

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



Mémoire de Fin d'Étude

Présenté à

L'Université Echahid Hamma Lakhdar d'El Oued

Faculté de Technologie

Département de Génie Electrique

En vue de l'obtention du diplôme de

MASTER ACADEMIQUE

En Télécommunications

Présenté par

Ouabri Hicham

Bekkouche Tedjani

Thème

**Rehaussement de signal parole par les méthodes de décomposition
en modes empirique – EMD (Empirical mode décomposition)**

Soutenu le 28 /05/2017. Devant le jury composé de :

Mr. M. Hettiri.

Maitre de conférences Président

Mr. R. Ajgou.

Maitre de conférences Rapporteur

Mr. A. Khelil.

Maitre de conférences Examineur

Remerciements

Je tiens à la fin de ce travail à remercier ALLAH le tout puissant de m'avoir donné la foi et de m'avoir permis d'en arriver là.

Nous adresse nos plus vifs remerciements à toutes les personnes qui nous ont aidés de près ou de loin de l'élaboration de ce modeste mémoire.

Nous tenons à remercier notre encadreur Mr. AJGOU Riadh pour ses conseils et ses directives.

Nos grands remercient aussi à tous les enseignants qui ont contribué à notre formation au département d'électronique.

Nous remercions vivement nos familles OUABRI et BEKKOUCHE et DEGACHI pour leur aide matérielle et moral durant toute la période de préparation.

✓ *OUABRI Hicham*

✓ *BEKKOUCHE Tedjani*

Dédicace

Je dédie ce modeste travail

A mes chers parents

A mes frères et mes sœurs

A toute ma famille

A mon encadreur Mr Ajgou Riadh

Je dédie aussi ce travail à mon collègue

Degachi Riadh.

A tous mes collègues et mes amis.

A tous ceux qui ont sacrifié leur temps pour la science

et à tous ceux qui ont utilisé la science pour le bien

et la prospérité de l'humanité.



Table des matières

Liste des figures.....	VII
Liste des tableaux.....	IX
Liste des symboles et abréviations.....	X
Résumé.....	XI
Introduction générale.....	XIV

Chapitre I Généralité sur le signal parole

I.1 Introduction.....	1
I.2 Qu'est-ce que la parole ?.....	1
I.3 Le signal de parole : de la production à la perception.....	2
I.3.1 Classification des sons de la parole.....	2
I.3.2 La production de la parole.....	2
I.3.3 Production de la parole et voix humaine.....	3
I.3.3.1 Physiologie des organes de la phonation.....	3
I.3.3.2 Les cavités résonantes.....	4
I.3.4 Mécanismes d'audition de la parole.....	4
I.3.4.1 L'appareil auditif périphérique.....	5
I.4 Phonétique et Phonologie.....	5
I.5 Phonétiques articulatoire et acoustique.....	7
I.6 Paramètres du signal de parole.....	8
I.6.1. La fréquence fondamentale.....	8
I.6.2. L'énergie.....	8
I.6.3. Le spectre.....	8
I.7 Analyse spectrale de la parole.....	9
I.8 La technique EMD.....	10
I.8.1 Définition 1.....	10
I.8.2 Définition 2.....	10
I.8.3 IMF (Intrinsic Mode Function).....	11
I.8.4 Principe de la méthode EMD.....	11
I.9 Algorithme de la décomposition modale empirique.....	12
I.9.1 Processus de tamisage : principe algorithmique.....	12

I.10 L'EMD appliquée au rehaussement de la parole.....	17
I.10.1 Méthode générale de réduction de bruit par EMD.....	17
I.11. Conclusion.....	20

Chapitre II Techniques améliorées d'EMD

II.1 Introduction.....	21
II.2 Principe d'EEMD (Ensemble Empirical mode decomposition).....	21
II.3 Annulation du mixage de mode.....	22
II.3.1 Le Choix des paramètres.....	23
II.3.1.1 Rapport signal/bruit (SNR).....	23
II.3.1.2 Nombre d'essai.....	24
II.4 Effet des paramètres EEMD.....	24
II.4.1 Effet de l'amplitude de bruit.....	24
II.4.2 Effet du nombre d'essai N_e	29
II.5 Evaluation de l'erreur par la méthode EEMD.....	34
II.6 Inconvénient de la méthode EEMD.....	37
II.7 La méthode MEEMD.....	38
II.7.1 Algorithme de la méthode MEEMD.....	38
II.7.2 Critères de choix des paramètres de la méthode MEEMD.....	38
II.7.2.1 le choix du type de filtre.....	38
II.7.2.2 Le rapport signal sur bruit.....	39
II.7.2.3 Le nombre d'essais N_e	40
II.7.3 La décomposition par la méthode MEEMD avec un filtre elliptique d'ordre 3.....	40
II.8 La méthode CEEMD.....	41
II.8.1 Algorithme de la méthode CEEMD.....	41
II.8.2 Critères de choix des paramètres.....	42
II.8.2.1 Le rapport signal sur bruit.....	42
II.8.2.2 Le nombre d'essais N_e	42
II.8.3 La décomposition par la méthode CEEMD.....	43
II.9 Les différentes techniques d'évaluation de la qualité de la parole.....	44
II.10 Conclusion.....	45

Chapitre III Application au débruitage	
III.1 Introduction.....	46
III.2 Simulation pour calculer les IMFs	46
III.3 La décompositions signal parole par EMD	56
III.4 Application d'EMD au débruitage	59
III.4.1 Première approche	59
III.4.2 Dèbruitage naïf.....	60
III.4.3 Débruitage par combinaison EMD-Ondelettes.....	62
III.4.3.1 Seuillage dépendant du niveau	63
III.5 Comparaison des deux approches de rehaussement de la parole en termes de PESQ, segSNR, en présence de bruit blanc gaussien... ..	64
III.5.1 En terme d'SNR.....	64
III.5.1 En terme de segSNR	65
III.5.3 En terme de PESQ	66
III.6 Conclusion	66
Conclusion générale.	67
Bibliographie	69

Liste de figure

Figure I.1: Traitement de la parole.....	1
Figure I.2 schéma de l'appareil phonatoire.....	3
Figure. I.3 Les résonateurs principaux.....	4
Figure I.4 Schéma des oreilles externe, moyenne et interne.....	5
Figure I.5 Spectre de puissance d'un segment acoustique de la voyelle [u] (signal échantillonné à 16kHz, fenêtre de Hamming de 32ms).....	10
Figure I.6 Organigramme de l'EMD.....	16
Figure I.7 Visualisation des IMFs 1 a 3 issues de la décomposition par EMD d'un signal de parole propre.....	17
Figure I.8 Visualisation des IMFs 4 a 7 issues de la décomposition par EMD d'un signal de parole propre.....	18
Figure I.9 Détail d'une région voisée des IMFs 1 a 7 issues de la décomposition par EMD d'un signal de parole propre.....	19
Figure I.10 Détail d'une région voisée des IMFs 1 a 7 issues de la décomposition par EMD d'un signal de parole corrompu par du bruit blanc gaussien.....	20
Figure II.1. L'organigramme de la méthode EEMD.....	22
Figure II.2. Le signal $x_1(t)$	23
Figure II.3 Le SNR et le nombre d'IMFs redondantes pour la décomposition du signal $x_1(t)$	23
Figure II.4 Relation entre le coefficient de corrélation et le nombre d'essai N_e	24
Figure II.5 Décomposition EEMD du signal $x_1(t)$ en utilisant un SNR = 37dB et un $N_e = 1000$	25
Figure II.6 Décomposition EEMD du signal $x_1(t)$ en utilisant un SNR = 57 dB et un $N_e = 1000$	26
Figure II.7 Décomposition EEMD du signal $x_1(t)$ en utilisant un SNR = 37 dB et un $N_e = 1000$	27
Figure II.8 Décomposition EEMD du signal $x_1(t)$ en utilisant un SNR = 11 dB et un $N_e = 1000$	29
Figure II.9 Décomposition EEMD du signal $x_1(t)$ en utilisant un SNR = 37 dB et un $N_e = 10$	31
Figure II.10 Décomposition EEMD du signal $x_1(t)$ en utilisant un SNR = 37 dB et un $N_e = 50$	33
Figure II.11 Décomposition EEMD du signal $x_1(t)$ en utilisant un SNR = 37 dB et un $N_e = 200$	33
Figure II.12 Décomposition EEMD du signal $x_1(t)$ en utilisant un SNR = 37 dB et un $N_e = 500$	34
Figure II.13 L'évaluation de l'erreur par la méthode EEMD pour $N_e = 10$	35
Figure II.14 L'évaluation de l'erreur par la méthode EEMD pour $N_e = 100$	35

Figure II.15 L'évaluation de l'erreur par la méthode EEMD pour $N_e = 500$	36
Figure II.16 L'évaluation de l'erreur par la méthode EEMD pour $N_e = 1000$	36
Figure II.17 Le temps de calcul en fonction de N_e pour la décomposition EEMD.....	37
Figure II.18 SNR et nombre d'IMFs redondantes pour la décomposition MEEMD du signal $x_1(t)$	39
Figure II.19 Décomposition MEEMD du signal $x_1(t)$ avec $SNR = 15dB$ et $N_e = 27$	41
Figure II.20 SNR et nombre d'IMFs redondantes pour la décomposition CEEMD du signal $x_1(t)$	42
Figure II.21 Relation entre le coefficient de corrélation et le nombre d'essais N_e	43
Figure II.22 La décomposition CEEMD du signal $x_1(t)$ avec $SNR = 47dB$ et $N_e = 40$	44
Figure III.1 Représentation du signal $x(t)$	46
Figure III.2 Interpole les maxima avec la méthode des splines cubiques	47
Figure III.3 Enveloppe supérieure et inférieure obtenues par interpolation des maxima et minima locaux	47
Figure III.4 Moyenne locale du signal $x(t)$	48
Figure III.5 Processus du tamisage pendant l'extraction de l'IMF1.	52
Figure III.6 Différentes composantes IMF et résidu du signal $x(t)$	55
Figure III.7 La décomposition modale empirique du signal $x(t)$	56
Figure III.8 signal parole originale.	57
Figure III.9 Bruit blanc gaussien	57
Figure III.10 signal parole bruité 5dB	57
Figure III.11 Détail d'une région voisée des IMFs 1 a 23 issues de la décomposition par EMD d'un signal de parole corrompu par du bruit blanc gaussien	59
Figure III.12 Signal d'origine et signal bruité à 5dB.	60
Figure III.13 IMFs de signal bruité	60
Figure III.14 Les IMFs bruité éliminé.	60
Figure III.15 Reconstructions partielles successives du signal bruité.....	61
Figure III.16 Débruitage naïf pour différents SNR en ne gardant que les IMFs de 3 à 10 et éliminer les autres IMFs du signal $x(t)$ bruité.....	62
Figure III.17 Résultat rehaussement par combinaison EMD-Ondelettes.	64
Figure III.18 Comparaison des performances en termes de segSNR en présence du bruit gaussien ($SNR = -5$ à 30 dB par pas de 5 dB)	65
Figure III.19 Comparaison des performances en termes de PESQ en présence du bruit gaussien ($SNR = -5$ à 30 dB par pas de 5 dB)..	66

Liste des Tableaux

Tab I.1 Les symboles de l'alphabet phonétique international utilisés en français	7
Tab I.2 Pseudo-code de l'EMD	14
Tab II.1 Choix de filtre passe bas	39
Tab III.1 Estimation le SNR initial et final pour le signal de parole bruité de la technique Naïf et technique combinaison EMD-Ondelette.....	65

Liste des Symboles et abréviations

EMD	Empirical Mode Decomposition
API	Phonetic Alphabet
EnvMin	The Minima envelope
EnvMax	The Maxima envelope
IMF	Intrinsic Mode Function
FI	Instantaneous Frequency
TH	Hilbert Transform
SD	Sum of the Difference
EEMD	Ensemble Empirical Mode Decomposition
Ne	Test Numbers
MEEMD	Modified Ensemble Empirical Mode Decomposition
RMSE	Root Mean Squared Error
CEEMD	Complementary Ensemble Empirical Mode Decomposition
PESQ	Perceptual Evaluation of Speech Quality
SNR	Signal –to Noise-Ratio
SegSNR	Segmental Signal –to Noise-Ratio

Résumé

L'amélioration de la parole vise à améliorer la qualité de parole, en utilisant divers algorithmes. Le dé-bruitage de signal parole est l'amélioration de l'intelligibilité et la qualité de perception globale du signal dégradé et cela on utilisant différentes techniques. L'objectif de ce travail est le rehaussement de signal parole bruité par les techniques de décomposition en modes empiriques. La décomposition en modes empiriques (en anglais : empirical mode decomposition ou EMD) consiste à décomposer un signal sur une base de fonctions comme le ferait une décomposition en séries de Fourier ou une décomposition en ondelettes. La particularité de l'EMD réside dans le fait que la base de fonctions n'est pas donnée a priori mais est construite à partir des propriétés du signal. Il a été observé que ces fonctions sont, avec une bonne précision, orthogonales entre elles et de somme égale au signal d'origine. La qualité de la parole reconstruite est mesurée avec différentes techniques d'évaluation (Perceptual Evaluation of Speech Quality – PESQ, segSNR,...). Dont, une amélioration du SNR et PESQ implique une meilleure intelligibilité de la parole. Les méthodes discutées dans ce travail ont été évaluée en présence bruit blanc gaussien en utilisant la base de données TIMIT.

Mots clés (4 à 6 au moins) : Décomposition, signaux non-stationnaires, EMD, PESQ, segSNR,

Abstract

Improvement of the speech aims to improve the quality of the sound by using various algorithms. The objective of débruitage of signal speech is the improvement of the intelligibility and quality of total perception of the degraded signal by using various techniques. The objective of this work is the enhancement of noisy speech signal by decomposition techniques in empirical modes. Empirical decomposition (EMD) consists in decomposing a signal on a basis of functions such as a Fourier series decomposition or a wavelet decomposition. The peculiarity of the EMD lies in the fact that the basis of functions is not given a priori but is constructed from the properties of the signal. It has been observed that these functions are, with good precision, orthogonal to one another and equal to the original signal. The quality of reconstructed speech is measured using different evaluation techniques (Perceptual Evaluation of Speech Quality (PESQ, segSNR, ...). Of which, an improvement of the SNR and PESQ implies a better speech intelligibility. The methods discussed in this work were evaluated in the presence Gaussian white noise by using database TIMIT.

Keywords (4 to 6 at least): Decomposition, non-stationary signals, EMD, PESQ, segSNR,

ملخص

يهدف تحسين الصوت إلى تحسين جودته و ذلك باستعمال خوارزميات مختلفة (متعددة) ، و بذلك يتم الحد أو التقليل من التشويش (الضوضاء). الهدف من هذا العمل هو تعزيز إشارة خطاب مشوشة من قبل التحلل التقني في الأساليب التجريبية (EMD) هو تحلل إشارة إلى وظائف قاعدة كما في تحليل سلسلة فورييه أو تحليل الموجات. خصوصية (EMD) هو حقيقة أن الوظائف الأساسية لا تعطى مسبقا ولكن هي التي شيدت من خصائص إشارة. وقد لوحظ أن هذه الوظائف هي، مع دقة جيدة، متعامدة مع بعضها البعض، ومساوية للإشارة الأصلية. يتم قياس نوعية الخطاب المعاد بناؤها باستخدام أساليب التقييم المختلفة (التقييم الإدراكي لجودة الكلام PESQ ، segSNR...). ، وتحسين SNR و PESQ ينطوي على وضوح الكلام أفضل. الأساليب في هذا العمل نوقشت تقييم في وجود الضوضاء البيضاء جاوس. نستعمل في هذه الدراسة قاعدة البيانات TIMIT.

كلمات مفتاحية (4-6 على الأقل): التحلل، وإشارات غير ثابتة، EMD، PESQ، segSNR،

Introduction générale

Le rehaussement de la parole est un domaine du traitement du signal qui prend de plus en plus d'ampleur. En effet, dans un monde où la télécommunication connaît un véritable essor, les technologies se doivent d'être de plus en plus performantes afin de satisfaire au plus grand nombre. Les applications qui nécessitent un rehaussement de la parole sont très nombreuses, la plus évidente étant sans doute celle de la téléphonie mobile, où de nombreux bruits environnants peuvent gêner la qualité et l'intelligibilité du signal de parole transmis.

Il existe à ce jour de nombreuses techniques de rehaussement de la parole. Celles-ci peuvent d'ores et déjà se décliner en deux catégories distinctes. En effet, certaines techniques utilisent plusieurs microphones et sont qualifiées de multi-capteur, tandis que d'autres techniques n'en utilisent qu'un seul et sont alors qualifiées de mono-capteur. Le présent sujet de recherche se situe dans la catégorie des techniques mono-capteurs qui utilisent principalement les propriétés statistiques de la parole et du bruit afin de réduire au mieux le signal de bruit.

Notre travail se divise en trois chapitres :

- ✓ Le premier chapitre : Dans ce chapitre, nous avons présenté la description et les caractéristiques générales du signal de parole, la production de la parole, paramètres du signal de parole, l'EMD appliquée au rehaussement de la parole.

- ✓ Le deuxième chapitre : on a exploré les techniques améliorées d'EMD, EEMD, MEEMD, CEEMD.

Le troisième chapitre : ce chapitre est consacré à la partie de simulation en menant une série de simulations à l'aide du logiciel " Matlab a2011R ". On a simulé quelques techniques de débruitage signal parole de la technique EMD. Les tests d'évaluation objectifs basés sur la mesure du rapport signal sur bruit et ceux subjectifs basés sur la mesure de l'indice de qualité PESQ.

A la fin et avec une conclusion générale et perspectives on termine notre travail.

Généralité sur le signal parole et technique

EMD

Sommaire :

I.1 Introduction

I.2 Qu'est-ce que la parole ?

I.3 Le signal de parole : de la production à la perception

I.4 Phonétique et Phonologie

I.5 Phonétiques articulatoire et acoustique

I.6. Paramètres du signal de parole

I.7 Analyse spectrale de la parole

I.8 La technique EMD

I.9 Algorithme de la décomposition modale empirique

I.10 L'EMD appliquée au rehaussement de la parole

I.11 Conclusion

I.1 Introduction [1]

Le traitement de la parole est une science située au croisement du traitement du signal numérique et du traitement du langage. La parole a la particularité, par rapport aux autres signaux du traitement de l'information, à être produite et perçue instantanément par le cerveau, et pour cela le traitement de la parole tend à remplacer ces fonctions par des systèmes automatiques (figure I.1):

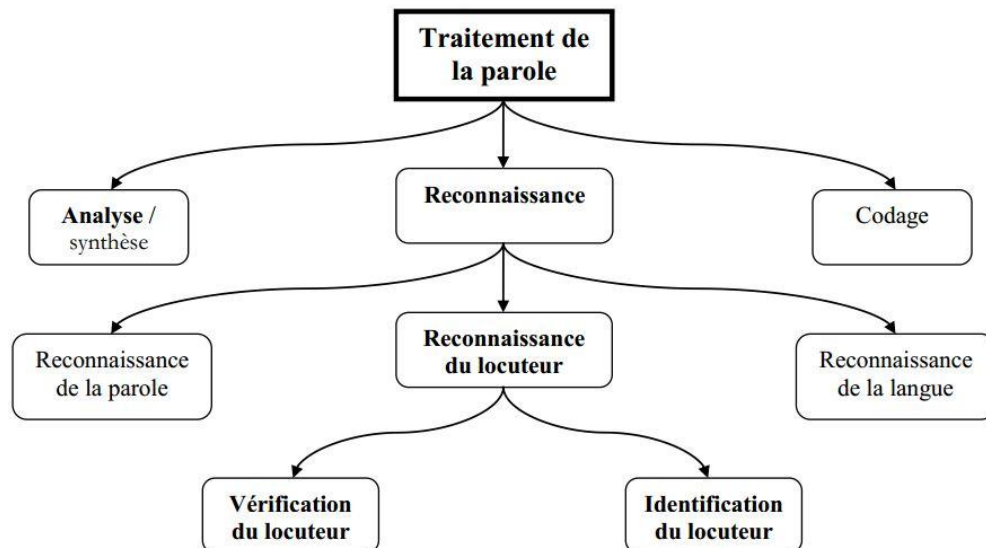


Figure I.1: Traitement de la parole.

Dans ce chapitre, nous présenterons généralités sur le signal de parole, et nous allons parler de technique EMD et aussi parler les techniques de rehaussement de la parole par EMD.

I.2 Qu'est-ce que la parole ?

La parole est un signal continu, d'énergie finie, non stationnaire. Sa structure est complexe et variable dans le temps [1]:

- * Tantôt périodique (plus exactement pseudopériodique) pour les sons voisés,
- * Tantôt aléatoire pour les sons fricatifs,
- * Tantôt impulsionnelle dans les phases explosives des sons occlusifs.

L'information portée par le signal de parole peut être analysée de bien des façons. On en distingue généralement plusieurs niveaux de description non exclusifs : acoustique, phonétique, phonologique, morphologique, syntaxique, sémantique, et pragmatique.

I.3 Le signal de parole : de la production à la perception

La parole s'oppose à la langue par son caractère concret, individuel et créatif. La parole est en effet la réalisation phonétique de la langue résultant d'un acte psychophysiologique et volontaire de la part d'un individu. Aucun mode de communication animal ne peut égaler la complexité du langage parlé. Celui-ci est rendu possible par une anatomie particulière et par l'existence de régions spécialisées dans le cerveau [2].

I.3.1 Classification des sons de la parole

Une décomposition simplifiée du signal de la parole doit ressortir deux types de sons: voisés et non voisés [3].

* **Les sons voisés**, tels que des voyelles, sont produits par le passage de l'air de poumons à travers la trachée qui met en vibration les cordes vocales. Ce mode, qui représente 80% du temps de phonation, est caractérisé en général par une quasi-périodicité, une énergie élevée et une fréquence fondamentale (pitch). Typiquement, la période fondamentale des différents sons voisés varie entre 2ms et 20ms [3].

* **Les sons non voisés**, comme certaines consonnes, dans ce cas les cordes vocales ne vibrent pas, l'air passe à haute vitesse entre les cordes vocales. Le signal produit est équivalent à un bruit blanc [3].

I.3.2 La production de la parole

La parole est un phénomène acoustique qui se distingue des autres sons par des caractéristiques liées aux mécanismes de sa production par l'appareil phonatoire. Ce dernier fait intervenir divers éléments : l'air, comme source d'énergie ; les cordes vocales, comme principal organe vibratoire ; la langue et les lèvres, comme organes vibratoires accessoires ; les cavités buccale et nasale, comme caisses de résonance ; et le système nerveux qui contrôle l'ensemble [4].

Le flux d'air est ensuite modulé en fréquence en passant dans une série d'articulateurs du conduit vocal. Différents sons sont ainsi produits en fonction de la configuration géométrique de ces articulateurs. Parmi ces articulateurs, nous pouvons distinguer la langue, la mandibule, les lèvres et le voile du palais. Ce dernier permet, selon sa position, le couplage ou non du conduit vocal avec les fosses nasales.

Le mode de fonctionnement précédent décrit la production des sons voisés (sonores).

Les sons non-voisés (sourds) sont quant à eux produits quand l'air passe librement à travers la glotte sans vibration des cordes vocales. Ainsi, les fricatives non-voisées

sont produites quand l'air passe à travers la glotte et à travers une constriction du conduit vocal, pour produire un flux d'air turbulent.

Les fricatives voisées combinent en revanche des composantes turbulente et périodique : les cordes vocales vibrent dans ce cas. Pour produire les occlusives, une forte pression d'air est créée en amont d'une occlusion en un point du conduit vocal (palais, dents ou lèvres). Cette occlusion est ensuite brusquement relâchée, causant un flux d'air impulsif et un son transitoire. On distingue aussi les occlusives voisées des occlusives non-voisées [4].

I.3.3 Production de la parole et voix humaine

I.3.3.1 Physiologie des organes de la phonation [5]

Trois groupes d'organes assument les fonctions essentielles dans l'acte de parole, ou phonation (voir la figure I.2) :

- **l'appareil respiratoire**, (diaphragme, poumons, trachées), soufflerie qui fournit l'énergie et la quantité d'air nécessaire.

- **le larynx**, organe vibrant, où naît le son.

- **Le conduit vocal**, formé des cavités résonantes supra-laryngées (pharynx, bouche, nez) où s'effectue l'articulation proprement dite par les changements de forme du tractus vocal. Ces changements résultent surtout des mouvements des lèvres, de la langue, du voile du palais (dont l'abaissement fait intervenir une cavité supplémentaire, les fosses nasales) et de la mâchoire inférieure.

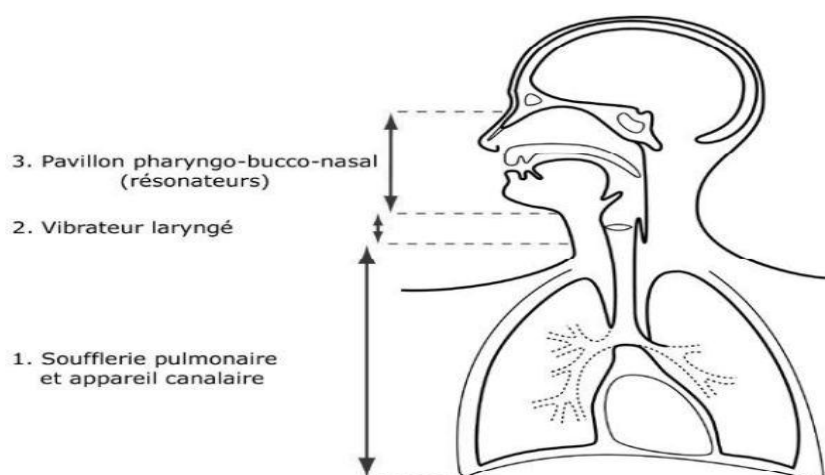


Figure I.2 schéma de l'appareil phonatoire

I.3.3.2 les cavités résonantes

La majorité des sons du langage sont le fait du passage d'une colonne d'air venant des poumons, qui traverse un ou plusieurs résonateurs de l'appareil phonatoire (voir la figure I.3).

Les résonateurs principaux sont :

- * le pharynx ;.....1
- * la cavité buccale ;.....2
- * la cavité labiale ;.....3
- * les fosses nasales.....4

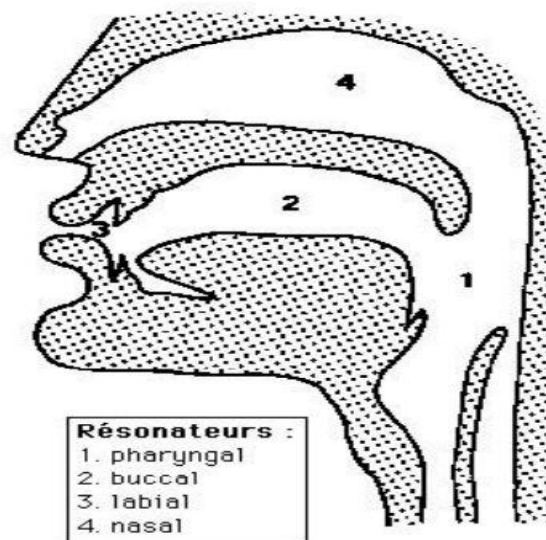


Figure. I.3 Les résonateurs principaux [5].

I.3.4 Mécanismes d'audition de la parole

La parole est un vecteur de transmission d'information d'une grande complexité. En tant que récepteur de ce vecteur, l'appareil auditif de l'être humain se caractérise par une grande finesse d'analyse de cette complexité et par une grande robustesse à l'environnement. Pour cette raison, de nombreux systèmes de traitement de la parole tentent de reproduire les fonctionnalités de cet appareil.

Les mécanismes physiologiques qui permettent l'audition d'un message oral sont classiquement séparés en deux parties : l'appareil auditif périphérique et le système auditif central. Dans ce qui suit, nous présentons succinctement l'appareil l'auditif périphérique chez l'être humain pour introduire la description d'importantes propriétés perceptives du système auditif en relation avec la psycho acoustique.

I.3.4.1 L'appareil auditif périphérique

Le système auditif périphérique est composé de trois parties qui sont l'oreille externe, l'oreille moyenne et l'oreille interne comme le montre le schéma de la Figure I.4 Chacun de ces trois sous-systèmes assure une étape dans la transmission du signal:

- l'oreille externe est responsable de la transmission aérienne au travers du conduit auditif externe ;
- l'oreille moyenne assure, au moyen des trois osselets (marteau, enclume, étrier), la transmission mécanique du tympan jusqu'à la fenêtre ovale ;
- l'oreille interne, quant à elle, permet la transmission hydromécanique au niveau de la membrane basilaire ainsi que la transmission électrochimique au niveau des cellules ciliées de l'organe de Corti.

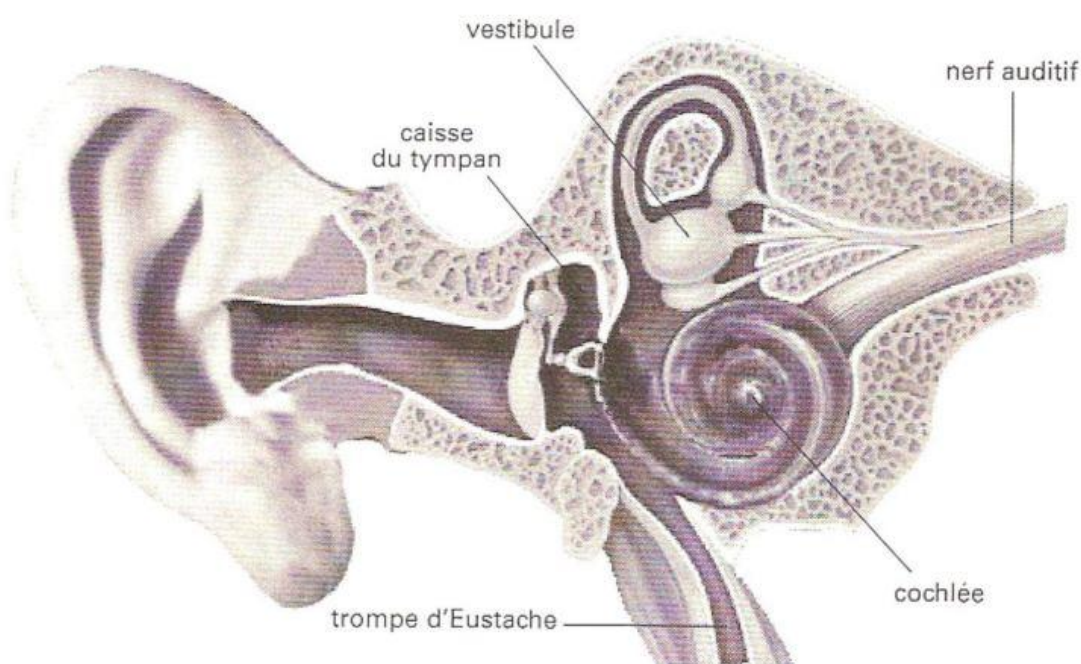


Figure I.4 Schéma des oreilles externe, moyenne et interne [6].

I.4 Phonétique et Phonologie

La phonétique et la phonologie sont deux branches de la linguistique qui interprètent le même matériau : la parole. La phonétique étudie les sons des langues du monde en tant que réalité physique (production, transmission et perception de ces sons), tandis que la phonologie recherche les principes qui régissent leur apparition et leur fonction de codage d'une langue particulière [2] [7].

En fonction de la configuration géométrique du conduit vocal, différents sons sont produits). Chacun des articulateurs du conduit vocal peut en effet prendre un nombre

considérable de positions. Le passage d'une position à une autre ne se fait pas de manière abrupte mais sur un continuum. Mais seulement un nombre restreint de configurations de l'appareil phonatoire sont utilisées linguistiquement pour la production des sons de parole que sont les phones.

La phonologie utilise le principe de regroupement des phones en fonction de leur capacité de différenciation linguistique pour une langue donnée. Ce principe conduit à la définition de classes de phones ayant la même fonction linguistique : les phonèmes [8].

L'unité de codage linguistique de la phonologie est en effet le phonème. Selon la théorie des traits distinctifs, chaque phonème se caractérise par un ensemble de traits qui le différencie des autres phonèmes. Ces traits distinctifs reflètent des propriétés de nature acoustico-auditive ou articulatoire [8] [9].

Le phonème est une unité phonologique distinctive et commutable. Exemples : Pour le Français, la substitution de la consonne occlusive labiale sonore du mot bain par une consonne occlusive labiale sourde (pain) provoque un changement de sens. Les consonnes (b) et (p) sont donc deux phonèmes différents. En revanche, la substitution d'une consonne (R) grasseyée (dorsale) par une consonne (r) roulée (apicale) dans un mot comme renard ne change pas le sens de ce mot. Les consonnes (R) et (r) sont donc deux variantes d'un unique phonème. La phonétique articulatoire peut donc être définie comme l'étude de l'articulation des phonèmes. Un phonème peut en effet avoir plusieurs variantes articulatoires (phonétiques), appelées les allophones.

L'apparition de ces variantes est due au phénomène de coarticulation : comme chaque articulateur du conduit vocal évolue de façon continue, les mouvements articulatoires peuvent être modifiés, suivant le contexte et sous l'effet de l'inertie mécanique de ces articulateurs, de manière à minimiser l'effort à produire pour réaliser une séquence de phonèmes donnée. Les consonnes liquides par exemple, se caractérisent par une forte sensibilité à la coarticulation d'où une grande variabilité acoustique.

Les réalisations physiques d'un phonème peuvent aussi varier en fonction d'autres facteurs tels que le locuteur et le dialecte.

La transcription phonétique d'un énoncé de parole consiste à noter la séquence des symboles représentant la séquence des sons phonétiques constituant cet énoncé. La représentation orthographique d'un énoncé n'est pas en effet suffisamment indicatrice de sa prononciation phonétique (Exemple : il a eu un peu...).

La transcription phonétique doit utiliser des symboles distincts pour tous les sons qui assurent des distinctions linguistiques. Mais la finesse de la notation peut être beaucoup plus grande en recensant les allophones des différents phonèmes.

L'API propose un alphabet phonétique qui associe des symboles aux sons phonétiques de la parole (définis par une description articulatoire) de façon à permettre la notation des prononciations des mots de chaque langue (voir par exemple le tableau I.1) , l'alphabet phonétique utilisé par l'équipe "synthèse de parole" de

France Télécom R&D pour transcrire phonétiquement le Français et les symboles phonétiques API équivalents).

API	EXEMPLES	API	EXEMPLES
i	idée, ami	p	patte, repas, cap
e	ému, ôté	t	tête, ôter, net
ɛ	perdu, modèle	k	carte, écaille, bec
a	alarme, patte	b	bête, habile, robe
ɑ	bâton, pâte	d	dire, rondeur, chaude
ɔ	obstacle, corps	g	gauche, égal, bague
o	auditeur, beau	f	feu, affiche, chef
u	coupable, loup	s	sœur, assez, passe
y	punir, élu	ʃ	chanter, machine, poche
ø	creuser, deux	v	vent, inventer, rêve
œ	malheur, peur	z	zéro, raisonner, rose
ə	petite, fortement	ʒ	jardin, manger, piège
ẽ	peinture, matin	l	long, élire, bal
ã	vantardise, temps	R	rond, chariot, sentir
õ	rondeur, bon	m	madame, aimer, pomme
œ̃	lundi, brun	n	nous, punir, bonne
j	piétiner, briller	ɲ	agneau, peigner, règne
w	oui, fouine	ŋ	jumping, smoking
ɥ	huile, nuire	h	hop (exclamations)

Tab I.1 Les symboles de l'alphabet phonétique international utilisés en français [10].

I.5 Phonétiques articulatoire et acoustique

Les sons de parole se distinguent en deux grandes classes en fonction de leur mode articulatoire : les voyelles et les consonnes.

Les voyelles : La caractéristique articulatoire principale des voyelles est le libre passage de l'air dans les cavités résonantes supra glottiques : buccale, labiale et nasale. Si le voile du palais est abaissé, l'air traverse le résonateur nasal outre le résonateur buccal. Si les lèvres sont projetées vers l'avant, il se forme un résonateur labial à la sortie du conduit buccal. Suivant les résonateurs mis en jeu, nous pouvons opposer les classes de voyelles suivantes :

- * les voyelles orales aux voyelles nasales.
- *les voyelles labialisées (arrondies) aux voyelles non-labialisées.

Les consonnes : Les consonnes se caractérisent d'un point de vue articulatoire par le passage complètement ou partiellement obstrué de l'air, en un ou plusieurs endroits du conduit vocal.

Suivant leur mode d'articulation, on peut décomposer les consonnes du Français en trois classes : les occlusives, les fricatives et les sonantes [11].

I.6. Paramètres du signal de parole

Le signal vocal est généralement caractérisé par trois paramètres: sa fréquence fondamentale, son énergie et son spectre.

I.6.1. La fréquence fondamentale

Elle représente la fréquence du cycle d'ouverture/fermeture des cordes vocales. Cette fréquence caractérise seulement les sons voisés, elle peut varier [1] :

- De 80Hz à 200Hz pour une voix masculine,
- De 150Hz à 450Hz pour une voix féminine,
- De 200Hz à 600Hz pour une voix d'enfant.

I.6.2. L'énergie

Elle est représentée par l'intensité du son qui est liée à la pression de l'air en amont du larynx. L'amplitude du signal de la parole varie au cours du temps selon le type de son, et son énergie dans une trame est donnée par :

$$E = \sum_{n=0}^{N-1} S^2(n) \quad (I.1)$$

I.6.3. Le spectre

L'enveloppe spectrale ou spectre représente l'intensité de la voix selon la fréquence, elle est généralement obtenue par une analyse de Fourier à court terme. La quasi stationnarité du signal de parole permet de mettre en œuvre des méthodes efficaces d'analyse et de modélisation utilisées pour le traitement à court terme du signal vocal sur des fenêtres de durée généralement comprise entre 20ms et 30ms appelées trames, avec un recouvrement entre ces fenêtres qui assure la continuité temporelle des caractéristiques de l'analyse.

I.7 Analyse spectrale de la parole

Propriétés fréquentielles des segments phonétiques de la parole. Cette analyse passe par la transformée de Fourier discrète à court terme de ce signal.

Ce traitement comporte principalement les étapes suivantes :

- Un segment de 20 à 30ms de parole est extrait du signal. Ce segment, appelé trame acoustique, est constitué de N échantillons de parole $\{s(0), \dots, s(N-1)\}$.

- Pour atténuer les distortions spectrales introduites par l'extraction de la trame du signal de parole, on pondère les échantillons de cette trame par une fonction (fenêtre) de pondération. La fenêtre de Hamming par exemple est définie comme suit [11] :

$$w(n) \begin{cases} 0.54 - 0.46 \cos \frac{2\pi n}{N-1} & 0 \leq n < N \\ 0 & \text{ailleurs} \end{cases} \quad (I.2)$$

On calcule enfin la transformée de Fourier discrète des échantillons de la trame pondérée [11]:

$$s(k) = s\left(f = \frac{k}{N}\right) = \sum_{n=0}^{N-1} s(n) w(n) \exp\left(\frac{-j2\pi nk}{N}\right); 0 \leq k < N \quad (I.3)$$

Le spectre de puissance de la transformée de Fourier est donné par :

$$S(k) = |s(k)|^2; 0 \leq k < \frac{N}{2} \quad (I.4)$$

La figure I.5 représente le spectre de puissance d'un segment stable d'une réalisation acoustique de la voyelle (u), sur lequel nous avons mis en évidence la structure formantique de cette voyelle.

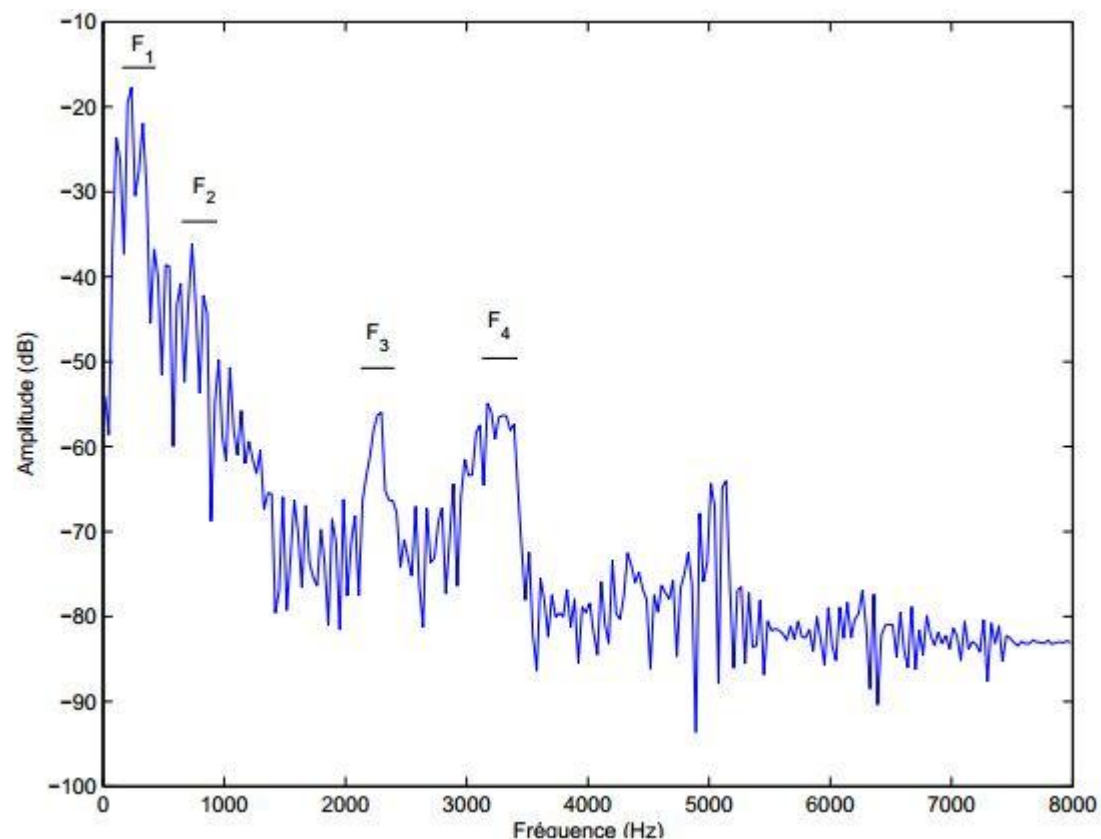


Figure I.5 Spectre de puissance d'un segment acoustique de la voyelle [u] (signal échantillonné à 16kHz, fenêtre de Hamming de 32ms).

I.8 La technique EMD

I.8.1 Définition 1

La décomposition modale empirique (Empirical Mode Decomposition "EMD" en anglais) est une méthode caractérisée par un processus appelé Tamisage (Sifting) permettant de décomposer temporellement un signal en une somme de composantes oscillantes appelées Modes Empiriques connues sous le nom de Intrinsic Mode Functions (IMF).

I.8.2 Définition 2

L'idée de base de l'EMD est de décomposer un signal en une somme de composantes notées IMF (Intrinsic Mode Function) plus un résidu. L'addition de toutes les IMF avec le résidu permet de reconstruire le signal original sans perte d'information. L'extraction des IMF est une opération non linéaire, mais leur recombinaison pour la reconstruction du signal est linéaire.

L'EMD décompose un signal d'une façon auto-adaptative en une somme de composantes oscillantes. Le principe de l'EMD est défini par un algorithme (voir tableau I.2) qui ne possède pas de fondement théorique. Cette méthode a été

appliquée dans des domaines très variés tels que l'océanographie, la sismologie, la biologie et plus généralement à des signaux non stationnaires [12].

I.8.3 IMF (Intrinsic Mode Function)

Un IMF est défini comme une fonction qui remplit les conditions suivantes:

- * Dans l'ensemble des données, le nombre d'extrema et le nombre de passages à zéro doivent être égaux ou diffèrent au plus d'un.
- * À tout moment, la valeur moyenne de l'enveloppe définie par les maxima locaux et l'enveloppe définie par les minima locaux est égale à zéro.

I.8.4 .Principe de la méthode EMD

L'EMD est une méthode algorithmique de décomposition des signaux. Elle se base sur le principe de décomposer le signal en une somme d'une composante locale haute fréquence (oscillation rapide) et d'une composante basse fréquence (tendance). Ce principe est illustré par l'équation (I.5):

$$x(t) = d(t) + m(t) \quad (I.5)$$

Où $x(t)$ constitue le signal à décomposer, $d(t)$ est l'oscillation rapide, $m(t)$ est le signal tendance et t indique le temps discret. De même le signal tendance peut être aussi décomposé en deux termes (I.6).

$$m(t) = d_1(t) + m_1(t) \quad (I.6)$$

Où $d_1(t)$ est la composante haute fréquence et $m_1(t)$ est la composante basse fréquence.

Pour calculer un mode relatif à un signal, on suit le principe suivant :

1. Identifier tous les extrema locaux de $x(t)$.
2. Interpoler les minima (resp. les maxima) de manière à construire une certaine enveloppe: EnvMin (resp. EnvMax).
3. Calculer la moyenne de deux enveloppes $m(t) = (\text{EnvMin}(t) + \text{EnvMax}(t))/2$.
4. Extraire le détail $d(t) = x(t) - m(t)$. Le signal $d(t)$ n'est considéré *IMF* qu'après un certain nombre d'itérations nécessaires afin que $d(t)$ obéisse à un critère d'arrêt donné.

En itérant ce principe, on obtient une décomposition du signal décrite comme suit [13] :

$$x(t) = \sum_{j=1}^N IMF_j(t) + r(t) \text{ avec } N \in \mathbb{N}^* \quad (I.7)$$

Où IMF_j est l'IMF d'ordre j qui est de type plus haute fréquence que l' IMF_{j+1} .

Le signal $r(t)$ est appelé résidu, il correspond à la composante la plus basse fréquence du signal.

D'après (I.7) et en supposant que N est fini, on reconstruit linéairement le signal original sans perte ou distorsion de l'information [14].

Toutefois, on ne parle d'une IMF que si elle vérifie les critères suivants [14]:

- a) Elle est de moyenne nulle.
- b) Les nombres d'extrema et de passages à zéro diffèrent au plus de un (en d'autres termes, cela signifie qu'entre un minimum et un maximum successif, une IMF passe par zéro).
- c) Elle suit une loi de modulation en amplitude et en fréquence (comportement oscillant) naturellement de type mono-composante.

La condition (a) est liée à la contrainte d'estimation de la fréquence instantanée (FI) d'un signal mono-composante avec la transformée de Hilbert (TH). La condition (b) repose sur la définition de la bande étroite d'un signal mono-composante. La dernière condition (c) est une conséquence due au processus d'estimation de l'IMF [12].

I.9 Algorithme de la décomposition modale empirique

La décomposition en mode empirique peut être décrite par l'organigramme de la figure (I.6). Les étapes durant lesquelles les IMF sont extraits forment un processus dit tamisage (Sifting Process). Il est important de noter qu'une condition d'arrêt est nécessaire pour arrêter le processus du tamisage [15].

Conceptuellement, la mise en place de cette méthode est assez simple :

- ♦ Il suffit de considérer le signal à l'échelle de ses oscillations locales, de soustraire l'oscillation la plus rapide et de réitérer le processus sur le résidu de manière à écrire le signal $x(t)$ comme une combinaison finie d'oscillations.

I.9.1 Processus de tamisage : principe algorithmique

Les différentes étapes du procédé de tamisage peuvent s'écrire sous la forme du pseudo-code développé dans le tableau I.2. L'organigramme de la méthode EMD est représenté dans la figure I.6. Il est important de mentionner que le tamisage est répété plusieurs fois (i) jusqu'à ce que $h_{j,i}(t)$ vérifie les conditions nécessaires d'une IMF. Nous commentons les différentes étapes du pseudo-code (Tableau I.2).

De manière plus littérale, l'extraction des IMFs suit le schéma général suivant (Tab. I.2).

- * Trouver les extrema locaux (maximum et minimum) du signal (étape: 3b).
- * Estimer les enveloppes supérieures et inférieures par interpolation respective des maxima et minima locaux (étape : 3c). L'interpolation utilisée dans ce cas est basée sur les splines cubiques [14], [16].
- * Estimer l'enveloppe moyenne locale à partir des enveloppes supérieures et inférieures (étape : 3d).

Etape 1) Fixer $\varepsilon, j \leftarrow 1$ (*jème IMF*) $\varepsilon = 0.3$

Etape 2) $r_{j-1}(t) \leftarrow x(t)$ (résidu)

Etape 3) Extraire la *jème IMF* :

(a) $h_{j,i-1}(t) \leftarrow r_{j-1}(t), i \leftarrow 1$, (*i*, itération de la boucle de sifting)

(b) Extraire les maxima et minima locaux de $h_{j,i-1}(t)$

(c) Calculer les enveloppes supérieure et inférieure : $U_{j,i-1}(t)$ et $L_{j,i-1}(t)$
par interpolation (splines cubiques)
avec les maxima et minima locaux de $h_{j,i-1}(t)$ respectivement.

(d) Calculer l'enveloppe moyenne :

$$\mu_{j,i-1}(t) \leftarrow \frac{U_{j,i-1}(t) + L_{j,i-1}(t)}{2}.$$

(e) Mettre à jour : $h_{j,i}(t) \leftarrow h_{j,i-1}(t) - \mu_{j,i-1}(t), i \leftarrow i + 1$.

(f) Calculer le critère d'arrêt (par exemple) :

$$SD(i) = \sum_{t=0}^T \frac{|h_{j,i-1}(t) - h_{j,i}(t)|^2}{(h_{j,i-1}(t))^2},$$

où T représente le nombre d'échantillons du signal.

(g) Décider : répéter l'étape (b)-(f) tant que $SD(i) < \varepsilon$

et alors mettre : $IMF_j(t) \leftarrow h_{j,i}(t)$ (*jème IMF*)

Etape 4) Mettre à jour le résidu : $r_j(t) \leftarrow r_{j-1}(t) - IMF_j(t)$

Etape 5) Répéter l'étape 3 avec $j \leftarrow j + 1$ jusqu'à ce que le nombre d'extrema dans $r_j(t)$ soit inférieur à 2.

Tab I.2 Pseudo-code de l'EMD [17].

Notations [13]:

ε : indique le seuil prédéfinie, c'est un critère de condition de la boucle indicée par i .

j : représente l'indice de l'IMF.

i : constitue l'indice de l'itération appliquée sur le résidu pour vérifier le critère d'une IMF.

r_j : désigne le résidu après l'obtention de la j^{eme} IMF

$h_{j,i}$: c'est une variable intermédiaire de calcul qui prend la valeur du nouveau résidu a la première itération, puis, elle prend la différence entre le résidu et la valeur de l'enveloppe moyenne aux itérations suivantes.

$U_{j,i}$: représente l'enveloppe supérieure de $h_{j,i}$, construite par interpolation des maxima.

$L_{j,i}$: représente l'enveloppe inférieure de $h_{j,i}$, construite par interpolation des minima.

$\mu_{j,i}$: désigne l'enveloppe moyenne, obtenu à partir des deux enveloppes de $h_{j,i}$.

SD (i) : indique le critère d'arrêt à la i^{eme} itération

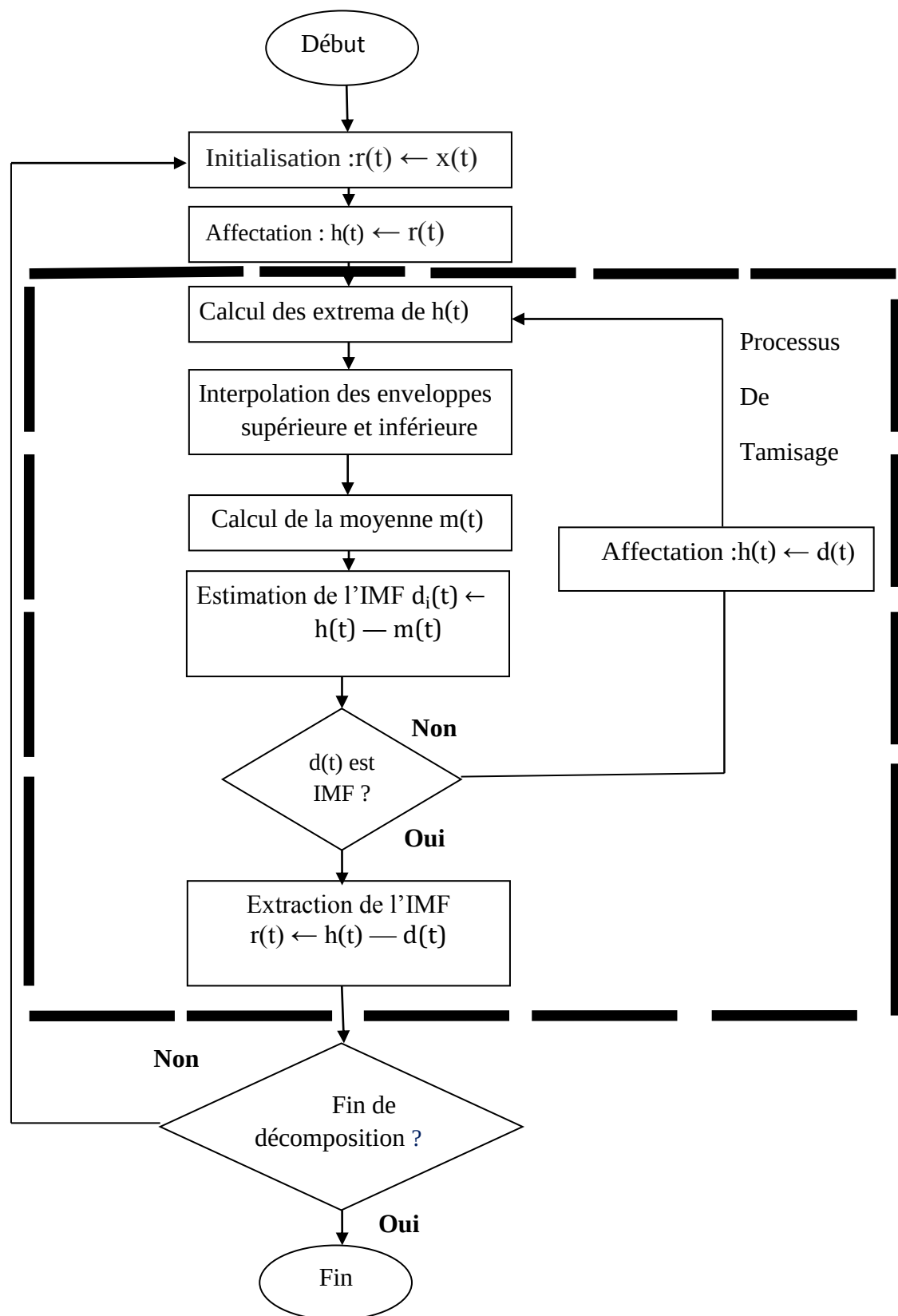


Figure I.6 Organigramme de l'EMD [18].

I.10 L'EMD appliquée au rehaussement de la parole

I.10.1 Méthode générale de réduction de bruit par EMD

Soit un signal de parole propre émit par une femme et qui exprime la phrase 'Ce bonbon contenait trop de sucre' La décomposition par EMD de ce signal de parole est réalisée dans le but d'illustrer les IMFs issues (voir les figures I.7 et I.8) d'un signal de parole propre. La décomposition tamise 10 IMFs et un résidu. Pour des raisons de clarté, seulement les 7 premières IMFs sont représentées [19] :

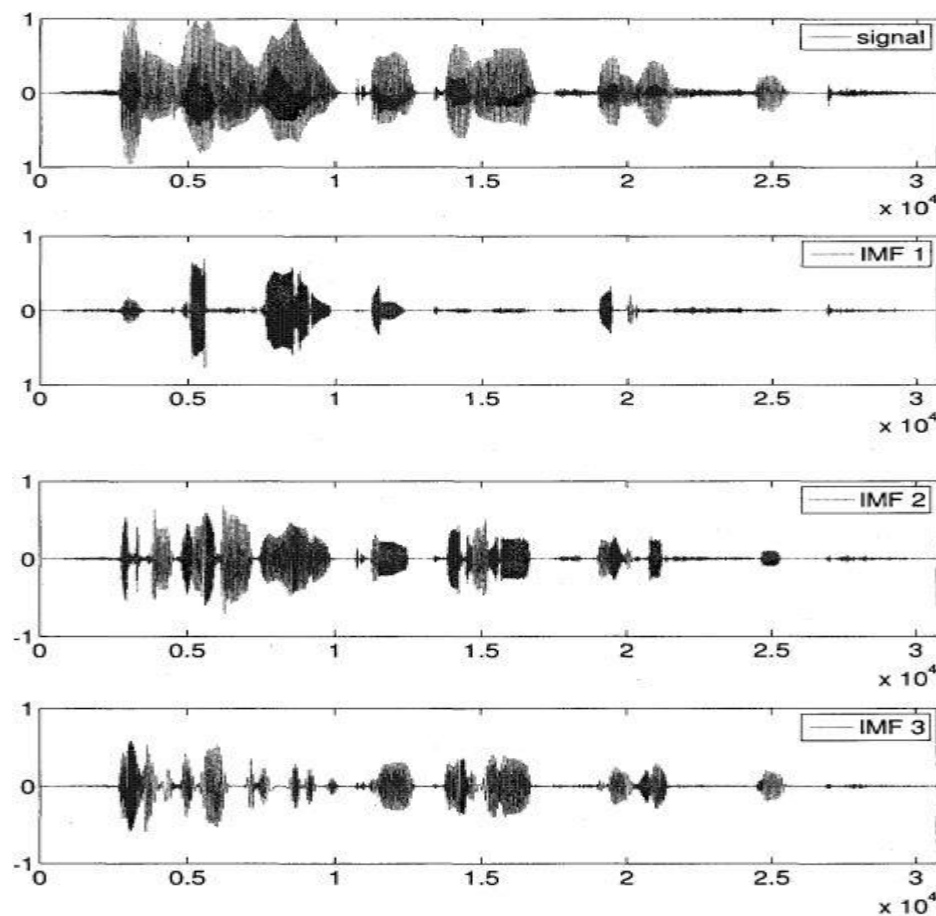


Figure I.7 Visualisation des IMFs 1 à 3 issues de la décomposition par EMD d'un signal de parole propre

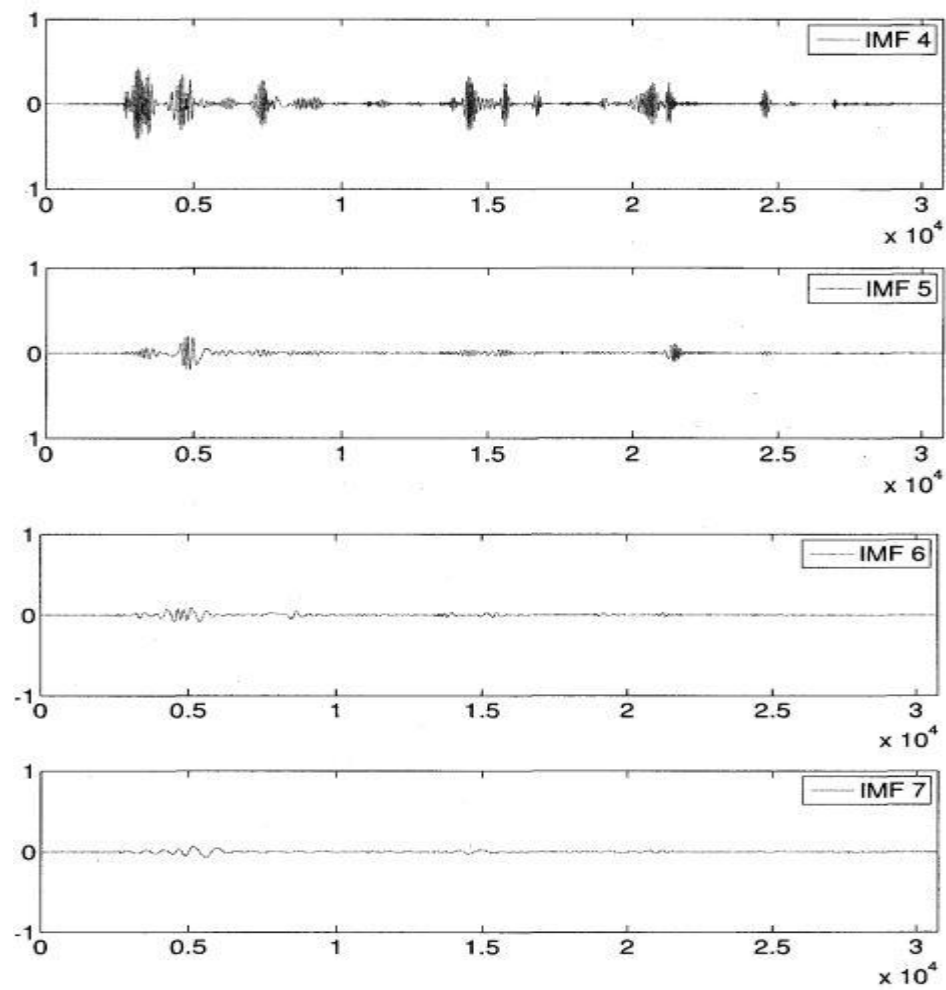


Figure I.8 Visualisation des IMFs 4 a 7 issues de la décomposition par EMD d'un signal de parole propre

La figure I.9 illustre le détail des IMFs dans une région voisée :

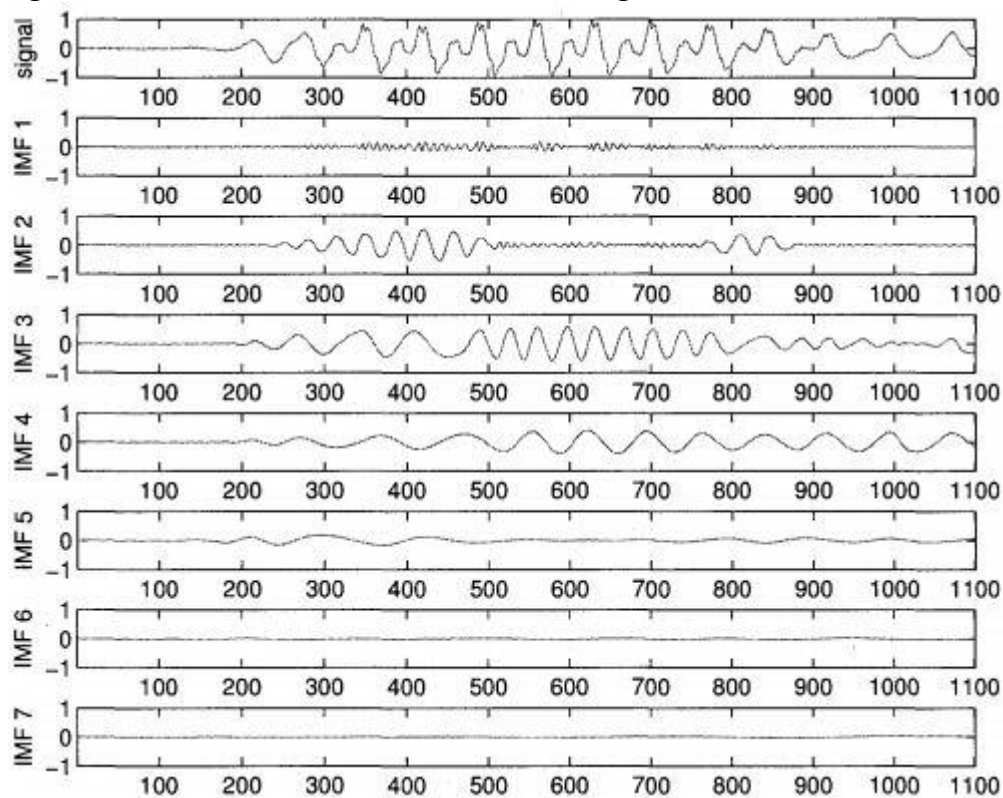


Figure I.9 Détail d'une région voisée des IMFs 1 à 7 issues de la décomposition par EMD d'un signal de parole propre.

Ainsi que le montre la figure I.9, l'EMD tamise d'abord les détails de la région voisée, donc les composantes de hautes fréquences, puis tamise ensuite les composantes de fréquences de plus en plus basses.

Soit maintenant le mélange additif entre ce même signal de parole et un bruit blanc gaussien, avec un rapport signal sur bruit, ou *RSBin* de 10 dB. La décomposition de ce signal par EMD est réalisée et cette fois-ci il en ressort 12 IMFs et un résidu, soit 2 IMFs de plus que lorsque le signal est propre. Pour les mêmes raisons, seulement les 7 premières IMFs sont représentées :

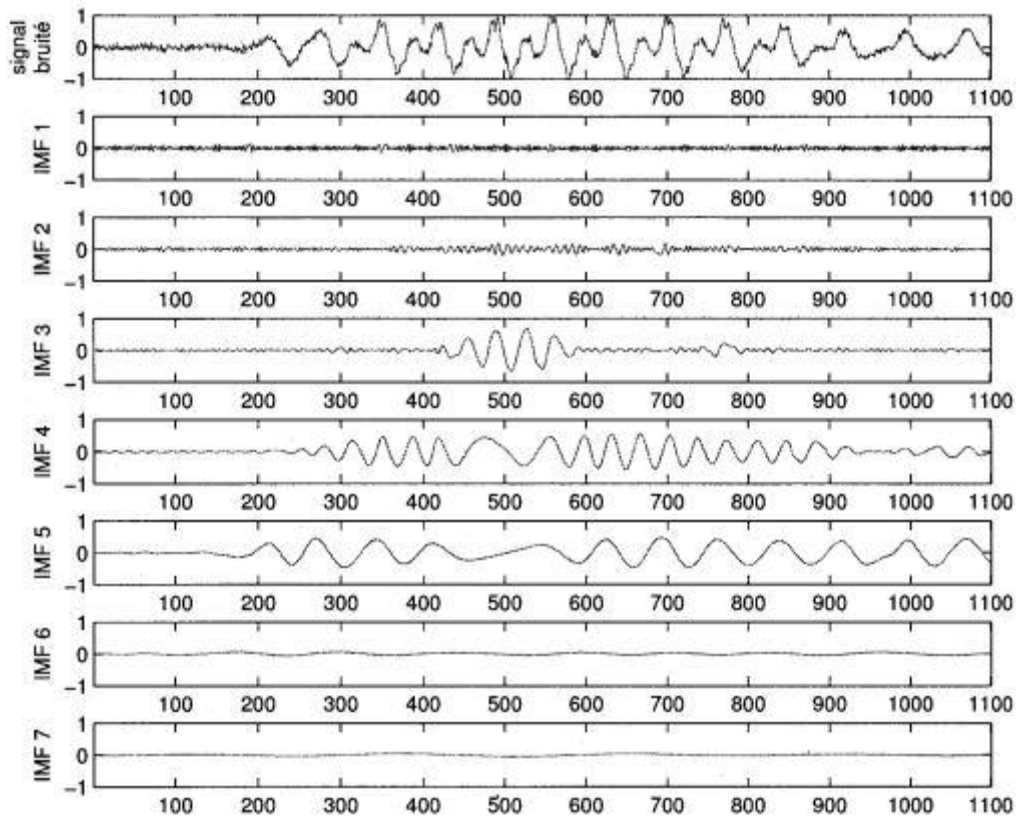


Figure I.10 Détail d'une région voisée des IMFs 1 à 7 issues de la décomposition par EMD d'un signal de parole corrompu par du bruit blanc gaussien.

Ainsi que le montre la figure I.10, l'EMD tamise d'abord les détails de la région voisée, donc les composantes de hautes fréquences, puis tamise ensuite les composantes de fréquences de plus en plus basses.

On remarque que le bruit blanc affecte les IMFs et on peut constater que les deux premières IMFs contiennent surtout du bruit. Cependant, les hautes fréquences de la région voisée du signal de parole illustre à la figure I.10 sont tamisées à partir de l'IMF 2. Ainsi, l'IMF 2 contient des composantes fréquentielles, et du signal de parole, et du bruit blanc.

I.11. Conclusion

Dans ce chapitre, nous avons étudié les caractéristiques d'un signal parole (sa production, modélisation, phonétiques...), ainsi on a étudié la décomposition modale empirique (EMD). Cette méthode de décomposition temporelle et non-linéaire est définie comme la sortie d'un algorithme. Nous avons montré que le signal peut se décomposer en un nombre fini de composantes ou oscillations (IMF) extraites des données signal. Les IMF sont des modes centres et de type AM-FM au sens large. Chaque IMF est obtenue par un processus, appelé « tamisage », itération...

Techniques améliorées d'EMD

Sommaire:

II.1 Introduction

II.2 Principe d'EEMD (Ensemble Empirical mode decomposition)

II.3 Annulation du mixage de mode

II.4 Effet des paramètres EEMD

II.5 Evaluation de l'erreur par la méthode EEMD

II.6 Inconvénient de la méthode EEMD

II.7 La méthode MEEMD

II.8 La méthode CEEMD

II.9 Les différentes techniques d'évaluation de la qualité de la parole

II.10 Conclusion

II.1 Introduction

Pour surmonter le problème de mélange de modes, la méthode EEMD (Ensemble EMD) a été proposée par Wu et Huang en 2009 [20]. L'EEMD est une méthode d'analyse de données assistée par du bruit dans laquelle les IMFs sont définis comme la moyenne d'un ensemble d'essais, chaque est la somme du signal à décomposer et d'un bruit blanc.

Dans ce chapitre nous allons parler des techniques améliorations d'EMD. Ainsi que les principes de ces méthodes.

II.2 Principe d'EEMD (Ensemble Empirical mode decomposition)

La méthode EEMD a été introduite [20] initialement pour résoudre le problème de mélange de modes. Etant donné un signal $x(t)$, son principe est le suivant [21] :

1. On génère N_e réalisations $b_i(t)$, $1 \leq i \leq N_e$, de bruit blanc gaussien de même variance σ^2 .
2. On calcule le signal bruité pour chaque réalisation,

$$s_i(t) = x(t) + b_i(t) \quad 1 \leq i \leq N_e \quad \text{II.1}$$

3. Les N_e réalisations donnent accès à N_e signaux bruités qui permettent l'extraction de N_e jeux de N IMFs : $IMF_{ki}(t)$, $1 \leq k \leq N$ et $1 \leq i \leq N_e$. Les IMFs de la méthode EEMD sont alors les moyennes d'ensemble de ces N_e jeux de N IMFs :

$$IMF_{EEMD_k}(t) = \frac{1}{N_e} \sum_{i=1}^{N_e} IMF_{ki}(t), 1 \leq k \leq N \quad \text{II.2}$$

L'amplitude du bruit blanc gaussien ε est pilotée par le taux d'erreur ε_n au final sur la reconstruction du signal $x(t)$ par sommation des IMFs et par le nombre N_e de moyennes effectuées par la relation [20].

$$\varepsilon_n = \frac{\varepsilon}{\sqrt{N_e}} \quad \text{II.3}$$

L'organigramme de la méthode EEMD est représenté dans la figure 2.1.

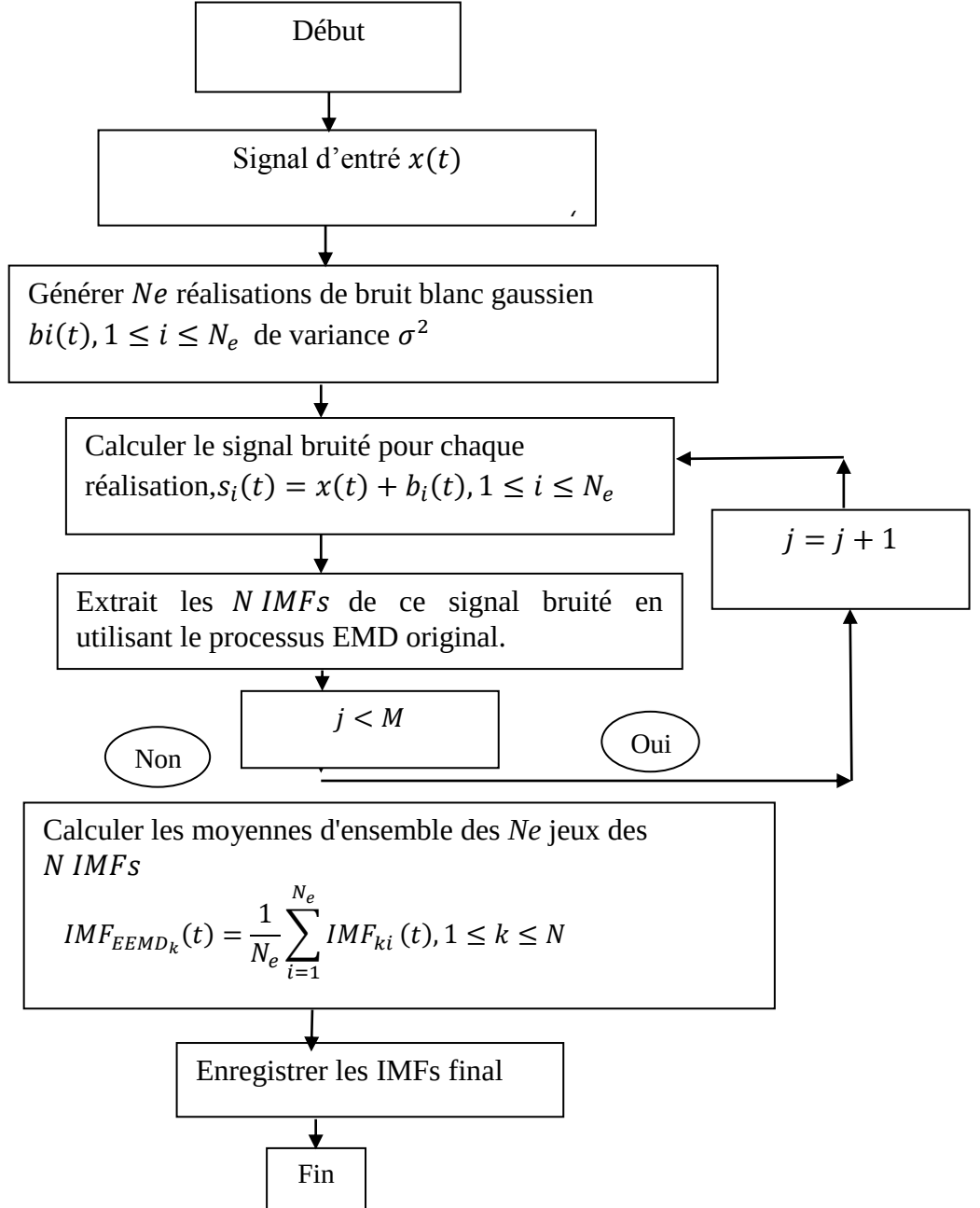


Figure II.1. L'organigramme de la méthode EEMD [17].

II.3 Annulation du mixage de mode

La méthode EEMD résout le problème de mélange de mode, mais l'amplitude du bruit et le nombre d'essai affectent les résultats de la décomposition. Les paragraphes suivants montrent l'effet de chacun des deux paramètres. Pour décomposer un signal par la méthode EEMD, il faut d'abord choisir le niveau du bruit et le nombre d'essai N_e adéquat. Le choix judicieux de ces deux paramètres assurera une décomposition EEMD sans mélange de modes [22] et [21].

II.3.1 Le Choix des paramètres

Nous avons considéré le signal $x_1(t)$ de l'équation (II.6) (401 points et un pas d'échantillonnage de 10^{-4} s) et nous l'avons décomposé par la méthode EEMD.

Il est à rappeler que le signal $x_1(t)$ composé de deux composantes (voir la figure II.2)

$$x_2(t) = \sin(2\pi f_1 t) \quad \text{II. 4}$$

$$x_3(t) = \lambda \sin(2\pi f_2 t) \left[e^{\frac{-(t-t_0)^2}{\sigma}} \right] \quad \text{II. 5}$$

$$x_1(t) = x_2(t) + x_3(t) \quad 0 \leq t \leq 0.04s \quad \text{II. 6}$$

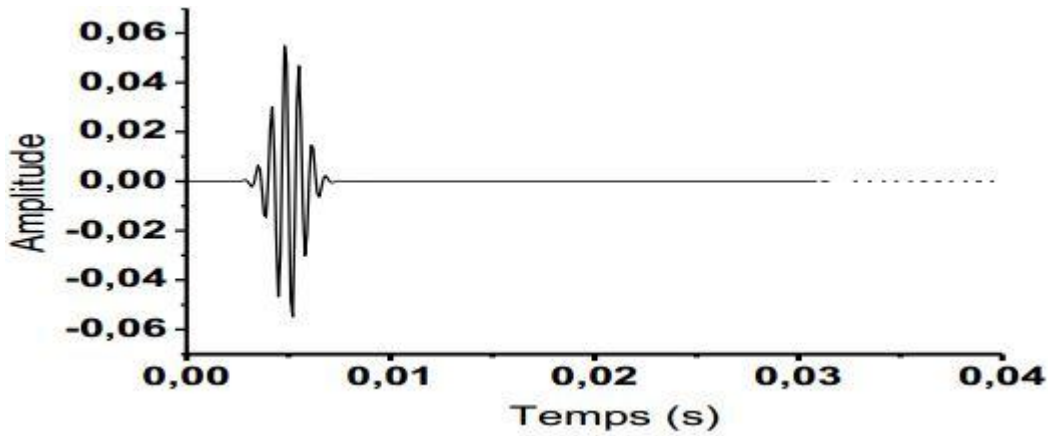


Figure II.2. Le signal $x_1(t)$.

II.3.1.1 Rapport signal/bruit (SNR)

Une approche systématique de la sélection de l'amplitude du bruit est obtenue par le SNR :

$$SNR = 10 \log_{10} \left(\frac{p_1}{p_2} \right) \quad \text{II. 7}$$

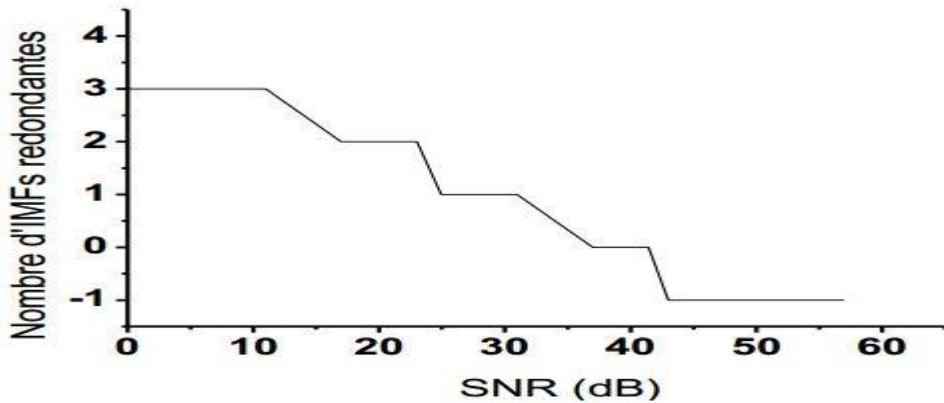


Figure II.3 Le SNR et le nombre d'IMFs redondantes pour la décomposition du signal $x_1(t)$.

Les valeurs des ordonnées correspondent aux différents résultats de décomposition [22]:

- a) Une valeur -1 représente le mélange de mode.
- b) La valeur 0 signifie qu'il n'y a pas de composantes IMF redondantes.
- c) Les valeurs positives représentent le nombre d'*IMFs* redondantes. Pour obtenir une bonne décomposition, il faut bruite le signal avec une puissance déduite du SNR appartenant à la plage de $[37dB - 41dB]$ qui assure la non redondance des *IMFs*.

II.3.1.2 Nombre d'essai

La figure II.4 montre qu'une décomposition EEMD avec un coefficient de corrélation élevé de l'ordre de 95% (moins affecté du bruit) est réalisée avec un nombre d'essai égal à 1000 [21].

Cette figure est obtenue en calculant le coefficient de corrélation entre l' IMF_1 et le signal $x_3(t)$ en fonction du nombre d'essais N_e [22].

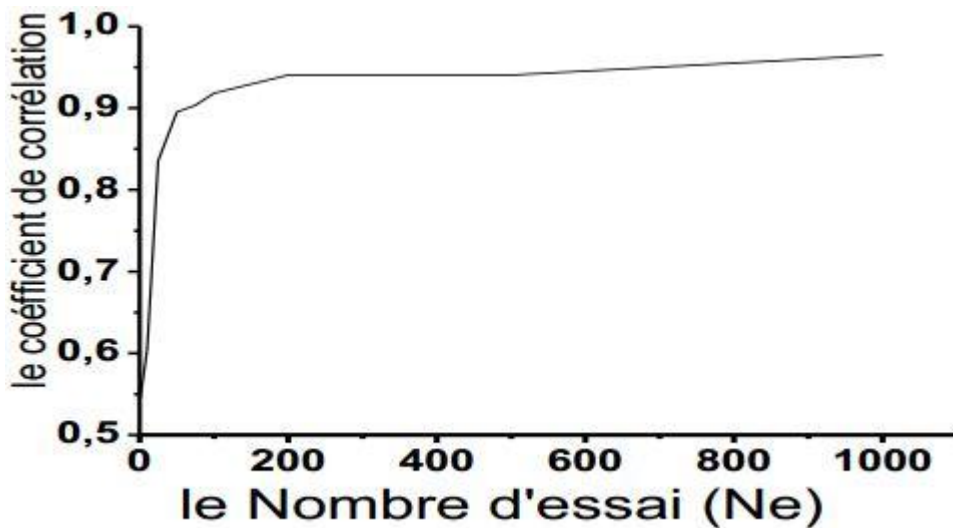


Figure II.4 Relation entre le coefficient de corrélation et le nombre d'essai N_e .

Après avoir fixé les paramètres qui correspondent au meilleur résultat, la figure II.5 représente le résultat de la décomposition EEMD. Et montre clairement l'annulation du problème de mixage de modes [18] [20].

II.4 Effet des paramètres EEMD

II.4.1 Effet de l'amplitude de bruit

Si l'amplitude du bruit est trop petite par rapport à l'amplitude du signal, alors le rajout du bruit n'aura aucun effet sur le mélange de modes. Mais, si l'amplitude du bruit est trop grande

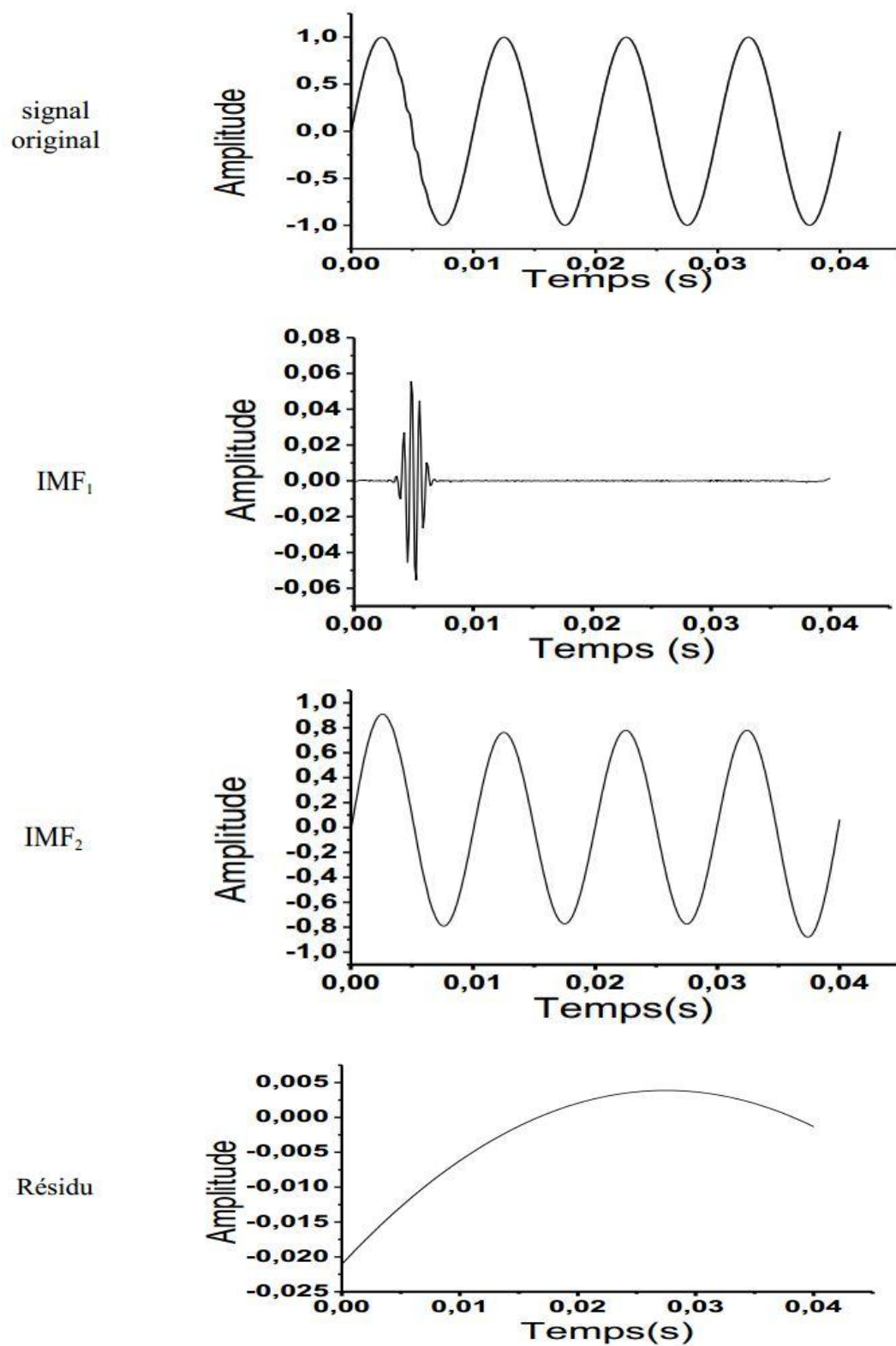


Figure II.5 Décomposition EEMD du signal $x_1(t)$ en utilisant un $SNR = 37dB$ et un $N_e = 1000$ [18].

par rapport à celle du signal, alors la décomposition EEMD donnera des composantes IMFs redondantes.

Les figures II.6, II.7 et II.8 illustrent l'effet de l'amplitude du bruit sur les résultats de la décomposition EEMD du signal test $x_1(t)$, le nombre d'essai a été fixé à 1000.

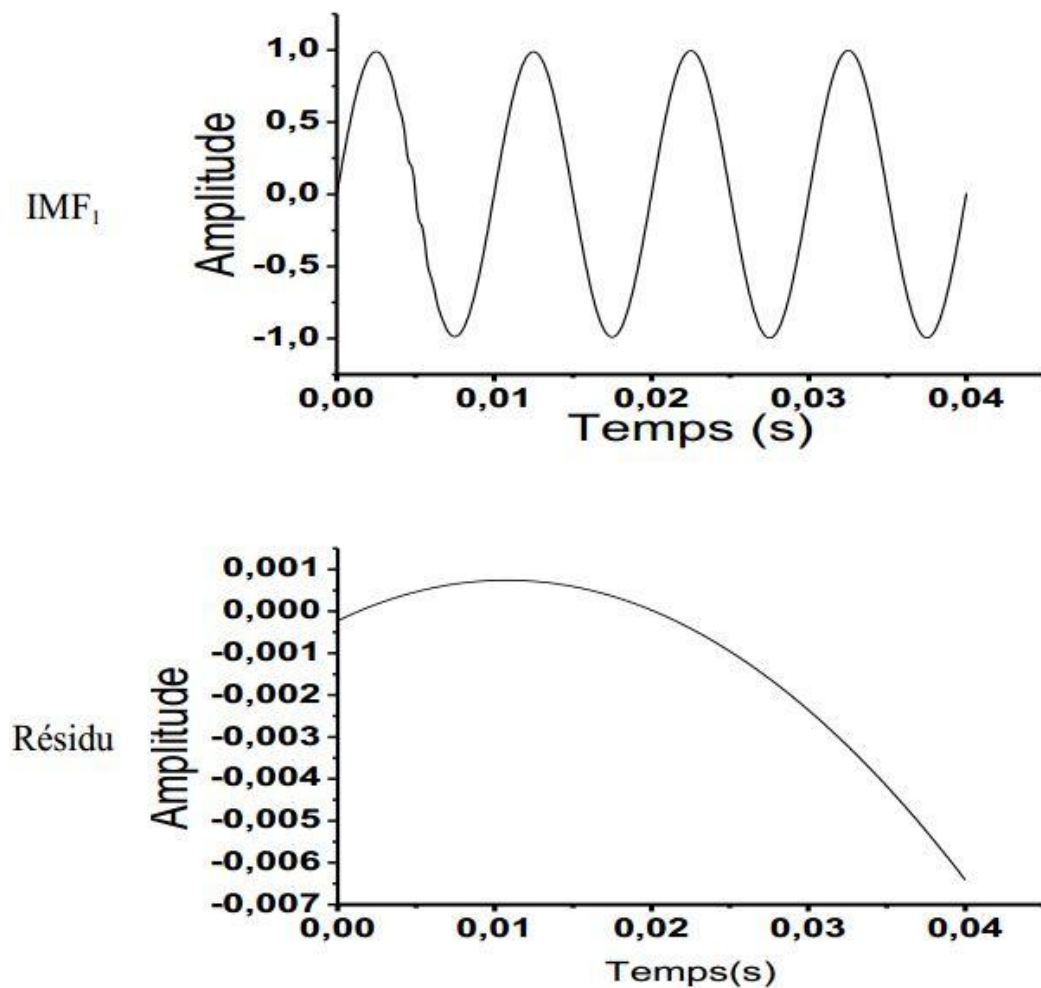


Figure II.6 Décomposition EEMD du signal $x_1(t)$ en utilisant un $SNR = 57 \text{ dB}$ et un $N_e = 1000$.

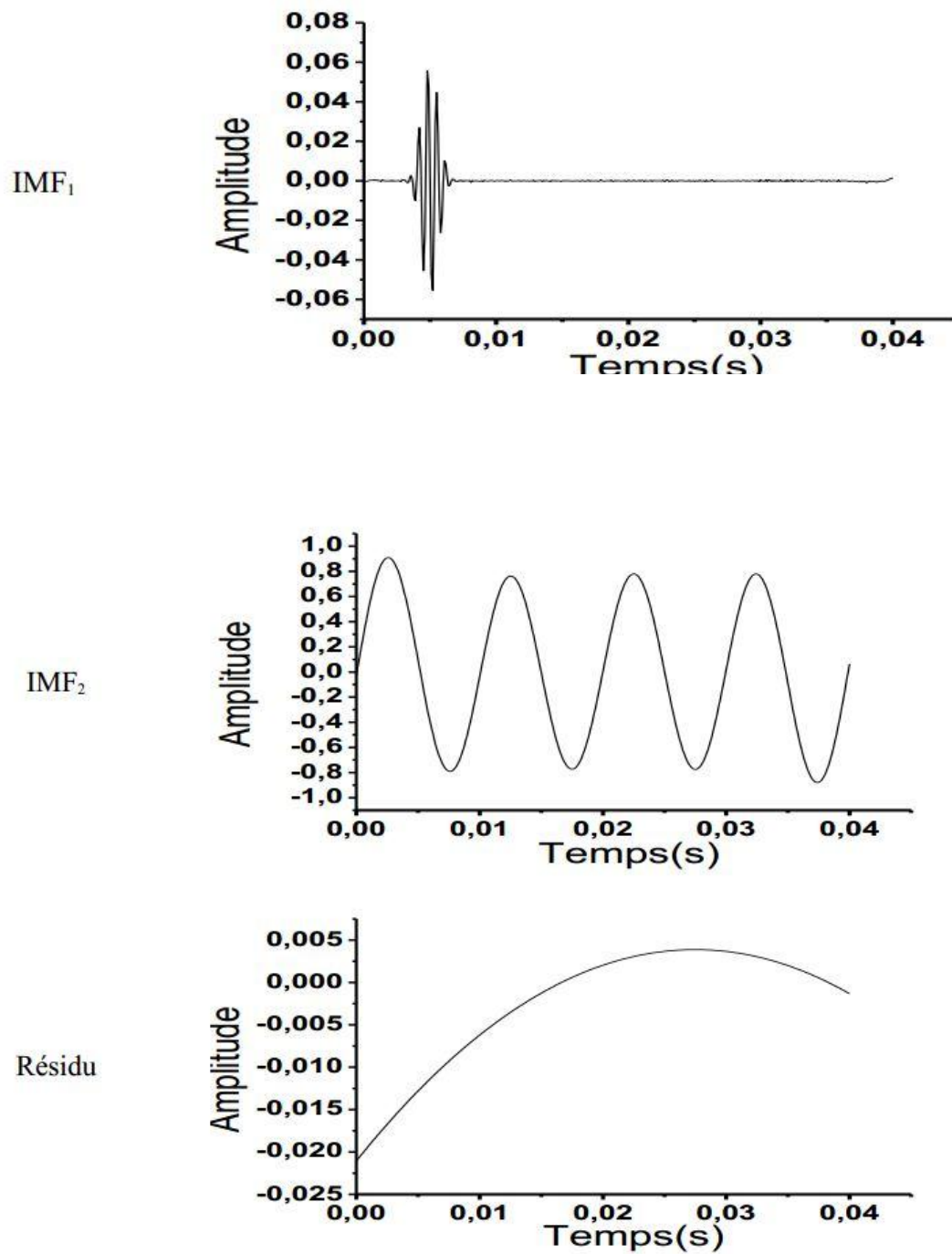
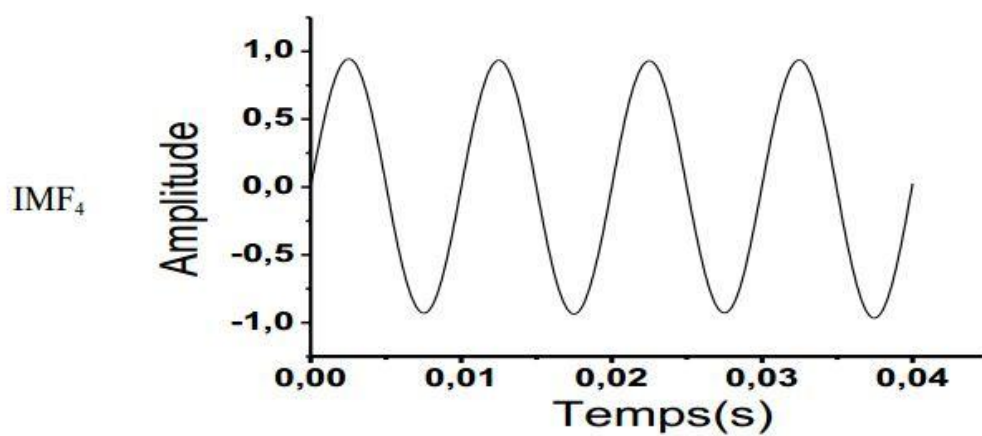
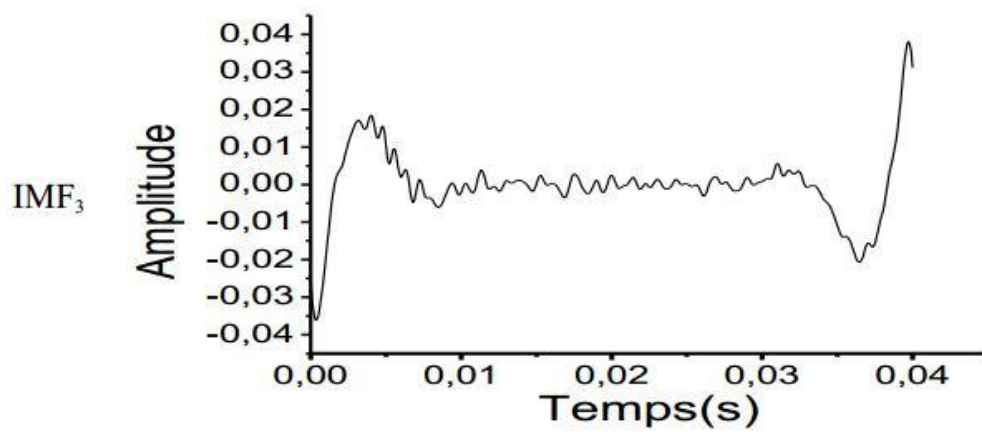
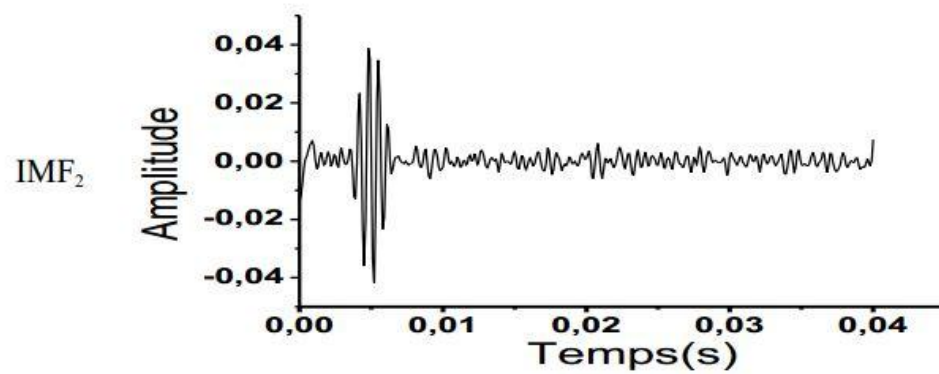
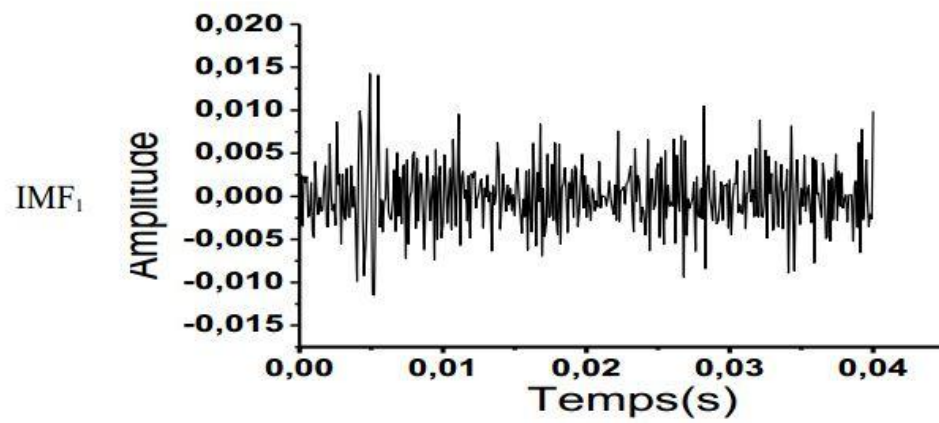


Figure II.7 Décomposition EEMD du signal $x_1(t)$ en utilisant un $SNR = 37 \text{ dB}$ et un $N_e = 1000$.



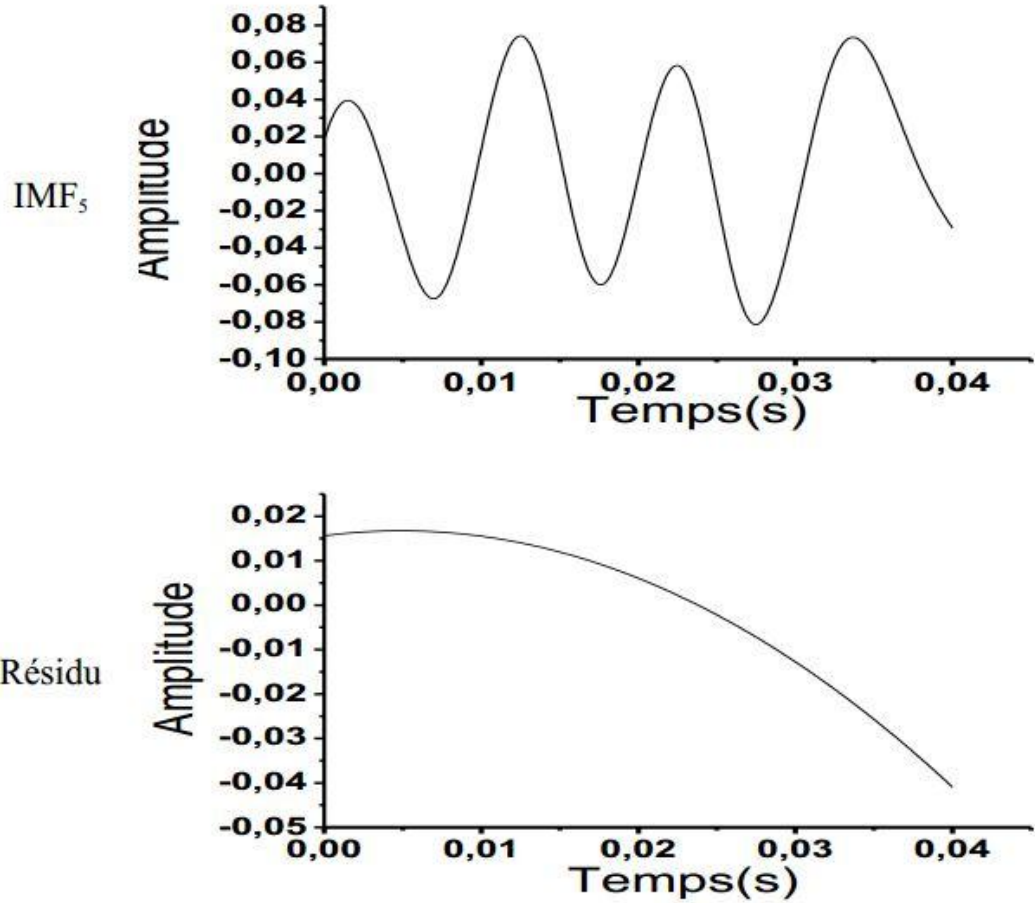


Figure II.8 Décomposition EEMD du signal $x_1(t)$ en utilisant un $SNR = 11 \text{ dB}$ et un $N_e = 1000$.

Ces figures montrent que la décomposition EEMD dépend de l'amplitude du bruit exprimée par la variance. Par exemple, pour un SNR important de l'ordre de 57 dB , les deux composantes du signal test n'ont pas été séparées et le problème de mélange de modes existe encore (figure II.6). Par contre pour un très faible SNR, la décomposition EEMD génère des *IMFs* redondantes (figure II.8). Cependant, pour un SNR approprié, dans cet exemple de 37 dB , la figure II.7 montre que le signal $x_1(t)$ a été efficacement décomposé en deux *IMFs*. Ces résultats démontrent que la variance du bruit blanc Gaussien doit être choisie convenablement pour assurer la performance de la méthode EEMD.

II.4.2 Effet du nombre d'essai N_e

Théoriquement, un nombre infini d'essai serait nécessaire pour complètement neutraliser l'effet du bruit blanc gaussien ajouté au signal original. Etant donné qu'une telle exigence ne peut pas être pratiquement respectée dans des applications

réalistes, une mesure de corrélation entre les $IMFs$ et les composantes du signal original est examinée pour déterminer un nombre d'essai faisable en pratique.

Quand le nombre d'essai est 1000, le coefficient de corrélation entre IMF_1 et la composante haute fréquence dans le signal $x_1(t)$ est élevé de 0.95 Cette valeur est prise comme une valeur seuil acceptable pour déterminé le nombre d'essai pratiquement significatif.

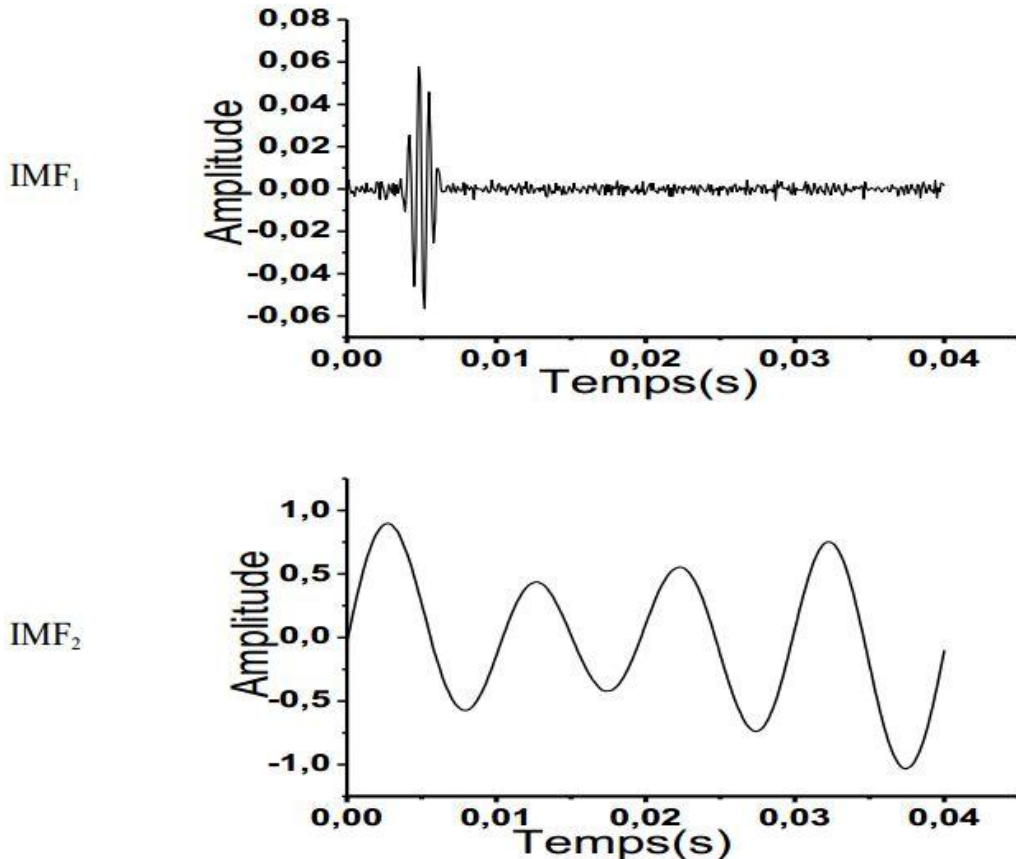
Les figure II.9- II.12 illustrent comment les résultats de la décomposition EEMD varient avec le nombre d'essai, où l'amplitude du bruit est fixée à 0.01 ($SNR = 37dB$).

Cette valeur a été choisie pour éviter les $IMFs$ redondantes selon le critère du rapport signal sur bruit.

L' IMF_2 qui représente la composante basse fréquence dans le signal $x_1(t)$ est stable pour tous les nombres d'essai N_e .

L' IMF_1 , correspondant à la composante haute fréquence de faible amplitude, est progressivement dé-bruité avec l'augmentation du nombre d'essai. L' IMF_1 est complètement dépourvue de bruit pour un nombre d'essai $N_e=1000$.

Ceci démontre que le coefficient de corrélation fournit une mesure efficace pour déterminer le nombre d'essai N_e dans la méthode EEMD.



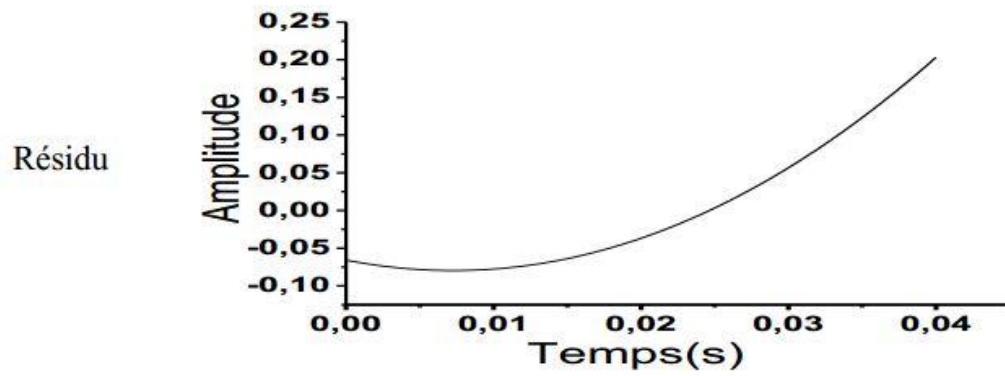
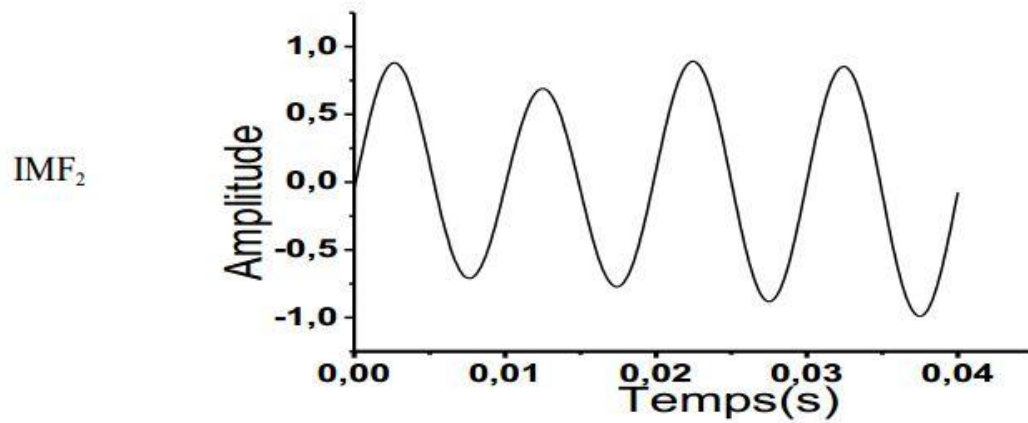
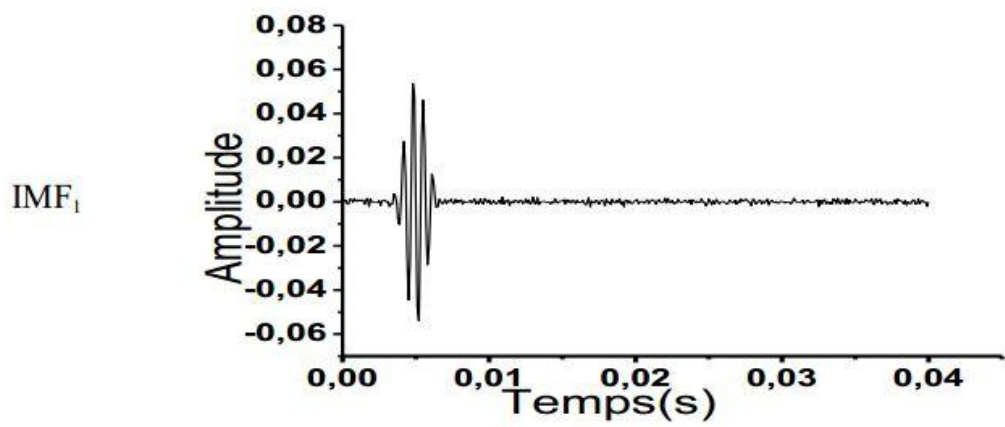


Figure II.9 Décomposition EEMD du signal $x_1(t)$ en utilisant un $SNR = 37 \text{ dB}$ et un $N_e = 10$.



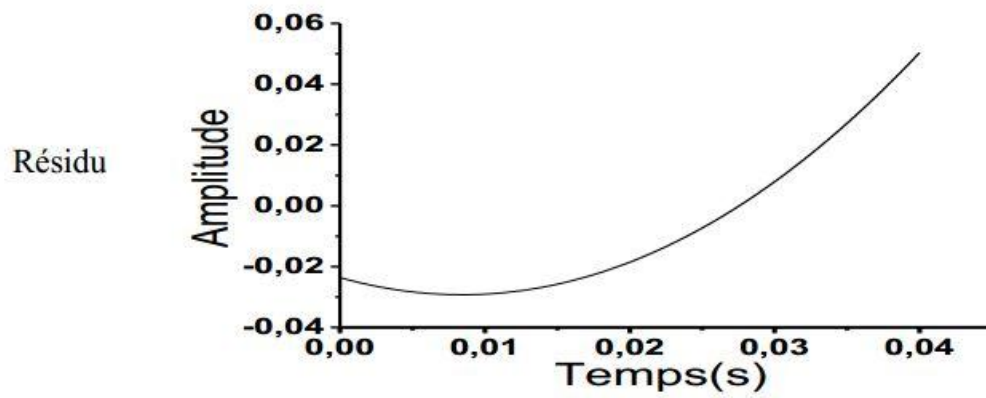
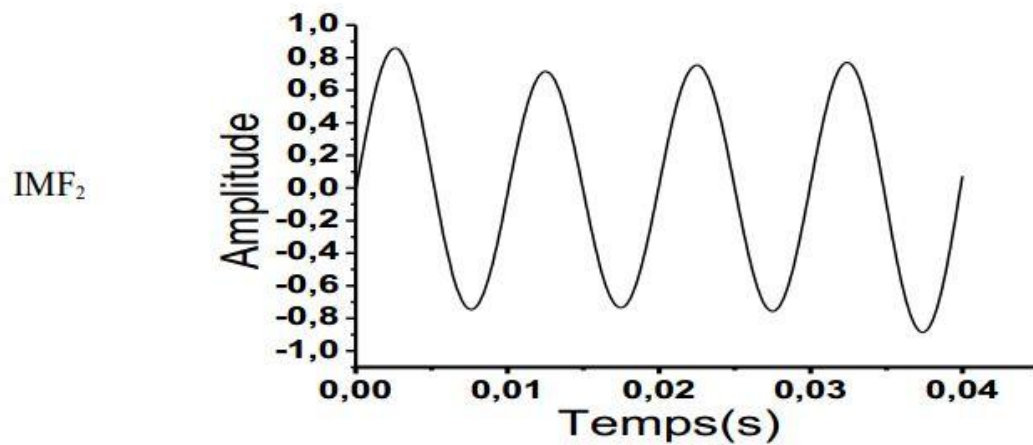
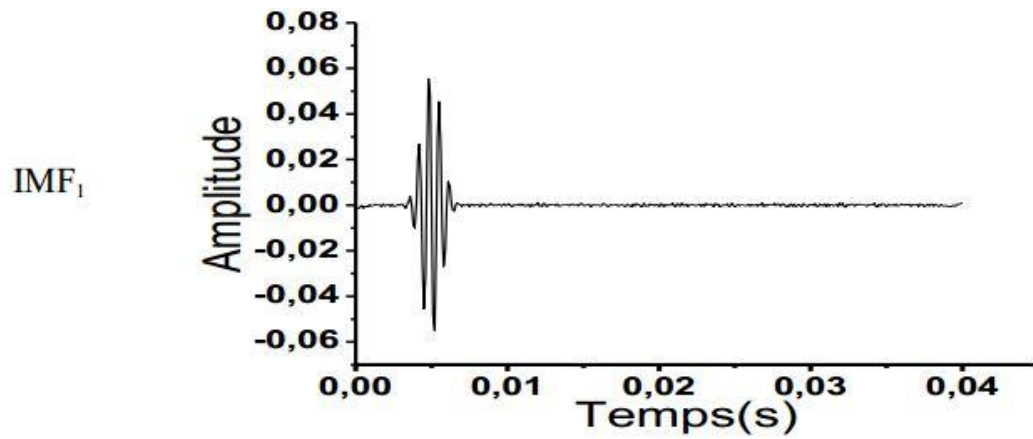


Figure II.10 Décomposition EEMD du signal $x_1(t)$ en utilisant un $SNR = 37 \text{ dB}$ et un $N_e = 50$ [20].



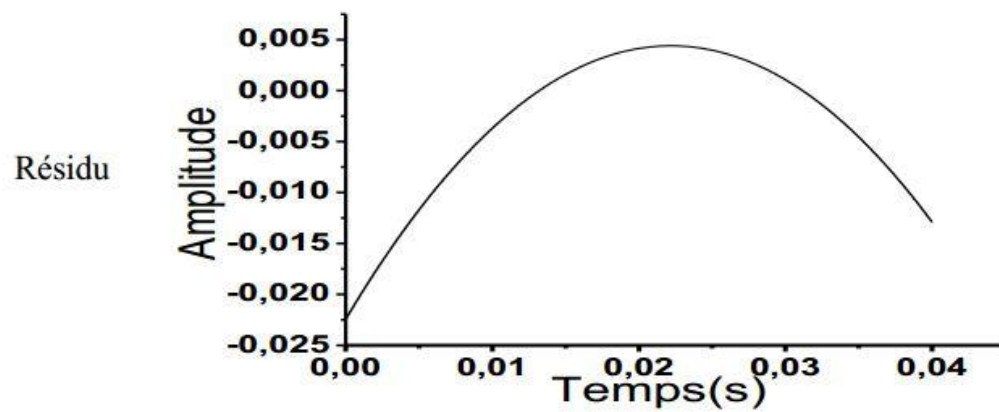
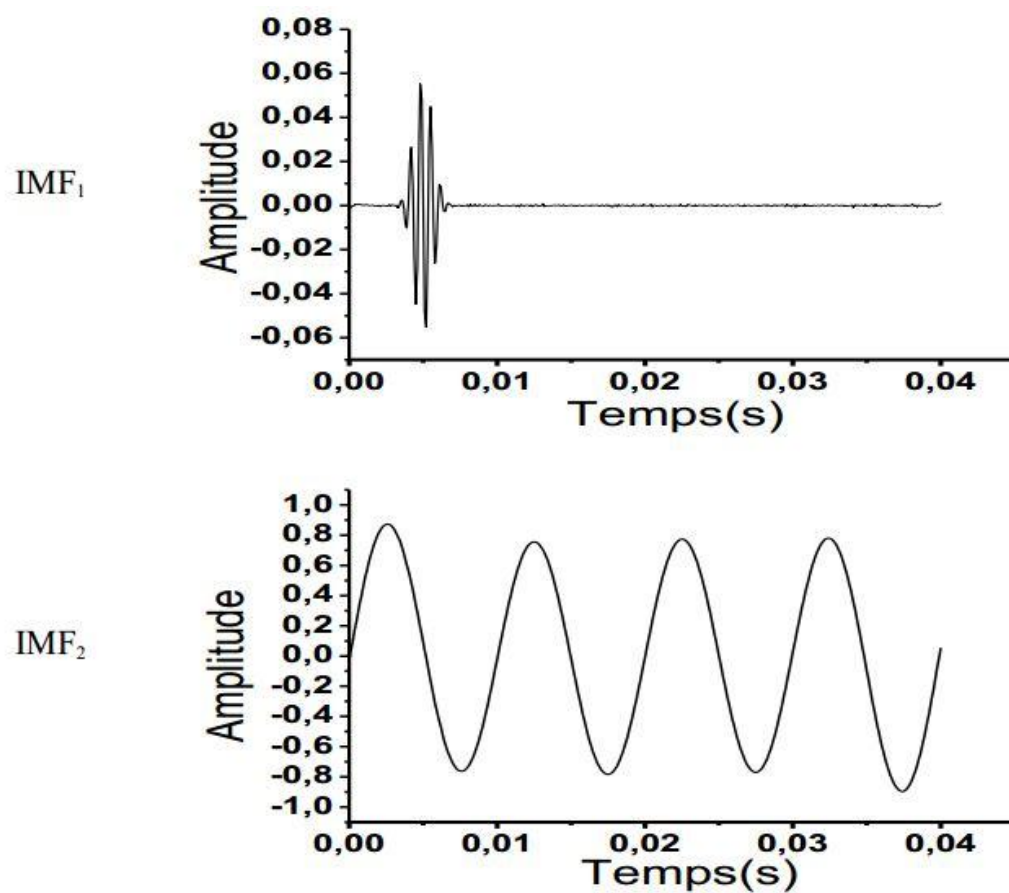


Figure II.11 Décomposition EEMD du signal $x_1(t)$ en utilisant un $SNR = 37 \text{ dB}$ et un $N_e = 200$.



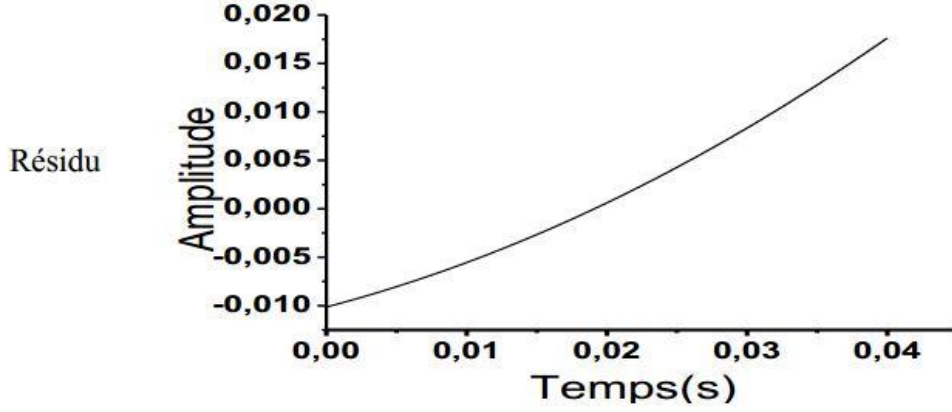


Figure II.12 Décomposition EEMD du signal $x_1(t)$ en utilisant un $SNR = 37 \text{ dB}$ et un $N_e = 500$ [18].

II.5 Evaluation de l'erreur par la méthode EEMD

Dans le but d'examiner la manière d'évaluation de l'erreur par la méthode EEMD en fonction du nombre d'essai N_e , nous avons décomposé le signal $x_1(t)$ avec un $SNR = 37 \text{ dB}$ et en faisant varier le nombre d'essai de 10 à 1000 et nous avons évalué l'erreur commise par la méthode EEMD. L'erreur est calculée en faisant la différence entre le signal original et le signal décomposé par la méthode EEMD [20].

$$l'erreur = signal \text{ original} - \sum_{K=1}^N IMF_{EEMD_K} + \text{résidu} \quad (\text{II.8})$$

Nous pouvons remarquer que lorsque le nombre d'essai augmente, la méthode EEMD diminue l'erreur sur toutes les $IMFs$ [18]. Mais l'augmentation du nombre d'essai N_e augmente le temps de calcul. Les résultats sont illustrés sur les figures II.13-II.16.

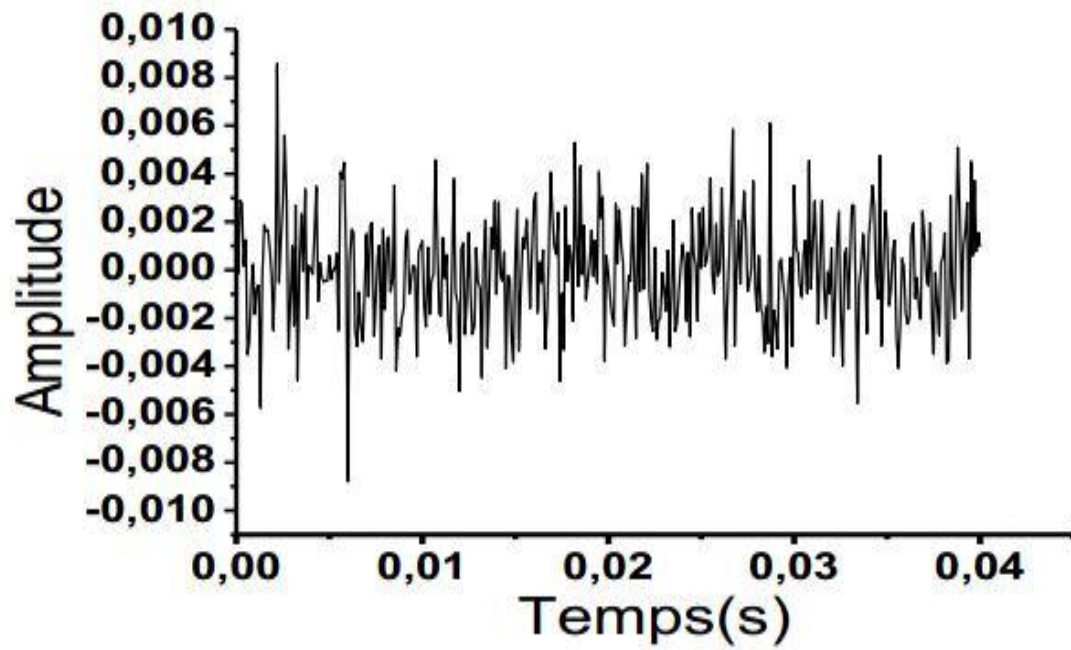


Figure II.13 L'évaluation de l'erreur par la méthode EEMD pour $N_e = 10$.

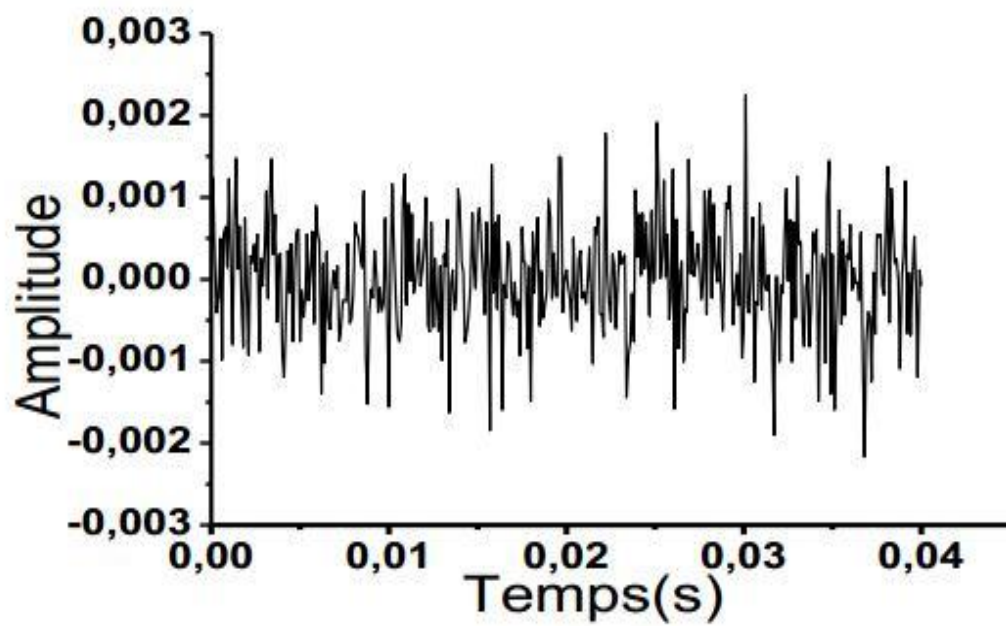


Figure II.14 L'évaluation de l'erreur par la méthode EEMD pour $N_e = 100$.

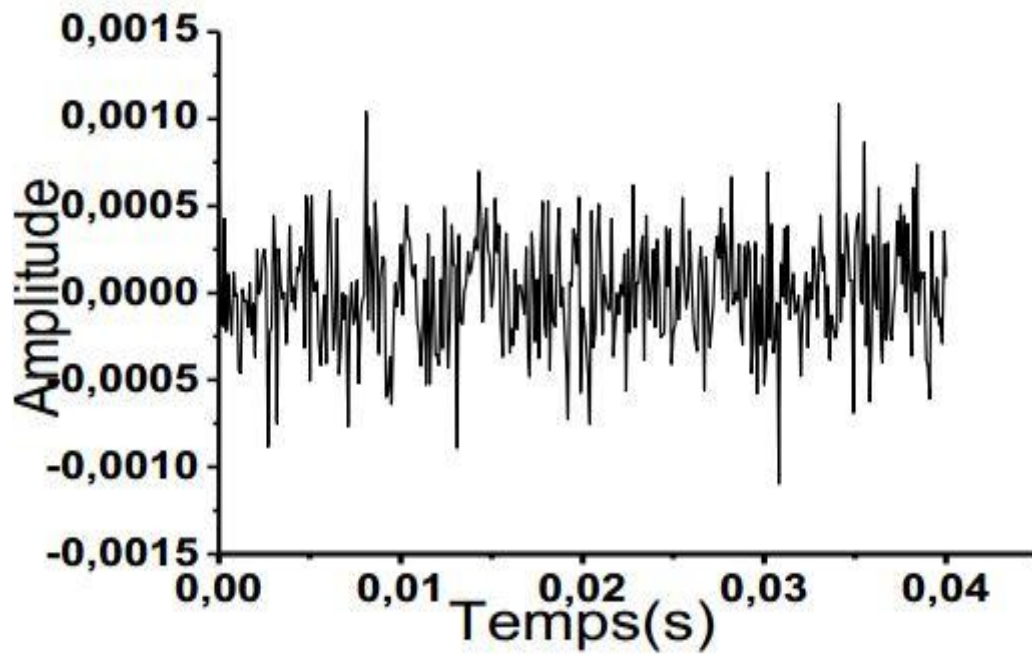


Figure II.15 L'évaluation de l'erreur par la méthode EEMD pour $N_e = 500$.

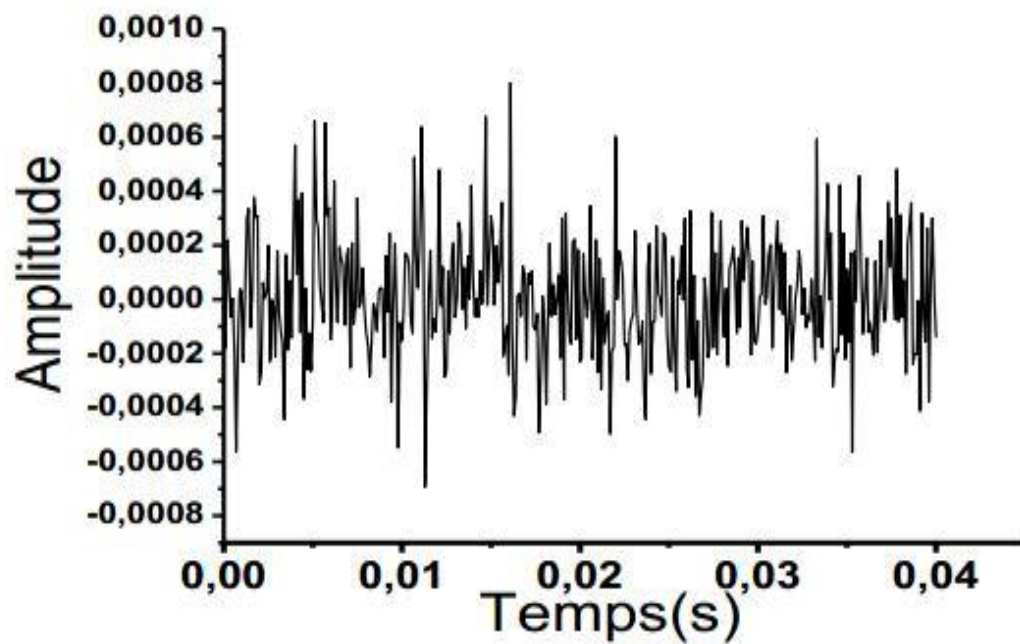


Figure II.16 L'évaluation de l'erreur par la méthode EEMD pour $N_e = 1000$.

II.6 Inconvénient de la méthode EEMD

Théoriquement, le bruit blanc ajouté est complétement éliminé par une décomposition utilisant un nombre d'essais très élevé (infini) qui pose un problème dans la méthode EEMD. Ainsi, cette méthode possède deux inconvénients:

- Difficulté d'éliminer complétement le bruit .
- Grand temps de calcul (figure II.17).

Par exemple, si le nombre d'essais Ne est fixé à 100, l'EEMD devrait exécuter 100 fois le processus de l'EMD et par conséquent EEMD est plus lente que l'EMD. Nous pouvons noter que le temps de calcul dans la méthode EEMD est fixé par le nombre d'essais Ne dont la figure II.17 donne quelques exemples. Nous avons utilisé un PC avec les caractéristiques suivantes processeurs : Intel pentium CPU B950, 2.1 GHz (dual core), avec une mémoire de 4GB.

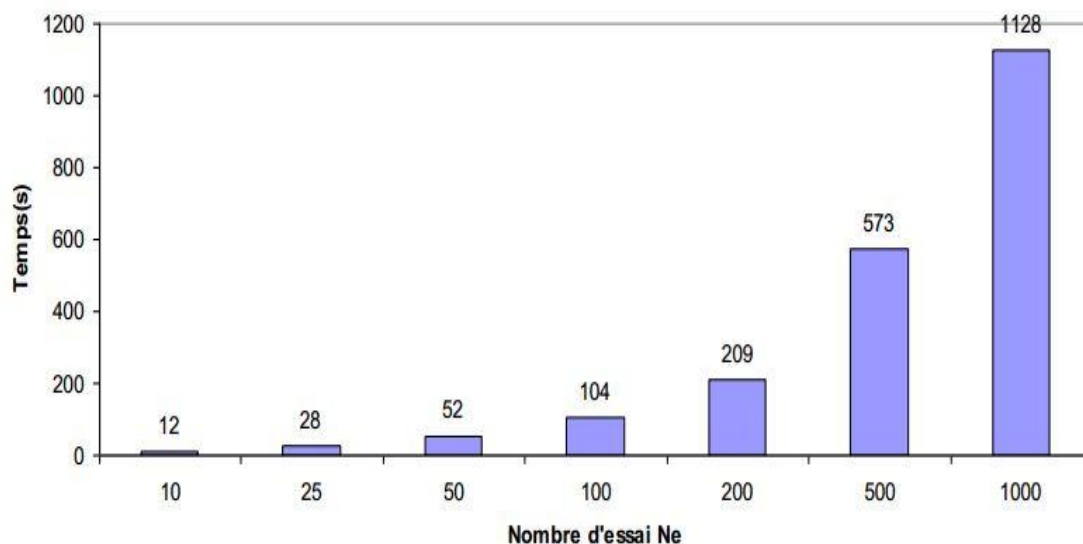


Figure II.17 Le temps de calcul en fonction de Ne pour la décomposition EEMD.

Nous avons confirmé que la méthode EEMD a résolu le problème de mixage de modes et par conséquent elle est plus performante que la méthode EMD. En plus, nous avons confirmé que l'erreur commise dans la reconstitution du signal diminue avec l'augmentation du nombre d'essai Ne .

Cependant, nous avons montré que l'augmentation du nombre d'essai accroît le temps de calcul.

II.7 La méthode MEEMD

Dans cette partie une amélioration de la méthode EEMD est présentée, c'est la méthode MEEMD (Modified EEMD). Cette amélioration est basée sur le remplacement du bruit blanc gaussien utilisé dans la méthode EEMD par un bruit à bande limité. Ce remplacement réduira efficacement le nombre d'essais N_e pour obtenir des *IMFs* significatives. Techniquement, le bruit blanc à bande limité est obtenu par un filtrage passe-bas du bruit blanc gaussien [18].

II.7.1 Algorithme de la méthode MEEMD

La méthode EEMD annule le problème de mélange de mode, mais avec un nombre de réalisations très élevé. Par conséquent une amélioration de l'EEMD est nécessaire. La partie basse fréquence du bruit ajouté affecte les résultats de la décomposition EEMD, alors que les hautes fréquences du bruit blanc ajouté n'ont aucun effet.

Les étapes de la méthode MEEMD sont décrites comme suit [22] :

- (1) Ajouter un bruit blanc filtré au signal original.
- (2) Décomposer l'ensemble (signal plus bruit blanc filtré) par la méthode EMD.
- (3) Répéter l'étape 1 et 2 en utilisant à chaque essai une autre réalisation de bruit blanc.
- (4) Calculer les différentes moyennes des différents ensembles des *IMFs* (une moyenne d'ensemble des *IMFs* de même ordre) de chaque décomposition EMD.

II.7.2 Critères de choix des paramètres de la méthode MEEMD

II.7.2.1 le choix du type de filtre

Nous avons considéré trois types de filtres passe bas (butterworth, Tchebychev 2 et elliptique) pour obtenir un bruit blanc à bande limité. Pour chaque type de filtre nous avons déterminé le SNR nécessaires pour la décomposition MEEMD du signal $x_1(t)$ (401 points et un pas d'échantillonnage de $10^{-4}s$) Le tableau II. 1 montre le nombre d'essais obtenu pour chaque type de filtre. Le meilleur résultat est obtenu avec le filtre elliptique d'ordre 3.

D'après le tableau II. 1, on peut constater que le filtre elliptique fournit un ordre minimal égal à 3 avec un nombre d'essais $N_e = 27$ et un faible RMSE : L'erreur quadratique moyenne (en anglais Root Mean Squared Error) (équation II. 9) par rapport aux deux autres filtres. Ensuite, nous avons décomposé le signal $x_1(t)$ par la méthode MEEMD en filtrant le bruit blanc en utilisant le filtre elliptique d'ordre 3.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - x'_i)^2} \quad (II. 9)$$

N est le nombre d'échantillons du signal test, x'_i est le signal estimé, x_i est le signal original [22].

Tableau II. 1 . Choix du filtre passe bas.

	Butterworth	Tchebychev2	Elliptique
SNR (dB)	15 dB	15 dB	15 dB
Ordre du filtre	7	3	3
N_e	40	27	27
RMSE	0.000671996	0.00061882	0.00060327

Tab II.1 Choix du filtre passe bas [18].

II.7.2.2 Le rapport signal sur bruit

La figure II.18 donne la relation entre le SNR et les résultats de la décomposition MEEMD. Le nombre d'essais a été fixé à $N_e = 27$ réalisations. Nous pouvons constater que la gamme du SNR [10dB – 15dB] permet d'avoir une bonne décomposition MEEMD du signal $x_1(t)$ (sans mélange de mode et sans *IMFs* redondantes).

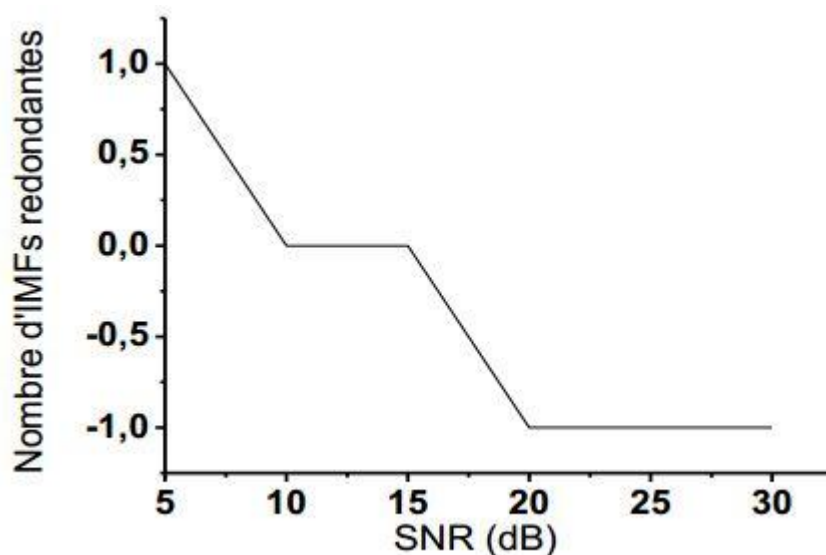


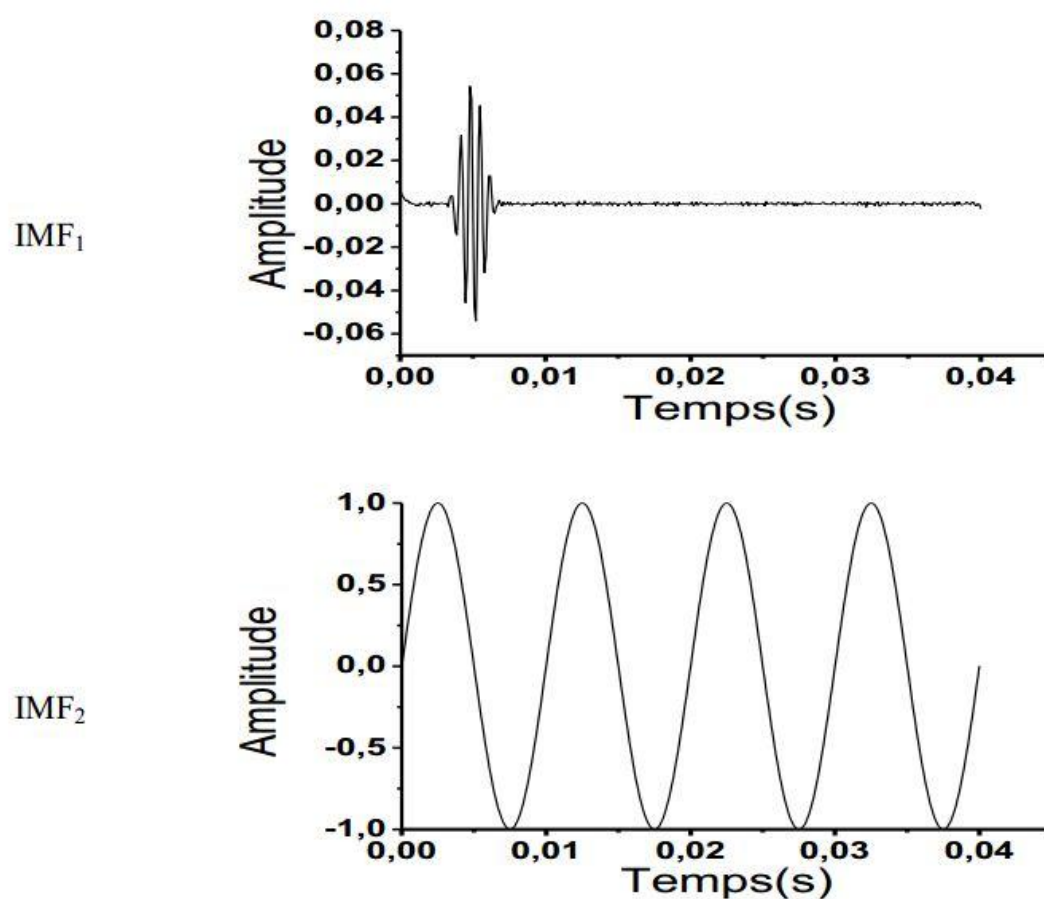
Figure II.18 SNR et nombre d'*IMFs* redondantes pour la décomposition MEEMD du signal $x_1(t)$.

II.7.2.3 Le nombre d'essais N_e

Pour obtenir un coefficient de corrélation, entre l'IMF₁ et la composante haute fréquence $x_3(t)$ du signal $x_1(t)$ élevé de l'ordre de 95%, Il faut que le nombre d'essais soit égal à 27.

II.7.3 La décomposition par la méthode MEEMD avec un filtre elliptique d'ordre 3

Après avoir choisi le meilleur filtre passe bas qui nécessite un ordre minimal pour décomposer le signal $x_1(t)$ avec un faible nombre d'essai et une faible erreur, et les bons paramètres nécessaires à la décomposition par le processus MEEMD. La figure II.19 montre le résultat de la décomposition MEEMD du signal $x_1(t)$ ($N_e = 27$, $SNR = 15dB$) avec l'utilisation d'un filtre elliptique d'ordre 3.



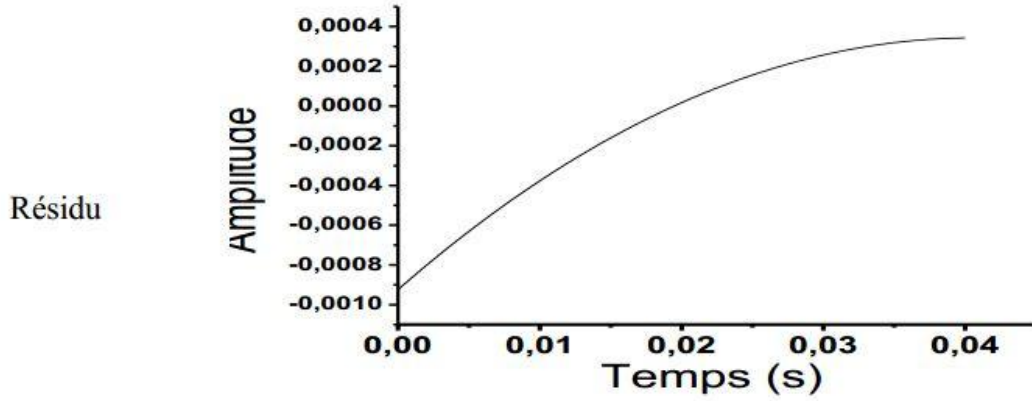


Figure II.19 Décomposition MEEMD du signal $x_1(t)$ avec $SNR = 15dB$ et $N_e = 27$.

La méthode EEMD, exige un nombre important de réalisations ($N_e > 1000$) pour réduire le bruit des *IMFs*. Par contre dans la méthode MEEMD, l'extraction de la composante haute fréquence (IMF_1) est réussie uniquement avec un très faible nombre d'essai $N_e = 27$. D'où l'avantage de la méthode MEEMD de réduire le temps de calcul.

II.8 La méthode CEEMD

Dans ce paragraphe, nous avons considéré la méthode CEEMD (Complementary Ensemble Empirical Mode Decomposition), pour améliorer l'efficacité de la méthode EEMD [18].

II.8.1 Algorithme de la méthode CEEMD

Dans la méthode CEEMD, du bruit blanc gaussien est ajouté et soustrait au signal original pour générer des ensembles d'*IMFs*. Par conséquent, deux mélanges composés des données originales sont obtenus par [23] :

$$\begin{bmatrix} M_1 \\ M_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} S \\ B \end{bmatrix} \quad (II. 10)$$

Où :

- *S est le signal original.
- *B est le bruit blanc gaussien ajouté.
- * M_1 est la somme du signal original avec le bruit blanc gaussien ajouté.
- * M_2 la différence entre le signal original et le bruit blanc gaussien.

Ainsi, les IMFs obtenues en additionnant du bruit blanc au signal contribuent à une série d'IMFs contaminées par des résidus dus au bruit blanc gaussien ajouté. De même, les IMFs obtenues en soustrayant du bruit blanc au signal contribuent à une autre série d'IMFs contaminées des résidus dus au bruit blanc soustrait. Les IMFs finales est la moyenne de l'ensemble des deux séries IMFs.

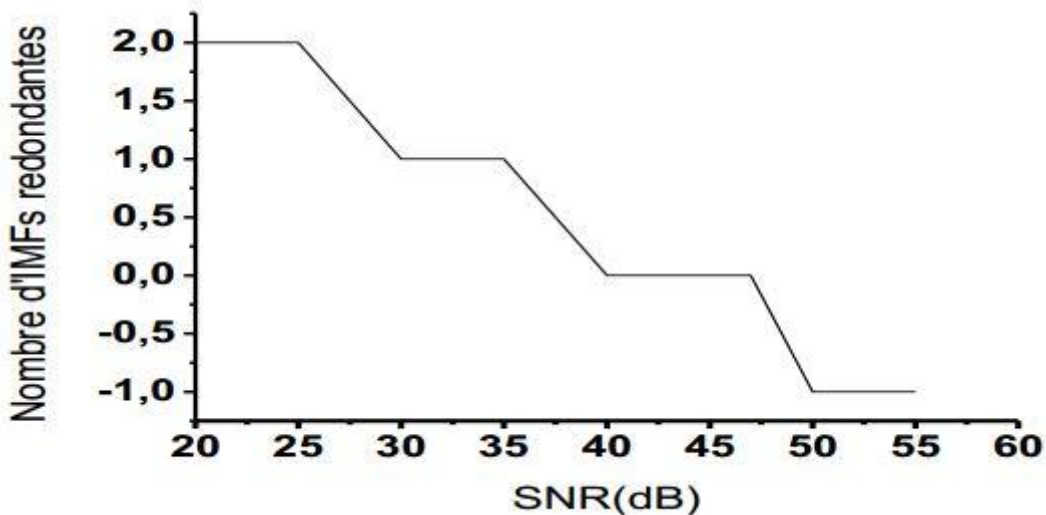
II.8.2 Critères de choix des paramètres

II.8.2.1 Le rapport signal sur bruit

Si le rapport signal sur bruit est très grand il y aurait un risque de mélange de modes. Par contre si le rapport signal sur bruit est faible, il y aurait un problème de redondance des IMFs. Il faut donc bien choisir la valeur du signal sur bruit qui évite ces deux phénomènes.

La figure II.20 représentant la relation entre le SNR et les résultats de la décomposition CEEMD du signal $x_1(t)$ montre que le SNR doit être compris dans la plage $[40 - 47.5dB]$ pour une avoir une bonne décomposition sans IMFs redondantes et sans mélange de mode.

Par conséquent, ce rapport signal sur bruit a été fixé à $47dB$ (résultat obtenu avec $N_e = 40$).



La figure II.20 SNR et nombre d'IMFs redondantes pour la décomposition CEEMD du signal $x_1(t)$ [18].

II.8.2.2 Le nombre d'essais N_e

Pour obtenir un coefficient de corrélation de 98% entre l'IMF1 et le signal $x_3(t)$ dans la décomposition CEEMD, le nombre d'essais N_e soit au moins

supérieur à 5 (figure II.21).

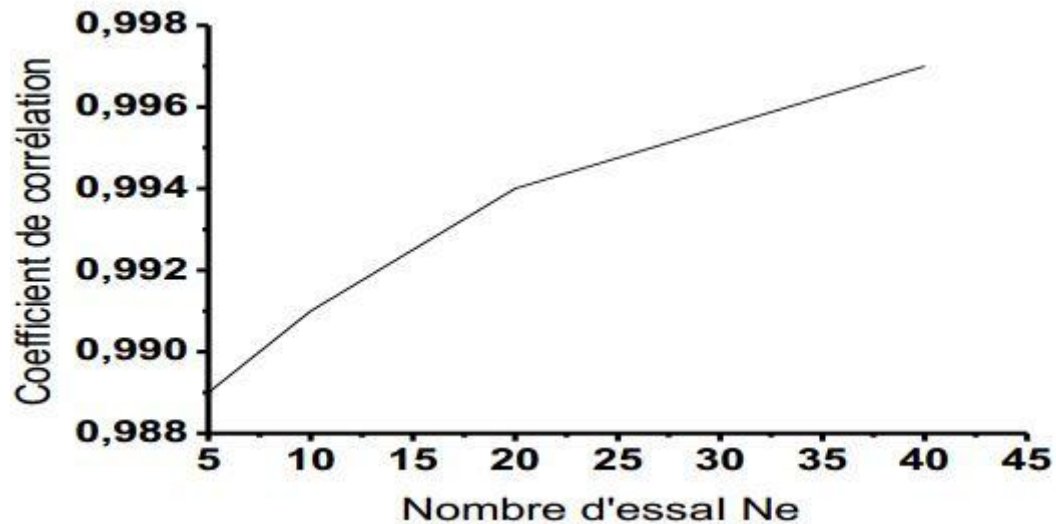
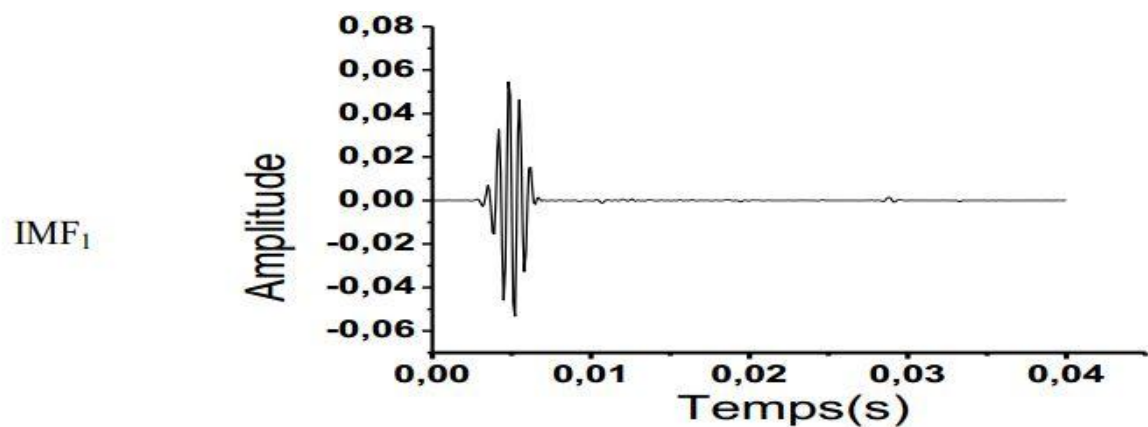


Figure II.21 Relation entre le coefficient de corrélation et le nombre d'essais N_e .

II.8.3 La décomposition par la méthode CEEMD

La figure II.22 représente la décomposition du signal $x_1(t)$ par la méthode CEEMD résultats de la décomposition réussie du signal $x_1(t)$ sont obtenus avec un $N_e = 40$ et un $SNR = 47dB$ a permis donc de réduire le nombre d'essais et par conséquent le temps de calcul. La méthode CEEMD constitue une amélioration de la méthode EEMD [23].



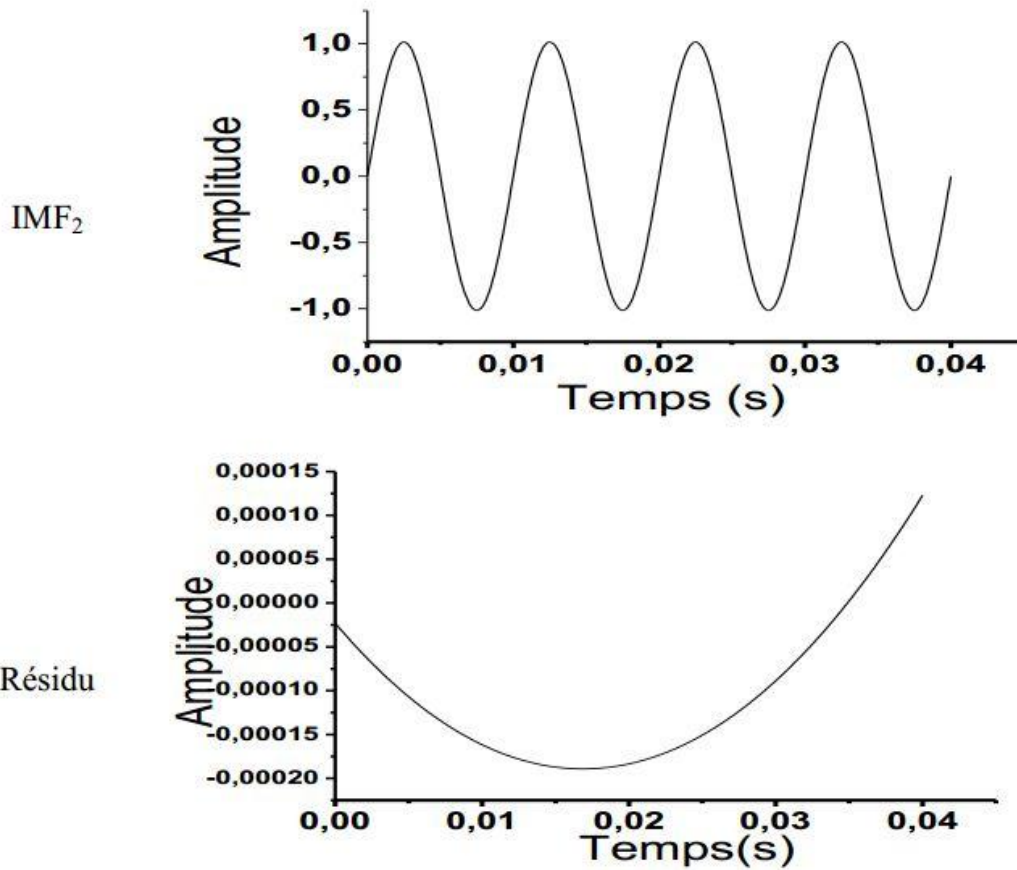


Figure II.22 La décomposition CEEMD du signal $x_1(t)$ avec $SNR = 47dB$ et $N_e = 40$.

II.9 Les différentes techniques d'évaluation de la qualité de la parole

L'évaluation subjective de la qualité de la parole est une étape indispensable dans tout processus de traitement, automatisé ou non. Elle permet de tenir compte du jugement humain à travers des essais d'écoute de laboratoire par plusieurs auditeurs.

✓ Seg SNR

Les mesures dans le domaine temporel sont surtout destinées à des systèmes de codage/transmission qui tentent de reproduire la forme d'onde d'origine (par exemple, dans le cas des techniques de codage de forme d'onde). Les mesures dans le domaine temporel, les plus simples et les plus communes, pour l'évaluation de la qualité du signal sont Signal-to-Noise-Ratio (SNR) et Segmental Signal –to Noise-Ratio (SSNR). SSNR est une variation du SNR qui fonctionne sur des segments courts (15 à 20 ms) de flux. Cela permet de faciliter l'alignement temporel et il fournit des résultats qui sont légèrement mieux (en termes de corrélation avec l'évaluation subjective) que ceux du SNR.

✓ **Perceptual Evaluation of Speech Quality (PESQ)**

Le PESQ (Perceptual Evaluation of Speech Quality) est l'évaluation de la qualité vocale perçue désignée dans la norme P.862 [24] comme moyen adapté aux codecs vocaux et aux mesures de bout en bout. De ce fait, d'autres facteurs supplémentaires sont pris en considération pour mieux simuler les conditions réelles, à savoir le temps de propagation, les distorsions dues aux erreurs de transmission, les pertes de paquets. Néanmoins, il existe bel et bien d'autres facteurs techniques et applications [24] pour lesquels la méthode d'évaluation PESQ n'a pas été encore validée à ce jour, notamment les artefacts causés par les algorithmes de réduction de bruit ainsi que les dégradations liées à l'interaction bidirectionnelle lors de la transmission comme par exemple l'effet d'écho.

II.10 Conclusion

Dans ce chapitre, nous avons étudié les différentes possibilités d'améliorations de la méthode EMD en vue de réduire le nombre d'essais. Pour chaque méthode d'amélioration, les paramètres (SNR , N_e et le coefficient de corrélation) nécessaires, pour une décomposition réussie du signal test, ont été calculés. Les résultats de simulations ont montrés que toutes les méthodes d'améliorations publiées sont plus performantes que la méthode classique EEMD. La méthode EMD décompose d'une façon adaptative un signal en une somme de composantes oscillantes dont la première contient sa composante haute fréquence pour terminer vers les plus basses.

Application au débruitage

Sommaire :

III.1 Introduction

III.2 Simulation pour calculer les IMFs

III.3 La décompositions signal parole par EMD

III.4 Application d'EMD au débruitage

**III.5 Comparaison des deux approches de rehaussement de la parole
en termes de PESQ, segSNR, en présence de bruit blanc gaussien**

III.6 Conclusion

III.1 Introduction

Il existe deux aspects de la qualité de la parole ; la qualité de la parole globale perçue et l'intelligibilité de la parole. La qualité globale reçue est l'impression générale de l'auditeur de la qualité de la parole. La définition de «bon» est laissée à l'auditeur. Cependant, étant donné que nous entendons la parole transmise par l'air naturel émis par un humain réel tous les jours, ce signal parole fournit un «point de référence» sur l'échelle de qualité. Les auditeurs évaluent le signal sous test par rapport à cette référence. Le rehaussement de la parole améliore la qualité et l'intelligibilité de la communication vocale pour une gamme d'applications, et afin d'améliorer la qualité de signal parole en présence du bruit, une étude comparative des techniques de rehaussement (d'élimination de bruits) de signal parole est réalisée dans ce chapitre.

III.2 Simulation pour calculer les IMFs

Les étapes permettant d'extraire les IMF sont :

Etape 0 : On initialise le signal $h_0(t)$ par $x(t)$ et le signal résidu $r(t) = x(t)$

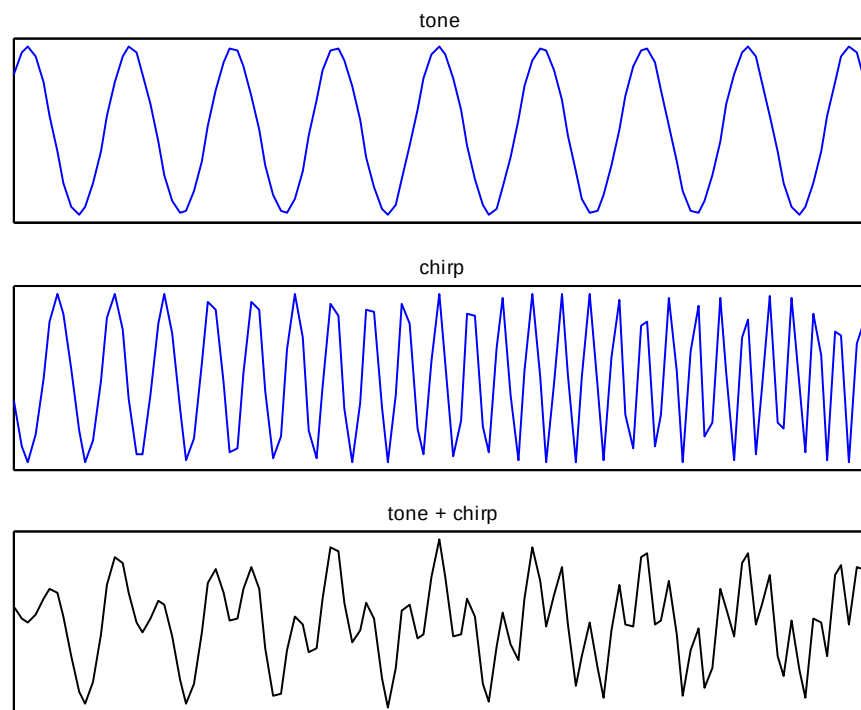


Figure III.1 Représentation du signal $x(t)$.

Etape 1 : Le premier pas de la décomposition consiste à identifier les extrema locaux du signal $h_0(t)$. C'est le début du processus de tamisage. On interpole ensuite tous les maxima avec la méthode **des splines cubiques** par exemple pour construire

l'enveloppe supérieure (env_{sup}) et on fait de même avec les minima pour construire l'enveloppe inférieure (env_{inf}). Notons que l'interpolation est une étape très importante dans le processus de décomposition car elle conditionne la forme des IMF. Plusieurs méthodes d'interpolation peuvent être utilisées. La méthode **des splines cubiques** reste cependant la plus utilisée [25].

Les figures III.2 et III.3 montrent les maxima de signal avec la méthode splines cubiques et l'enveloppe supérieur et inférieur par interpolation respectivement.

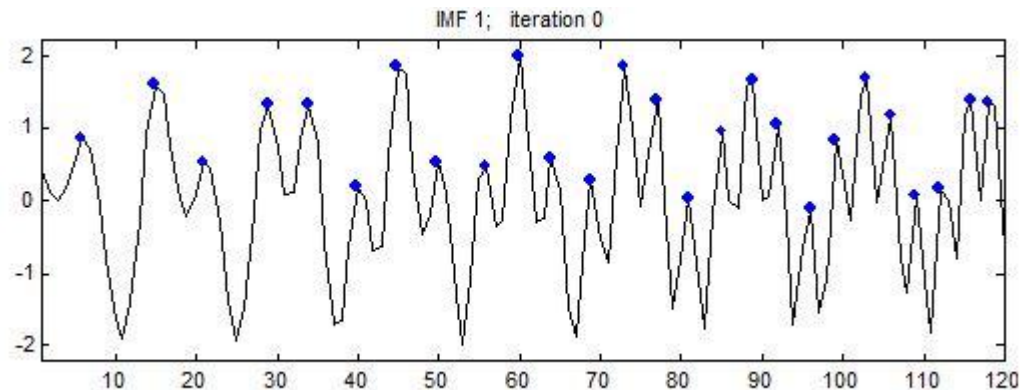


Figure III.2 Interpole les maxima avec la méthode des splines cubiques.

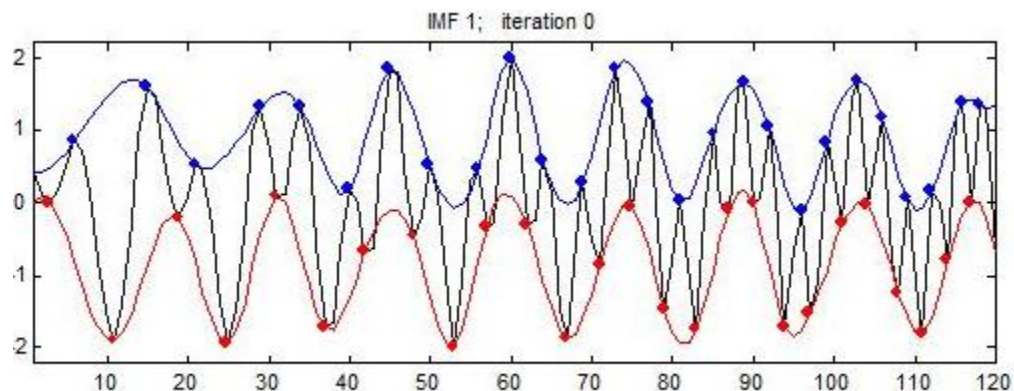


Figure III.3 Enveloppe supérieure et inférieure obtenues par interpolation des maxima et minima locaux.

Etape 2 : Dans cette étape on calcule la moyenne locale du signal suivant la formule suivante :

$$m_1(t) = (env_{sup} + env_{inf})/2 \quad (III.1)$$

La figure III.4 montre la moyenne locale de signal $x(t)$ des minimas et maxima locaux du signal présenté dans la figure III.1.

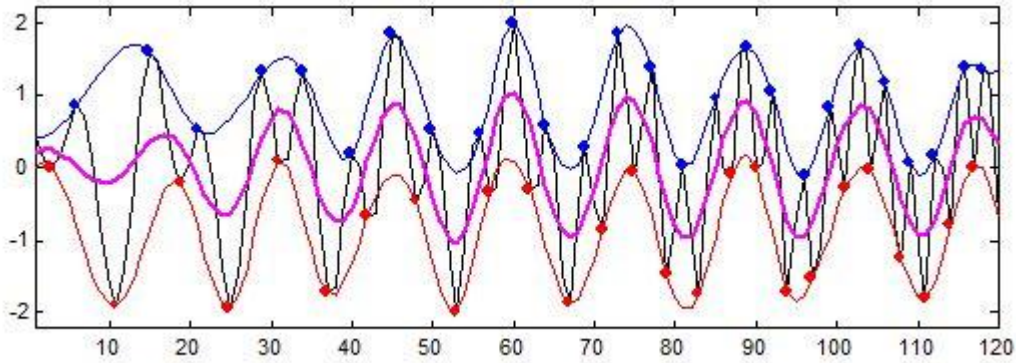


Figure III.4 Moyenne locale du signal $x(t)$.

Etape 3 : On calcule la différence entre le signal et le moyenne $m_1(t)$ afin de déterminer la première composante IMF notée $h_1(t)$

$$h_1(t) = h_0(t) - m_1(t) \quad (\text{III. 2})$$

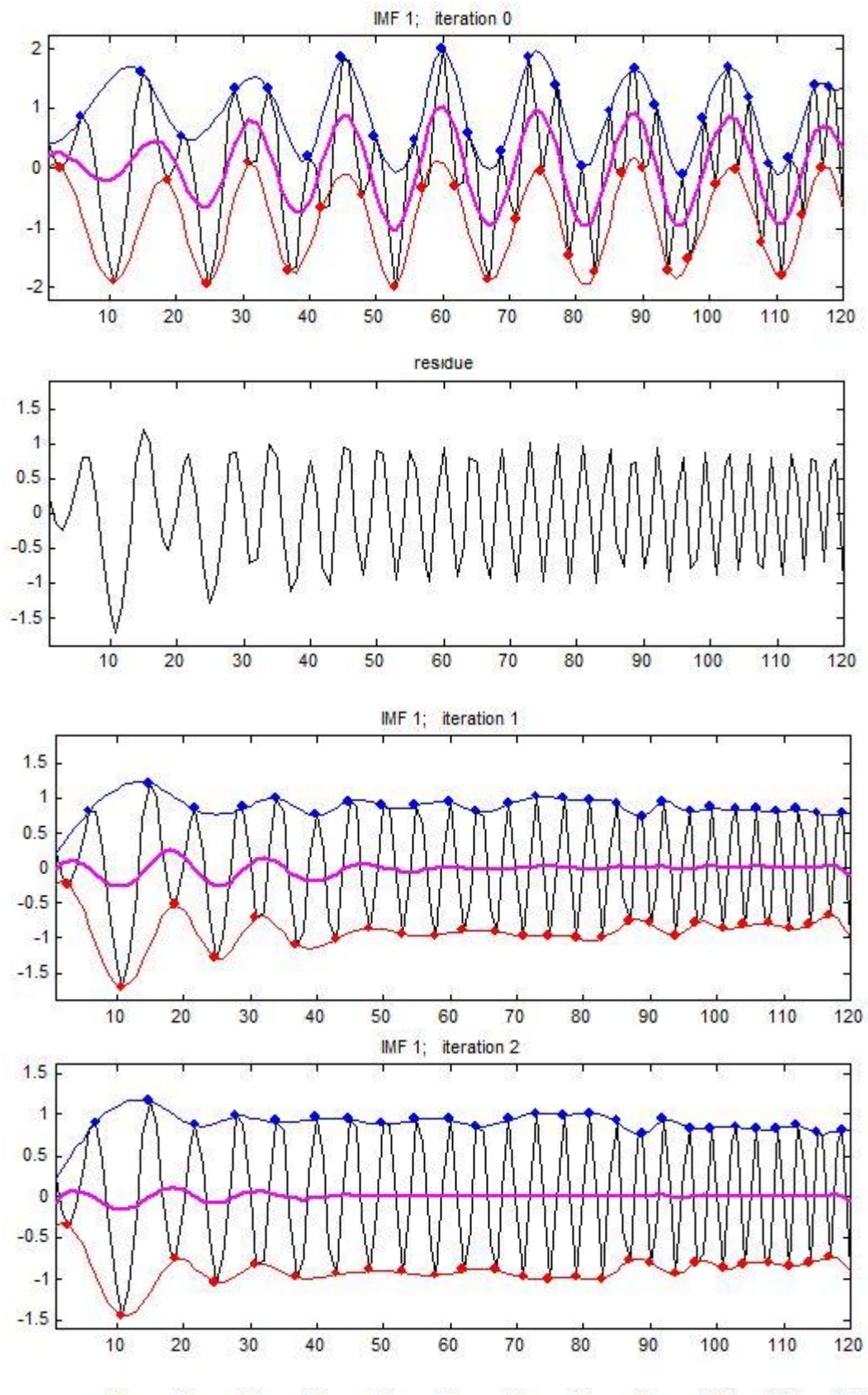
La figure III.5 montre le résultat de cette différence. On remplace ensuite le signal $h_0(t)$ par le signal $h_1(t)$ et on refait les étapes de 1 à 3 jusqu'à satisfaire un critère donné.

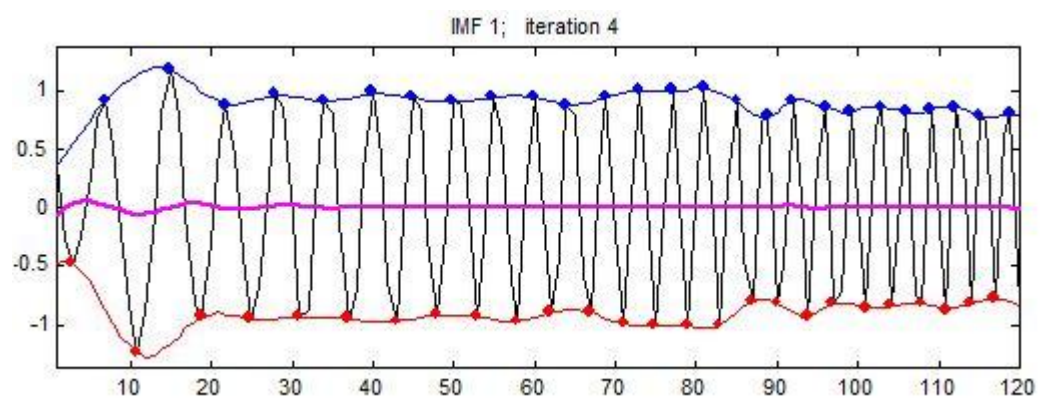
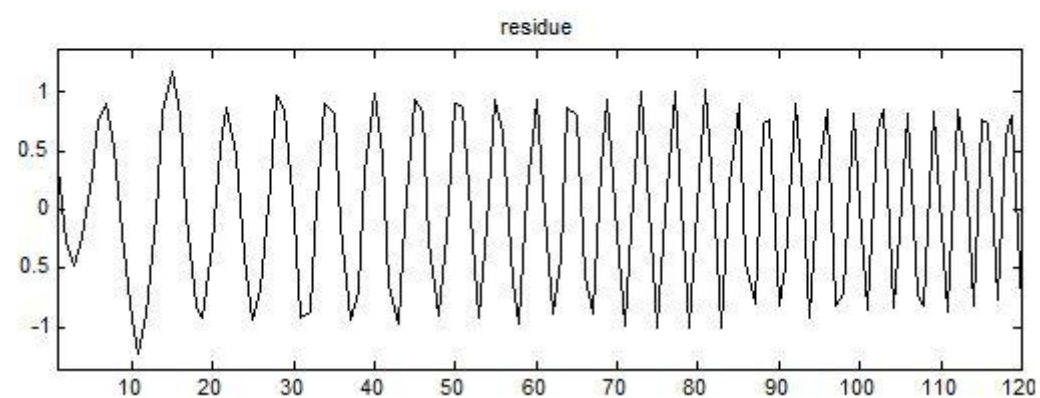
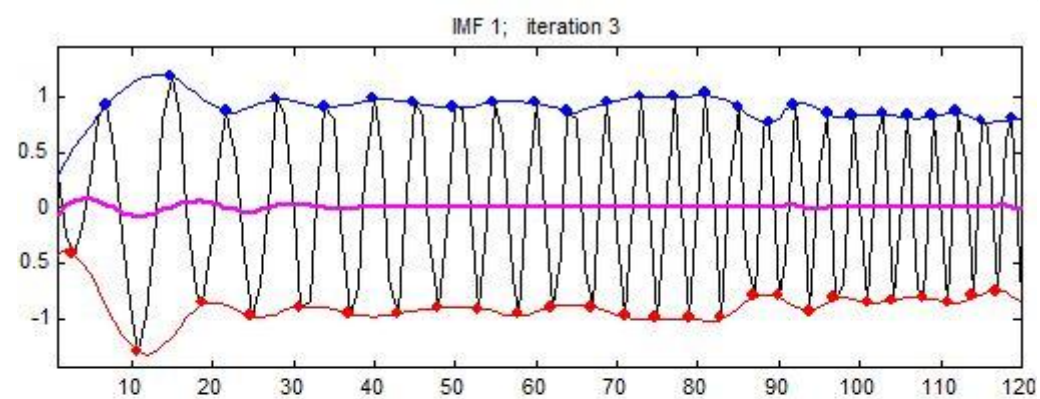
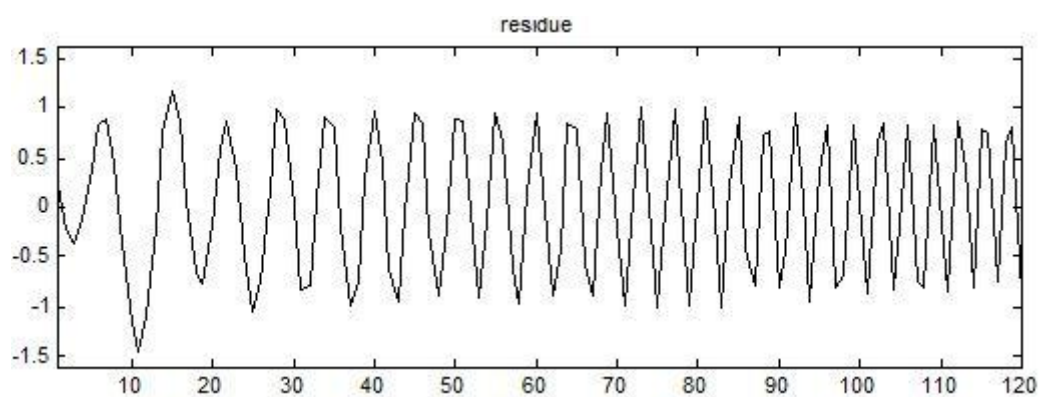
Etape 4 : A l'issue de l'étape 3 qui indique la fin du processus de tamisage, on recueille la première composante IMF $d_1(t) = h_1(t)$ et on calcule le signal résidu.

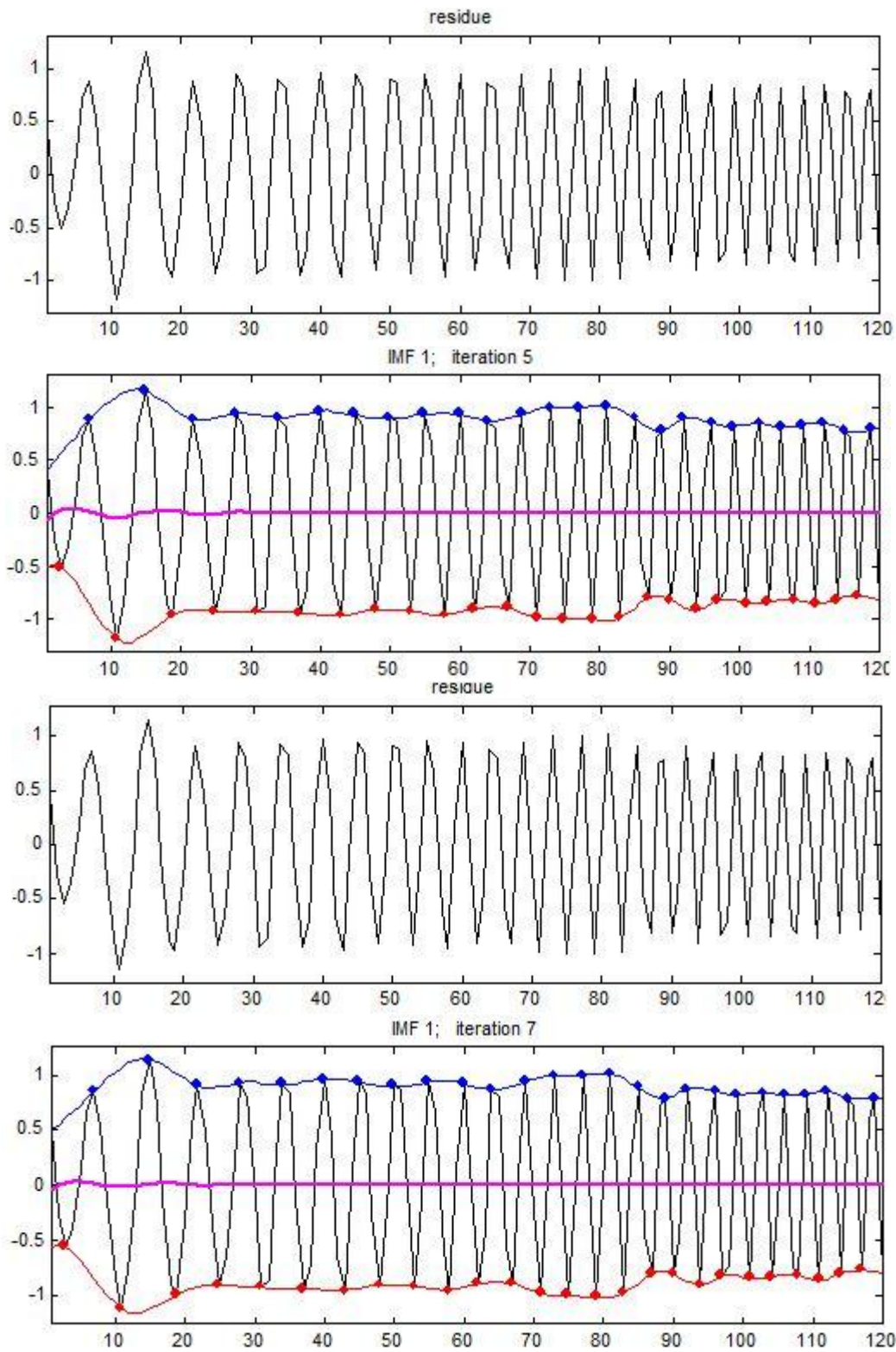
$$r_1(t) = r_0(t) - d_1(t) \quad (\text{III. 3})$$

Cette première IMF forme la composante haute fréquence du signal et le résidu $r_1(t)$ contient les composantes des plus basses fréquences. La figure (III.5) montre la première composante $h_1(t)$ et le premier résidu $r_1(t)$.

Etape 5 : On remplace $r_0(t)$ par $r_1(t)$ et $h_0(t)$ par $r_1(t)$ et on refait les étapes 1 à 5 afin d'extraire les autres composantes IMF. Ces étapes sont répétées jusqu'à atteindre le nombre d'IMF désiré ou le nombre d'extrema de $r_0(t)$ est inférieur à 2. La figure (III.6) montre les autres composantes IMF et leurs résidus.







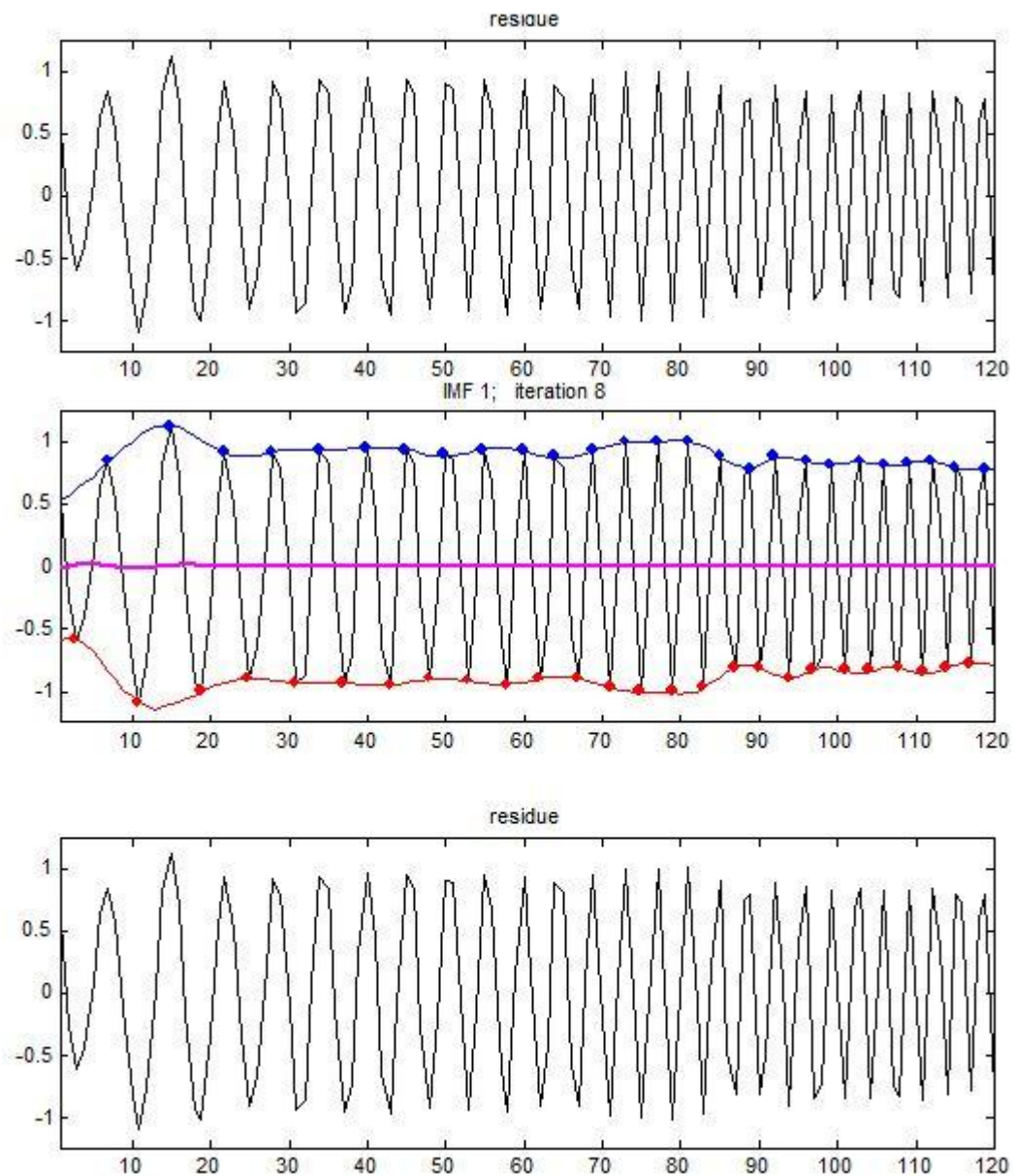
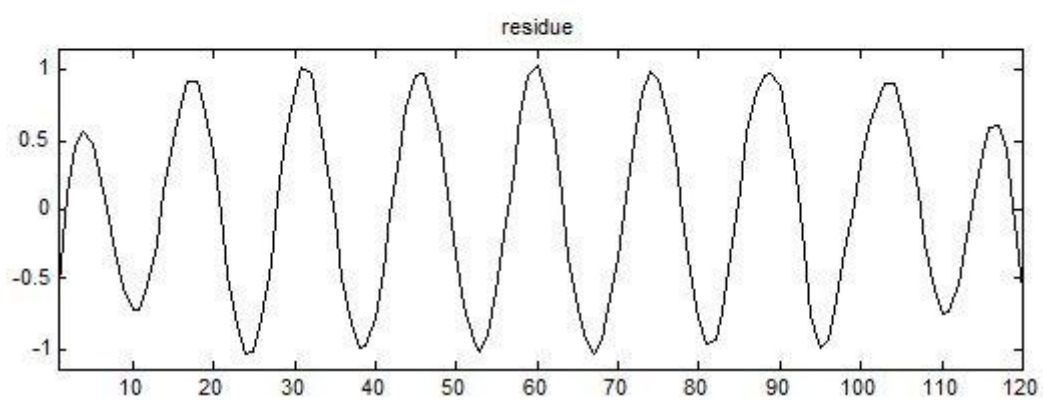
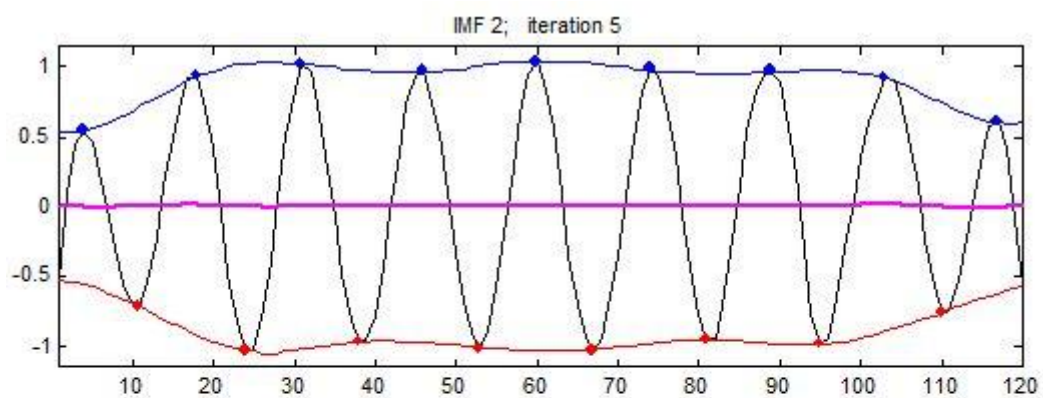
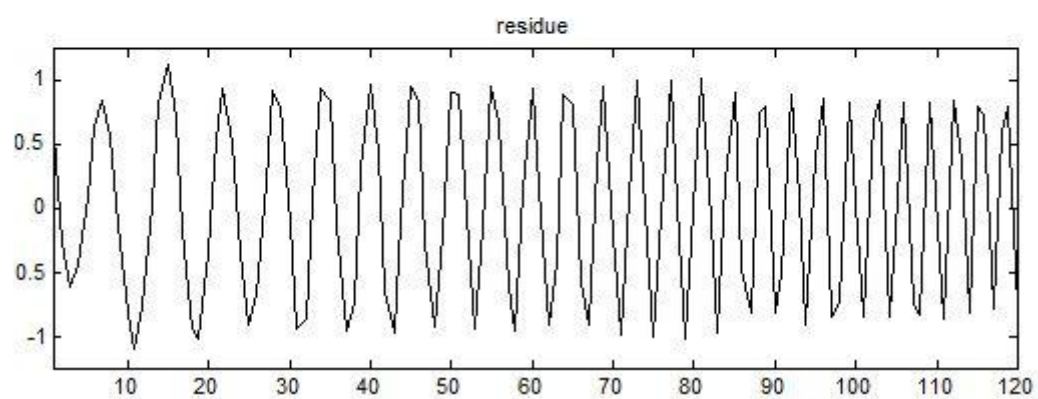
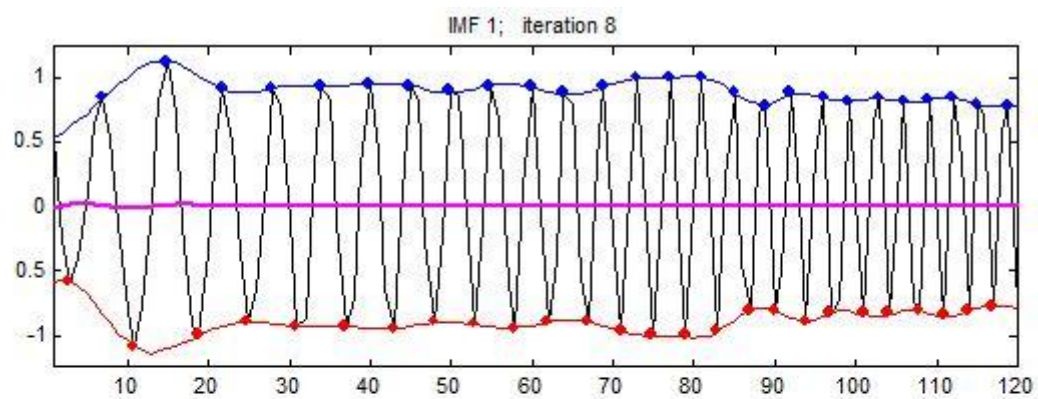
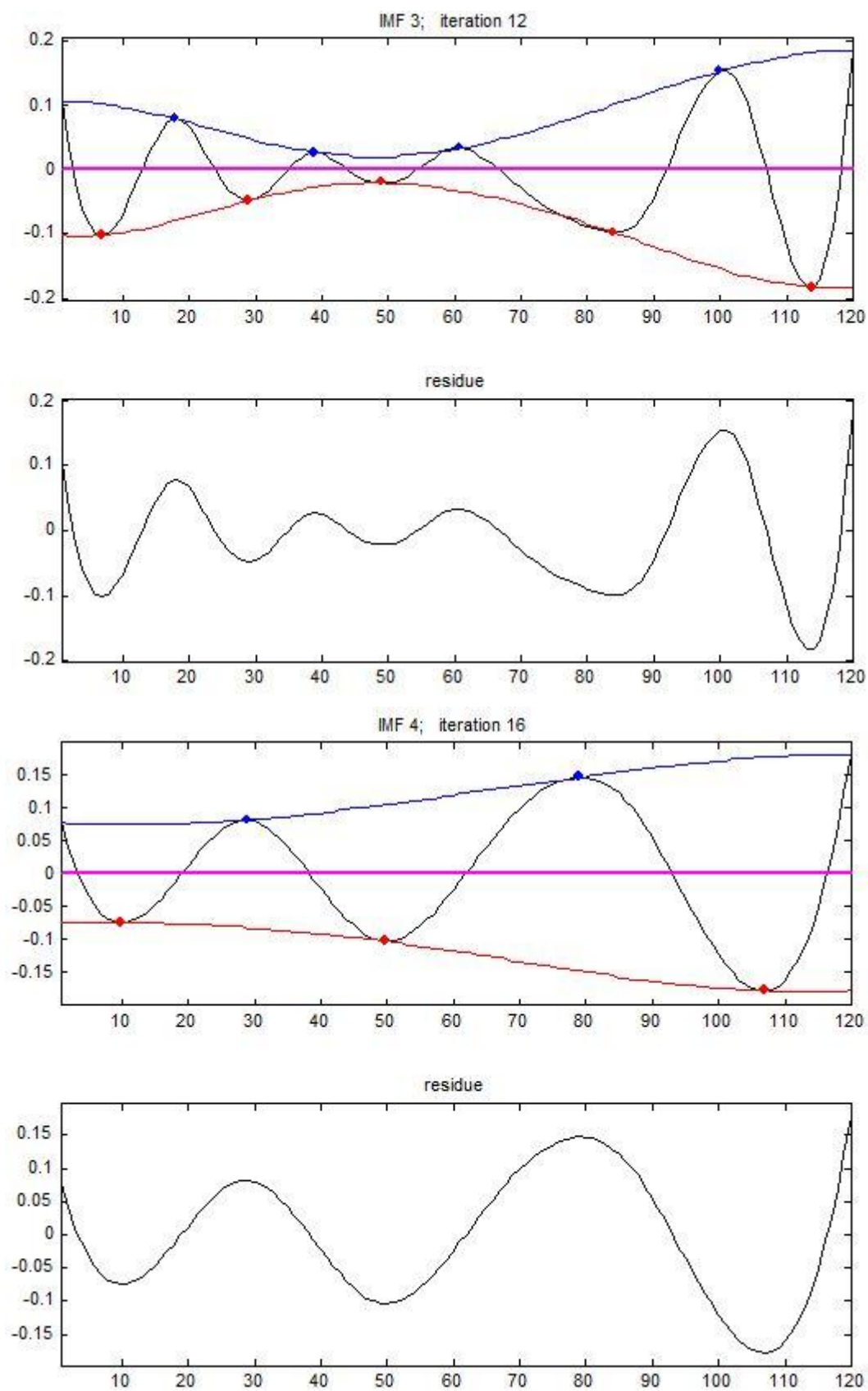


Figure III.5 Processus du tamisage pendant l'extraction de l'IMF1





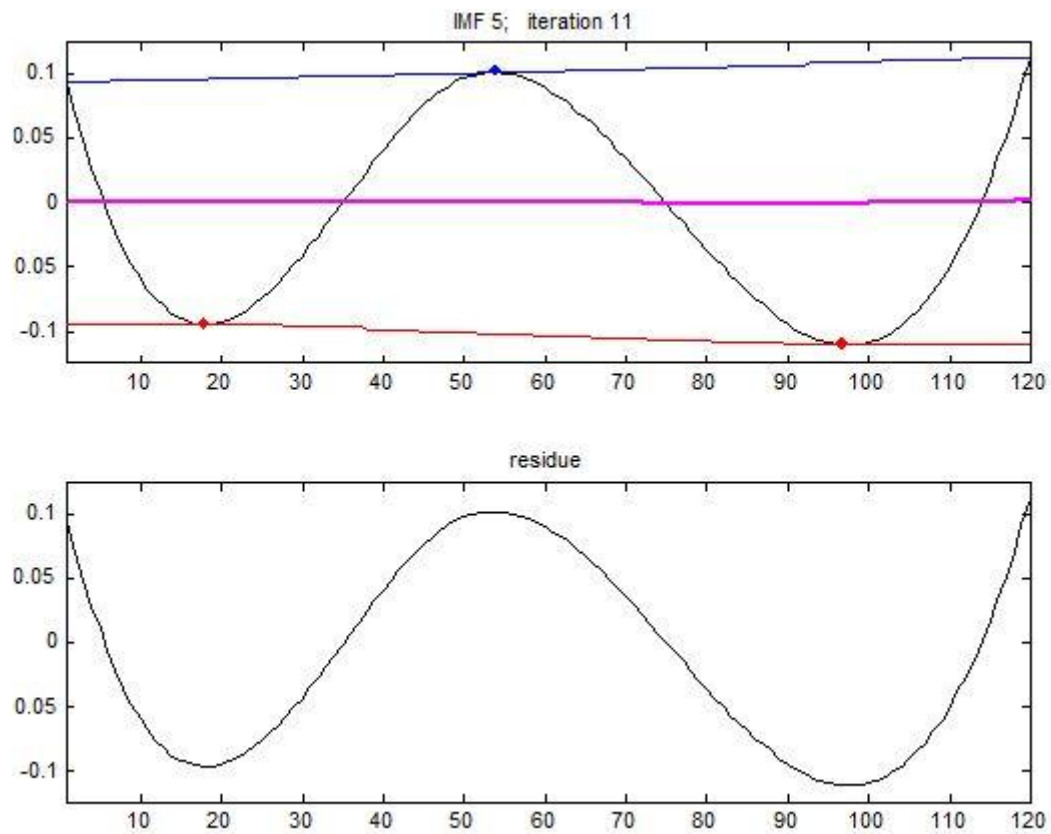


Figure III.6 Différentes composantes IMF et résidu du signal $x(t)$.

Remarque

A chaque itération du processus de tamisage, la moyenne se rapproche de zéro (Figure III.6). En sommant toutes les IMF ainsi le résidu, on retrouve le signal d'origine.

Le résultat de décomposition est donné par la figure III.7

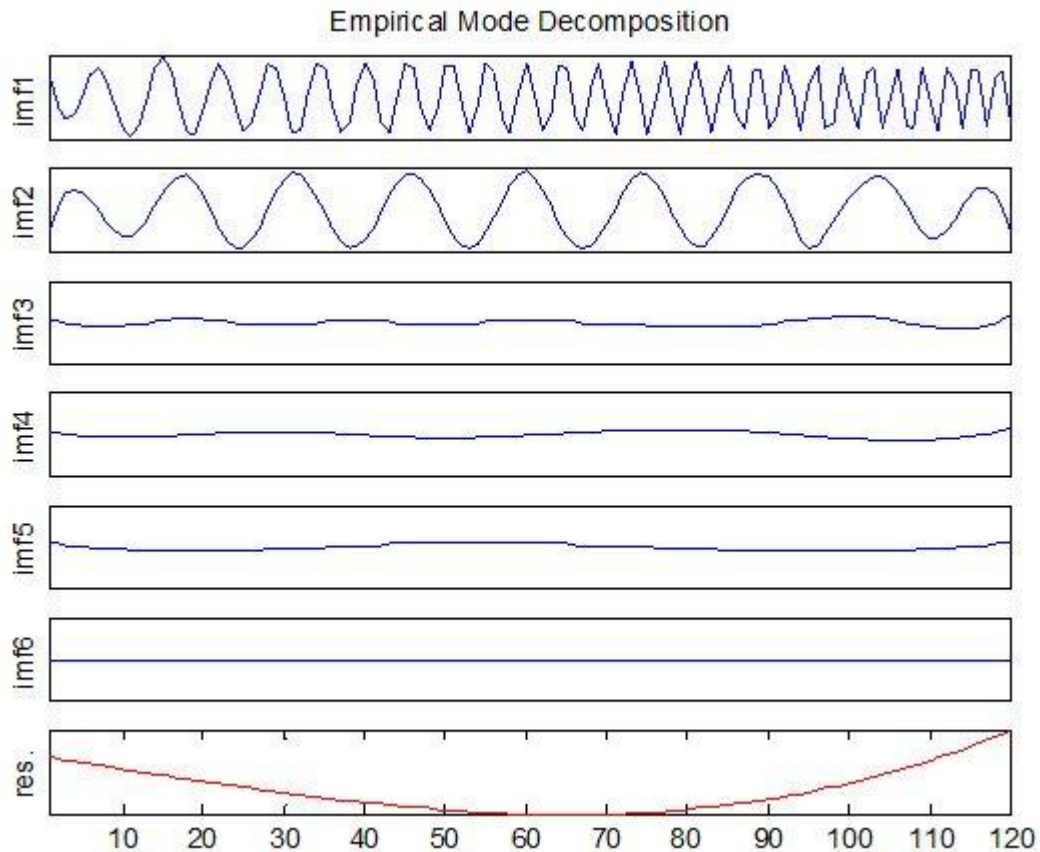


Figure III.7 La décomposition modale empirique du signal $x(t)$.

III.3 La décompositions signal parole par EMD

En utilisant la base de données TIMIT [26], on prend un signal propre (voir figure III.8) et on le noyer dans un bruit blanc gaussien de 5dB dont, les figures III.9 et figure III.10 représentent un bruit blanc gaussien et un signal parole bruité de 5dB respectivement par la suite on applique notre algorithme EMD sur ce signal bruité à fin de le décomposer en plusieurs signaux pour un objectif de débruitage dans la suite de ce chapitre. Le résultat de décomposition est donné par la figure III.11 dont il y a 23 IMF_s pus un résidu.

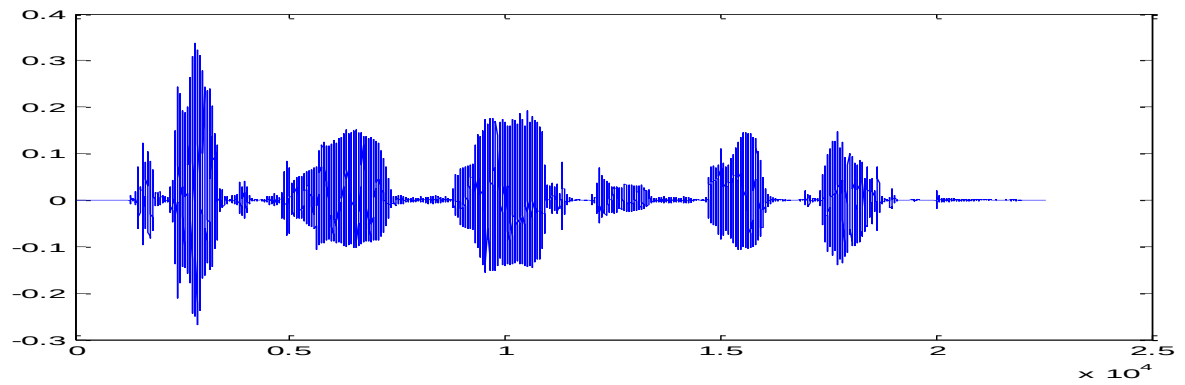


Figure III.8 Signal parole originale

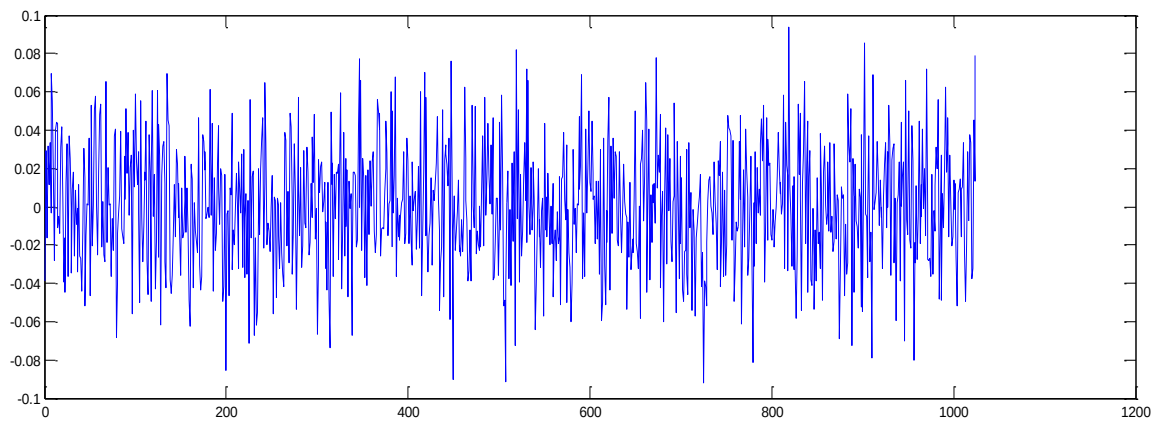


Figure III.9 Bruit blanc gaussien

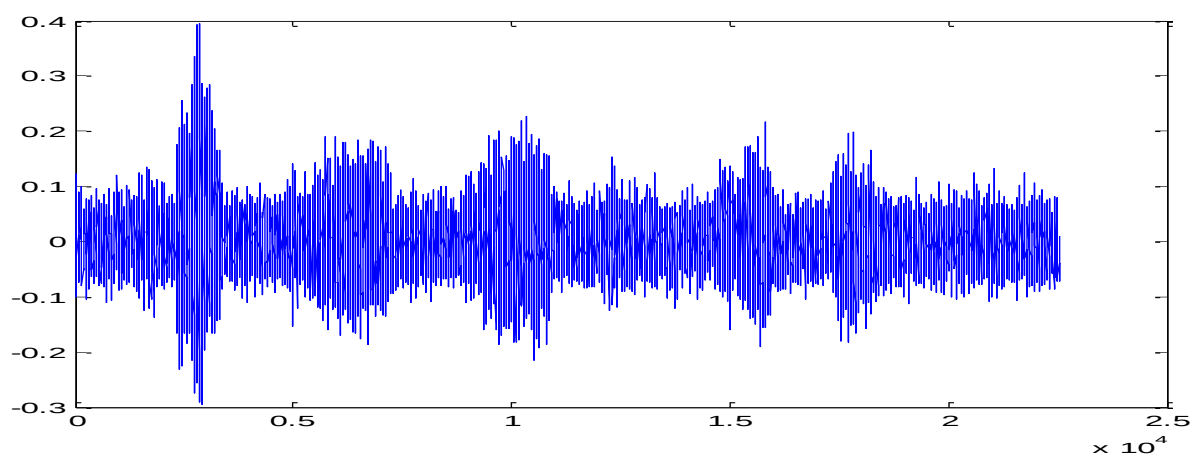
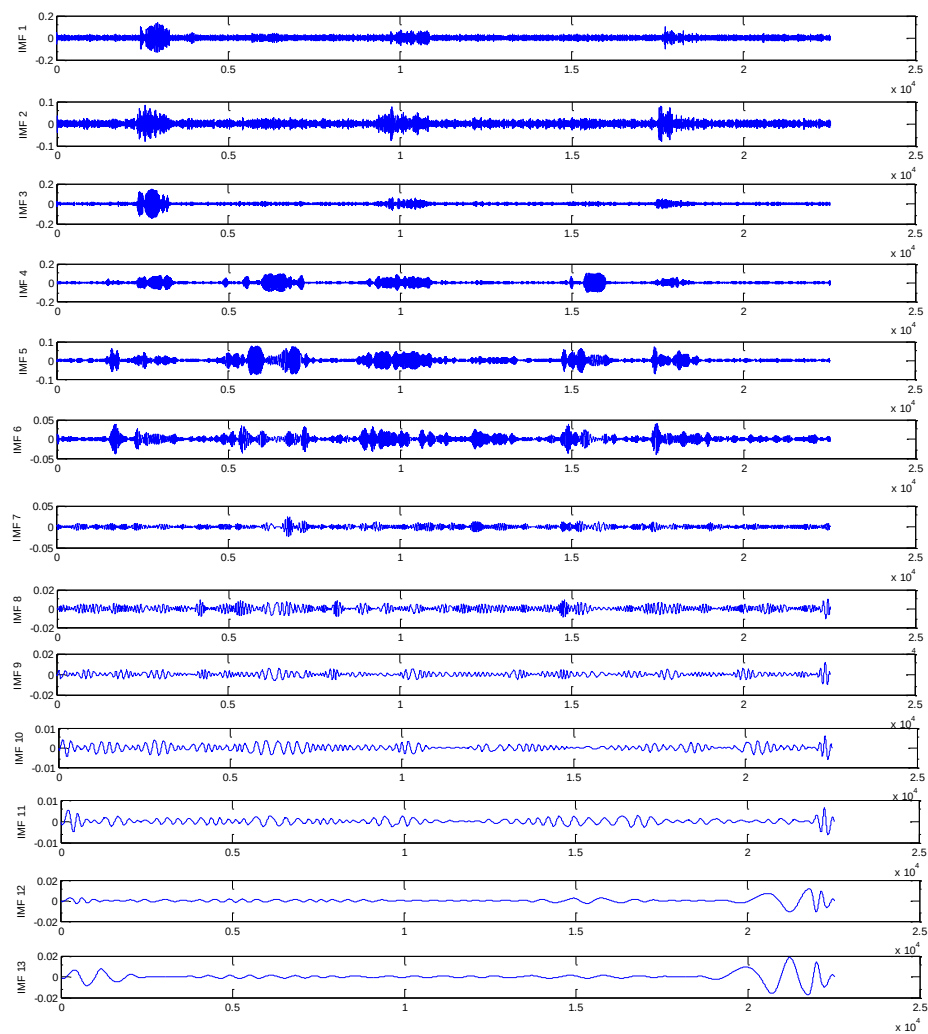


Figure III.10 Signal parole bruité 5dB

- La décompositions par EMD du signal parole bruité :



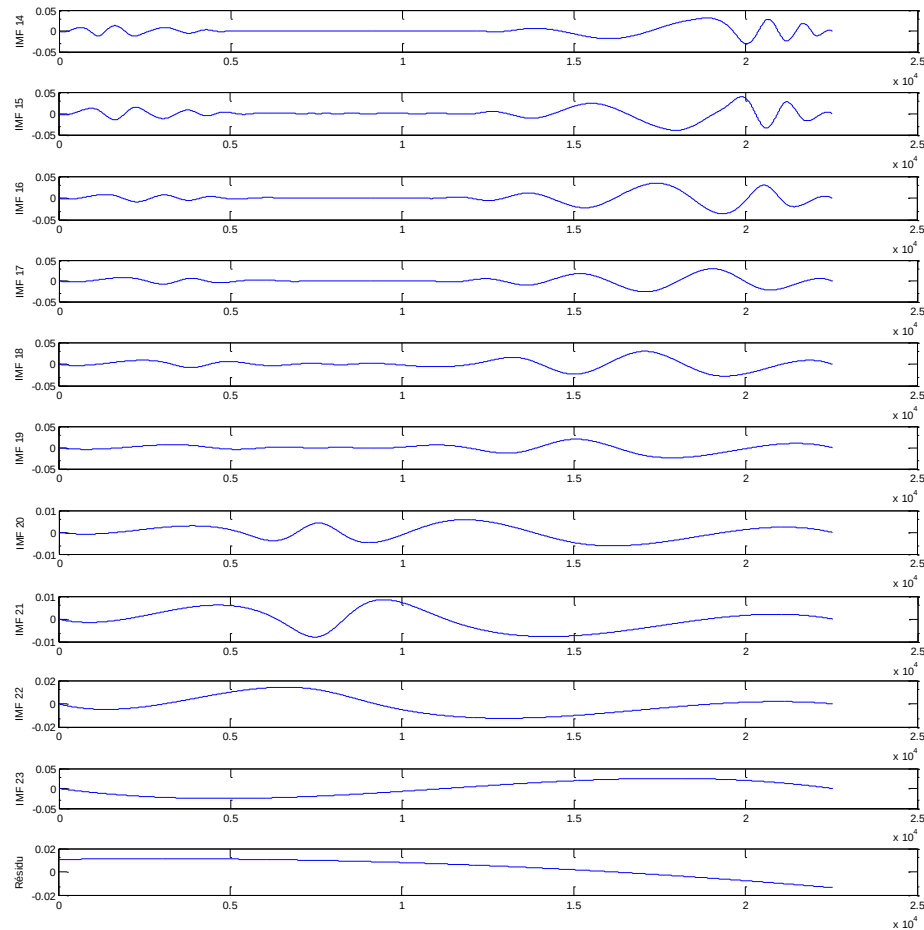


Figure III.11 Détail d'une région voisée des IMFs de 1 à 23 issues de la décomposition par EMD d'un signal de parole corrompu par du bruit blanc gaussien de 5dB.

III.4 Application d'EMD au débruitage

Dans cette partie on applique l'EMD au débruitage par deux méthode l'une est l'approche Naïf, l'autre est une approche s'appuie sur une combinaison de la décomposition par EMD et ondelette.

III.4.1 Première approche

On teste ici les applications potentielles de l'EMD pour le débruitage. Comme chaque IMF est sensé avoir un sens physique en capturant une certaine gamme de fréquences, on peut supposer que le bruit est contenu dans les premières IMFs, qui contiennent les fréquences les plus élevées.

On a donc bruité le signal par un bruit blanc gaussien un signal extrait de la base TIMIT, comme on peut le voir dans la figure III.12. La décomposition en IMFs est présentée dans la figure III.13. A l'œil, on a effectivement la confirmation que le bruit supplémentaire est contenu dans les cinq premières IMFs, mais pas dans les cinq dernières. Il faudrait donc supprimer le bruit qui semble présent dans les premières IMFs (les premières IMFs présentes les hautes fréquences qui correspondent au bruit).

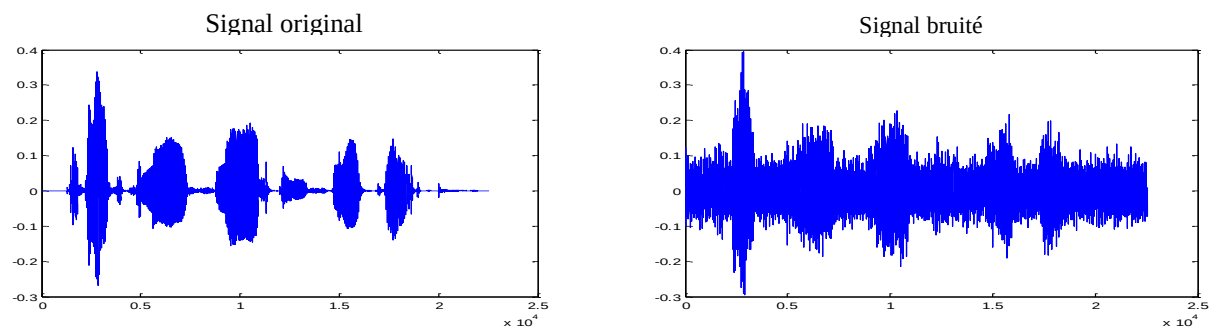


Figure III.12 Signal d'origine et signal bruité à 5dB.

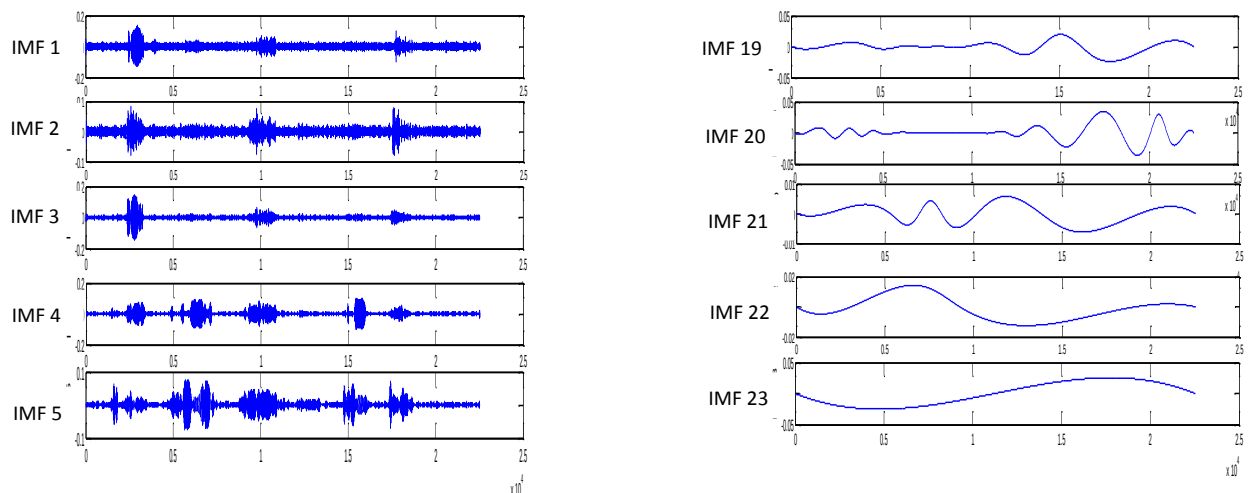


Figure III.13 Les cinq premiers et les cinq derniers IMFs de signal bruité à 5dB

III.4.2 Débruitage naïf

Pour cela, une première idée a été de tout simplement supprimer les premières IMFs. Le résultat est présenté dans la figure III.15. Dans l'ordre, chaque courbe représente la somme partielle des IMFs du signal bruité, ou l'on enlève 5, puis 4, puis 3, puis 2, puis 1 IMF, en partant toujours de la plus oscillante. A l'œil, la figure III.15 montre que la qualité de signal débruité dépend des IMFs éliminés lors de la reconstitution du signal rehaussé (le signal rehaussé est égal à la somme des IMFs sauf les IMFs qu'on a éliminés plus le résidu).

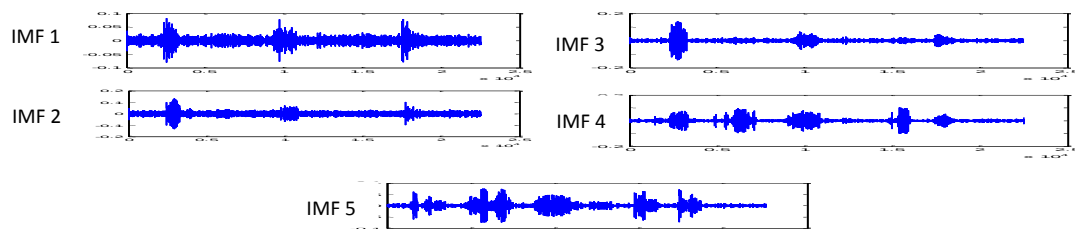


Figure III.14 Les IMFs à éliminer

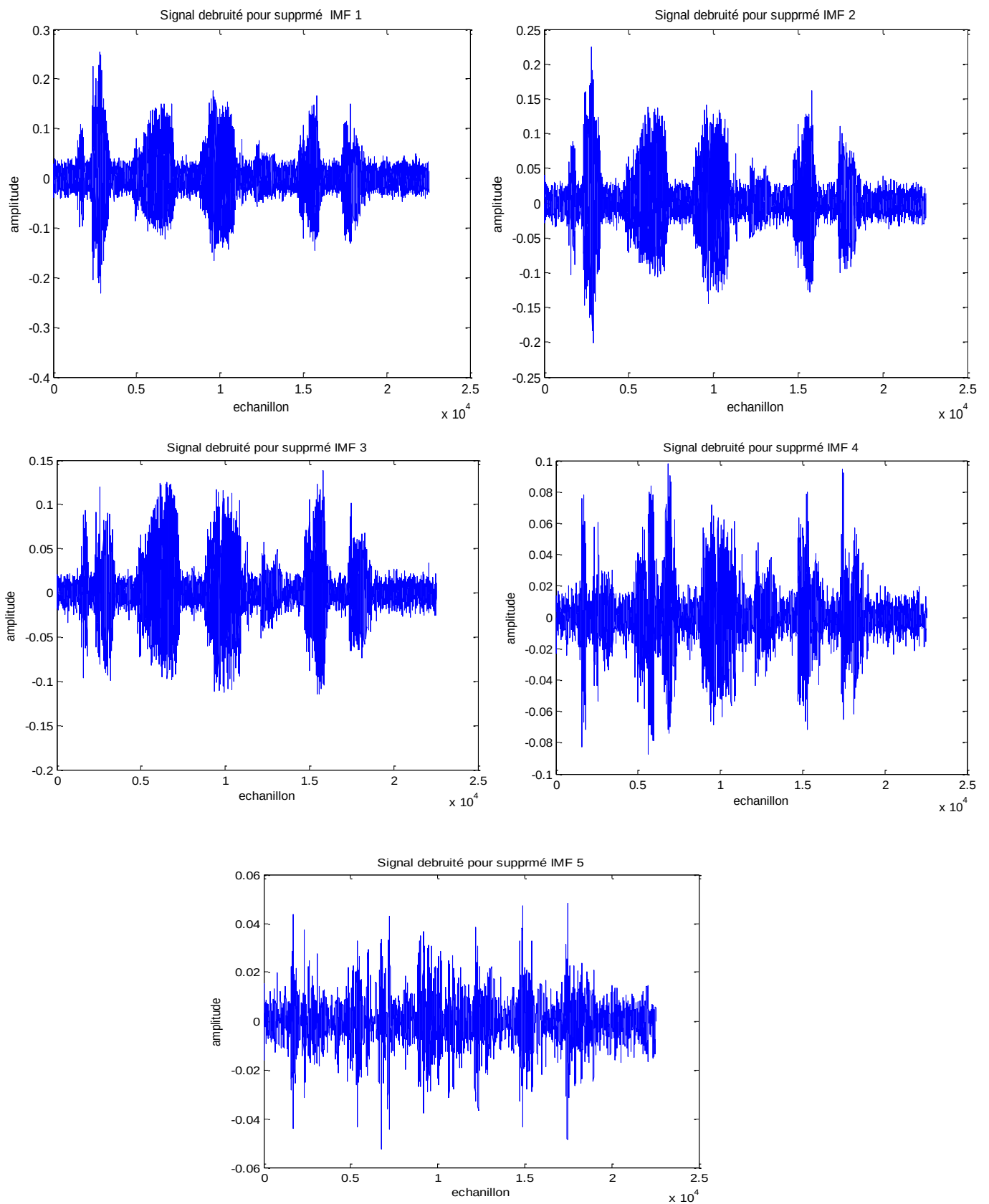


Figure III.15 Reconstructions partielles successives du signal bruité.

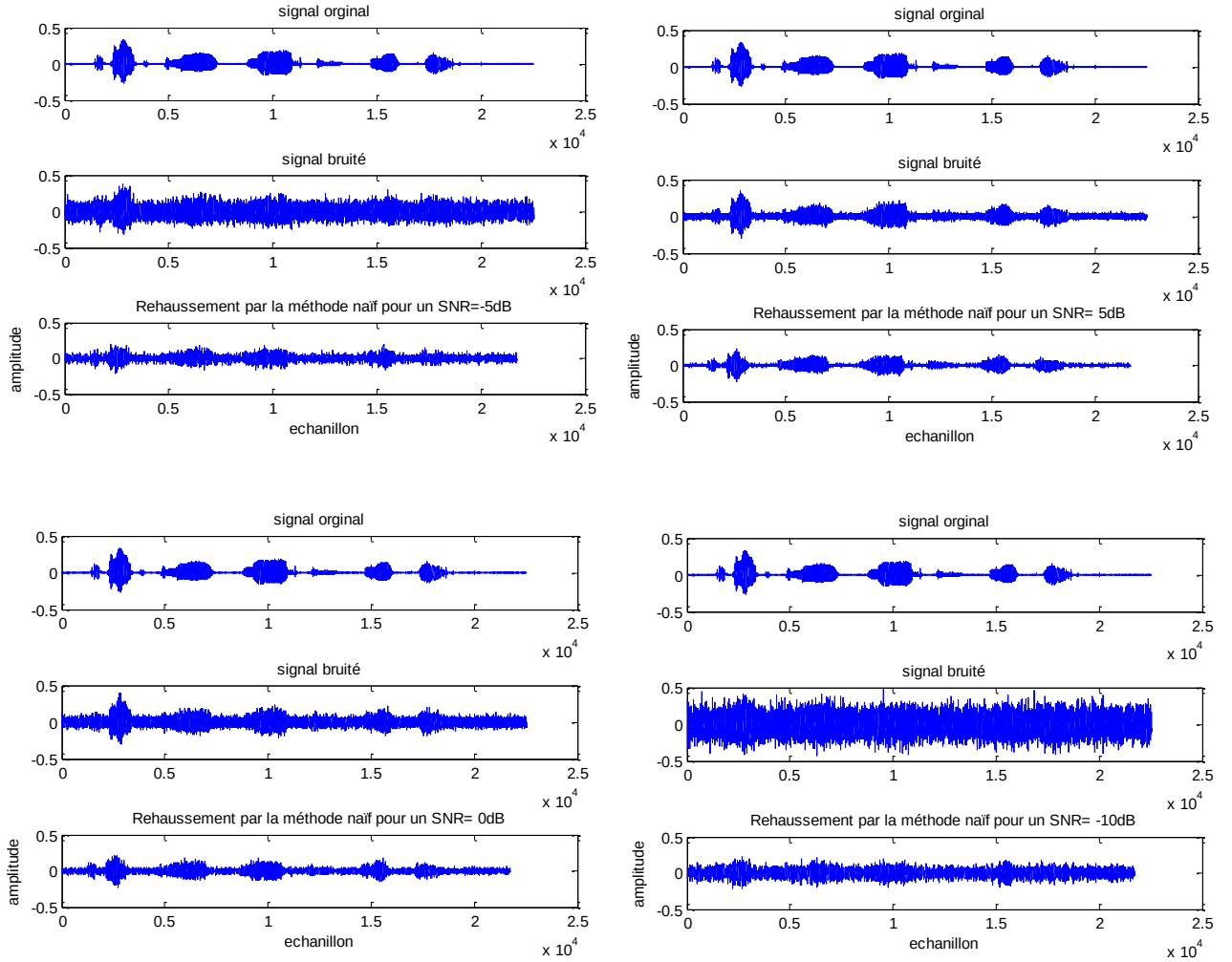


Figure III.16 Débruitage naïf pour différents SNR en ne gardant que les IMFs de 3 à 10 et éliminer les autres IMFs du signal $x(t)$ bruité.

On observe d'après la figure III.16 que la technique Naïf a rehaussé le signal bruité par un bruit blanc gaussien mais cette technique reste n'est pas performante à cause que les IMFs éliminés susceptibles d'avoir des informations utiles.

III.4.3 Débruitage par combinaison EMD-Ondelettes

Dans cette partie on procède à utiliser une autre technique de décomposition s'appelle la décomposition en ondelette [27] et la combiner avec la technique EMD.

La décomposition en ondelette est un processus itératif qui décompose le signal $x(t)$ à l'aide d'un filtre passe haut et d'un filtre passe bas en quadrature. Le filtre passe haut est réalisé à l'aide de la convolution du signal et de l'ondelette choisie. La méthode **EMD-Ondelettes** est proposée pour obtenir un signal rehaussé avec une distorsion minimale et faible perte d'informations. Tout d'abord, le signal bruité est décomposé dans

plusieurs IMFs par EMD de manière adaptative. Ensuite, le seuil d'ondelettes dépendant du niveau est appliqué au maximum des IMFs comportant les premières IMFs, mais cette fois-ci on élimine les IMFs les moins oscillants y compris le résidu, dont avec cette technique **EMD-Ondelettes** procède à traiter chaque IMFs en le débruiter par une ondelette en utilisant un seuillage dépendant de niveau [28] et par la suite on reconstitue le signal rehaussé en rassemblant tous les IMFs soumis au débruitage par l'ondelette. Enfin, le signal est construit par les IMFs traités. Le résultat de débruitage est donné par la figure III.17.

III.4.3.1 Seuillage dépendant du niveau

Johnstone et Silverman ont proposé un seuil dépendant du niveau donné par [28] :

$$T_j = \sigma_j \sqrt{2 \log(N)} \quad (\text{III. 4})$$

N : Les nombres d'échantillonnages.

Où j représente le niveau de décomposition et σ_j l'écart-type du bruit calculé dépendamment du niveau en question qui est donné par :

$$\sigma_j = \frac{\text{mediane}(|d_1^j|)}{0.6745} \quad (\text{III. 5})$$

Où d_1^j représentent les coefficients des détails obtenus au niveau j correspondant au nœud de la plus haute résolution.

✓ **Paramètre de simulation :**

- La décomposition est par EMD dont le résultat a donné 23 IMFs.
- Les IMF à débruiter sont de 1 à 10.
- Ondelette utilisé est Daubechies 1.
- Le nombre de décomposition est de 5 niveaux.
- Le seuillage utilisé est le seuillage doux dépend de niveau.

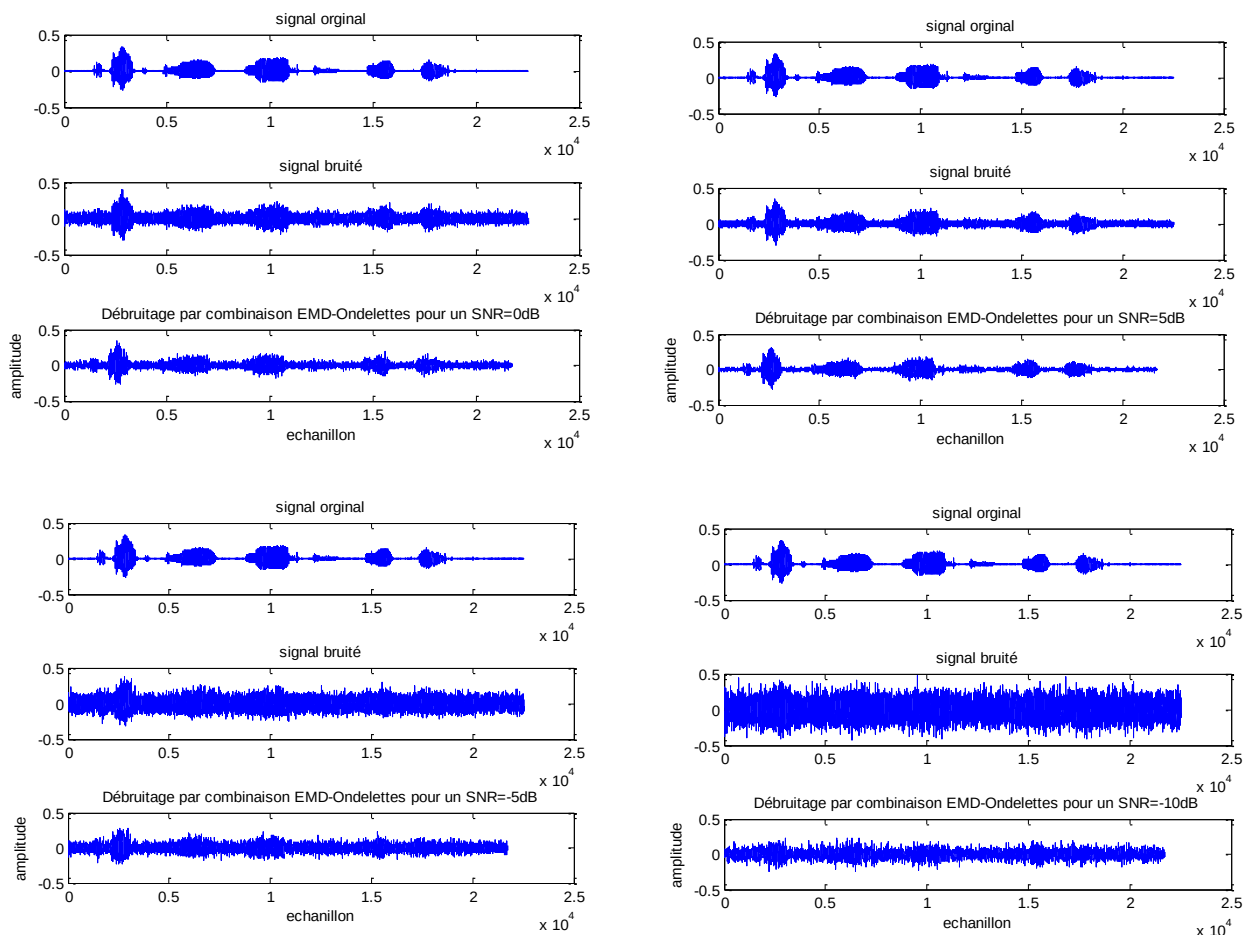


Figure III.17 Résultat rehaussement par combinaison EMD-Ondelettes.

On observe d'après la figure III.17 que la technique qui amélioré le signal bruité par un bruit blanc gaussien est EMD-Ondelettes.

III.5 Comparaison des deux approches de rehaussement de la parole en termes de PESQ, segSNR, en présence de bruit blanc gaussien.

III.5.1 En terme d'SNR

On prend un signal parole de la base de donner TIMIT et en le noyer dans un bruit blanc gaussien de différents niveaux (de -5 à 30dB avec un pas de 5dB). Le tableau III.1 résume les résultats de simulation de la comparaison des deux techniques (la méthode Naïf et la méthode EMD-Ondelettes) point du la vue SNR initial et SNR final après rehaussement. On éclaire que pour la technique Naïf on a pris en considération les IMFs de 3 à 10, pour la technique d'EMD-Ondelettes on a pris en considération les IMF de 1 à 10.

Technique Naïf		Technique EMD-Ondelette	
SNR _{int}	SNR _{fin}	SNR _{int}	SNR _{fin}
-5dB	2.3839 dB	-5 dB	5.8068 dB
0 dB	8.7069 dB	0 dB	11.9897 dB
5 dB	16.9798 dB	5 dB	15.2742 dB
10 dB	22.2837 dB	10 dB	18.3766 dB
15 dB	27.0335 dB	15 dB	22.4848 dB
20 dB	29.8265 dB	20 dB	23.8196 dB
25 dB	32.2316 dB	25 dB	27.0068 dB
30 dB	34.1728 dB	30 dB	28.5063 dB

Tab III.1 Estimation le SNR initial et final pour le signal de parole bruité de la technique Naïf et technique combinaison EMD-Ondelette.

A partir du tableau III.1 nous constatons la technique Naïf donner de bons résultats sur la technique combinaison EMD-Ondelette, on remarque que pour des niveaux de SNR de - 5 dB à 0 dB la technique par combinaison EMD-Ondelette fournit meilleur résultat.

III.5.2 En terme de SegSNR

La figures ci- dessous représentent la comparaison des performances en termes de d' SNR, de -5 à 30 dB par un pas de 5 dB.

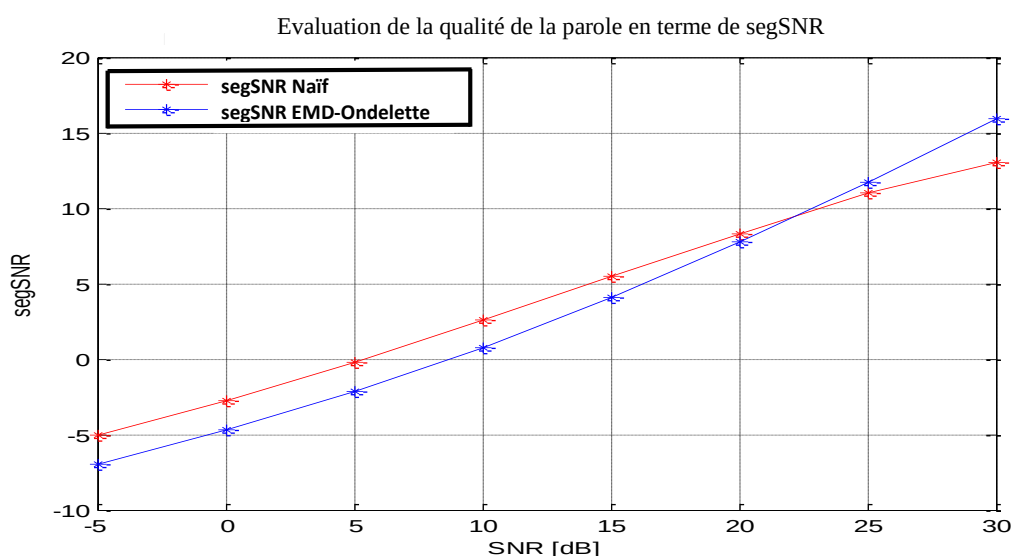


Figure III.18 Comparaison des performances en termes de segSNR en présence du bruit gaussien (SNR=-5 à 30 dB par pas de 5 dB)

Cette figure montre que les méthodes naïf, combinaison EMD-Ondelette seuillage dépendant du niveau, presque le même segSNR pour les SNR allant de -5 dB à 0 dB, tandis que, la méthode a présenté combinaison EMD-Ondelette seuillage dépendant du niveau le mauvais score de segSNR. En d'autre terme, on remarque que pour des niveaux de SNR de 5 dB à 20 dB la technique naïf le meilleur résultat. Pour des niveaux de SNR de 20 dB à 30 dB la méthode combinaison EMD-Ondelette seuillage dépendant du niveau fournit le meilleur résultat, en outre, a présenté le mauvais score de segSNR.

III.5.3 En terme de PESQ

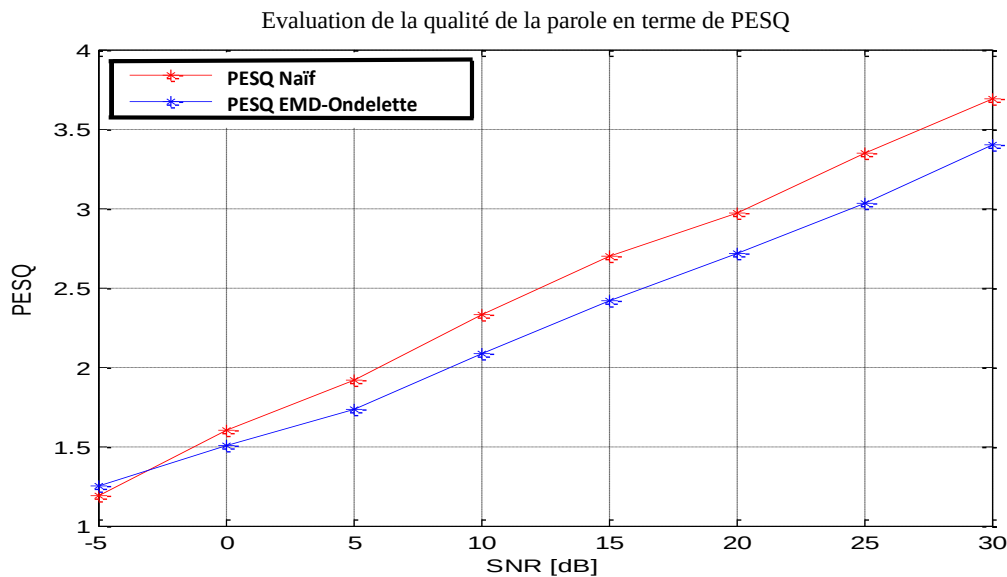


Figure III.19 Comparaison des performances en termes de PESQ en présence du bruit gaussien (SNR=-5 à 30 dB par pas de 5 dB).

La Figure III.19 présente une mesure de PESQ pour les deux approches mentionnées plus haut en présence du bruit blanc gaussien. On remarque que la méthode de débruitage naïf et combinaison EMD-Ondelette dépendant du niveau ont presque le même PESQ pour SNR plus de -5 dB, mais, pour de faibles niveaux de SNR (0 dB à 5 dB) naïf fournit le meilleur résultat. En outre, la méthode de débruitage par combinaison EMD-Ondelette dépendant du niveau présente le mauvais score de PESQ au-dessous de 0 dB à 30 dB.

III.6 Conclusion

Dans ce chapitre nous avons fait une comparaison entre les différentes méthodes d'amélioration du signal parole (éliminé le bruit) en termes de PESQ, segSNR, et à l'aide de logiciel Matlab. On peut résumer les résultats de ce travail :

- On observe d'après le calcul de critère PESQ que la technique naïf est la meilleure pour améliorer le signal parole bruité.
- On observe aussi d'après les calculs des critères, segSNR, que la méthode la plus efficace pour éliminer le bruit est la méthode de naïf est ce dernier à moins de temps d'exécution que le autre technique.

Conclusion générale

Le débruitage de la parole a pour objectif d'améliorer la qualité et l'intelligibilité du signal vocal et par conséquent améliorer les performances des applications en relation. Les techniques de débruitage sont très nombreuses et plusieurs approches ont été proposées dans la littérature telles que les techniques d'estimation du spectre à court-terme, les techniques de sous-espace, les techniques paramétriques, les techniques à base de modèles statistiques, les techniques de filtrage spatio-temporel et les techniques perceptuelles. Dans ce travail, nous avons présenté un projet de recherche sur le débruitage de la parole en utilisant les méthodes de décompositions mode empirique

Notre travail est divisé en trois chapitres, le premier sert de présenter en général que est ce qu'un signal parole, comment fonctionner l'appareil vocal et la production de signal parole et la méthode de la décompositions mode empirique EMD.

Dans le deuxième chapitre nous avons présenté les différentes techniques d'amélioration d'EMD.

Enfin le dernier chapitre contient les résultats de simulation à l'aide logiciel MATLAB et contient aussi la comparaison entre les différentes méthodes en mesurant les critères objectifs tel que PESQ, segSNR, où ses résultats démontrent à nous que la meilleure technique pour améliorer le signal parole bruité (Blanc gaussienne) est la technique naïf en termes de PESQ.

Afin d'améliorer la qualité du signal transmis au correspondant distant, d'accroître son intelligibilité et de réduire la fatigue de ce dernier, il s'avère important de développer des systèmes de réduction de bruit dont le but consiste à extraire l'information utile en effectuant un traitement sur le signal d'observation bruité. En plus de ces applications de communication parlée, l'amélioration de la qualité du signal de parole s'avère également nécessaire pour la reconnaissance vocale, dont les performances sont fortement altérées lorsque l'utilisateur est plongé dans un environnement bruité.

La réduction de bruit s'applique principalement au domaine du traitement du signal sonore (parole ou musique) dont voici une liste non exhaustive:

- ✓ téléconférence et visioconférence en milieu bruité (en salle dédiée ou bien à partir d'un ordinateur multimédia, etc..)
- ✓ téléphonie : traitement au niveau du terminal (terminaux classiques et terminaux mobiles, et au sein du réseau de transport)
- ✓ terminaux mains-libres (bureau, terminaux mobiles, terminaux fixes embarqués en véhicules)
- ✓ prise de son dans les lieux publics (gare, aéroport, rue, etc..)

- ✓ prise de son mains-libres dans les véhicules,
- ✓ reconnaissance de parole robuste à l'environnement acoustique,
- ✓ restauration d'enregistrements anciens,
- ✓ prise de son pour le cinéma et les médias (radio, télévision, par exemple pour le journalisme sportif ou les concerts, etc.).

Le principe de la réduction de bruit est également applicable à tous les domaines où l'on cherche à extraire une information utile à partir d'une observation bruitée. On peut notamment envisager les domaines suivants : imagerie sous-marine, télédétection sous-marine, traitement des signaux biomédicaux (EEG, ECG, imagerie biomédicale, etc.).

Références Bibliographiques

- [1] R. Boite. Traitement de la parole. Collection Electricité. Presses Polytechniques et Universitaires Romandes, 2000.
- [2] J.-L. Duchet, La phonologie, Collection "Que sais-je ?", Presses Universitaires de France, 1981.
- [3] Y. AZIZA, "Modélisation AR et ARMA de la Parole pour une Vérification Robuste du Locuteur dans un Milieu Bruité en Mode Dépendant du Texte" thèse magister de l'universite ferhat abbas –setif1-ufas (algerie) 2013.
- [4] Calliope, La parole et son traitement automatique, Édition Masson et CNET-ENST, Paris, 1989.
- [5] M. Campedel Oudot. Étude du modèle "sinusoïdes et bruit" pour le traitement des signaux de parole. estiation robuste de l'enveloppe spectr ale. PhD thesis, École Nationale Supérieure des Télécommunications, 1998.
- [6] A. J. Gallimard, Extrait de L'oreille musicienne, planche n°III.7.2001.
- [7] F. Dell, Les règles et les sons, Éditions Hermann, 1985.
- [8] N. Chomsky and M. Halle, The sound pattern of English, Harper and Row, 1968.
- [9] R. Jakobson, C.G.M. Fant and M. Halle, "Preliminaries to speech analysis: The distinctive features and their correlates", Technical Report n° 13, Acoustics Laboratory, Massachussetes institute of technology, 1952.
- [10] R. Boite et all, Traitement de la parole, presses polytechnique et universitaires Romandes, Novembre 1999.
- [11] S. Nefti, "Segmentation automatique de parole en phones. Correction d'étiquetage par l'introduction de mesures de confiance" Thèse Docteur de l'Université de Rennes 1, 2004.
- [12] S. J. Andaloussi. Indexation de l'information médicale : Application à la recherche d'images et de vidéos par le contenu. Thèse doctorat, Université européenne de Bretagne, 2010.
- [13] K. Khaldi" Traitement Et analyse Des Signaux Sonores Par Transformée De Huang (EMD)", Thèse de Doctorat l'Université européenne de Bretagne, 2012 .
- [14] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu. The empirical mode decomposition and Hilbert spectrum for non-

linear and non-stationary time series analysis. Proceedings of the Royal Society of London, 454(1971):903–995, 1998.

[15] H S. Diop. Modèles AM-FM et approche par équations aux dérivées partielles de la décomposition modale empirique pour l'analyse des signaux et des images. Thèse doctorat, université de Rennes1, 2009.

[16] P. Flandrin, G. Rilling, et Gonçalves P, Empirical mode decomposition as a filter bank. IEEE Signal Processing Letters, 11(2) : 112–114, 2004.

[17] W. Mohguen , " Améliorations de la méthode EEMD", thèse de magister, universite ferhat abbas-setif-1-ufas(algerie),2014 .

[18] S. GADA," Décomposition Modale Empirique : Application à l'analyse de la texture", Mémoire De Magister, Université Mouloud Mammeri de Tizi-Ouzou, 2013.

[19] A. Girard," Réduction de bruit des signaux de parole mono-capteur basée sur la modélisation par EMD", Mémoire De Magister, Université de Sherbrooke, 2010.

[20] Wu _Z.Huang N. E., Ensemble empirical mode decomposition : a noise-assisted data analysis method. Adv. Adapt. Data Anal., 1 (1) : 1–41, 2009.

[21] <https://www.kaggle.com/wiki/RootMeanSquaredError>.

[22] J. Zhang, R Yan, R Gao. X., Z Feng, Performance enhancement of ensemble empirical mode decomposition, Mechanical Systems and Signal Processing., 24 (7): 2104-2123, 2010.

[23] N. Huang . E., Yeh J. R., and Shieh J. S., Complementary ensemble empirical mode decomposition a novel noise enhanced data analysis method, Advances in Adaptive Data Analysis., Vol. 2, No. 2, 135–156, 2010.

[24] Y. Hu & P. Loizou. Evaluation of objective Measures for speech enhancement. in Proc. Interspeech, pages 1447–1450, 2006.

[25] H. Sharabaty. Diagnostic de La Somnolence d'un Operateur: Analyse automatique de signaux physiologiques. Thèse doctorat, Université de Toulouse. 2007.

[26] J. S Arofolo , Lamel, L. F, Fisher, W. M., Fiscus, J. G., Pallett, D. S., and Dahlgren, N.L., "DARPA TIMIT Acoustic Phonetic Continuous Speech Corpus CDROM," NIST, 1993.

[27] S. Mallat, A wavelet tour of signal processing. New York: Academie Press.1998.

[28] I. M Johnstone, and Silverman, B.W., Wavelet threshold estimators for data with correlated noise, J. Roy. Statist. Soc. B.1997.