

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة سعد دحلب البلدية
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Master

en Télécommunication
Spécialité : Réseaux & Télécommunications

Présenté par

BELALIA Nassim

&

HADJALLA HEDJALLA Mohamed Amine

Développement d'un algorithme RNQ efficace pour la réduction bi-capteurs du bruit acoustique

Proposé par : Pr. DJENDI Mohamed

Année Universitaire 2021-2022

Remerciements

La réalisation de ce mémoire a été possible grâce au concours de plusieurs personnes à qui nous voudrions témoigner toute notre gratitude.

Nous voudrions tout d'abord adresser toute notre reconnaissance au directeur de ce mémoire, Monsieur DJENDI Mohamed, qui fut le premier à nous faire découvrir le sujet qui a guidé notre mémoire, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter notre réflexion.

Nous désirons aussi remercier les professeurs de l'université de Blida 1, qui nous ont fourni les outils nécessaires à la réussite de nos études universitaires.

Nous voudrions aussi exprimer notre reconnaissance envers les amis et collègues qui ont apporté leur soutien moral et intellectuel tout au long de notre démarche.

Un grand merci à notre professeur BENSEBNTI Messaoud pour ses conseils concernant notre style d'écriture, ils ont grandement facilité notre travail.

A mes parents qui m'ont soutenu et encouragé durant ces années d'études.

Qu'ils trouvent ici le témoignage de ma profonde reconnaissance.

A mes frères, mes grands-parents et Ceux qui ont partagé avec moi tous les moments d'émotions lors de la réalisation de ce travail. Ils m'ont chaleureusement supporté et encouragé tout au long de mon parcours.

A ma famille, mes proches et à ceux qui me donnent de l'amour et de la vivacité.

A mes amis qui m'ont toujours encouragé, et à qui je souhaite plus de succès.

A tous ceux que J'aime.

Mohamed

Je dédie ce travail, à mes chers parents, aucune dédicace ne saurait exprimer mon respect, mon amour pour les efforts qu'ils ont fourni pour mon instruction et mon bien être. J'espère que je pourrais leurs rendre un peu de ce qu'ils ont fait pour moi un jour,

A mon cher frère qui était toujours présent dans les moments difficiles.

A tous mes amis, et à mon binôme pour sa volonté et pour son soutien durant tout ce mémoire.

Merci à vous.

Nassim

ملخص: يهدف مشروع نهاية الدراسة هذا إلى تطوير خوارزمية تكيفية جديدة تهدف إلى تحسين الجودة المتصورة للكلام عن طريق تقليل الضوضاء الصوتية. تم تأكيد كفاءة هذه الخوارزمية في نتيجة المحاكاة بين هذه الخوارزمية المقترحة وخوارزمية التدرج العشوائي الكلاسيكية.

كلمات المفاتيح: خوارزمية تكيفية تلقائية.

Résumé : Ce projet de fin d'étude a pour objectif de développer un nouvel algorithme RNQ automatique dont le but est l'amélioration de la qualité perçue de la parole par la réduction du bruit acoustique.

L'efficacité de cet algorithme est confirmée à travers les résultats de simulation obtenus par la comparaison entre l'algorithme proposé et l'algorithme de gradient stochastique normalisé (NLMS).

Mots clés : RNQ ; algorithme adaptatif ; NLMS.

Abstract : This end-of-study project aims to develop a new automatic RNQ adaptive algorithm to improve the perceived quality of the speech by reducing the acoustic noise.

The efficiency of this algorithm is confirmed in the simulation results between this proposed algorithm and the normalized stochastic gradient algorithm NLMS.

Keywords : RNQ ; NLMS ; adaptive algorithm.

Listes des acronymes et abréviations

- DAV : Détection d'Activité Vocale.
- EQM : Erreur Quadratique Moyenne.
- LMS : Least Mean Squares.
- MSE : Mean Squared Error.
- NLMS : Normalized Least Mean Squares.
- RSB : Rapport Signal à Bruit.
- SAS : Séparation aveugle de source.
- SegMSE: Segmental Mean Squared Error.
- SegSNR: Segmental Signal to Noise Ratio.
- SM : System Mismatch.
- USASI : United States of America Standard Institute.
- FBSS : Forward Blind Source Separation.
- BBSS : Backward Blind Source Separation.
- RNQ : Récursif Non-Quadratic.
- dB : décibel.
- GN : Gaussian Noise.

Table des matières

Introduction générale	1
------------------------------------	----------

Chapitre 1 Généralités sur le filtrage et les algorithmes adaptatifs

1.1	Introduction	3
1.2	Notions de base sur la parole et le bruit	3
1.2.1	Qu'est-ce qu'un signal ?	3
1.2.2	Signal de parole	4
1.2.3	Bruit acoustique	5
1.2.4	Sources de bruit	5
1.2.5	Types de bruits	6
1.2.6	Rapport signal à bruit (RSB)	6
1.3	Filtrage de Wiener	7
1.3.1	Erreur quadratique moyenne	8
1.3.2	Principe d'orthogonalité	9
1.4	Filtrage adaptatif	10
1.4.1	Présentation des algorithmes de filtrage adaptatif	10
1.4.2	Choix des algorithmes adaptatifs	10
1.4.3	Algorithme du gradient stochastique (Least Mean Square – LMS)	11
1.4.4	Algorithme LMS normalisé (Normalized LMS – NLMS)	12
1.5	Conclusion	13

Chapitre 2 Techniques de réduction du bruit

2.1	Introduction	14
2.2	Réduction de bruit acoustique	14
2.2.1	Notion de canal acoustique	14
2.2.2	Réponse impulsionnelle acoustique	15
2.3	Méthode mono-capteur	16

2.4	Méthode bi-capteurs	17
2.4.1	Méthode bi-capteurs simplifiée	18
2.5	Méthode multi-capteurs	19
2.6	Réduction de bruit acoustique par la séparation de source aveugle	19
2.6.1	Caractéristiques des signaux source	20
2.6.2	Structure directe (FBSS)	20
2.6.3	Structure récursive (BBSS)	22
2.7	Conclusion	23

Chapitre 3 Proposition d'un nouvel algorithme Récursif Non-Quadratic automatique

3.1	Introduction	24
3.2	Réduction de bruit par l'algorithme NLMS	24
3.3	Réduction de bruit par un nouvel algorithme RNQ automatique	25
3.4	Conclusion	32

Chapitre 4 Résultats de simulation

4.1	Introduction	33
4.2	Signaux utilisés aux simulations	33
4.2.1	Signal de la parole	33
4.2.2	Bruit blanc	34
4.2.3	Réponses impulsionnelles	34
4.2.4	Signaux de mélange	35
4.3	Détection d'activité vocale (DAV)	36
4.4	Evaluation du nouvel algorithme proposé	36
4.5	Conclusion	53
Conclusion générale		54
Bibliographie		55

Liste des figures

Figure 1.1 :	Appareil phonatoire humain.	3
Figure 1.2 :	Autocorrélation et densité spectrale théoriques d'un bruit blanc.	5
Figure 1.3 :	Représentation du filtrage de Wiener.	6
Figure 1.4 :	Principe du filtre adaptatif.	9
Figure 2.1 :	Propagation du son dans une salle acoustique.	14
Figure 2.2 :	Réponse impulsionnelle dispersive ($L = 64$).	14
Figure 2.3 :	Principe de débruitage mono-capteur.	15
Figure 2.4 :	La structure complète d'un mélange convolutif bi-capteurs.	16
Figure 2.5 :	La structure simplifiée d'un mélange convolutif bi-capteurs.	17
Figure 2.6 :	La structure d'un mélange convolutif multi capteurs (cas général).	18
Figure 2.7 :	Structure directe de la SAS.	19
Figure 2.8 :	Structure récursive de la SAS.	21
Figure 3.1 :	Structure d'annulation adaptative du bruit par l'algorithme NLMS.	23
Figure 3.2 :	Schéma bloc simplifié de la nouvelle structure de réduction de bruit basée sur l'algorithme RNQ et un détecteur d'activité vocale automatique.	25
Figure 3.3 :	Schéma détaillé de la nouvelle structure de réduction de bruit basée sur l'algorithme RNQ et un détecteur d'activité vocale automatique.	25

Figure 3.4 :	Schéma du principe du fonctionnement de l'algorithme RNQ.	29
Figure 4.1 :	Signal de la parole utilisé et son spectrogramme.	33
Figure 4.2 :	Signal du bruit blanc et son spectrogramme.	34
Figure 4.3 :	La réponse impulsionnelle h_{21} (à droite) et la réponse impulsionnelle h_{12} (à gauche) avec $L = 64$.	34
Figure 4.4 :	Signal du premier mélange $m_1(n)$ et son spectrogramme.	35
Figure 4.5 :	Signal du deuxième mélange $m_2(n)$ et son spectrogramme.	35
Figure 4.6 :	Signal de la parole avec sa DAV.	36
Figure 4.7 :	Evolution temporelle des signaux pour un $RSB = -3 dB$ et $L = 64$.	38
Figure 4.8 :	Critère SegMSE pour un $RSB = -3 dB$ et $L = 64$.	38
Figure 4.9 :	Critère SegSNR pour un $RSB = -3 dB$ et $L = 64$.	39
Figure 4.10 :	Critère SM pour un $RSB = -3 dB$ et $L = 64$.	39
Figure 4.11 :	Evolution temporelle des signaux pour un $RSB = 0 dB$ et $L = 64$.	40
Figure 4.12 :	Critère SegMSE pour un $RSB = 0 dB$ et $L = 64$.	40
Figure 4.13 :	Critère SegSNR pour un $RSB = 0 dB$ et $L = 64$.	40
Figure 4.14 :	Critère SM pour pour un $RSB = 0 dB$ et $L = 64$.	40

Figure 4.15 :	Evolution temporelle des signaux pour un $RSB = 3 \text{ dB}$ et $L = 64$.	41
Figure 4.16 :	Critère SegMSE pour un $RSB = 3 \text{ dB}$ et $L = 64$.	41
Figure 4.17 :	Critère SegSNR pour un $RSB = 3 \text{ dB}$ et $L = 64$.	41
Figure 4.18 :	Critère SM pour un $RSB = 3 \text{ dB}$ et $L = 64$.	41
Figure 4.19 :	Evolution temporelle des signaux pour un $RSB = -3 \text{ dB}$ et $L = 128$.	42
Figure 4.20 :	Critère SegMSE pour un $RSB = -3 \text{ dB}$ et $L = 128$.	42
Figure 4.21 :	Critère SegSNR pour un $RSB = -3 \text{ dB}$ et $L = 128$.	43
Figure 4.22 :	Critère SM pour un $RSB = -3 \text{ dB}$ et $L = 128$.	43
Figure 4.23 :	Evolution temporelle des signaux pour pour un $RSB = 0 \text{ dB}$ et $L = 128$.	44
Figure 4.24 :	Critère SegMSE pour pour un $RSB = 0 \text{ dB}$ et $L = 128$.	44
Figure 4.25 :	Critère SegSNR pour pour un $RSB = 0 \text{ dB}$ et $L = 128$.	44
Figure 4.26 :	Critère SM pour pour un $RSB = 0 \text{ dB}$ et $L = 128$.	44
Figure 4.27 :	Evolution temporelle des signaux pour un $RSB = 3 \text{ dB}$ et $L = 128$.	45
Figure 4.28 :	Critère SegMSE pour un $RSB = 3 \text{ dB}$ et $L = 128$.	45
Figure 4.29 :	Critère SegSNR pour un $RSB = 3 \text{ dB}$ et $L = 128$.	45

Figure 4.30 :	Critère SM pour un $RSB = 3 \text{ dB}$ et $L = 128$.	45
Figure 4.31 :	Signal du bruit USASI et son spectrogramme.	46
Figure 4.32 :	Evolution temporelle des signaux pour un $RSB = -3 \text{ dB}$ et $L = 64$.	47
Figure 4.33 :	Critère SegMSE pour un $RSB = -3 \text{ dB}$ et $L = 64$.	47
Figure 4.34 :	Critère SegSNR pour un $RSB = -3 \text{ dB}$ et $L = 64$.	48
Figure 4.35 :	Critère SM pour un $RSB = -3 \text{ dB}$ et $L = 64$.	48
Figure 4.36 :	Signal du bruit d'automobile.	48
Figure 4.37 :	Signal du bruit de la pluie.	48
Figure 4.38 :	Spectrogramme du signal du bruit d'automobile.	49
Figure 4.39 :	Spectrogramme du signal du bruit de la pluie.	49
Figure 4.40 :	Evolution temporelle des signaux pour un $RSB = -3 \text{ dB}$ et $L = 64$ (automobile).	50
Figure 4.41 :	Evolution temporelle des signaux pour un $RSB = -3 \text{ dB}$ et $L = 64$ (pluie).	50
Figure 4.42 :	Critère SegMSE pour un $RSB = -3 \text{ dB}$ et $L = 64$ (automobile).	50
Figure 4.43 :	Critère SegMSE pour un $RSB = -3 \text{ dB}$ et $L = 64$ (pluie).	50
Figure 4.44 :	Critère SegSNR pour un $RSB = -3 \text{ dB}$ et $L = 64$ (automobile).	51

Figure 4.45 :	Critère SegSNR pour un $RSB = -3 \text{ dB}$ et $L = 64$ (pluie).	51
Figure 4.46 :	Critère SM pour un $RSB = -3 \text{ dB}$ et $L = 64$ (automobile).	51
Figure 4.47 :	Critère SM pour un $RSB = -3 \text{ dB}$ et $L = 64$ (pluie).	51
Figure 4.48 :	Critère SegMSE pour différentes valeurs de j .	52
Figure 4.49 :	Critère SegMSE pour différentes valeurs de λ .	52

Introduction générale

L'essor fantastique des télécommunications de ces vingt dernières années a permis au grand public de bénéficier d'outils de communication mobile. Il est désormais devenu possible et courant de téléphoner de partout (ou presque) dans des environnements aussi divers et variés que la rue, une gare ou bien encore une voiture. Cependant, tous ces lieux ne bénéficient pas du calme du salon ou le téléphone fixe était autrefois cantonné.

La gêne due à la perturbation qualifiée de bruit est généralement source d'inconfort et de fatigue pour les correspondants, il s'avère donc 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 bruité, de plus les équipements d'acquisition et de traitement des signaux peuvent générer un bruit supplémentaire lors de l'amplification, ce qui donne au bruit une présence inaltérable.

L'objectif des travaux effectués dans le cadre de ce mémoire est de fournir une solution efficace de réduction du bruit, ce qui nous a conduit à proposer un algorithme récursif non-quadratique bi-capteurs appliqué à la structure de séparation aveugle de source, et basé sur un détecteur automatique d'activité vocale en utilisant le RSB comme paramètre de détection.

Pour cela notre travail sera présenté comme suit :

Le premier chapitre définit quelques généralités sur la parole et le bruit acoustique en premier lieu, pour la deuxième partie nous illustrerons la formulation mathématique des deux algorithmes adaptatifs LMS et NLMS.

Le deuxième chapitre présente les méthodes de mélange convolutif linéaire mono-capteur, bi-capteurs et multi-capteurs, ainsi que la structure directe et récursive de la technique de séparation aveugle de source.

Le troisième chapitre expose notre nouvel algorithme RNQ automatique et son développement mathématique ainsi que son principe de fonctionnement.

Le quatrième chapitre et le dernier est réservé pour les résultats de simulation afin d'étudier les performances du nouvel algorithme proposé en utilisant les critères objectifs de qualité.

Enfin nous clôturons notre mémoire par une conclusion générale.

Chapitre 1 Généralités sur le filtrage et les algorithmes adaptatifs.

1.1 Introduction

Avant de parler sur le chapitre des méthodes de réduction de bruit acoustique, il est mieux, voir nécessaire d'éclairer quelques notions qui sont essentiellement connus pour tous ingénieurs électronicien. La présence d'un bruit superposé au signal utile dégrade la qualité et l'intelligibilité du signal de la parole dans de nombreuses applications, et ça peut donc poser beaucoup de problèmes.

Le but des algorithmes de débruitage de la parole est d'avoir une qualité d'information (la parole dans notre cas) améliorée en réduisant le bruit.

Dans ce chapitre, nous allons définir quelques notions de base sur la parole et le bruit pour une bonne compréhension du travail, avant de parler sur les différents algorithmes de base du filtrage adaptatifs.

1.2 Notions de base sur la parole et le bruit

1.2.1 Qu'est-ce qu'un signal ?

Un signal est une quantité (grandeur, fonction) physique mesurable qui évolue en fonction d'une ou plusieurs variables. Il correspond le plus souvent à une modélisation du comportement de la quantité physique observable. Cependant, il est peut-être extrêmement difficile d'obtenir une forme mathématique simple et concise pour un signal donné.

Les signaux les plus couramment utilisés sont les signaux électriques. Mais ces signaux sont les plus souvent des traductions de signaux physiques comme des signaux acoustiques, sismiques, de température ou de pression ... L'obtention des signaux électriques à partir des variations d'une grandeur naturelle se fait à l'aide d'un capteur ou d'un transducteur.

La mesure est utile en physique et comme toute mesure est entachée de bruit, d'incertitude voire d'erreur, le traitement du signal peut intervenir pour réduire l'effet de ces nuisances. C'est pourquoi, le traitement du signal peut se révéler utile dans presque tous les domaines. [10]

1.2.2 Signal de parole

La parole est un signal structuré qui permet aux êtres humains de communiquer entre eux. Elle se distingue des autres sons par ces caractéristiques acoustiques qui ont leurs origines dans les mécanismes de production, L'information d'un message parlé est transmise par les fluctuations de la pression de l'air qui sont émises par l'appareil phonatoire.

La parole est un signal réel, continu, de puissance moyenne finie et non stationnaire. Sa structure est complexe et variable avec le temps [1].

La parole est produite par le système articulatoire présenté dans la figure ci dessous :

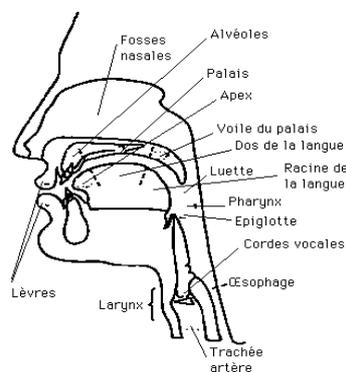


Fig. 1.1 - *Appareil phonatoire humain.*

La présence ou l'absence d'obstacle sur le parcours de la colonne d'air modifie la nature du son produit. C'est, entre autre en classant ces obstacles éventuels que la phonétique articulatoire dégage des différentes classes de son (consonnes, voyelles etc...) [2].

1.2.3 Bruit acoustique

Le bruit correspond à tout phénomène perturbateur gênant la transmission ou l'interprétation d'un signal. On peut également définir le bruit comme un son indésirable, qui rend difficilement perceptible l'information, il constitue donc une gêne dans la compréhension du signal utile qui est dans notre cas, la parole [1].

Le mécanisme de génération de bruit est toujours le même. Une excitation est produite par la source de bruit, il y a ensuite une phase de transmission durant laquelle le bruit ou les vibrations engendrées par la source se propagent respectivement dans l'air ou dans les solides. L'excitation de départ peut alors être amplifiée ou atténuée. La dernière phase du processus est l'émission du bruit vers le récepteur.

Entre la source et le récepteur, l'excitation peut être amenée à traverser successivement plusieurs milieux solides ou fluides. Aux interfaces entre ces milieux plusieurs phénomènes peuvent intervenir :

- Transmission
- Réflexion
- Absorption

Ces phénomènes complexifient l'identification des sources de bruit [3].

1.2.4 Sources de bruit

- **Bruits additifs**

Les bruits additifs ou les bruits externes sont causés par des agents agresseurs extérieurs au trinôme source-voie-destinataire [11].

Les origines de ses bruits sont généralement des perturbations naturelles comme le bruit atmosphérique et des perturbations artificielles.

- **Bruits multiplicatifs**

Les bruits multiplicatifs (convolutionnels) ou les bruits internes sont dus à la distorsion induite par la voie de communication (bruit de fond) [11].

1.2.5 Types de bruits

- **Bruit blanc**

Un bruit blanc est un signal aléatoire stationnaire dont la densité spectrale de puissance $S(f)$ est constante et la fonction d'autocorrélation $R(\tau)$ est l'impulsion de Dirac.

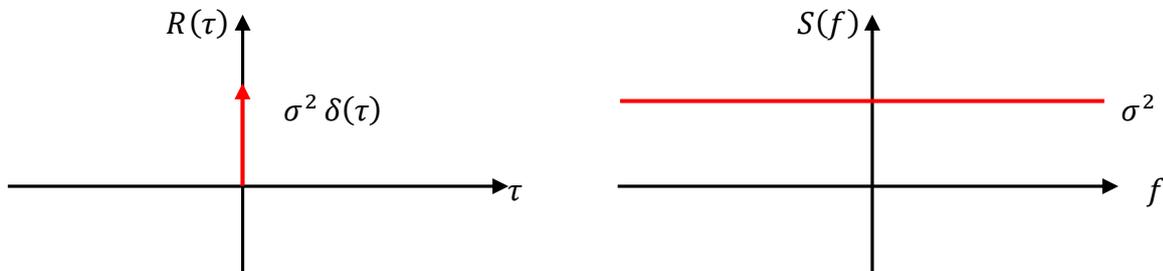


Fig. 1.2 - Autocorrélation et densité spectrale théoriques d'un bruit blanc.

- **Bruit coloré**

Un bruit coloré est un bruit blanc filtré, ainsi la fonction d'autocorrélation d'un tel bruit ne sera pas une impulsion de Dirac, mais plutôt une courbe étroite. Un bruit blanc de basse fréquence est parfois appelé bruit rose car il ne conserve que les grandes longueurs d'onde [5].

- **Bruit USASI** (*United States of American Standard Institut*)

C'est un signal aléatoire stationnaire dont les échantillons sont des variables aléatoires corrélés. Il est utilisé dans le domaine de rehaussement de la parole et la réduction du bruit pour évaluer la vitesse de convergence des algorithmes adaptatif [6].

1.2.6 Rapport signal à bruit (RSB)

Le RSB ou le SNR (en anglais : **S**ignal to **N**oise **R**atio) mesure la qualité du signal. C'est une caractéristique de la dégradation d'un signal, par définition informatif, par un bruit non informatif. C'est un moyen pour caractériser un système de transmission en comparant sa valeur à l'entrée du

système avec sa valeur en sortie de ce dernier, ou pour comparer la qualité des diverses méthodes de traitement des signaux [1].

Soient P_s et P_b la puissance du signal et la puissance du bruit respectivement, le RSB est donné par la relation :

$$RSB = \frac{P_s}{P_b} \quad (1.1)$$

En décibel :

$$RSB(dB) = 10 \log\left(\frac{P_s}{P_b}\right) \quad (1.2)$$

1.3 Filtrage de Wiener

Le filtre de WIENER est un filtre linéaire utilisé pour la réduction de bruit, il consiste à minimiser l'erreur quadratique moyenne entre le signal désiré et le signal de sortie, le filtre de Wiener est réalisable seulement dans le cas où le bruit et la parole sont stationnaires.

Le principe de ce filtre est de trouver en sortie une réponse plus proche d'une réponse désirée $d(n)$.

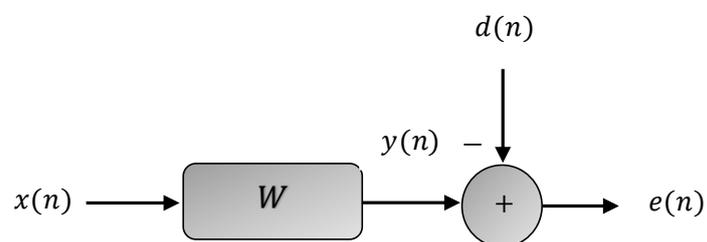


Fig. 1.3 - Représentation du filtrage de Wiener.

$x(n)$: Le signal d'entrée.

$w(n)$: Le filtre d'estimation.

$y(n)$: La sortie du filtre adaptatif.

$e(n)$: Le signal d'erreur.

$d(n)$: Le signal désiré (référence).

La sortie du filtre est donnée par :

$$y(n) = \sum_{k=0}^{L-1} w_k x(n-k), \quad n = 0, 1, 2, \dots \quad (1.3)$$

Nous pouvons écrire en notation vectorielle :

$$y(n) = \mathbf{w}(n)^T \mathbf{x}(n) = \mathbf{x}(n)^T \mathbf{w}(n) \quad (1.4)$$

Où \mathbf{w} désigne un vecteur de L coefficients avec :

$$\mathbf{w}(n) = [w_0, w_1, \dots, w_{L-1}]^T$$

1.3.1 Erreur quadratique moyenne

Dans les systèmes de communication numérique l'objectif est de minimiser l'erreur lors de la transmission d'information, afin de quantifier cette erreur on peut utiliser l'erreur quadratique moyenne, la fonction coût est la fonction de référence pour minimiser l'erreur avec les filtres optimaux.

$$J = E\{e^2(n)\} \quad (1.5)$$

J : La fonction coût

E : L'espérance mathématique

On note que le signal d'erreur est donné par :

$$e(n) = d(n) - y(n) = d(n) - \sum_{k=0}^{L-1} w_k x(n-k) \quad (1.6)$$

1.3.2 Principe d'orthogonalité

Le vecteur optimum \mathbf{w}_{opt} est celui qui annule le gradient :

$$\frac{\partial J}{\partial \mathbf{w}} = \mathbf{0}_{L \times 1} \quad (1.7)$$

$$\frac{\partial J}{\partial \mathbf{w}} = 2E \left\{ e(n) \frac{\partial e(n)}{\partial \mathbf{w}} \right\} = -2E \{ e(n) \mathbf{x}(n) \} \quad (1.8)$$

A l'optimum nous aurons :

$$E \{ e_{min}(n) \mathbf{x}(n) \} = \mathbf{0}_{L \times 1} \quad (1.9)$$

Où $e_{min}(n)$ est l'erreur minimale, donc l'optimum est atteint lorsque $e(n)$ est orthogonal à $\mathbf{x}(n)$.

Ainsi, dans le cas optimal, nous avons :

$$e_{min} = d(n) - [\mathbf{x}(n)^T \mathbf{w}_{opt}] \quad (1.10)$$

En développant l'équation (1.9) on obtient :

$$E[\mathbf{x}(n)d(n)] = E[\mathbf{x}(n)\mathbf{x}(n)^T] \mathbf{w}_{opt} \quad (1.11)$$

$$\mathbf{R}_{xd} = \mathbf{R}_{xx} \mathbf{w}_{opt} \quad (1.12)$$

Ou encore, l'équation de **Wiener-Hopf** :

$$\mathbf{w}_{opt} = \mathbf{R}_{xx}^{-1} \mathbf{R}_{xd} \quad (1.13)$$

Avec :

$$\text{La matrice d'autocorrélation : } \mathbf{R}_{xx} = E\{\mathbf{x}(n)\mathbf{x}^T(n)\} \quad (1.14)$$

$$\text{Le vecteur d'intercorrélation : } \mathbf{R}_{xd} = E\{\mathbf{x}(n)d(n)\} \quad (1.15)$$

1.4 Filtrage adaptatif

Le filtrage adaptatif est une technique utilisée pour l'ajustement automatique des paramètres d'un filtre lorsque ces derniers sont difficiles à déterminer ou variants dans le temps. Les coefficients sont ajustés selon l'algorithme d'adaptation en vue de minimiser l'erreur associée à l'algorithme du gradient, ce que les filtres classiques ne peuvent pas faire. Ce dernier est un filtre numérique utilisé de façon répandue dans la réduction du bruit et d'écho acoustique.

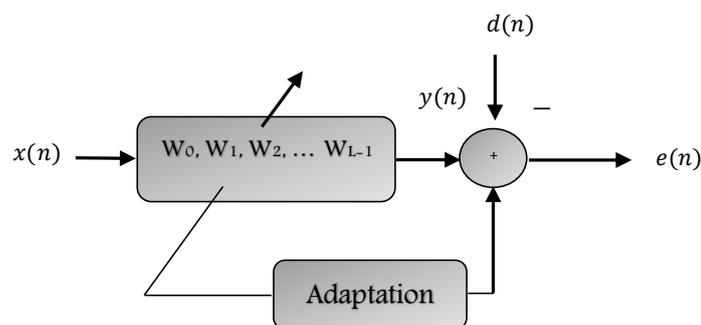


Fig. 1.4 – Principe du filtre adaptatif.

1.4.1 Présentation des algorithmes de filtrage adaptatif

Un algorithme adaptatif repose sur le même principe qu'un algorithme classique. Il se traduit par une suite finie de valeurs et d'instructions. Comme son nom l'indique, l'algorithme adaptatif s'adapte automatiquement et modifie son comportement au moment où il est exécuté, suivant le contexte.

1.4.2 Choix des algorithmes adaptatifs

Le choix des algorithmes adaptatifs se fait en fonction de :

- ❖ La structure.
- ❖ La complexité.

- ❖ La robustesse au bruit.
- ❖ Les propriétés numériques.
- ❖ La capacité de poursuite des variations.
- ❖ La rapidité de convergence qui est le nombre d'itération nécessaires pour converger vers la solution optimale.

1.4.3 Algorithme du gradient stochastique (Least Mean Squares – LMS)

L'algorithme LMS est l'un des algorithmes adaptatifs les plus populaires pour le calcul des coefficients d'un filtre adaptatif à cause de sa robustesse et sa simplicité de calcul. C'est un algorithme qui minimise l'erreur quadratique moyenne (EQM) entre un signal désiré et celui d'estimation, c'est-à-dire l'erreur.

L'équation de l'algorithme du gradient déterministe est donnée par :

$$\mathbf{w}(n + 1) = \mathbf{w}(n) - \frac{1}{2} \mu \mathbf{g}(n) \quad (1.16)$$

Où :

- $\mathbf{g}(n)$: est le gradient de la fonction coût.
- μ : est le pas d'adaptation.

Dans la résolution de l'équation de Wiener Hopf avec l'algorithme du gradient on rencontre un problème à savoir que les grandeurs \mathbf{R}_{xx} et \mathbf{R}_{xd} sont inconnues. C'est alors que l'algorithme du gradient stochastique LMS (Least Mean Squares) a été introduit par Widrow-Hoff [8]. L'LMS découle de l'algorithme du gradient qui consiste à remplacer les grandeurs \mathbf{R}_{xx} et \mathbf{R}_{xd} par une estimation à l'instant n .

On peut approcher les grandeurs des équations (1.14) et (1.15) par :

$$\widehat{\mathbf{R}}_{xx}(n) = \mathbf{x}(n)\mathbf{x}^T(n) \quad (1.17)$$

$$\widehat{\mathbf{R}}_{xd}(n) = \mathbf{x}(n)d(n) \quad (1.18)$$

En remplaçant $\widehat{\mathbf{R}}_{xx}(n)$ et $\widehat{\mathbf{R}}_{xd}(n)$ dans l'algorithme du gradient déterministe, on obtient l'équation de mise à jour de l'algorithme LMS donnée par :

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \mathbf{x}(n)e(n) \quad (1.19)$$

Avec μ est le pas d'adaptation tel que :

$$0 < \mu < \frac{1}{M\sigma_x^2}$$

Et :

M est la taille du filtre adaptatif.

σ_x^2 est l'énergie du signal d'entrée $x(n)$.

1.4.4 Algorithme LMS normalisé (Normalized LMS – NLMS)

Pour les signaux non stationnaires (énergie du signal $x(n)$ variante dans le temps), l'algorithme LMS aura du mal à fonctionner correctement car μ est constant. La création d'un algorithme LMS normalisé (NLMS) résout ce problème en remplaçant μ un pas μ_n défini à chaque itération :

$$\mu_n = \frac{\mu}{\mathbf{x}(n)^T \mathbf{x}(n)} \quad (1.20)$$

Donc l'équation de mise à jour des coefficients du filtre adaptatif par l'algorithme NLMS est donnée par :

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\mathbf{x}(n)^T \mathbf{x}(n) + \epsilon} \mathbf{x}(n)e(n) \quad (1.21)$$

Ou ϵ désigne une constante positive pour éviter la division par zéro.

L'étude de la stabilité de NLMS se fait en posant l'erreur a priori $e(n) = d(n) - \mathbf{w}(n)^T \mathbf{x}(n)$

soit supérieur à l'erreur a posteriori $\epsilon(n) = d(n) - \mathbf{w}(n+1)^T \mathbf{x}(n)$.

$$|\epsilon(n)| < |e(n)|$$

On remplace la formule de mise à jour de L'NLMS dans $\epsilon(n)$ et en supposant que $\epsilon = 0$ on trouve :

$$\epsilon(n) = d(n) - (\mathbf{w}(n) + \frac{\mu}{\mathbf{x}(n)^T \mathbf{x}(n)} \mathbf{x}(n) e(n))^T \mathbf{x}(n)$$

$$\epsilon(n) = d(n) - \mathbf{w}^T(n) \mathbf{x}(n) - \mu e(n)$$

$$\epsilon(n) = e(n)(1 - \mu)$$

Donc :

$$|\epsilon(n)| < |e(n)|$$

$$e(n)(1 - \mu) < |e(n)|$$

$$|1 - \mu| < 1$$

$$0 < \mu < 2$$

Qui est la condition de stabilité de NLMS [8].

1.5 Conclusion

Dans ce chapitre, nous avons exposé en premier lieu quelques notions de base sur la parole et le bruit acoustique, pour la deuxième partie nous avons présenté une étude détaillée sur les différents algorithmes de filtrage adaptatif. Cette partie théorique sera complétée dans le chapitre suivant par une étude approfondie sur les différentes techniques de réduction de bruit acoustique.

Chapitre 2 Techniques de réduction du bruit

2.1 Introduction

Dans un monde où la télécommunication joue un rôle très important dans notre vie quotidienne, les technologies doivent être les plus performantes possibles pour la satisfaction de tout le monde. Le rehaussement de la parole est nécessaire aujourd'hui pour plusieurs applications comme la téléphonie, les enregistrements audio, le domaine biomédical, etc...

De nombreuses techniques de débruitage sont donc proposées pour transmettre un signal de parole le moins bruité possible. Certaines techniques utilisent plusieurs microphones (multi-capteurs) ou juste deux microphones (bi-capteurs), il s'agit d'un réseau de capteurs, d'autres utilisent un seul microphone (mono-capteur, qui est la plus ancienne). Dans ce chapitre, nous allons voir les méthodes de débruitage de la parole mono-capteur, bi-capteurs, et multi-capteurs.

2.2 Réduction de bruit acoustique

Avant la présentation des méthodes de débruitage de la parole, il est nécessaire de définir quelques notions fondamentales.

2.2.1 Notion de canal acoustique

On peut considérer qu'un signal sonore est transmis de l'émetteur au récepteur par un canal acoustique. Le canal acoustique se comporte comme un système linéaire invariant dans le temps. La représentation du canal acoustique par sa réponse impulsionnelle dans le domaine numérique est comme suite :

$$y(n) = \sum_{k=0}^{+\infty} h_k x(n-k) \quad (2.1)$$

Où $x(n)$ et $y(n)$ représentent respectivement le signal d'entrée et le signal de sortie, h_k étant une fonction qui dépend de la géométrie de la salle, des matériaux des murs et des obstacles, de la position de la source et de son entendue, etc... [7].

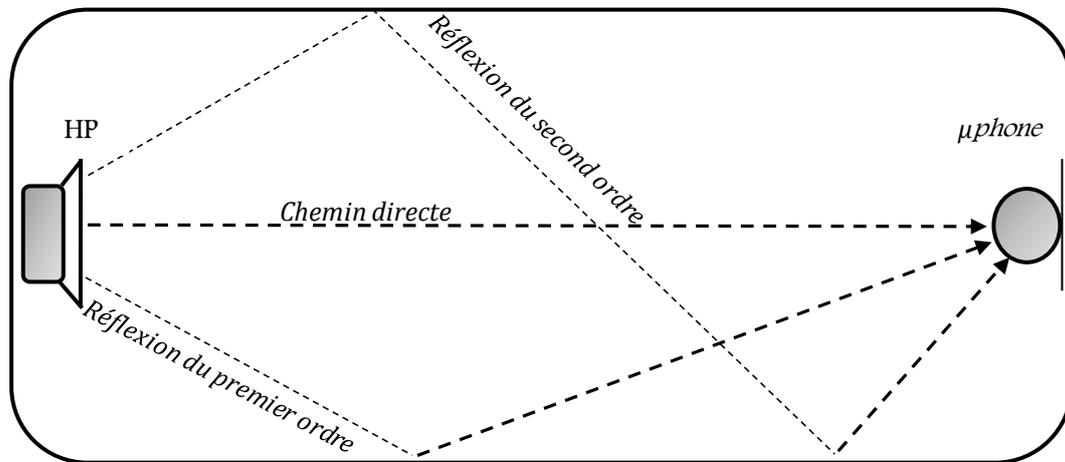


Fig. 2.1 - Propagation du son dans une salle acoustique.

2.2.2 Réponse impulsionnelle acoustique

Dans une salle, le premier son à être entendu est celui venant directement de la source. Puis, après un court instant, l'auditeur entendra les reflets atténués du son (échos) sur les murs, lorsque le son est réfléchi à plusieurs reprises par les murs, chaque réflexion sera encore retardée et atténuée. Un plus grand nombre d'inspections de la réponse impulsionnelle de la pièce a permis de constater que le son décroît de manière exponentielle [7].

La réponse impulsionnelle dispersive de la salle montrée ci-dessus peut être similaire à la figure suivante :

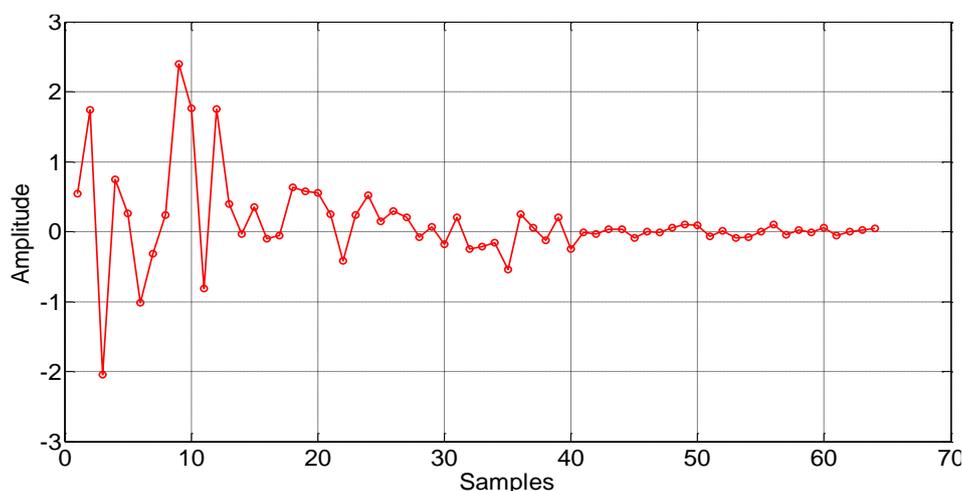


Fig. 2.2 - Réponse impulsionnelle dispersive ($L=64$).

2.3 Méthode mono-capteur

Dans cette technique, une seule observation est disponible pour séparer la parole du bruit par un algorithme adaptatif. Le signal arrivé au microphone est considéré comme une somme d'un signal utile et un bruit qu'on doit éliminer. La technique de réduction de bruit mono-capteur peut être schématisée dans la figure ci-dessous :

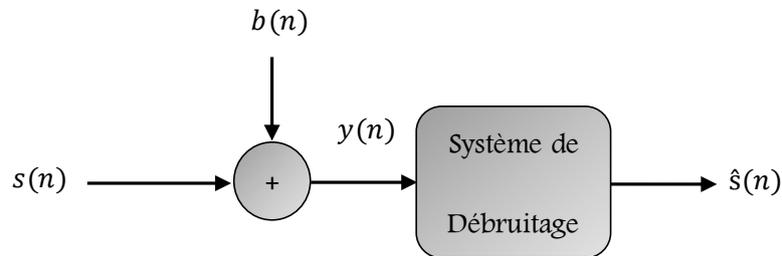


Fig. 2.3 - Principe de débruitage mono-capteur.

Où :

- $s(n)$ est le signal de la parole.
- $b(n)$ est le bruit.
- $y(n)$ est le signal de la parole bruité tel que : $y(n) = s(n) + b(n)$ (signal observé).
- $\hat{s}(n)$ est le signal de la parole après débruitage.

Les méthodes diffèrent en fonction de l'atténuation requise, de la méthode d'estimation du niveau du bruit et du niveau de parole sur chaque composante spectrale. Les techniques supposent que le bruit et le signal sont décorrélés sur la durée de la trame, que le signal de parole utile est intermittent, et que l'oreille humaine est insensible à la phase de signal [4]. Ces techniques se basent sur les transformées des signaux à temps discret comme :

- La transformée en cosinus discret (DCT).
- La transformée en ondelette.
- La transformée de Fourier discrète (DFT).

2.4 Méthode bi-capteurs

La méthode bi-capteurs nécessite l'utilisation de deux microphones pour annuler le bruit, contrairement à la méthode présentée avant qui nécessite un seul microphone. Il s'agit donc d'un mélange convolutif qui dépend de la propagation des signaux dans le milieu, de la position des microphones et des sources, et les caractéristiques de la salle [5]. Le mélange convolutif linéaire entre le signal de parole et le bruit peut être représenté par un modèle complet donné par la figure suivante :

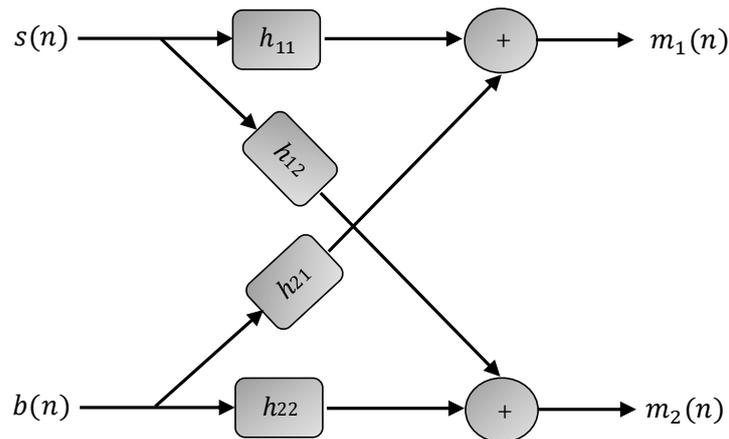


Fig. 2.4 - La structure complète d'un mélange convolutif bi-capteurs.

Les équations des signaux observés à la sortie du mélange sont données par :

$$m_1(n) = h_{11}(n) * s(n) + h_{21}(n) * b(n) \quad (2.2)$$

$$m_2(n) = h_{22}(n) * b(n) + h_{12}(n) * s(n) \quad (2.3)$$

En supposant que le bruit est ponctuel, et la distance entre la source de la parole et la source du bruit et les microphones est petite, on peut dire alors que $h_{11}(n) = h_{22}(n) = \delta(n)$.

2.4.1 Méthode bi-capteurs simplifiée

En supposant que $h_{11}(n) = h_{22}(n) = \delta(n)$, le schéma de la méthode bi-capteurs simplifiée sera comme suit :

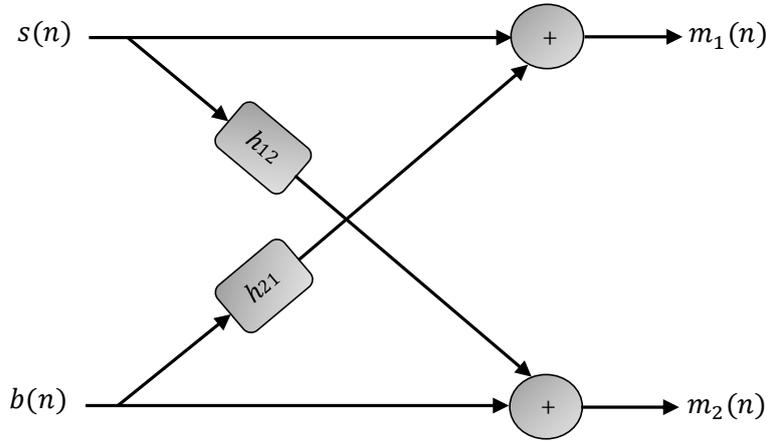


Fig. 2.5 - La structure simplifiée d'un mélange convolutif bi capteurs.

Les équations des signaux observés à la sortie du mélange sont données par :

$$m_1(n) = s(n) + \alpha_1 b(n) * h_{21}(n) \quad (2.4)$$

$$m_2(n) = b(n) + \alpha_2 s(n) * h_{12}(n) \quad (2.5)$$

Où $h_{12}(n)$ et $h_{21}(n)$ représentent les effets du couplage mutuel entre les signaux, $s(n)$ et $b(n)$ sont respectivement, deux sources ponctuelles de parole et de bruit.

α_1 : Coefficient pour contrôler l'interférence du bruit dans la voie 1.

α_2 : Coefficient pour contrôler l'interférence de la parole dans la voie 2.

Avec :

$$RSB_1 = 10 \log_{10} \left(\frac{E_s}{\alpha_1^2 E_b} \right) \quad \Rightarrow \quad \alpha_1 = \sqrt{\frac{E_s}{E_b} 10^{-RSB_1/10}} \quad (2.6)$$

$$RSB_2 = 10 \log_{10} \left(\frac{\alpha_2^2 E_s}{E_b} \right) \quad \Rightarrow \quad \alpha_2 = \sqrt{\frac{E_b}{E_s} 10^{+RSB_2/10}} \quad (2.7)$$

2.5 Méthode multi-capteurs

C'est le cas général car le capteur reçoit plusieurs signaux venus des sources différentes, parole ou bruit. Dans ce cas, φ signaux de sources $s(n) = [s_1(n), s_2(n) \dots \dots s_\varphi(n)]$ subissent une convolution avec φ canaux de $h_{ip}(k)$ (réponses impulsionnelles). Les relations entre les signaux de source et les observations sont données par la relation suivante :

$$m_p(n) = \sum_{i=1}^{\varphi} \sum_{k=0}^{M-1} h_{ip}(k) s_i(n-k) \quad , \quad p = 1, 2 \dots, \Psi \quad (2.8)$$

Où Ψ , φ et M représentent respectivement, le nombre de mélanges, le nombre de sources et la longueur de réponses d'impulsionnelles.

Voici ci-dessous le schéma général du mélange :

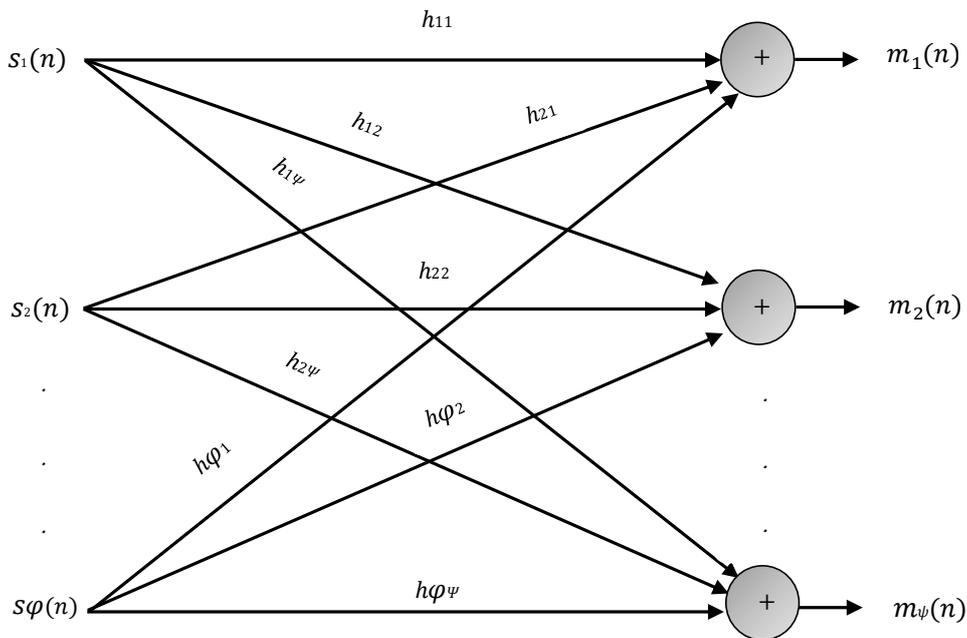


Fig. 2.6 - La structure d'un mélange convolutif multi capteurs (cas général).

2.6 Réduction de bruit acoustique par la séparation de source aveugle

La séparation aveugle de source ou BSS en anglais (**B**lind **S**ource **S**eparation) a été toujours un sujet de recherche actif depuis plusieurs années, parce qu'elle est utilisée dans de nombreuses application comme l'acoustique, le domaine biomédical, etc...

L'idée derrière la SAS c'est qu'on peut, sans la connaissance d'aucun paramètre d'un mélange donné, extraire, ou récupérer toutes les sources ou au moins isoler des sources particulières.

L'objectif est donc, de concevoir un système qui est capable de récupérer des signaux de sortie y_i égaux aux signaux source $S\varphi$ à partir de Ψ signaux de mélanges.

Il existe deux structures de la SAS, la première est appelée directe ou FBSS (**F**orward **B**lind **S**peech **S**eparation) et la deuxième est appelée récursive ou bien BBSS (**B**ackward **B**lind **S**peech **S**eparation).

2.6.1 Caractéristiques des signaux source

Les signaux – en général – peuvent être classés en différentes modes de classifications, c'est pour faciliter l'étude. Certains signaux sont stationnaires, d'autres sont non-stationnaires, certains sont à énergie finie, d'autres sont à puissance moyenne finie. Il est nécessaire de connaître les caractéristiques du signal pour le développement mathématique des algorithmes de la SAS [10].

2.6.2 Structure directe (FBSS)

C'est la structure la plus utilisée, chaque sortie du système est une combinaison linéaire des observations [4], elle est schématisée par la figure ci-dessous :

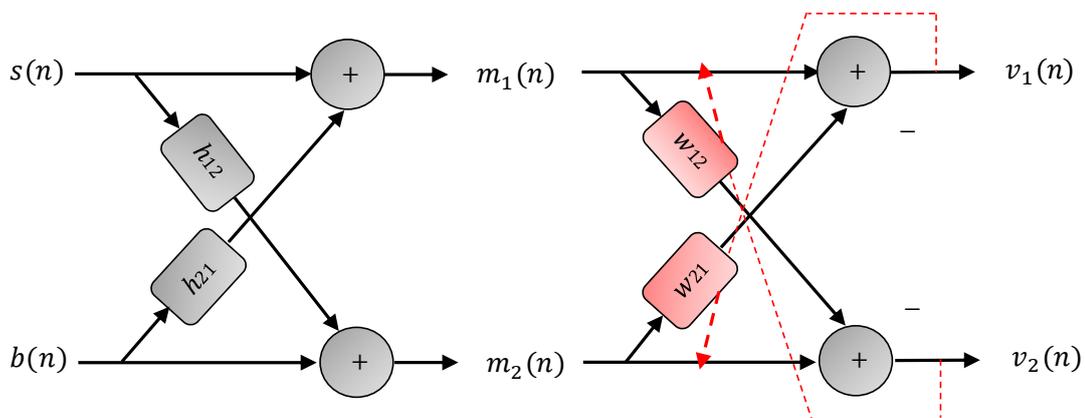


Fig. 2.7 – Structure directe de la SAS.

Avec w_{12} et w_{21} sont les deux filtres adaptatifs, h_{12} et h_{21} sont les réponses impulsionnelles du système.

Les équations d'entrée de la structure directe sont :

$$m_1(n) = s(n) + b(n) * h_{21}(n) \quad (2.9)$$

$$m_2(n) = b(n) + s(n) * h_{12}(n) \quad (2.10)$$

Et les sorties $v_1(n)$ et $v_2(n)$ sont données par les équations (2.11) et (2.12) :

$$v_1(n) = m_1(n) - m_2(n) * w_{21}(n) \quad (2.11)$$

$$v_2(n) = m_2(n) - m_1(n) * w_{12}(n) \quad (2.12)$$

En remplaçant les expressions de $m_1(n)$ et $m_2(n)$ dans les expressions de $v_1(n)$ et $v_2(n)$, nous trouvons :

$$v_1(n) = b(n) * [h_{21}(n) - w_{21}(n)] + s(n) * [\delta(n) - h_{12}(n) * w_{21}(n)] \quad (2.13)$$

$$v_2(n) = s(n) * [h_{12}(n) - w_{12}(n)] + b(n) * [\delta(n) - h_{21}(n) * w_{12}(n)] \quad (2.14)$$

En utilisant l'hypothèse d'optimalité pour les deux filtres adaptatifs :

$$w_{21}^{opt} = h_{21} \text{ et } w_{12}^{opt} = h_{12} \quad (2.16)$$

On obtient les sorties $v_1(n)$ et $v_2(n)$ selon les deux expressions suivantes :

$$v_1(n) = s(n) * [\delta(n) - h_{12}(n) * w_{21}(n)] \quad (2.17)$$

$$v_2(n) = b(n) * [\delta(n) - h_{21}(n) * w_{12}(n)] \quad (2.18)$$

Nous notons que les solutions optimales des filtres adaptatifs sont obtenues en utilisant les critères du minimum de l'erreur quadratique moyenne (MMSE ou bien EQMM).

On remarque aussi que les signaux $v_1(n)$ et $v_2(n)$ convergent vers les signaux initiaux de source mais avec une petite distorsion, c'est l'inconvénient majeur de cette structure. Pour pallier à ce problème il a été proposé une solution qui consiste à mettre des post-filtres $PF_1(n)$ et $PF_2(n)$ à la sortie de la structure après $v_1(n)$ et $v_2(n)$ respectivement [4]. Ils sont donnés par :

$$PF_1(n) = \frac{1}{\delta(n) - h_{12}(n) * w_{21}(n)} \quad \text{et} \quad PF_2(n) = \frac{1}{\delta(n) - h_{21}(n) * w_{12}(n)} \quad (2.19)$$

A la fin, les deux signaux de sortie des post-filtres convergent respectivement vers $s(n)$ et $b(n)$ sans aucune distorsion.

2.6.3 Structure réursive (BBSS)

C'est la deuxième structure utilisée dans la SAS. Cette structure est une solution très efficace car elle élimine le besoin d'utiliser des post-filtres cependant elle mène à obtenir de bon résultat du coté rapidité de convergence, cout en mémoire et complexité algorithmique [4].

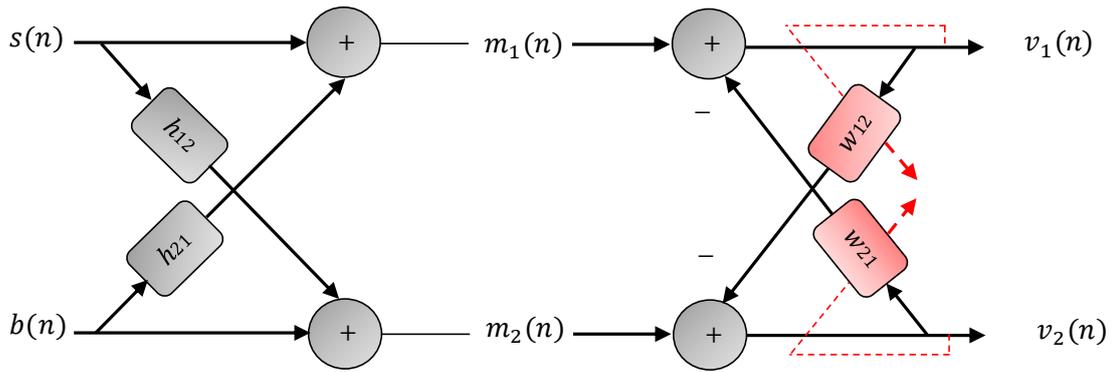


Fig. 2.8 – Structure réursive de la SAS.

La structure réursive résulte les relations suivantes :

$$v_1(n) = m_1(n) - v_2(n) * w_{21}(n) \quad (2.20)$$

$$v_2(n) = m_2(n) - v_1(n) * w_{12}(n) \quad (2.21)$$

En remplaçant les expressions $m_1(n)$ et $m_2(n)$ dans les expressions de $v_1(n)$ et $v_2(n)$, nous trouvons :

$$v_1(n) = [s(n) * (\delta(n) - h_{12}(n) * w_{21}(n)) + b(n) * (h_{21}(n) * w_{21}(n))] * (\delta(n) - w_{12}(n) * w_{21}(n))^{-1} \quad (2.22)$$

$$v_2(n) = [b(n) * (\delta(n) - h_{21}(n) * w_{12}(n)) + s(n) * (h_{12}(n) * w_{12}(n))] * (\delta(n) - w_{21}(n) * w_{12}(n))^{-1} \quad (2.23)$$

En utilisant l'hypothèse d'optimalité pour les deux filtres adaptatifs :

$$w_{21}^{opt} = h_{21} \text{ et } w_{12}^{opt} = h_{12}$$

On obtient à la sortie :

$$v_1(n) = s(n) \tag{2.24}$$

$$v_2(n) = b(n) \tag{2.25}$$

On remarque le signal de parole estimé $v_1(n)$ égale exactement au signal originale $s(n)$.

2.7 Conclusion

Dans cette partie, nous avons abordé les principales méthodes et techniques de réduction du bruit avec acquisition mono-capteur, bi-capteurs et multi-capteurs suivis par les deux structures de séparation aveugle de source directe et récursive (Forward et Backward). Les objectifs de ce chapitre sont de définir les notions que nous utiliserons dans notre travail.

Chapitre 3 Proposition d'un nouvel algorithme

Recuratif Non-Quadratic automatique

3.1 Introduction

La réduction de bruit comprend l'ensemble des techniques qui réduisent le bruit acoustique contenu dans un signal. Le bruit étant un paramètre indésirable, doit être éliminé sans dégrader la qualité du signal utile. Le but est donc, de concevoir un système qui est capable de restituer un signal identique au signal d'entrée.

Dans ce chapitre, nous allons présenter la structure de réduction de bruit basée sur l'algorithme NLMS, ainsi que la nouvelle structure qui est basée sur l'algorithme RNQ et sur un détecteur d'activité vocal automatique utilisant le RSB comme paramètre de décision.

3.2 Réduction de bruit par l'algorithme NLMS

Dans cette partie, nous présentons le bref principe de la réduction de bruit acoustique par l'algorithme NLMS.

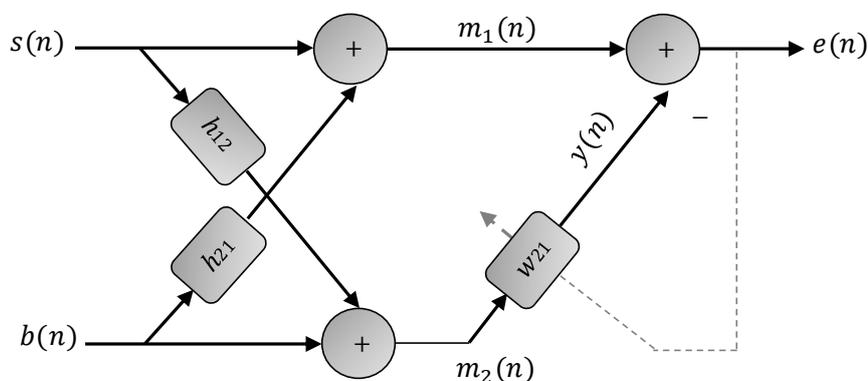


Fig. 3.1 – Structure d'annulation adaptative du bruit par l'algorithme NLMS.

Où :

$s(n)$: Signal de la parole.

$b(n)$: Signal du bruit.

$m_1(n), m_2(n)$: Signaux de mélanges.

$h_{12}(n), h_{21}(n)$: Les réponses impulsionnelles du système.

$y(n)$: La sortie du filtre.

$w(n)$: Le filtre adaptatif.

$e(n)$: Le signal d'erreur.

La sortie $y(n)$ du filtre $w(n)$ est utilisée pour calculer l'estimation du signal utile $e(n)$ par la relation suivante :

$$y(n) = \mathbf{w}(n)^T \mathbf{m}_2(n) \quad (3.1)$$

Le signal d'erreur est donné par :

$$e(n) = m_1(n) - \mathbf{w}(n)^T \mathbf{m}_2(n) \quad (3.2)$$

L'équation de la m-à-j des coefficients du filtre est donnée par :

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu \frac{\mathbf{m}_2(n)e(n)}{\mathbf{m}_2(n)^T \mathbf{m}_2(n) + \varepsilon} \quad (3.3)$$

Ou $0 < \mu < 2$ et $\varepsilon \ll \mathbf{m}_2(n)^T \mathbf{m}_2(n)$.

3.3 Réduction de bruit par un nouvel algorithme RNQ automatique

Dans cette partie, nous présentons le principe de la réduction de bruit acoustique par l'algorithme RNQ qui est décrit dans la figure 3.2 :

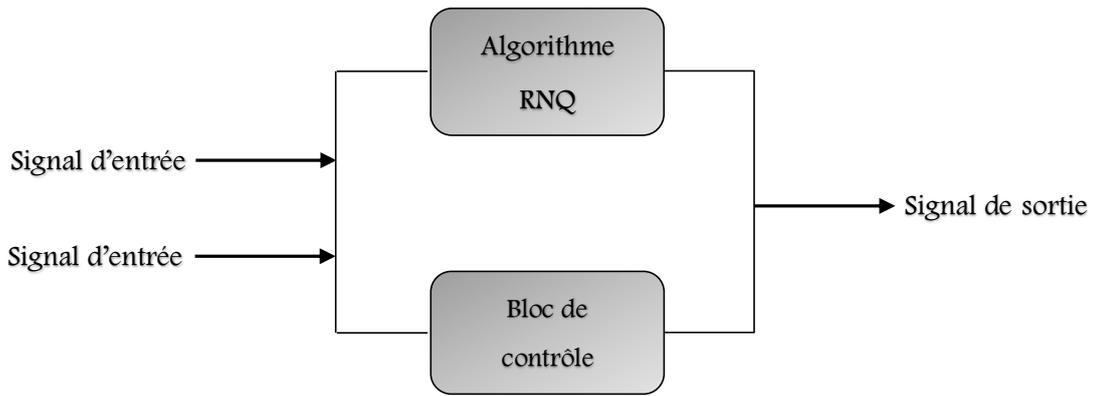


Fig. 3.2 – Schéma bloc simplifié de la nouvelle structure de réduction de bruit basée sur l'algorithme RNQ et un détecteur d'activité vocale automatique.

Le bloc de contrôle possède un mécanisme qui à l'aide de certains calculs et comparaisons, peut décider que sera la sortie (activité vocale ou silence) afin d'automatiser le système. Pour que notre système soit clair, il est mieux d'illustrer le détail de chaque bloc :

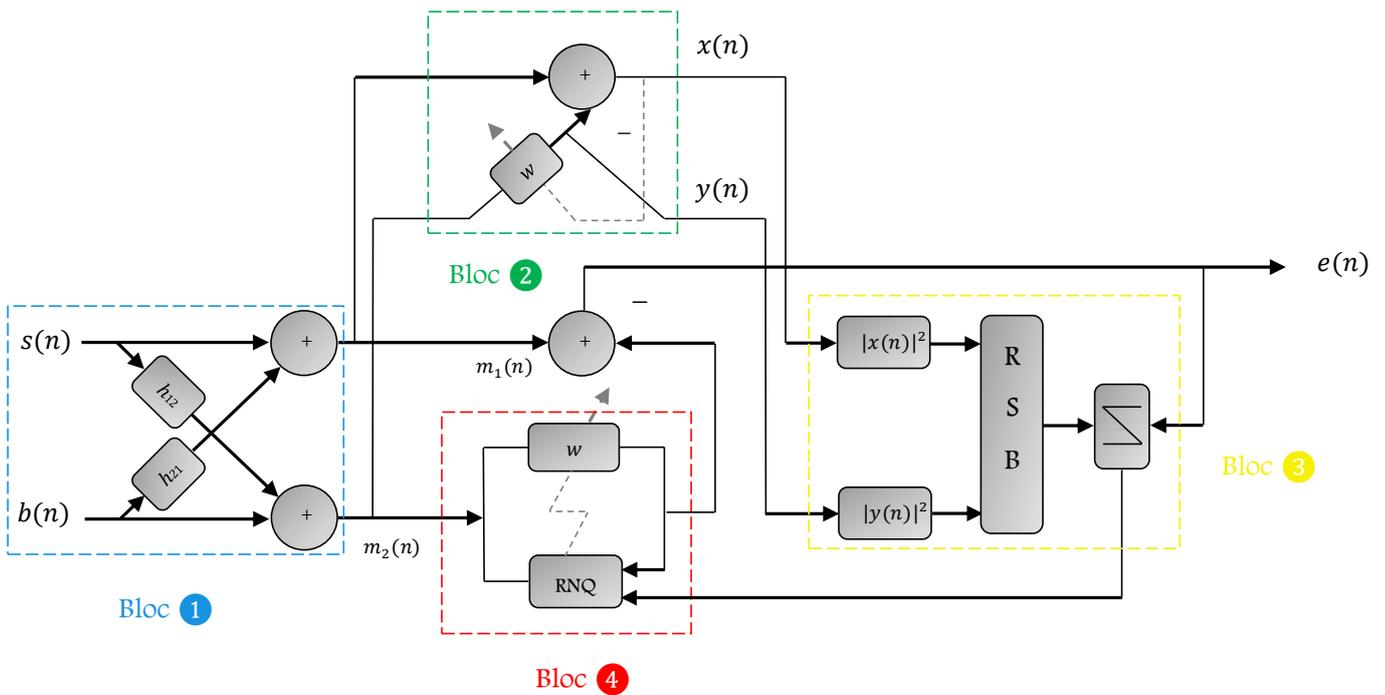


Fig. 3.3 – Schéma détaillé de la nouvelle structure de réduction de bruit basée sur l'algorithme RNQ et un détecteur d'activité vocale automatique.

- Bloc ①

Le premier bloc représente le mélange convolutif bi-capteurs dont les signaux de sortie sont donnés par les équations suivantes :

$$m_1(n) = s(n) + b_1(n) \quad (3.4)$$

$$m_2(n) = b(n) + s_1(n) \quad (3.5)$$

Avec :

$$b_1(n) = \alpha_1 b(n) * h_{21}(n) \quad (3.6)$$

$$s_1(n) = \alpha_2 s(n) * h_{12}(n) \quad (3.7)$$

Donc :

$$m_1(n) = s(n) + \alpha_1 b(n) * h_{21}(n) \quad (3.8)$$

$$m_2(n) = b(n) + \alpha_2 s(n) * h_{12}(n) \quad (3.9)$$

Où $h_{12}(n)$ et $h_{21}(n)$ représentent les effets du couplage mutuel entre les signaux, $s(n)$ et $b(n)$ sont respectivement, deux sources ponctuelles de parole et de bruit.

Les réponses impulsionnelles des canaux directes h_{11} et h_{22} étant négligées, c.-à-d. : $h_{11} = h_{22} = \delta(n)$.

Cette supposition n'a pas d'impact sur l'utilité pratique de ce modèle, nous pouvons aussi supposer que l'orateur n'est pas très loin des microphones et nous n'avons pas *a priori* d'informations sur la source de bruit [12].

α_1 et α_2 sont deux facteurs de régularisation de la voie 1 et la voie 2 respectivement, tel que:

$$RSB_1 = 10 \log_{10} \left(\frac{E_s}{\alpha_1^2 E_b} \right) \quad \Rightarrow \quad \alpha_1 = \sqrt{\frac{E_s}{E_b} 10^{-RSB_1/10}} \quad (3.10)$$

$$RSB_2 = 10 \log_{10} \left(\frac{\alpha_2^2 E_s}{E_b} \right) \quad \Rightarrow \quad \alpha_2 = \sqrt{\frac{E_b}{E_s} 10^{+RSB_2/10}} \quad (3.11)$$

Et:

E_s : L'énergie du signal de la parole $s(n)$.

E_b : L'énergie du signal du bruit $b(n)$.

- Bloc ②

Le deuxième bloc c'est l'algorithme NLMS classique de longueur L . Les deux signaux de sortie de ce filtre $x(n)$ et $y(n)$ sont donnés par :

$$x(n) = m_1(n) - y(n) \quad (3.12)$$

Où :

$$y(n) = \mathbf{w}(n)^T \mathbf{m}_2(n) \quad (3.13)$$

On remplace (3.12) dans (3.13), on obtient :

$$x(n) = m_1(n) - \mathbf{w}(n)^T \mathbf{m}_2(n) \quad (3.14)$$

Ce filtre met à jour ces coefficients par :

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu \frac{\mathbf{m}_2(n)x(n)}{\mathbf{m}_2(n)^T \mathbf{m}_2(n) + \varepsilon} \quad (3.15)$$

$\mathbf{m}_2(n) = [m_2(n), \dots, m_2(n - (L - 1))]^T$ est le vecteur d'entrée de L échantillons, μ est le pas d'adaptation qui doit être compris entre 0 et 2 comme condition de convergence de l'algorithme.

Les deux composantes de sortie $x(n)$ et $y(n)$ qui sont proportionnelles aux signaux de la parole et de bruit, sont utilisées par la suite dans le 3^{eme} bloc pour estimer le *RSB*.

- Bloc ③

Ce bloc est utilisé pour l'estimation d'énergie et le calcul du *RSB*, et la prise de décision par la suite (activité vocale ou silence).

Le RSB est donné par la relation suivante :

$$RSB|_{dB} = 10 \log_{10} \left(\frac{E_x}{E_y} \right) \quad (3.16)$$

E_x et E_y sont les valeurs d'énergie des signaux de sortie du 2^{ieme} bloc, il sont calculées comme suit:

$$E_x = \sum_{i=0}^{L-1} |x(i)|^2 \quad (3.17)$$

$$E_y = \sum_{i=0}^{L-1} |y(i)|^2 \quad (3.18)$$

Après l'estimation du RSB , un RSB_{seuil} est fixé. Pour des valeurs supérieures on a une activité vocale, et pour des valeurs inférieures on a le silence, c.-à-d. :

$$\left\{ \begin{array}{l} RSB \leq RSB_{seuil} \rightarrow \theta = 1 \\ RSB > RSB_{seuil} \rightarrow \theta = 0 \end{array} \right. \quad (3.19)$$

Donc la valeur du paramètre du circuit de décision θ a un effet sur l'erreur du filtrage de l'algorithme RNQ, pour donner une nouvelle valeur qui est donnée par :

$$\hat{e}(n) = \theta e(n) \quad (3.20)$$

Cette nouvelle valeur de l'erreur du filtrage est utilisée par la suite par l'algorithme RNQ dans le 4^{ieme} bloc.

- Bloc ④

Le dernier bloc est dédié à l'algorithme que nous avons automatisé précédemment, RNQ.

L'algorithme **Recursive Non-Quadratic** ou RNQ, est un type d'algorithme adaptatif basé sur des fonctions non quadratiques de l'erreur, et qui converge plus rapidement que les algorithmes cités précédemment (LMS, NLMS) [13]. La figure ci-dessous représente le principe de l'algorithme RNQ :

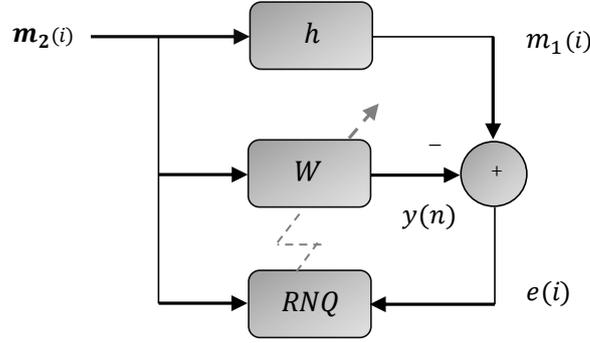


Fig. 3.4 – Schéma du principe de fonctionnement de l'algorithme RNQ.

Où $d(n)$ est le signal désiré, $x(n)$ est le signal d'entrée défini par :

$$\mathbf{x}(i) = [x(i), \dots, x(i - (L - 1))]^T \quad (3.21)$$

Et une erreur $e(i)$, qui est utilisé pour mettre à jour le vecteur $w(n)$:

$$\mathbf{w}(n) = [w_0(n), w_1(n), \dots, w_{L-1}(n)]^T \quad (3.22)$$

L'objectif est de récupérer $d(n)$ par estimation du signal de sortie:

$$y(n) = \mathbf{w}(n)^T \mathbf{m}_2(n) \quad (3.23)$$

Après avoir calculé l'erreur :

$$e(i) = m_1(n) - y(n) \quad (3.24)$$

Où n représente le nombre d'itérations et $1 \leq i \leq n$ et L c'est la longueur du filtre.

Dans cet algorithme, nous allons utiliser la fonction cout suivante :

$$J(n) = \sum_{i=1}^n \{\lambda^{n-i} [e(i)]^{2j}\} \quad (3.25)$$

Où j et n sont des nombres entiers positifs. Le facteur de coefficient exponentiel, ou le facteur d'oubli, λ est une constante proche, mais plus petite que 1 [14].

Afin d'obtenir le coefficient optimal, nous calculons le gradient de $J(n)$, prenons l'expansion binomiale de $[m_1(i) - \mathbf{w}(i)^T \mathbf{m}_2(i)]^{2j-1}$ et faisons l'approximation suivante :

$$\begin{aligned}
& \nabla J(n) \\
&= -2j \left[\sum_{i=1}^n [\lambda^{n-i} m_1(i)^{2j-1} \mathbf{m}_2(i)] - (2 \right. \\
&\quad \left. - 1) \sum_{i=1}^n [\lambda^{n-i} m_1(i)^{2j-2} \mathbf{m}_2(i) \mathbf{m}_2(i)^T] \mathbf{w}(n) \right] \tag{3.26}
\end{aligned}$$

On suppose que $e(i)$ est approximativement égal à zéro lorsque les poids atteignent ses valeurs minimales.

Pour simplifier l'équation (3.26) on pose :

$$Z(n) = \sum_{i=1}^n [\lambda^{n-i} m_1(i)^{2j-1} \mathbf{m}_2(i)] \tag{3.27}$$

Et :

$$\Phi(n) = (2j - 1) \sum_{i=1}^n [\lambda^{n-i} m_1(i)^{2j-2} \mathbf{m}_2(i) \mathbf{m}_2(i)^T] \tag{3.28}$$

L'équation (3.26) sera donc :

$$\nabla J(n) \approx -2j[Z(n) - \Phi(n)\mathbf{w}(n)] \tag{3.29}$$

Maintenant, en égalisant à zéro l'équation (3.29), on obtient l'équation suivante :

$$\mathbf{w}(n) = Z(n)\Phi(n)^{-1} \tag{3.30}$$

En isolant le terme correspondant à $i = n$ de $\Phi(n)$, on obtient :

$$\Phi(n) = \lambda\Phi(n-1) + [(2j-1)m_1(n)^{2j-2}] \mathbf{m}_2(n) \mathbf{m}_2(n)^T \tag{3.31}$$

Appliquant le lemme de la matrice inverse, nous obtenons le vecteur de gain $L * 1$ donné par :

$$\mathbf{G}(n) = \frac{\lambda^{-1} \mathbf{P}(n-1) \mathbf{m}_2(n)}{[(2j-1)m_1(n)^{2j-2}]^{-1} + \lambda^{-1} \mathbf{m}_2(n)^T \mathbf{P}(n-1) \mathbf{m}_2(n)} \tag{3.32}$$

Où :

$$\mathbf{P}(n) = \Phi(n)^{-1} = \lambda^{-1} \mathbf{P}(n-1) - \lambda^{-1} \mathbf{G}(n) \mathbf{m}_2(n)^T \mathbf{P}(n-1) \tag{3.33}$$

Pour faciliter la prochaine manipulation mathématique, nous réorganisons (3.32) comme suit :

$$[(2j - 1)m_1(n)^{2j-2}]^{-1}G(n) = P(n)m_2(n) \quad (3.34)$$

En remplaçant l'expression de $P(n)$ dans l'équation (3.30), on obtient :

$$\mathbf{w}(n) = \mathbf{w}(n - 1) + G(n) \left[\frac{m_1(n)}{(2j-1)} - \mathbf{m}_2(n)^T \mathbf{w}(n - 1) \right] \quad (3.35)$$

Le terme entre crochets dans (3.35) est l'estimation *a priori* de l'erreur, que nous désignons par $\epsilon(n)$.

Enfin, l'équation finale de la mise à jour de notre algorithme proposé RNQ est donnée par :

$$\mathbf{w}(n) = \mathbf{w}(n - 1) + G(n)\epsilon(n) \quad (3.36)$$

3.4 Conclusion

Dans ce chapitre nous avons présenté le principe du fonctionnement de l'algorithme que l'on a automatisé, c'est à dire qu'on n'aura pas besoin de détecter manuellement l'activité vocale, cela lui rend pratique et très performant.

Par la suite, dans le chapitre suivant nous allons présenter nos résultats de simulations, et évaluer les performances de l'algorithme proposé en utilisant certains critères.

4.1 Introduction

Dans ce chapitre, nous allons présenter les résultats de simulations de la réduction de bruit obtenus par le nouvel algorithme RNQ automatique.

Des critères sont utilisés au cours de cette partie comme le System Mismatch, Segmental SNR, et le Segmental MSE, cela a pour objectif de comparer cet algorithme avec d'autres comme le NLMS classique.

4.2 Signaux utilisés aux simulations

4.2.1 Signal de la parole

Dans ce travail, on a introduit une séquence de 4 secondes (dupliquée), échantillonnée à une fréquence d'échantillonnage $f_e = 8000 \text{ Hz}$ et prononcée par un locuteur masculin dont le scénario est le suivant : « Un loup s'est jeté immédiatement sur la petite chèvre ». La figure ci-dessous présente le signal de la parole utilisé et son spectrogramme.

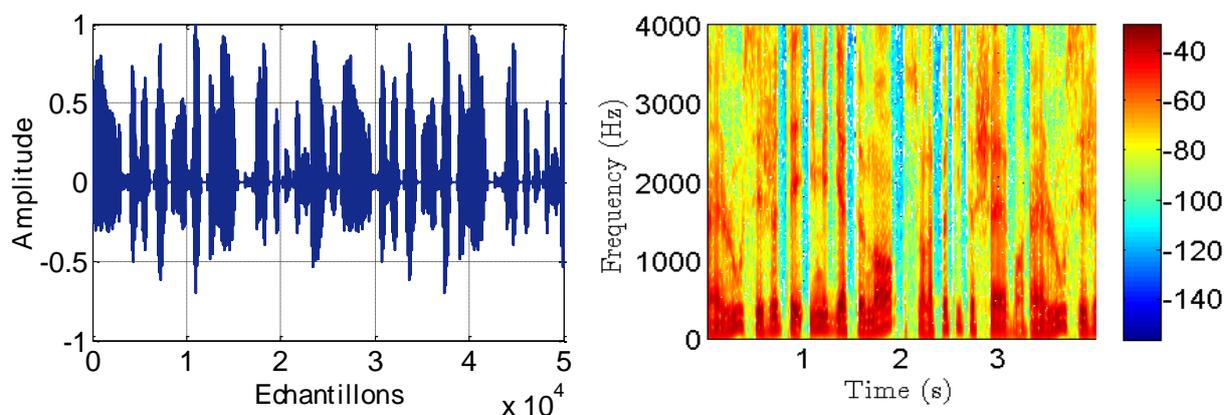


Fig. 4.1 - Signal de la parole utilisé et son spectrogramme.

4.2.2 Bruit blanc

Le bruit blanc est une réalisation d'un processus aléatoire dont lequel la densité spectrale est la même (constante) pour toutes les fréquences et sert surtout à vérifier la stabilité de l'algorithme utilisé [5]. La figure ci-dessous représente le signal du bruit blanc utilisé et son spectrogramme.

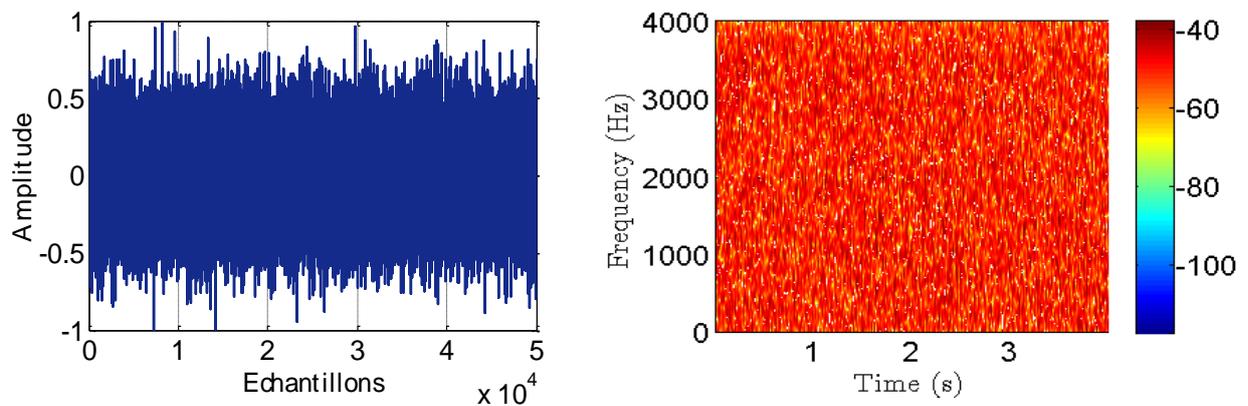


Fig. 4.2 - Signal du bruit blanc et son spectrogramme.

4.2.3 Réponses impulsionnelles

La figure ci-dessous montre la modélisation des réponses impulsionnelles des filtres h_{12} et h_{21} . Ces réponses impulsionnelles sont générés à partir du bruit gaussien (GN) modulé par une fonction exponentielle à un support temporel d'environ $L = 64$ à la fréquence d'échantillonnage $f_e = 8 \text{ kHz}$ [15], avec ces deux modèles on peut mieux représenter les conditions réelles de propagation des signaux à travers une salle quelconque et montrer l'effet de la distance.

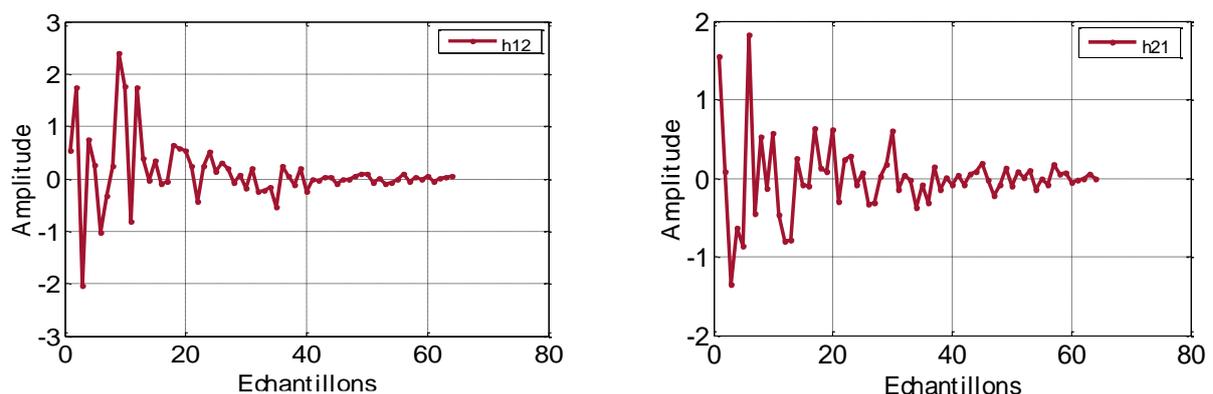


Fig. 4.3 – La réponse impulsionnelle h_{21} (à droite) et la réponse impulsionnelle h_{12} (à gauche) avec $L = 64$.

4.2.4 Signaux de mélange

Les signaux présentés ici sont les signaux de sortie de la structure du mélange convolutif $m_1(n)$ et $m_2(n)$ et leurs spectrogrammes, sachant que les signaux d'entrée sont le signal de la parole et le signal du bruit avec un rapport signal à bruit égale à 6 dB sur les deux voies et une longueur du filtre $L = 64$.

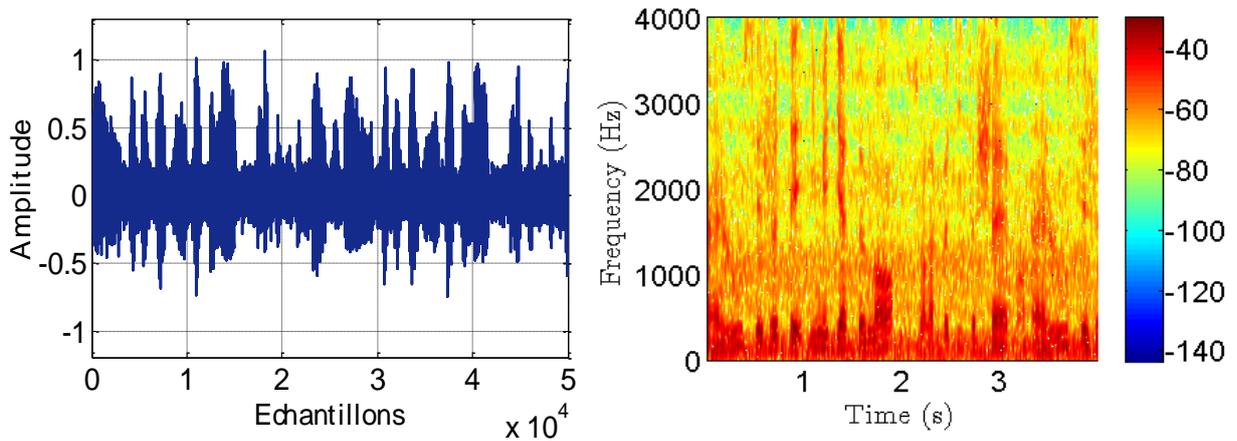


Fig. 4.4 – Signal du premier mélange $m_1(n)$ et son spectrogramme.

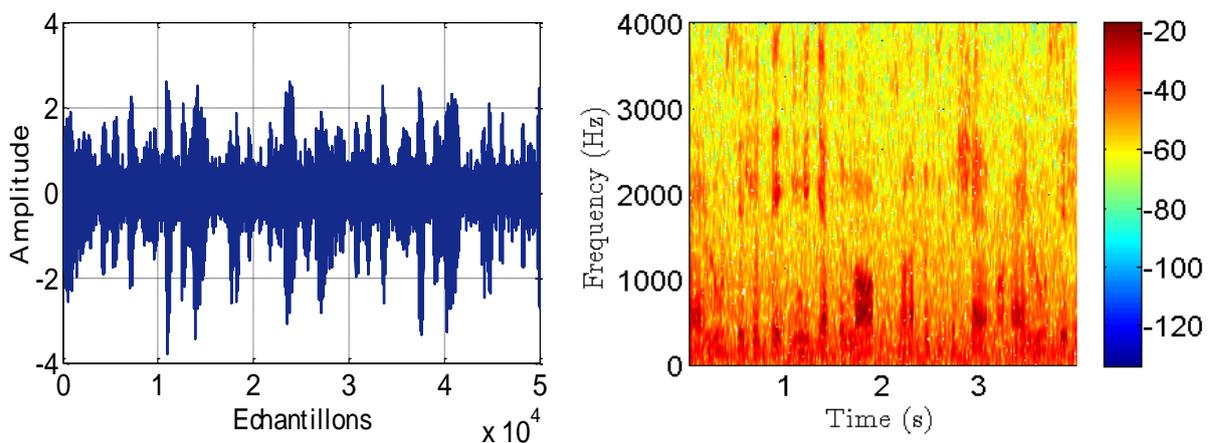


Fig. 4.5 – Signal du deuxième mélange $m_2(n)$ et son spectrogramme.

4.3 Détection d'activité vocale (DAV)

La fonction de de la détection d'activité vocale est utilisée pour la distinction entre l'activité vocale et le silence. C'est un mécanisme pour contrôler l'adaptation des filtres w_{12} et w_{21} , le filtre w_{21} est adapté uniquement à la période du bruit seul, alors que le filtre w_{12} s'adapte pendant la présence du signal de la parole utile. La figure 4.8 illustre un exemple de système DAV manuelle.

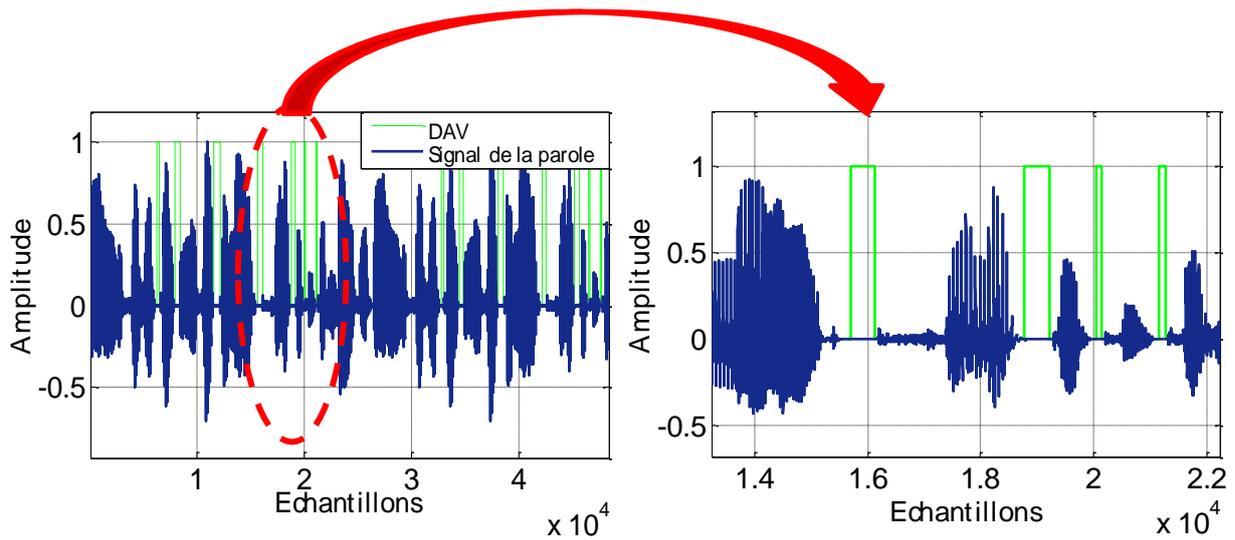


Fig. 4.6 – Signal de la parole avec sa DAV.

4.4 Evaluation du nouvel algorithme proposé

Dans cette partie, nous allons évaluer les performances du nouvel algorithme proposé RNQ automatique par rapports aux différentes tailles du filtre, différents niveaux du RSB d'entrée, différentes valeurs de j , et en utilisant trois (3) critères objectifs de qualité qui sont :

- L'erreur quadratique moyenne segmentée (Segmental-MSE) qui indique le bruit résiduel qui reste dans le signal après le rehaussement, il se calcule pendant le silence. L'estimation du $SegMSE$ est basée sur la relation suivante :

$$SegMSE|_{dB} = 20 \log_{10} \sum_{k=0}^{M-1} |e(k)| DAV(k) \quad (4.1)$$

Ou M est la longueur du trame de sortie $e(n)$.

- Le rapport signal à bruit segmenté (Segmental-SNR) qui indique le niveau de l'amplification du signal par rapport au bruit résiduel. L'estimation du $SegSNR$ est basée sur la relation suivante :

$$SegSNR|_{dB} = \frac{10}{M} \sum_{m=0}^{M-1} 10 \log_{10} \left\{ \frac{\sum_{n=Lm}^{Lm+L-1} \tilde{s}(n)^2}{\sum_{n=Lm}^{Lm+L-1} \tilde{b}(n)^2} \right\} \quad (4.2)$$

Avec $\tilde{s}(n)$ et $\tilde{b}(n)$ sont les signaux de la parole et bruit. L est le nombre d'échantillons, et M est le nombre de trame dans le signal.

- Le system mismatch (SM) qui mesure la convergence de l'algorithme vers les coefficients optimaux. L'estimation de SM est basée sur la relation suivante :

$$SM|_{dB} = 10 \log_{10} \left\{ \frac{\|h_{z1} - w_{z1}(k)\|^2}{\|h_{z1}\|^2} \right\} \quad (4.3)$$

Les paramètres intervenants dans le comportement du RNQ automatique sont le pas d'adaptation μ , le facteur d'oubli λ , et la constante j . Pour la simulation, on a fixé ces derniers comme suit :

- $\mu = 0.6$
- $\lambda = 0.99$
- $j = 2$

L'effet du RSB avec une taille de filtre $L = 64$

En utilisant un bruit blanc, et pour une longueur de filtre fixée à $L = 64$, nous avons varié le RSB d'entrée, et fixé un $RSB_{seuil} = -30$ dB.

- $L = 64$ et $RSB = -3$ dB

La figure (4.7) représente une évolution temporelle où nous faisons une comparaison à l'échelle temporelle entre le signal de la parole initial et la sortie de l'RNQ automatique.

Nous remarquons dans la figure (4.7) une qualité acceptable du signal de sortie de l'RNQ proposé par rapport au signal original de la parole.

La figure (4.8) représente le critère SegMSE des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé converge plus rapide et atteint directement une meilleure valeur de MSE qui est autour de -50 dB.

La figure (4.9) représente le critère SegSNR des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé atteint un gain de 40 dB dès le début du traitement.

La figure (4.10) représente le critère System Mismatch des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Nous remarquons dans cette figure que notre algorithme proposé converge dans un petit laps de temps, et reste stable vers la valeur de -40 dB alors que les autres algorithmes restent instables.

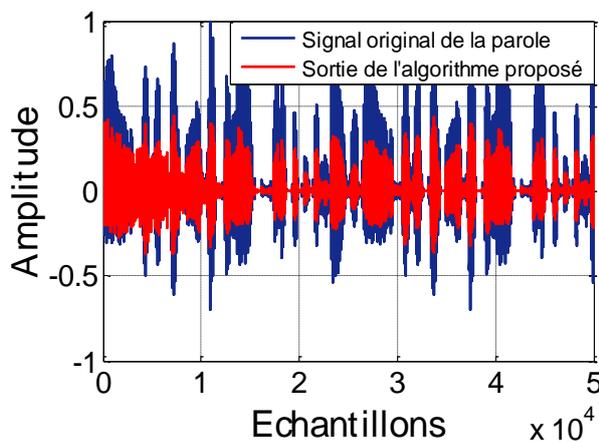


Fig. 4.7 – Evolution temporelle des signaux pour un $RSB = -3$ dB et $L = 64$.

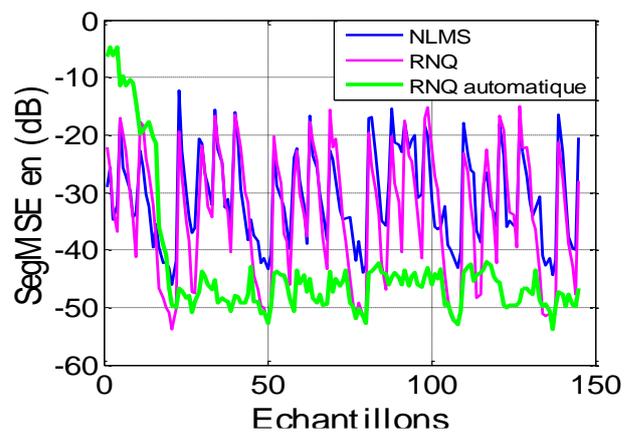


Fig. 4.8 – Critère SegMSE pour un $RSB = -3$ dB et $L = 64$.

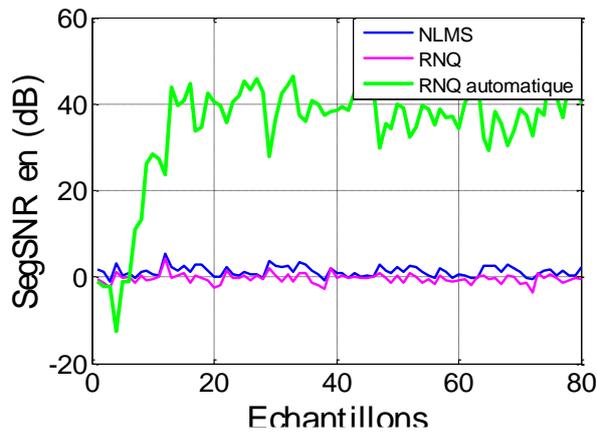


Fig. 4.9 – Critère SegSNR pour un RSB = -3 dB et L = 64.

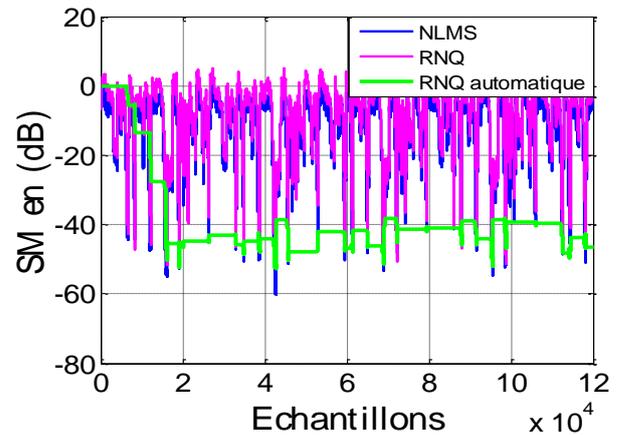


Fig. 4.10 – Critère SM pour un RSB = -3 dB et L = 64.

- L = 64 et RSB = 0 dB

Nous représentons dans la figure (4.11) une comparaison à l'échelle temporelle entre le signal de la parole initial et la sortie de l'RNQ automatique.

Nous remarquons dans cette figure la bonne qualité de la sortie de l'RNQ automatique par rapport au signal original de la parole.

Nous représentons dans la figure (4.12) le critère SegMSE des algorithmes NLMS, RNQ, et RNQ automatique superposés.

D'après cette figure nous remarquons que notre algorithme proposé atteint une faible valeur de SegMSE qui a pu toucher -50 dB plusieurs fois.

La figure (4.13) représente le critère SegSNR des algorithmes NLMS, RNQ, et RNQ automatique.

Nous remarquons dans cette figure que l'RNQ automatique dépasse directement un gain de 40 dB.

Et enfin, dans la figure (4.14) nous représentons le critère System Mismatch des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Nous remarquons bien dans cette figure que notre algorithme proposé dépasse la valeur de -40 dB plusieurs fois.

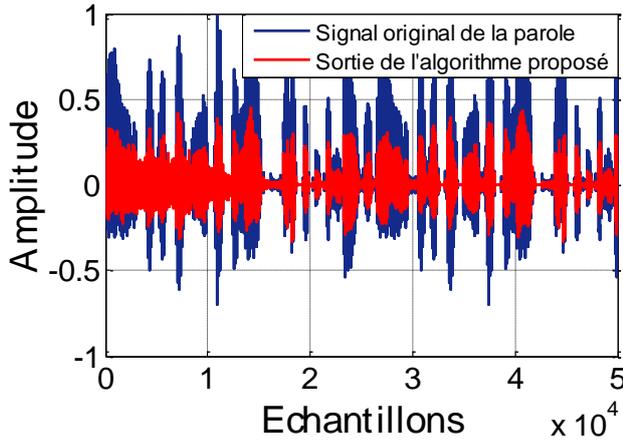


Fig. 4.11 – Evolution temporelle des signaux pour un RSB = 0 dB et $L = 64$.

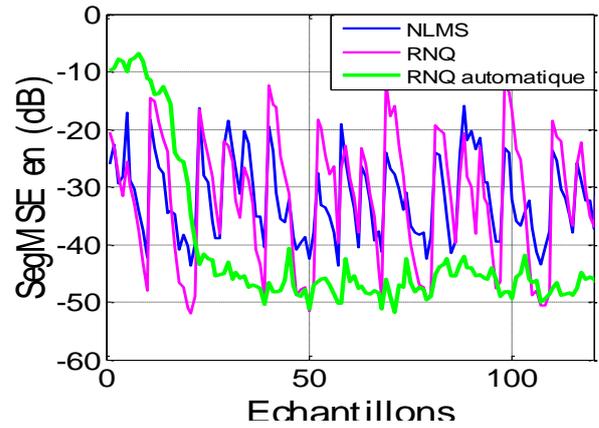


Fig. 4.12 – Critère SegMSE pour un RSB = 0 dB et $L = 64$.

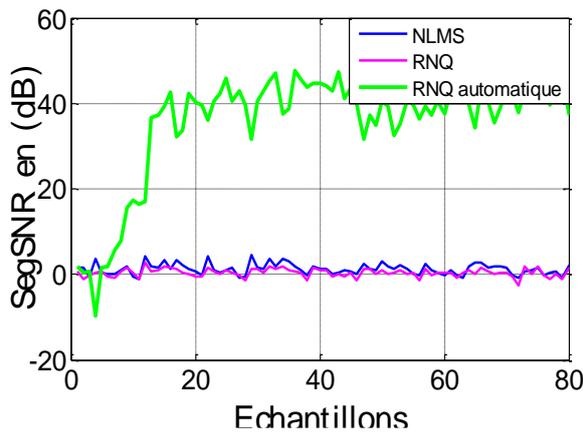


Fig. 4.13 – Critère SegSNR pour un RSB = 0 dB et $L = 64$.

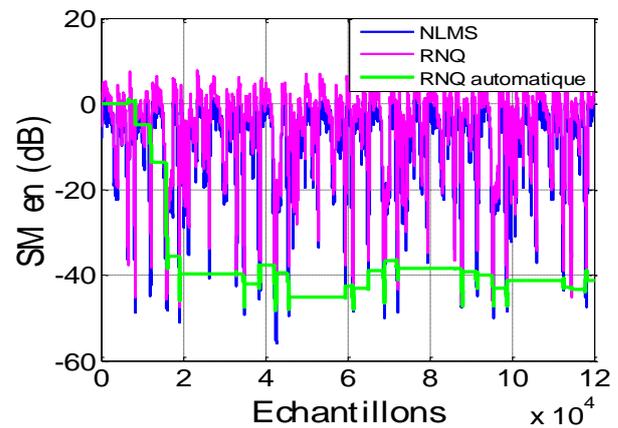


Fig. 4.14 – Critère SM pour un RSB = 0 dB et $L = 64$.

○ $L = 64$ et RSB = 3 dB

Nous représentons dans la figure (4.15) une comparaison à l'échelle temporelle entre le signal de la parole initial et la sortie de l'algorithme proposé.

Nous remarquons bien la bonne qualité de la sortie de l'RNQ automatique par rapport au signal original de la parole.

Nous représentons dans la figure (4.16) le critère SegMSE des algorithmes NLMS, RNQ, et RNQ automatique superposés.

D'après cette figure nous remarquons que notre algorithme proposé maintient une faible valeur de SegMSE qui a pu atteindre -50 dB plusieurs fois.

La figure (4.17) représente le critère SegSNR des algorithmes NLMS, RNQ, et RNQ automatique.

Nous remarquons dans cette figure que l’RNQ automatique dépasse directement un gain de 40 dB.

Et enfin, dans la figure (4.18) nous représentons le critère System Mismatch des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Nous remarquons que notre algorithme proposé dépasse la valeur de -40 dB plusieurs fois.

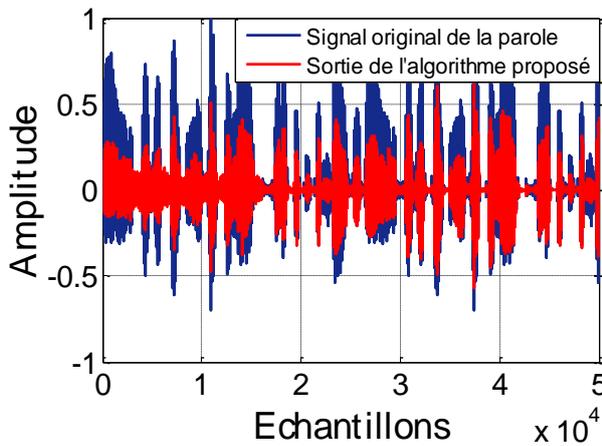


Fig. 4.15 – Evolution temporelle des signaux pour un RSB = 3 dB et $L = 64$.

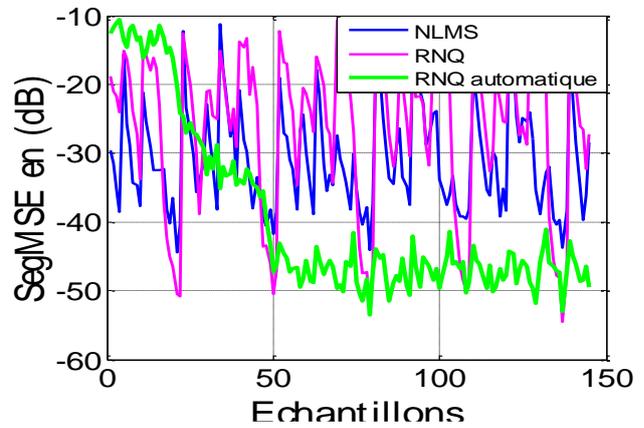


Fig. 4.16 – Critère SegMSE pour un RSB = 3 dB et $L = 64$.

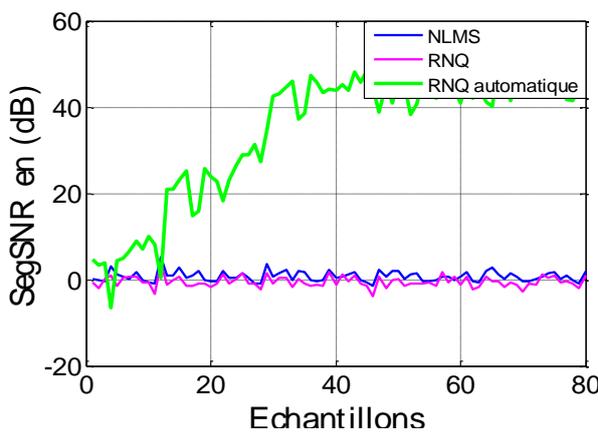


Fig. 4.17 – Critère SegSNR pour un RSB = 3 dB et $L = 64$.

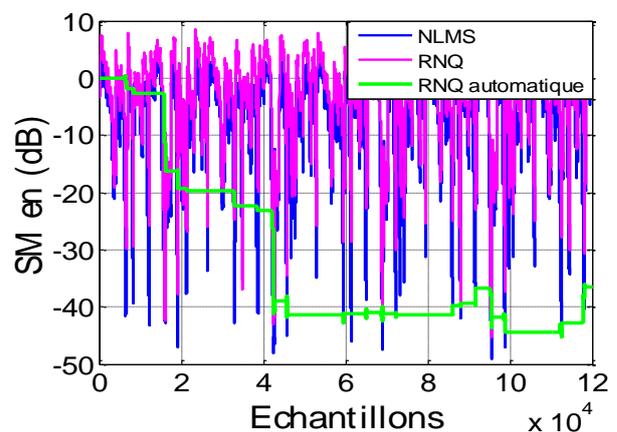


Fig. 4.18 – Critère SM pour un RSB = 3 dB et $L = 64$.

L'effet du RSB avec une taille de filtre $L = 128$

En utilisant toujours un bruit blanc, et pour une longueur de filtre fixée à $L = 128$, nous avons varié le RSB d'entrée, et fixé un $RSB_{seuil} = -20$ dB.

- $L = 128$ et $RSB = -3$ dB

La figure (4.19) représente une évolution temporelle où nous faisons une comparaison entre le signal de la parole initial et la sortie de l'RNQ automatique.

Nous remarquons dans cette figure une qualité acceptable du signal de sortie de l'RNQ proposé par rapport au signal original de la parole.

La figure (4.20) représente le critère SegMSE des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé converge vers une valeur de MSE qui dépasse -20 dB et garde une meilleure stabilité.

La figure (4.21) représente le critère SegSNR des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé atteint plusieurs fois le gain de 20 dB.

La figure (4.22) représente le critère System Mismatch des algorithmes NLMS, RNQ, et RNQ automatique, ou nous remarquons que notre algorithme reste stable autour de la valeur de -20 dB.

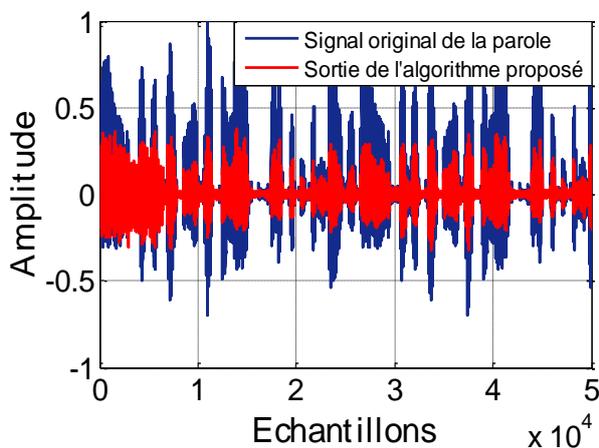


Fig. 4.20 – Evolution temporelle des signaux pour un $RSB = -3$ dB et $L = 128$.

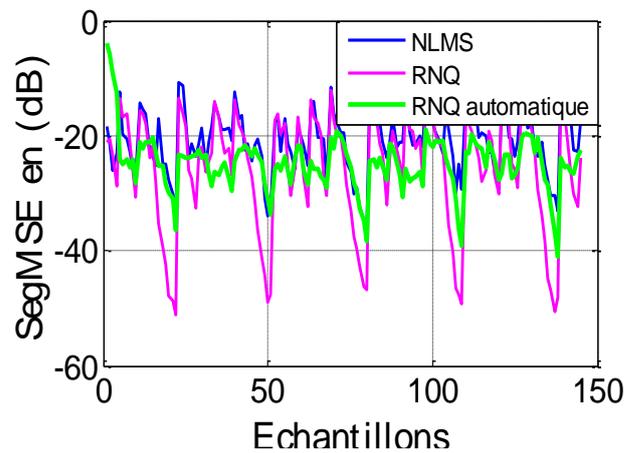


Fig. 4.19 – Evolution temporelle des signaux Critère SegMSE pour un $RSB = -3$ dB et $L = 128$.

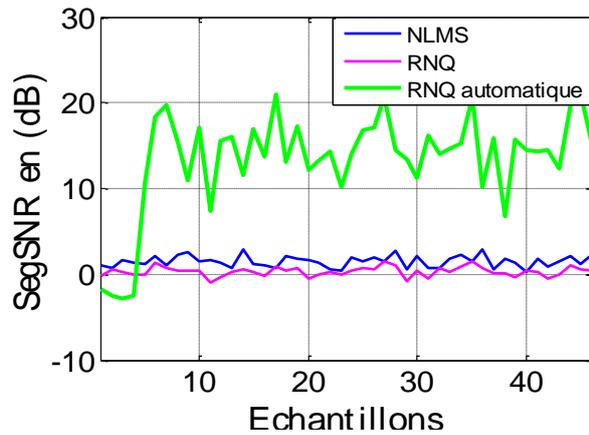


Fig. 4.21 – Critère SegSNR pour
un RSB = -3 dB et L = 128.

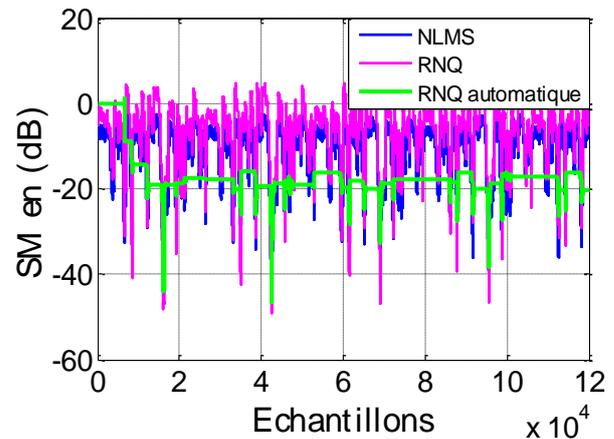


Fig. 4.22 – Critère SM pour
un RSB = -3 dB et L = 128.

- L = 128 et RSB = 0 dB

La figure (4.23) représente une évolution temporelle où nous faisons une comparaison entre le signal de la parole initial et la sortie de l’RNQ automatique.

Nous remarquons dans la figure (4.23) une qualité acceptable du signal de sortie de l’RNQ proposé par rapport au signal original de la parole.

La figure (4.24) représente le critère SegMSE des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé garde une meilleur stabilité et converge vers une valeur de MSE qui dépasse -20 dB et touche - 30 dB.

La figure (4.25) représente le critère SegSNR des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé dépasse plusieurs fois le gain de 20 dB.

La figure (4.26) représente le critère System Mismatch des algorithmes NLMS, RNQ, et RNQ automatique, ou nous remarquons que notre algorithme reste toujours stable autour de la valeur de -20 dB.

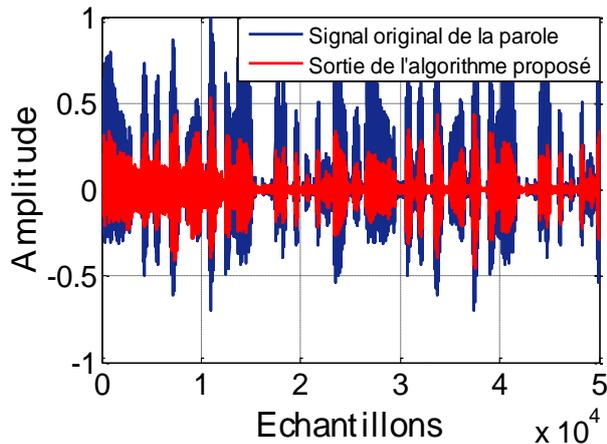


Fig. 4.23 – Evolution temporelle des signaux pour un $RSB = 0$ dB et $L = 128$.

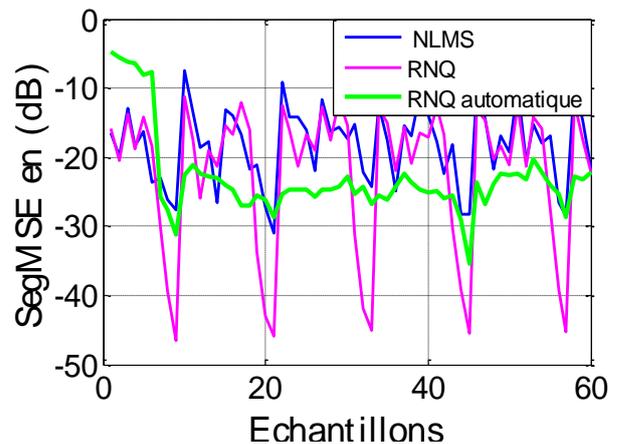


Fig. 4.24 – Critère SegMSE pour un $RSB = 0$ dB et $L = 128$.

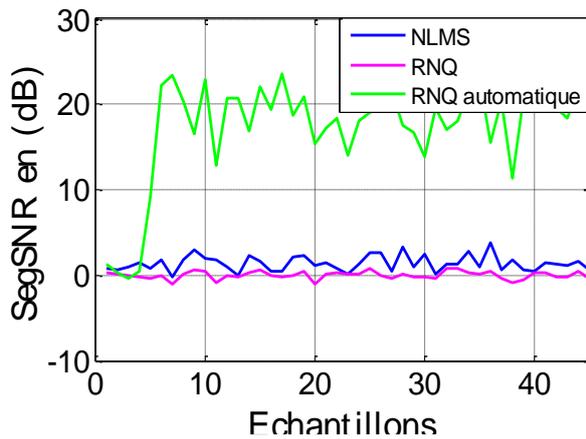


Fig. 4.25 – Critère SegSNR pour un $RSB = 0$ dB et $L = 128$.

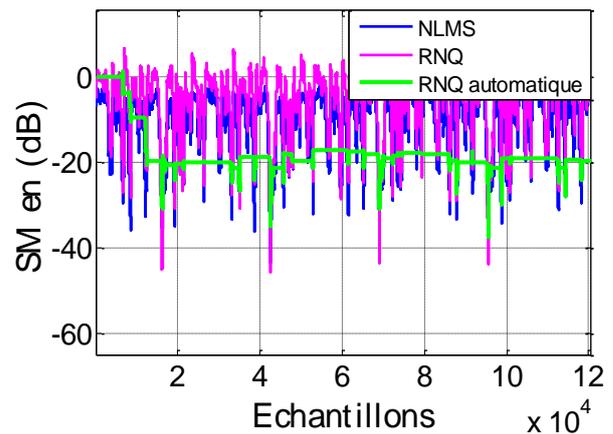


Fig. 4.26 – Critère SM pour un $RSB = 0$ dB et $L = 128$.

- $L = 128$ et $RSB = 3$ dB

La figure (4.27) représente une évolution temporelle où nous faisons une comparaison entre le signal de la parole initial et la sortie de l’RNQ automatique.

Nous remarquons dans la figure (4.27) une qualité toujours acceptable du signal de sortie de l’RNQ proposé par rapport au signal original de la parole.

La figure (4.28) représente le critère SegMSE des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé garde une meilleure stabilité et converge vers une valeur de MSE qui dépasse -25 dB et touche -30 dB plusieurs fois.

La figure (4.29) représente le critère SegSNR des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé dépasse plusieurs fois le gain de 20 dB.

La figure (4.30) représente le critère System Mismatch des algorithmes NLMS, RNQ, et RNQ automatique, ou nous remarquons que notre algorithme reste toujours stable autour de la valeur de -20 dB.

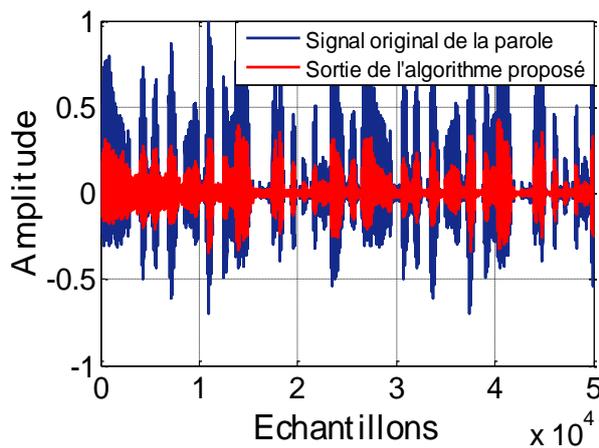


Fig. 4.27 – Evolution temporelle des signaux pour un $RSB = 3$ dB et $L = 128$.

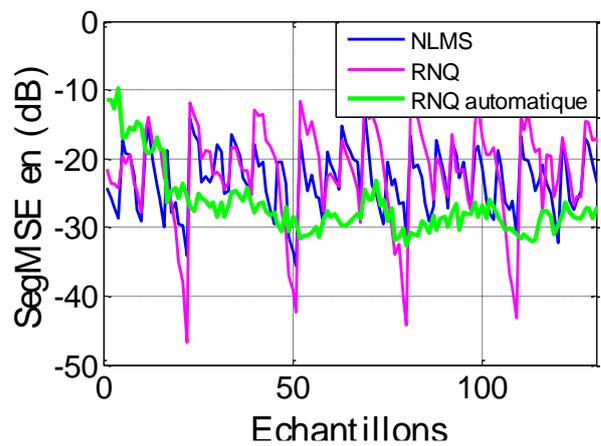


Fig. 4.28 – Critère SegMSE pour un $RSB = 3$ dB et $L = 128$.

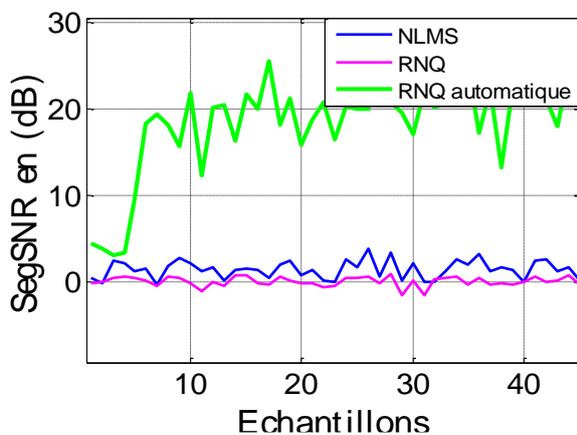


Fig. 4.29 – Critère SegSNR pour un $RSB = 3$ dB et $L = 128$.

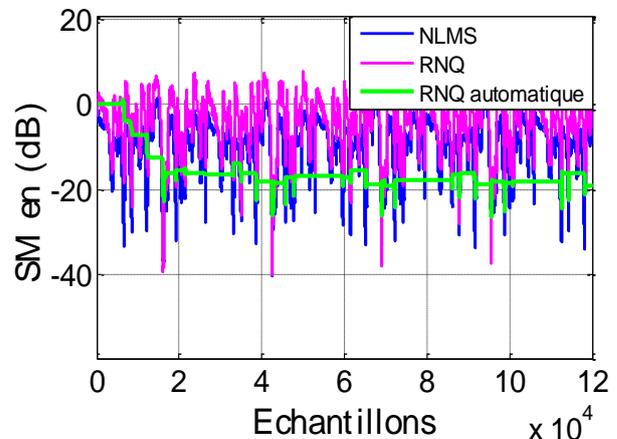


Fig. 4.30 – Critère SM pour un $RSB = 3$ dB et $L = 128$

Résultats de simulations avec d'autres types de bruit

Dans cette partie, nous allons présenter les résultats de simulation avec d'autres types de bruits, pour cela nous proposons en premier lieu, d'utiliser le bruit USASI qui est utilisé souvent pour étudier la vitesse de convergence puisqu'il est stationnaire et corrélé en même temps, c'est la combinaison de la théorie avec la réalité. La figure ci-dessous présente le signal du bruit USASI et son spectrogramme.

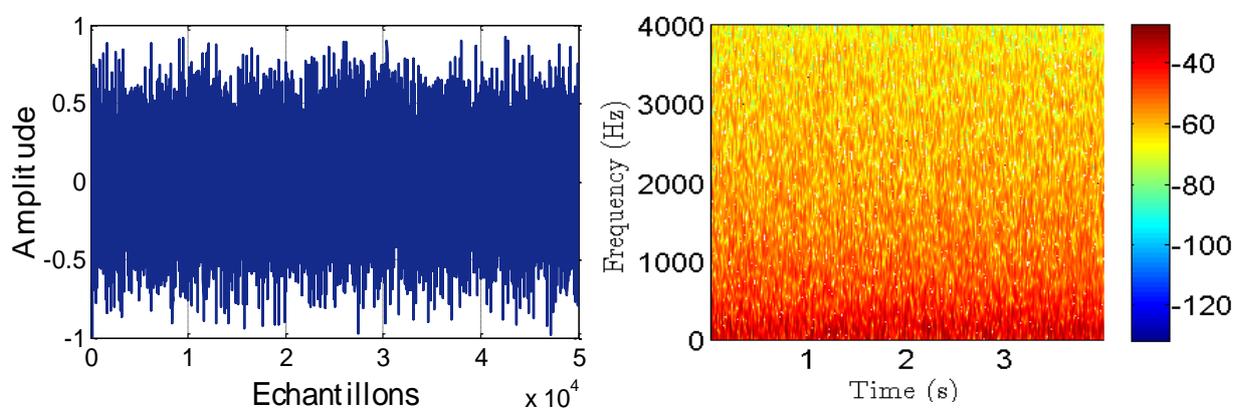


Fig. 4.31 - Signal du bruit USASI et son spectrogramme.

Pour la simulation, on a fixé les paramètres qui interviennent au comportement de l'RNQ comme suit :

- $RSB_1 = RSB_2 = -3 \text{ dB}$ (C.-à-d. à mi-puissance)
- $RSB_{seuil} = -30 \text{ dB}$
- $L = 64$
- $\mu = 0.6$
- $\lambda = 0.99$
- $j = 2$

La figure (4.32) représente une évolution temporelle où nous faisons une comparaison à l'échelle temporelle entre le signal de la parole initial et la sortie de l'RNQ automatique.

Nous remarquons dans la figure (4.32) une bonne qualité du signal de sortie de l'RNQ proposé par rapport au signal original de la parole.

La figure (4.33) représente le critère SegMSE des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé converge plus rapide et atteint directement une meilleure valeur stable de MSE qui est autour de -40 dB.

La figure (4.34) représente le critère SegSNR des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Dans cette figure nous remarquons que notre algorithme proposé est stable autour d'un gain de 30 dB.

La figure (4.35) représente le critère System Mismatch des algorithmes NLMS, RNQ, et RNQ automatique superposés.

Nous remarquons dans cette figure que notre algorithme proposé converge et reste stable vers la valeur de -40 dB alors que les autres algorithmes restent instables.

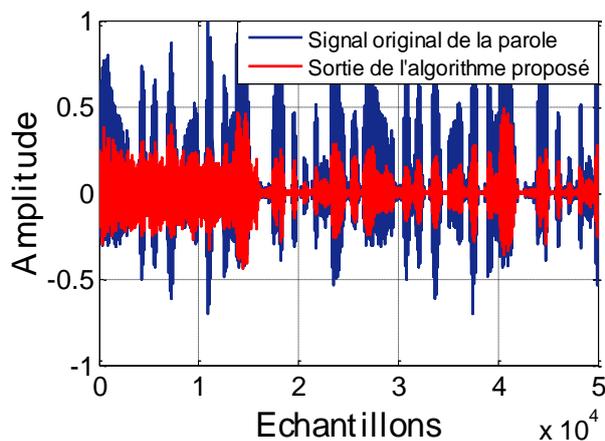


Fig. 4.32 – Evolution temporelle des signaux pour un $RSB = -3$ dB et $L = 64$.

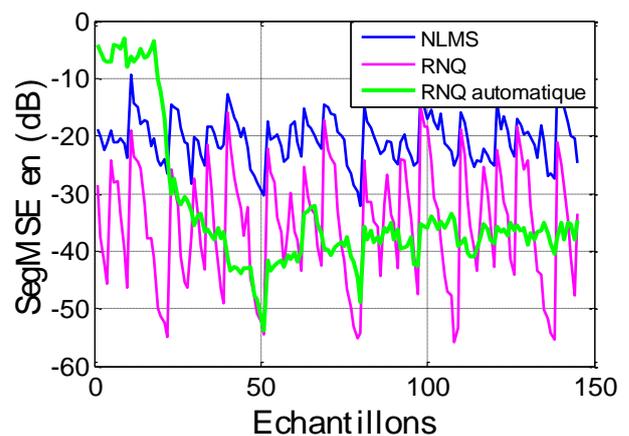


Fig. 4.33 – Critère SegMSE pour un $RSB = -3$ dB et $L = 64$.

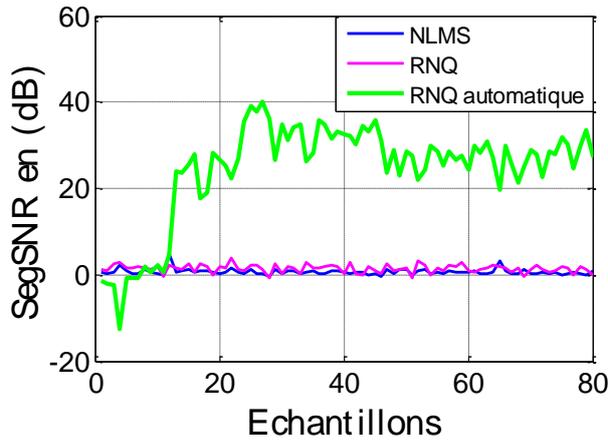


Fig. 4.34 – Critère *SegSNR* pour un $RSB = -3\text{ dB}$ et $L = 64$.

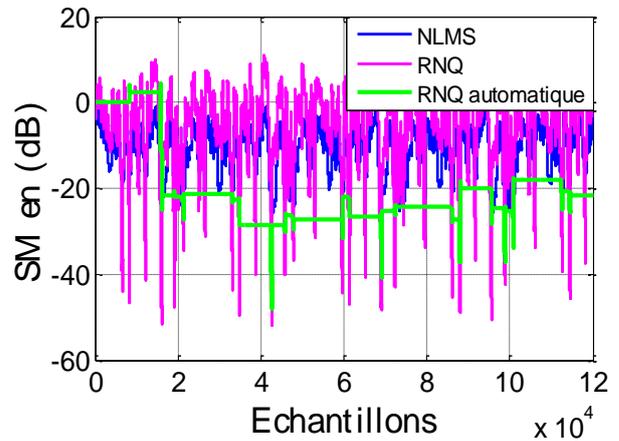


Fig. 4.35 – Critère *SM* pour un $RSB = -3\text{ dB}$ et $L = 64$.

Pour se rapprocher plus de la réalité, nous allons tester maintenant notre algorithme avec d'autres types de bruit qui perturbent la communication mobile le plus fréquemment, pour cela nous choisissons d'utiliser un bruit d'automobile en mouvement, et un bruit d'un phénomène naturel qui est celui de la pluie.

Les figures ci-dessous présentent les signaux du bruit d'une ancienne automobile (à gauche) et de la pluie (à droite) choisis.

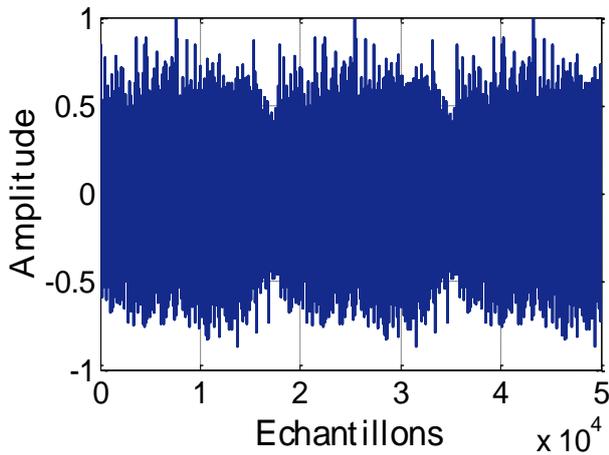


Fig. 4.36 – Signal du bruit d'automobile.

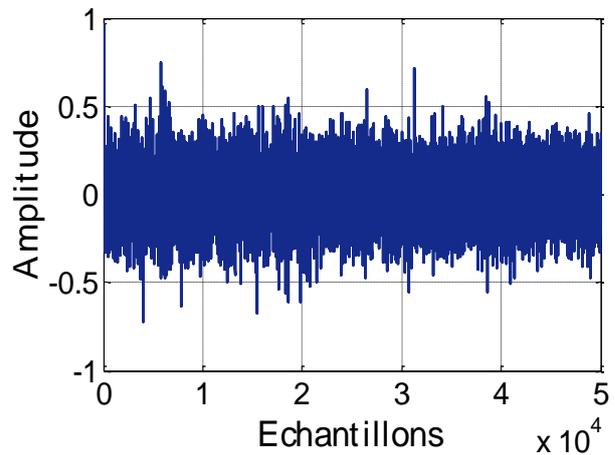


Fig. 4.37 – Signal du bruit de la pluie.

Les figures ci-dessous présentent les spectrogrammes des signaux du bruit d'une ancienne automobile (à gauche) et de la pluie (à droite) choisis.

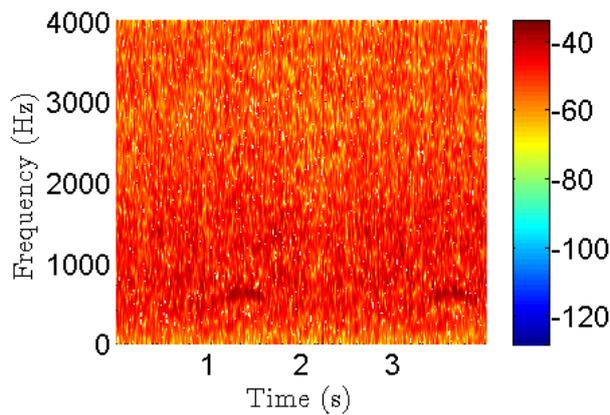


Fig. 4.38 – Spectrogramme du signal du bruit d'automobile.

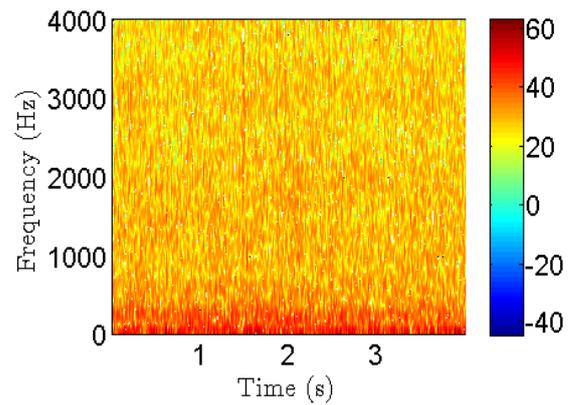


Fig. 4.39 – Spectrogramme du signal du bruit de la pluie.

La figure (4.40) et la figure (4.41) représentent une évolution temporelle où nous faisons une superposition de la sortie de l'RNQ automatique sur le signal de la parole original, et ça en utilisant un bruit de la pluie (à droite), et un bruit d'automobile (à gauche).

Nous remarquons dans ces figures une qualité satisfaisante du signal de sortie de l'RNQ proposé par rapport au signal initial de la parole dans les deux cas.

La figure (4.42) et la figure (4.43) représentent le critère SegMSE des algorithmes NLMS, RNQ, et RNQ automatique superposés, et ça en utilisant un bruit de la pluie (à droite), et un bruit d'automobile (à gauche).

Nous remarquons d'après ces figures, que notre algorithme garde une meilleure stabilité par rapport aux autres algorithmes et maintient une faible valeur de MSE qui dépasse -30 dB.

La figure (4.44) et la figure (4.45) représentent le critère SegSNR des algorithmes NLMS, RNQ, et RNQ automatique superposés et ça en utilisant un bruit de la pluie (à droite), et un bruit d'automobile (à gauche).

Nous remarquons que l'RNQ automatique touche le gain de 30 dB instantanément et maintient sa stabilité autour de cette valeur.

La figure (4.46) et la figure (4.47) représentent le critère SM des algorithmes NLMS, RNQ, et RNQ automatique superposés et ça en utilisant un bruit de la pluie (à droite), et un bruit d'automobile (à gauche).

Nous remarquons que l'RNQ automatique converge rapidement vers la valeur de -30 dB et reste stable en maintenant cette valeur.

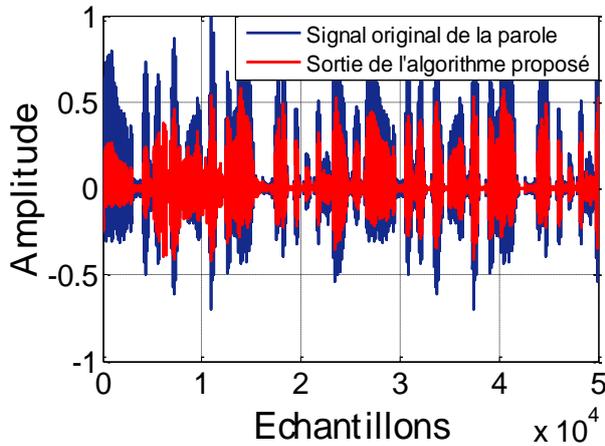


Fig. 4.40 – Evolution temporelle des signaux pour un RSB = -3 dB et $L = 64$ (automobile).

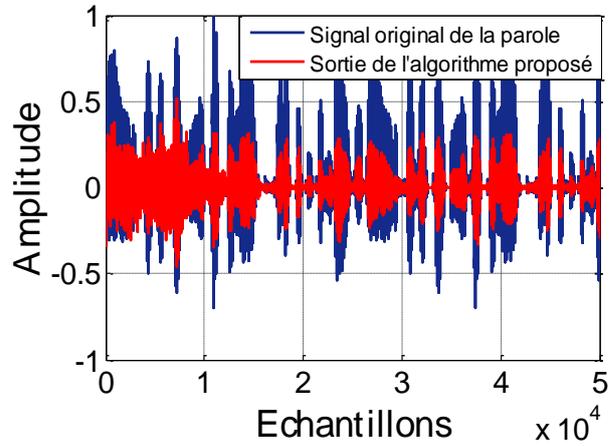


Fig. 4.41 – Evolution temporelle des signaux pour un RSB = -3 dB et $L = 64$ (pluie).

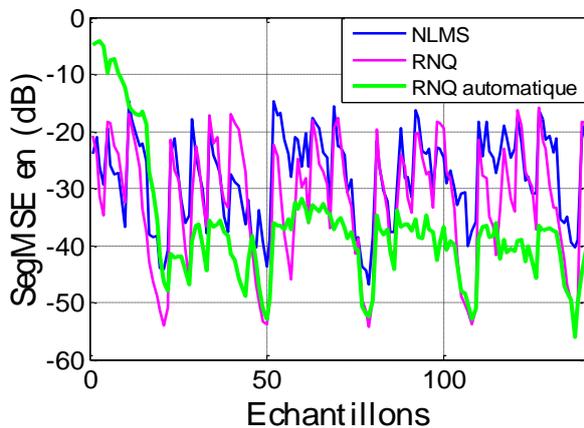


Fig. 4.42 – Critère SegMSE pour un RSB = -3 dB et $L = 64$.

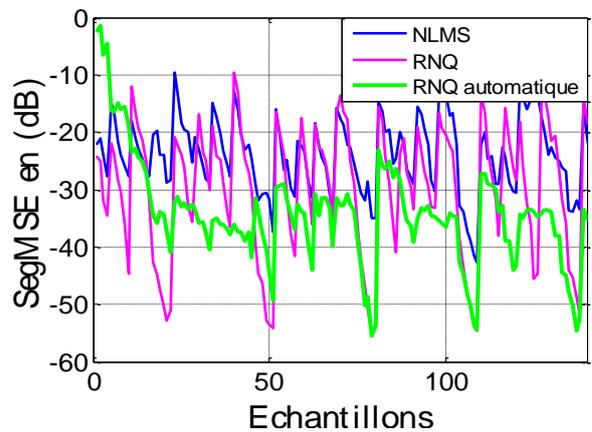


Fig. 4.43 – Critère SegMSE pour un RSB = -3 dB et $L = 64$.

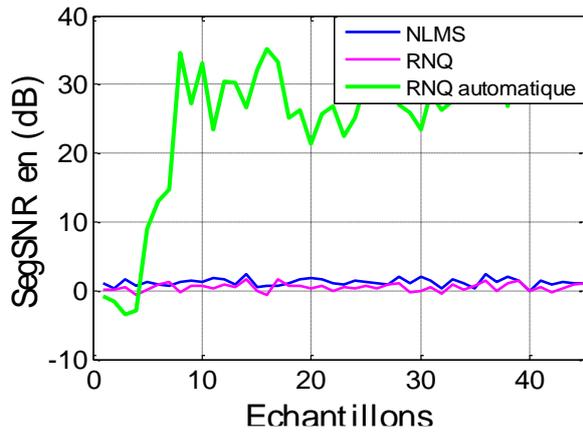


Fig. 4.44 – Critère SegSNR pour un RSB = -3 dB et $L = 64$.

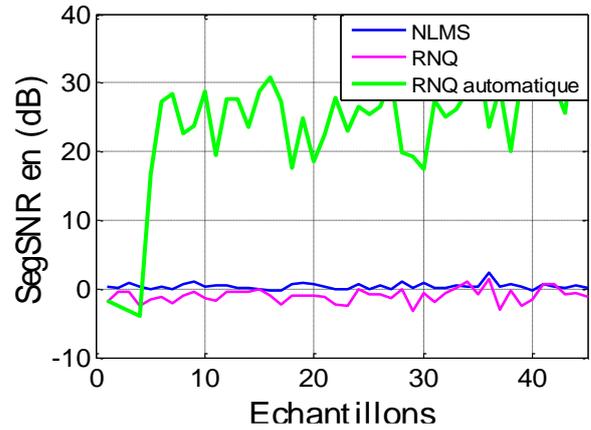


Fig. 4.45 – Critère SegSNR pour un RSB = -3 dB et $L = 64$.

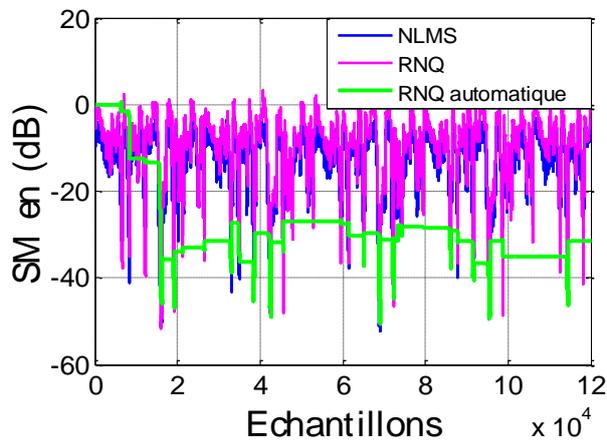


Fig. 4.46 – Critère SM pour un RSB = -3 dB et $L = 64$.

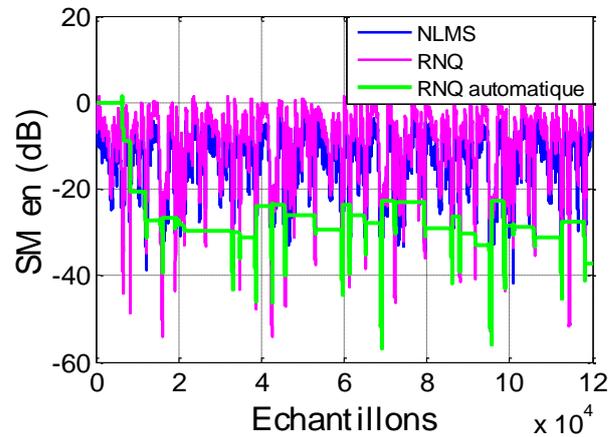


Fig. 4.47 – Critère SM pour un RSB = -3 dB et $L = 64$.

Les figures (4.7), (4.11), (4.15), (4.19), (4.23), (4.27), (4.32), (4.40) et (4.41) montrent la qualité satisfaisante que délivre notre algorithme proposé à la sortie.

Les figures (4.8), (4.12), (4.16), (4.20), (4.24), (4.28), (4.33), (4.42) et (4.43) montrent que les performances de l’RNQ automatique en terme de MSE sont presque les mêmes en changeant la taille du filtre sauf que la vitesse de convergence diminue, on constate alors que la vitesse de convergence de l’algorithme adaptatif est inversement proportionnelle à la taille du filtre L .

Les figures (4.9), (4.13), (4.17), (4.21), (4.25), (4.29), (4.34), (4.44) et (4.45) montrent un niveau très important de l’amplification de la parole par rapport au bruit résiduel atteint par l’RNQ

automatique pour différentes valeurs du rapport signal à bruit d'entrée et différentes tailles du filtre L .

Et finalement, les figures (4.10), (4.14), (4.18), (4.22), (4.26), (4.30), (4.35), (4.46) et (4.47) montrent bien que notre l'algorithme RNQ proposé converge plus rapidement vers les coefficients optimaux que les autres algorithmes.

L'effet de l'ordre de puissance et du facteur d'oubli sur L'EQM

Toujours pour $RSB = -3 \text{ dB}$ et $L = 64$:

La figure (4.48) présente le critère SegMSE pour différentes valeurs de j (ordre de puissance de la fonction d'erreur).

On remarque que la vitesse de convergence et la quantité du bruit réduit par notre algorithme sont proportionnelles à j .

La figure (4.49) présente le critère SegMSE pour différentes valeurs de facteur d'oubli λ .

On remarque que pour $\lambda = 0.6, 0.7, 0.8$ l'algorithme perd sa stabilité, pour

$0.9 \leq \lambda \leq 0.99$ l'algorithme fourni un meilleur résultat en terme de stabilité, et pour $\lambda = 1$ l'algorithme ne fonctionne pas.

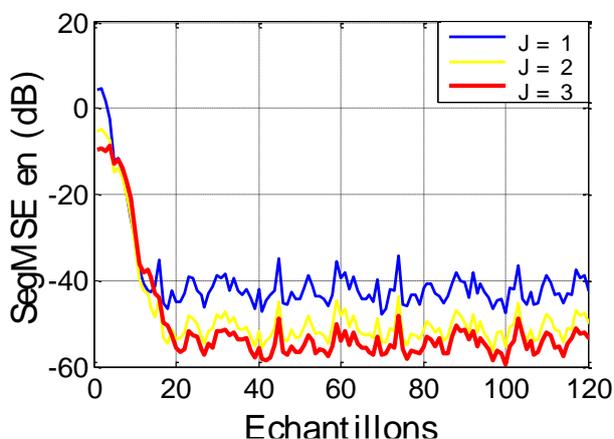


Fig. 4.48 – Critère SegMSE pour différentes valeurs de j .

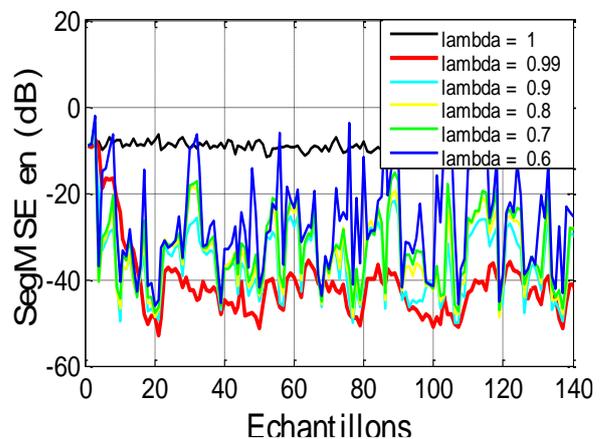


Fig. 4.49 – Critère SegMSE pour différentes valeurs de λ .

4.5 Conclusion

Dans ce chapitre, on a comparé l'RNQ automatique avec l'RNQ et l'NLMS classique, nous avons présenté nos résultats en introduisant un signal de parole dans un environnement bruité (bruit blanc, bruit USASI, bruit d'automobile et bruit de la pluie) et en utilisant différentes valeurs de paramètres de simulation.

Ces résultats de simulation ont bien montré le bon comportement de l'algorithme proposé en terme de convergence et réduction de bruit, ce qui permet à ce dernier d'avoir une vitesse importante pour atteindre la solution optimale du filtrage adaptatif.

Conclusion générale

Le but principal de ce projet de fin d'étude était de remédier aux problèmes causés par le bruit acoustique durant les transmissions du signal utile (téléphonie mobile) par les différentes techniques de réduction du bruit.

Nous avons commencé par une étude théorique sur le filtrage et les algorithmes de filtrage adaptatif tel que LMS et NLMS, ensuite nous avons abordé les principales méthodes et techniques de réduction du bruit avec acquisition mono-capteur, bi-capteurs et multi capteurs via la technique de séparation aveugle de source.

Dans une deuxième partie nous avons présenté le développement de notre algorithme proposé RNQ et son fonctionnement, qui est basé sur la détection automatique de l'activité vocale, en utilisant le rapport signal sur bruit RSB comme paramètre de détection.

Après plusieurs tests et évaluations des critères de performances (SegMSE, SegSNR, SM) pour différentes valeurs des paramètres de simulation (longueur du filtre, type du bruit, facteur d'oubli) sous le logiciel MATLAB, nous avons pu démontrer que notre algorithme est plus performant en terme de convergence, stabilité et réduction du bruit en le comparant avec l'algorithme du gradient stochastique NLMS.

En fin, notre travail nous a permis de découvrir un volet du domaine de réduction de bruit et de rehaussement de la parole qui est un domaine vaste et en plein développement.

Bibliographie

[1] A. Amrane, K. Ould Ammar, "Nouvelle technique automatique de réduction de bruit acoustique basée sur le principe de séparation aveugle de source", mémoire de master, USDB 2019.

[2] Cour de phonétique, Université de Lausanne UNIL.

[3] P. Loiseau, "Contrôle actif acoustique du bruit large bande dans un habitacle automobile", Thèse de doctorat, l'Ecole national supérieur des mines de Nantes, 2016.

[4] R. Boumegouas, et M. Izri, "Annulation du bruit acoustique par un nouvel algorithme APA Backward bi-capteurs", mémoire de master, USDB 2020.

[5] I. Hammouda, et F. Chemmai, "Nouvelle structure de séparation aveugle de source pour le débruitage automatique du signal de parole en sous bandes", mémoire de master, USDB 2020.

[6] M. Omari, et S. Abdat, "Nouvelle version proportionnelle de l'algorithme Backward bi-capteurs pour la réduction du bruit acoustique", mémoire de master, USDB 2019.

[7] D. Boucherf, "Annulation d'écho acoustique bi-voie par les algorithmes de filtrage adaptatif", mémoire de magister, USDB 2007.

[8] B. Widrow, et M.E. Hoff, "Adaptive switching circuits", WESCOM Conv. Rec., V4, (1960), 96-140.

[9] S. Haykin, "Adaptive Filter Theory", 4th edition, Upper Saddle River, New Jersey: Prentice Hall, 2002.

[10] M. Arezki, cour de théorie de signal, USDB 2019.

[11] X. Huang, A. Acero, et H.-W. Hon, "Spoken language processing: a guide to theory, algorithm, and system development", Upper Saddle River, New Jersey: Prentice Hall, 2001.

[12] L. Belley, "Séparation de source aveugle utilisant la transformée en ondelettes continues", mémoire pour l'obtention de maîtrise en génie électrique, université de Québec, 2008

[13] I. Gueraini, et R. Abdi, "Egalisation des canaux de transmission numérique par un algorithme récursif non quadratique", mémoire de master, USDB 2018.

[14] M. Djendi, A. Bentahar, Y. Benghanem, "3rd CISTEM'18 - Algiers, Algeria, October 29-31, 2018".

[15] M. Djendi, A. Gilloire, P. Scalart, "Noise Cancellation Using Two Closely Spaced Microphones, Experimental Study with a specific Model and Two Adaptive Algorithms", Proceedings - ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing. 3.10.1109/ICASSP.2006.1660761, Toulouse, 2006.