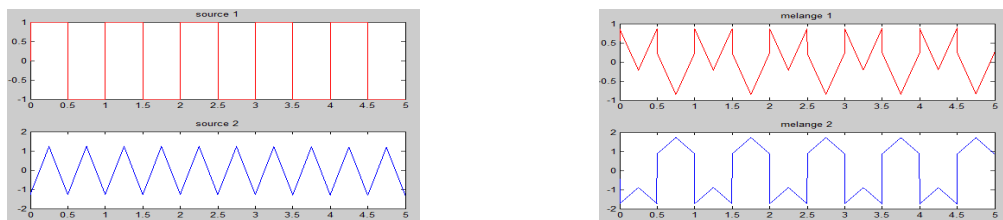
Le problème pour cette technique est que le temps d'exécution est très grand, même si la recherche de l'angle de rotation pour avoir le Kurtosis maximal a été prise dans l'intervalle de  $[0 \ \Pi/2]$  pour chaque rotation. Mais les résultats sont satisfaisants et on a pu atteindre la séparation.

**IV.3.2: Méthode basée sur les cumulants croisés (Ali Mansour [19])**

**IV.3.2.1: Résultats de l'application de l'algorithme**

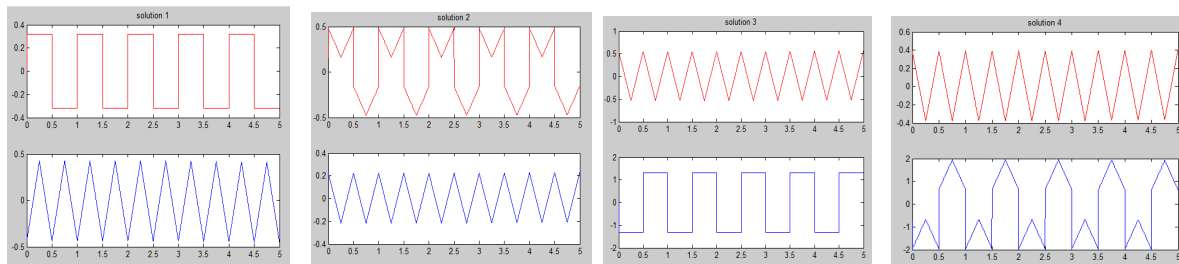
La matrice de mélange: 
$$A = \begin{bmatrix} 0.3188 & -0.4336 \\ -1.3077 & 0.3426 \end{bmatrix}$$

Il a été mentionné au chapitre 3 que la séparation est atteinte par la résolution du polynôme du quatrième de degré de l'équation (3.12) et les résultats sont les suivants:



(a) les signaux sources

(b) les signaux mélanges



(Solution 1)

(Solution 2)

(Solution 3)

(Solution 4)

**Figure IV.14:** Résultats de l'application (les 4 solutions)

*Résultats de l'application de l'algorithme basé sur les Cumulants d'Ali Mansour sur le mélange d'un signal triangulaire et un signal carré.*

Et les coefficients du modèle (3.10) pour chaque solution sont:

Les valeurs de  $\alpha$  : [-1.2598 1.7475 -0.2433 -0.2445]

Les valeurs de  $\beta$  : [-4.1108 -4.0905 -0.7938 0.5722]

Les valeurs des Kurtosis des signaux sources et les signaux mélanges sont

Kurtosis des sources= 1.5994, Kurtosis des mélanges= 1.3360

Les valeurs des Kurtosis pour chaque solution : [1.5994 1.4569 1.5994 1.4569]

IV.3.2.2:Décision

Donc à partir des solutions du polynôme, on doit choisir la valeur du Kurtosis maximale pour la décision, par exemple dans notre cas la première solution et la troisième, et voici le résultat en appliquant cette décision sur le même model de mélange mais avec d'autres signaux.

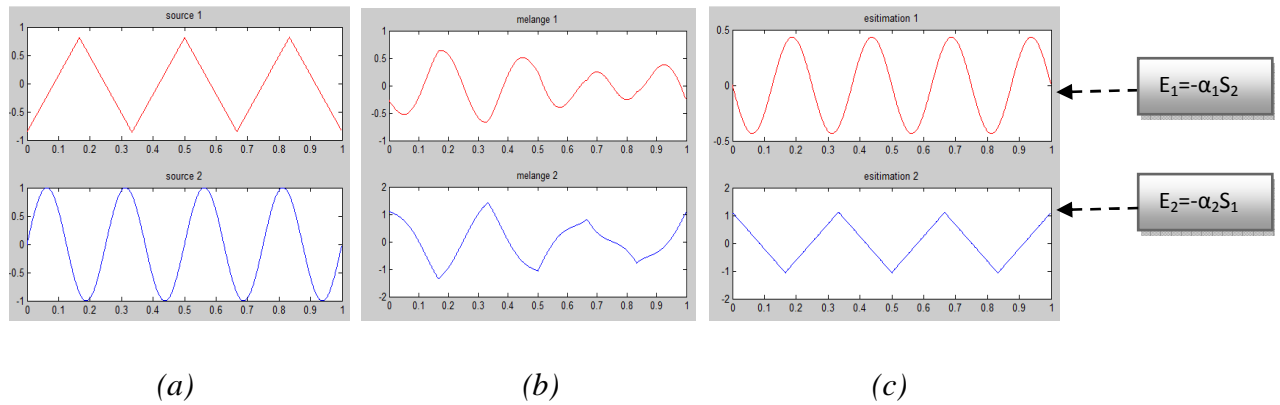


Figure IV.15: Résultats de l'application (décision)

Résultats de l'application de l'algorithme basé sur les Cumulants de Ali Mansour sur le mélange d'un signal triangulaire et un signal sinusoïdal, (a): signaux sources, (b):signaux mélanges,(c): signaux estimés après décision( $\alpha_i > 0$ )

Si on applique cette technique sur différentes matrices prises aléatoirement, les valeurs des Kurtosis des résultats comparés avec ceux des signaux sources sont ci-dessous:

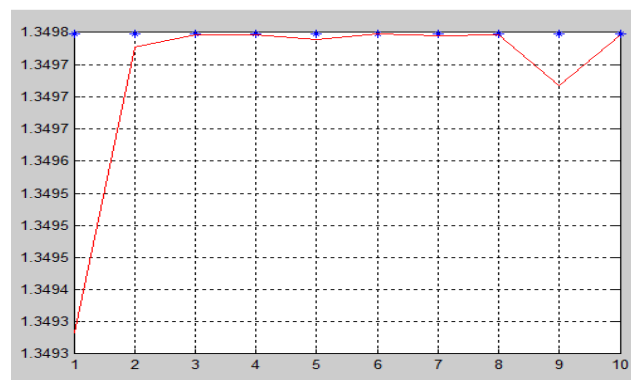


Figure IV.16: Valeur finale du Kurtosis (décision)

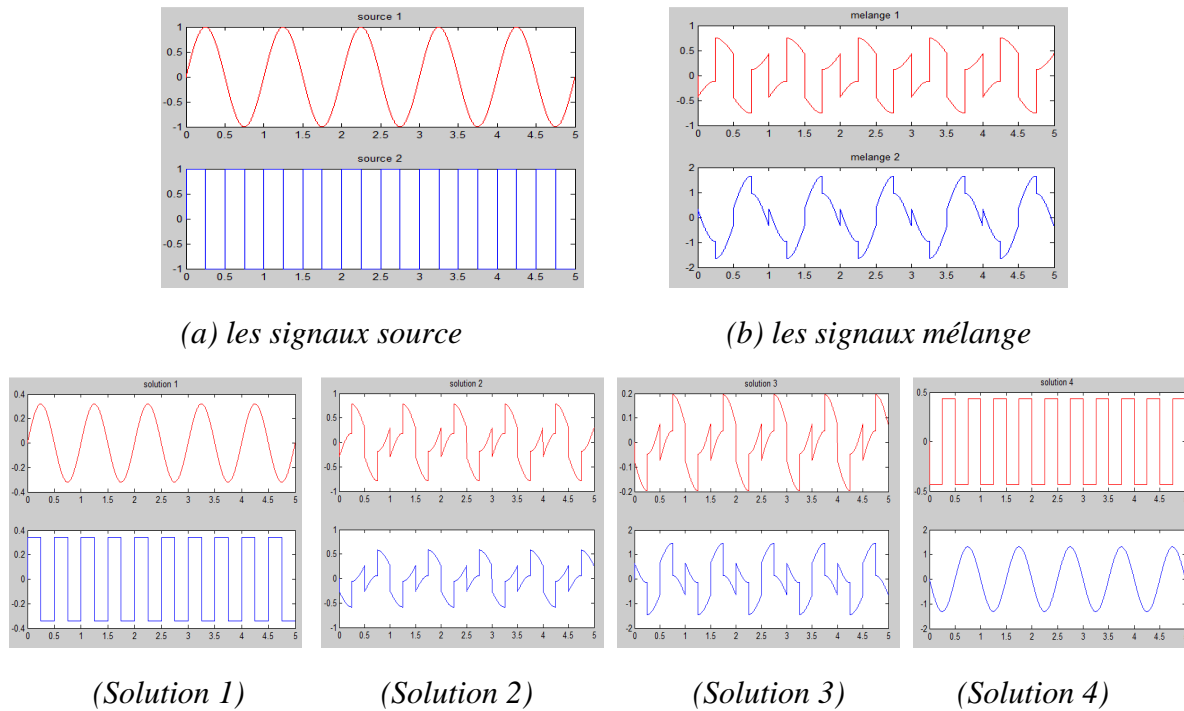
Graphe continu représente les valeurs du Kurtosis des estimés pour différentes matrices de mélange, le graphe discontinu en bleu la valeur du Kurtosis des signaux sources

On remarque que les valeurs des Kurtosis des signaux estimés ont presque la même valeur de Kurtosis des signaux sources, pour n'importe quelle matrice non singulière.

### IV.3.3: Méthode basée sur les cumulants croisés (Taro Yamaguchi [28])

#### IV.3.3.1: Résultats de l'application de l'algorithme

De la même façon pour l'application de l'algorithme de Ali Mansour, on applique l'algorithme de Taro Yamaguchi au polynôme de l'équation (3.18) du chapitre 3, les solutions pour la matrice (Page.73) sont les suivantes:



**Figure IV.17:** Résultats de l'application (les 4 solutions)

*Résultats de l'application de l'algorithme basé sur les Cumulants de Taro Yamaguchi sur le mélange d'un signal sinusoïdal et un signal carré*

Et les coefficients du séparateur proposé en (3.16) et (3.15) sont.

Les valeurs de  $\alpha$  : [-1.2646 0.5213 -0.5498 -0.2438]

Les valeurs de  $\beta$  : [-4.0971 -2.0328 4.1788 -0.7921]

Les valeurs des Kurtosis des signaux sources et signaux mélanges:

Kurtosis des sources= 1.749 , Kurtosis des mélanges= 1.2482

Les valeurs des Kurtosis pour chaque solution : [1.7497 0.8925 0.8924 1.7497]

IV.3.3.2:Décision

De la même manière pour la décision de l'algorithme précédent, on doit choisir la valeur du Kurtosis maximal pour la décision par exemple dans notre cas la première solution et la quatrième, et voici le résultat en appliquant cette décision sur le même model de mélange mais avec des signaux différents.

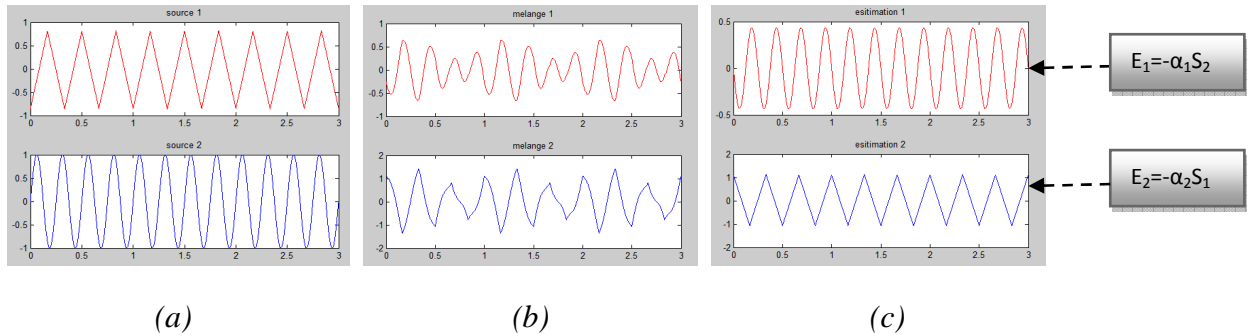


Figure IV.18: Résultats de l'application (décision)

Résultats de l'application de l'algorithme basé sur les Cumulants de Taro Yamaguchi sur le mélange d'un signal triangulaire et un signal sinusoidal, (a) signaux sources, (b) signaux mélanges, (c) signaux estimés après décision ( $\alpha_i > 0$ )

La comparaison des valeurs des Kurtosis des estimés avec ceux des signaux sources pour différentes matrices de mélange prises aléatoirement est la suivante:

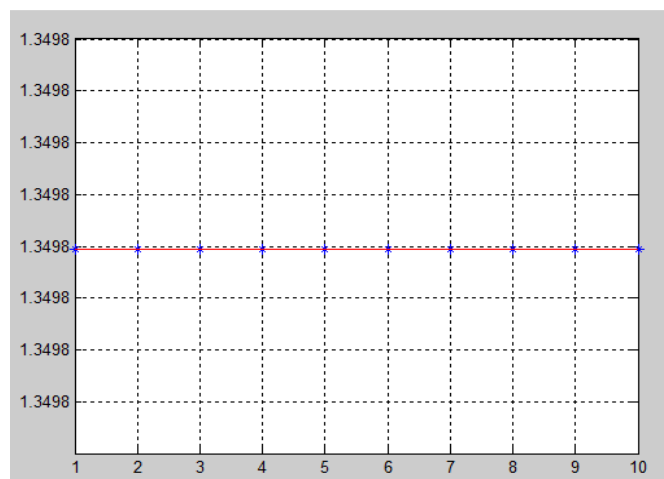


Figure IV.19: Valeur finale du Kurtosis (décision)

En rouge (graphe continu) represente les valeurs du Kurtosis des estimés pour différentes matrices de mélange, en bleu (graphe discontinu) représente la valeur du Kurtosis des signaux sources

**IV.4: Comparaison des résultats des méthodes directes**

Afin de mieux comparer les résultats des méthodes directes, on va mettre les résultats dans un tableau et la comparaison est faite selon trois différentes matrices de mélange.

$$A_1 = \begin{matrix} 0.3188 & -0.4336 \\ -1.3077 & 0.3426 \end{matrix} \quad A_2 = \begin{matrix} 0.8577 & 0.4494 \\ -0.6912 & 0.1006 \end{matrix} \quad A_3 = \begin{matrix} 0.8261 & 0.8979 \\ 0.5362 & -0.1319 \end{matrix}$$

Et les symbolisations des algorithmes sont les suivantes:

AG (Ali mansour géometrique), AC(Ali Manssour Cumulants),YC( Taro Yamaguchy )

<i>A1</i>	<i>600 échantillons</i>			<i>800 échantillons</i>			<i>1000 échantillons</i>		
<i>Méthode</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>
$\frac{ \Delta K }{K} \times 100$	5.92%	15.81%	6.64%	1.33%	4.20%	2.38%	0.0004%	0.0001%	$10^{-13}$ %
<i>Temps(s)</i>	1.97	0.54	0.22	2.80	0.76	0.27	3.71	1.0	0.3

<i>A2</i>	<i>600 échantillons</i>			<i>800 échantillons</i>			<i>1000 échantillons</i>		
<i>Méthode</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>
$\frac{ \Delta K }{K} \times 100$	5.92%	9.6%	6.62%	1.33%	3.64%	2.38%	0.0009%	0.00002%	$10^{-13}$ %
<i>Temps(s)</i>	1.92	0.55	0.17	2.82	0.80	0.27	3.78	1	0.3

<i>A3</i>	<i>600 échantillons</i>			<i>800 échantillons</i>			<i>1000 échantillons</i>		
<i>Méthode</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>	<i>AG</i>	<i>AC</i>	<i>YC</i>
$\frac{ \Delta K }{K} \times 100$	5.92%	10.86%	6.62%	1.32%	4.88%	2.38%	0.0008%	0.00005%	$10^{-13}$ %
<i>temps</i>	1.95	0.52	0.17	2.82	0.8	0.27	3.8	1	0.28

**TABLEAU 1:** Comparaison des méthodes directes

La première remarque concernant ces tableaux est que chaque méthode mène à des résultats satisfaisants pour n'importe quelle matrice non singulière, on déduit aussi que les méthodes basés sur l'estimation du système de mélange sont plus rapides et plus exactes à partir de 1000 échantillons par rapport à la méthode géométrique. Et comme on a mentionné pour les méthodes qui utilisent les cumulants croisés des observations la méthode de Tomaguchi est plus rapide et plus performante que celle d'Ali Mansour. L'avantage des méthodes géométrique est qu'on peut extraire plus que deux sources contrairement aux méthodes basées sur les cumulants croisés.

## IV.5: Implémentation des Méthodes Adaptatives

Dans cette partie nous appliquons les techniques adaptatives pour la SAS discutées dans le chapitre 3, et afin de montrer comment ces techniques peuvent atteindre la séparation, on s'est basé sur les conditions nécessaires concernant chaque algorithme pour qu'il puisse avoir des résultats acceptables, ces conditions sont posées par le concepteur de chaque algorithme.

### IV.5.1: Algorithme de (Herault et Jutten [26][15]) réseau direct

#### IV.5.1.1: (Deux sources)

En se référant à la figure(III.6) du chapitre 3. La matrice de mélange utilisée pour cette application est:

$$A_I = \begin{bmatrix} 0.5 & 0.6 \\ 0.8 & 0.2 \end{bmatrix}$$

Et les poids du réseau dans le cas de deux sources sont pris initialement:  $C_{12}=C_{21}=0.1$

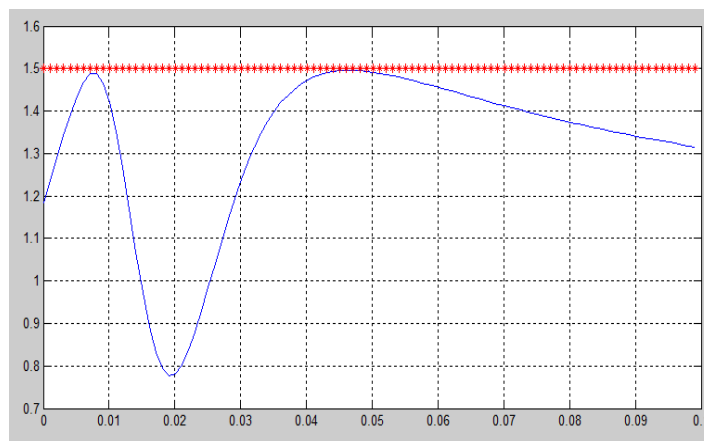
Les fonctions d'activations utilisées pour l'application sont:

$f(y)=y^3$ , et  $g(y)=\arctang(y)$ .

Et le gain d'apprentissage est exponentiellement décroissant.

#### IV.5.1.1.1: Convergence et influence du gain d'apprentissage

Pour cette matrice de mélange et ces valeurs des poids initiaux et en appliquant les relations (3.22) et (3.26) sur le modèle précédent, on doit tout d'abord vérifier l'influence de différents gains d'apprentissage sur la valeur finale du Kurtosis, afin de montrer est-ce que l'application de l'algorithme converge vers une solution ou non et le résultat est le suivant:



**Figure IV.20:** Valeurs finales du Kurtosis pour chaque valeur du gain d'apprentissage( $A_1$ )

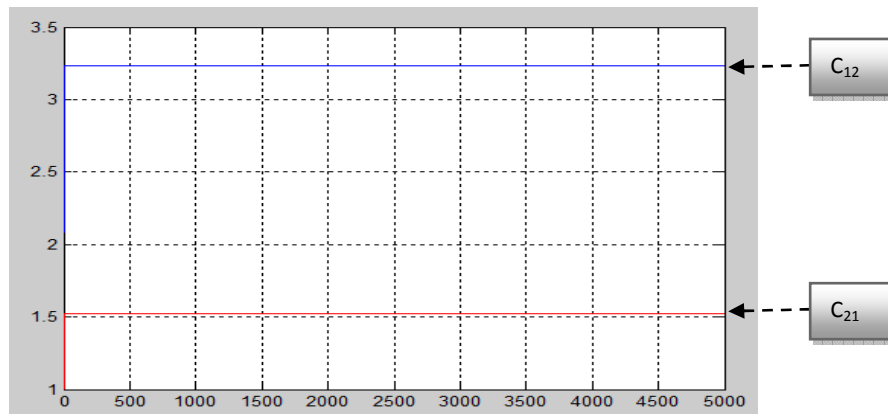
*Graphique continu (en bleu) représente les valeurs finales du Kurtosis pour chaque gain d'apprentissage, le graphique discontinu (en rouge) représente le Kurtosis des signaux sources.*

À partir de la figure précédente, on a deux gains qui convergent pour une meilleure séparation, mais est-ce qu'on a obtenu l'une des solutions du réseau directe, et qui sont déduites à partir des équations (3.24) et (3.25), et avec la matrice  $A_1$  normalement, on doit avoir une de ces solutions:

Solution 1 :  $C_{12}= 3$ , et  $C_{21}= 1.6$

Solution 2 :  $C_{12}= 0.6250$ , et  $C_{21}= 0.3333$

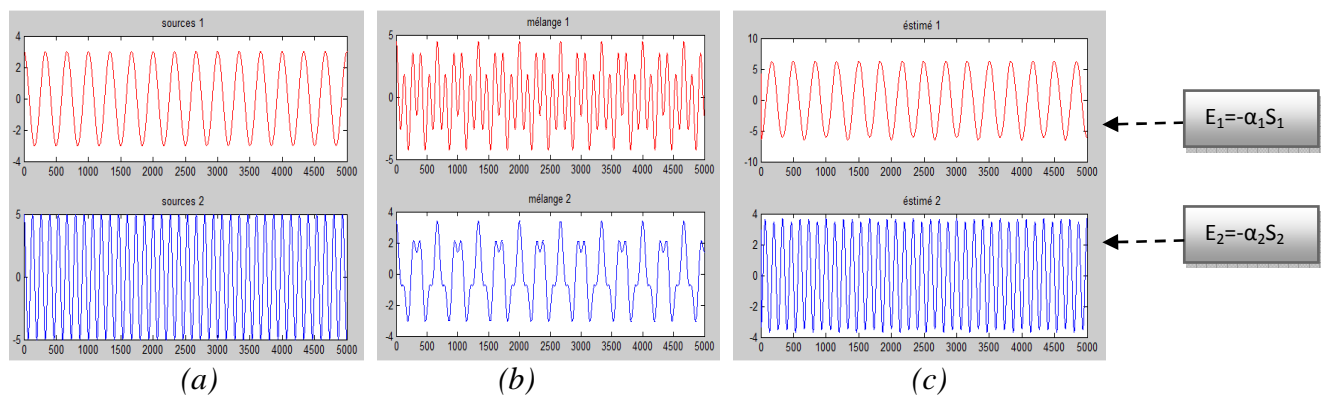
et le résultat de l'adaptation des poids est sur la figure ci-dessous :



**Figure IV.21:** Valeurs des poids du réseau( $A_1$ )

*Valeurs des poids du réseau direct en fonction d'échantillon pour le meilleur gain d'apprentissage (pour  $A_1$ ), en bleu  $C_{12}$ , en rouge  $C_{21}$*

Donc l'algorithmme a convergé vers la première solution, et les résultats de la séparation sont les suivants:



**Figure IV.22:** Résultat de séparation ( $A_1$ )

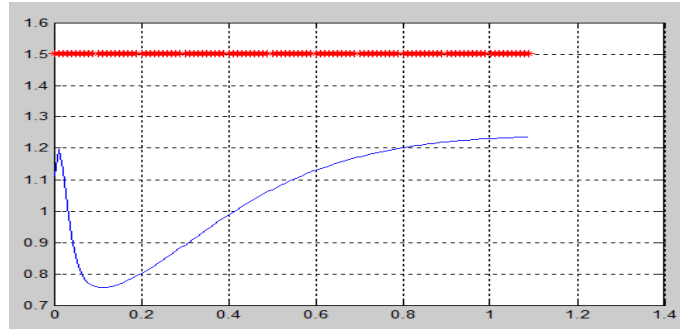
*Résultat de séparation en utilisant le réseau directe pour la matrice  $A_1$ , (a): signaux source, (b): signaux mélange, (c): signaux estimés ( $\alpha_i > 0$ )*



Si on change la matrice de mélange par :

$$A_2 = \begin{bmatrix} -0.5 & 0.6 \\ 0.8 & 0.2 \end{bmatrix}$$

L'influence de différents gains d'apprentissage sur la valeur finale du Kurtosis est sur la figure suivante:



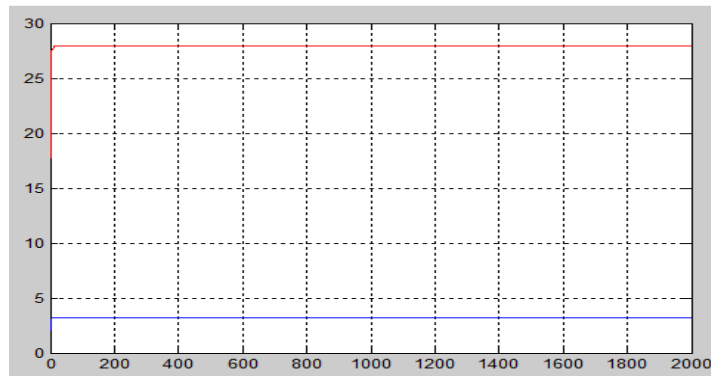
**Figure IV.23:** Valeurs finales du kurtosis pour chaque valeur du gain d'apprentissage( $A_2$ )

*Grphe continu(en bleu) représente les valeurs finales du Kurtosis pour chaque gain d'apprentissage, le graphe discontinu (en rouge) représente le Kurtosis des signaux sources.* On remarque qu'on n'a pas pu atteindre la valeur du Kurtosis des signaux sources, et avec la matrice  $A_2$  normalement on doit avoir une de ces solutions:

Solution 1 :  $C_{12}= 3$ , et  $C_{21}= -1.6$

Solution 2 :  $C_{12}= -0.6250$ , et  $C_{21}= 0.3333$

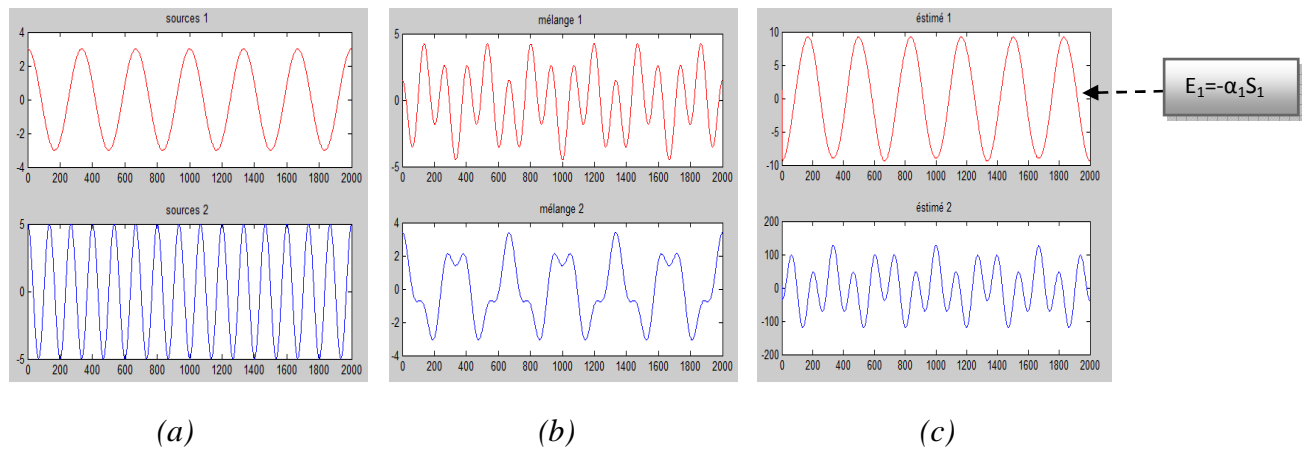
et le résultat de l'adaptation des poids est sur la figure ci-dessous:



**Figure IV.24:** Valeurs des poids du réseau( $A_2$ )

*Valeurs des poids du réseau direct en fonction d'échantillon pour le meilleur gain d'apprentissage (pour  $A_2$ ), en bleu  $C_{12}$ , en rouge  $C_{21}$*

D'après cette figure on remarque que la convergence est atteinte que pour le poids  $C_{12}$  de la première solution, mais le poids  $C_{21}$  a divergé vers une autre solution, donc on peut recouvrer que la première source, et les résultats de la séparation sont les suivants.

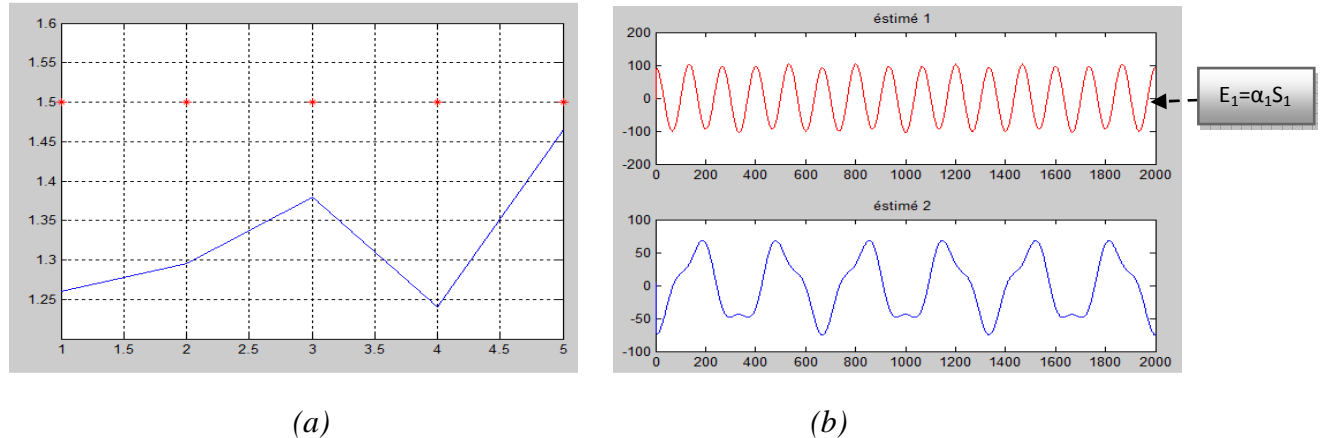


**Figure IV.25:** Résultat de séparation(A2)

Résultat de séparation en utilisant le réseau directe pour la matrice  $A_2$ , (a) signaux sources, (b) signaux mélanges, (c) signaux estimés ( $\alpha_i > 0$ )

**IV.5.1.1.2: Valeur du Kurtosis pour différentes matrices**

Les résultats des Kurtosis des signaux pour des différentes matrices de mélange prises aléatoirement sont les suivants:



**Figure IV.26:** Résultat pour différentes matrices de mélanges

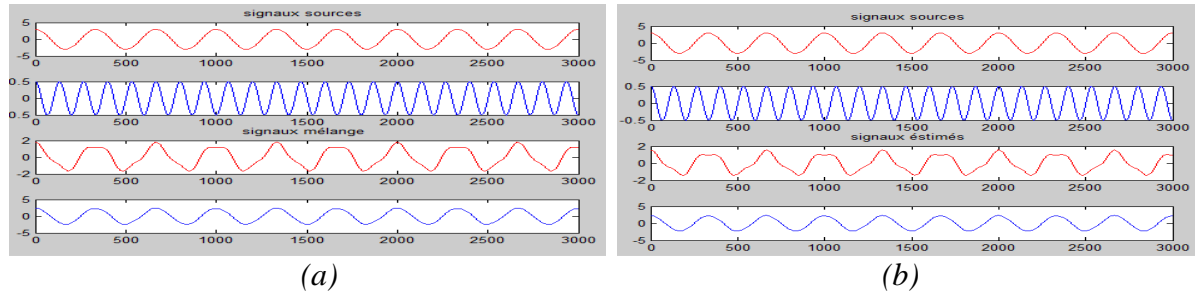
(a): valeur finale du Kurtosis pour chaque matrice de mélange prise aléatoirement

(b) : Résultat de séparation pour la cinquième matrice  $A = \begin{bmatrix} -0.86 & -0.16 \\ -0.03 & 0.62 \end{bmatrix}$

On déduit de cette figure qu'avec les mêmes conditions on n'aura pas toujours des solutions désirables et comme exemple, pour la cinquième matrice du graphe on a pu obtenir que la deuxième source.

**IV.5.1.1.3 : Problème des échelles mal conditionnées**

On a mentionné aussi que l'un des problèmes de cet algorithme est que si les sources ont des valeurs amplitudes petites ou mal conditionnées la séparation n'est pas atteinte comme le montre la figure suivante:



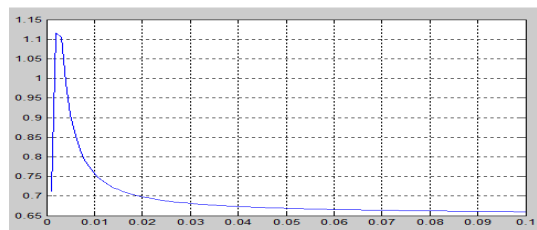
**Figure IV.27:**Problème des échelles mal conditionnées

(a): signaux sources et signaux mélanges, (b): signaux sources et signaux estimés

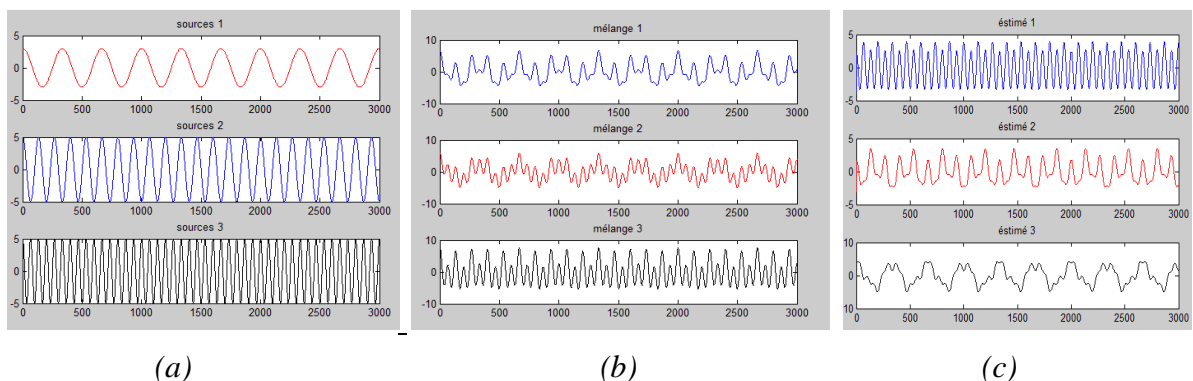
**IV.5.1.2:(Trois sources)**

On suivant les mêmes étapes de l'application de l'algorithme, le meilleur résultat qu'on a pu obtenir pour des signaux sinusoïdaux (Kurtosis=1.5) et avec la matrice ci dessous sont les suivants:

$$A = \begin{bmatrix} 0.7 & 0.6 & 0.3 \\ 0.8 & 0.2 & 0.5 \\ 0.2 & 0.5 & 0.9 \end{bmatrix}$$



**Figure IV.28:** Valeurs finales du Kurtosis pour chaque valeur du gain d'apprentissage



**Figure IV.29:** Résultat de séparation

Résultat de séparation en utilisant le réseau directe pour le cas de trois sources, (a): signaux sources, (b):signaux mélanges, (c):signaux estimés.

**IV.5.2:Algorithme de (Herault et Jutten [26] [15]) réseau récurrent**

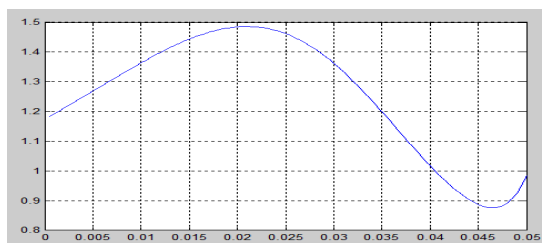
**IV.5.1.1:(Deux sources)**

Avec les mêmes conditions du réseau directe, l'application du réseau récurrent doit nous mener à une des deux solutions déduites à partir des équations (3.29) et (3.30), donc à partir de la matrice  $A_1$  (page.78), les poids du réseau doivent converger vers une de ces solution:

Solution 1 : $C_{12}= 3$ , et  $C_{21}= 1.6$

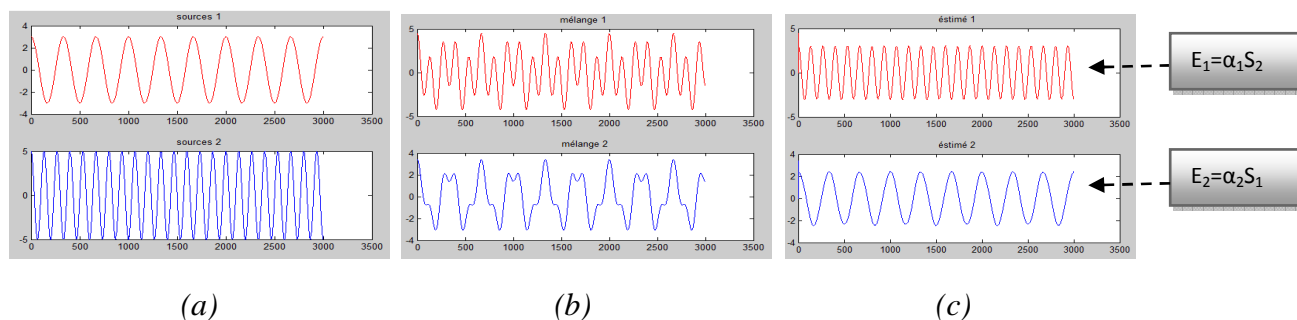
Solution 2 : $C_{12}= 0.6250$ , et  $C_{21}= 0.3333$

L'influence de différents gains d'apprentissage sur la valeur finale du kurtosis est présentée sur la figure suivante pour le cas des signaux sinusoïdaux (Kurtosis=1.5).



**Figure IV.30:** Valeurs finales du Kurtosis pour chaque valeur du gain d'apprentissage

Les valeurs finales des poids du réseau sont,  $C_{12}= 0.6247$ , et  $C_{21}= 0.3385$ . Donc les poids ont convergé vers la deuxième solution, et en se référant au troisième chapitre on doit avoir une permutation de position des sources et les résultats de la séparation sont les suivants:



**Figure IV.31:** Résultat de séparation

Résultat de séparation pour (A1) en utilisant le réseau récurrent, (a): signaux sources, (b): signaux mélanges, (c): signaux estimés ( $\alpha_i > 0$ ).

**IV.5.2.2:Remarques concernant les résultats des deux réseaux**

À l'application de ce dernier réseau sur différentes matrices de mélange et en suivant les mêmes étapes du réseau direct on a remarqué que le réseau est plus stable et donne de

meilleurs résultats comparant au réseau direct, mais malgré ça dans certains cas, l'application de cet algorithme nous mène qu'à l'extraction d'une seule source et il diverge aussi pour le cas de trois sources.

**IV.5.3:Algorithme de (A-Cichocki et R -Unbehauen [20][24])**

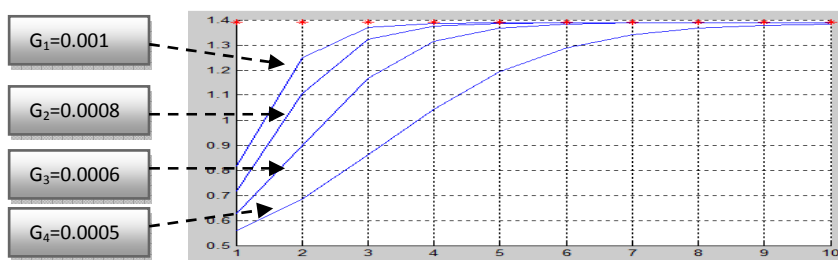
**IV.5.3.1:(Deux sources)**

En appliquant les relations (3.31) et (3.32) de l'algorithme de (A-Cichocki et R-Unbehauen) du troisième chapitre sur un modèle de mélange linéaire instantané. la matrice de mélange utilisée pour le cas de deux sources est :

$$A = \begin{bmatrix} 0.1097 & -0.2900 \\ 1.1287 & 1.2616 \end{bmatrix}$$

Les fonctions d'activation sont :  $f(y)=y^3$ , et  $g(y)=y^5$

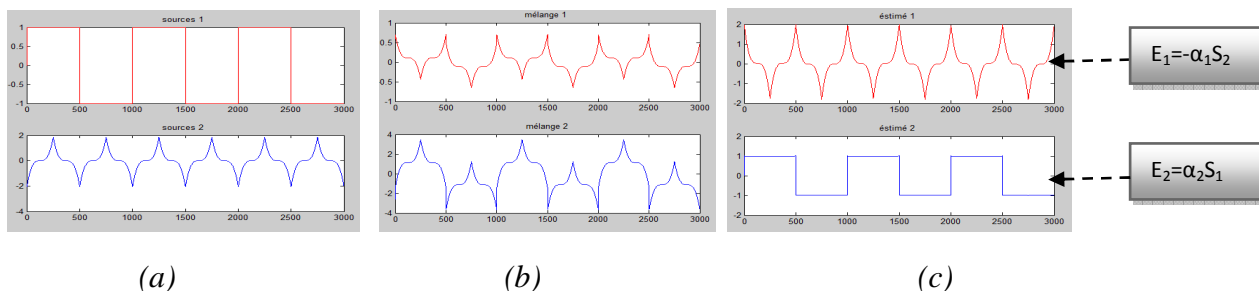
Tout d'abord il faut s'avoir l'influence du gain d'apprentissage sur la convergence de l'algorithme, vers une séparation de l'ensemble des sources et vers la valeur du Kurtosis finale pour chaque gain et selon le nombre d'itérations. Les résultats sont les suivants:



**Figure IV.32:** Valeurs du Kurtosis en fonction du nombre d'itérations

(a) : les graphes (continus) présentent l'évolution de la valeur du kurtosis pour 10 itérations de quatre gains d'apprentissage, (b): le graphe (discontinu) représente la valeur du Kurtosis des signaux sources.

Et le résultat de séparation est le suivant:



**Figure IV.33:** Résultat de séparation

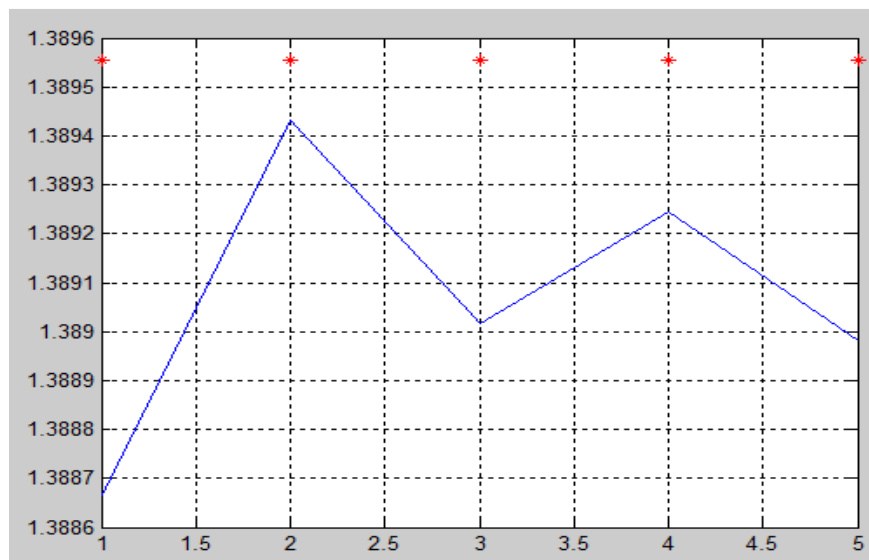
Résultat de séparation de deux sources en utilisant le réseau Cichoki et Unbehauen, (a): signaux sources, (b): signaux mélanges, (c): signaux estimés ( $\alpha_i > 0$ ).

à partir des deux figures précédentes on déduit qu'avec suffisamment d'échantillons tous les gains convergent approximativement vers le même point, et pour n'importe quel gain, et dix itérations sont suffisantes pour atteindre ce point, la seule différence est que certains gains mènent vers une séparation plus rapidement que les autres.

#### IV.5.3.1.1: Valeur du Kurtosis pour différentes matrices

Maintenant il faut vérifier si l'algorithme mène aux mêmes résultats des Kurtosis pour différentes matrices de mélange prises aléatoirement, à condition que ces matrices ne soient pas singulières.

Les signaux utilisés pour cette application sont les mêmes qui sont utilisés pour le cas de deux sources et Le graphe ci dessous nous montre les résultats.



**Figure IV.34:** Valeurs du Kurtosis pour différentes matrices

(a): le graphe (continu) en bas les valeurs du Kurtosis pour chaque matrice, (b) le graphe (discontinu) la valeur du Kurtosis des signaux sources.

On déduit de ce graphe que l'algorithme est convergent pour toutes les matrices non singulières et mène vers une séparation satisfaisante. Pour le cas où on a la présence de la matrice mal conditionnée la séparation est atteinte sauf que pour avoir un résultat satisfaisant l'algorithme nécessite plus d'itérations.

IV.5.3.2:(Trois sources)

La matrice de mélange utilisée pour trois sources est :  $A = \begin{bmatrix} 0.5377 & 0.8622 & -0.4336 \\ 1.8339 & 0.3188 & 0.3426 \\ -2.2588 & -1.3077 & 3.5784 \end{bmatrix}$

Pour la convergence, nous avons la même remarque, tous les gains mènent à une valeur désirable mais avec des petites erreurs légères, et on remarque aussi que les gains de petites valeurs sont lents pour atteindre la séparation, mais ils donnent de meilleurs résultats comme le montre la figure ci-dessous :

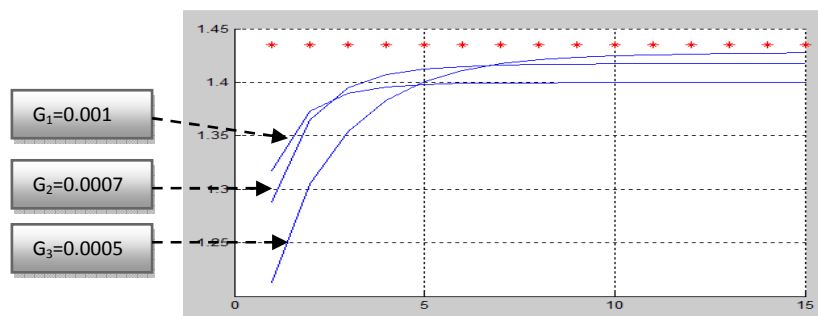


Figure IV.35: Valeurs du Kurtosis en fonction du nombre d'itérations

(a): les graphes continus présentent l'évolution de la valeur du kurtosis pour 15 itérations de trois gains d'apprentissage, (b): le graphe discontinu c'est la valeur du Kurtosis des signaux sources.

Et le résultat de la séparation pour trois sources est le suivant :

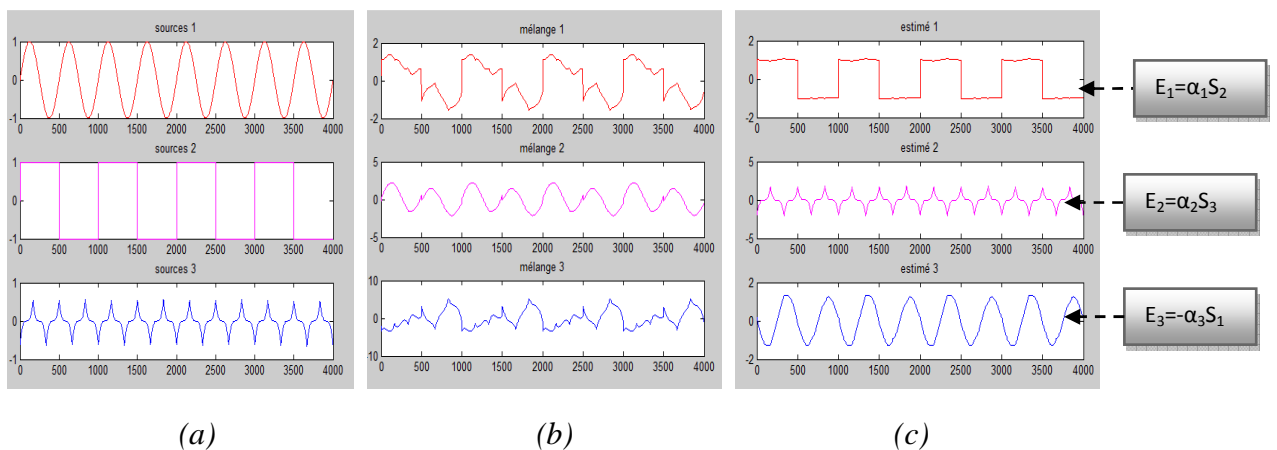


Figure IV.36: Résultat de séparation

Résultat de séparation de trois sources en utilisant le réseau Cichoki et Unbehauen,

(a) : signaux sources, (b) signaux mélanges, (c) signaux estimés ( $\alpha_i > 0$ ).

IV.5.3.3:(Quatre sources)

En appliquant la même procédure pour quatre sources et avec plus d'échantillons les remarques déduites sont les mêmes pour le cas de trois sources, et les résultats sont les suivants pour la matrice de mélange:

$$A = \begin{bmatrix} 1.7119 & 1.3546 & 1.4367 & 2.9080 \\ -0.1941 & -1.0722 & -1.9609 & 0.8252 \\ -2.1384 & 0.9610 & -0.1977 & 1.3790 \\ -0.8396 & 0.1240 & -1.2078 & -1.0582 \end{bmatrix}$$

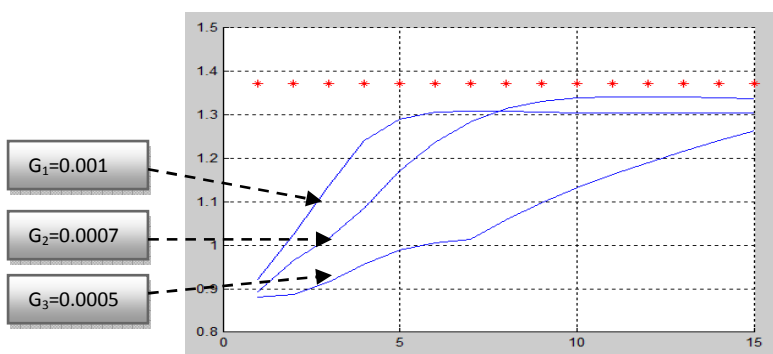


Figure IV.37: Valeurs du Kurtosis en fonction du nombre d'itérations

(a): les graphes (continus) représentent l'évolution de la valeur du Kurtosis pour 15 itérations de trois gains d'apprentissage, (b): le graphe (discontinu) c'est la valeur du Kurtosis des signaux sources.

Et le résultat de la séparation pour quatre sources est le suivant :

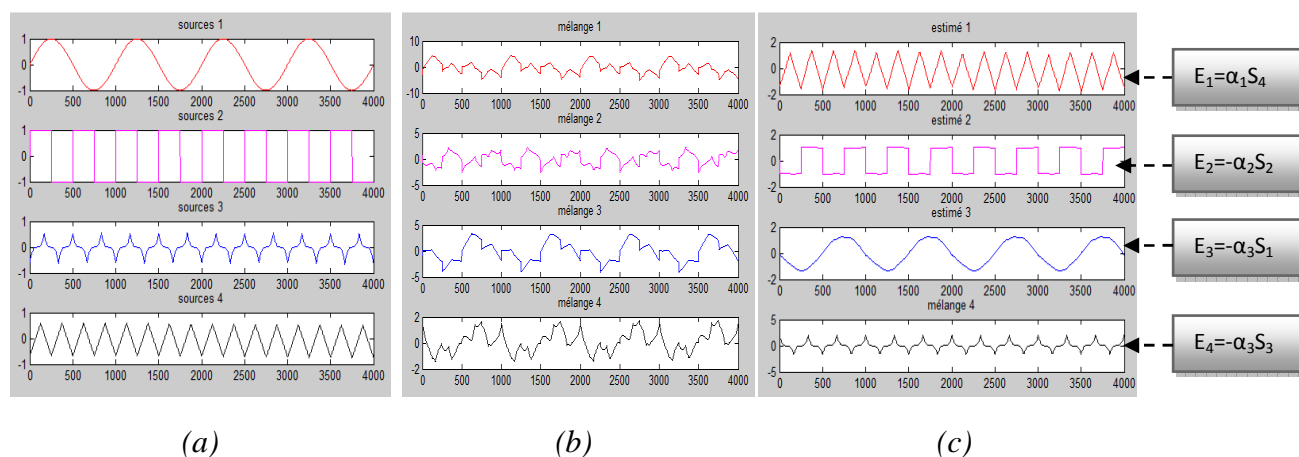


Figure IV.38: Résultat de séparation

Résultat de séparation de quatre sources en utilisant le réseau Cichoki et Unbehauen, (a): signaux sources, (b): signaux mélanges, (c): signaux estimés ( $\alpha_i > 0$ ).

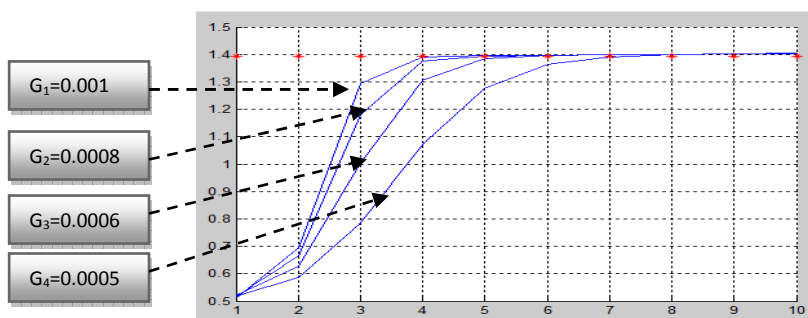


Et lors de l'implémentation on a remarqué aussi que le problème des valeurs d'amplitude mal conditionnées des sources n'est pas présent dans cet algorithme, et on a pu même extraire deux sources qui ont le rapport des longueurs des intervalles égal à 500.

**IV.5.4:Algorithme DE (A-Cichoki et S-Amari [21][25])**

**IV.5.4.1:(Deux sources)**

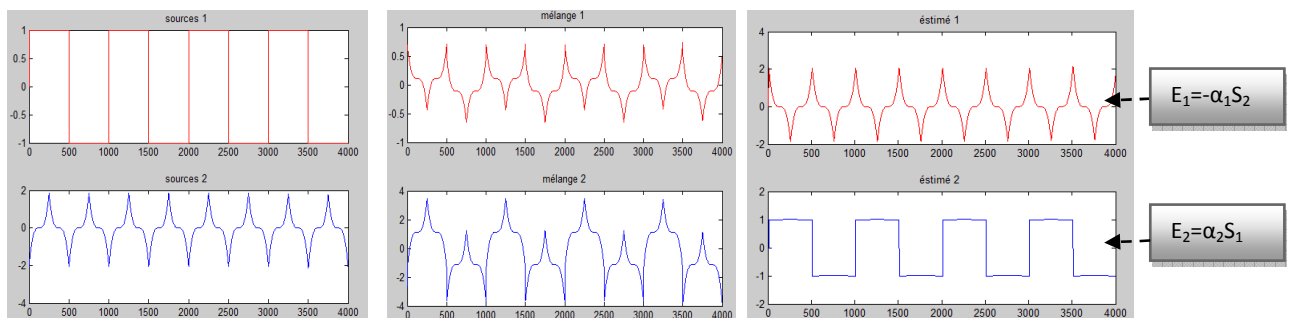
En appliquant les équations (3.33) et (3.34) du chapitre 3 et en suivant les mêmes étapes pour l'application de l'algorithme de (Cichoki et Unbehauen), et avec la même matrice de mélange, les résultats pour le cas de deux sources sont les suivants



**Figure IV.39:** Valeurs du Kurtosis en fonction du nombre d'itérations

(a) : les graphes (continus) présentent l'évolution de la valeur du kurtosis pour 10 itérations de quatre gains d'apprentissage, (b) : le graphe (discontinu) représente la valeur du Kurtosis des signaux sources.

Et le résultat de séparation de deux sources est le suivant:



(a)

(b)

(c)

**Figure IV.40:** Résultat de séparation

Résultat de séparation de deux sources en utilisant le réseau A-Cichoki et S-Amari,

(a): signaux sources, (b): signaux mélanges, (c): signaux estimés ( $\alpha_i > 0$ ).

IV.5.4.2:(Trois sources)

Les résultats obtenus par l'application de l'algorithme de Cichoki et S-Amari pour le cas de trois sources sont ci dessous, où la matrice de mélange utilisée et la même qui a été prise dans le cas d'application de l'algorithme de Cichoki et Unbehauen dans le cas de trois sources.

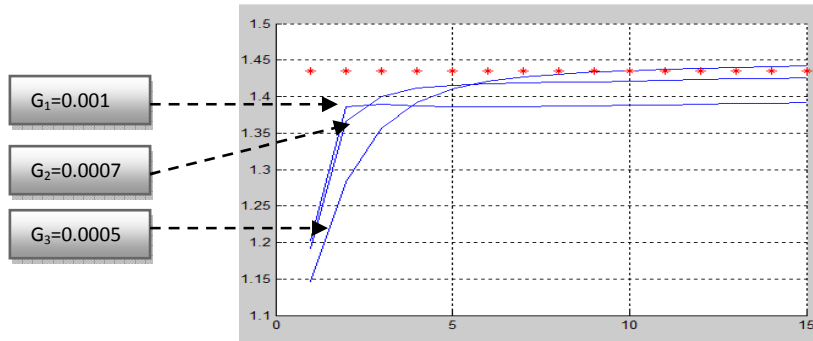


Figure IV.41: Valeurs du Kurtosis en fonction du nombre d'itérations

(a): les graphes (continus) représentent l'évolution de la valeur du kurtosis pour 15 itérations de trois gains d'apprentissage, (b): le graphe (discontinu) représente la valeur du Kurtosis des signaux sources.

À partir de cette figure et en se référant à la figure (IV.35), on peut déduire facilement que cet algorithme est plus rapide et plus performant que celui de (Cichoki et Unbehauen), et le résultat de séparation de trois sources est le suivant:

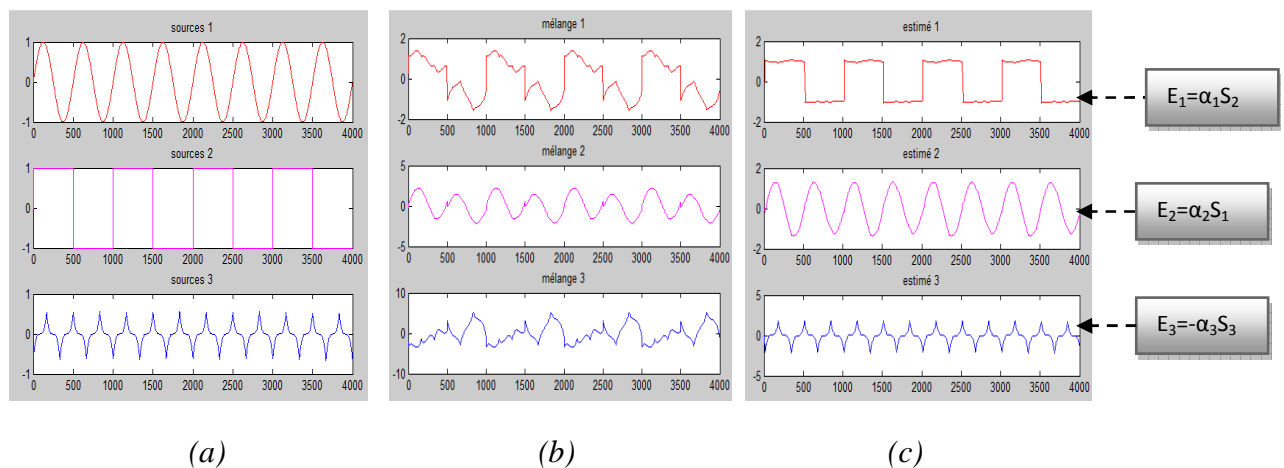


Figure IV.42: Résultat de séparation

Résultat de séparation de trois sources en utilisant le réseau A-Cichoki et S-Amari , (a) signaux source, (b) signaux mélange, (c) signaux estimés ( $\alpha_i > 0$ ).

IV.5.4.3:(Quatre sources)

La matrice de mélange utilisée est la même qui a été prise dans le cas d'application de l'algorithme de Cichoki et Unbehauen dans le cas de quatre sources et les résultats sont les suivants.

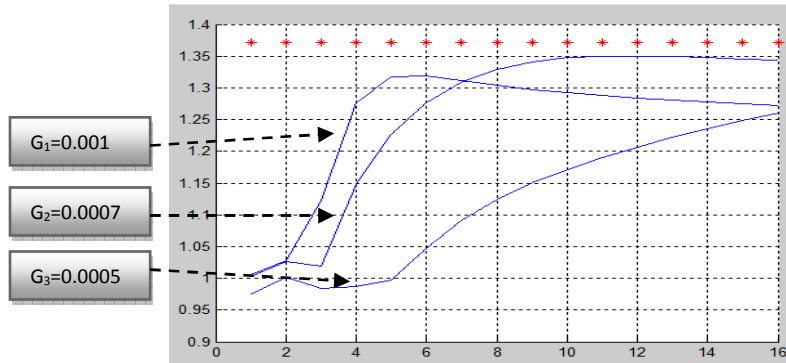


Figure IV.43: Valeurs du Kurtosis en fonction du nombre d'itérations

(a): les graphes (continus) présentent l'évolution de la valeur du kurtosis pour 15 itérations de trois gains d'apprentissage, (b): le graphe (discontinu) représente la valeur du Kurtosis des signaux sources.

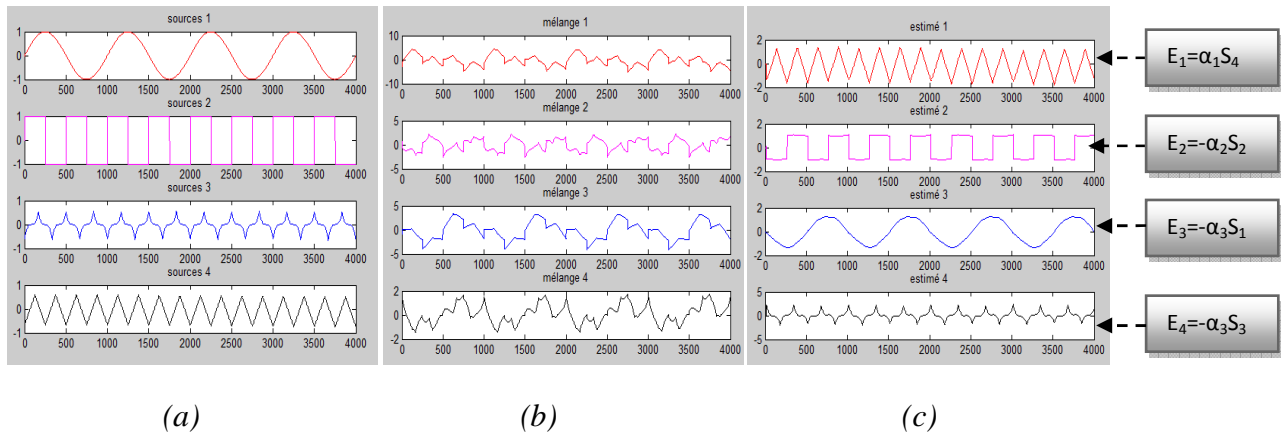


Figure IV.44: Résultat de séparation

Résultat de séparation de quatre sources en utilisant le réseau A-Cichoki et S-Amari, (a) signaux source, (b) signaux mélange, (c) signaux estimés ( $\alpha_i > 0$ ).

IV.6: Comparaison des méthodes adaptatives

On doit maintenant comparer les résultats des méthodes adaptatives selon le nombre de sources et pour une seule matrice de mélange non singulière. Pour les matrices de mélanges, dans le cas de deux sources on a pris la matrice utilisée dans le cas de deux sources de l'algorithme d'Hérault et Jutten, mais pour trois et quatre sources on a choisi les deux matrices

utilisées pour l'algorithme de Cichoki et Unbehauen. Les symbolisations des algorithmes sont les suivantes:

HJ: (Hérault et Jutten), CR: (Cichoki et R. Unbehauen), CA:( Cichoki et Amari)

Remarque: pour la comparaison, les algorithmes sont exécutés de façon à avoir la meilleure approximation avec les gains utilisés, pour cela on a impliqué même l'étape d'apprentissage pour le calcul du temps d'exécution.

<u>2 sources</u>	2000 échantillons			3000 échantillons			4000 échantillons		
Méthode	HJ	CR	CA	HJ	CR	CA	HJ	CR	CA
$\frac{ \Delta K }{K} \times 100$	0.26%	0.13%	1.74%	0.25%	0.04%	1.09%	0.24%	0.16%	0.72%
Temps(s)	9.59	5	3.78	17.49	7.65	5.65	43.19	10.58	7.83

<u>3 sources</u>	2000 échantillons			3000 échantillons			4000 échantillons		
Méthode	HJ	CR	CA	HJ	CR	CA	HJ	CR	CA
$\frac{ \Delta K }{K} \times 100$	29.1%	7.9%	0.027%	27.15%	0.94%	0.35%	25.56%	0.58%	0.46%
Temps(s)	13.8	8	10.19	27.06	21.45	18.50	44.13	33.81	29.36

<u>4 sources</u>	2000 échantillons			3000 échantillons			4000 échantillons		
Méthode	HJ	CR	CA	HJ	CR	CA	HJ	CR	CA
$\frac{ \Delta K }{K} \times 100$	-	4.63	3.70	-	2.39	1.78	-	0.69	0.46
Temps(s)	-	15.21	10.83	-	27.5	24	-	46.14	42.73

**TABLEAU 2:** Comparaison des méthodes adaptatives

(I): Pour le cas de deux sources tous les algorithmes sont performants et donnent de meilleurs résultats pour plus d'échantillons. Sauf pour le cas où la matrice est mal conditionnée, les résultats de HJ divergent beaucoup, et comme exemple pour la matrice

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0.99 \end{bmatrix} \text{ mal conditionnée, } \det(A) = -0.01$$

Les résultats du pourcentage d'erreur du Kurtosis pour les trois algorithmes sont les suivants:

$$HJ: \frac{|\Delta K|}{K} \times 100 = 359 \%$$

$$CR: \frac{|\Delta K|}{K} \times 100 = 0.0136 \%$$

$$CA: \frac{|\Delta K|}{K} \times 100 = 0.6875 \%$$

(2): pour le cas de trois sources on remarque que l'algorithme de H.J diverge et le pourcentage d'erreur du Kurtosis est supérieur à 25%, qui est un résultat indésirable pour la séparation, même avec plus d'échantillon on ne peut avoir de meilleurs résultats.

(3): en ce qui concerne les algorithmes de (Cichoki et R. Unbehauen.) et (Cichoki et Amari), les algorithmes donnent des résultats satisfaisants pour les trois cas 2, 3,4 sources, et on remarque aussi que le premier algorithme donne un résultat meilleur pour le cas de deux sources, mais à partir de trois sources l'algorithme de (Cichoki et Amari), devient plus rapide et plus performant.

#### IV.7: Comparaison des méthodes adaptatives et directes

Afin de montrer les avantages et les inconvénients de chaque technique, il faut faire une comparaison entre ces deux catégories. Cette comparaison est faite pour le cas de deux sources entre l'algorithme de Taro Yamaguchy pour les méthodes directes et celui de (Cichoki et R. Unbehauen.) pour les méthodes adaptatives, et la matrices de mélange est prise d'une façon aléatoire pour les deux algorithmes, et les résultats de comparaison sont les suivants:

	2000 échantillons		3000 échantillons	
Méthode	YC	CA	YC	CA
$\frac{ \Delta K }{K} \times 100$	$3.9 \times 10^{-14} \%$	0.1343%	$1.9 \times 10^{-14} \%$	0.0464%
Temps(s)	0.7	5	1.3	21.45

**TABLEAU 3:** Comparaison des méthodes directes et adaptatives

L'avantage des algorithmes directs est qu'ils sont plus rapides et performants que ceux des méthodes adaptatives pour le cas de deux sources. Au-delà de deux sources les méthodes directes nécessitent un autre calcul analytique contrairement aux méthodes adaptatives, avec l'application du même algorithme on peut séparer plus que trois sources et avec des résultats satisfaisants, mais l'inconvénient de ces techniques est que le temps d'exécutions est très grand et elles nécessitent plus d'échantillons pour atteindre la séparation.

**IV.8:Présence d'une source gaussienne**

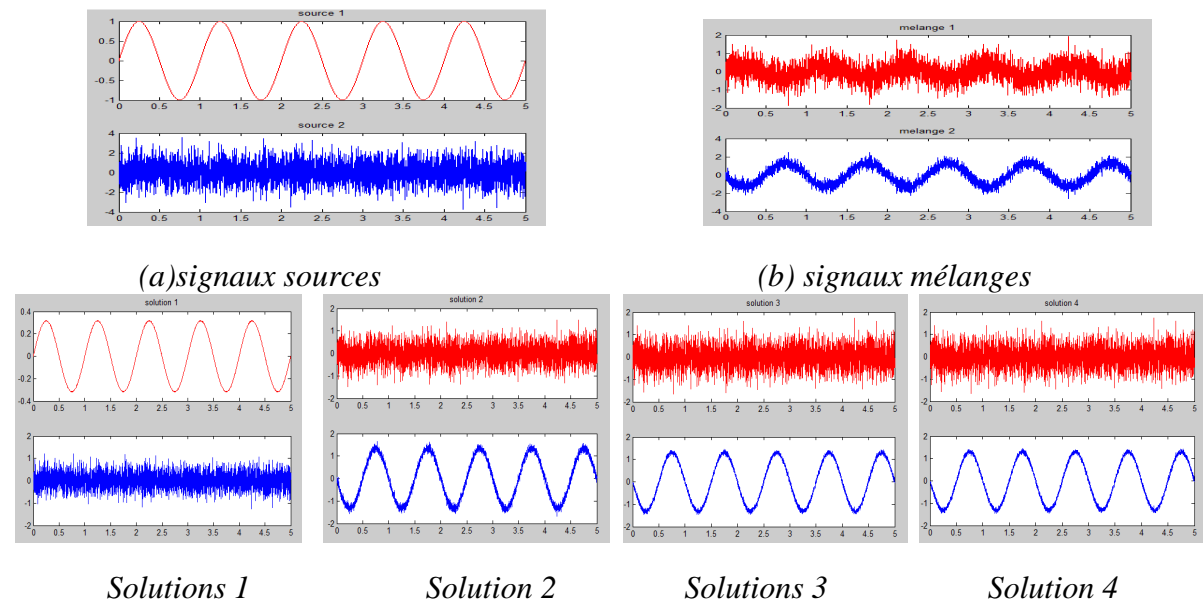
on a supposé dans l'application des algorithmes que toutes les sources sont non gaussiennes, et on a obtenu des résultats satisfaisants et la séparation est atteinte, maintenant si on a parmi ces sources une qui a la densité de probabilité gaussienne tel que le bruit blanc aditif, si on se base sur le théorème de séparabilité du deuxième chapitre dans la section (II.6.3) qui stipule que pour atteindre la séparation en utilisant l'ICA il faut que les signaux sources contiennent tout au plus une source gaussienne.

**IV.8.1:Présence d'une source gaussienne (méthodes directes)**

**IV.8.1.1:Solutions**

Pour les méthodes directes on a choisi l'application de l'algorithme de Tymaguchi et les résultats de la solution du polynôme (3.18) sont les suivants pour la matrice de mélange.

$$A = \begin{bmatrix} 0.3188 & -0.4336 \\ -1.3077 & 0.3426 \end{bmatrix}$$



**Figure IV.45:** Résultats de l'application (les 4 solutions)

*Résultats de l'application de l'algorithme basé sur les Cumulants de Taro Yamaguchi sur le mélange d'un signal sinusoïdal et un signal gaussien*

Les valeurs de  $\alpha$  : [-1.2774    -0.2814 - 0.1112i    -0.2814 + 0.1112i    -0.1984]

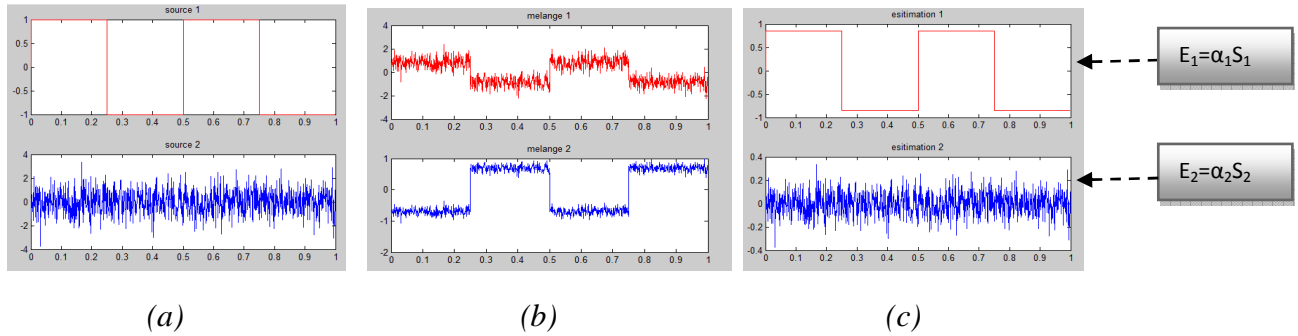
Les valeurs de  $\beta$  : [-4.1275    -0.7487 - 0.5732i    -0.7487 + 0.5732i    -0.9658]

Le Kurtosis des sources= 0.7899, et le kurtosis des mélange= 0.5929

Les Kurtosis des quatre solutions : [0.7901    0.9509    0.9509    0.7742]

**IV.8.1.2:Décision**

Alors pour la décision il faut choisir les solutions dont les coefficients sont réels et avec le Kurtosis maximal, et voici un exemple pour la séparation d'un signal carré et d'une source gaussienne après décision:



**Figure IV.46:** Résultats de l'application (décision)

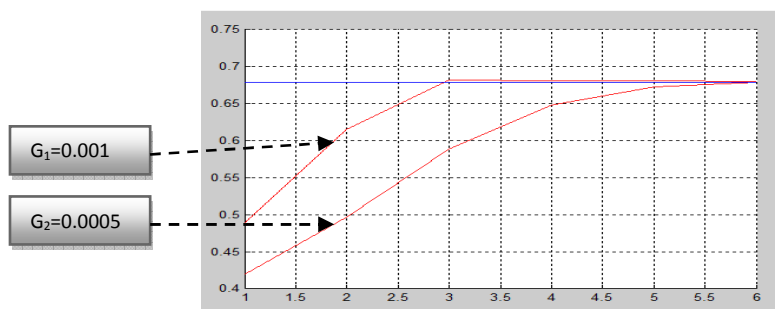
*Résultats de l'application sur le mélange d'un signal carré et un signal gaussien,*

*(a): signaux sources, (b) signaux mélanges, (c) : signaux estimés après décision ( $\alpha_i > 0$ ).*

**IV.8.2:Présence d'une source gaussienne (méthodes adaptatives)**

**IV.8.2.1:Deux sources**

Pour les méthodes adaptatives on a choisi l'application de l'algorithme de (Cichoki et Amari) pour une matrice de mélange prise d'une façon aléatoire, et les résultats sont les suivants:



**Figure IV.47:** Valeurs du Kurtosis en fonction du nombre d'itérations

*En rouge l'évolution de la valeur du Kurtosis pour 6 itérations de deux gains d'apprentissage, en bleu (la droite en haut) la valeur du Kurtosis des signaux sources*

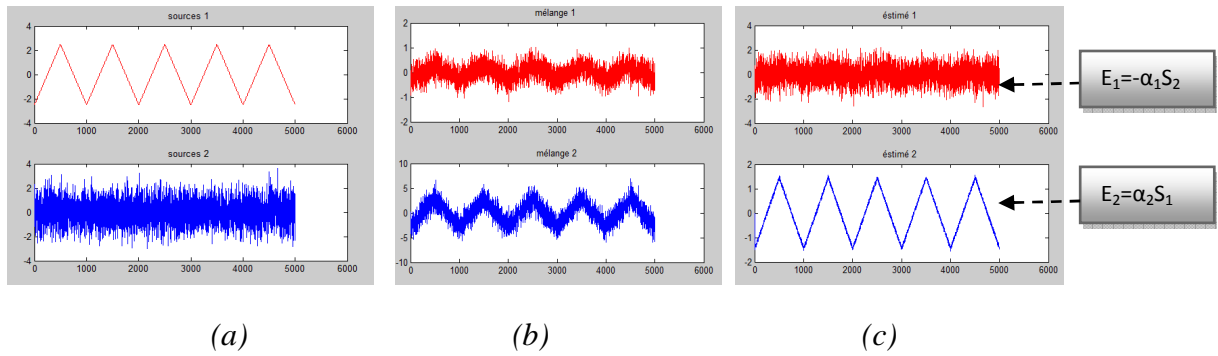


Figure IV.48: Résultat de séparation

Résultat de séparation de deux sources où une d'elles est gaussienne en utilisant le réseau (Cichoki et Amari), (a): signaux sources, (b): signaux mélanges, (c): signaux estimés ( $\alpha_i > 0$ ).

IV.8.2.2: Trois sources

Même la présence d'une source gaussienne parmi trois sources, on peut atteindre la séparation en utilisant les techniques adaptatives et les résultats de la séparation sont les suivants:

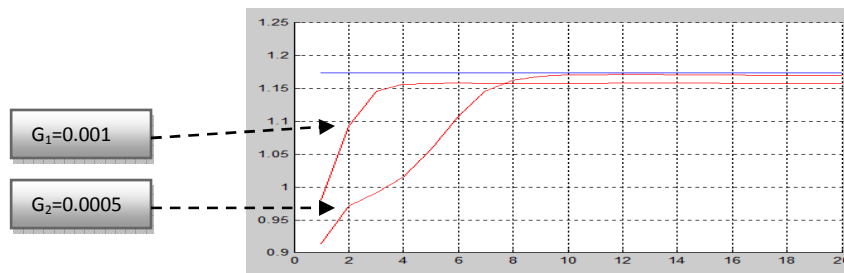


Figure IV.49: valeurs du Kurtosis en fonction du nombre d'itérations

En rouge l'évolution de la valeur du Kurtosis pour 20 itérations de deux gains d'apprentissage, en bleu (la droite en haut) la valeur du Kurtosis des signaux sources

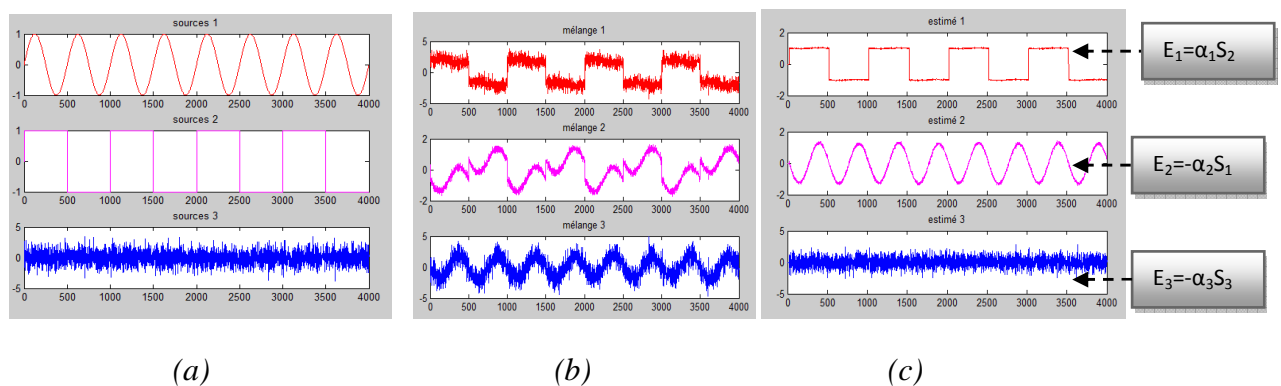


Figure IV.50: résultat de séparation

Résultat de séparation de trois sources dont l'une d'elles est gaussienne en utilisant le réseau (Cichoki et Amari), (a): signaux source, (b): signaux mélange, (c): signaux estimés ( $\alpha_i > 0$ ).

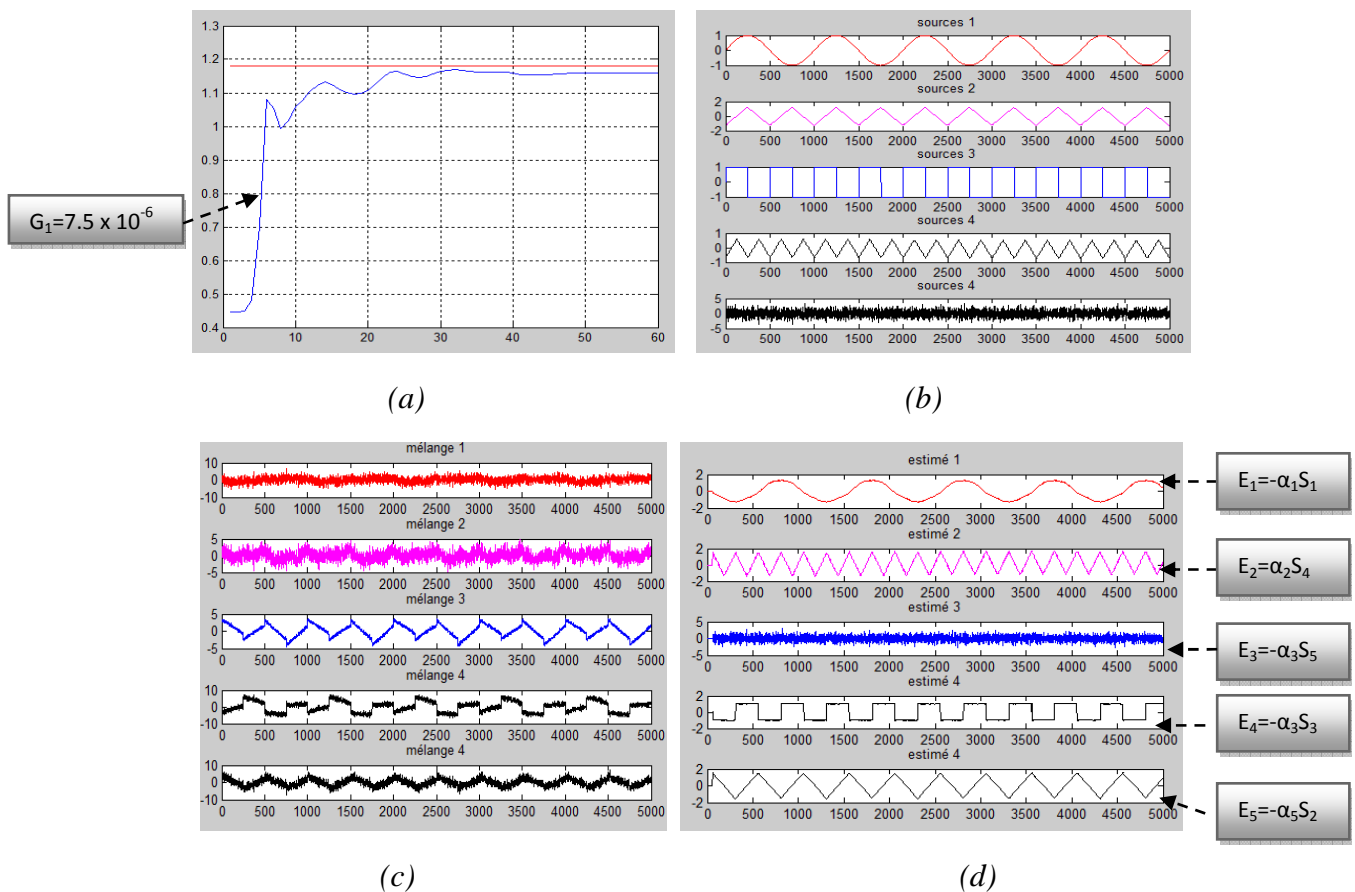


**IV.9:Influence du gain d'apprentissage sur le nombre de sources**

Jusqu'ici on a discuté la séparation en utilisant les deux techniques, et on a conclu que les méthodes adaptatives dépendent de plusieurs facteurs, L'un de ces facteurs et le gain d'apprentissage que sa valeur joue un rôle très important pour le nombre de sources à séparer Lors de l'implémentation de l'algorithme de (Cichoki et Amari), on a pu séparer cinq sources, mais il nous a fallu le changement du gain.

**IV.9.1:Cinq sources**

Dans cette application, on a séparé cinq sources avec l'algorithme de (Cichoki et Amari), mais avec un changement de l'intervalle du gain d'apprentissage et une augmentation du nombre d'itérations et d'échantillons. Et pour l'accomplissement de ce travail, on a pris aussi le cas où on a la présence d'une source gaussienne et les résultats sont les suivants:



**Figure IV.51:**Résultat de séparation de cinq sources

*Résultat de séparation de cinq sources par le réseau de (Cichoki et Amari) où la matrice de mélange a été prise d'une façon aléatoire. (a):l'évolution du Kurtosis en fonction du nombre d'itérations, (b):siganux sources,(c): signaux mélanges, (d): signaux estimés*

**IV.10: Conclusion**

Nous avons prouvé par simulation que le problème de la séparation de sources peut être résolu en deux façons, l'une par les techniques algébriques et l'autre par les techniques adaptatives, bien que les deux techniques atteignent la séparation, mais chacune d'elles a ses avantages et ses inconvénients. Pour les méthodes hors ligne l'application de la formule mathématique est simple et donne de bons résultats et l'exécution est rapide, mais pour séparer un autre nombre de sources la séparation est atteinte que par l'application d'une autre formule calculée analytiquement et qui est une tâche difficile. Pour cela les algorithmes adaptatifs sont une solution alternative. Avec les méthodes adaptatives, l'application d'un seul modèle et avec des conditions nécessaires on peut atteindre la séparation au delà de deux sources mais ces techniques dépendent de plusieurs facteurs. Lors de l'implémentation des algorithmes on a pu même recouvrer cinq sources, mais il nous a fallu plus de temps et plus d'échantillons, même le gain d'apprentissage a joué un rôle très important pour le nombre de sources à séparer.

## *Conclusion générale*

Nous avons présenté dans ce travail le problème de la séparation de sources, qui a reçu une très grande considération dans beaucoup d'applications du domaine de l'extraction de l'information, et en raison de l'augmentation d'échange d'informations dans les systèmes MIMO, le problème devient plus évident, car le chevauchement et la détérioration des signaux par les canaux de transmission engendrent une ambiguïté de l'information, alors pour extraire l'information utile on doit recourir au traitement aveugle du signal, c'est pour cela on a consacré le premier chapitre pour la représentation des notions de base sur la théorie de probabilité et les statistiques d'ordre supérieur, et qui sont les outils indispensables pour le traitement du signal aveugle, et on a conclu dans ce chapitre, que il est possible de caractériser les processus stochastiques en se basant sur leurs descriptions probabilistes et leurs comportements, après avoir construit leurs histogrammes.

Le deuxième chapitre avait pour but de présenter le problème de la séparation de sources d'une façon simple où on a mentionné quelques critères utilisés pour résoudre le problème, et on a aussi mentionné que le problème se présente sous différentes formes tel que le mélange non linéaire et le mélange convolutif, mais on a focalisé le chapitre que pour le cas du mélange linéaire instantané, et que ce dernier peut être résolu par deux méthodes, les méthodes directes et les méthodes adaptatives.

Le troisième chapitre a été consacré pour présenter quelques algorithmes célèbres pour la résolution du problème de la séparation de sources pour le cas d'un mélange linéaire instantané, et on a exposé les deux techniques et avec différents algorithmes.

Le quatrième chapitre inclut l'implémentation des algorithmes où on a conclu que les techniques se différencient selon la procédure et selon les suppositions concernant le mélange et les signaux sources. Pour les techniques directes elles sont plus rapides et plus performantes par rapport aux techniques adaptatives, mais elles sont limitées par le nombre de sources à extraire, dans notre travail on a pu extraire que deux sources en utilisant les algorithmes qui se basent sur les cumulants croisés, et pour extraire plus que deux sources il nous faut une autre expression algébrique. Contrairement aux méthodes directes, les méthodes adaptatives et qui sont une solution alternative, on a pu extraire jusqu'à cinq sources avec le même algorithme, mais l'inconvénient de ces techniques est qu'elles nécessitent d'avoir plus d'échantillons et de temps pour atteindre une séparation désirable.

Ces conclusions nous ouvrent quelques perspectives concernant l'utilisation des techniques directes et adaptatives pour la résolution du problème de la séparation de sources

- \* une formule générale valable pour l'extraction de deux sources et plus pour le cas des méthodes directes.
- \* amélioration des résultats et réduire la complexité des algorithmes adaptatifs.
- \* résolution du problème pour le cas convolutif avec la réalisation (cas des signaux de la parole)

## ***Bibliographie***

### **A:LIVRES:**

- [1]:Alfioquarteroni, Riccardosacco, Faustosaleri:(*'Méthodes Numériques,algorithmes,analyse et applications'*) (2004)
- [2]Andrzej Cichoki and Shun-ichi Amari :(*'Adaptive Blind signal and image processing'*) (October 2005)
- [3]: Alexander D, Poularikas Zayed M. Ramadan: (*'Adaptive filtering primer with Matlab'*) (2006)
- [4]: Hwei Hsu :(*'Probability,Random variables, and Random processes'*) (1997)
- [5]:Hwei Hsu :(*'analogue and digital communications'*) (2003)
- [6]:Jean-Louis Lacoume, Pierre-Olivier Amblard, Pierre Comon :(*'Statistiques d'ordre supérieur pour le traitement du signal'*) (1997)
- [7]:James V. Stone (*'Independent Component Analysis A Tutorial Introduction'*)
- [8]:Joa~o M. T. Romano,Romis R. de F. Attux,Charles C. Cavalcante,Ricardo Suyama (*'unsupervised signal processing-Channel Equalization and Source Separation'*) (2011)
- [9]: Peyton Z. Peebles Jr.(*'Probability Random Variables and Random Signal Principles'*) (1987)
- [10]: Pierre comon Christian Jutten(*'Hand book of blind source separation- Independent component Analysis and applications'*) (2011)
- [11]: Richard Bronson :(*'theory and problems of Matrix operations'*) (1989)
- [12]: Seymour Lipschutz: (*'Probabilittes cours et problemes'*) (1973)

### **B:TEHSES ET MEMOIRES**

- [13]: Chahinez Chaouchi (*'Méthodes de séparation aveugle de sources non linéaires, étude du modèle quadratique  $2x2$ '*) l'Université Toulouse III - Paul Sabatier (thèse- doctorat-2011)
- [14]: Frederic ABRARD (*'Méthodes de séparation aveugle de sources et applications :Des statistiques d'ordre supérieur `a l'analyse temps-fréquence'*)- l'Université e Paul Sabatier. (thèse- doctorat-2003)
- [15]: Fiad rabah(*'La séparation aveugle de sources par les techniques neuronales'*) Université Hadj Lakhder -batna. (mémoire- magistère-2011)

- [16]: Hicheme GHennioui ('Séparation aveugle de mélanges linéaires et convolutifs de sources corrélées') école doctorale (Sciences et Technologies de l'Information et de la Communication) de Toulon-Var. (thèse- doctorat-2008)
- [17]: Hakim boumaraf ('séparation aveugle de mélanges convolutifs de sources') l'université joseph fourier. (thèse- doctorat-2005)
- [18]: Valeriu Vrabie ('statistiques d'ordre supérieur : applications en géophysique et électrotechnique') l'Université POLITEHNICA de Bucarest.(thèse- doctorat)

### **C: PUBLICATIONS**

- [19]: A. Mansour and C.Jutten. ('A direct solution for blind separation of sources') IEEE transaction on signal processing Vol.44 No 3.MARCH 1996
- [20]: Andrzej Cichoki and Rolf Unbehauen ('Robust neural Networks with on-line learning for Blind identification and blind separation of sources'), IEEE (1057-7122(96)07598-8. 1996
- [21]: Andrzej Cichoki, Włodzimierz Kasprzak, Shun-ichi Amari ('Multi-layer neural networks with a local adaptive learning rule for blind separation of sources signals') international Symposium of nonlinear theory and its applications(Nolta'95) NTA research society of IEICE Tokyo, Japan 1995, pp 61 -65)
- [22]: A.Mansour, C.G Puntonet and N.Ohnishi ('A simple ICA algorithm based on geometrical approach') IEEE 0-august 2001
- [23]: A.mansour and C.jutten ('fourth order criteria for blind sources separation'), IEEE transaction on signal processing Vol 43,No8 August 1995
- [24]: A.Cichoki, Runbehaun, L Moszczyński and E.Rummert ('a new on-line adaptive learning algorithm for blind separation of source signals') Wiley -Teubner 1994, PP 461-471
- [25]: Andrzej Cichoki, Shun ichi Amari, Msaharu Adachi, Włodzimierz Kasprzak(' self adaptive neural network for blind separation of sources') IEEE vol 2,157-160 1996
- [26]: Christian Jutten and Jeanny Herault ('Blind separation of sources: an adaptive algorithm based on neuromimetic architecture') Elsevier Science Publisher B.V 1991
- [27]: Eric Moreau and Odile Macchi ('two novel architectures for the self adaptive separation of signals') laboratoire des signaux et systèmes, CNRS-ESE. 91192 Gif sur Yvette, France et groupement de recherche TDSI
- [28]: Taro Yamaguchi, Kazuyoshi Itoh ('An algebraic Solution to independent component analysis'), Optics Communications. 178 (2000) 59-6

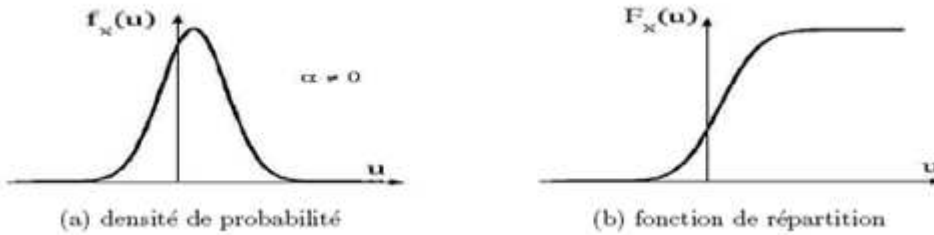
# *Appendice*

**A.1: Variable gaussienne**

soit x est variable gaussienne de moyenne  $\alpha$  et variance  $\sigma^2$

La densité de probabilité  $f_x(\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\pi(\mu-\alpha)^2}{2\sigma^2}}$

Les moments et les cumulants d'ordre 4 sont [18].

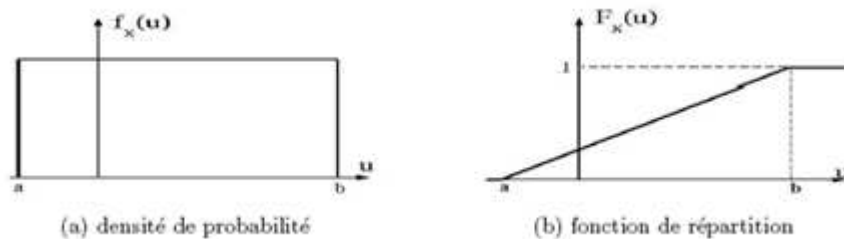


l'ordre r	moment $\mu'_{x(r)}$	moment centré $\mu_{x(r)}$	cumulant $\kappa_{x(r)}$	skewness $\mathcal{K}_{x(3)}$	kurtosis $\mathcal{K}_{x(4)}$
1	$\alpha$	0	$\alpha$		
2	$\alpha^2 + \sigma^2$	$\sigma^2$	$\sigma^2$		
3	$\alpha(\alpha^2 + 3\sigma^2)$	0	0	0	
4	$\alpha^4 + 6\alpha^2\sigma^2 + 3\sigma^4$	$3\sigma^4$	0		0

**A.2: Variable uniforme continue**

Soit x est variable aléatoire uniformément répartie dans l'intervalle [a, b]

La densité de probabilité  $f_x(\mu) = \begin{cases} \frac{1}{b-a} & a \leq u \leq b \\ 0 & u < a, u > b \end{cases}$



Les moments et les cumulants d'ordre 4 sont [18].

l'ordre r	moment $\mu'_{x(r)}$	moment centré $\mu_{x(r)}$	cumulant $\kappa_{x(r)}$	skewness $\mathcal{K}_{x(3)}$	kurtosis $\mathcal{K}_{x(4)}$
1	$\frac{1}{2}(a + b)$	0	$\frac{1}{2}(a + b)$		
2	$\frac{1}{3}(a^2 + ab + b^2)$	$\frac{1}{12}(b - a)^2$	$\frac{1}{12}(b - a)^2$		
3	$\frac{1}{4}(a + b)(a^2 + b^2)$	0	0	0	
4	$\frac{1}{5}(a^4 + a^3b + a^2b^2 + ab^3 + b^4)$	$\frac{1}{80}(b - a)^4$	$-\frac{1}{120}(b - a)^4$		$-\frac{6}{5}$

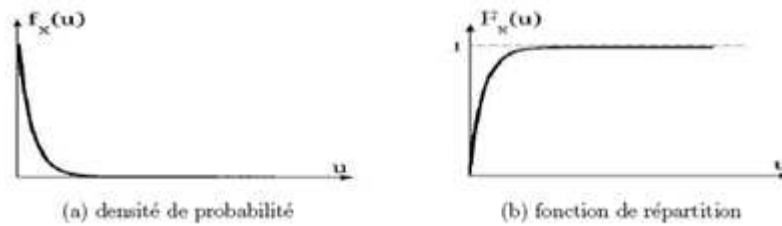


### A.3: Variable exponentielle

Soit  $x$  variable aléatoire positive distribuée selon la loi exponentielle

$$\text{La densité de probabilité } f_x(u) = \begin{cases} \lambda e^{-u\lambda} & 0 \leq u \\ 0 & u < 0 \end{cases}$$

Les moments et les cumulants d'ordre 4 sont [18].



l'ordre	moment	moment centré	cumulant	skewness	kurtosis
$r$	$\mu'_{x(r)}$	$\mu_{x(r)}$	$\kappa_{x(r)}$	$\mathcal{K}_{x(3)}$	$\mathcal{K}_{x(4)}$
1	$\sigma$	0	$\sigma$		
2	$2\sigma^2$	$\sigma^2$	$\sigma^2$		
3	$6\sigma^3$	$2\sigma^3$	$2\sigma^3$	2	
4	$24\sigma^4$	$9\sigma^4$	$6\sigma^4$		6

**B.1: La décomposition LU**

Une matrice carrée A est dite triangulaire supérieur si tous les éléments en dessus de la diagonale principale égaes à zéro; et elle est dite triangulaire inférieure si tous les éléments en dessous de la diagonale principale sont des zéros.

Et tous les éléments restants dans les deux cas peuvent prendre n'import quelle valeur sauf le zéro.

Dans la plupart des cas la matrice A peut être écrite sous la forme de deux matrices triangulaire, l'une inférieure L et l'autre supérieure U [1] [11].

$$A=LU \tag{B.1}$$

Où

$$L = \begin{bmatrix} l_{11} & 0 & 0 & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ l_{31} & l_{32} & l_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & l_{n3} & \dots & l_{nn} \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 1 & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & 1 & u_{23} & \dots & u_{2n} \\ 0 & 0 & 1 & \dots & u_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

**Exemple:**

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & -1 & 1 \\ 4 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 2 & -2 & 0 \\ 4 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

**B.2:Matrice hermitienne**

On définit la transposée hermitienne de la matrice A, par la matrice conjuguée complexe de la matrice A dénotée A<sup>H</sup> [1] [11].

$$A^H = \bar{A}^T \tag{B.2}$$

Si A A<sup>H</sup> = A<sup>H</sup> A on dit que la matrice A est une matrice normale

On dit que A est une matrice hermitienne si A = A<sup>H</sup> (B.3)

Remarque:

Toutes les valeurs propres d'une matrice hermitienne sont réelles [1] [11]

**B.3: Matrice définie positive**

On dit que la matrice A est définie positive seulement si ces valeurs propres sont toutes positives. [1] [11]

**B.4: La décomposition de (Cholesky [1] [11])**

N'importe quelle matrice A définie positive peut être décomposée en cette forme

$$A=LL^H \quad (B.4)$$

Ou L est une matrice triangulaire inferieur et ces éléments de la diagonale sont tous positifs

Les étapes pour avoir cette matrice triangulaire inferieur sont les suivants

Étape 1: Initialisation mettre tous les éléments en dessous de la diagonale à Zéro et poser

$l_{11}=\sqrt{a_{11}}$ , le reste de la de la première colonne resterons les mêmes, et on pose  $j=2$

Étape 2: si  $j = n+1$  (n la dimension de la matrice carré A), on stop l'algorithme

Étape 3: pour  $L_j$  de ( $j=i, j+1, \dots, n$ ) on calcul

$$l_{jj}=\sqrt{a_{jj} - \langle l_j | l_j \rangle} \quad (B.5)$$

Étape 4: si  $j=n$  on passe directement à l'étape 5 si non chaque élément suivant doit etre clculé par cette formule

$$l_{ij}=\frac{a_{ij}-\langle l_i | l_j \rangle}{l_{jj}} \quad (B.6)$$

étape 5: augmenter j par 1 et revenir à l'étape2

# Résumé

## Résumé

Dans ce travail nous avons exposé le problème de la séparation de sources (signaux) par sa forme la plus simple qui est la séparation de sources pour le cas d'un mélange linéaire instantané, et on a présenté les différentes techniques utilisées pour atteindre la séparation, et qui se divisent en deux catégories, ces techniques sont les techniques directes et les techniques adaptatives. Pour chaque catégorie on a implémenté trois algorithmes afin de comparer et de présenter les avantages et les inconvénients de chaque technique.

**Mots clefs:** Statistiques d'ordre supérieur, indépendance statistique, cumulants croisés, processus stochastique, théorème de la limite centrale, histogramme, réseau neuronal, Kurtosis.

## Abstract

In this work we exposed the problem of source (signal) separation by its simplest form, this form is the case when the sources are mixed by a linear instantaneous mixture system, and we presented the various techniques used to achieve the separation and which are divided into two categories, these techniques are the direct techniques and the adaptive techniques, each category has been implemented with three algorithms in order to compare and present the advantages and disadvantages of each technique.

**Key words:** Higher order statistics, statistical independence, cross cumulants, stochastic process, central limit theorem, histogram, neural network, Kurtosis.

## ملخص

في هذا العمل قمنا بعرض مشكلة فصل المصادر (الإشارات) بالشكل البسيط التي تظهر به ألا وهو حالة امتزاج المصادر بطريقة خطية آنية و قدمنا أيضا طرق حل لمشكل فصل المصادر والتي تنقسم إلى قسمين، هذان القسمان هما الطرق المباشرة و الطرق التكيفية، وفي كل قسم قمنا بمحاكاة ثلاث خوارزميات و ذلك لهدف مقارنة القسمين وعرض مساوي ومحاسن كل طريقة

**كلمات مفتاحية:** إحصائيات ذات الرتبة العالية، الاستقلالية الإحصائية، التراكمات المتصالية، العملية العشوائية، نظرية النهاية المركزية، الشبكة العصبية، التفرطح