

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



**Université de Batna -2-
Faculté de Technologie
Département d'Électronique**



THÈSE

**Présentée pour l'obtention du diplôme de
DOCTORAT EN SCIENCES
Spécialité : Électronique
Option: Contrôle**

Par

ZERDOUMI Zohra

Thème

Estimation des Filtres de Restauration des Signaux en Communications Numériques

Soutenue le :03/05 /2018

Devant le jury :

Dr. FORTAKI Tarek	Prof.	Université Batna 2	Président
Dr. BENATIA Djamel	Prof.	Université Batna 2	Rapporteur
Dr. CHIKOUCHE Djamel	Prof.	Université M'Sila	Co-rapporteur
Dr. ATHAMENA Noureddine	M.C.A.	Université Batna 2	Examineur
Dr. KHEDROUCHE Djamel	Prof.	Université M'Sila	Examineur
Dr. BELATTAR Mounir	M.C.A.	Université Skikda	Examineur

Année universitaire :2017/2018

A ma chère mère, à mon cher père

*Pour leur soutien et leur patience
Qu'Allah leur accorde santé et prospérité*

A mes chères sœurs, à mes chers frères

Pour m'avoir soutenu et encourager

A toutes mes amies et collègues

Pour m'avoir incité à progresser

Remerciements

Louange à Allah le plus miséricordieux pour ses innombrables bénédictions et pour m'avoir donné la force d'accomplir ce travail.

Je voudrais exprimer ma plus grande reconnaissance à mon encadreur monsieur **Djamel Benatia**, professeur à l'université de Batna et mon Co-encadreur, monsieur **Djamel Cikouche**, professeur à l'université de M'sila qui ont dirigé cette thèse avec diplomatie, efficacité et rigueur. Je leur en témoigne mes sincères reconnaissances pour leurs encadrements, leurs conseils et leurs orientations.

J'exprime mes remerciements aux membres de jury, d'avoir bien voulu consacrer du temps et d'attention à mon travail de thèse à savoir :

Monsieur **Fortaki Tarak**, professeur à l'université de Batna pour l'honneur qui il me fait de présider le jury de cette thèse.

Monsieur **Khadrouche Djamel**, professeur à l'université de M'sila,

Monsieur **Athamena Noureddine** MCA à l'université de Batna,

Monsieur **Bellatar Mounir** MCA à l'université de Skikda,

d'avoir accepté de faire partie du jury en tant qu'examinateurs. Je leurs témoigne ici ma profonde reconnaissance.

Je tiens à saluer ardemment Abdou Latifa, MCA à l'université de Biskra, pour son amitié, sa disponibilité et ses conseils précieux. Je tiens à remercier vivement Kebaili Farida et Benmeddour Fadila, MCB à l'université de M'sila pour leurs conseils avisés.

Je voudrais exprimer ma plus grande reconnaissance à toute ma famille, mes parents, mes frères et sœurs pour leurs encouragements et leurs soutiens, je leur en témoigne ici l'expression de ma profonde gratitude.

Mes remerciements s'adressent aussi à toutes mes amies et tous mes collègues, ainsi que l'ensemble des personnes qui ont contribué à l'aboutissement de ce travail.

Table des matières

Dédicaces.....	i
Remerciements	ii
Résumé.....	iii
Table des matières	vi
Liste des figures.....	x
Liste des tableaux	xiii
Liste des acronymes.....	xiv
Introduction générale	1
Chapitre 1 Notions sur les communications numériques.....	8
1.1 Introduction.....	8
1.2 Structure d'un système de communication numérique	8
1.2.1 La source d'information.....	9
1.2.2 L'émetteur	10
1.2.2.1 Le codage source.....	10
1.2.2.2 Le codage canal	10
1.2.2.3 Le modulateur.....	10
1.2.3 Le canal de transmission	11
1.2.4 Le récepteur.....	11
1.2.4.1 Le démodulateur.....	12
1.2.4.2 Le décodeur de canal.....	12
1.2.4.3 Le décodeur de source.....	12
1.3 Modèles de canaux de communication.....	12
1.3.1 Canal Gaussien.....	12

1.3.2 Canal linéaire	13
1.3.3 Canal non linéaire	15
1.4 Les effets perturbateurs.....	16
1.4.1 Le bruit	16
1.4.2 Propagation multi-trajets.....	18
1.4.3 Dispersion du canal	18
1.4.4 Sélectivité en fréquence	19
1.4.5 Les interférences inter-symboles	19
1.5 Mesures de performances en communications numériques	21
1.5.1 Le rapport signal sur bruit.....	21
1.5.2 L'erreur quadratique moyenne	22
1.5.3 La probabilité d'erreur	23
1.5.4 Le taux d'erreur.....	25
1.6 Les filtres numériques.....	26
1.6.1 Les filtres à réponse impulsionnelle finie	27
1.6.2 Les filtres à réponse impulsionnelle infinie	27
1.6.3 Choix entre un filtre RIF et un filtre RII	29
1.7 Conclusion	29
Chapitre 2 Filtrage adaptatif et algorithmes d'estimation.....	30
2.1 Introduction.....	30
2.2 Les filtres adaptatifs.....	30
2.3 Critères d'estimation.....	31
2.3 .1 Critère du maximum de vraisemblance.....	31
2.3 .2 Critère du maximum à postériori MAP.....	32
2.4 Structures des filtres adaptatifs	32
2.5 Erreur quadratique moyenne.....	33
2.6 Les algorithmes d'estimation.....	35
2.6.1 L'algorithme des moindres carrés moyens	35
2.6.2 L'algorithme des moindres carrés récurrents	37
2.7 Applications des filtres adaptatifs	39
2.8 Filtres adaptatifs en égalisation.....	39
2.8 .1 Principe de l'égalisation.....	40

2.8 .2 Modes opératoires des filtres égaliseurs	41
2.8.2.1 Mode supervisé.....	41
2.8.2.2 Mode de décision.....	42
2.8 .3 Choix de la structure du filtre égaliseur	43
2.9 Types des filtres égaliseurs	44
2.9.1 Filtre égaliseur par forçage à zéro	44
2.9.2 Filtre égaliseur transversal	47
2.9.3 Filtre égaliseur en treillis.....	50
2.10 Conclusion	53
Chapitre 3 Filtrage adaptatif non linéaire à base des réseaux de neurones.....	55
3.1 Introduction.....	55
3.2 Aperçu général sur les réseaux neuronaux	55
3.2.1 Le neurone artificiel	56
3.2.2 Théorème de l'approximation universelle	59
3.2.3 Apprentissage des réseaux de neurones	60
3.2.3 .1 Apprentissage supervisé	60
3.2.3 .2 Apprentissage non supervisé	60
3.3 Types de réseaux de neurone	61
3.3.1 Les réseaux neuronaux de type feed-forward	61
3.3.1.1 Le perceptron multicouche.....	62
3.3.1.2 Le réseau à fonction de base radiale.....	63
3.3.1.3 Comparaison entre le réseau MLP et le réseau RBF	65
3.3.2 Les réseaux neuronaux récurrents (de type feedback)	65
3.4 Les filtres égaliseurs neuronaux	67
3.5 Les filtres égaliseurs à base du Perceptron multicouches.....	68
3.5.1 Le filtre égaliseur transversal à base du perceptron multicouches.....	68
3.5.2 Le filtre égaliseur en treillis à base du perceptron multicouche.....	69
3.6 Algorithme d'adaptation du MLP	70
3.6.1 Fonctions d'activations sigmoïdales	70
3.6.2 Algorithme de la back propagation avec fonction d'activation adaptative ..	72
3.7 Conclusion	77

Chapitre 4 Etude des performances des filtres égaliseurs neuronaux	78
4.1. Introduction.....	78
4.2. Modèles des canaux de communication	78
4.2.1 Canaux linéaires	79
4.2.2 Canaux non linéaires	80
4. 3. Etude des performances des filtres égaliseurs	82
4.3.1 Qualité des signaux restaurés	82
4.3.2 Les courbes de convergence MSE	88
4.3.3 Etude du taux d'erreur par bit (BER).....	93
4.4. Étude des capacités de convergence et de poursuite des filtres égaliseurs dans un environnement non stationnaire.....	97
4.4.1 Changement abrupt dans l'environnement.....	97
4.4.2 Canal linéaire non stationnaire.....	99
4.4.3 Canal non linéaire non stationnaire.....	101
4.5. Conclusion	104
Conclusion générale.....	105
Annexes.....	107
Annexe A: L'algorithme GAL.....	107
Annexe B: Les effets de distorsions des canaux et code MATLAB.....	110
B.1 Les effets de distorsions des canaux.....	110
B.2 Code MATLAB.....	111
Annexe C: Caractéristiques du canal et du filtre égaliseur linéaire.....	112
C.1 Réponses impulsionnelles.....	112
C.2 Réponses fréquentielles.....	113
Bibliographie	114

Liste des figures

1.1	Schéma Synoptique d'un système de communication numérique	9
1.2	Modélisation du canal: (a) canal Gaussien (b) canal linéaire.....	15
1.3	Canal discret linéaire.....	15
1.4	Modèle d'un canal non linéaire.....	16
1.5	Les statistiques du bruit blanc Gaussien: (a) Fonction densité de probabilité (b) Densité spectrale de puissance (c) Fonction d'auto-corrélation.....	18
1.6	Circuit de décision (a); Probabilité d'erreur (b)	25
1.7	Structure d'un filtre à réponse impulsionnelle finie (RIF)	27
1.8	Architecture d'un filtre récursif.....	28
2.1	Configuration générale d'un filtre adaptatif.....	31
2.2	Configuration d'un filtre adaptatif en égalisation.....	40
2.3	Modes opératoires d'un filtre égaliseur.....	42
2.4	Filtres égaliseurs et algorithmes correspondants.....	44
2.5	Filtre égaliseur par forçage à zéro.....	46
2.6	Filtre égaliseur linéaire transversal.....	58
2.7	Filtre égaliseur en treillis.....	51
3.1	Modèle de base du neurone artificiel.....	57
3.2	Fonctions d'activations courantes.....	58
3.3	Modes d'apprentissage des réseaux de neurone (a) supervisé(b) non supervisé.....	60
3.4	Le perceptron multicouche MLP.....	62
3.5	Architecture du réseau RBF.....	63
3.6	Réseau récurrent NARX	67
3.7	Filtre égaliseur à base du perceptron multicouche (TMLP).....	69

3.8	Filtre égaliseur en treillis à base du perceptron multicouche (LMLP).....	70
3.9	Sigmoïde bipolaire à paramètre variable.....	72
4.1	Les caractéristiques des canaux:(a) Réponses en amplitude (b) Réponses en phase (c) Les zéros de Ch_1 (d) Les zéros de Ch_2	80
4.2	Canal non linéaire.....	81
4.3	Signaux restaurés, TMLP, canal Ch_1 : (a) BP, (b) BPAAF: signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)	84
4.4	Signaux restaurés, TMLP, canal Ch_2 : (a) BP, (b) BPAAF : signal transmis (Transig), signaux reçu (Recsig), signal restauré (Resig).....	84
4.5	Signaux restaurés, TMLP, canal NCh_1 (a) BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig).....	85
4.6	Signaux restaurés, TMLP, canal NCh_2 (a) BP, (b) BPAAF: signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)	85
4.7	Signaux restaurés, LMLP, canal Ch_1 : (a) BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)	86
4.8	Signaux restaurés, LMLP, canal Ch_2 : (a) BP, (b) BPAAF: signal transmis (Transig), signaux reçu (Recsig), signal restauré (Resig).....	86
4.9	Signaux restaurés, LMLP, canal NCh_1 : (a) BP, (b) BPAAF: signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig).....	87
4.10	Signaux restaurés, LMLP, canal NCh_2 : (a) BP, (b) BPAAF: signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)	87
4.11	Courbes de convergence des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal ch_1	89
4.12	Courbes de convergence des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal ch_2	89
4.13	Courbes de convergence des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal Nch_1	91
4.14	Courbes de convergence des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal Nch_2	91
4.15	Courbes de taux d'erreur des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal ch_1	94
4.16	Courbes de taux d'erreur des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal ch_2	94
4.17	Courbes de taux d'erreur des filtres égaliseurs TMLP et LMLP; algorithmes de BP	

	et BPAAF, canal Nch_1	95
4.18	Courbes de taux d'erreur des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal Nch_2	95
4.19	Capacité de poursuite des filtres égaliseurs sous un changement abrupt dans l'environnement.....	98
4.20	Variations des coefficients du canal linéaire	99
4.21	Courbes de convergence, filtre égaliseur TMLP, algorithmes de BP et BPAAF, canal linéaire variable.....	100
4.22	Courbes de convergence, filtre égaliseur LMLP, algorithmes de BP et BPAAF, canal linéaire variable.....	101
4.23	Variations des coefficients du canal non linéaire.....	102
4.24	Courbes de convergence, filtre égaliseur TMLP, algorithmes de BP et BPAAF, canal non linéaire variable.....	103
4.25	Courbes de convergence, filtre égaliseur LMLP, algorithmes de BP et BPAAF, canal non linéaire variable.....	103
B.1	Les effets de distorsions du canal Ch_1 : (a) Canal idéal (b) Effet de l'ISI (c) Effets de l'ISI+NL (d) effets de l'ISI + NL+bruit, SNR 20dB.....	110
B.2	Les effets de distorsions du canal Ch_2 : (a) Canal idéal (b) Effet de l'ISI (c) Effets de l'ISI+NL (d) effets de l'ISI + NL+bruit, SNR 20dB.....	111
C.1	Réponses impulsionnelles: (a) du canal (b) du filtre égaliseur (c) du système (canal, filtre égaliseur).....	112
C.2	Réponses fréquentielles: (a) du canal (b) du filtre égaliseur (c) du système (canal, filtre égaliseur).....	113

Liste des Tableaux

2.1	L'algorithme LMS.....	36
2.2	L'algorithme RLS.....	39
2.3	L'algorithme GAL.....	53
4.1	Analyse des performances MSE des filtres égaliseurs.....	92
4.2	Analyse des performances en taux d'erreur BER des filtres égaliseurs.....	96

Liste des acronymes

ANN	Artificial Neural network
ABC	Artificial Bee Colony
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate (taux d'erreur par bit)
BP	Back Propagation
BPAAF	Back Propagation with Adaptive Activation Function
BPSK	Binary phase-shift keying
CMA	Constant Modulus Algorithm
DA	Data-Aided,
DD	Directed Decision
DFE	Decision Feedback Equalizer
DSO	Directed Search Optimization
ELM	Extreme Learning machine
FFF	Feed Forward Filter
FB-TDL	Feedback Time Delay Line
FF-TDL	Feed Forward Time Delay Line
FBF	Feed Back Filter
GAL	Gradient Adaptive Lattice
ISE	Instantaneous Square Error
ISI	Inter Symbol Interference
LMLP	Lattice Multilayer Perceptron
LMS	Least mean square
LTE	Linear Transversal Equalizer
LIT	Linéaire Invariant dans le Temps

LSL	Least Square Lattice
MAP	Maximum a Posteriori
MLP	Multi Layer Perceptron
MP	Minimum Phase
MSE	Mean Square Error
MV	Maximum de vraisemblance
NARX	Nonlinear Auto-Regressive with exogenous
NDA	Non-Data-Aided
NMP	Nom Minimum Phase
PD	Peak Distortion
PDF	Probability Density Function
PE	Preset Equalizer
PSD	Power Spectral Density
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
RIF	Réponse Impulsionnelle Finie
RII	Réponse Impulsionnelle Infinie
RLS	recursive least square
RNN	Recurrent Neural Networks
RTRL	Real-Time Recurrent Learning
TDL	Time Delay Line
TDNN	Time Delay Neural Networks
TMLP	Transversal Multilayer Perceptron
SNR	Signal to Noise Ratio
SVM	Support Vector Machine
WLS	Weighted Least Square
WNN	Wavelet Neural Network
ZF	Zero Forcing
ZFA	Zero Forcing Algorithm

Introduction générale

De nos jours, la communication par des moyens techniques prend plus d'ampleur et devient si courante dans nos vies quotidiennes que nous ne considérons pas la complexité des systèmes impliqués. Les améliorations faites dans ce domaine sont dues aux techniques numériques qui ont permis d'atteindre des meilleures performances au niveau de la fiabilité, la flexibilité, et la qualité de transmission. Par ailleurs, le développement de la microélectronique et les capacités de miniaturisation ont permis la mise en œuvre des techniques complexes dans des appareils de taille réduite.

Que ce soit la téléphonie, la télévision ou l'internet, il est évident que nous sommes constamment entourés par des systèmes de communications. Ce fait incite d'ailleurs à augmenter continuellement les capacités de ces systèmes qui de plus en plus requièrent la transmission de quantités importantes d'information dans des bandes de fréquences les plus étroites possibles [1].

L'efficacité de la largeur de bande est devenue un intérêt en croissance avec l'augmentation du taux de données dans un réseau de communication en expansion. C'est cette quête qui est la motivation principale des recherches effectuées en communications [2,3]. Toutefois, la limite physique des médiums de communication constitue un obstacle à la transmission fiable de l'information, à cause des limitations en largeur de bande et des propagations multi-trajets qui engendrent des interférences entre symboles (ISI-inter symbol interference) [1,4].

Ces interférences se produisent pour des canaux de transmission ayant l'amplitude de la réponse en fréquence non constante sur la bande passante et une phase non linéaire [1]. Dans,

le cas du canal radio micro-ondes et téléphone cellulaire, les vitesses de transmission sont très grandes, ce qui rend l'ISI très élevée. De plus l'environnement fluctue continuellement et comporte en général divers obstacles qui entravent la libre propagation des ondes électromagnétiques, au récepteur on reçoit des versions décalées dans le temps, du même signal [1,4].

Dans le cas d'un canal satellitaire, les effets de la propagation multi trajet sont peu importants, toutefois, l'utilisation d'amplificateurs haute puissance entraîne de la distorsion non-linéaire [5,6]. L'ISI non-linéaire est cependant plus difficile à corriger que l'ISI linéaire d'un canal téléphonique ou radio [5,6].

Toutes ces perturbations altèrent le signal d'origine d'une façon constructive ou destructive, il est alors fondamental de restaurer l'entrée du canal en observant sa sortie [2, 3,7]. Une simple approche consiste à concevoir au niveau du récepteur un filtre spécifique, appelé égaliseur qui peut être un filtre à réponse impulsionnelle finie (RIF). Ce filtre utilise les observations du canal pour éliminer les distorsions induites par la transmission et restaurer le signal transmis.

Cette technique dite d'égalisation est très étudiée et constitue une recherche très axée dans le domaine du filtrage adaptatif pour les communications numériques [1,3, 8-11]. D'ailleurs, l'égalisation est employée pour tout type de système de communication y compris les communications optiques [12], sous-marines [13] et même en séismologie [14].

Des filtres adaptatifs sont mis en œuvre, en raison de la variété des canaux et de leur nature variante dans le temps. Ces techniques adaptatives permettent de concevoir des algorithmes estimant en permanence, les paramètres du filtre en minimisant un certain critère d'erreur. Les algorithmes les plus utilisés sont l'algorithme du gradient stochastique (LMS -Least Mean Square) et l'algorithme des moindres carrés récursifs (RLS - Récursive Least Squares) [15].

Pour un canal peu dispersif, les performances réalisables sont souvent adéquates. Mais, lorsque la distorsion du canal devient sévère, les performances de ces filtres deviennent sous optimales [2, 7, 16,17].

Les réseaux de neurones artificiels (ANN-Artificial Neural network) semblent la solution toute indiquée, avec leur non-linéarité intrinsèque qui réside dans leurs fonctions d'activations, ils sont aptes à résoudre des problèmes non linéaires complexes, ils sont d'ailleurs considérés comme des approximateurs universels [18].

Ces réseaux neuronaux permettent de réaliser une relation complexe entre les espaces d'entrée et de sortie. Beaucoup de recherches ont montré que les filtres égaliseurs basés sur les réseaux de neurones peuvent fournir des meilleures performances en comparaison avec les filtres conventionnels [16, 17, 19-21].

Trois architectures neuronales sont principalement investies en égalisation à savoir; le perceptron multicouches MLP [2, 16, 22-24], les fonctions de base radiale RBF [25-27] et les réseaux récurrents RNN [28-30].

Une revue des différentes architectures des filtres égaliseurs neuronaux, leurs méthodes d'apprentissage sont présentées dans [17]. Parmi toutes ces architectures, la plus communément et profondément étudiée est celle du perceptron multicouches (MLP), en raison de sa stabilité, sa structure flexible, sa paramétrisation finie et sa mise en œuvre simple [2, 16, 23]. Le réseau MLP incite de plus en plus l'intérêt de la communauté scientifique, des recherches récentes lui accordent des intérêts particuliers [31-34].

L'algorithme de rétropropagation du gradient (BP-back Propagation) est une méthode d'apprentissage supervisée pour le réseau MLP. Jusque à présent, cette simple méthode constitue un moyen d'apprentissage central des algorithmes, pour les réseaux neuronaux [35]. L'un des avantages les plus intéressants de l'algorithme de BP est son implémentation matérielle aisée. En dépit de son succès remarqué dans de nombreuses applications, son désavantage majeur et le tau d'apprentissage lent.

Des recherches ont été reportées dans la littérature pour améliorer la capacité d'apprentissage du MLP en proposant d'autres techniques autre que la BP, tel que l'algorithme de Gauss-Newton [22], des algorithmes se basant sur l'utilisation des critères non quadratiques [10] ou des algorithmes basés sur la linéarisation du réseau neuronal, tel que l'algorithme dit machine d'apprentissage extrême (ELM-Extreme Learning Machine) [36]. D'autres approches, inspirées de la nature telles que : l'optimisation par l'essaim de particule (PSO-Particle Swarm Optimisation) [37] et l'optimisation par la recherche dirigée (DSO-Directed Search Optimization) [38], des méthodes inspirées des lucioles [39] ou des sauts de grenouilles [40] ou d'autres basés sur les colonies des abeilles artificielles (ABC- Artificial Bee Colony) [41].

Toutes les améliorations apportées par ces techniques sont généralement au prix d'une complexité accrue [35].

D'autres recherches se sont penchées sur l'algorithme de BP lui-même afin d'améliorer son efficacité. A savoir, l'utilisation du moment (momentum) [42], l'adaptation du pas d'apprentissage avec le nombre d'itérations [43] ou l'utilisation d'un pas d'apprentissage à contrôle flou [44], ou encore, l'utilisation d'un pas et d'un moment adaptatifs [45,46].

Une modification du critère utilisé dans l'algorithme de BP a été aussi proposée dans [47]. L'amélioration de la performance de la BP en se basant sur la décorrélation des données d'entrée de la structure MLP a été développée dans [16]. De nombreuses améliorations des algorithmes basés sur la descente de gradient tel que l'algorithme de BP sont fournies dans [35, 48,49].

Des recherches se sont orientées vers l'exploitation des propriétés des fonctions d'activations en créant des non linéarités adaptatives dans celles-ci [31, 33, 50-56]. Ces travaux confirment que les algorithmes d'apprentissage qui adaptent les fonctions d'activations conduisent à des convergences plus rapides que ceux qui utilisent des fonctions d'activations fixes.

Pour un MLP avec une fonction d'activation fixe, la non-linéarité finale est déterminée uniquement par l'ajustement des poids des connexions. Par conséquent, la flexibilité qui en résulte, peut être insuffisante pour atteindre le comportement collectif non linéaire requis. Par contre l'utilisation d'une fonction d'activation adaptative régit les propriétés non linéaires de chaque neurone et améliore la non linéarité et la flexibilité du MLP et accroît largement ses propriétés de convergence [31, 33, 51, 53, 56]. Il a été pareillement établi qu'il existe une relation entre l'adaptation de la fonction d'activation et l'ajustement du tau d'apprentissage et des poids synaptiques [45, 51,56].

Parmi les recherches améliorant les capacités d'apprentissage du MLP à travers les fonctions d'activations, une classe de fonctions d'activations sigmoïdales à paramètre changeant d'une façon empirique non adaptable avec les paramètres du réseau a été proposée dans [55]. Une suggestion d'adapter ce paramètre pour éviter la saturation des poids a été réalisé selon le principe de la descente de gradient dans [31, 56].

Une autre approche a proposé une architecture neuronale flexible, non standard utilisant deux fonctions d'activation pour chaque neurone, l'une adaptée d'une façon supervisée, l'autre adaptée via un modèle paramétrique statistique, estimé par un apprentissage semi supervisé [33]. Une suggestion d'utiliser une fonction d'activation à deux paramètres

ajustables selon la descente de gradient dans la couche de sortie, et préserver une fonction fixe dans la couche cachée a été réalisé dans [50]. Des études ont proposé des méthodes hybrides combinant une stratégie de gradient ainsi que d'autres méthodes d'optimisation tel que la logique flou [51] et la SVM (Support Vector Machine) [53]. Une autre étude a proposé une fonction d'activation à plusieurs paramètres ajustables [52].

Les travaux présentés pour l'adaptation des fonctions d'activations, constituent des recherches en réseaux neuronaux et ne privilégie pas un domaine d'application spécifique, d'autres destinés aux problèmes d'approximation des fonctions ou de classification. Parmi ces recherches celles qui ont préservé la simplicité de la BP et celles qui ont proposé des approches, d'une complexité remarquable en comparaison avec la BP.

Dans ce contexte d'amélioration des capacités d'apprentissage de la BP nous avons proposé trois approches. La première approche consiste en l'amélioration de la performance de la BP en se basant sur la décorrélation des données d'entrée de la structure MLP à travers l'utilisation d'un filtre à structure treillis. L'utilisation de cette structure constitue une amélioration d'apprentissage pour le filtre égaliseur à base du MLP. Les algorithmes utilisés pour cette structure possèdent en général, une vitesse de convergence plus rapide. Cette approche est une modification de celle utilisée en [16].

La deuxième approche consiste à modifier l'algorithme de la rétropropagation du gradient, en créant une non linéarité adaptative dans sa fonction d'activation. L'algorithme est baptisé : rétro propagation du gradient avec une fonction d'activation adaptative (BPAAF-back propagation with adaptive activation function) [57, 58]. L'algorithme d'apprentissage BPAAF performe l'ajustement d'un seul paramètre, additionnement aux poids et aux biais du réseau MLP, conservant ainsi la similitude avec la BP conventionnelle.

La troisième approche consiste en une combinaison des deux premières approches, nous procédons à l'utilisation conjointe d'une fonction d'activation adaptative et d'un filtre de structure treillis, ce qui a pour effet d'améliorer considérablement, les performances des filtres égaliseurs. Nous utilisons l'algorithme (GAL- Gradient Adaptive Lattice) pour adapter le filtre de la structure treillis, au lieu de l'algorithme (LSL- Least Square Lattice) adopté par [16]. Nous utilisons l'algorithme de la BPAAF pour adapter la structure MLP au lieu de la BP adopté par [16].

La présente thèse est organisée en quatre chapitres aux quels nous donnons une brève description, en plus de l'introduction générale et la conclusion générale.

Le chapitre 1 présente le contexte générale de communications numériques telles que la modélisation des canaux de communications et un bref rappel sur les différents phénomènes perturbateurs, comme le bruit et les interférences inter-symboles. Il donne aussi un bref rappel concernant les mesures de performances dans les systèmes de communication tels que l'erreur quadratique moyenne et le taux d'erreur. En dernier lieu il présente un aperçu sur les filtres numériques.

Le chapitre 2 scrute d'abord la notion du filtrage adaptatif et les algorithmes d'optimisation correspondants, tels que les algorithmes LMS et RLS. Il présente quelques applications du filtrage adaptatif et accorde une attention particulière à l'application de ces filtres en égalisation. Il présente aussi, les différentes architectures des filtres égaliseurs, à savoir les filtres égaliseurs linéaires à structure transversale et à structure treillis. Les critères et les algorithmes d'optimisation et d'adaptation des coefficients de ces filtres sont également revus.

Le chapitre 3 donne un aperçu général sur les réseaux de neurones avec une focalisation principale sur le perceptron multicouche MLP (Multi Layer Perceptron) toutefois le réseau à fonction de base radiale RBF (Radial Basis Fonction) et le réseau récurrent sont brièvement discutés. L'utilisation du MLP comme filtre égaliseur est également scrutée. Une attention particulière et accordée à la fonction d'activation de type sigmoïdale et son rôle dans l'adaptation du MLP. Nos contributions se situent au niveau de ce chapitre où nous proposons d'améliorer la capacité d'apprentissage du MLP à travers l'utilisation d'une structure treillis au filtre à base du MLP, l'adaptation de la non linéarité matérialisée par la fonction d'activation et la combinaison des deux approches pour augmenter d'avantage les performances des filtres égaliseurs à base du MLP.

Le chapitre 4 s'intéresse aux simulations des différents filtres égaliseurs neuronaux et l'évaluation des approches proposées. Il décrit en premier lieu, les modèles des canaux de communication adoptés en simulations, leurs caractéristiques et leurs effets sur les signaux transmis. Il présente les performances des filtres égaliseurs à base du MLP à structures,

transverse et treillis, adaptés par les algorithmes de la BP et la BPAAF. Il évalue leurs performances à travers des exemples de simulation, sur des canaux, linéaire à phase non minimale, et non linéaires, en considérant des environnements, linéaire stationnaire et non linéaire non stationnaire. Les mesures de performances utilisées sont : la qualité du signal restauré, l'erreur quadratique moyenne (MSE-Mean square Error) en sortie des filtres égaliseurs et le taux d'erreur par bit (BER-Bit Error Rate).

Pour conclure, nous présentons une synthèse des travaux effectués, puis nous proposons quelques extensions possibles à notre travail, comme perspectives.

Chapitre 1

Notions sur les communications numériques

1.1 Introduction

Les communications numériques font partie des technologies modernes qui ont révolutionnées notre mode de vie. Les progrès en la matière sont spectaculaires, les dispositifs comme les téléphones portables ont dominé l'interaction sociale et d'affaire et sont devenus indispensables pour la plupart des personnes de la société moderne. Nous donnerons quelques notions de base dans le contexte des communication numériques, qui vont servir à la bonne compréhension de cette thèse. Nous introduisons brièvement la structure de base d'un système de communication numérique, ainsi que leurs différents constituants, nous nous intéressons plus particulièrement à la partie canal et sa modélisation. Les phénomènes perturbateur qui nuisent aux signaux utiles et les mesures de performances utilisées en communication numériques seront aussi présentées.

1.2 Structure d'un système de communication numérique

Les systèmes de communications modernes sont très complexes utilisant des milliers de composants. Pour obtenir une vue générale nous regardons le système d'un haut niveau d'abstraction comme présenté par le schéma synoptique de la figure 1.1.

La tâche d'un système de communication est d'acheminer l'information de la source vers le destinataire par l'intermédiaire d'un canal de transmission avec plus de fiabilité que possible. Les constituants de base d'un système de communication numérique sont décrits ci-dessous. D'une manière générale, un système de communication comprend un émetteur, un récepteur et un support ou milieu de transmission appelé canal.

L'émetteur est chargé d'adapter la source au milieu de transmission alors que le récepteur effectue l'opération inverse à partir du signal reçu. L'émetteur est généralement

constitué de trois fonctions principales, ayant leurs duales au récepteur. Nous allons décrire d'une façon succincte ces différents éléments [1,59, 60].

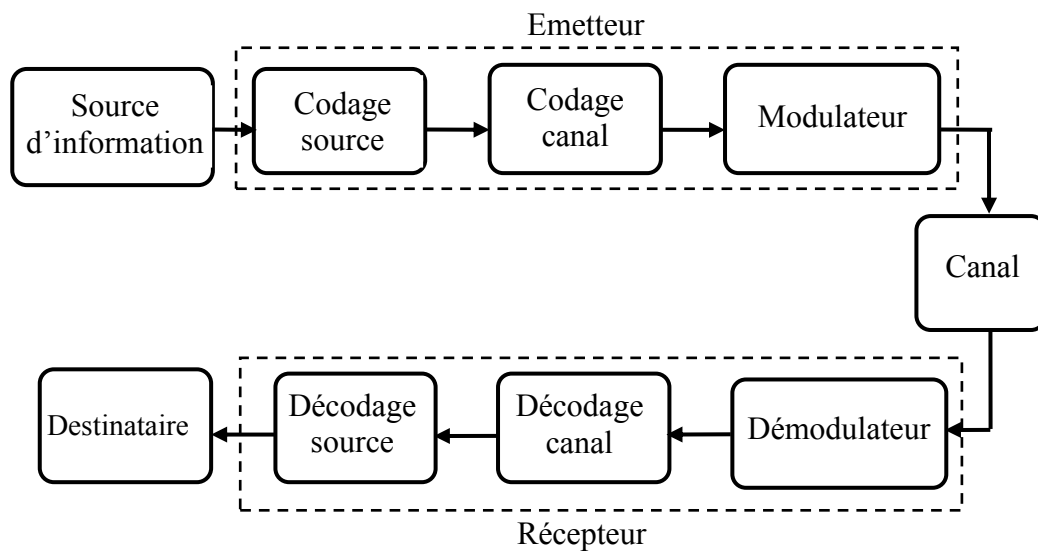


Figure 1.1 Schéma Synoptique d'un système de communication numérique

1.2.1 La source d'information

La source peut être de différentes natures telles que la voix, l'image, le texte, ou des données informatiques. La sortie d'une source n'est pas connue à priori par le destinataire elle est par nature aléatoire. Néanmoins certains messages déterministes, tels que les séquences pilotes, générées par l'émetteur pour servir au bon fonctionnement du récepteur.

La source peut être soit de forme analogique qu'on numérise ensuite (telle que la voix pour un téléphone mobile) soit directement de type numérique. Dans le cas d'une information analogique, elle est échantillonnée et numérisée à travers un étage de conversion analogique numérique. La taille du message binaire original ainsi produit est en général très importante et contient de la redondance.

La source délivre des symboles prenant leurs valeurs dans un alphabet fini. Dans le cas fréquent où cet alphabet est constitué uniquement de deux valeurs '0' et '1', le symbole est dit binaire. En groupant sous forme de blocs, n éléments binaires indépendants, on obtient un alphabet de M éléments ($M=2^n$), on parlera alors de symbole M -aire. Le débit est une caractéristique de la source, défini comme étant le nombre d'éléments binaires émis par unité de temps. Si une source délivre des symboles indépendants dont les valeurs sont

équiprobables, on dit que la source est idéale dans le sens où elle maximise le débit d'information.

1.2.2 L'émetteur

L'émetteur est chargé d'adapter la source au milieu de transmission, il est essentiellement composé des éléments suivants : Un codeur de source, un codeur de canal et un modulateur numérique permettant d'adapter le signal émis aux propriétés spectrales du canal.

1.2.2.1 Le codage source

Sa fonction principale est de contrôler la redondance naturelle éventuellement présente dans la séquence binaire et mettre ainsi le message sous forme réduite. Etant une opération de compression, le message est alors sous forme concise. Le codage source représente donc la source avec un minimum de bits sans en diminuer la quantité d'information, c'est à dire de délivrer une source aussi proche que possible d'une source idéale. Cette séquence binaire en sortie du codeur de source est appelée séquence d'information.

Notons que le code utilisé doit permettre d'effectuer l'opération inverse de décodage au récepteur. Le codage source peut aussi comporter une étape de cryptage dans le cas où l'on souhaite sécuriser le transfert des données et leur archivage.

1.2.2.2 Le codage canal

Il consiste à ajouter volontairement et d'une manière contrôlée de la redondance au signal afin de le protéger contre les différentes perturbations. On retrouve par exemple l'ajout de codes détecteurs ou correcteurs d'erreurs. Le codage de canal est réalisé uniquement en bande de base. Une fois le codage est réalisé le signal numérique est transformé en un signal physique capable de transiter sur le canal de transmission, ce qui constitue la tâche du modulateur.

1.2.2.3 Le modulateur

Etant donné que le milieu de transmission physique autorise uniquement la transmission de signaux analogiques, il est nécessaire de convertir l'information numérique en un signal analogique avant sa transmission effective à travers le milieu de transmission.

Le modulateur numérique permet alors d'interfacer la séquence binaire au canal de transmission. Il a pour rôle d'adapter le spectre du signal au milieu physique sur lequel il sera transmis. Etant donné qu'en pratique tous les canaux sont capables de transmettre des signaux électriques (une forme d'onde), il est donc nécessaire d'associer au message numérique une représentation physique, sous forme d'un signal électrique. Pour ce faire, il est possible de transmettre les bits ' n ' un à la fois ou en les regroupant pour transmettre un groupe de bits en même temps. Cette dernière méthode est connue sous le nom de modulation *M-aire* où M représente le nombre de groupes possibles ($M=2^n$). Le modulateur associe à chaque i ème groupe de ' n ' éléments binaires, un signal $S_i(t)$, $i=1, \dots, M$, choisi parmi $M=2^n$ signaux.

1.2.3 Le canal de transmission

Le canal est un élément important dans la chaîne de transmission, il constitue le support physique de l'information. Il peut être un support guidé (câble coaxial, paire torsadée ou fibre optique) ou l'espace libre utilisant la propagation d'une onde électromagnétique dans l'atmosphère. Ce milieu est généralement réservé aux transmissions par satellite ou par faisceaux hertziens. Une caractéristique commune des milieux de transmission est l'atténuation de la puissance du signal : plus la distance entre l'émetteur et le récepteur est grande, plus l'atténuation est élevée. Les variations d'atténuation en fonction de la distance des liaisons guidées sont généralement plus élevées que celles en espace libre¹ [61]. Le canal dont les caractéristiques exactes peuvent être inconnues de l'utilisateur, introduit généralement des distorsions linéaires ou non linéaires du signal utile et un délai de propagation dus aux perturbations aléatoires en provenance de phénomènes physiques indépendants du signal utile. Ces perturbations plus ou moins importantes sont créées par les conditions de propagation (imperfection des équipements, les propagations multi-trajets, présence de bruiteurs, affaiblissements...).

Ces dernières provoquent une dégradation du signal émis qui se traduit par l'apparition d'erreurs de transmission. Ces erreurs peuvent être très gênantes pour la restitution fidèle de l'information au destinataire.

¹ Les liaisons guidées provoquent des pertes exponentielles en fonction de la distance (d) alors que celles en espace libre, sont polynomiales (en d^2).

1.2.4 Le récepteur

Alors que l'émetteur adapte la source au milieu de transmission le récepteur effectue l'opération inverse. Les opérations duales sont effectuées au niveau du récepteur afin de retrouver le message d'origine. Il est essentiellement composé des éléments suivants : Un démodulateur qui ramène le spectre du signal de la bande passante vers la bande de base. Un décodeur de canal et un décodeur de source pour corriger les erreurs introduites par le canal.

1.2.4.1 Le démodulateur

Le démodulateur ramène le signal en bande de base. Il traite la forme d'onde transmise et la réduit en une séquence de nombres représentant les estimés des symboles transmis. La démodulation est généralement plus complexe car elle fait intervenir des opérations d'estimation et de détection. Elle peut être constituée, d'étages de récupération de porteuse et de rythme d'échantillonnage, d'estimation des distorsions subies par le signal lors de la transmission.

1.2.4.2 Le décodeur de canal

Connaissant le code utilisé par le codeur canal et la redondance contenue dans les données reçues, le décodeur canal exploite ces informations pour détecter et minimiser le taux d'erreur entre le message émis et le message décodé pour reconstituer la séquence transmise.

1.2.4.3 Le décodeur de source

Le décodeur connaissant la méthode de codage source recrée un signal numérique, qui puisse être compris par le destinataire. La différence entre le signal original et celui reconstruit est une mesure de la qualité d'un système de communication donné.

1.3 Modèles de canaux de communication

La modélisation du canal de communication peut être simple ou très complexe selon la nature du milieu de propagation, qui peut se comporter comme un filtre linéaire, ou présenter des non linéarités. Il peut être aussi stationnaire ou variable dans le temps.

1.3.1 Canal Gaussien

Le modèle mathématique le plus simple d'un canal de communication est le canal à bruit blanc Gaussien additif [1] généralement connu sous l'acronyme AWGN pour (Additif White Gaussian Noise) (figure 2.5 (a)). Le bruit gaussien blanc additif est un modèle de canal dans lequel la seule altération de la communication est une addition linéaire de bruit blanc avec une densité spectrale constante (exprimée en watts par hertz de bande passante) et une distribution Gaussienne d'amplitude. Le modèle ne tient pas compte de la sélectivité en fréquence, de l'interférence, de la non-linéarité ou de la dispersion.

Le signal transmis est perturbé par un bruit additif, qui est une perturbation aléatoire dont les origines sont le milieu de transmission (bruit externe tels que les rayonnements divers captés par l'antenne), ou les dispositifs électroniques utilisés (bruit interne) [61-63]. Le signal à la sortie du canal est donné par :

$$x(t) = s(t) + \eta(t) \quad (1.1)$$

$S(t)$ dénote le signal transmis et $\eta(t)$ désigne le bruit blanc Gaussien additif.

1.3.2 Canal linéaire

On considère souvent que le canal de communication n'ajoute qu'un effet de filtrage linéaire obéissant au principe de superposition c'est-à-dire, la sortie correspondant à une combinaison linéaire de signaux d'entrée est la même combinaison linéaire des sorties qui auraient été produites si chaque excitation se présentait seule à l'entrée.

Une telle situation peut se produire par exemple lorsqu'on est dans le cas de la propagation par trajets multiples.

La sortie d'un canal modélisé par un filtre linéaire et invariant dans le temps (LTI) pour (linear and time invariant) peut être décrite par une relation de convolution entre la réponse impulsionnelle $h(t)$ du canal et l'excitation présentée à son entrée :

$$x(t) = s(t) * h(t) \quad (1.2)$$

* dénote le produit de convolution.

L'équation (1.2) peut être exprimée par une intégrale de convolution:

$$x(t) = \int_{-\infty}^{+\infty} s(\tau) h(t - \tau) d\tau \quad (1.3)$$

A ce niveau il est bénéfique de parler de la réponse en fréquence du canal, ce qui conduit à une représentation équivalente de la convolution dans le domaine fréquentiel et ceci grâce au théorème de Plancherel.

D'une façon équivalente, (1.2) peut être décrite dans le domaine fréquentiel par :

$$X(f) = S(f)H(f) \quad (1.4)$$

Avec $X(f)$, $S(f)$ définissent les transformés de Fourier de $x(t)$ et $s(t)$ respectivement et $H(f)$ désigne la transformée de Fourier de $h(t)$ exprimée par :

$$H(f) = F(h(t)) = \int_{-\infty}^{+\infty} h(t)e^{-i2\pi ft} dt \quad (1.5)$$

Les canaux dont la réponse en fréquence $H(f)$ n'est pas uniformément plane à toutes les fréquences sont dits canaux sélectifs en fréquence. Dans le cas contraire ils sont nommés canaux non sélectifs en fréquence [61-63].

La plupart des canaux de communication sont caractérisés par une bande passante présentant des évanouissements d'amplitude et des non-linéarités de phase [1]. De ce fait, pour l'étude de la distorsion apportée par le canal, on peut caractériser chaque canal par sa réponse en fréquence $H(f)$ faisant apparaître son module et son argument comme suit:

$$H(f) = |H(f)|e^{j\theta(f)} \quad (1.6)$$

$|H(f)|$ étant la réponse en amplitude et $\theta(f)$ la réponse en phase du canal.

Un canal est dit idéal si la réponse en amplitude du canal est constante et la réponse en phase est linéaire pour toutes les fréquences. Ce canal ne provoque pas de distorsion de phase ou d'amplitude. Si les réponses d'amplitude et de phase ne sont pas constantes, la distorsion subie par le signal est dite d'amplitude ou de phase respectivement.

Pour les canaux bruités, linéaires invariants dans le temps, le modèle ci-dessous a été largement adopté [1,59]. Dans ce cas, le bruit est modélisé comme un processus aléatoire, indépendant du signal d'entrée et qui est ajouté à la sortie du canal. La sortie du canal bruité peut être écrite comme suit:

$$x(t) = s(t) * h(t) + \eta(t) \quad (1.7)$$

Ou par une intégrale de convolution :

$$x(t) = \int_{-\infty}^{+\infty} s(\tau)h(t-\tau)d\tau + \eta(t) \quad (1.8)$$

Un modèle typique du canal à filtre linéaire avec un bruit blanc gaussien additif est représenté à la figure 1.2.

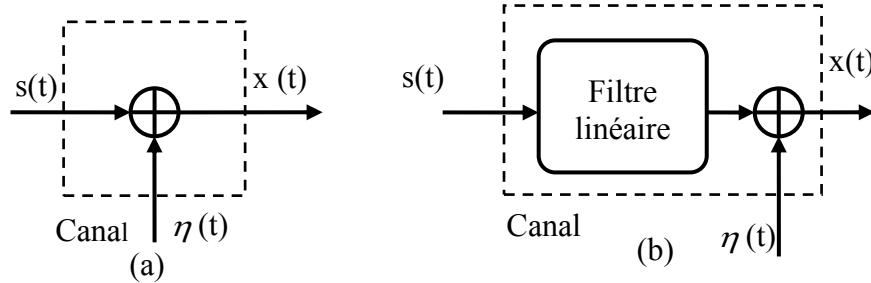


Figure 1.2 Modélisation du canal : (a) canal Gaussien (b) canal linéaire

Le filtre linéaire peut être un filtre discret linéaire de réponse impulsionnelle finie (RIF). Ce modèle discret utilisé pour décrire les effets de distorsion du canal est représenté par la fonction de transfert dans le domaine Z [1,59], définie par l'équation suivante :

$$H(z) = \sum_{i=0}^{L-1} h_i z^{-i} = h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots + h_{L-1} z^{L-1} \quad (1.9)$$

Ce modèle est illustré à la figure 1.3.

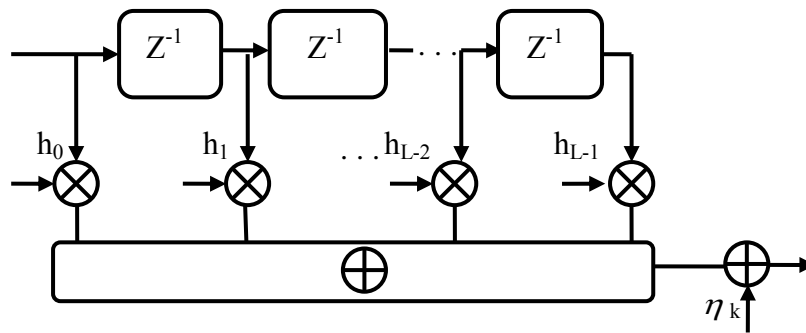


Figure 1.3 Canal discret linéaire

1.3.3 Canal non linéaire

Souvent on inclut les imperfections de l'émetteur et du récepteur dans la modélisation du canal ce qui peut entraîner une modélisation non linéaire. Dans ce cas, le canal ne peut plus être décrit par une réponse impulsionnelle.

La non linéarité est due essentiellement aux amplificateurs de puissance qui constituent les éléments clé d'un système embarqué (téléphones mobiles, satellites...), car ce sont les principaux consommateurs d'énergie. Les meilleurs rendements pour un

amplificateur de puissance sont obtenus en zone saturée, c'est-à-dire lorsque l'amplificateur est proche de son niveau de puissance maximum. Cependant dans de telles conditions, la non linéarité introduite par l'amplificateur dégrade l'information transportée par le signal [5,6].

Le canal non linéaire est essentiellement composé de la mise en cascade d'une partie linéaire qui reflète la réponse impulsionnelle du canal, généralement modélisée par un filtre linéaire, et d'une partie non-linéaire due aux autres composantes du canal de communication [16,7]. En plus, on aura un bruit additionnel inévitable s'additionnant au signal. Le signal à la sortie du canal est donné par :

$$x(t) = g[s(t) * h(t)] + \eta(t) \quad (1.10)$$

Où $g(\cdot)$ représente la fonction non linéaire

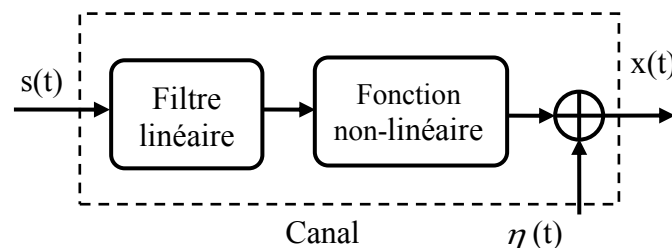


Figure 1.4 Modèle d'un canal non linéaire

1.4 Les effets perturbateurs

Au cours de son passage à travers les différentes parties d'un système de communication, le signal transmis est sujet à de nombreux phénomènes dont la plupart ont souvent des effets dégradants sur la qualité du signal. Ces effets peuvent être de nature interne au système ou externe, émanant du milieu de propagation. Ils engendrent des distorsions de nature linéaires ou non linéaires et l'addition de signaux non désirés qui se traduisent en pratique par des erreurs dans les messages reçus entraînant des pertes d'informations.

On cite dans ce qui suit certains effets perturbateurs souvent rencontrés dans les systèmes de communication.

1.4.1 Le bruit

L'une des perturbations les plus courantes qui est quasiment présente pour tout type de transmission est le bruit additif, il est de nature aléatoire et s'ajoute au signal utile. Le

bruit engendre une dégradation de l'amplitude du signal reçu ce qui augmente le risque d'erreur de détection de ce signal. Ce bruit modélise toutes les imperfections d'origine externes (rayonnement de dispositifs électriques, décharges électriques, rayonnement cosmique, autres émetteurs) ou d'origine internes (bruit thermique) [61-63]. Ce type de bruit est souvent modélisé par un processus stochastique blanc gaussien de moyenne nulle et de variance σ_b^2 .

Les caractéristiques du bruit blanc additif sont sa variance, sa fonction d'autocréation et sa densité spectrale de puissance PSD (Power Spectral Density) figure (1.5).

Soit x la réalisation de ce processus stochastique Gaussien à un instant de temps t_i , la fonction densité de probabilité de x ou PDF pour (PDF Probability Density Function) suit une distribution gaussienne, exprimée par [1,59] :

$$f_X(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(x-m_x)^2}{2\sigma_x^2}\right) \quad (1.11)$$

$m_x = E[x]$, la moyenne de la variable aléatoire x et $\sigma_b^2 = E(x^2) - (E(x))^2$, la variance.

Pour x variable aléatoire centrée, la PDF gaussienne s'écrit comme suit :

$$f_X(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \quad (1.12)$$

La densité spectrale de puissance du bruit blanc Gaussien est uniforme pour toute valeur de fréquence, elle est exprimée par [1,59]:

$$s_X(f) = \frac{1}{2} N_0, |f| < \infty \quad (1.13)$$

La fonction d'auto-corrélation du bruit blanc est liée à la densité spectrale de puissance par le théorème de Wiener-Kintchine [59], elle s'exprime par les relations suivantes:

$$R_X(\tau) = F^{-1}\{s_X(f)\} \quad (1.14)$$

$$R_X(\tau) = \int_{-\infty}^{+\infty} s_X(f) \exp(j2\pi f\tau) df \quad (1.15)$$

$$R_X(\tau) = \frac{1}{2} N_0 \delta(\tau) \quad (1.16)$$

Les statistiques du bruit blanc Gaussien sont montrées à figure (1.5), la partie (a) montre sa fonction densité de probabilité PDF qui est de forme Gaussienne, la partie (b) sa densité spectrale de puissance qui est uniforme pour toute fréquences et la partie (c) sa fonction d'auto-corrélation.

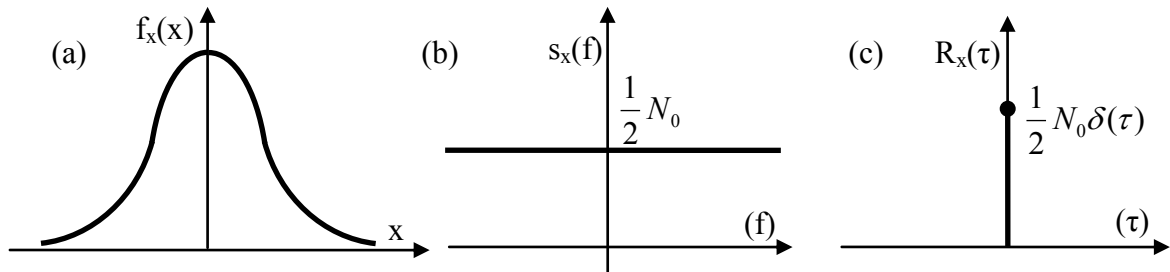


Figure 1.5 Les statistiques du bruit blanc Gaussien : (a) Fonction densité de probabilité (b) Densité spectrale de puissance (c) Fonction d'auto-corrélation

Néanmoins, le bruit n'est pas la seule source de perturbations, la fonction de transfert du canal introduit une distorsion au signal lors de sa propagation.

1.4.2 Propagation multi-trajets

Le phénomène de multi-trajets ou (multi-paths) est une caractéristique liée au canal de propagation radioélectrique mobil qui se manifeste lorsque le signal émis se propage suivant plusieurs chemins de l'émetteur au récepteur. Cette propagation multi-trajets est due aux obstacles rencontrés par le signal comme les réflexions spéculaires (propriétés d'un miroir) ou diffuses sur des obstacles tels que des montagnes ou des bâtiments, ainsi qu'à la position relative de l'émetteur et du récepteur s'ils sont en mouvement, ce qui cause une variation temporelle du canal. Cet environnement de propagation riche en échos et donc caractérisé provoque des atténuations de la puissance du signal [61-63].

1.4.3 Dispersion du canal

Lorsque le phénomène de dispersion existe dans le canal, les composantes du signal reçu arrivent suivant des chemins distincts, directs et indirects avec des délais différents et des caractéristiques en amplitude et en phase propres à chacun des trajets. Ceci est équivalent à transmettre le signal à travers un nombre de canaux séparés, chacun possédant une atténuation et un délai différent [61-63]. Au niveau du récepteur, le signal reçu est constitué de plusieurs répliques du signal transmis. Un signal correspondant au trajet

direct et des signaux ayant suivi des trajets différents. Une fois combinés, ces signaux multiples s'ajoutant d'une façon destructive, le signal reçu dans ce cas, sera très faible ou pratiquement nul. Ces répliques peuvent également s'additionner de façon constructive, le signal reçu est alors plus puissant que le signal du seul trajet direct [61-63].

L'étalement temporel, noté T_m appelé encore (delay spread) est défini par la différence entre le plus grand et le plus court des retards produits par le signal et ses répliques lors de la propagation dans le canal, il permet de caractériser par une seule variable la dispersion temporelle du canal. Dans ce cas, nous dirons que cet étalement temporel est la cause d'une dispersion dans le temps [61-63].

1.4.4 Sélectivité en fréquence

Soit B_c , la bande passante de cohérence du canal, qui correspond à la gamme de fréquences sur laquelle les amplitudes des composantes fréquentielles du signal, fortement corrélées, subissent des atténuations semblables. Elle représente un seuil en fréquence où au-delà de cette bande de fréquence, les distorsions du signal deviennent non négligeables ainsi, deux composantes fréquentielles du signal émis subiront des atténuations différentes [61-64].

En général, la bande de cohérence d'un canal est du même ordre de grandeur que l'inverse de son étalement temporel, soit $B_c = \frac{1}{T_m}$. Notons B_s la largeur de bande du signal transmis. Si cette bande est suffisamment étroite de telle façon que le signal n'est pas distordu, dans ce cas, il n'y a pas de sélection en fréquence [61-64]. Plus la largeur de bande B_s augmente, plus les distorsions prennent des valeurs plus élevées.

Cela veut dire que si $B_s \ll B_c$, toutes les composantes fréquentielles du signal subissent des atténuations semblables, et le canal est dit non sélectif en fréquence.

Dans le cas contraire c'est-à-dire $B_s \gg B_c$ aux moins deux composantes fréquentielles subissent des atténuations indépendantes, et le canal est dit sélectif en fréquence.

1.4.5 Les interférences inter-symboles

Connu sous l'acronyme (ISI) pour (inter-symboles interférence) résultant généralement de deux causes principales : de la bande limitée allouée au canal et de la présence des distorsions dues à la propagation multi trajet. L'interférence inter-symbole se produit dans les canaux pour tout type de modulation [1,4, 59].

Les caractéristiques temporelles du canal tendent à étaler le temps de transmission d'un symbole, augmentant le risque de chevauchement de plusieurs symboles adjacents et limitant le débit de transmission admissible sur ce canal.

La nature de la réponse du canal de transmission influence la facilité avec laquelle le récepteur détectera les signaux.

Soit une séquence de symboles s_j transmise à travers un canal discret équivalent de réponse impulsionnelle $h(t)$, le signal reçu $x(t)$ est une superposition de la réponse du canal pour chaque symbole émis plus le bruit gaussien additif $\eta(t)$.

$$x(t) = \sum_{n=0}^{\infty} s_n h(t - nT) + \eta(t) \quad (1.17)$$

Le signal reçu est ensuite échantillonné avec un taux de $1/T$ échantillons par seconde aux instants² $t=kT$, avec $k=0,1,\dots$; on obtient :

$$x(kT) = \sum_{n=0}^{\infty} s_n h(kT - nT) + \eta(kT) \quad (1.18)$$

La relation (1.18) s'écrit d'une façon équivalente comme suit :

$$x_k = \sum_{n=0}^{\infty} s_n h_{k-n} + \eta_k \quad (1.19)$$

$$x_k = h_0 \left(s_k + \frac{1}{h_0} \sum_{\substack{n=0 \\ n \neq k}}^{\infty} s_n h_{k-n} \right) + \eta_k \quad (1.20)$$

h_0 est un facteur d'échelle arbitraire qu'on pose égal à '1' pour convenance, on aura donc:

$$x_k = s_k + \sum_{\substack{n=0 \\ n \neq k}}^{\infty} s_n h_{k-n} + \eta_k \quad (1.21)$$

On voit clairement que l'échantillon reçu à un instant donné est la somme de la contribution du symbole transmis correspondant à l'instant d'échantillonnage (le terme s_k ou l'information désirée à l'instant d'échantillonnage k) et une contribution cumulative des symboles transmis aux autres instants (second terme). Cette contribution non désirée des symboles voisins constitue l'interférence inter symboles ISI qui est très commune dans les communications numériques [1,59, 65]. Le troisième terme est la contribution du bruit additif. L'ISI constitue une perturbation qui engendre des erreurs de détection au niveau du récepteur.

² [1] ajoute un délai aux instants d'échantillonnage ($t=kT+\tau_0$), avec $k=0,1,\dots$; Où τ_0 est le délai de transmission à travers le canal néanmoins, le raisonnement reste équivalent.

De l'équation (1.21), découle la condition d'interférence inter symboles (ISI) nulle que doit satisfaire la réponse impulsionnelle h_k , cette dernière doit être équivalente à une impulsion de Dirac discrète :

$$h_k = \begin{cases} 1 & \text{pour } k = 0 \\ 0 & \text{pour } k \neq 0 \end{cases} \quad (1.22)$$

Dans ce cas le signal reçu sans ISI s'écrit comme suit:

$$x_k = s_k + \eta_k \quad (1.23)$$

La condition exprimée par (1.22) se traduit dans le domaine fréquentiel par la condition nécessaire et suffisante sur $H(f)$ pour que $h(t)$ satisfait la condition ci-dessus:

$$H(f) = \frac{1}{T} \sum_{m=-\infty}^{\infty} H(f + \frac{m}{T}) = 1 \quad (1.24)$$

$H(f)$ est la transformée de Fourier de $h(t)$. L'équation (1.24) sera seulement satisfaite si la largeur de bande B de $H(f)$ est supérieure ou égale au taux symbole $\frac{1}{T}$ c'est-à-dire, $B \geq \frac{1}{T}$

L'équation (1.24) est connue comme étant la condition de Nyquist pour une interférence inter symboles (ISI) nulle [1,65].

1.5 Mesures de performances en communications numériques

Les performances d'une transmission numérique dépendent de la fidélité avec laquelle les éléments binaires du message transmis sont restitués au destinataire. Lorsque le signal émis est analogique, la première mesure de performance est le rapport signal sur bruit SNR. Cependant en communication numérique cette mesure de performance s'avère insuffisante en outre plusieurs mesures de performances peuvent être utilisées. Ces mesures de performance permettent d'évaluer et comparer les performances des algorithmes. Dans ce qui suit, on cite quelques critères de performances couramment utilisés en communications numériques.

1.5.1 Le rapport signal sur bruit

Le rapport signal sur bruit, plus connu sous l'acronyme SNR pour Signal to Noise Ratio, est la mesure de performance la plus répandue dans les systèmes de communication. Il permet de mesurer la contribution du bruit [1,60]

Celui-ci va donc permettre d'apprécier la qualité de transmission d'un signal. En effet, plus le rapport signal à bruit est faible, plus le signal est dégradé par le bruit et plus il sera difficile de supprimer l'influence du bruit sur le signal. Il est nécessaire de garantir un rapport signal à bruit important pour s'assurer que le signal reçu reste une copie fidèle du signal transmis. Cette mesure du niveau du bruit, évaluée par le rapport signal sur bruit est exprimée par:

$$SNR = \frac{E\{s(t)^2\}}{E\{n(t)^2\}} \quad (1.25)$$

Pour un signal et un bruit centrés de variances respectives σ_s^2 de variance σ_n^2 l'équation (1.25) s'écrit d'une façon équivalente par :

$$SNR = \frac{\sigma_s^2}{\sigma_n^2} \quad (1.26)$$

Le SNR est généralement exprimé en décibel (dB) comme suit:

$$SNR_{dB} = 10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_n^2} \right) \quad (1.27)$$

1.5.2 L'erreur quadratique moyenne

L'erreur quadratique moyenne, plus connue sous l'acronyme MSE pour (Mean Square Error) est utilisée comme un paramètre de mesure de performance. Elle permet de mesurer la moyenne du carré de l'erreur entre le signal émis et le signal estimé [1,59]. Si $s_k = [s_1, s_2 \dots s_M]^T$ et $\hat{s}_k = [\hat{s}_1, \hat{s}_2 \dots \hat{s}_M]^T$ sont les vecteur transmis et estimé respectivement, l'erreur est donnée par :

$$e(k) = s(k) - \hat{s}(k) \quad (1.28)$$

Evidemment, plus $e(k)$ sera faible, plus l'estimation sera bonne. Il est pratique de chercher à minimiser $e^2(k)$ car c'est une fonction quadratique facilement dérivable. L'erreur quadratique moyenne MSE est définie par :

$$MSE = E(s(k) - \hat{s}(k))^2 \quad (1.29)$$

$$MSE = E(e^2(k)) \quad (1.30)$$

Une estimation du MSE est donnée par :

$$MSE = \frac{1}{N} \sum_{k=0}^N (e(k))^2 \quad (1.31)$$

Où N représente le nombre de symboles utilisés.

D'une façon générale une conception à base d'une minimisation de l'erreur quadratique moyenne MSE conduit à de bons résultats [1].

1.5.3 La probabilité d'erreur

La probabilité d'erreur constitue l'un des paramètres de qualité en communication numérique. Ce critère est le plus adapté pour mesurer la performance des algorithmes, puisque le but est de transmettre le maximum d'information avec le minimum d'erreurs possibles. La probabilité d'erreur est définie par élément binaire, elle exprime la probabilité de prendre une décision erronée sur un élément binaire. Cette probabilité n'est jamais strictement nulle, mais cela ne signifie pas pour autant que la qualité de transmission est mauvaise; en effet il suffit qu'elle prenne une valeur suffisamment faible pour satisfaire un certain critère de fidélité [1, 59,60].

Pour déterminer l'expression de la probabilité d'erreur considérant le cas d'une transmission binaire antipodale sur un canal sans distorsion qui ajoute seulement un bruit blanc additif gaussien (AWGN)[59].

On caractérise la variable aléatoire y par une distribution Gaussienne de moyenne m_y et de variance σ_y^2 , exprimées par:

$$m_y = E[y] = \begin{cases} -A & \text{pour 0 transmis} \\ A & \text{pour 1 transmis} \end{cases} \quad (1.32)$$

$$\sigma_y^2 = \frac{N_0}{2T_b} \quad (1.33)$$

Avec $\frac{N_0}{2}$, densité spectrale de puissance du bruit et T_b durée symbole.

Les densités de probabilité conditionnelle de la variable aléatoire y (sachant que 0 ou 1 sont transmis), sont exprimées respectivement par :

$$f_Y(y/0) = \frac{1}{\sigma_y \sqrt{2\pi}} \exp\left(-\frac{(y+A)^2}{2\sigma_y^2}\right) \quad (1.34)$$

$$f_Y(y/1) = \frac{1}{\sigma_y \sqrt{2\pi}} \exp\left(-\frac{(y-A)^2}{2\sigma_y^2}\right) \quad (1.35)$$

L'échantillon y prélevé à l'instant T_b est comparé à un seuil au niveau du détecteur (figure 1.6 (a)) et une décision est prise selon la règle suivante :

Lorsque $y > s$ alors la sortie est égale à 1 et lorsque $y < s$ alors la sortie est égale à 0. Le récepteur peut commettre alors deux types d'erreurs ; prendre une décision égale à 1 alors que c'est un 0 qui est transmis, avec une probabilité conditionnelle de l'erreur de décision P_{e0} , ou prendre une décision égale à 0 alors que c'est un 1 qui est transmis, avec une probabilité P_{e1} .

$$P_{e0} = P\{y > s / 0\} \quad \text{et} \quad P_{e1} = P\{y < s / 1\}. \quad \text{La probabilité d'erreur } P_e \text{ est exprimée par :}$$

$$P_e = p_0 \cdot P_{e0} + p_1 \cdot P_{e1} \quad (1.36)$$

Sachant que p_0 et p_1 les probabilités associées aux éléments binaires 0 et 1 respectivement, qu'on suppose équiprobables avec :

$$p_0 + p_1 = 1 \quad \text{et} \quad p_0 = p_1 = \frac{1}{2} \quad (1.37)$$

$$P_e = p_0 \int_s^{+\infty} f_Y(y/0) dy + p_1 \int_{-\infty}^s f_Y(y/1) dy \quad (1.38)$$

Le seuil est choisi de façon à égaliser les deux types de probabilités d'erreurs ($P_{e0} = P_{e1}$) c'est-à-dire ($s = 0$). De ce fait et en vertu de (1.35) et (1.36), on en déduit que :

$$P_e = P_{e0} = P_{e1} \quad (1.39)$$

$$P_{e0} = \frac{1}{\sqrt{\pi N_0 / T_b}} \int_0^{+\infty} \exp\left(-\frac{(y + A)^2}{N_0 / T_b}\right) dy \quad (1.40)$$

Cette intégrale ne peut être évaluée car la primitive n'existe pas.

On peut cependant utiliser le calcul numérique et tabuler les valeurs obtenues. On applique un changement de variable pour obtenir une fonction normalisée (standard) ; on pose

$$z = \frac{y + A}{\sqrt{N_0 / T_b}}, \quad \text{on obtient alors :}$$

$$P_{e0} = \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \exp(-z^2) dz \quad (1.41)$$

Avec, E_b l'énergie par bit transmis défini par : $E_b = A^2 T_b$

A ce niveau on introduit la fonction d'erreur de Gauss complémentaire $erfc(x)$, telle que : $erfc(x) = 1 - erf(x)$. Les fonctions d'erreur et d'erreur complémentaire sont définies respectivement par :

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-u^2) du \quad (1.42)$$

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-u^2) du \quad (1.43)$$

De (1.38), (1.41) et (1.43) on obtient l'expression de la probabilité d'erreur en fonction de l'erreur complémentaire :

$$P_e = \frac{1}{2} erfc \sqrt{\frac{E_b}{N_0}} \quad (1.44)$$

La probabilité d'erreur P_e , ne dépend que du rapport E_b/N_0 .

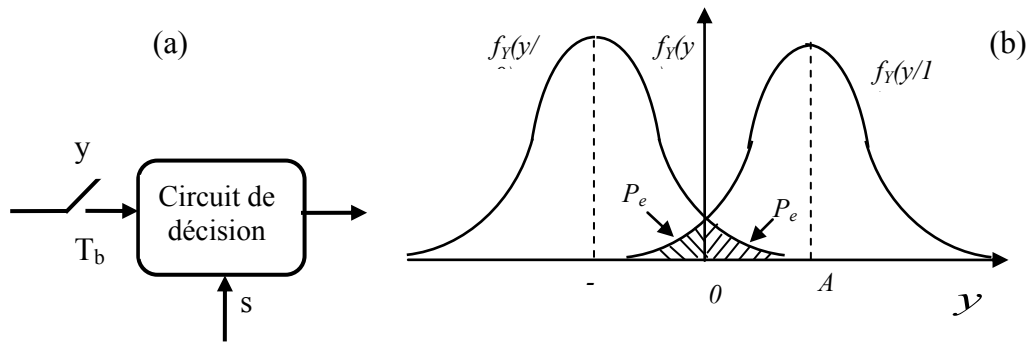


Figure 1.6 Circuit de décision (a) ; Probabilité d'erreur (b)

En pratique, il est toutefois presque impossible de concevoir des systèmes basés sur le principe de la probabilité d'erreur [59]. Néanmoins il est toujours important d'évaluer les performances en termes du taux d'erreur par bit.

1.5.4 Le taux d'erreur

Il est à noter que la probabilité d'erreur (P_e) est une valeur théorique dont une bonne estimation non biaisée au sens statistique est réalisée par la mesure du taux d'erreur par bit noté BER (Bite Error Rate). Ce dernier est le rapport entre le nombre d'éléments binaires erronés et le nombre d'éléments binaires transmis. On notera que le BER tend vers P_e si le nombre de bits transmis tend vers l'infini. Le taux d'erreur par bit est exprimé par :

$$BER = \frac{\text{nombre de 'eb' erronés}}{\text{nombre de 'eb' émis}} \quad (1.45)$$

eb : abréviation d'élément binaire, le nombre d'élément binaire erronés correspond au comptage de nombre d'élément non nuls, résultant de la différence entre l'élément binaire transmis x_k et ceux estimés.

$$BER = \frac{\text{nombre de } (x(k) - \hat{x}(k)) \neq \text{zero}}{\text{nombre 'eb' émis}} \quad (1.46)$$

Dans les cas des modulations M-aire où les bits sont groupés en symboles on définit le taux d'erreurs symboles (SER : Symbole Error Rate) qui consiste à comptabiliser le nombre de symboles différents entre x et $\hat{x}$. Notons que dans le cas des signaux binaire, le SER est assimilé au taux d'erreurs bits (BER).

1.6 Les filtres numériques

Etant donné que les filtres de restauration des signaux sont conçus essentiellement à partir des filtres numériques, nous allons présenter une description succincte de ces filtres. Un filtre numérique est un système linéaire invariant dans le temps (LIT). Ainsi, il est complètement déterminé par la connaissance de sa réponse impulsionnelle, qui joue un rôle de pondération des entrées passées et futures [66]. Un filtre numérique effectue la transformation linéaire suivante :

$$y(k) = f[x(k)] = \sum_i h(i)x(k-i) = h(k) * x(k) \quad (1.47)$$

La formule (1.47) représente l'équation de convolution qui caractérise le système linéaire invariant dans le temps (LIT), et $*$ dénote le produit de convolution. Un tel système est donc complètement défini par les coefficients de la réponse impulsionnelle $h(k)$. Alors que $x(k)$ et $y(k)$ sont respectivement, les séquences d'entrée et de sortie. Ce système possède les propriétés de causalité et de stabilité qui impliquent que $h(k) = 0$, pour $k < 0$ et

$\sum_k h(k) < \infty$ respectivement. Les caractéristiques des systèmes LIT sont étudiées à l'aide de

la transformation en Z . En vertu de cette transformée on retrouve une analogie avec relations de Plancherel pour la transformée de Fourier. La relation (1.47) s'écrit donc comme suit :

$$Y(z) = H(z)X(z) \quad (1.48)$$

Avec $H(z)$, $X(z)$ et $Y(z)$ les transformée en Z de $h(k)$, $x(k)$ et $y(k)$ respectivement.

On distingue deux types de filtres numériques suivant le support de la réponse impulsionnelle [66].

1.6.1 Les filtres à réponse impulsionnelle finie

La réponse impulsionnelle d'un tel filtre est nulle en dehors d'un certain intervalle de longueur T ou comportant $N-1$ échantillons. Dans ce cas l'équation déterminant la sortie du filtre est de la forme :

$$y(k) = \sum_{i=0}^{N-1} h(i)x(k-i) \quad (1.49)$$

La fonction de transfert en Z du filtre s'écrit :

$$H(Z) = \frac{Y(z)}{X(z)} = \sum_{i=0}^{N-1} h(i)Z^{-i} \quad (1.50)$$

La structure du filtre RIF est donnée à la figure 1.7

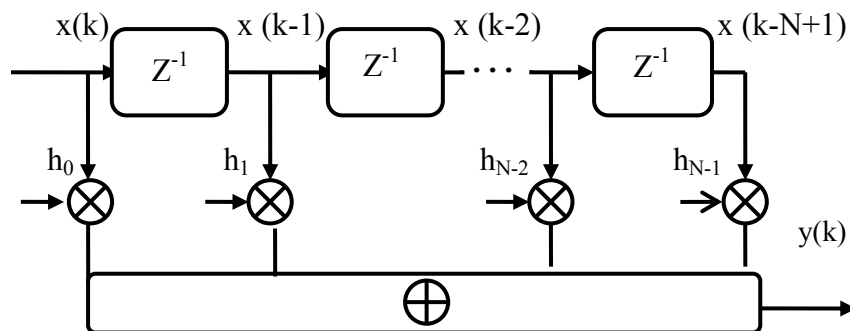


Figure1.7 Structure d'un filtre à réponse impulsionnelle finie (RIF)

Un filtre RIF est stable et causal et donc physiquement réalisable. Pour cette raison les filtres RIF sont très souvent utilisés.

1.6.2 Les filtres à réponse impulsionnelle infinie

Les filtres à réponse impulsionnelle infinie (RII) conservent une trace des échantillons qui leurs ont été appliqués pendant une durée infinie, ils sont de ce fait à mémoire infinie. Une telle mémoire est réalisée en utilisant une boucle de rétroaction de la sortie sur l'entrée d'où la dénomination de filtre récursif. La sortie $y(k)$ du filtre RII est une

combinaison linéaire des échantillons $x(k)$ présents à son entrée et de ceux précédemment déterminées en sa sortie.

$$y(k) = \sum_{l=0}^L a(l)x(k-l) - \sum_{m=1}^M b(m)y(k-m) \quad (1.51)$$

La fonction de transfert en Z du filtre s'écrit :

$$H(z) = \frac{\sum_{l=0}^L a(l)z^{-l}}{1 + \sum_{m=1}^M b(m)z^{-m}} \quad (1.52)$$

Ou d'une façon équivalente pour faire apparaître les pôles et les zéros de $H(z)$:

$$H(z) = A_0 \frac{\prod_{l=0}^L (1 - z_l z^{-1})}{\prod_{m=1}^M (1 - p_m z^{-1})} \quad (1.53)$$

z_l et p_m : Dénotent les zéros et les pôles de $H(z)$ respectivement, A_0 est le facteur d'échelle qui définit le gain.

Un filtre RII est physiquement réalisable, s'il est causal et stable (c'est-à-dire que les pôles de sa fonction de transfert en Z sont tous à l'intérieur du cercle unité).

Notons qu'un filtre dont tous les pôles et les zéros sont à l'intérieur du cercle unité est dit à phase minimale, tandis qu'un filtre dont tous les pôles sont à l'intérieur du cercle unité et des zéros à l'extérieur de ce cercle est dit à phase non minimale.

Le filtre RII permet d'obtenir en général des fonctions de filtrage beaucoup plus sélectives que celles des filtres RIF à quantité de calculs équivalente. Cependant, la boucle de réaction complique l'étude des propriétés et la conception de ces filtres et amène des phénomènes parasites.

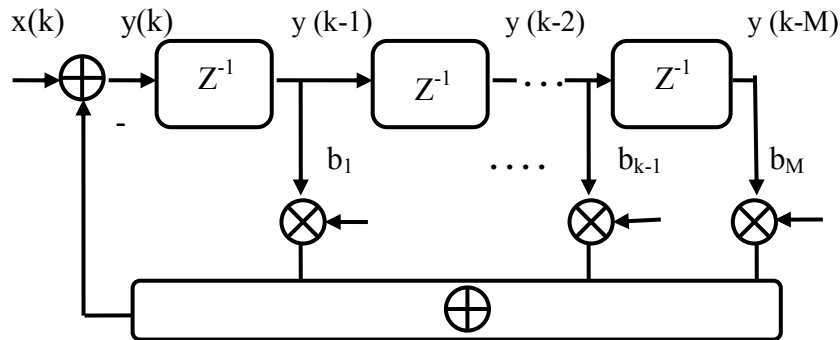


Figure 1.8 Architecture d'un filtre récursif

Le filtre RII est réalisé par la mise en cascade d'un filtre purement récursif et d'un filtre RIF [66]. On donne à la figure 1.11, la représentation du filtre purement récursif dont la fonction de transfert est celle donnée par l'équation (1.52), le numérateur réduit à l'unité.

1.6.3 Choix entre un filtre RIF et un filtre RII

Le choix entre les deux structures des filtres RII et RIF se pose fréquemment aux concepteurs de systèmes. Les RIF sont toujours stables et facile à réaliser, les filtres RII sont plus délicats à mettre en œuvre. Cependant ils permettent de réaliser des fonctions de filtrage beaucoup plus sélectives que celles des RIF à quantité de calcul équivalente [66].

1.7 Conclusion

Dans ce chapitre nous avons commencé par présenter un aperçu général sur les systèmes de communication numérique. Nous avons donné un intérêt particulier au canal de transmission et à sa modélisation. Parmi les modèles du canal étudiés, le modèle non linéaire englobe tout les effets rencontrés à savoir l'ISI, le bruit et la non linéarité. Nous avons également étudié les types de perturbation introduites par les canaux de communication et qui altèrent le signal transmis, en particulier le phénomène d'interférence inter symbole ISI. Nous avons également présenté un certain nombre de modulations numériques ainsi que quelques mesures de performances, qui sont utilisées individuellement ou conjointement pour évaluer les qualités d'un système. Les filtres adaptatifs sont conçus à partir des filtres numériques dont nous avons donné un aperçu à la fin de ce chapitre. Le choix d'un type de filtre relativement à l'autre est basé sur un compromis entre qualité souhaitée et complexité.

Chapitre 2

Filtrage adaptatif et algorithmes d'estimations

2.1 Introduction

Le filtrage adaptatif constitue un outil puissant en traitement du signal et en communications numériques. Il modélise, d'une manière interactive, la relation entre un signal se présentant à son entrée et un autre signal en sa sortie. le filtrage adaptatif comporte une mise à jour récursive des paramètres (coefficients) du filtre, pour s'adapter à l'environnement.

Nous allons présenter les différents algorithmes d'estimation des coefficients de ces filtres qui ont suscité un intérêt majeur dans ce domaine. Nous présenterons aussi l'application du filtrage adaptatif à la restauration des signaux altérés par les imperfections des canaux de communication numériques. Les filtres adaptatifs dans ce contexte sont des filtres égaliseurs, nous présenterons les différentes structures de ces filtres ainsi que leurs critères d'optimisation.

2.2 Les filtres adaptatifs

Un filtre adaptatif est considéré comme un filtre numérique dont les coefficients sont auto-ajustables, évoluant en fonction des signaux reçus. Ces coefficients sont estimés par des algorithmes récursifs, au sens d'un certain critère. Ce critère, également appelé fonction de coût, est une mesure de distance entre le signal de référence ou désiré et la sortie du filtre adaptatif [67].

La configuration de base d'un filtre adaptatif est illustrée dans la figure 2.1. Dans un tel schéma, le signal d'entrée est noté par $x(k)$, le signal de référence $d(k)$ représente le signal désiré, $y(k)$ est la sortie du filtre adaptatif. Le signal d'erreur $e(k)$ est utilisé par

l'algorithme d'adaptation pour mettre à jour les coefficients du filtre adaptatif en fonction de certains critères de performance.

En général, l'ensemble du processus d'adaptation vise à minimiser une certaine métrique du signal d'erreur, forçant le signal de sortie à approximer le signal de référence [68].

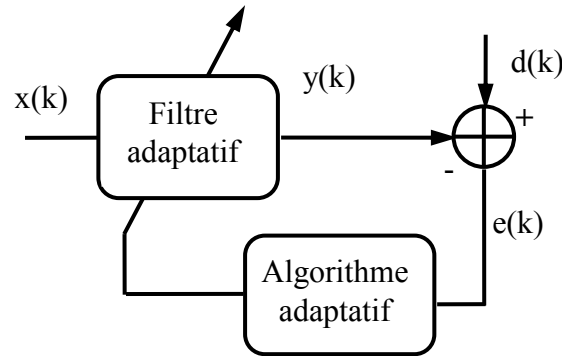


Figure 2.1 Configuration générale d'un filtre adaptatif

On constate que le signal de référence, sert de guide général pour l'ensemble du processus d'adaptation. Les algorithmes de filtrage adaptatif, constituent le mécanisme d'ajustement des coefficients du filtre, et sont en effet étroitement liés aux techniques classiques d'optimisation. Par ailleurs, un filtre adaptatif, en raison de sa caractéristique d'auto-ajustement, est parfois appelé à suivre des variations lentes de l'environnement [67]. Cette non-stationnarité, ou variation lente au cours du temps des propriétés statistiques, justifie le besoin de l'adaptatif [69].

2.3 Critères d'estimations

L'estimation optimale se présente selon différents critères tels que la moyenne quadratique, le maximum de vraisemblance (MV) ou bien encore l'estimation bayésienne avec le critère du maximum à postériori (MAP).

Bien que l'estimation en moyenne quadratique soit de très loin la technique la plus employée.

2.3.1 Critère du maximum de vraisemblance

De point de vue de l'estimation paramétrique, une observation aléatoire vectorielle Y possède une loi de probabilité qui dépend d'un vecteur de paramètres à estimer θ . Dans le contexte signal, θ représente les paramètres du modèle de filtrage qui relie l'observation Y au signal de référence.

La vraisemblance de l'observation Y est alors : $L(\theta) = P(y | \theta)$, $\theta \in \Theta \subset \mathbb{R}$.

On choisit comme estimateur de θ , la valeur qui maximise la vraisemblance des observations tel que [3, 61, 63, 69]:

$$\hat{\theta}_{MV} = \text{Arg max} \{L(\theta), \theta \in \Theta\} \quad (2.1)$$

Ce critère, qui ne privilégie a priori aucune des valeurs de θ , est naturellement adapté aux paramètres déterministes mais aussi à l'estimation des paramètres aléatoires, lorsque la distribution a priori est non informative, en quelque sorte uniforme sur un intervalle borné. L'estimateur du maximum de vraisemblance, s'il existe, est asymptotiquement efficace [3, 69].

2.3.2 Critère du maximum à postérieur MAP

Ce critère dérive de l'approche bayésienne qui consiste à introduire une information a priori sur les valeurs possibles du paramètre θ , en le considérant lui-même comme aléatoire de loi de probabilité a priori connue $P(\theta)$ [3, 61, 69].

On peut alors choisir l'estimateur $\hat{\theta}(Y)$ qui maximise la probabilité a posteriori de l'observation Y :

$$\hat{\theta}_{MAP} = \text{Arg max} \{P(\theta | y), \theta \in \Theta\} = \text{Arg max} \{P(\theta)P(y | \theta), \theta \in \Theta\} \quad (2.2)$$

En communication numérique ce critère du maximum à posteriori (MAP) est fortement relié à la probabilité d'erreur par bit, Il nécessite la connaissance de la densité de probabilité du bruit du canal [3].

Le critère de l'estimation en moyenne quadratique sera étudié ultérieurement lorsque on considérera le filtre adaptatif de structure RIF.

2.4 Structures des filtres adaptatifs

Les filtres adaptatifs peuvent être à réponse impulsionnelle finie (RIF) ou infinie (RII), la structure est soit transverse, soit en treillis. Si le filtre adaptatif est implémenté par un filtre RII, la fonction objective est une fonction non-quadratique par rapport aux paramètres du filtre, ce qui rend le problème d'optimisation plus difficile [67]. Les minimas locaux sont susceptibles d'exister, rendant certaines solutions obtenues par des algorithmes basés sur la méthode du gradient inacceptables. Malgré ces inconvénients, les filtres RII adaptatifs sont nécessaires dans un certain nombre d'applications où l'ordre d'un filtre RIF convenable est trop élevé [67].

En raison de sa simplicité et de son efficacité, la structure adaptative la plus employée est la structure transverse (ou ligne à retard) associée aux filtres RIF [67, 68]. Les filtres en treillis sont aussi utilisés en raison leur structure modulaire plus adaptées en implémentation, et de leur moindre sensibilité aux effets de quantification et de représentation en précision finie [69].

Donnons, la relation d'entrée-sortie du filtre adaptatif RIF transverse:

$$y(k) = w_0 x(k) + w_1 x(k-1) + \dots + w_{N-1} x(k-(N-1)) \quad (2.3)$$

$$y_k = \sum_{i=0}^{N-1} w_i x(k-i) \quad (2.4)$$

$$y_k = \mathbf{w}^T \mathbf{x}(k) \quad (2.5)$$

Où N est l'ordre du filtre, $\mathbf{x}(k) = [x(k) \ x(k-1) \ \dots \ x(k-(N-1))]^T$ et $\mathbf{w} = [w_0 \ w_1 \ \dots \ w_{N-1}]^T$ sont respectivement des vecteurs composés des échantillons du signal d'entrée et des coefficients du filtre.

2.5 Erreur quadratique moyenne

L'estimation des coefficients du filtre suit une procédure de minimisation d'une fonction objective ou coût. Cette fonction est généralement définie comme une norme du signal d'erreur $e(k)$. Les critères d'estimation les plus couramment employées en filtrage adaptatif sont : l'erreur quadratique moyenne connue sous son acronyme MSE pour (Mean Square Error), l'erreur quadratique instantanée ou ISE pour (Instantaneous Square Error) et les moindres carrés pondérés ou WLS pour (Weighted Least Square) [67,68]. Dans ce paragraphe on se focalisera sur le critère de l'erreur quadratique moyenne, définie par:

$$\xi(k) = E[e^2(k)] = E[d(k) - y(k)]^2 \quad (2.6)$$

En écrivant le signal de sortie $y(k)$ tel que donné dans l'équation (2.3), on obtient

$$\begin{aligned} \xi(k) &= E[d(k) - \mathbf{w}^T \mathbf{x}(k)]^2 \\ &= E[d^2(k)] - 2E[d(k)\mathbf{w}^T \mathbf{x}(k)] + E[\mathbf{w}^T \mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}] \end{aligned} \quad (2.7)$$

Ou encore:

$$\begin{aligned} \xi(k) &= E[d^2(k)] - 2\mathbf{w}^T E[d(k)\mathbf{x}(k)] + \mathbf{w}^T E[\mathbf{x}(k)\mathbf{x}^T(k)]\mathbf{w} \\ &= E[d^2(k)] - 2\mathbf{w}^T \mathbf{P} + \mathbf{w}^T \mathbf{R}\mathbf{w} \end{aligned} \quad (2.8)$$

Où R représentent la matrice d'auto-corrélation ($N \times N$) du signal d'entrée alors que P et le vecteur d'inter-corrélation ($N \times 1$) entre le signal de référence et le signal d'entrée. R et P sont définis respectivement par :

$$R = E[x(k)x^T(k)] \quad (2.9)$$

$$P = E[d(k)x(k)] \quad (2.10)$$

Ces notations supposent la stationnarité des signaux d'entrée et de référence [67,68]. A partir de l'équation (2.8), le vecteur gradient de la fonction MSE par rapport au vecteur des coefficients du filtre adaptatif est donné par:

$$\frac{\partial \xi}{\partial \mathbf{w}} = \left(\frac{\partial \xi}{\partial w_0}, \dots, \frac{\partial \xi}{\partial w_{N-1}} \right) = -2P + 2\mathbf{w}R \quad (2.11)$$

En égalisant le vecteur de gradient dans l'équation (2.11) à zéro et en supposant que R est non singulière, on obtient la solution de Wiener¹ donnant les valeurs optimales des coefficients minimisent la fonction de coût MSE:

$$\mathbf{w}_o = R^{-1}P \quad (2.12)$$

La solution de Wiener requiert la connaissance de la matrice d'auto-corrélation R et d'inter-corrélation P . Ceci nécessite la connaissance complète des statistiques de second ordre des signaux $x(k)$ et $d(k)$, ce qui rend l'équation (2.12) inadaptée pour la plupart des applications pratiques. Naturellement, la solution de Wiener peut être approchée par une estimation correcte de R et P sur la base d'intervalles de temps suffisamment longs, en supposant également l'ergodicité des signaux considérés [67, 69]. Dans ce cas, on peut utiliser des moyennes temporelles pour estimer R et P , ce qui est implicitement réalisé par la plupart des algorithmes adaptatifs [67].

En remplaçant la solution optimale de l'équation (2.12) dans l'expression MSE, on peut calculer l'erreur quadratique minimale (MMSE) pour (Minimum Mean Square Error) fournie par la solution de Wiener:

$$\begin{aligned} \xi_{\min} = MSE_{\min} &= E[d^2(k)] - 2\mathbf{w}^T E[d(k)x(k)] + \mathbf{w}^T E[x(k)x^T(k)]\mathbf{w} \\ &= E[d^2(k)] - 2\mathbf{w}_0^T P + \mathbf{w}_0^T R R^{-1} P \\ &= E[d^2(k)] - \mathbf{w}_0^T P \end{aligned} \quad (2.13)$$

¹ Norbert Wiener, fondateur du filtrage optimal.

2.6 Les algorithmes d'estimation

2.6.1 L'algorithme des moindres carrés moyens

La solution de Wiener nécessite l'inversion de la matrice R , ce qui constitue une charge de calcul. On peut alors estimer la solution de Wiener, d'une manière plus efficace, en ajustant itérativement le vecteur coefficient w à chaque instant k , de telle sorte que la séquence résultante $w(k)$ converge vers la solution désirée w_0 , éventuellement après un nombre d'itérations suffisamment petits.

L'algorithme à base de la descente de gradient pour atteindre le minimum de la fonction de l'erreur dans la direction opposée du vecteur de gradient associé est connu sous le nom de l'algorithme des moindres carrés moyens ou LMS² pour (Least Mean Square) dont l'équation de mise à jour est donnée par :

$$w(k) = w(k-1) - \frac{\mu}{2} \frac{\partial \xi(k)}{\partial w} \quad (2.14)$$

L'erreur quadratique moyenne MSE est une fonction de coût qui requiert la connaissance de la fonction d'erreur $e(k)$ en tout temps k . A cette fin, l'erreur quadratique moyenne MSE ne peut être déterminée précisément, elle est généralement approchée par d'autres fonctions de coûts. La forme la plus simple pour estimer la fonction MSE est l'erreur quadratique instantanée donnée par:

$$\hat{\xi}(k) = e^2(k) \quad (2.15)$$

Dans ce cas, le vecteur de gradient associé par rapport au vecteur de coefficient est déterminé comme suit :

$$\begin{aligned} \frac{\partial \hat{\xi}}{\partial w} &= 2e(k) \frac{\partial e(k)}{\partial w} \\ &= 2e(k) \frac{\partial}{\partial w} [d(k) - w^T x(k)] \\ &= -2e(k)x(k) \end{aligned} \quad (2.16)$$

Dans ce cas l'équation (2.14) s'écrit :

$$\begin{aligned} w(k) &= w(k-1) - \frac{\mu}{2} \frac{\partial \hat{\xi}(k)}{\partial w} \\ &= w(k-1) + \mu e(k)x(k) \end{aligned} \quad (2.17)$$

² Cet algorithme a été proposé par Widrow et Hoff

Où μ est le pas de convergence, ce paramètre joue un rôle important dans les caractéristiques de convergence ainsi que la condition de stabilité de l'algorithme.

Pour garantir la convergence des coefficients, μ doit être choisi dans la plage $0 < \mu < \frac{2}{\lambda_{\max}}$,

avec $\lambda_{\max}$, la plus grande valeur propre de la matrice d'auto corrélation du signal d'entrée [67, 68]. Cependant, calculer les valeurs propres de la matrice de corrélation R est aussi compliqué que d'inverser cette dernière. Une mesure de compromis consiste à prendre

$\mu = \frac{2}{\text{Trace}(R)}$ ou pour des raisons pratiques choisir μ comme suit :

$$0 < \mu < \frac{2}{N\sigma_x^2} \quad (2.18)$$

Avec, N le nombre des coefficients du filtre et σ_x^2 , la variance du signal d'entrée.

L'algorithme LMS³ est très populaire et largement utilisé dans le domaine du filtrage adaptatif en raison de son extrême simplicité, sa faible complexité de calcul, sa convergence évidente dans un environnement stationnaire, sa convergence non biaisée vers la solution de Wiener et un comportement stable lors de son implémentation. Cependant la convergence du LMS dépend de la dispersion des valeurs propres de la matrice d'auto-corrélation du signal d'entrée [67].

L'algorithme LMS est résumé au tableau 2.1.

LMS
Initialisation de μ $\mathbf{w}(0) = [0 \dots 0]^T$ Faire pour chaque k { $e(k) = d(k) - \mathbf{w}^T(k)\mathbf{x}(k)$; $\mathbf{w}(k) = \mathbf{w}(k-1) + \mu e(k)\mathbf{x}(k)$; } }

Tableau 2.1 L'algorithme LMS

³ L'appellation se réfère au but de l'algorithme (minimiser l'écart quadratique moyen) et pas à sa nature. Néanmoins l'appellation d'algorithme de gradient stochastique (GS) explicite la double origine de l'algorithme : méthode de gradient et approximation stochastique [69].

2.6.2 L'algorithme des moindres carrés récursifs

Cet algorithme connu sous la nomination RLS pour (*Recursive Least Squares*) dont le critère statistique de l'erreur quadratique moyenne MSE est remplacé par un critère de moindres carrés pondérés dit WLS pour (*Weighted Least Squares WLS*): qui est un critère déterministe minimisant la somme pondérée des erreurs quadratiques.

Ce critère est exprimé par :

$$\begin{aligned}\xi_d(k) &= \sum_{i=0}^k \lambda^{k-i} \varepsilon^2(i) \\ &= \sum_{i=0}^k \lambda^{k-i} [d(i) - x^T(i)w(k)]^2\end{aligned}\quad (2.19)$$

Le paramètre λ est un facteur de pondération exponentiel qui doit être choisi dans la plage $0 < \lambda < 1$ il est également appelé facteur d'oubli puisque l'information du passé lointain a un effet de plus en plus négligeable sur la mise à jour des coefficients du filtre. Le paramètre λ^{k-i} met en évidence les échantillons d'erreurs les plus récents (où $i \approx k$). Lorsque $\lambda = 1$, ξ_d est simplement la somme des erreurs quadratiques.

Si $\lambda < 1$, les erreurs passées sont pondérées avec un poids (facteur d'oubli) qui décroît exponentiellement. Asymptotiquement, si le signal est ergodique, les coefficients du filtre tendent vers la solution de Wiener [67, 69].

Puisque le critère (2.19) est basé sur plusieurs échantillons d'erreur, sa nature stochastique se réduit dans le temps, quand k augmente [68].

$\varepsilon(k)$ est l'erreur a posteriori qui consiste en la différence entre le signal désiré et la sortie du filtre, en utilisant les coefficients $w(k)$ les plus récents. En différenciant $\xi_d(k)$ par rapport à $w(k)$, il s'ensuit que :

$$\frac{\partial \xi_d(k)}{\partial w(k)} = -2 \sum_{i=0}^k \lambda^{k-i} x(i) [d(i) - x^T(i)w(k)] \quad (2.20)$$

En égalisant le gradient de l'erreur à zéro, on peut trouver le vecteur optimal $w(k)$:

$$\begin{aligned}w(k) &= \left[\sum_{i=0}^k \lambda^{k-i} x(i)x^T(i) \right]^{-1} \sum_{i=0}^k \lambda^{k-i} x(i)d(i) \\ &= R_d^{-1}(k)P_d(k)\end{aligned}\quad (2.21)$$

Où $R_d(k) = \sum_{i=0}^k \lambda^{k-i} x(i)x^T(i)$ et $P_d(k) = \sum_{i=0}^k \lambda^{k-i} x(i)d(i)$ sont appelées respectivement : la matrice d'auto-corrélation déterministe du signal d'entrée et le vecteur d'inter-corrélation

croisée déterministe entre le signal d'entrée et le signal désiré. La matrice $R_d(k)$ est supposé non singulière. L'équation (2.21) est réécrite sous la forme suivante :

$$\left[\sum_{i=0}^k \lambda^{k-i} x(i)x^T(i) \right] w(k) = \lambda \left[\sum_{i=0}^{k-1} \lambda^{k-i-1} x(i)d(i) \right] + x(k)d(k) \quad (2.22)$$

En considérant que $R_d(k-1)w(k-1) = P_d(k-1)$, il s'ensuit que :

$$\begin{aligned} \left[\sum_{i=0}^k \lambda^{k-i} x(i)x^T(i) \right] w(k) &= \lambda P_d(k-1) + x(k)d(k) \\ &= \lambda R_d(k-1)w(k-1) + x(k)d(k) \\ &= \left[\sum_{i=0}^k \lambda^{k-i} x(i)x^T(i) - x(k)x^T(k) \right] w(k-1) + x(k)d(k) \end{aligned} \quad (2.23)$$

Dans la dernière égalité, la matrice $x(k)x^T(k)$ a été ajoutée et soustraite du côté droit de l'équation (2.23). A ce niveau, on définit l'erreur a priori comme suit :

$$e(k) = [d(k) - x^T(k)w(k-1)] \quad (2.24)$$

En exprimant $d(k)$ en fonction de l'erreur a priori et en remplaçant le résultat dans l'équation (2.23), on obtient l'équation de mise à jour des coefficients du filtre adaptatif:

$$w(k) = w(k-1) + e(k)R_d^{-1}(k)x(k) \quad (2.25)$$

Dans cette expression, la charge de calcul pour déterminer $R_d^{-1}(k)$ peut être réduite de façon significative en utilisant le lemme l'inversion matricielle⁴, qui, dans ce cas, donne l'expression suivante :

$$R_d^{-1}(k) = \frac{1}{\lambda} \left[R_d^{-1}(k-1) - \frac{R_d^{-1}(k-1)x(k)x^T(k)R_d^{-1}(k-1)}{\lambda + x^T(k)R_d^{-1}(k-1)x(k)} \right] \quad (2.26)$$

En posant : $\psi(k) = R_d^{-1}(k-1)x(k)$, l'équation (2.26) s'écrit :

$$R_d^{-1}(k) = \frac{1}{\lambda} \left[R_d^{-1}(k-1) - \frac{\psi(k)\psi^T(k)}{\lambda + \psi^T(k)x(k)} \right] \quad (2.27)$$

⁴ $[A + BCD]^{-1} = A^{-1} - A^{-1}B[DA^{-1}B + C^{-1}]^{-1}DA^{-1}$

L'algorithme de type RLS possède une vitesse de convergence élevée qui est indépendante de la dispersion des valeurs propres de la matrice de corrélation d'entrée. Cet algorithme est également très utile dans des applications où l'environnement varie lentement. Le prix de tous ces avantages est une augmentation considérable de la complexité des calculs et certains problèmes de stabilité, qui ne sont pas aussi critique dans les algorithmes LMS [67, 68].

L'algorithme RLS est récapitulé au tableau 2.2.

RLS
Initialisation de $0 < \lambda < 1$, $\varepsilon \approx 1/\sigma_x^2$, and $\mathbf{R}^{-1}(0) = \varepsilon \mathbf{I}$ $\mathbf{x}(-1) = \mathbf{w}(-1) = [0 \ 0 \ \dots \ 0]^T$ Pour chaque k $e(k) = d(k) - \mathbf{x}^T(k)\mathbf{w}(k-1)$ $\boldsymbol{\psi}(k) = \mathbf{R}_d^{-1}(k-1)\mathbf{x}(k)$ $\mathbf{R}_d^{-1}(k) = \frac{1}{\lambda} \left[\mathbf{R}_d^{-1}(k-1) - \frac{\boldsymbol{\psi}(k)\boldsymbol{\psi}^T(k)}{\lambda + \boldsymbol{\psi}^T(k)\mathbf{x}(k)} \right]$ $\mathbf{w}(k) = \mathbf{w}(k-1) + e(k)\mathbf{R}_d^{-1}(k)\mathbf{x}(k)$

Tableau 2.2 L'algorithme RLS

2.7 Applications des filtres adaptatifs

Les filtres adaptatifs constituent des outils puissants en traitement du signal et en communications numériques. Les applications de ces filtres sont diverses, on cite à titres d'exemples : la modélisation directe ou identification (identification du canal), la modélisation inverse (égalisation), la prédiction linéaire et l'annulation d'écho [68].

Dans ce qui suit, on se concentrera sur leur application à la restauration des signaux altérés par les canaux de communication numérique. Dans ce contexte de tels filtres adaptatifs sont nommés égaliseurs.

2.8 Filtres adaptatifs en égalisation

Les dispositifs utilisés pour restaurer le signal et combattre les imperfections du canal de transmission ont pour effet d'égaliser l'enveloppe des signaux reçus ; ils sont pour cela nommés égaliseurs [1]. La configuration du filtre adaptatif en égalisation est donnée à la figure 2.2 [67, 68].

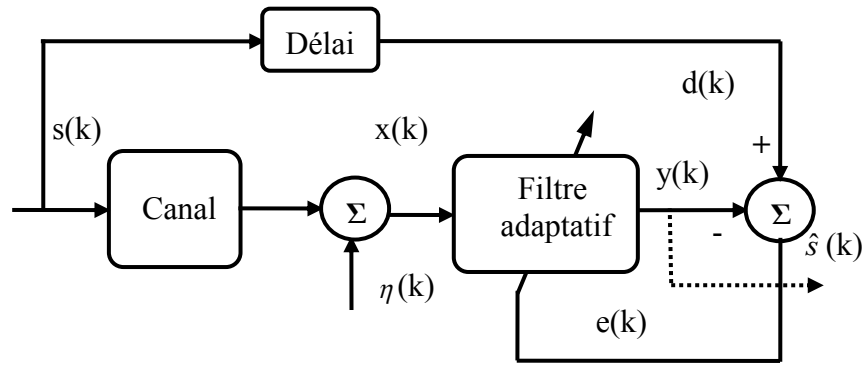


Figure 2.2 Configuration d'un filtre adaptatif en égalisation

2.8 .1 Principe de l'égalisation

Dans les systèmes de communication pratiques, les signaux transmis sont sujet à des distorsions diverses, telles que le bruit, les interférences inter-symboles dues à la propagation multi-trajet et les effets non linéaires des amplificateurs. Tous ces facteurs altèrent le signal d'origine, il est alors fondamental de restaurer l'entrée du canal en observant sa sortie [2, 3, 7].

Le processus de restauration du signal d'origine utilise un filtre adaptatif tel qu'un filtre à réponse impulsionnelle finie comme égaliseur du canal, (figure 2.2.) [67, 68]. Le rôle du filtre égaliseur est d'utiliser les observations du canal pour restaurer le signal transmis et réduire au mieux les distorsions causées par le canal [1, 4, 7-10].

La question qui se pose est de savoir comment estimer les valeurs des coefficients du filtre égaliseur pour pouvoir restaurer le signal transmis, avec une faible probabilité d'erreur possible. Pour accomplir ce fait, il existe deux approches: l'approche indirecte où le modèle du canal est d'abord estimé et en utilisant cette estimation, le réglage des coefficients du filtre égaliseur est alors optimisé. L'autre est l'approche directe où les paramètres du filtre égaliseur sont ajustés directement sans l'utilisation d'une estimation du canal [3]. C'est cette dernière approche qu'on va adopter. Pour commencer on s'intéresse aux algorithmes qui utilisent une séquence d'apprentissage pour la mise à jour initiale des coefficients du filtre égaliseur. Dans cette situation, une séquence d'apprentissage $d(k)$ connue par le récepteur est envoyée via un canal donné générant un signal déformé. La séquence $s(k)$, est utilisée après un décalage temporel approprié comme signal de référence du filtre adaptatif dont l'entrée est le signal déformé. Lorsque la fonction d'erreur se rapproche de zéro, le signal de sortie $y(k)$ ressemble au signal transmis $s(k)$, indiquant que le filtre adaptatif compense les distorsions du canal de transmission et restaure

l'information transmise. Après ce processus d'adaptation, les informations souhaitées peuvent être envoyées par le canal, qui est correctement égalisé par le filtre adaptatif.

Dans le cas idéal, $\hat{s}(k)$ sont identiques à $s(k)$, cependant, des erreurs de transmission peuvent se produire. Une optimisation des coefficients du filtre égaliseur minimisant au mieux la probabilité d'apparition de ces erreurs est fortement désirée.

Plusieurs critères d'optimisation de ces coefficients sont proposés, le plus utilisé est celui de la minimisation de l'erreur quadratique moyenne.

2.8.2 Modes opératoires des filtres égaliseurs

Pour pouvoir restaurer un signal au niveau du récepteur, il est nécessaire d'avoir à sa disposition en plus du signal reçu les informations partagées entre émetteur et récepteur qui peuvent être une partie du signal émis ou également des statistiques du signal comme la moyenne, le moments d'ordre deux ou d'ordre supérieur selon le mode opérationnel de l'égaliseur. Généralement un filtre égaliseur opère en deux modes: Le mode dit supervisé ou Data-Aided (DA), et le mode avec décision ou Directed Decision (DD) [1-3, 59].

2.8.2.1 Mode supervisé

Dans le mode supervisé une partie du signal émis est à priori connue du récepteur est émise et le processus d'égalisation est uniquement basée sur ces données. En effet l'optimisation de la fonction de coût se fera uniquement pendant la réception du signal connu. Pour les données non connues par le récepteur, les paramètres du filtre égaliseur ne sont pas modifiés [1,59].

Dans ce mode supervisé on retrouve aussi la technique d'égalisation par préambule qui consiste à insérer au début de chaque trame à émettre une séquence de données qui sera connue par le récepteur, l'optimisation se fera ainsi sur chaque début de trame [3, 61].

Pour des canaux variant lentement, il peut seulement être nécessaire de transmettre ce signal d'apprentissage au démarrage (start-up).

Ce mode de fonctionnement présente l'avantage de converger vite vers la valeur optimale du critère utilisé, et également d'avoir une complexité raisonnable. Par contre le principal inconvénient de ce processus est la diminution du débit utile au profit des données de références qui est souvent très handicapant surtout lorsque nous sommes en présence de canal sévère nécessitant un grand nombre de données de références[2, 3, 63].

Pour éviter une perte en débit utile du signal et en même temps obtenir des convergences assez rapides, le mode auto-adaptatif ou mode par retour de décision est souvent utilisé [1, 3, 59].

2.8.2.2 Mode de décision

Lorsque le canal de transmission n'est plus invariant dans le temps, l'adaptation doit assurer la poursuite des variations du canal, les coefficients du filtre égaliseur adaptatif doivent être ajustés continûment ; après la phase d'apprentissage, le processus d'égalisation commute vers le mode de décision qui est dirigé par le circuit de décision voir figure 2.3.

Dans la phase de poursuite, les décisions sur les symboles en sortie du filtre égaliseur sont utilisées pour actualiser les coefficients du filtre et l'adapter à la variation temporelle du canal. Dans le cas normal les décisions du récepteur sont correctes avec une grande probabilité pour permettre au filtre égaliseur de maintenir une égalisation précise [1, 4, 59].

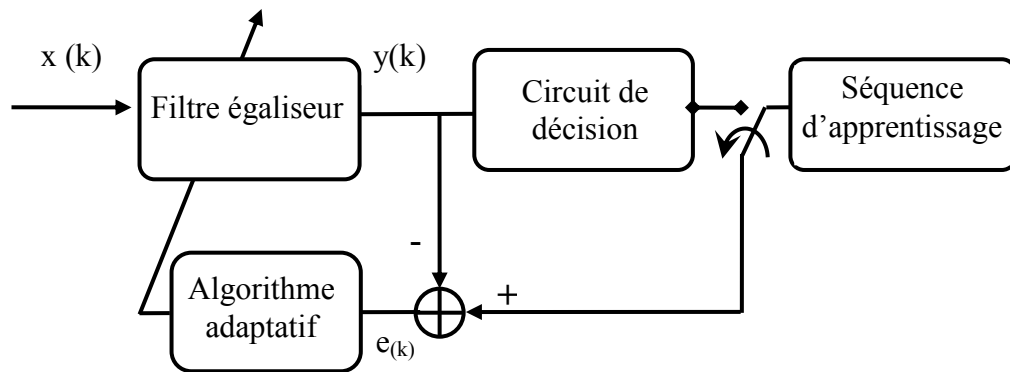


Figure 2.3 : Modes opératoires d'un filtre égaliseur

Il arrive que le filtre égaliseur opère en un mode unique dit mode non supervisé, autodidacte ou nommé encore NDA pour (Non-Data-Aided) sans l'aide d'une séquence d'apprentissage [3, 63, 70, 71]. Dans ce mode aucune information à priori du signal émis et du canal est supposée connue du récepteur, sauf éventuellement leurs statistiques d'ordre supérieur. Les filtres égaliseurs opérant dans ce mode sont généralement plus complexes à mettre en œuvre et convergent lentement vers la valeur optimale de la fonction de coût, mais ils ont un intérêt majeur qui est la non dégradation du débit utile du signal transmis.

Les algorithmes classiques d'adaptations dans ce domaine sont l'algorithme du module constant, le CMA pour (constant modulus algorithm), celui de Sato et celui de Godard [63, 70, 71].

2.8 .3 Choix de la structure du filtre égaliseur

Le choix de la structure utilisée pour le filtre égaliseur, est lié au critère de qualité de transmission que l'on souhaite réaliser. Plusieurs critères existent, on peut citer entre autres le critère du maximum de vraisemblance MV ou ML pour (maximum likelihood), le critère de probabilité d'erreur et le critère du Minimum d'erreur quadratique, MMSE. etc.

Parmi tous les critères que l'on peut rencontrer, le critère de probabilité d'erreur par bit est certainement le meilleur. Son rôle est de déterminer à partir d'un bit reçu le bit le plus probable émis. Cela nécessite des structures très complexes à mettre en œuvre [1, 3, 69]. Ce critère est souvent utilisé dans les simulations comme une référence pour les différents autres critères.

Le critère étant choisi, vient ensuite le moyen d'optimisation du critère. La théorie de l'optimisation offre là aussi un choix varié d'algorithmes : on peut citer par exemple les algorithmes du gradient, l'algorithme de Viterbi etc.

Le choix d'un algorithme est essentiellement lié au choix du critère. L'algorithme de Viterbi par exemple est appliqué à un critère du maximum de vraisemblance, alors que les algorithmes du gradient sont reliés aux critères linéaires tels que l'erreur quadratique moyenne MSE pour (Mean Square Error) [1, 3, 59, 69].

Le choix de la structure quant à lui est déterminé à la fois par le critère et l'algorithme d'optimisation. Deux structures fondamentales sont présentées dans la littérature : la structure transversale et la structure en treillis. Ces structures sont conçues à partir des filtres à réponse impulsionnelle finie RIF, ou à réponse impulsionnelle infinie RII [67]. Des structures de type non-linéaire sont à base de filtre mais ayant un comportement non linéaire comme par exemple les architectures avec retour de décision [3, 67].

La structure la plus employée est la structure transverse associée aux filtres RIF en raison de sa simplicité et de son efficacité. Contrairement à celle associée aux filtres RII, qui rend le problème d'optimisation plus difficile. La figure (2.4) donne une classification générale des différents types des filtres égaliseurs conventionnels rencontrés ainsi que quelques algorithmes d'optimisation [72] sans prétention toutefois, d'en fournir une liste exhaustive.

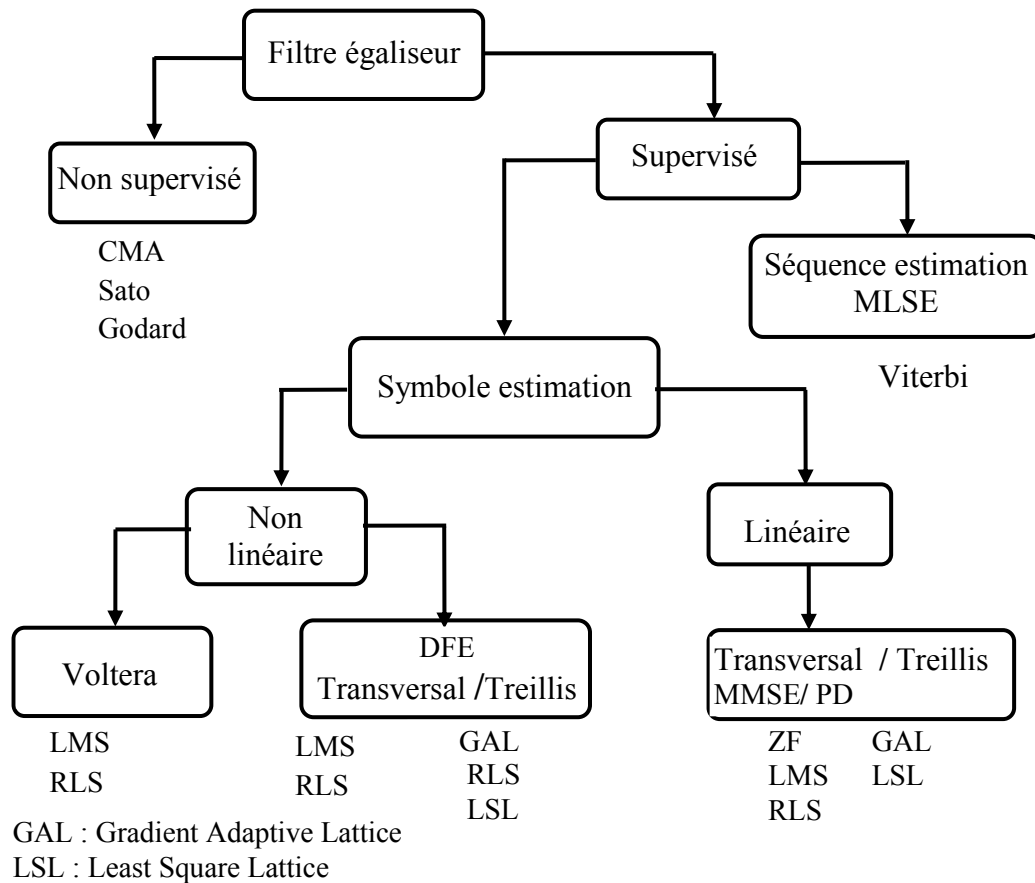


Figure 2.4 Filtres égaliseurs et algorithmes correspondants

2.9 Types des filtres égaliseurs

Comme on l'a évoqué antérieurement, il existe plusieurs variétés de filtres égaliseurs ainsi que plusieurs critères d'optimisation de ces derniers [1, 59.] Parmi les filtres égaliseurs existants en littérature on présentera: le filtre égaliseur par forçage à zéro et le filtre égaliseur de structure transversale et en treillis.

2.9.1 Filtre égaliseur par forçage à zéro

Ce filtre connu par son acronyme ZF (Zero Forcing) est un filtre qui tente d'inverser exactement la fonction de transfert du canal, ce qui est a priori précisément le but recherché, par l'égalisation [1, 4]. Ce faisant, l'interférence inter-symboles est exactement compensée, et l'on dit que l'interférence est forcée à zéro. Ce type de filtre utilise le critère de la distorsion crête ou PD pour (Peak Distortion) [1, 59].

Pour illustrer ceci, on considère la réponse du filtre équivalent noté $\{q_n\}$, résultant de la mise en cascade du filtre égaliseur représenté par ses coefficients $\{w_n\}$, et du canal dont les coefficients de la réponse impulsionnelle sont donnés par $\{h_n\}$.

$$q_n = \sum_{j=-\infty}^{\infty} w_j h_{n-j} \quad (2.29)$$

Le filtre égaliseur est supposé avoir un nombre infini de coefficients, sa sortie à la $k^{\text{ième}}$ instant, peut être exprimée comme suit:

$$\hat{s}_k = q_0 s_k + \sum_{n \neq k} s_n q_{k-n} \quad (2.30)$$

Le premier terme dans (2.30) représente le symbole désiré à un coefficient multiplicatif près; le second terme correspond à celui de l'interférence inter-symboles [1,59]. La valeur crête de l'interférence noté par PD (2.31), qui détermine le critère de la distorsion crête est défini comme étant l'énergie du filtre équivalent répartie sur les symboles interférant, comme explicité ci-dessous:

$$\begin{aligned} PD &= \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} |q_n| \\ &= \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \left| \sum_{j=-\infty}^{\infty} w_j h_{n-j} \right| \end{aligned} \quad (2.31)$$

Le PD est en fonction des coefficients du filtre égaliseur, par conséquent, on cherche les coefficients qui permettraient d'annuler totalement le critère PD, c'est à dire l'ISI, ce qui se traduit par:

$$q_n = \sum_{j=-\infty}^{\infty} w_j h_{n-j} = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \quad (2.32)$$

On retrouve ainsi la condition d'ISI nul qu'on a évoqué au chapitre un, au moyen de l'équation (1.22) du paragraphe 1.4.5.

En prenant la transformée en Z de l'équation (2.32), on obtient:

$$Q(z) = W(z)H(z) = 1 \quad (2.33)$$

Où $W(z)$, $H(z)$ et $Q(z)$ représentent respectivement les transformées en Z du filtre égaliseur, du canal de transmission et du filtre équivalent résultant de leur cascade [1,59]. Le filtre égaliseur qui réalise cette condition sera donc obtenu en inversant directement la transformé en Z de la réponse impulsionnelle du canal comme c'est exprimé ci-dessous:

$$W(z) = \frac{1}{H(z)} \quad (2.34)$$

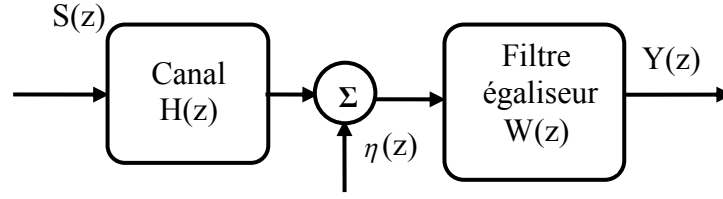


Figure 2.5 Filtre égaliseur par forçage à zéro

Dans le cas où le filtre égaliseur utilisé soit un filtre de réponse impulsionnelle finie de longueur $2N+1$. On peut déterminer une solution rapprochée permettant de minimiser le critère PD en acceptant une certaine erreur résiduelle. Cette solution correspond à celle qui annule complètement les $2N+1$ premières équations du système $PD = 0$, de sorte que :

$$q_n = \sum_{j=-N}^N w_j h_{n-j} = \begin{cases} 1, & n = 0 \\ 0, & -N \leq n \leq N, n \neq 0 \end{cases} \quad (2.35)$$

Les autres coefficients q_n définis pour $n < -N$ et $n > N$ n'étant pas pris en compte, ils constitueront l'erreur résiduelle obtenue.

L'équation (2.32) peut être réécrite sous notation vectorielle comme suit [59]:

$$\begin{bmatrix} h_0 & \dots & h_{-N+1} & h_{-N} & h_{-N-1} & \dots & h_{-2N} \\ \vdots & \vdots & & & & & \\ h_{N-1} & \dots & h_0 & h_{-1} & h_{-2} & \dots & h_{-N-1} \\ h_N & \dots & h_1 & h_0 & h_{-1} & \dots & h_{-N} \\ h_{N+1} & \dots & h_2 & h_1 & h_0 & \dots & h_{-N+1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \\ h_{2N} & \dots & h_{N+1} & h_N & h_{-N+1} & \dots & h_0 \end{bmatrix} \begin{bmatrix} w_{-N} \\ \vdots \\ w_{-1} \\ w_0 \\ w_1 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.36)$$

Ou d'une façon équivalente par :

$$H \cdot W = Q \quad (2.37)$$

On obtient alors les coefficients du filtre égaliseur :

$$W = H^{-1}Q \quad (2.38)$$

Une solution itérative pour ce critère peut être envisageable sous certaine condition. En effet, le critère de distorsion crête est une fonction convexe des coefficients du filtre égaliseur, il possède un minimum global. En général, il n'existe pas d'algorithme simple

pour performer cette optimisation à l'exception du cas spécial où la distorsion crête à l'entrée du filtre égaliseur défini par PD_{in} de l'équation (2.36) est inférieure à l'unité :

$$PD_{in} = \frac{1}{|h_0|} \sum_{n=1}^L |h_n| \quad (2.39)$$

Avec L , la mémoire du canal. Dans ce cas il existe un algorithme simple appelé algorithme de forçage à zéro ou ZFA (Zero Forcing Algorithm) reposant sur une méthode de descente de gradient itérative dont les caractéristiques sont similaires à celle de l'algorithme LMS [1].

Le critère PD permet donc de supprimer totalement l'interférence inter symbole induite par le canal hormis le bruit additif. Ce critère est moins complexe à mettre en œuvre que les critères MAP et MV, il est suggéré dans beaucoup de systèmes de transmission comme le wifi, wimax [3]. On notera en outre que dans cette méthode supervisée, le canal est supposé parfaitement connu; ce qui n'est pas toujours le cas pratiquement et par conséquent une bonne estimation du canal s'avère nécessaire. Étant donné que le bruit additif n'est pas pris en compte dans ce critère, son niveau peut augmenter du fait de l'inversion de la fonction de transfert du canal ce qui peut entraîner une détérioration des performances obtenues [1,59].

2.9.2 Filtre égaliseur transversal

L'égaliseur le plus simple à mettre en œuvre est un filtre adaptatif transversal de type RIF dont la fonction de transfert $W(z)$, en l'absence de bruit converge vers $H^{-1}(z)$.

Pour pouvoir fonctionner de manière optimale, il est nécessaire, lors de la phase d'initialisation du filtre égaliseur, de commencer la communication par une séquence connue. L'adaptation des paramètres prendra un certain nombre de cycles afin de converger vers un minimum d'erreur.

Pour les canaux avec une réponse en fréquence inconnue mais invariante dans le temps, les paramètres de l'égaliseur sont ajustés en fonction des caractéristiques du canal. Ces paramètres restent fixes durant la transmission des données. Ce type d'égaliseur est appelé filtre égaliseur préétabli [1, 3].

Lorsque le canal de transmission n'est plus invariant dans le temps, l'adaptation doit assurer la poursuite des variations du canal par l'ajustement des paramètres du filtre égaliseur, les techniques adaptatives de traitement permettent de concevoir des algorithmes estimant en

permanence, les paramètres pour le filtre égaliseur en minimisant un certain critère. On parle alors d'égalisation adaptative [1, 4,59].

La figure 2.6 présente un filtre égaliseur transversal, la séquence de données provenant de l'émetteur traversent un canal qui y introduit des déformations dues à ces caractéristiques, la sortie du canal est corrompu par un bruit additif $\eta(k)$ est le signal $x(k)$. A partir du signal d'erreur $e(k)$ les paramètres du filtre sont optimisés par un algorithme de minimisation d'erreur.

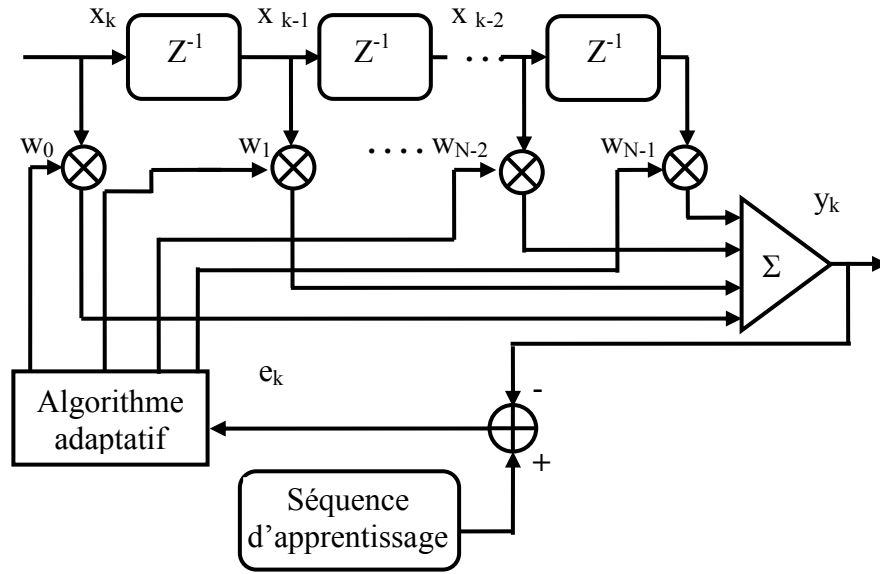


Figure 2.6 Filtre égaliseur transversal

Le signal reçu $x(k)$ présenté à l'entrée du filtre peut être déterminé par une convolution discrète entre le signal transmis et la réponse impulsionnelle du canal, discret équivalent additionné au bruit comme suit [1, 2, 4,59]:

$$x_k = \sum_{i=0}^L h_i s_{k-i} + \eta_k \quad (2.40)$$

Avec L , la longueur de la réponse impulsionnelle du canal, $x_k = [x_k \ x_{k-1} \ \dots \ x_{k-N+1}]^T$ le vecteur des observations bruitées, $s_k = [s_k \ s_{k-1} \ \dots \ s_{k-l-N+1}]^T$ le vecteur à l'entrée du canal et $\eta_k = [\eta_k \ \eta_{k-1} \ \dots \ \eta_{k-N+1}]^T$, le vecteur d'échantillons du bruit. T représente l'opération de transposition.

L'équation (2.40) peut être réécrite d'une façon équivalente sous la forme suivante:

$$x_k = H s_k + \eta_k \quad (2.41)$$

H étant la matrice de convolution du canal de dimension $N \times (L+N)$, ou sous forme matricielle comme suit :

$$\begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-N+1} \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 & \cdots & h_L & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_{L-1} & h_L & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & h_0 & h_1 & \cdots & h_L \end{bmatrix} \times \begin{bmatrix} s_k \\ s_{k-1} \\ \vdots \\ s_{k-L-N+1} \end{bmatrix} + \begin{bmatrix} \eta_k \\ \eta_{k-1} \\ \vdots \\ \eta_{k-N+1} \end{bmatrix} \quad (2.42)$$

La $k^{\text{ième}}$ sortie du filtre égaliseur notée $y(k)$ est la somme linéaire de la valeur courante et des valeurs passées du signal reçu qui sont respectivement multipliées par les coefficients du filtre égaliseur comme à l'équation suivante :

$$y_k = \sum_{i=0}^{N-1} w_i x_{k-i} = \mathbf{w}^T \mathbf{x}_k \quad (2.43)$$

Compte tenu du vecteur d'observations bruitées $\mathbf{x}(k)$ se présentant à l'entrée du filtre égaliseur, l'objectif est la conception d'un algorithme pour l'estimation du vecteur de paramètre $\mathbf{w} = [w_0 w_1 \dots w_{N-1}]^T$.

Une optimisation de ces coefficients minimisant au mieux la probabilité d'apparition des erreurs est fortement souhaitée. Plusieurs critères d'optimisation ont été proposés dans la littérature. Un des critères les plus utilisés est celui de la minimisation de l'erreur quadratique moyenne MMSE comme on l'a évoqué plus haut. On cherche le vecteur de coefficients optimum du filtre égaliseur qui minimise la fonction de coût.

En se basant sur l'équation (2.41), la matrice d'auto-corrélation du signal à l'entrée du filtre égaliseur $\mathbf{R}$ de l'équation (2.9) s'écrit alors comme suit :

$$\mathbf{R} = \mathbf{H}^T \mathbf{E}[\mathbf{s}_k \mathbf{s}_k^T] \mathbf{H} + 2 \mathbf{H}^T \mathbf{E}[\mathbf{s}_k^T \mathbf{n}_k] + \mathbf{E}[\mathbf{n}_k \mathbf{n}_k^T] \quad (2.44)$$

En considérant que la séquence binaire aléatoire $s(k)$ présentée à l'entrée du canal est formée de symboles équiprobables, de moyenne nulle et de variance normalisée σ_s^2 donnée par : $\sigma_s^2 = \mathbf{E}[s_k^2] = 1$, et que le canal est perturbé par un bruit blanc Gaussien additif $\eta(k)$, de moyenne nulle et de variance σ_n^2 définie par: $\sigma_n^2 = \mathbf{E}[\eta_k^2]$.

En remplaçant chaque expression par sa valeur correspondante et sachant que les données et le bruit sont dé-corrélés, on obtient :

$$\mathbf{R} = \mathbf{H}^T \mathbf{H} + \sigma_n^2 \mathbf{I} \quad (2.45)$$

En remplaçant l'expression de R dans (2.12), on obtient l'expression des coefficients optimums du filtre égaliseur:

$$W_o = (HH^T + \sigma_n^2 I)^{-1} P = \frac{P}{HH^T + \sigma_n^2 I} \quad (2.46)$$

Cette expression nécessite l'inversion de la matrice R , les coefficients sont aussi déterminés par des procédures itératives qui évitent l'inversion matricielle ; tels que les algorithmes itératifs LMS (2.17) et RLS (2.25) remémorés aux paragraphes précédents.

2.9.3 Filtre égaliseur en treillis

Les structures adaptatives transversales sont principalement basées sur des récurrences temporelles, les structures en treillis font appel à des récurrences temporelles et sur l'ordre. Ces filtres combinent plusieurs opérations de filtrage des erreurs de prédiction directe (avant) et rétrograde (arrière) dans une structure unique. Ils se composent d'une connexion en cascade d'unités élémentaires ou étages [66, 67,69,73]. A l'étage m , les erreurs de prédiction linéaire directe $f_m(k)$ et rétrograde $b_m(k)$ sont liées par les relations suivantes :

$$f_m(k) = f_{m-1}(k) - \Gamma_m b_{m-1}(k-1) \quad (2.47)$$

$$b_m(k) = b_{m-1}(k-1) - \Gamma_m f_{m-1}(k) \quad (2.48)$$

Où les coefficients Γ_m sont dit coefficients de réflexion ou encore coefficients de corrélation partielle exprimés par :

$$\Gamma_m = \frac{E[b_{m-1}(k-1)f_{m-1}(k)]}{E[|f_{m-1}(k)|^2]} \quad (2.49)$$

Ces relations permettent de calculer les erreurs de prédiction directe $f_m(k)$ et rétrograde $b_m(k)$, par récurrence sur l'ordre m . A l'initialisation pour $m = 0$, on choisit : $f_0(k) = b_0(k) = x(k)$. Cette procédure, itérée pour $(m = 0, \dots, M)$ conduit à la structure symétrique en treillis, seule la connaissance de l'ensemble des coefficients de réflexion, Γ_m est nécessaire. La structure en treillis est illustrée à la figure 2.7.

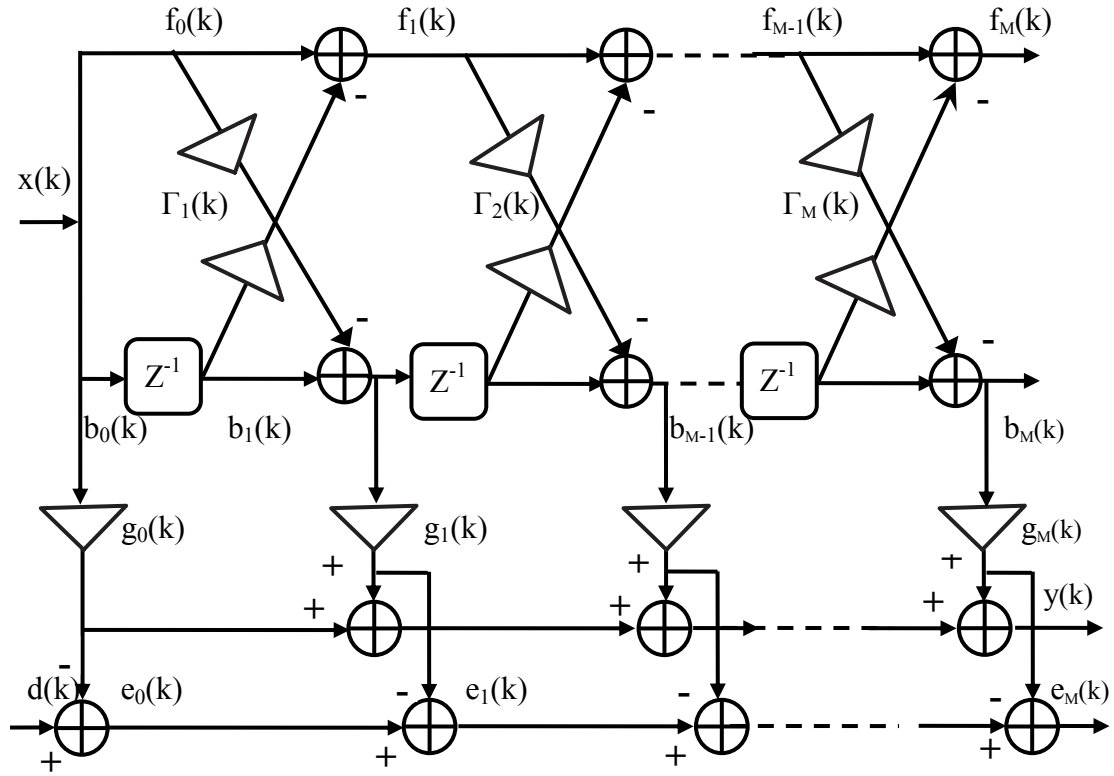


Figure 2.7 Filtre égaliseur en treillis

Le filtre égaliseur en treillis est constitué de m étages, dont les coefficients de réflexion $\Gamma_1, \Gamma_2, \dots, \Gamma_M$, transformant la séquence des observations corrélées $x(k)$, en une séquence d'erreurs de prédiction arrière non corrélées, $b_0(k), b_1(k) \dots, b_M(k)$. Cette dernière est ensuite utilisée comme entrée du filtre transverse de coefficients $g(k)$ pour déterminer la sortie $y(k)$ qu'on exprime par :

$$y(k) = \sum_{m=0}^M g_m(k) b_m(k) \quad (2.50)$$

La solution de Wiener donne les valeurs optimales des coefficients $g(k)$:

$$g_{opt} = D^{-1}Z \quad (2.51)$$

Où $D = E[b(k)b^T(k)]$ et $Z = E[d(k)b(k)]$ représentent respectivement la matrice d'auto-corrélation de l'erreur de prédiction et le vecteur d'inter-corrélation entre le signal de référence et l'erreur de prédiction.

Du fait l'orthogonalité des erreurs de prédiction, la matrice D^{-1} est une matrice diagonale [69]. L'obtention des coefficients optimum est relativement simple, étant donné qu'il n'est plus question d'inversion matricielle.

La structure du filtre égaliseur en treillis peut être adaptée selon une méthode itérative basée sur l'algorithme du type gradient nommé GAL (Gradient Adaptive Lattice) [69, 73]. Cet algorithme est récursif en temps et en ordre, à chaque itération dans le temps, les erreurs de prédiction avant et arrière d'ordre m , ainsi que les coefficients du treillis d'ordre m , sont calculés à partir de ces mêmes quantités d'ordre $m-1$.

A l'étage m , les coefficients de réflexion sont obtenus par une méthode itérative par l'expression suivante :

$$\Gamma_m(k+1) = \Gamma_m(k) + \mu_m(k)[f_m(k)b_{m-1}(k-1) + f_{m-1}(k)b_m(k)] \quad (2.52)$$

Dans un but de normalisation du pas d'adaptation μ_m , la valeur suivante est adoptée :

$$\mu_m(k) = \frac{\mu}{\varepsilon_{m-1}(k)} \quad (2.53)$$

Où μ représente un pas d'adaptation ordinaire et $\varepsilon_{m-1}(k)$ représente l'énergie totale des erreurs de prédiction, exprimée par méthode itérative par :

$$\varepsilon_{m-1}(k) = \alpha \varepsilon_{m-1}(k-1) + (1-\alpha)(f_{m-1}^2(k) + b_{m-1}^2(k-1)) \quad (2.54)$$

Avec, α est une constante positive comprise entre ($0 < \alpha < 1$)

Les coefficients du filtre transverse sont exprimés d'une manière itérative par :

$$g_m(k+1) = g_m(k) + \mu_{m,g}(k)e_m(k)b_m(k) \quad (2.55)$$

Le pas d'adaptation dans ce cas est exprimé comme suit :

$$\mu_{m,g}(k) = \frac{\mu}{\varepsilon_{m,b}(k)} \quad (2.56)$$

Où $\varepsilon_{m,g}(k)$ représente l'énergie totale des erreurs de prédiction arrière, exprimé par :

$$\varepsilon_{m,g}(k) = (1-\alpha)\varepsilon_{m,g}(k-1) + b_{m-1}^2(k) \quad (2.57)$$

Les coefficients de réflexion et les coefficients du filtre transverse (équations (2.52) et (2.55)) sont adaptés en parallèle.

Plus de détails sur l'algorithme adaptatif de la structure en treillis (GAL) sont donnés à l'annexe A. L'implémentation de l'algorithme GAL est récapitulée au tableau 2.3.

GAL
Initialisation : $\hat{f}_m(0) = b_m(0) = 0$; $\varepsilon_{m-1}(0) = \varepsilon_{m,g}(0) = \delta$
δ étant une petite constante
$\Gamma_m(0) = 0$; $g_m(0) = 0$
Pour chaque k
$\hat{f}_0(k) = b_0(k) = x(k)$
Pour chaque m
$\hat{f}_m(k) = f_{m-1}(k) - \Gamma_m b_{m-1}(k-1)$
$b_m(k) = b_{m-1}(k-1) - \Gamma_m f_{m-1}(k)$
$\varepsilon_{m-1}(k) = \alpha \varepsilon_{m-1}(k-1) + (1-\alpha)(f_{m-1}^2(k) + b_{m-1}^2(k-1))$
$\Gamma_m(k+1) = \Gamma_m(k) + \frac{\mu}{\varepsilon_{m-1}(k)} [f_m(k)b_{m-1}(k-1) + f_{m-1}(k)b_m(k)]$
$e_m(k) = d(k) - y(k)$
$g_m(k+1) = g_m(k) + \frac{\mu}{\varepsilon_{m,b}(k)} e_m(k)b_m(k)$
$\varepsilon_{m,g}(k) = (1-\alpha)\varepsilon_{m,g}(k-1) + b_{m-1}^2(k)$

Tableau 2.3 L'algorithme GAL

Les avantages principaux des filtres égaliseurs en treillis sont leur insensibilité au niveau de corrélation des entrées, due à l'orthogonalisation des étages successifs du treillis et la modularité de mise en œuvre. Par conséquent, l'ordre du filtre adaptatif peut être augmenté ou diminué sans affecter les solutions d'ordre inférieur [67,69].

2.10 Conclusion

Dans ce chapitre nous avons discuté quelques concepts fondamentaux sur le filtrage adaptatif et les différents algorithmes d'estimation des coefficients de ces filtres, tels que les algorithmes LMS, RLS et GAL. Nous avons étudié l'application de ces filtres adaptatifs dans le domaine des communications numériques en tant que filtre égaliseur permettant de supprimer les distorsions dues aux canaux de transmission. Nous avons également donné un aperçu sur les types des filtres égaliseur à savoir, les filtres égaliseur par forçage à zéro, transversal et en treillis.

L'algorithme LMS est très populaire en raison de sa simplicité, sa faible complexité de calcul, sa convergence évidente dans un environnement stationnaire, son comportement stable lors de son implémentation. Cependant sa convergence dépend du conditionnement de la matrice de corrélation du signal d'entrée. L'algorithme RLS s'affranchi de cette dépendance et possède une vitesse de convergence élevée. Cet algorithme est également très utile dans des applications où l'environnement varie lentement. Le prix de ses avantages est une augmentation considérable de la complexité des calculs et certains problèmes de stabilité, qui ne sont pas aussi critique pour le LMS. La convergence de l'algorithme GAL est aussi moins dépendante au conditionnement de la matrice de corrélation du signal d'entrée, en raison des propriétés d'orthogonalisation des étages successifs du treillis. Cet algorithme possède une vitesse de convergence plus rapide que son équivalent sous forme transversale.

Chapitre 3

Filtrage adaptatif non linéaire à base des réseaux de neurones

3.1 Introduction

Le filtrage adaptatif non linéaire utilise des structures non linéaires, pour le filtre adaptatif. Les réseaux de neurone disposent d'une multitude de propriétés qui les rendent très attrayants pour le filtrage adaptatif non linéaire, en communications numériques. Plus particulièrement, leur capacité de généralisation, leur non-linéarité intrinsèque due aux fonctions d'activations, les rendent avantageux en tant que filtres égaliseurs non linéaires. Beaucoup de recherches ont montré que les filtres égaliseurs non linéaires à base des réseaux de neurones peuvent fournir des meilleures performances en comparaison avec les filtres égaliseurs conventionnels [16, 17, 19-21]. Dans ce chapitre, nous utilisons les réseaux de neurones en tant que filtres égaliseurs pour la restauration des signaux altérés par les imperfections des canaux de communications. Nous présentons les structures neuronales des filtres égaliseurs les plus utilisées, à savoir, le perceptron multicouches MLP [2, 16, 22-24], la fonction de base radiale RBF [25-27] et le réseau récurrent RNN [28-30]. On se focalise principalement sur les filtres égaliseurs non linéaire à base du réseau MLP.

3.2 Aperçu général sur les réseaux neuronaux

Les réseaux de neurones artificiels constituent des structures s'inspirant du cerveau humain. Les neurones qui forment les centres de décisions sont connectés de façon à accomplir une tâche bien déterminée. Les réseaux de neurones constituent une méthode de programmation qui consiste à établir les paramètres de chaque neurone et peuvent ainsi faire l'approximation d'une fonction de façon précise.

Les ANN cherchent à hériter de certaines propriétés intéressantes des réseaux de neurones biologiques, telles que le parallélisme dans l'activité des unités nerveuses, la forte puissance de calcul, la tolérance envers des défaillances [74, 75]. Un réseau neuronal possède les caractéristiques privilégiées représentées par sa capacité à apprendre des informations et à généraliser. L'emploi des réseaux de neurones plutôt que des techniques classiques se justifie par les arguments suivants:

- ❖ Simplicité de mise en œuvre.
- ❖ Leur propriété d'approximation universelle prouvée.
- ❖ La possibilité de considérer le processus comme étant boîte noire.
- ❖ Leur robustesse par rapport à des défaillances internes du réseau.
- ❖ Leur capacité d'adaptation aux conditions imposées par un environnement quelconque.
- ❖ La facilité de rechanger ses paramètres (poids, nombre de neurones cachés, nombre de couches cachées...) lors d'une modification éventuelle.
- ❖ Leur réalisation matérielle parallèle.

3.2.1 Le neurone artificiel

Un neurone artificiel est une modélisation simplifiée du neurone biologique, c'est un processeur élémentaire qui reçoit un nombre d'entrées en provenance des autres neurones ou de l'extérieur. A chacune de ces entrées est associé un poids représentant la force de la connexion.

Tous les types de neurones ont le même principe de base. On associe chaque entrée à un poids w , c'est-à-dire un facteur multiplicatif. Pour garder l'analogie avec le cerveau humain, l'adaptation des poids est appelée apprentissage. Les produits sont ensuite sommés pour constituer la base du processus de décision. C'est ce que nous appellerons la somme des produits. Le résultat passe ensuite dans une fonction de décision $\varphi(.)$ dite fonction d'activation afin d'évaluer la valeur de sortie. La figure 3.1 montre le neurone artificiel.

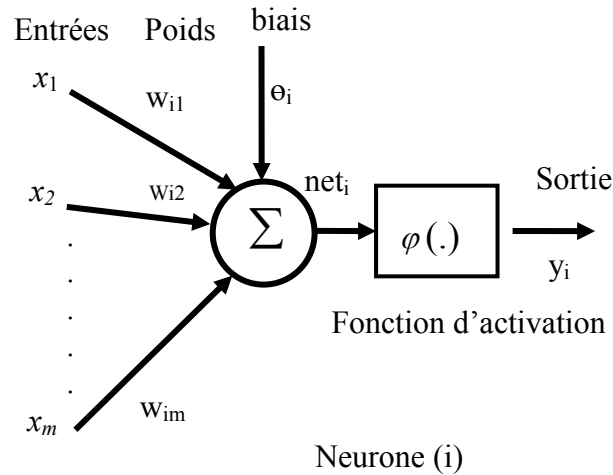


Figure 3.1 Modèle de base du neurone artificiel

La sortie du neurone est une somme pondérée de ses entrées plus un biais, multipliée par une fonction d'activation qui peut être linéaire ou non linéaire (identité, sigmoïde, tangente hyperbolique, etc.) [74-76]. Une fonction d'activation linéaire aboutira à une structure linéaire, qui défait le but d'une mise en œuvre des réseaux de neurones. La somme pondérée du neurone est exprimée par:

$$net_i = \sum_{k=1}^m w_{ik} x_k + \theta_i \quad (3.1)$$

La sortie du neurone est donnée par :

$$y_i = \varphi(net_i) = \varphi\left(\sum_{k=1}^m w_{ik} x_k + \theta_i\right) \quad (3.2)$$

x_k est la valeur de la k^{ieme} entrée du neurone, w_{ik} est le poids synaptique correspondant à l'entrée k du neurone i , θ_i désigne le biais ou seuil propre au neurone qui est un nombre réel représentant la limite à partir de laquelle le neurone s'activera.

$\varphi(.)$ Représente la fonction d'activation qui limite en général l'amplitude de la sortie du neurone entre $[-1, 1]$. Plusieurs fonctions d'activations sont possibles, les plus utilisées sont illustrées à la figure 3.2

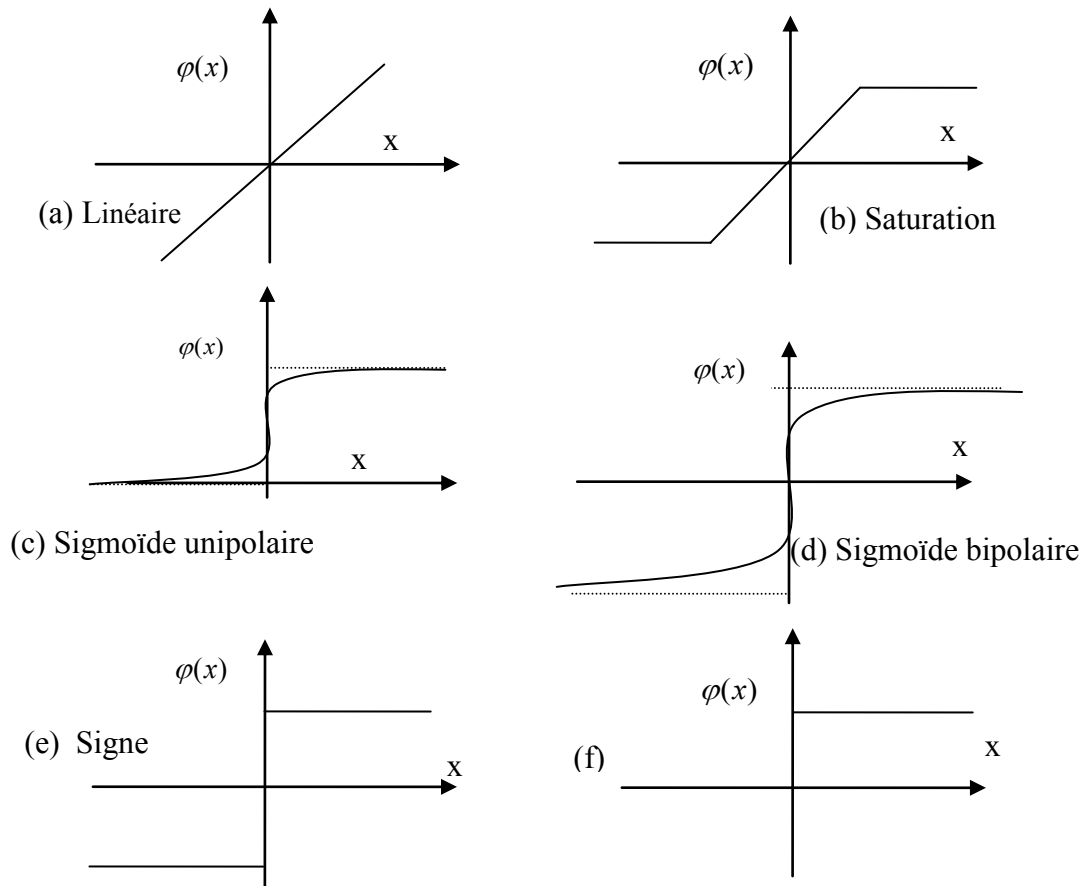


Figure 3.2 Fonctions d'activations courantes.

On peut faire les commentaires suivants:

- ❖ L'utilisation d'une fonction linéaire dans un réseau multicouche, le réduit à un filtre FIR conventionnel.
- ❖ La fonction doit être bornée, sinon un débordement pourrait survenir dans un calculateur.
- ❖ Les dérivées de $\varphi(\cdot)$ doivent exister et être bornées, afin d'être utilisables par l'algorithme d'apprentissage du réseau neuronal tel que la rétro-propagation du gradient ou back propagation (BP) qui exige des fonctions d'activation différentiables en tout point. Pour cela on utilise en général des fonctions continues, telles les fonctions de type sigmoïde.
- ❖ Si le signal que l'on désire obtenir est symétrique, alors la fonction d'activation doit l'être aussi.

3.2.2 Théorème de l'approximation universelle

Ce théorème indique le nombre minimal de couches cachées que doit avoir un perceptron multicouches pour fournir une réalisation approximative de toute fonction continue. Il présente aussi un ensemble de conditions que doit satisfaire une fonction d'activation pour que le réseau multicouche MLP correspondant jouisse de cette propriété d'approximation universelle [18, 75-78].

Ce théorème s'énonce comme suit :

Etant donné $\varphi(.)$ une fonction non-constante, bornée, monotone croissante et continue

Soit I_m l'hyper-cube unitaire de dimension m ; $[0,1]^m$

Soit $C(I_m)$ l'ensemble des fonctions continues définies sur I_m . Ainsi, pour toute fonction f appartenant à $C(I_m)$, et pour tout $\varepsilon > 0$, il existe un entier p et des constantes $\alpha_i, \theta_i, w_{ik}$, avec $i = \overline{1, p}$ et $k = \overline{1, m}$, de sorte que nous puissions définir :

$$F(x_1, \dots, x_m) = \sum_{i=1}^p \alpha_i \varphi \left(\sum_{k=1}^m w_{ik} x_k + \theta_i \right) \quad (3.3)$$

Comme une approximation de la fonction $f(.)$; de telle façon que

$$|F(x_1, \dots, x_m) - f(x_1, \dots, x_m)| < \varepsilon \quad (3.4)$$

Pour tout $x_1, x_2, \dots, x_m$ de l'espace de l'entrée (appartenant à I_m)

Notons que l'équation (3.3) représente la sortie d'un perceptron multicouche ayant m entrées et une seule couche cachée de p neurones et une couche de sortie possédant un seul neurone.

Le théorème d'approximation universelle est un théorème d'existence dans le sens qu'il fournit une justification mathématique pour approximer n'importe quelle fonction arbitraire. Il indique qu'une seule couche cachée est suffisante pour un perceptron multicouches pour performer cette approximation. Le théorème informe uniquement sur l'existence d'une fonction aux caractéristiques précédentes cependant il ne donne pas d'indications sur la façon de l'optimiser.

3.2.3 Apprentissage des réseaux de neurones

L'apprentissage est la propriété la plus intéressante des réseaux neuronaux. Vu que l'information que peut acquérir un réseau de neurones est représentée dans les poids des connexions entre les neurones, l'apprentissage consiste donc à ajuster ces poids en fonction d'un jeu de données présentées en entrée de telle manière que le réseau manifeste certains comportements désirés [74 -76].

L'apprentissage peut être supervisé ou non supervisé selon la présence ou l'absence de la réponse désirée comme c'est montré à la figure 3.3.

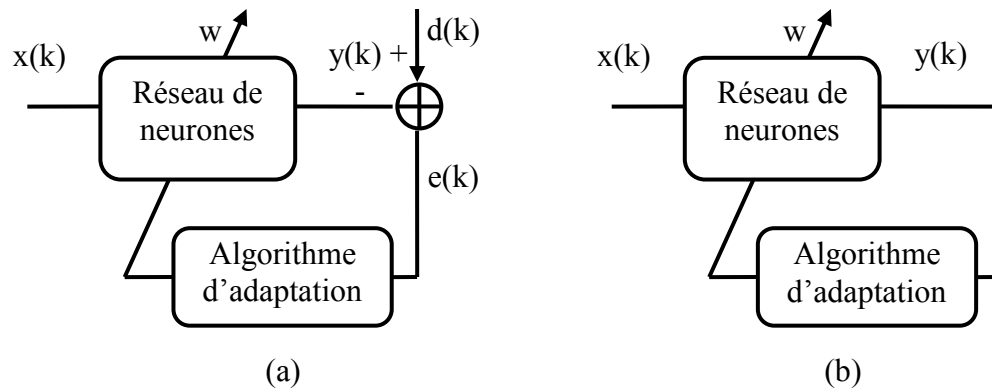


Figure 3.3 Modes d'apprentissage des réseaux de neurones

(a) supervisé, (b) non supervisé

3.2.3.1 Apprentissage supervisé

Ce mode d'apprentissage dispose d'un comportement de référence, vers lequel il tente de faire converger le réseau, ce qui nécessite la connaissance à priori de la réponse désirée $d(k)$. Les valeurs d'entrée sont présentées au réseau et la sortie actuelle correspondante est calculée. Les poids sont ensuite ajustés de façon à réduire l'écart entre la réponse désirée et celle du réseau, pour satisfaire un certain critère de performance tel que l'erreur quadratique moyenne.

3.2.3.2 Apprentissage non supervisé

Dans ce mode, seules les valeurs d'entrée sont disponibles, pendant la phase d'apprentissage, le réseau neuronal construit une topologie de l'espace des vecteurs d'entrées. Dans ce cas, la connaissance à priori de sortie désirée n'est pas nécessaires et la procédure d'apprentissage est basée uniquement sur les valeurs d'entrées.

Le réseau s'auto-organise de manière à optimiser une certaine fonction de coût, sans qu'on lui présente la réponse désirée.

L'intérêt de cette approche est que l'on n'a pas besoin de disposer d'exemple de problèmes résolus, c'est à dire, de vecteur de sortie correspondant à chaque vecteur d'entrée utilisée pour l'apprentissage.

3.3 Types des réseaux de neurones

Un réseau de neurones artificiels est un ensemble de neurones associés en couches et fonctionnant en parallèle. Ce réseau neuronal possède la capacité de stocker de la connaissance empirique et de la rendre disponible à l'usage. Les habiletés de traitement et la connaissance du réseau vont être stockées dans les poids synaptiques, obtenus par des processus d'adaptation ou d'apprentissage. La structure des connexions entre les différents neurones détermine la topologie du réseau. On peut classer les réseaux de neurones artificiels dans deux grandes catégories :

- ❖ Réseaux neuronaux non bouclé de type Feed- Forward.
- ❖ Réseau neuronaux récurrent ou de type Feed-back.

3.3.1 Les réseaux neuronaux de type Feed-Forward

Ces réseaux sont formés de couches constituées de neurones de mêmes types. Plusieurs couches peuvent composer un réseau ayant plus de flexibilité afin de répondre à un plus grand nombre d'applications.

Dans ce type de réseau, l'information se propage couche en couche sans que le retour en arrière soit possible. La connectivité des neurones dans les réseaux à couches est locale, où ils ne sont reliés qu'à leur plus proche voisin de la couche précédente et la couche suivante. La couche de sortie est celle qui produit la sortie du réseau, il y a un neurone par sortie.

Ce réseau est considéré comme un système neuronal non linéaire statique. Il existe trois types de réseaux neuronaux multicouches de type Feed Forward qui sont: Le perceptron multicouche MLP, le réseau RBF (Radial- Basis- Function) et le réseau d'ondelettes WNN (Wavelet Neural Network). On se limitera à présenter les deux premiers types seulement.

3.3.1.1 Le perceptron multicouche

L'architecture du réseau multicouche MLP est basée sur l'idée de grouper des neurones dans une couche, ensuite plusieurs couches sont rangées bout à bout et les neurones de deux couches adjacentes sont connectés. La couche de transmission dite couche d'entrée est une couche passive qui n'effectue aucun traitement, elle transmet les entrées au réseau son rôle est de recevoir et transmettre les configurations à mémoriser.

Les couches internes sont appelées les couches cachées elles se situent entre la couche de transmission et la couche de sortie, leur nombre est variable et dépend de l'application. Le filtrage non linéaire du MLP dépend énormément du nombre de couches cachées et du nombre des neurones dans ces couches. Trop de neurones compliquera l'apprentissage et augmentera le coût de calcul de la structure. Trop peu de neurones seront inadéquats pour créer la non linéarité suffisante [74-76]

Notons que chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante (à l'exception des couches d'entrée et de sortie) et il n'y a pas de connexions entre les neurones d'une même couche. Les sorties des neurones dans une couche forment les entrées de la couche suivante. La figure 3.4 montre l'architecture d'un réseau multicouche MLP.

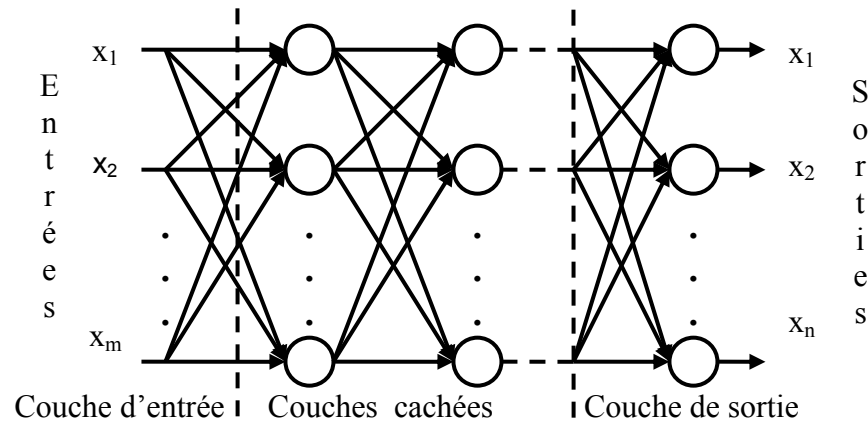


Figure 3.4 Le perceptrons multicouche MLP

Le réseau MLP contient une couche d'entrée avec m entrées, L couches cachées de $p(N)$ neurones chacune, avec $1 \leq N \leq L$ et une couche de sortie contenant n neurones.

Selon le théorème d'approximation universelle évoqué ci-dessus, un réseau MLP avec seulement une couche cachée dont la fonction d'activation est de type sigmoïdale et une sortie

avec une fonction d'activation linéaire peut approximer n'importe quelle fonction continue pourvu qu'un nombre suffisant de neurones soit fourni [18,75].

3.3.1.2 Le réseau à fonction de base radiale

Le réseau à fonction de base radiale (RBF : Radial Basis Function) est basé sur une architecture multicouche ; une couche d'entrée passive constituée de nœuds ayant la taille du vecteur d'entrée (M). Une couche cachée constituée de (N_1) neurones RBF qui sont connectés à ceux de la couche d'entrée par des connexions directes non pondérées contrairement au MLP. Cette couche effectue une transformation non linéaire de l'espace d'entrée. La couche de sortie possède (N_2) neurones connectés à ceux de la couche cachée par des connexions pondérées. Cette couche calcule une combinaison linéaire des sorties de la couche cachée.

L'architecture d'un réseau RBF est représentée à la figure 3.5. L'utilisation habituelle des RBF conserve une fonction d'activation linéaire en sortie mais l'utilisation d'une fonction non linéaire reste possible [75, 76, 79].

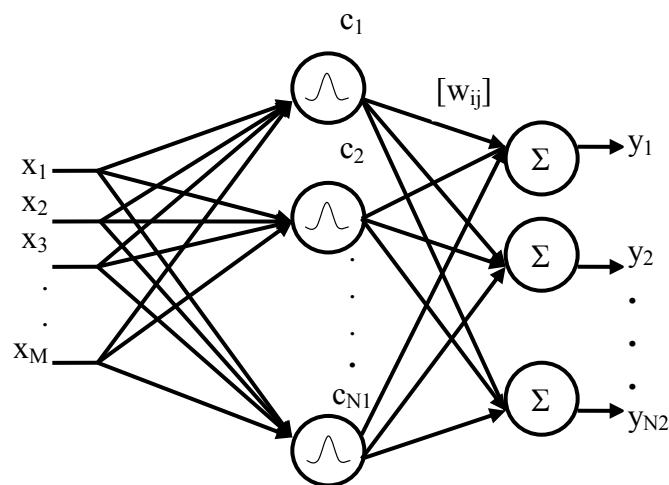


Figure 3.5 Architecture du réseau RBF

L'apprentissage d'un réseau RBF consiste à déterminer les paramètres inconnus de ce réseau. Cela signifie, déterminer le nombre de fonctions de base (unités cachées), les centres et les largeurs de chaque fonction de base, et les poids de la couche de sortie. D'une manière générale l'apprentissage est composé de deux étapes [75, 76, 79].

Une étape de paramétrisation des fonctions de base qui consiste en une règle d'apprentissage non supervisée dans la couche cachée. Elle a pour rôle de trouver la forme

exacte des fonctions de base qui consiste à déterminer le nombre et la position des centres. Le choix de la position des centres et le nombre de neurones reste généralement arbitraires. Il existe des méthodes de positionnement automatique des centres, ces méthodes consistent à déplacer les centres vers les points où l'erreur est grande. L'algorithme des K-means et la méthode souvent utilisée pour la sélection des centres. L'apprentissage des poids de la couche de sortie qui consiste en une règle d'apprentissage supervisé pour l'adaptation des poids tels que les algorithmes RLS ou LMS. Chaque neurone élémentaire calcule la distance entre l'entrée et son centre. Le résultat est ensuite multiplié par une fonction d'activation, généralement de type gaussienne, exprimée par :

$$\phi(x) = \exp\left(-\frac{x^2}{2\rho^2}\right) \quad (3.8)$$

La sortie de la couche cachée notée x_{1i} est donnée par :

$$x_{1i} = \phi_i(\|x - c_i\|) \quad (3.9)$$

Avec, x le vecteur d'entrée, c_i est le centre associé au neurone i et $\|\cdot\|$ dénote la norme euclidienne. La sortie du réseau est donnée par :

$$y_j = \sum_{i=1}^{N_1} w_{ij} \phi_i(\|x - c_i\|) \quad (3.10)$$

Où N_1 est le nombre de neurones RBF et w_{ij} , avec $j = \overline{1, N_2}$, sont les poids associés à la couche de sortie. Le paramètre ρ associé à ϕ est une mesure de la largeur de la i^{eme} fonction radiale de centre c_i . Généralement, tous les noyaux gaussiens cachés disposent d'une largeur commune, dans une telle situation, les paramètres qui les distinguent les uns des autres sont les centres c_i . La valeur que prend la sortie du noyau gaussien est d'autant plus importante que l'entrée est plus proche de son centre et tend vers zéro, lorsque la distance, entrée-centre devienne importante.

D'autres fonctions autre que la gaussienne peuvent aussi être utilisées telles que:

- ❖ La fonction multiquadrique exprimée par : $\phi(x) = (x^2 + \rho^2)^{\frac{1}{2}}$
- ❖ La fonction inverse multiquadrique donnée par : $\phi(x) = \frac{(x^2 + \rho^2)^{-\frac{1}{2}}}{\rho}$

A l'instar du réseau MLP, le réseau RBF possède aussi la propriété d'approximation universelle c'est à dire que le réseau est capable d'approximer n'importe quelle fonction continue avec une précision donnée pourvu qu'on fournisse un nombre de neurones suffisant [75,76]. Cette propriété garantit que le réseau RBF aura au moins les mêmes capacités théoriques que le réseau multicouche, avec des non linéarités sigmoïdales [76].

3.3.1.3 Comparaison entre le réseau MLP et le réseau RBF

Certains travaux ont été effectués pour montrer la relation entre les réseaux RBF et les MLP. Particulièrement, si l'on considère qu'un MLP est un approximateur universel, il peut approximer un réseau RBF et vice versa [76]. Il a été montré que pour les entrées normalisées, le MLP peut être considéré comme un réseau RBF avec des fonctions de base irrégulières [76].

La différence entre les deux réseaux porte sur l'architecture qui est figée en deux couches pour les RBF et peut comporter plusieurs couches pour le MLP. En plus le réseau MLP jouit de la flexibilité de la structure, de bonnes capacités de représentation et une large utilisation pour les applications industrielles et un grand nombre d'algorithmes d'apprentissage [31].

Le réseau RBF quand à lui souffre d'une difficulté dans le calcul de la distribution optimale des centres, si ces derniers ne sont pas identifiés avec une grande précision, il se produira des dégradations dans les performances du système [75]. En outre la complexité de calcul augmente exponentiellement avec le nombre d'entrée. Cependant l'addition d'une couche cachée supplémentaire améliore la généralisation du MLP et peut réduire le nombre de neurones requis [75]. Toutefois, il faut garder à l'esprit que les deux types de réseaux jouissent d'une capacité d'approximation universelle, la principale raison de considérer un réseau par rapport à l'autre est sa performance d'apprentissage sur des ensembles de données particuliers [76].

3.3.2 Les réseaux neuronaux récurrents (de type feedback)

Appelés aussi réseaux bouclés ou dynamiques, ils possèdent des interconnexions des sorties avec les entrées. Ce sont des réseaux dans lesquels il y a retour en arrière de l'information, c'est-à-dire des connexions apparaissent entre la sortie du réseau et les neurones

se trouvant dans des couches amont. Ce type de réseau est caractérisé par la présence d'au moins une boucle de rétroaction au niveau des neurones ou entre les couches. L'architecture d'un réseau récurrent prend beaucoup de formes distinctes, la manière la plus simple consiste à réinjecter la sortie du réseau statique, à travers une ligne de retard, à son entrée [75]. Cette structure particulière est présentée à la figure 3.6. Ce modèle possède une seule entrée qui est appliquée à une ligne à retard FF-TDL (Feed Forward time Delay line) et une seule sortie qui est renvoyée à l'entrée à travers une autre ligne à retard FB-TDL (feedback time Delay line). Les sorties de ces deux lignes à retard sont utilisées pour alimenter la couche d'entrée d'un réseau statique tel que le MLP. La TDL fournit une forme simple qui donne au réseau l'habilité de modéliser une classe de systèmes dynamiques non linéaires. Ce réseau récurrent est nommé modèle non linéaire autorégressif avec entrée exogènes NARX pour (nonlinear autoregressive with exogenous) inputs [75, 80, 81].

Ainsi le comportement dynamique du modèle NARX est décrit par un vecteur appliqué à la couche d'entrée du MLP constitué des valeurs actuelles et passées de l'entrée, à savoir $x(k), x(k-1), \dots, x(k-m)$, représentant les entrées exogènes provenant de l'extérieur du réseau et les valeurs retardées de la sortie, à savoir $y(k), y(k-1), \dots, y(k-m)$.

En effet, en utilisant cette architecture, il est possible de modéliser tous les systèmes décrits par une équation de la forme :

$y(k) = F(x(k), x(k-1), \dots, x(k-m); y(k-1), \dots, y(k-m))$, avec F est une fonction non linéaire.

L'algorithme d'apprentissage récurrent en temps réel RTRL (Real-Time Recurrent Learning) est utilisé pour adapter le réseau récurrent [75,80,81]. A l'instar de l'algorithme de la rétro-propagation du gradient BP, le RTRL est également un algorithme basé sur la méthode de descente du gradient, minimisant un critère d'erreur quadratique relativement aux poids synaptiques du réseau.

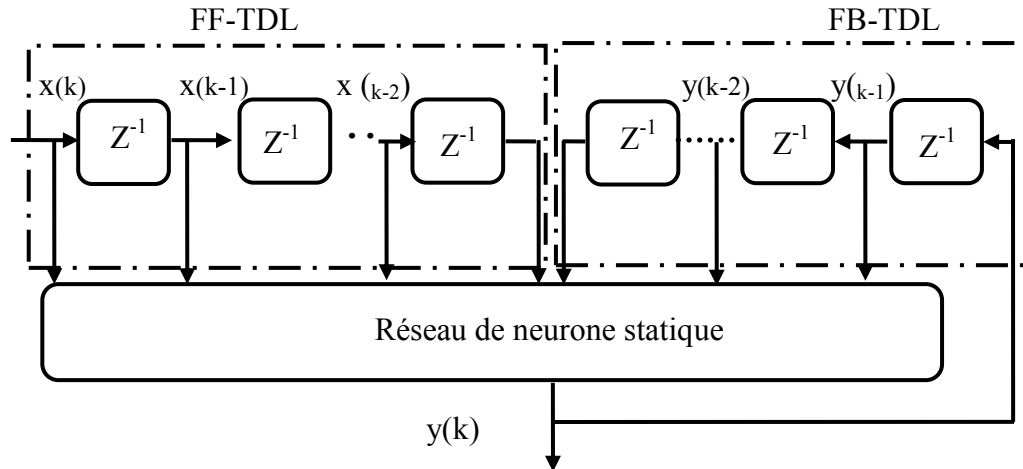


Figure 3 .6 Réseau récurant NARX

Le réseau récurant jouit aussi de la propriété d'approximation universelle pour tout système dynamique non linéaire, pourvu qu'un nombre suffisant de neurones cachés soit fourni [75]

3.4 Les filtres égaliseurs neuronaux

Les réseaux de neurones artificiels (ANN) ont suscité une grande attention en raison de leurs propriétés attrayantes, ils sont considérés comme des estimateurs universels [75,76]. Ces réseaux peuvent performer des relations complexes entre les espaces d'entrée et de sortie et ils sont capables de créer des frontières de décision non linéaire.

Les réseaux neuronaux ont été privilégiés en tant que filtres égaliseurs grâce à leur traitement non linéaire, leur capacité d'apprentissage et de généralisation et leur implémentation aisée et efficace [16, 17, 19-20, 79]. Le principal attrait des réseaux de neurone comme se trouve dans les propriétés de leurs fonctions d'activations [31, 33].

Beaucoup de recherches ont montré que les filtres égaliseurs basés sur les ANN peuvent fournir les meilleures performances relativement aux filtres conventionnels [2, 7,16-17, 19-21].

Le point clé pour une utilisation efficace de ces réseaux est de trouver une architecture adaptée au problème donnant des résultats efficaces. Les architectures des ANN en tant que filtres égaliseurs de canaux sont vraiment variées. Les connections, le nombre de couches

cachées, les fonctions d'activation, toutes ces modifications peuvent être apportées pour obtenir une nouvelle architecture.

Principalement, trois architectures sont proposées dans la littérature à savoir ; le perceptron multicouche MLP [2,16, 23-24, 82,83], les fonctions de base radiale RBF [25-27] et les réseaux récurrents RNN [28-30]. Une revue des différentes architectures des filtres égaliseurs neuronaux, leurs méthodes d'apprentissage sont présentées dans [17].

Parmi toutes ces architectures, la plus communément et largement utilisée est celle du perceptron multicouche (MLP) en raison de sa stabilité, sa structure flexible, sa paramétrisation finie et de sa simplicité de mise en œuvre [16, 23, 24, 31]. Les filtres égaliseurs basés sur les RNN sont considérés comme des filtres à réponse impulsionnelle infinies (RII), surpassent les autres filtres neuronaux ; Cependant, ils sont très instables en raison de la structure RII [17].

Les filtres égaliseurs basés sur le réseau RBF souffrent des difficultés à trouver la répartition optimale des centres, en outre, si le nombre des centres est insuffisant ou sont mal placés, les performances peuvent être considérablement affectées [24, 75]. En plus, le nombre de centres RBF nécessaire augmente avec le nombre d'entrée ce qui entraîne une charge de calcul [27].

Dans ce qui suit nous focaliserons notre étude sur le MLP en tant que filtre égaliseur nous proposerons des méthodes pour améliorer ses capacités de convergence.

3.5 Les filtres égaliseurs à base du Perceptron multicouche

Le réseau MLP incite de plus en plus l'intérêt de la communauté scientifique, il constitue le réseau neuronal le plus profondément étudié [31-34, 53, 82, 83]. Nous présenterons deux structures à base du MLP la structure transversale et la structure en treillis.

3.5.1 Le filtre égaliseur transversal à base du perceptron multicouche

Le filtre égaliseur à base du perceptron multicouche est réalisé en reliant les échantillons du signal reçu à une ligne à retard ou (TDL-Time Delay Line). Les sorties de cette dernière sont utilisées pour alimenter la couche d'entrée du réseau MLP [57,58] (figure 3.7). Ce type de réseau neuronal avec ligne à retard ou TDNN (Time Delay Neural Networks) constitue

une forme simple qui procure au réseau l'habilité de modéliser une classe de systèmes dynamiques non linéaires[17,75].

L'architecture du réseau MLP est composée d'une couche d'entrée, une couche cachée et une couche de sortie. La couche de sortie du réseau MLP est composée d'un seul neurone dont la sortie constitue la sortie du filtre égaliseur. On désigne ce filtre égaliseur par TMLP pour (Transversal Multilayer Perceptron).

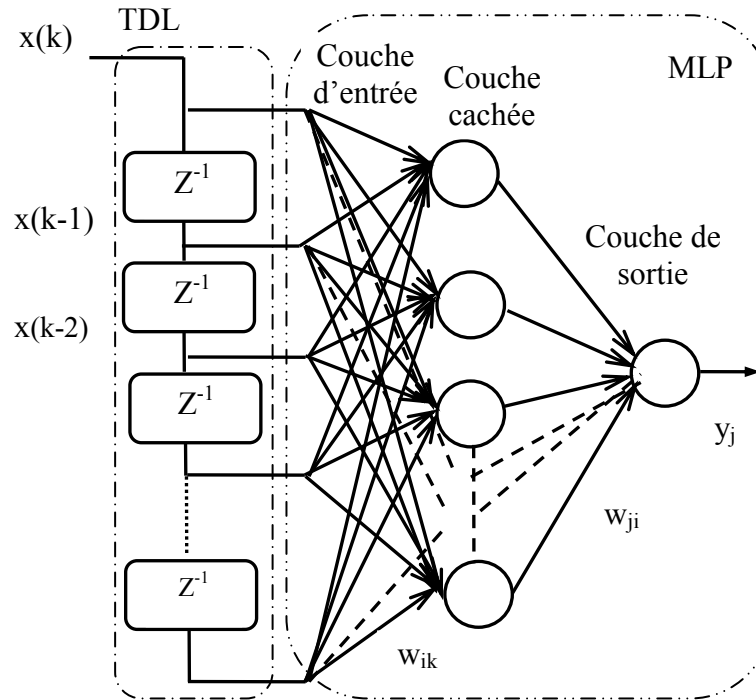


Figure 3.7 Filtre égaliseur transverse à base du perceptron multicouche (TMLP)

3.5.2 Le filtre égaliseur treillis à base du perceptron multicouche

La structure de ce type de filtre égaliseur est conçu en connectant les échantillons du signal reçu au treillis, les sorties du treillis alimentent la couche d'entrée du MLP comme dans le cas de la ligne à retard ou (TDL). La construction du treillis correspond à l'orthogonalisation de l'espace des observations. De ce fait, les observations corrélés $x(k)$ sont transformées en une séquence d'erreurs de prédiction rétrograde qui constituent les sorties du treillis alimentent la couche d'entrée du réseau MLP (figure 3.8). Nous adoptons cette structure comme méthode d'amélioration de la convergence du MLP. Notre approche constitue en une modification de celle adoptée en [16]. Le filtre égaliseur neuronal nommé

LMLP pour (Lattice Multilayer Perceptron) est formé en connectant les sorties du treillis à la couche d'entrée du MLP.

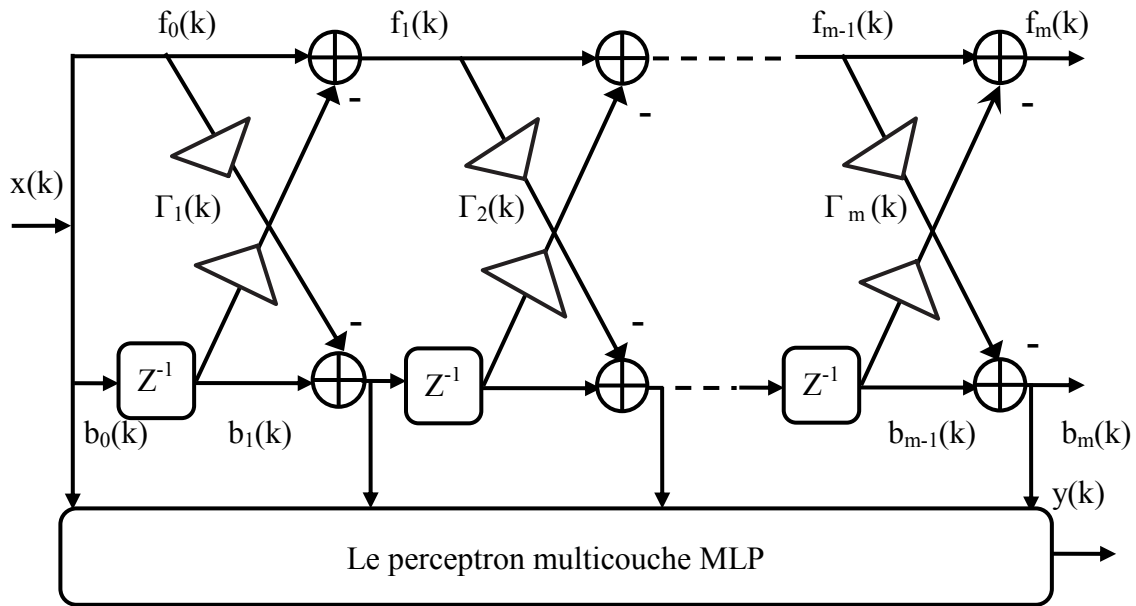


Figure 3.8 Filtre égaliseur treillis à base du perceptron multicouche (LMLP)

L'algorithme utilisé pour la structure treillis possède en général une vitesse de convergence plus rapide que son équivalent sous forme transversale [16, 67, 69]. Deux algorithmes peuvent être utilisés conjointement pour l'adaptation de ce type de filtre égaliseur. Un algorithme destiné pour ajuster les coefficients de réflexion de la structure en treillis, qui peut être un algorithme de type gradient et un algorithme pour adapter les paramètres de la structure MLP tel que l'algorithme de rétro-propagation du gradient (BP).

3.6 Algorithme d'adaptation du MLP

3.6.1 Fonctions d'activations sigmoïdales

Les fonctions d'activations jouissent d'une place importante dans les réseaux neuronaux multicouches [53, 78, 84]. Cela a conduit à l'utilisation d'une expansion de fonctions d'activations [33, 84, 85]. Etant donné qu'elles gouvernent la non linéarité du réseau et influencent directement la vitesse de convergence des algorithmes d'adaptation, une intention particulière leur été accordée [31, 33, 53, 55]. Les fonctions d'activations couramment utilisées sont de type sigmoïdal [59, 78, 86].

Une fonction sigmoïdale [84, 86, 87] $\varphi(\cdot)$ est une application $\varphi: \mathbb{R} \rightarrow \mathbb{R}$

Où $\mathbb{R}$ est l'ensemble des nombres réels, ayant les limites suivantes pour l'argument (x) tendant à $\pm \infty$:

$$\lim_{x \rightarrow +\infty} f(x) = B \quad \lim_{x \rightarrow -\infty} f(x) = A, \text{ avec } A < B \quad (3.11)$$

Les valeurs généralement considérées pour A et B sont: $A \in \{-1, 0\}$ et $B = 1$.

Ainsi, pour qu'une fonction quelconque soit considérée comme une fonction d'activation, nous devons vérifier les propriétés de continuité, de monotonie ainsi que l'existence des limites exprimées par la relation (3.11) [84].

En outre, l'existence de ces limites, pour toute fonction, affirme en plus, si la fonction est monotone et continue quelle est aussi bornée [84].

La classe générale des fonctions sigmoïdes comprend des fonctions discontinues comme la fonction échelon et la fonction signe, ainsi que des fonctions continues comme les fonctions: sigmoïde unipolaire, bipolaire et la fonction tangente hyperbolique. Dans notre travail, nous utilisons une sigmoïde bipolaire paramétrique en tant que fonction d'activation non linéaire des neurones cachés de la structure MLP, une telle fonction est définie comme suit:

$$\varphi(x) = \frac{1 - e^{(-\beta x)}}{1 + e^{(-\beta x)}} \quad (3.12)$$

Où β est le paramètre de la sigmoïde dit aussi gain, il contrôle la pente non linéaire de la sigmoïde (figure 3.9) [57]. On obtient des degrés de non linéarité variés en agissant sur ce paramètre. Lorsque $\beta = 1$, on obtient la fonction sigmoïde ordinaire non paramétrée. La dérivée $\varphi'(x)$ de la sigmoïde est donnée par:

$$\varphi'(x) = \frac{\partial \varphi(x)}{\partial x} = 2\beta \frac{e^{-\beta x}}{[1 + e^{-\beta x}]^2} = \frac{\beta}{2} (1 - (\varphi(x))^2) \quad (3.13)$$

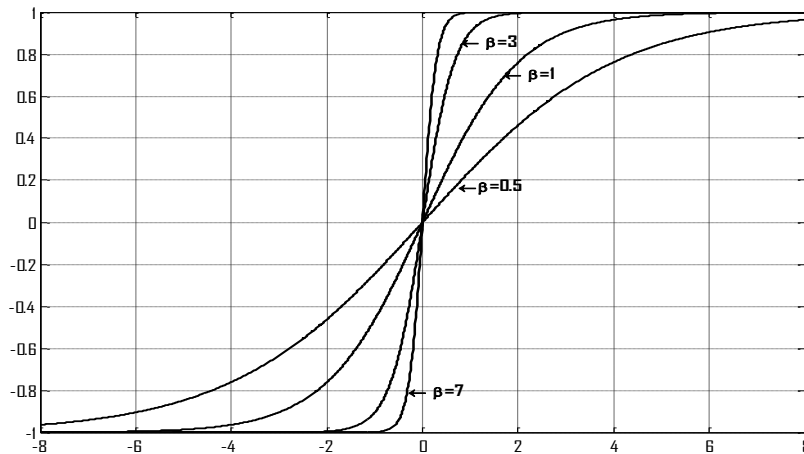


Figure 3.9 Sigmoïde bipolaire à paramètre variable

3.6.2 Algorithme de la back propagation avec fonction d'activation adaptative

L'algorithme de rétro-propagation du gradient (BP) est une méthode d'apprentissage supervisée pour le MLP [59]. Jusqu'à présent, cette procédure constitue un moyen d'apprentissage central des algorithmes pour les réseaux neuronaux [35].

Malgré son succès accentué dans de nombreuses applications et son implémentation matérielle aisée la BP possède cependant un tau de convergence lent. Des recherches ont été reportées dans la littérature afin d'accroître les performances de la BP [16, 31, 33, 45-47, 50-56]. De nombreuses améliorations des algorithmes basés sur la descente de gradient tel l'algorithme de BP sont présentés dans [35, 48, 49]. Parmi les approches proposées dans la littérature pour accroître les performances de la BP, ceux qui utilisent l'adaptation de la fonction d'activation jouent un rôle décisif [33, 53, 55,56].

Le rôle des fonctions d'activations dans l'adaptation du réseau multicouche a été bien reconnu et les algorithmes d'apprentissage utilisant une fonction d'activation adaptative conduisent à une convergence plus rapide que ceux qui utilisent une fonction d'activation fixe [31, 33, 52-54,56].

Les procédures d'ajustement des fonctions d'activation augmentent considérablement la flexibilité et la capacité d'approximation non linéaire du réseau multicouche, le processus

d'apprentissage présente une meilleure performance, l'état final du réseau sera éloigné des régions de saturation indésirables [51, 32,33].

Pour un réseau avec une fonction d'activation fixe, la forme finale de la non-linéarité est déterminée exclusivement par l'ajustement des poids des connexions. En effet, la flexibilité résultante, peut être insuffisante pour aboutir au comportement non linéaire requis. Par contre, l'utilisation d'une fonction d'activation adaptative gère le comportement non linéaire de chaque neurone et améliore d'une façon significative, la capacité non linéaire et la flexibilité du réseau MLP ce qui permet d'obtenir des meilleurs propriétés de convergence [31, 33, 51 53, 56].

Il a été pareillement confirmé qu'il existe une relation entre l'adaptation de la fonction d'activation et l'ajustement correspondant des autres paramètres du MLP, plus explicitement l'adaptation de la fonction d'activation est équivalente à l'adaptation du taux d'apprentissage et du module du vecteur de poids synaptique [45, 51,56].

Notre approche consiste à améliorer l'algorithme conventionnel de la rétropropagation du gradient, en créant une non linéarité adaptative dans sa fonction d'activation. L'algorithme est baptisé : rétro propagation du gradient avec une fonction d'activation adaptative BPAAF (back propagation with adaptive activation function) [57, 58]. L'algorithme utilise un seul paramètre adaptatif additionnel en plus des poids et des biais du MLP conservant ainsi la similitude avec la BP conventionnelle et la complexité de calcul est aussi évitée.

Nous utilisons une sigmoïde bipolaire paramétrique en tant que fonction d'activation non linéaire des neurones cachés de la structure MLP, (équation (3.12))

La sigmoïde bipolaire est paramétrée par un seul paramètre réel positif qui contrôle le degré de la non linéarité. Ainsi, l'algorithme d'apprentissage proposé (BPAAF) performe l'ajustement du gain de la sigmoïde additionnement aux poids et aux biais du réseau MLP [58]. Pour l'algorithme proposé, chaque neurone possède sa propre capacité pour ajuster le paramètre de la sigmoïde.

On considère un perceptron multicouche avec une couche d'entrée, une couche cachée et une couche de sortie (figure 3.7). Notons w_{ik} , les poids connectant les entrées à la couche cachée et w_{ji} les poids connectant la couche cachée à la sortie; avec $k = \overline{1, m}$, l'indice des entrées; $i = \overline{1, p}$, l'indice de la couche cachée et $j = \overline{1, n}$ l'indice de la couche de sortie .

Afin d'adapter le gain de la fonction d'activation ainsi que les autres paramètres du MLP, une fonction de coût, telle que la somme des erreurs quadratiques entre la sortie du réseau et la sortie désirée est introduite:

$$J = \frac{1}{2} \sum_{j=1}^n (d_j - y_j)^2 \quad (3.14)$$

On définit les entités suivantes :

$$net_i = \sum_{k=1}^m w_{ik} x_k + \theta_i \quad (3.15)$$

$$net_j = \sum_{i=1}^p w_{ji} y_i + \theta_j \quad (3.16)$$

Les sorties du neurone i de la couche cachée et celle du neurone j de la couche de sortie sont données respectivement par :

$$y_i = f(net_i) = \frac{1 - e^{-\beta_i net_i}}{1 + e^{-\beta_i net_i}} \quad (3.17)$$

$$y_j = f(net_j) = \frac{1 - e^{-\beta_j net_j}}{1 + e^{-\beta_j net_j}} \quad (3.18)$$

Pour chaque neurone, l'algorithme BPAAF effectue une descente de gradient sur la fonction de coût afin d'atteindre le minimum, en ajustant les paramètres réglables (poids, biais et gains) du MLP, comme suit:

$$w(n+1) = w(n) - \eta \frac{\partial J(n)}{\partial w(n)} \quad (3.19)$$

$$\theta(n+1) = \theta(n) - \eta \frac{\partial J(n)}{\partial \theta(n)} \quad (3.20)$$

$$\beta(n+1) = \beta(n) - \eta \frac{\partial J(n)}{\partial \beta(n)} \quad (3.21)$$

Avec η le pas d'apprentissage nommé aussi gain d'apprentissage.

En considérant le neurone de sortie, l'évaluation des dérivées partielles de la fonction de coût J relativement aux paramètres réglables, est donnée respectivement par les équations suivantes:

$$\begin{aligned}\frac{\partial J}{\partial w_{ji}} &= \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial w_{ji}} = \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} \\ &= \frac{\beta_j}{2} (y_j - d_j) (1 - y_j^2) y_i\end{aligned}\quad (3.22)$$

De la même manière, le gradient de la fonction du coût relativement au biais est exprimé comme suit :

$$\begin{aligned}\frac{\partial J}{\partial \theta_j} &= \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial \theta_j} \\ &= \frac{\beta_j}{2} (y_j - d_j) (1 - y_j^2)\end{aligned}\quad (3.23)$$

Le gradient de la fonction du coût relativement au gain de la fonction d'activation est donné par :

$$\begin{aligned}\frac{\partial J}{\partial \beta_j} &= \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial \beta_j} \\ &= (y_j - d_j) \frac{net_j}{2} (1 - y_j^2)\end{aligned}\quad (3.24)$$

Evaluons les différents gradients de la fonction de coût relativement aux paramètres du réseau, en considérons le neurone caché.

$$\frac{\partial J}{\partial w_{ik}} = \frac{\partial J}{\partial net_i} \cdot \frac{\partial net_i}{\partial w_{ik}} = \frac{\partial J}{\partial net_i} \cdot x_k \quad (3.25)$$

$$\begin{aligned}\frac{\partial J}{\partial net_i} &= \frac{\partial J}{\partial y_i} \cdot \frac{\partial y_i}{\partial net_i} \\ &= \left(\sum_{j=1}^n \frac{\partial J}{\partial net_j} \cdot \frac{\partial net_j}{\partial y_i} \right) \frac{\partial y_i}{\partial net_i} \\ &= \left(\sum_{j=1}^n \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial y_i} \right) \frac{\partial y_i}{\partial net_i}\end{aligned}\quad (3.26)$$

En exprimant chaque gradient par sa valeur correspondante on obtient :

$$\frac{\partial J}{\partial net_i} = \frac{\beta_i}{2} \left(\sum_{j=1}^n \frac{\beta_j}{2} (y_j - d_j) (1 - y_j^2) w_{ji} \right) (1 - y_i^2) \quad (3.27)$$

En substituant l'équation (3.27) dans l'équation (3.25), on obtient l'expression suivante :

$$\frac{\partial J}{\partial w_{ik}} = \frac{\beta_i}{2} \left(\sum_{j=1}^n \frac{\beta_j}{2} (y_j - d_j) (1 - y_j^2) w_{ji} \right) (1 - y_i^2) \cdot x_k \quad (3.28)$$

De la même manière, le gradient de la fonction du coût relativement au biais, est évalué comme suit :

$$\frac{\partial J}{\partial \theta_i} = \frac{\partial J}{\partial net_i} \cdot \frac{\partial net_i}{\partial \theta_i} \quad (3.29)$$

En utilisant l'équation (3.27), on obtient :

$$\frac{\partial J}{\partial \theta_i} = \frac{\beta_i}{2} \left(\sum_{j=1}^n \frac{\beta_j}{2} (y_j - d_j) (1 - y_j^2) w_{ji} \right) (1 - y_i^2) \quad (3.30)$$

En considérant le neurone cachée, le gradient de la fonction du coût relativement au gain de la fonction d'activation est donné par:

$$\frac{\partial J}{\partial \beta_i} = \frac{\partial J}{\partial y_i} \cdot \frac{\partial y_i}{\partial \beta_i} \quad (3.31)$$

En exprimant chaque gradient par sa valeur correspondante et en utilisant l'expression de $\frac{\partial J}{\partial y_i}$ évalué dans l'équation (3.26), on obtient :

$$\begin{aligned} \frac{\partial J}{\partial \beta_i} &= \left(\sum_{j=1}^n \frac{\partial J}{\partial y_j} \cdot \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial y_i} \right) \frac{\partial y_i}{\partial \beta_i} \\ &= \left(\sum_{j=1}^n \frac{\beta_j}{2} (y_j - d_j) (1 - y_j^2) w_{ji} \right) \frac{net_i}{2} (1 - y_i^2) \end{aligned} \quad (3.32)$$

L'application de l'algorithme proposé est réalisée en commençant par définir l'architecture du réseau MLP, initialiser ces paramètres à savoir les poids, les biais et les gains des fonctions d'activations. Présenter répétitivement les entrées, calculer les gradients correspondants en utilisant les équations (3.22), (3.23) et (3.24) cas du neurone de sortie ou les équations (3.28), (3.30) et (3.32) dans le cas du neurone caché. Ensuite mettre à jour les paramètres en utilisant les équations (3.19), (3.20) et (3.21) [58].

L'algorithme ainsi dérivé (BPAAF) est utilisé pour adapter le filtre égaliseur transverse à base du MLP (TMLP) de la figure 3.7. Nous utilisons aussi cette approche pour une structure treillis qui constitue une modification de celle adoptée par [16], pour accroître d'avantage, la capacité de convergence du MLP (LMLP) de la figure 3.8. L'approche développée dans [16] utilise un filtre à retour de décision à structure treillis pour la decorrélation des entrées du MLP, cette procédure a pour effet d'accroître la convergence du MLP. Le filtre treillis est adapté par l'algorithme LSL et le MLP est adapté par la BP. Dans notre approche nous utilisons un filtre de structure treillis adapté par l'algorithme de gradient adaptatif en treillis GAL et une structure MLP adaptée par l'algorithme BPAAF. Cette approche jouit des avantages de l'algorithme de BPAAF et ceux de la structure treillis.

3.7 Conclusion

Dans ce chapitre, nous avons en premier lieu, donné un aperçu général sur les réseaux de neurones ainsi que les différentes structures neuronales couramment recherchées en littérature, à savoir les réseaux MLP, RBF et RNN. L'utilisation des réseaux neuronaux en tant que filtres égaliseurs des canaux de communications est également présentée, une attention particulière a été accordée à la structure MLP. Le principal inconvénient des réseaux neuronaux est le tau d'apprentissage lent. L'utilisation de la structure treillis améliore la convergence du MLP relativement à la structure transverse. Nous avons proposé d'améliorer d'avantage la convergence du MLP à travers l'utilisation d'une structure en treillis et en créant une non linéarité adaptative dans la fonction d'activation. L'adaptation de la fonction d'activation a pour effet d'accroître considérablement les capacités du traitement non linéaire du MLP ainsi que son tau de convergence. Les performances de nos filtres égaliseurs à base du MLP seront prouvées au moyen des simulations extensives que nous présenterons au chapitre suivant.

Chapitre 4

Etude des performances des filtres égaliseurs neuronaux

4.1. Introduction

Des études en simulation, utilisant le logiciel MATLAB, permettront de valider et d'évaluer les performances des filtres égaliseurs ainsi que les approches présentées. Nous illustrons les performances des filtres égaliseurs à structure transverse et treillis à savoir: TMLP-BP, TMLP-BPAAF, LMLP-BP, LMLP-BPAAF. Les algorithmes de BP et de BPAAF sont utilisés pour l'adaptation du réseau MLP en addition de l'algorithme GAL utilisé pour adapter les coefficients du treillis. Nous exposerons les modèles et les caractéristiques des canaux de communications utilisés dans les simulations et leurs impacts sur les signaux transmis. Des mesures de performances telles que la qualité du signal restauré, l'erreur quadratique moyenne minimale (MMSE) en sortie des filtres égaliseurs ainsi que le taux d'erreur binaire BER seront utilisées. Nous évaluons nos approches dans plusieurs contextes, manifestés par des canaux de type linéaire à phase minimale et non minimale et de type non linéaire, pour un environnement stationnaire et non stationnaire.

4.2. Modèles des canaux de communications

Nous sélectionnons plusieurs types de canaux de communications, afin d'évaluer les performances des filtres égaliseurs à base du MLP à structure transverse et treillis. Deux modèles linéaires de canaux sont considérés, l'un est à phase minimale, l'autre est à phase non minimale. Deux modèles non linéaires sont aussi utilisés. Ces quatre canaux sont fixes matérialisant un environnement stationnaire. Le cinquième modèle est un canal variable, concrétisant un environnement non stationnaire.

4.2.1 Canaux linéaires

Les modèles des canaux linéaires présentés sont largement utilisés dans la littérature pour évaluer la performance des filtres égaliseurs [2, 16, 28, 30].

Les deux canaux linéaires sont décrits par leurs réponses impulsionnelles respectives $Ch_1(k)$ et $Ch_2(k)$.

$$Ch_1(k) = h_1 d(k) + h_2 d(k-1) + h_3 d(k-2) \quad (4.1)$$

$$Ch_2(k) = h'_1 d(k) + h'_2 d(k-1) + h'_3 d(k-2) \quad (4.2)$$

Avec, $h_i = [0.870 \quad 0.440 \quad 0.230]$ et $h'_i = [0.3482 \quad 0.8704 \quad 0.3482]$, où h_i ; h'_i sont les coefficients des réponses impulsionnelles respectives de Ch_1 et Ch_2 .

$i = \overline{1,3}$. Et $d(k)$ dénote l'impulsion de Dirac. Les deux canaux linéaires sont normalisés, c'est-à-dire que : $\sum_i |h_i|^2 = 1$ et $\sum_i |h'_i|^2 = 1$

En particulier le canal Ch_2 présente une distorsion sévère [28], il est substantiellement plus dur à égaliser, dû à la forte corrélation entre les amplitudes de ses parcours [71]. Les transformées en Z des réponses impulsionnelles de ces canaux sont données par :

$$Ch_1(z) = 0.870 + 0.440 z^{-1} + 0.230 z^{-2} \quad (4.3)$$

$$Ch_2(z) = 0.3482 + 0.8704 z^{-1} + 0.3482 z^{-2} \quad (4.4)$$

Les zéros de Ch_1 sont donnés par : $Z_1 = -0.2529 + 0.4477i$; $Z_2 = -0.2529 - 0.4477i$

Les zéros de Ch_2 sont donnés par : $Z'_1 = -1.9996 + 0i$; $Z'_2 = -0.5001 + 0i$

Les caractéristiques de ces canaux sont illustrées à la figure 4.1. Comme on peut voir, la partie (a) représente les réponses en amplitude des canaux considérés. Celles ci sont non linéaires et présentent des creux profonds aux différentes fréquences normalisées en raison de la propagation multi-trajet. Le canal Ch_1 présente deux creux d'environ -5 dB aux fréquences angulaires normalisées de 0,35 et 0,65. Alors que le canal Ch_2 présente un creux plus profond, d'environ -15dB, ce qui est extrêmement difficile à égaliser, ce canal résulte des interférences inter-symboles sévères. Les réponses en phase des canaux sont représentées en parties (b), Ch_1 possède une réponse en phase non linéaire alors que le canal Ch_2 présente une réponse en phase linéaire. Les zéros des canaux Ch_1 et Ch_2 , dans le plan Z sont illustrés respectivement

en (partie (c) et (d)). Les zéros de la fonction de transfert du canal Ch_1 sont localisés à l'intérieur du cercle unité, ce qui justifie que ce modèle est un canal à phase minimale (MP-Minimum Phase). Tandis que ceux du canal Ch_2 , l'un est à l'intérieur du cercle unité tandis que le second est à l'extérieur, d'où Ch_2 est un canal à phase non minimale (NMP-Nom Minimum Phase). Le zéro hors du cercle unité peut provoquer une instabilité ce qui constitue un défi pour les filtres égaliseurs.

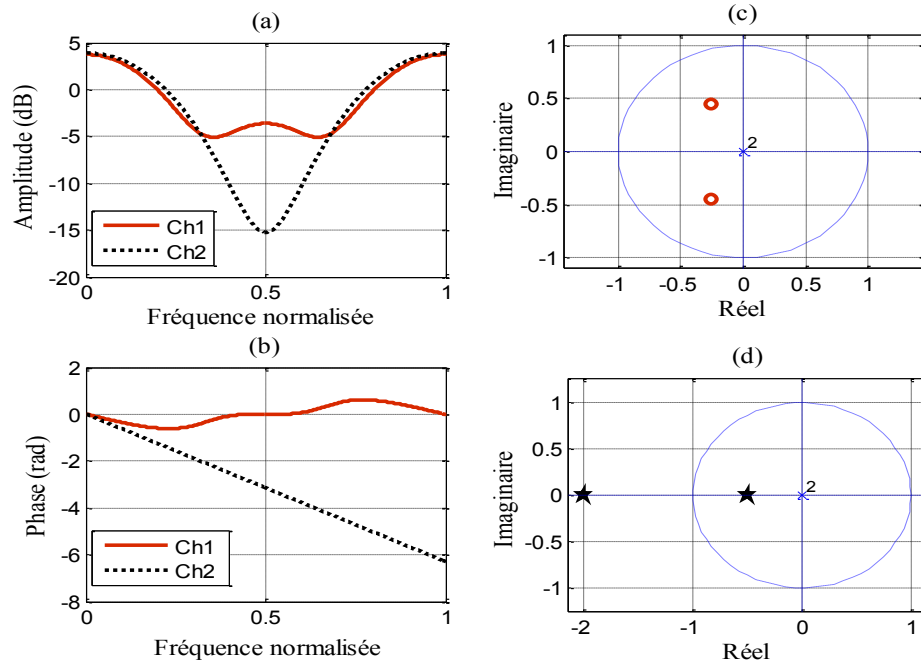


Figure 4.1. Les caractéristiques des canaux Ch_1 et Ch_2 : (a) Réponses en amplitude (b) Réponses en phase (c) Les zéros de Ch_1 (d) Les zéros de Ch_2

D'autres simulations illustrant les effets de distorsion engendrés par ces canaux ainsi que les réponses impulsionnelles et fréquentielles du canal, du filtre égaliseurs et du système formé par le canal-filtre égaliseur sont donnés aux annexes B et C respectivement. Le code MATLAB donnant ces réponses (Figure 4.1) est aussi donné à l'annexe B.

4.2.2 Canaux non linéaires

Les canaux non linéaires sont communément rencontrés dans la pratique; les distorsions non linéaires peuvent survenir en raison des phénomènes de saturation des amplificateurs [5,6], ou provenir des processus de modulation [7]. L'effet non linéaire du canal introduit des distorsions non linéaires d'amplitude et de phase. Un model d'un canal linéaire est donné à la figure 4.1.

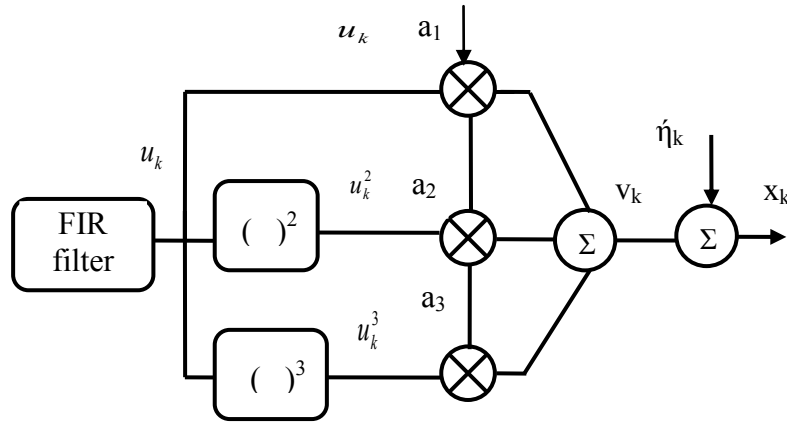


Figure 4.2 Canal non linéaire

Ce model est composé par la mise en cascade d'un canal linéaire modélisé par un filtre FIR et d'une non linéarité de type polynomiale [6, 7, 37]. La sortie $v(k)$ est liée à l'entrée $u(k)$ à travers l'équation (4.5) :

$$v_k = \sum_{i=0}^3 a_i u_k^i \quad (4.5)$$

Où u_k est la sortie du filtre RIF et v_k la sortie de la partie non linéaire. Un tel model est utilisé dans la littérature pour décrire les canaux non linéaires [28, 37]. Les coefficients a_i avec, $i = \overline{1,3}$ sont des scalaires, qui contrôlent le degré de la distorsion non linéaire.

Nous considérons deux canaux non linéaires composés d'une partie linéaire définie par les réponses impulsionnelles des équations ci-dessus (4.3) et (4.4), suivies d'une non linéarité de type polynomiale décrite précédemment par l'équation (4.5).

Les canaux non linéaires sont désignés par Nch_1 et Nch_2 . Le canal non linéaire Nch_1 , utilise Ch_1 comme partie linéaire. Alors que le canal non linéaire Nch_2 , c'est Ch_2 qui est utilisé comme partie linéaire du canal. Les paramètres contrôleurs de la non-linéarité a_1 , a_2 et a_3 de (4.5) sont définis respectivement à 1, 0,2 et -0,1, comme indiqué dans [28].

4. 3. Etude des performances des filtres égaliseurs

Pour l'étude des performances des filtres égaliseurs neuronaux, nous utilisons une structure MLP avec une seule couche cachée de 11 neurones et une seule couche de sortie d'un neurone. Les neurones dans la couche cachée possèdent chacun une fonction d'activation

de type sigmoïde bipolaire avec un paramètre adaptatif, en addition des poids et des biais. Alors que la couche de sortie avec un seul neurone utilise une sigmoïde bipolaire ordinaire non adaptative.

Les expériences de simulation sont toutes réalisées au moyen du logiciel MATLAB. La séquence numérique appliquée à l'entrée du canal est choisie à partir des nombres aléatoires bipolaires uniformément répartis sur l'alphabet $\{-1,1\}$. La sortie du canal est perturbée par un bruit blanc Gaussien de moyenne nulle et de variance σ_n^2 . Le SNR est de 20 dB. Les algorithmes de la BP et la BPAAF sont utilisés pour l'apprentissage du MLP. L'algorithme GAL est utilisé pour adapter les coefficients de réflexion du treillis. Les entrées du MLP, pour les deux structures transverse et treillis et pour les deux algorithmes sont au nombre de 5. Le pas d'adaptation pour l'algorithme GAL est $\mu = 0,015$, comme dans [73]. Le pas d'apprentissage pour les algorithmes BP et BPAAF est $\eta = 0,07$ [57,58].

Dans ce qui suit, nous présenterons les performances des filtres égaliseurs TMLP-BP, LMLP-BP, TMLP-BPAAF et LMLP-BPAAF en terme de qualité du signal restauré, la convergence MSE et du taux d'erreur par bit BER. Les capacités de convergences et de poursuites des filtres égaliseurs seront testées dans des environnements non stationnaires.

4.3.1 Qualité des signaux restaurés

On examine à travers ces simulations la qualité des signaux restaurés par les filtres égaliseurs à structure transversale TMLP et à treillis LMLP adaptés par les algorithmes de en la BP et la BPAAF. Les figures (4.3, 4.4) et (4.5, 4.6) montrent les signaux restaurés par le filtre égaliseur TMLP en considérant les canaux linéaires (Ch_1 , Ch_2) et non linéaires (Nch_1 , Nch_2) respectivement. Les parties (a) et (b) désignent, dans l'ordre, les algorithmes de BP et de BPAAF, le SNR est de 20 dB. Nous avons réduit l'échelle du signal restauré pour une comparaison adéquate avec celui transmis.

Les signaux reçus sont complètement altérés par les effets des canaux, certaines zones sont amplifiées, l'amplitude dépassent celle des signaux transmis.

D'autres zones sont atténuées, leurs amplitudes sont au voisinage du niveau zéro. Ces distorsions s'accroissent d'avantage dans le cas du canal à phase non minimale Ch_2 et des canaux non linéaires Nch_1 et Nch_2 .

Indépendamment des altérations des canaux, les signaux restaurés par le filtre égaliseur TMLP adapté par l'algorithme de BPAAF ne contiennent pas de distorsions résiduelles, ils sont identiques aux signaux transmis. Par contre, les signaux restaurés le même filtre adapté par l'algorithme de BP présentent encore certaines distorsions résiduelles. L'écart entre les signaux restaurés en considérant les algorithmes de BP et de BPAAF devient perceptible pour les canaux non linéaires Nch_1 et Nch_2 . L'algorithme de la BPAAF restaure les signaux avec succès malgré la non-linéarité des canaux; l'amplitude des distorsions est réduite au niveau requis et les fortes atténuations sont mises à niveau.

Dans le même contexte, les figures (4.7, 4.8) et (4.9, 4.10) montrent les signaux restaurés par le filtre égaliseur en treillis LMLP en considérant les canaux linéaires (Ch_1 , Ch_2) et non linéaires (Nch_1 , Nch_2) respectivement. Les parties (a) et (b) désignent, dans l'ordre, les algorithmes de BP et de BPAAF, le SNR est de 20 dB. On note que les filtres égaliseurs à structure en treillis présentent des améliorations remarquables relativement à ceux de structures transversales. La qualité des signaux restaurés par le filtre égaliseur LMLP adapté par la BP est meilleurs que celle de son équivalent transversal. Du même pour l'algorithme de BPAAF dont la qualité des signaux restaurés surpassent celle de son équivalent transversal, et celle de la BP adaptant le filtre égaliseur LMLP.

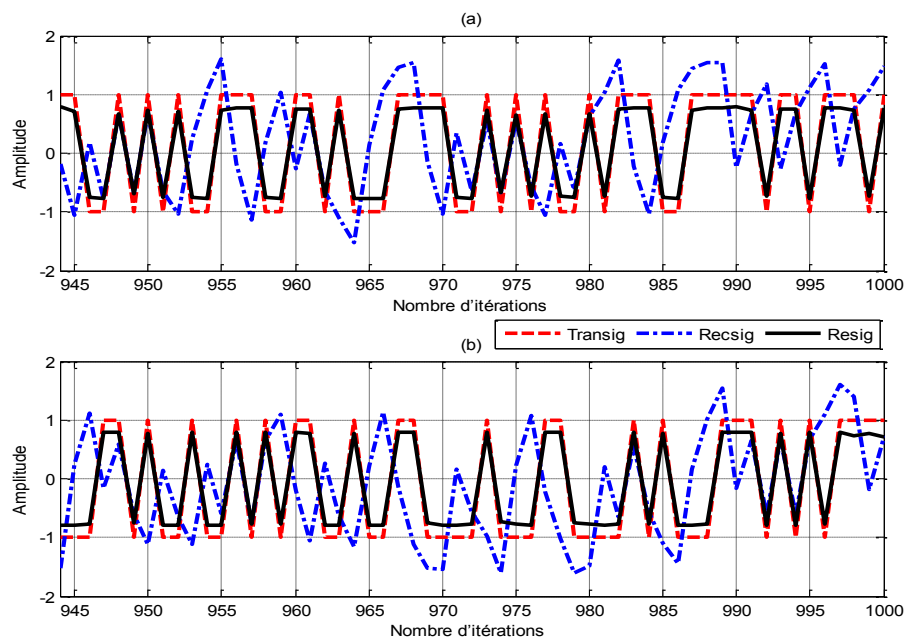


Figure 4.3. Signaux restaurés, TMLP, canal Ch₁ : (a) BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)

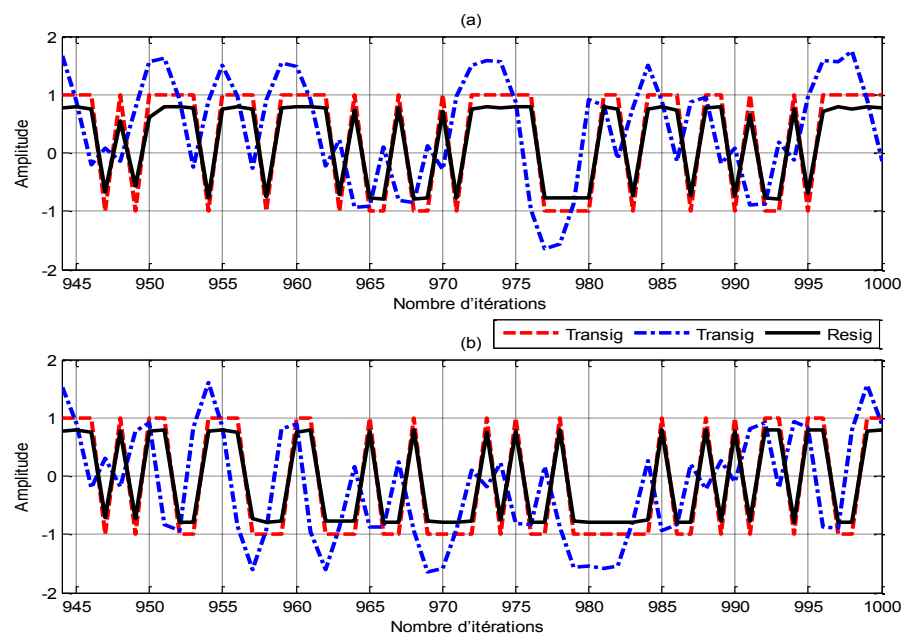


Figure 4.4. Signaux restaurés, TMLP, canal Ch₂ (a) BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)

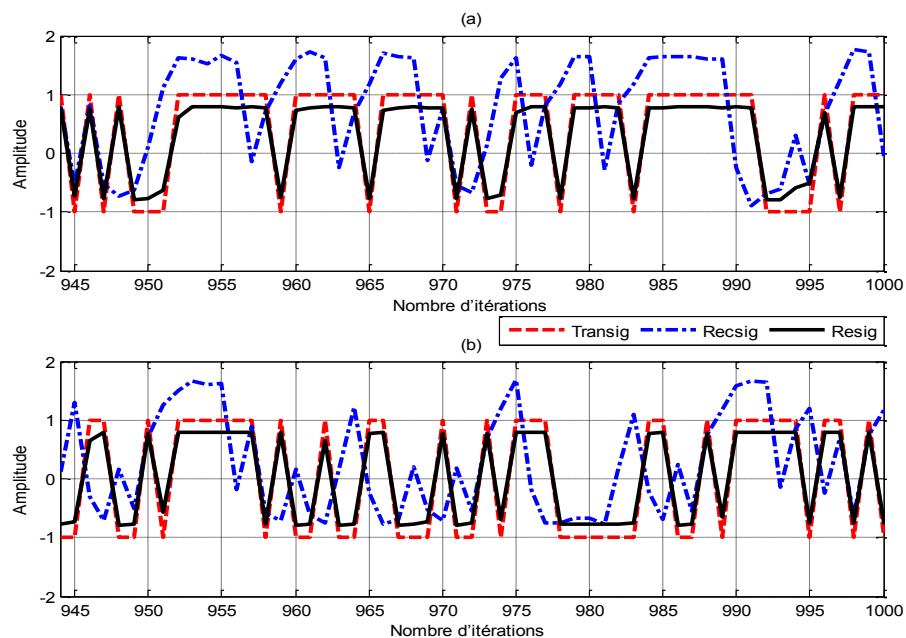


Figure 4.5. Signaux restaurés, TMLP, canal Nch₁ (a) BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)

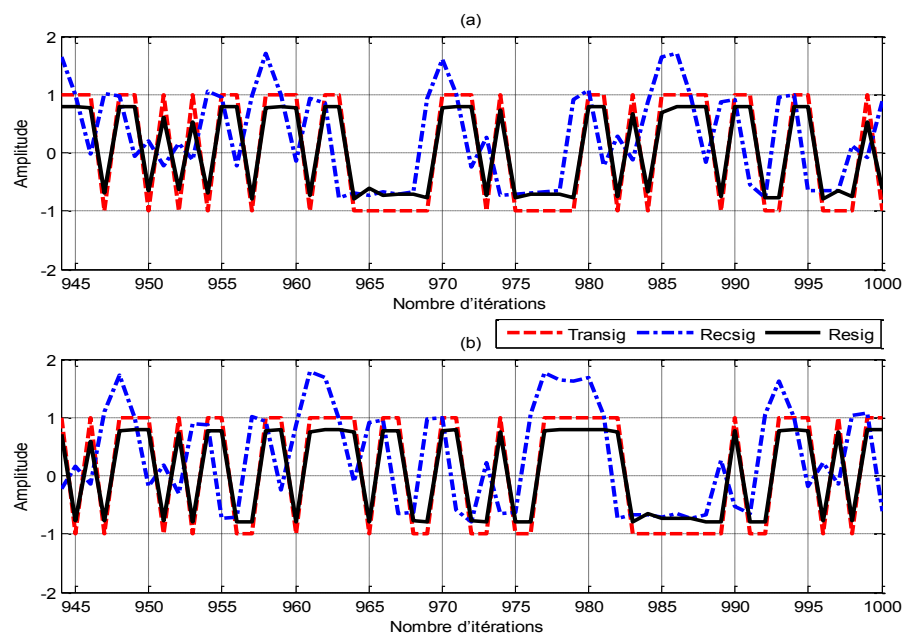


Figure 4.6 Signaux restaurés, TMLP, canal Nch₂ (a) BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)

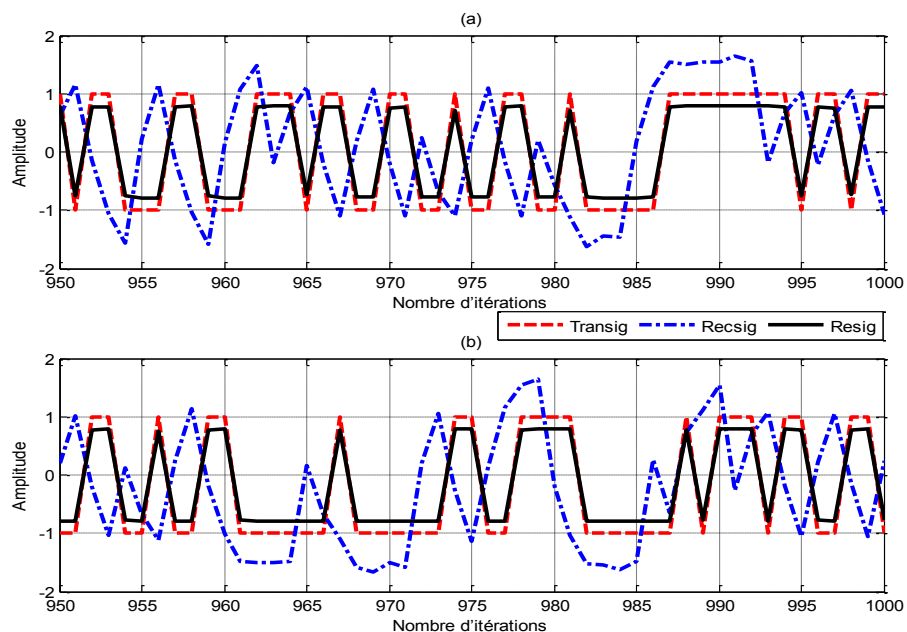


Figure 4.7. Signaux restaurés, LMLP canal Ch_1 : (a)BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)

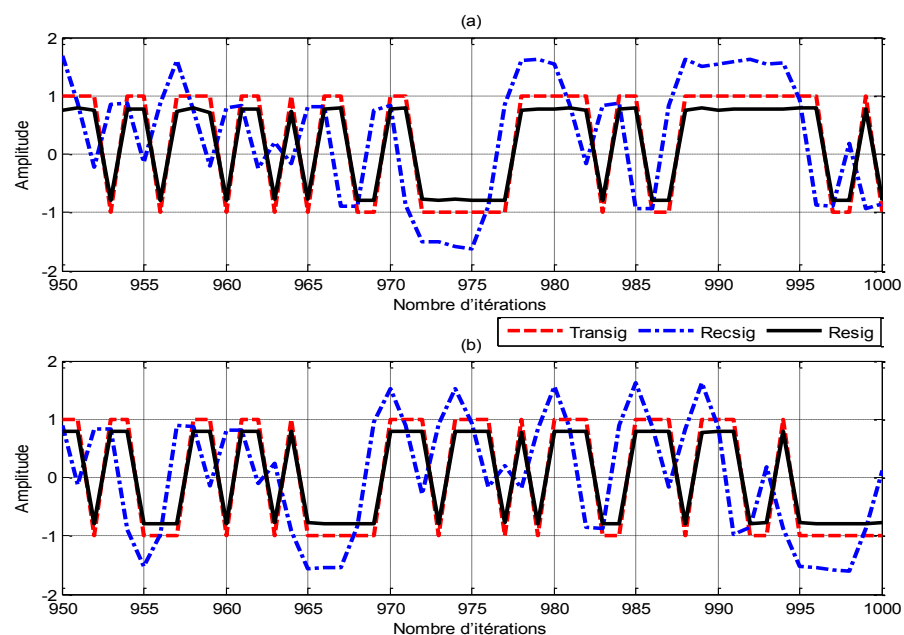


Figure 4.8. Signaux restaurés, LMLP canal Ch_2 : (a)BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)

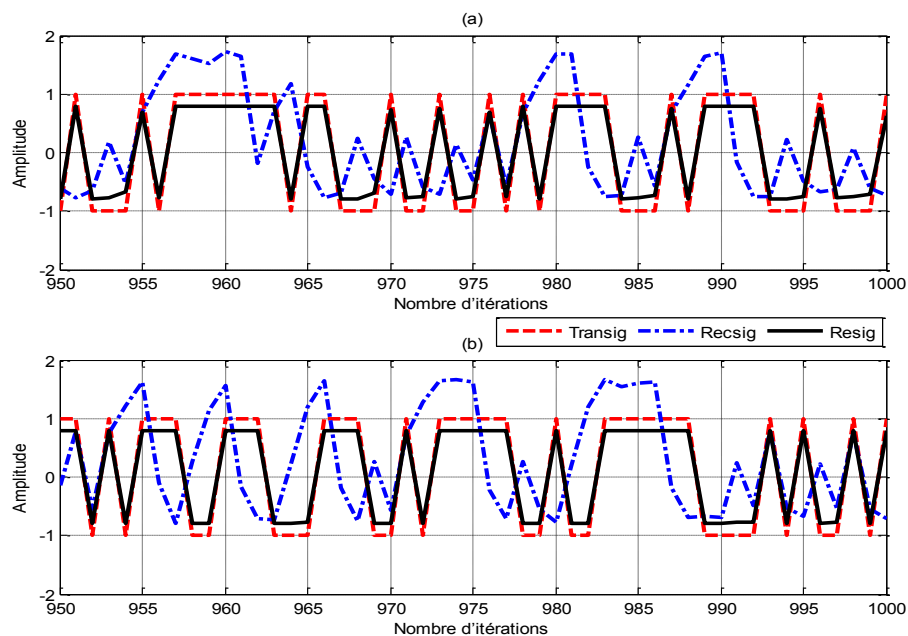


Figure 4.9. Signaux restaurés, LMLP, canal Nch₁ (a) BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)

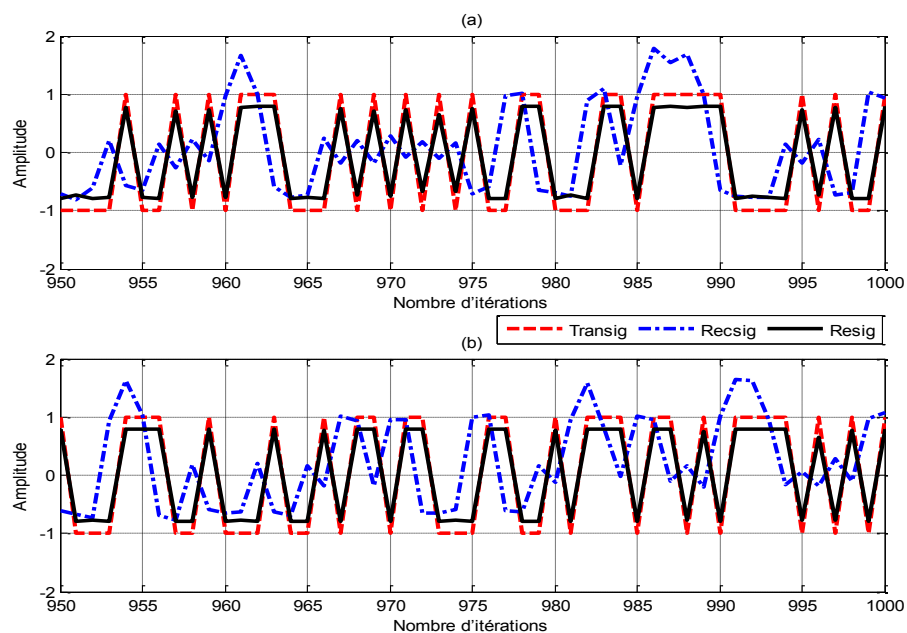


Figure 4.10. Signaux restaurés, LMLP, canal Nch₂ (a) BP, (b) BPAAF : signal transmis (Transig), signal reçu (Recsig), signal restauré (Resig)

4.3.2 Les courbes de convergence MSE

Les courbes de convergence décrivent l'évolution de l'erreur quadratique moyenne (MSE) à la sortie des filtres égaliseurs en considérant un SNR d'environ 20 dB. Les courbes MSE sont calculées en moyennant 300 simulations indépendantes, chacune avec des échantillons aléatoires différents et des paramètres initiaux aléatoires pour le MLP.

Les figures 4.11 et 4.12 illustrent les courbes de convergence des algorithmes BP et BPAAF en considérant les filtres égaliseurs TMLP et LMLP, cas des canaux linéaires Ch_1 et Ch_2 respectivement. L'algorithme de BPAAF, adaptant la structure transverse ou treillis montre une amélioration consistante en temps de convergence et en niveau MSE produit. Pour le filtre égaliseur TMLP, l'algorithme de BPAAF atteint un niveau MSE de -20dB en environ 100 itérations pour Ch_1 et 217 itérations pour Ch_2 , tandis que la BP atteint le même niveau MSE en environ 540 itérations pour Ch_1 et 1240 itérations pour Ch_2 . Dans le cas du filtre égaliseur en treillis LMLP, l'algorithme BPAAF opère encore mieux que la BP de la même structure ou la BPAAF de la structure transverse. Le niveau MSE de -20dB est atteint par la BPAAF de structure treillis au bout de 40 itérations pour Ch_1 et Ch_2 . Le même niveau est atteint par la BP en 237 itérations et 251 itérations pour les canaux Ch_1 et Ch_2 respectivement.

La BPAAF effectue également une amélioration considérable en état MSE produit relativement à la BP de même structure. La BPAAF de la structure transverse atteint un état MSE de -33,7dB et -31dB pour Ch_1 et Ch_2 respectivement. Alors que l'état MSE atteint par la BP est d'environ -27dB et -22 dB pour Ch_1 et Ch_2 respectivement.

Pour la structure treillis, la BPAAF fournit un état MSE de -36,5dB et -36dB pour Ch_1 et Ch_2 respectivement. Alors que le niveau MSE atteint par la BP est d'environ -30,6dB et -30dB pour Ch_1 et Ch_2 respectivement.

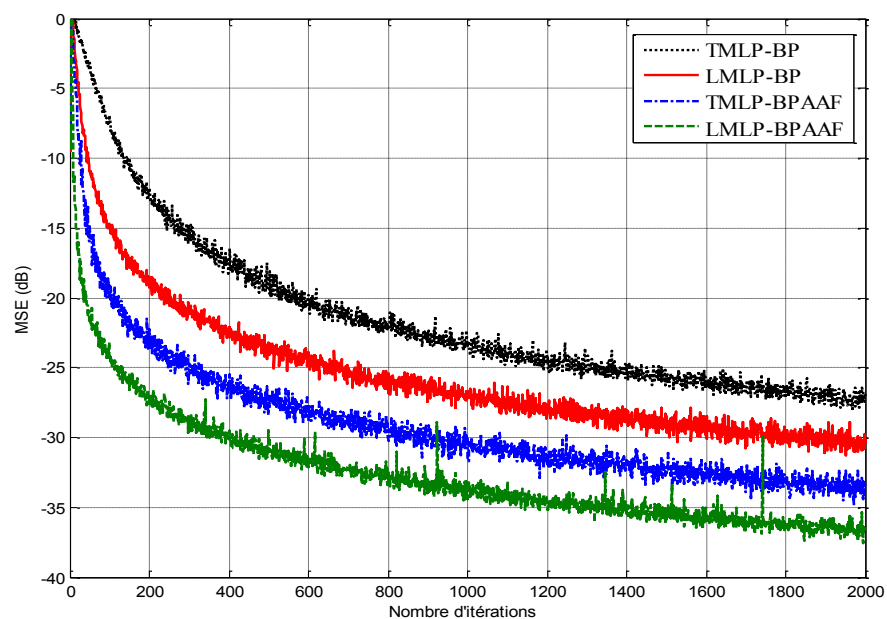


Figure4.11. Courbes de convergence des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal ch_1

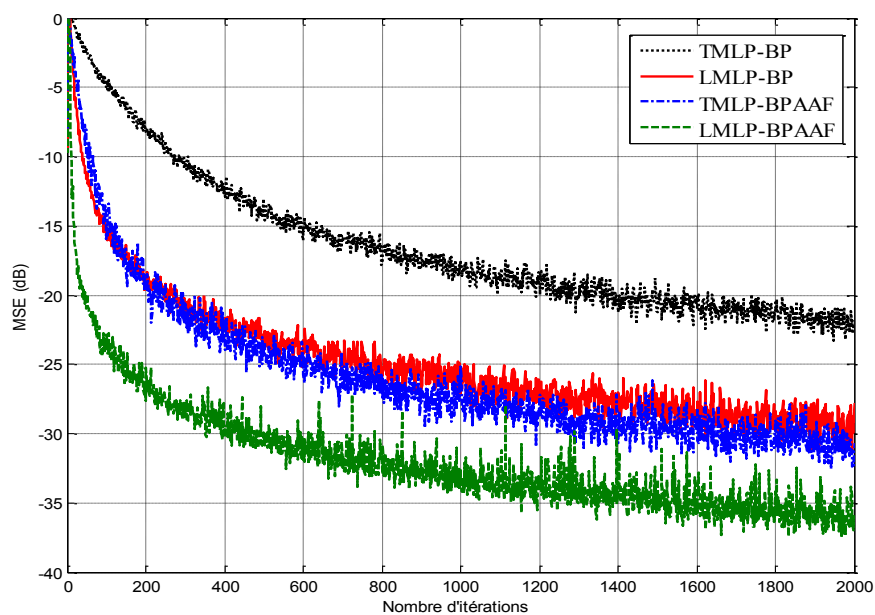


Figure4.12. Courbes de convergence des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal ch_2

Des améliorations similaires sont obtenues aux figures 4.13 et 4.14 où les canaux non linéaires Nch_1 et Nch_2 sont considérés. Concernant le filtre égaliseur transversal TMLP, l'algorithme de BPAAF atteint le niveau MSE de -20 dB à environ 104 itérations et 239 itérations, alors que la BP l'atteint à 800 itérations et 1247 itérations pour Nch_1 et Nch_2 respectivement. En considérant le filtre égaliseur en treillis LMLP, le niveau MSE de 20dB est atteint par la BPAAF en environ 47 itérations pour Nch_1 et Nch_2 . Cependant, le même niveau est atteint par la BP de même structure au bout de 331 itérations et 335 itérations pour Nch_1 et Nch_2 respectivement.

En ce qui concerne les états MSE finaux atteints par les filtres égaliseurs à structure transverse, la BPAAF atteint des niveaux MSE de -33,5dB et -30dB, par contre, la BP atteint des niveaux MSE de -25 dB et -22 dB pour Nch_1 et Nch_2 , dans l'ordre. Pour les égaliseurs en treillis, la BPAAF fournit continuellement les niveaux MSE les plus réduits en comparaison avec la BP de même structure ou la BPAAF de structure transverse. Des états MSE de -36 dB et -35,5dB sont atteints par la BPAAF, en comparaison avec des états MSE de -29dB et -28 dB fournis par la BP, pour Nch_1 et Nch_2 respectivement.

On remarque que le filtre égaliseur LMLP adapté par l'algorithme de la BPAAF est plus robuste vis-à-vis du changement de la nature du canal. Cet égaliseur fournit des états MSE similaires à savoir : -36,5 ; -36 ; -36 et -35,5 pour les canaux Ch_1 , Ch_2 , Nch_1 et Nch_2 respectivement.

On peut pareillement remarquer à partir de ces figures qu'à mesure que la sévérité du canal s'accroisse, les états MSE engendrés par les algorithmes de BP et BPAAF augmentent également. Bien que, la BPAAF conserve les états MSE les plus réduits pour les deux structures et pour tous les canaux considérés. Cela résulte de l'adaptation du paramètre de la sigmoïde qui procure plus de souplesse et de capacité non linéaire au MLP et de la structure en treillis qui offre une insensibilité au conditionnement de la matrice d'auto-corrélation du signal d'entrée.

Les performances en MSE, des filtres égaliseurs (figures 4.11-4.14) sont résumées au tableau 4.1. Nous comparons la vitesse de convergence en nombre d'itérations pour atteindre le niveau MSE de -20dB (Convergence (iter) MSE (-20dB)), par les filtres égaliseurs TMLP et LMLP. Ces derniers sont adaptés par les algorithmes de la BPAAF et la BP, pour les canaux linéaires et non linéaires.

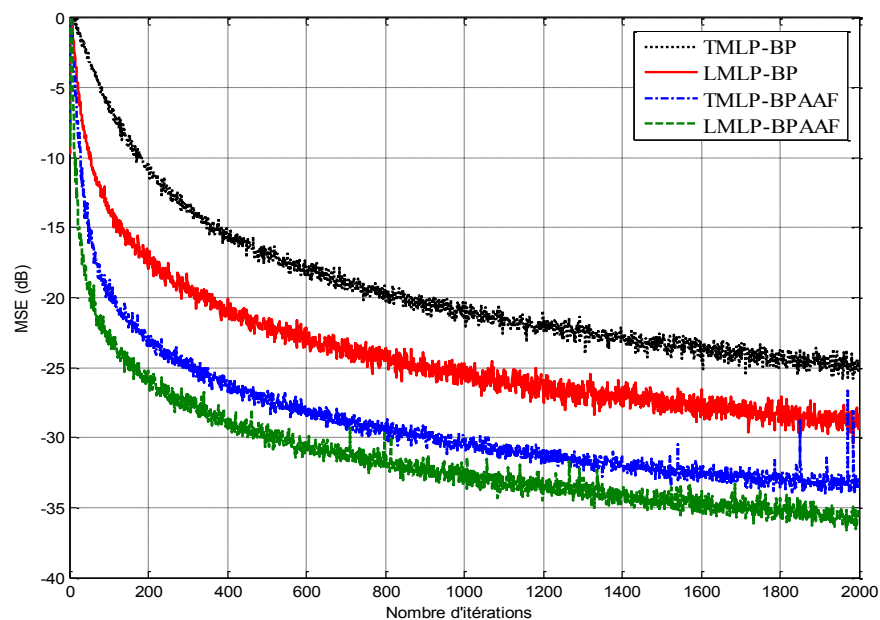


Figure 4.13. Courbes de convergence des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal Nch_1

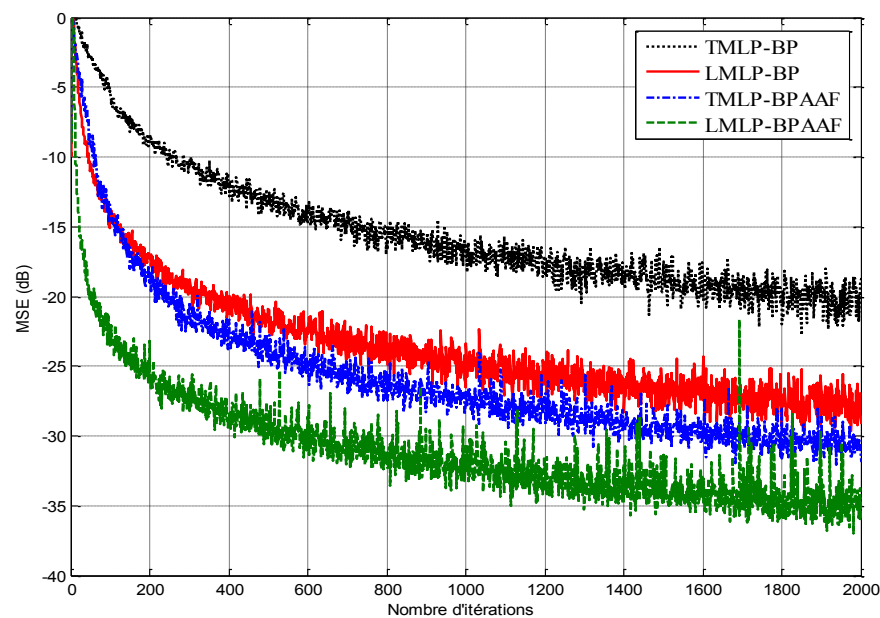


Figure 4.14. Courbes de convergence des filtres égaliseurs TMLP et LMLP, algorithmes de BP et BPAAF, canal Nch_2

Nous confrontons nos résultats pour le canal Ch_2 , avec ceux obtenus en [16], utilisant un filtre égaliseur à retour de décision DFE à base du MLP adapté par la BP(LDFE-MLP-BP). Ce filtre est de structure treillis dont les coefficients sont adaptés par l'algorithme LSL (Least Squares Lattice) au lieu de l'algorithme GAL que nous avons utilisé.

Le filtre égaliseur LDFE-MLP-BP atteint le niveau MSE de -20dB en environ 400 itérations pour Ch_2 . Tandis que les filtres égaliseurs que nous avons proposés surpassent de loin ce résultat. Le filtre égaliseurs TMMLP-BPAAF atteint ce même niveau en 217 itérations alors que le filtre égaliseur LMLP-BPAAF l'accède en 40 itérations seulement. Nous examinons aussi les niveaux MSE minimums achevés par ces filtres (MMSE (dB)). Le filtre égaliseur LDFE-MLP-BP atteint un niveau MSE minimal de -28 dB. Cependant nos filtres égaliseurs réalisent des améliorations consistantes, par apport à ce résultat, ils aboutissent à des niveaux MSE minimums de -31dB et de -36 dB, pour les filtres égaliseurs TMMLP-BPAAF et LMLP-BPAAF respectivement.

Canal/Nature	Ch1/ MP Linéaire	Ch2/ NMP Linéaire	Nch ₁ / MP Non-linéaire	Nch ₂ /NMP Non-linéaire
Convergence(iter) MSE(-20dB) TMMLP-BP	540	1240	800	1247
Convergence(iter) MSE(-20dB) TMMLP-BPAAF	100	217	104	239
Convergence(iter) MSE(-20dB) LMLP-BP	237	251	331	335
Convergence(iter) MSE(-20dB) LMLP-BPAAF	40	40	47	47
Convergence(iter) MSE(-20dB) LDFE-MLP-BP [16]	-	400	-	-
MMSE (dB) TMMLP-BP	-27	-22	-25	-22
MMSE (dB) TMMLP-BPAAF	-33,7	-31	-33,5	-30
MMSE (dB) LMLP-BP	-30	-30	-29	-28
MMSE (dB) LMLP-BPAAF	-36,5	-36	-36	-35,5
MMSE (dB) LDFE-MLP-BP [16]	-	-28	-	-

Tableau 4.1. Analyse des performances MSE des filtres égaliseurs

4.3.3 Etude du taux d'erreur par bit (BER)

Nous étudions les performances des filtres égaliseurs à travers les courbes de taux d'erreur par bit (BER), pour les canaux linéaires Ch_1 et Ch_2 (figures 4.15 et 4.16).

Le filtre égaliseur TMLP-BPAAF achève des améliorations d'environ 3dB et 2,7 dB à la valeur 10^{-2} du taux d'erreur, relativement à la BP, pour Ch_1 et Ch_2 respectivement. De même, le filtre égaliseur LMLP-BPAAF réalise, des améliorations de 3dB et 4dB à la valeur 10^{-2} du taux d'erreur, relativement à la BP, cas de Ch_1 de Ch_2 , respectivement.

Nous remarquons également que le filtre égaliseur adapté par la BPAAF accomplit continuellement le taux d'erreur le plus minimal relativement à la BP et que la BPAAF adaptant une structure treillis opère mieux que celle adaptant une structure transverse.

Le filtre égaliseur LMLP-BPAAF accomplit des taux d'erreur de $0,10 \times 10^{-3}$ et $0,23 \times 10^{-3}$, à la valeur 14 dB du rapport E_b/N_0 , alors que le filtre égaliseur LMLP-BP réalise des taux d'erreur de $0,97 \times 10^{-3}$ et $1,90 \times 10^{-3}$, cas des canaux Ch_1 et Ch_2 respectivement.

Dans le même contexte, le filtre égaliseur TMLP- BPAAF achève des taux d'erreur de $0,27 \times 10^{-3}$ et $0,73 \times 10^{-3}$, à la valeur 14 dB de E_b/N_0 , alors que le filtre égaliseur TMLP- BP réalise des taux d'erreur de $2,60 \times 10^{-3}$ et $4,0 \times 10^{-3}$, pour les canaux Ch_1 et Ch_2 respectivement.

De la même manière, en considérant les canaux non linéaires Nch_1 et Nch_2 , cas des figures 4.17 et 4.18 respectivement. Pour le filtre égaliseur TMLP, des améliorations de 2,40dB et 2.60dB sont réalisées par la BPAAF, relativement à la BP, à la valeur 10^{-2} du BER, pour Nch_1 et Nch_2 respectivement. Pour le filtre LMLP, des gains de 3,20dB et 2.70dB sont réalisés par la BPAAF relativement à la BP, à la valeur 10^{-2} du BER, pour Nch_1 et Nch_2 respectivement.

Nous constatons aussi que la BPAAF accomplit constamment le BER le plus minimal en comparaison avec la BP. Dans le cas du filtre égaliseur LMLP, la BPAAF accomplit des BER de $0,76 \times 10^{-3}$ et $1,0 \times 10^{-3}$, à la valeur 14 dB de E_b/N_0 , alors que la BP réalise des BER de $2,50 \times 10^{-3}$ et $5,90 \times 10^{-3}$, cas des canaux Nch_1 et Nch_2 respectivement.

Dans le cas du filtre égaliseur TMLP, la BPAAF achève des BER de $1,30 \times 10^{-3}$ et $2,48 \times 10^{-3}$, à la valeur E_b/N_0 de 14 dB, alors que la BP réalise des BER de $5,30 \times 10^{-3}$ et $12,0 \times 10^{-3}$.

A travers ces courbes, nous constatons que l'algorithme de BPAAF réalise les meilleures performances en taux d'erreur BER, pour les deux structures des filtres égaliseurs et pour tous les canaux étudiés. Ces résultats sont également résumés dans le tableau 4.2.

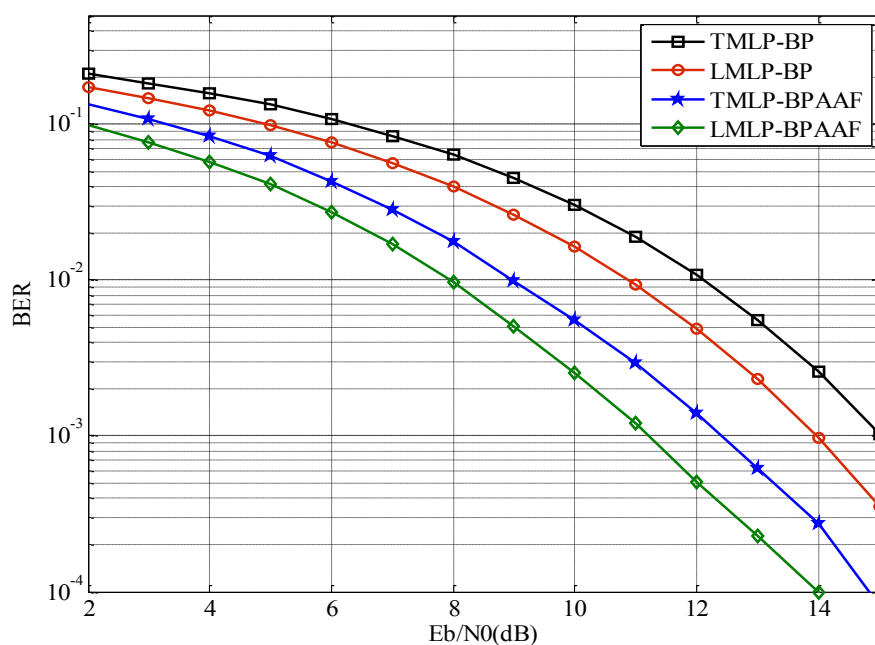


Figure 4.15. Courbes de taux d'erreur des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal ch_1

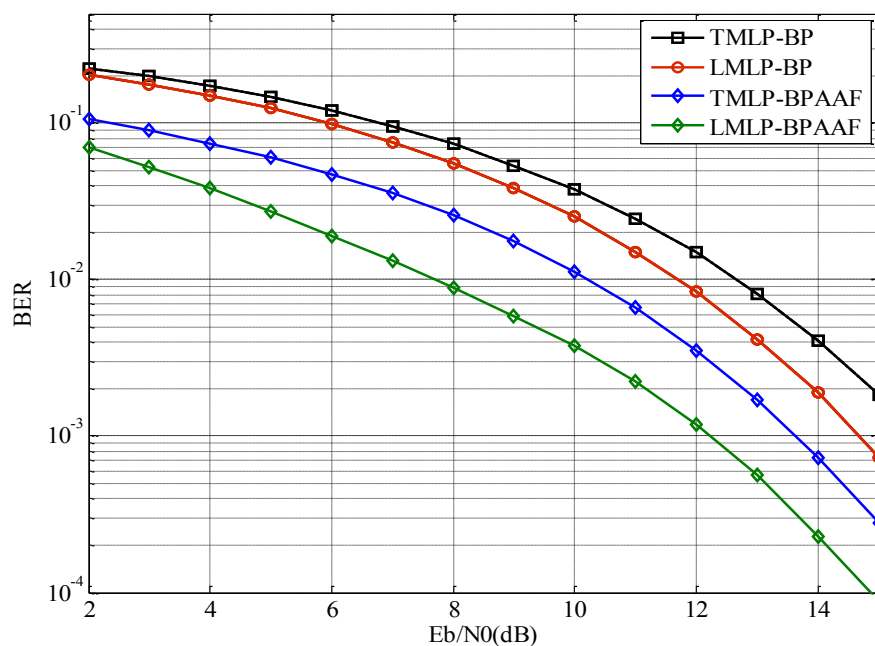


Figure 4.16. Courbes de taux d'erreur, des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal ch_2

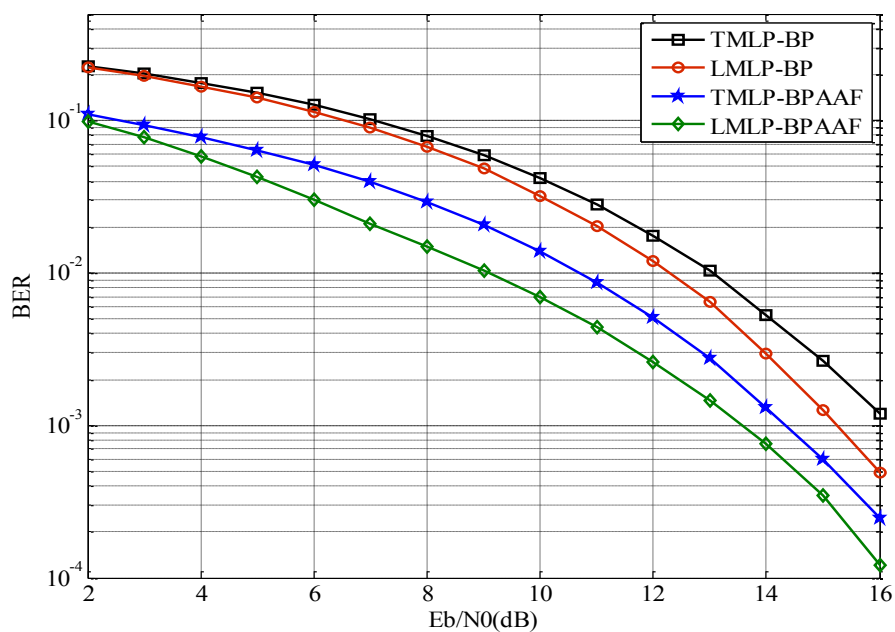


Figure 4.17. Courbes de taux d'erreur, des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal N_{ch1}

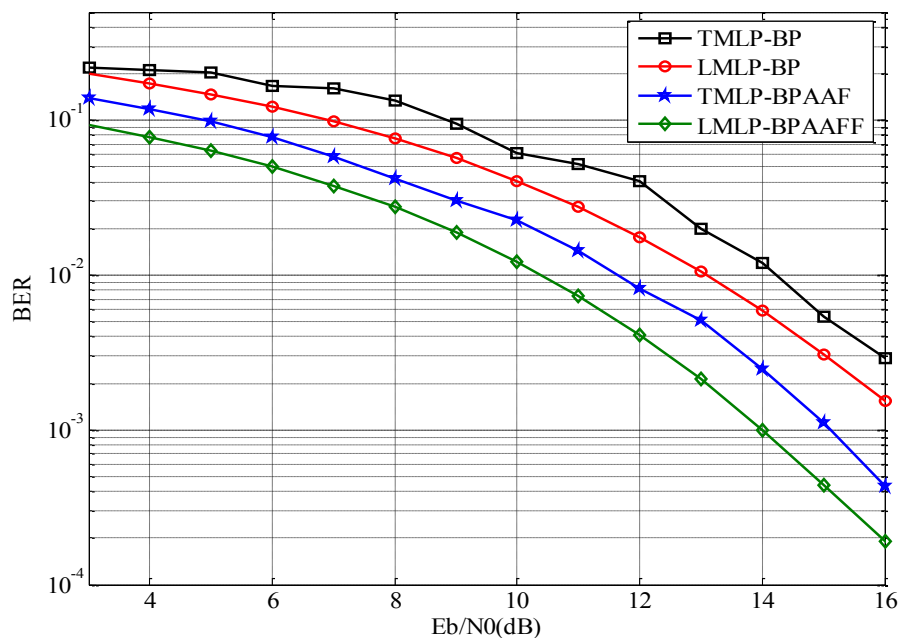


Figure 4.18. Courbes de taux d'erreur, des filtres égaliseurs TMLP et LMLP; algorithmes de BP et BPAAF, canal N_{ch2}

Le tableau 4.2 récapitule les performances en BER, présentées par les courbes 4.15-4.18. Nous comparons les taux d'erreurs achevés par les filtres égaliseurs, TMLP et LMLP adaptés par la BP et la BPAAF, à la valeur 14 dB de E_b/N_0 .

Nous comparons aussi, les performances en taux d'erreur (BER), des filtres égaliseurs proposés avec celles obtenus en [16], pour Ch₂. Le filtre égaliseur LDFE-MLP-BP réalise un BER de $0,39 \times 10^{-3}$ à la valeur 14 dB de E_b/N_0 , pour Ch₂ qui est inférieur à la valeur $0,73 \times 10^{-3}$ de BER accompli par le filtre égaliseur TMLP-BPAAF. Le filtre LMLP-BPAAF atteint la valeur $0,23 \times 10^{-3}$ de BER qui constitue une amélioration par apport au résultat réalisé dans [16]. Nous récapitulons aussi, les améliorations accomplies par la BPAAF en comparaison avec la BP de même structure pour un taux d'erreur de 10^{-2} , pour les canaux linéaires et non linéaires.

Canal/Nature	Ch1/ MP Linéaire	Ch2/ NMP Linéaire	Nch ₁ / MP Non-linéaire	Nch ₂ /NMP Non-linéaire
BER réalisé par TMLP-BP E_b/N_0 de 14dB ($\times 10^{-3}$)	2,57	4,00	5,32	11,97
BER réalisé par LMLP-BP E_b/N_0 de 14dB ($\times 10^{-3}$)	0,97	1,91	2,97	5,90
BER réalisé par TMLP-BPAAF E_b/N_0 de 14dB ($\times 10^{-3}$)	0,27	0,73	1,30	2,48
BER réalisé par LMLP-BPAAF E_b/N_0 de 14dB ($\times 10^{-3}$)	0,10	0,23	0,76	1,00
BER réalisé par LDFE-MLP-BP [16], E_b/N_0 de 14dB ($\times 10^{-3}$)	-	0,39	-	-
Gain (dB) réalisé par TMLP-BPAAF/ TMLP-BP BER de 10^{-2}	3,00	2,70	2,40	2,60
Gain (dB) réalisé par LMLP-BPAAF/LMLP-BP BER de 10^{-2}	3,00	4,00	3,20	2,70

Tableau 4.2. Analyse des performances en taux d'erreur BER des filtres égaliseurs

4.4. Étude des capacités de convergence et de poursuite des filtres égaliseurs dans un environnement non stationnaire

Dans cette phase de simulation nous présentons les capacités de convergence et de poursuite des filtres égaliseurs dans un environnement non stationnaire. Pour cela, nous considérons deux scénarios distincts : en premier lieu on procède à l'étude des performances des filtres égaliseurs transverse et treillis, adaptés par la BP et la BPAAF, en considérant un changement abrupt dans l'environnement [10, 57, 88]. Par la suite on présente les performances pour un environnement non stationnaire matérialisé par les variations des coefficients de la réponse impulsionnelle du canal [10, 16, 22, 28, 89, 90]. Ces deux scénarios sont explicités ci dessous.

4.4.1 Changement abrupt dans l'environnement

Dans cet exemple, nous procédons à créer un changement abrupt dans l'environnement qu'on matérialise par une permutation du modèle du canal. Nous examinons, la capacité des filtres égaliseurs à converger vers un état stable après introduction du nouveau model du canal.

Au cours des 1800 premières itérations, la réponse impulsionnelle du canal est choisie à être celle du canal Ch_3 , et à l'itération 1801, on passe brusquement à la réponse impulsionnelle du canal Ch_4 . Les canaux Ch_3 et Ch_4 sont respectivement à phase minimale MP et non minimale NMP. Les transformées en Z de leurs réponse impulsionnelles respectives sont décrites par les équations (4.6) et (4.7) :

$$Ch_3(z) = 0,6963 + 0,6964z^{-1} + 0,1741z^{-2} \quad (4.6)$$

$$Ch_4(z) = 0,260 + 0,930z^{-1} + 0,260z^{-2} \quad (4.7)$$

Le canal Ch_3 possède deux zéros, donnés par : $Z''_1 = -0,5061 + 0i$; $Z''_2 = -0,4941 + 0i$.

Le canal Ch_4 possède deux zéros, donnés par : $Z''_1 = -3,2712 + 0i$; $Z''_2 = 0,3057 + 0i$.

La figure 4.19, représente les courbes de convergence MSE des filtres égaliseurs, avant et après la perturbation. Ces courbes mettent en évidence, les capacités des filtres égaliseurs à s'adapter au changement abrupt dans l'environnement.

Au moment de l'introduction du changement du model du canal, les performances en MSE, des filtres égaliseurs se dégradent énormément.

Les niveaux MSE d'environ -26dB, -30dB, -33dB et -36dB, accomplis respectivement, par les filtres égaliseurs TMLP-BP, LMLP-BP, TMLP-BPAAF et LMLP-BPAAF, dans le cas du canal Ch_3 , augmentent brusquement et dépassent le niveau 0dB, en basculant vers Ch_4 .

Les filtres égaliseurs opèrent convenablement malgré le changement introduit dans l'environnement tout en convergeant vers des états stables, néanmoins avec des vitesses différentes.

Le filtre égaliseur TMLP-BP, tout en présentant une convergence moins rapide, n'est pas si bien adapté à cette situation. Contrairement aux autres filtres LMLP-BP, TMLP-BPAAF et LMLP-BPAAF qui manifestent des meilleures adaptations, en convergeant plus rapidement et en aboutissant à des états MSE plus réduit. Toute fois le filtre LMLP-BPAAF réalise la meilleure performance, en manifestant une rapidité en convergence et un état MSE minimal, en comparaison avec les autres filtres.

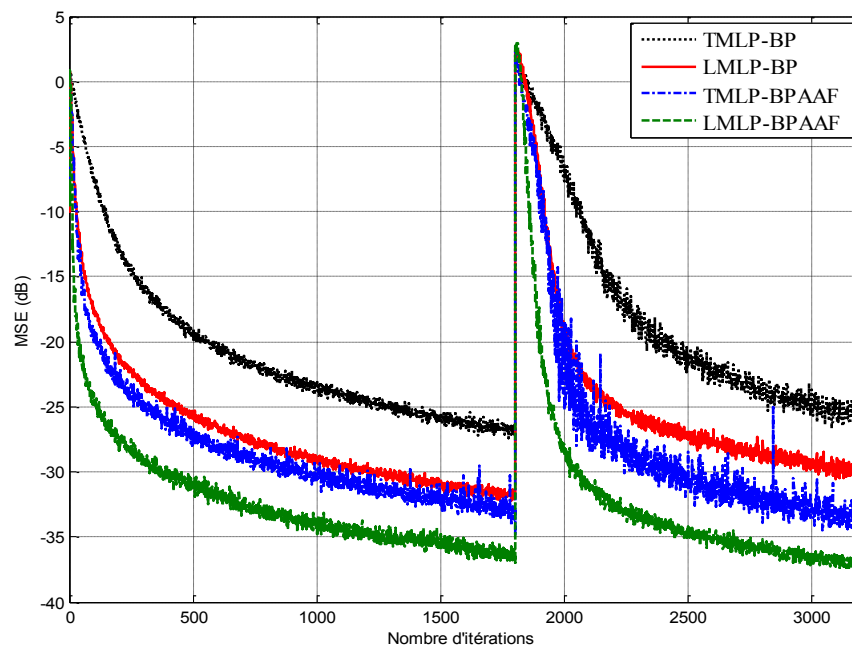


Figure 4.19 Capacité de poursuite des filtres égaliseurs sous un changement abrupt dans l'environnement

Pour mieux mettre en valeur la capacité de convergence et de poursuite des algorithmes, le canal est choisi à varier continument dans le temps, matérialisant ainsi un environnement non stationnaire. Deux environnements non stationnaires sont considérés ci-après.

4.4.2 Canal linéaire non stationnaire

Cet environnement non stationnaire est générer en faisant varier les coefficients de la réponse impulsionnelle du canal. Le modèle du canal variable est décrit par la fonction suivante :

$$Ch_v(z) = (h_1 + \alpha_1(k)) + (h_2 + \alpha_2(k)) z^{-1} + (h_3 + \alpha_3(k)) z^{-2} \quad (4.8)$$

Avec $[h_1 \ h_2 \ h_3]$, les coefficients fixe du canal Ch_1 (4.3)) et $[\alpha_1(k) \ \alpha_2(k) \ \alpha_3(k)]$, les coefficients variables dans le temps. Ceux-ci sont générés par filtrage d'un bruit blanc Gaussien au moyen d'un filtre de Butterworth de second ordre avec une fréquence de coupure normalisée. Ce modèle de canal a été proposé dans [10, 16, 22, 28, 89, 90].

L'équation (4.8) peut être réécrite d'une façon équivalente comme suit :

$$Ch_v(z) = c_1(k) + c_2(k) z^{-1} + c_3(k) z^{-2} \quad (4.9)$$

Avec $c_i(k) = h_i + \alpha_i(k)$; $i = \overline{1,3}$. Les coefficients variables sont indépendamment générés avec différents écart types du bruit. Les courbes représentant les variations des coefficients du canal sont représentées sur la figure 4.20.

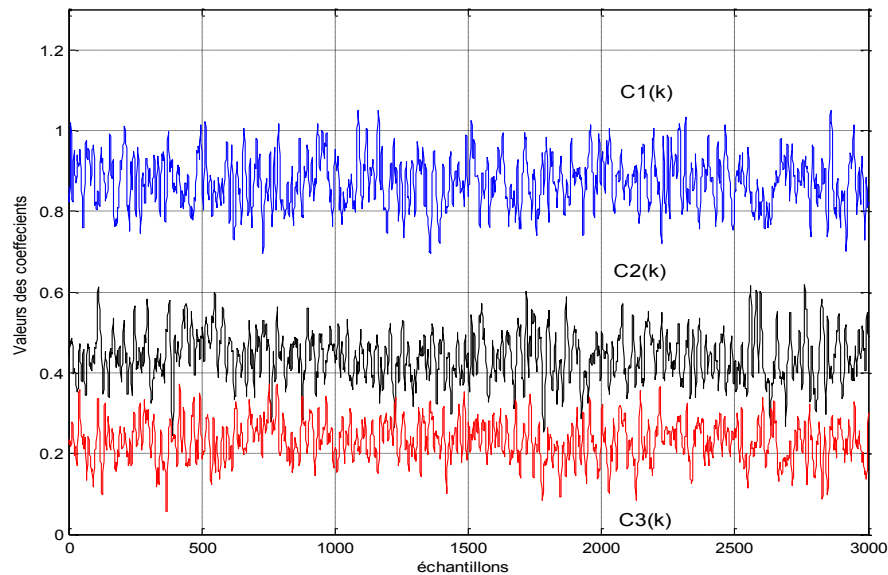


Figure 4.20 Variations des coefficients du canal linéaire

Les figures 4.21 et 4.22 montrent les courbes de convergence en MSE des filtres égaliseurs transversaux et treillis, pour le canal linéaire à coefficients variable. Nous observons des dégradations dans les performances MSE, les courbes de convergence présentent des fluctuations dues aux variations des coefficients du canal cependant ces fluctuations sont entretenues.

Nous remarquons aussi une augmentation de l'état MSE pour les deux algorithmes, toutefois, l'algorithme de la BPAAF opère mieux et présente une convergence rapide et un état MSE minimal relativement à la BP pour les deux structures.

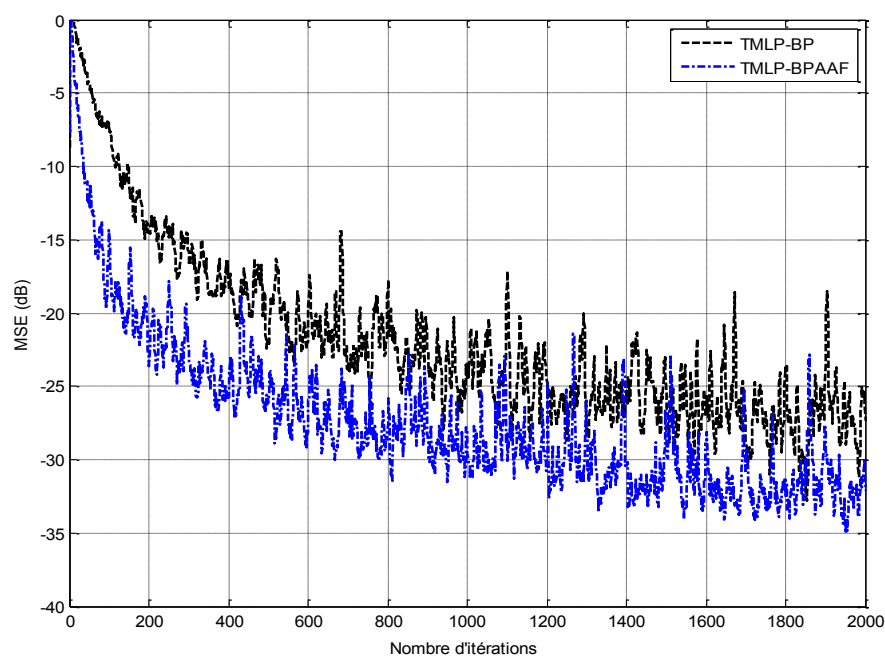


Figure 4.21 Courbes de convergence du filtre égaliseur TMLP, algorithmes de BP et BPAAF, canal linéaire variable

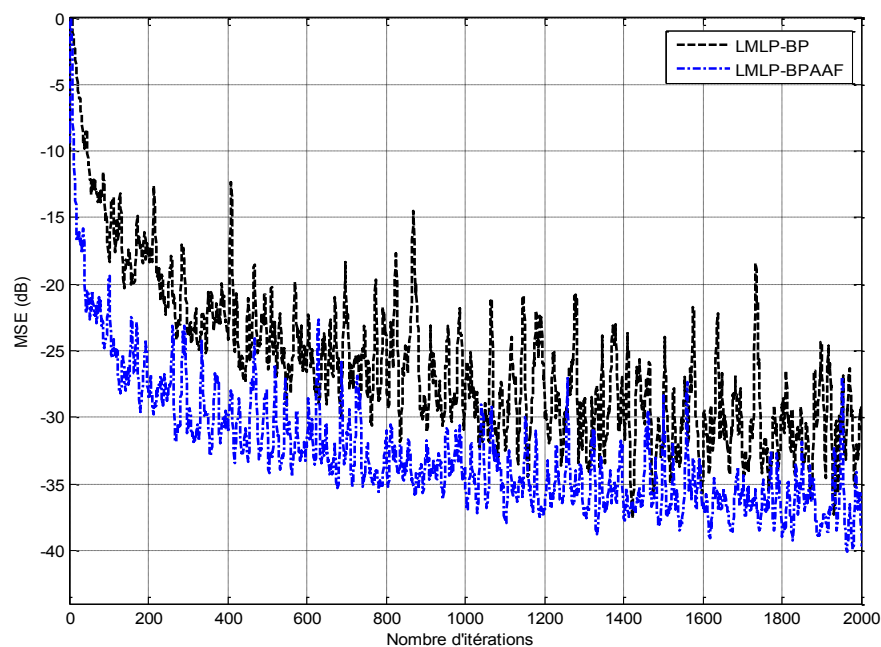


Figure 4.22 Courbes de convergence, filtre égaliseur LMLP, algorithmes de BP et BPAF, canal linéaire variable

4.4.3 Canal non linéaire non stationnaire

Cet environnement non linéaire et non stationnaire est généré en faisant passer les coefficients variables du canal linéaire (décrit précédemment par l'équation (4.9)), à travers une non linéarité de type polynomiale décrite par l'équation (4.5). Les variations des coefficients du canal non linéaire sont représentées à la figure 4.23

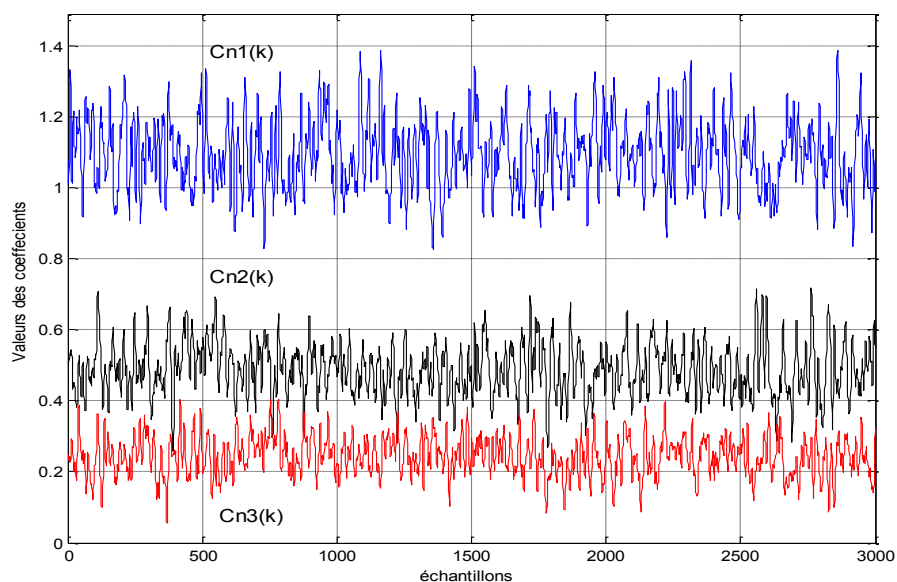


Figure 4.23 Variations des coefficients du canal non linéaire

Les figures 4.24 et 4.25 montrent les courbes de convergence en MSE des filtres égaliseurs TMLP et LMLP adaptés par les algorithmes de BP et BPAAF, pour le canal non linéaire à coefficients variables. La dégradation dans les performances en MSE des filtres égaliseurs est clairement perceptible, les fluctuations des courbes de convergence augmentent d'avantage sous les variations non linéaires des coefficients du canal.

A travers ces simulations, on s'aperçoit que l'algorithme de la BPAAF s'adapte en mieux aux différentes variations de l'environnement et fournit les meilleurs performance en terme de vitesse d'adaptation et du niveau MSE obtenu, relativement à la BP pour les deux structures.

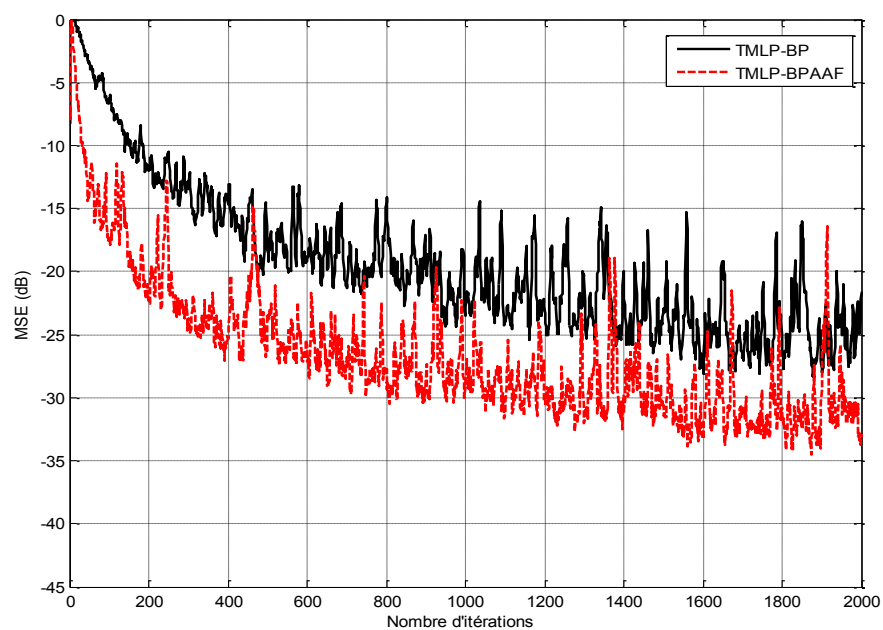


Figure 4.24 Courbes de convergence du filtre égaliseur TMLP, BP et BPAAF, canal non linéaire variable

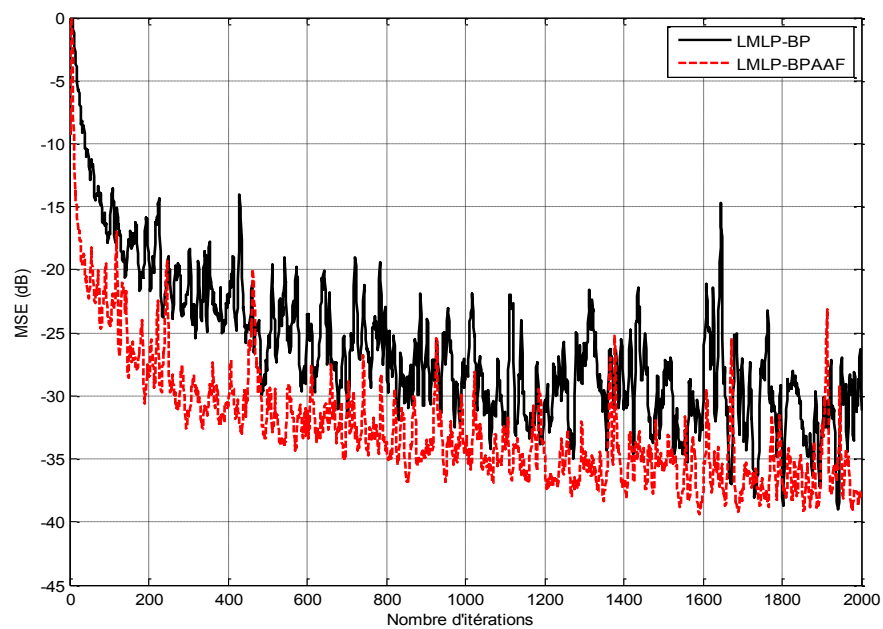


Figure 4.25 Courbes de convergence du filtre égaliseur LMLP, BP et BPAAF, canal non linéaire variable

4.5. Conclusion

A travers les résultats de simulations, nous avons met en évidence les capacités des filtres égaliseurs proposés, à structure transverse et treillis, adaptés par l'algorithme de BPAAF face à ceux adaptés par la BP. La consistance dans les performances des structures proposées est très distinctive pour tous, les canaux présentés.

Les filtres égaliseurs à base de MLP de structure treillis présentent des améliorations considérables relativement à ceux de structure transverse, pour le même algorithme d'adaptation. Néanmoins, les filtres égaliseurs adaptés par l'algorithme de BPAAF dépassent de loin, ceux adaptés par la BP pour la même structure. Le filtre égaliseur LMLP-BPAAF accomplit toujours les meilleures performances, en qualités des signaux restaurés, de vitesse de convergence, des états MSE réalisés et même des taux d'erreurs (BER) achevés. Dans le cas des environnements non stationnaires, manifestés par un changement abrupt dans l'environnement ou par les variations des coefficients des canaux linéaire et non linéaires, le filtre égaliseur LMLP-BPAAF montre une meilleure adaptation, une convergence rapide et un état stable MSE minimal relativement aux filtres égaliseurs TMLP-BPAAF, LMLP-BP et TMLP-BP.

Toutes ces améliorations sont dues à l'utilisation d'une fonction d'activation non linéaire adaptative qui procure plus de souplesse et de capacité non linéaire à la structure MLP, en plus de l'utilisation de la structure treillis qui fournit une indépendance et une insensibilité du filtre égaliseur aux conditionnements de la matrice d'autocorrélation du signal d'entrée.

Conclusion Générale

Le travail que nous avons présenté dans cette thèse porte essentiellement, sur le filtrage adaptatif, non linéaire à base des réseaux de neurones et son application à la restauration des signaux, perturbés par les canaux de communication.

Les réseaux de neurone constituent un outil puissant dans le domaine du filtrage adaptatif non linéaire, pour les communications numériques. Beaucoup de recherches se sont consacrées à développés des techniques permettant d'accélérer le taux d'apprentissage de ces réseaux.

Dans ce contexte, nous avons proposé des améliorations des capacités d'apprentissage des filtres égaliseurs adaptatifs à base du perceptron multicouche. Nos contributions se sont focalisées plus particulièrement au niveau de la structure du filtre à base du MLP et au niveau des fonctions d'activations constituant les noyaux ultimes de la structure MLP. Des recherches ont attesté que l'utilisation d'une structure treillis pour le filtre adaptatif constitue une manière d'accélérer la convergence de l'algorithme correspondant, étant donné qu'elle assure la décorrélation de l'espace des entrées. Du ce fait, nous avons utilisé un filtre de structure treillis adapté par l'algorithme GAL alimentant les entrées du MLP qui est adapté par la BP, le filtre résultant à fourni des améliorations substantielles relativement à la structure transverse.

Nous avons proposé d'adapter les fonctions d'activations des neurones cachés de la structure MLP ce qui a pour effet d'améliorer considérablement, ses capacités d'apprentissage et de convergence.

Nous avons également combiné les deux approches, à savoir, l'utilisation de la structure treillis et l'adaptation des fonctions d'activations de la couche cachée du MLP, pour bénéficier de leurs avantages respectifs.

Nous avons évalué les performances des différents filtres égaliseurs (TMLP-BP, TMLP-BPAAF, LMLP-BP et LMLP-BPAAF) de structures, transverse et treillis adaptés par les algorithmes de BP et de BPAAF, à travers des simulations intensives et des environnements d'études diversifiés. Nous avons confirmé que l'utilisation des fonctions d'activations adaptatives procure plus de souplesse et de capacité non linéaire à la structure MLP, ce qui a permis d'atteindre des améliorations substantielles, relativement à l'algorithme de BP. L'étude des performances des filtres égaliseurs, dans des environnements non stationnaires a montré la capacité de ces filtres à s'adapter aux variations induites et à converger vers des états stables. Néanmoins le filtre égaliseur de structure treillis LMLP-BPAAF a manifesté une grande capacité à poursuivre les fluctuations des canaux, linéaire et non linéaire, à coefficients variables en donnant les meilleures performances, pour toutes les expériences présentées.

Les travaux effectués dans cette thèse ouvrent plusieurs perspectives de futurs travaux. Tout d'abord il est possible de travailler sur les fonctions d'activations elles même qui peuvent être non sigmoïdale tout en respectant le théorème d'approximation universelle. Il est aussi envisageable d'utiliser d'autres techniques d'optimisations intelligentes autre que les réseaux de neurones telles que les PSO ou les lucioles. Il est toujours possible d'étendre ce travail en utilisant les modulations à grand nombre d'états. Nous avons utilisé les méthodes supervisées celle-ci utilisent une partie de la bande passante ce qui entraîne une diminution de l'efficacité spectrale et donc du débit. Ce travail peut être prolongé également sous un autre volé en considérant des algorithmes autodidactes dit (Non-data aided) basés sur les statistiques d'ordre supérieur à deux. Ces algorithmes jouissent d'une complexité remarquable cependant économisent la bande passante et permettent d'augmenter l'efficacité des systèmes de communication.

Annexes

Annexe A : L'algorithme GAL

L'algorithme GAL pour (Gradient Adaptive Lattice), constitue une extension de l'algorithme LMS pour un filtre en treillis. A l'étage m , les coefficients de réflexion de la structure en treillis sont déterminés par une méthode itérative minimisant le critère d'optimisation suivant:

$$J_m(k) = E[f_m^2(k) + b_m^2(k)] \quad (\text{A.1})$$

La méthode du gradient appliquée aux paramètres du filtre en treillis à l'étage m donne l'expression suivante :

$$\frac{\partial J_m(k)}{\partial \Gamma_m} = -2E[f_m(k)b_{m-1}(k-1) + f_{m-1}(k)b_m(k)] \quad (\text{A.2})$$

L'expression (A.2) peut être approximée comme dans le cas de l'algorithme LMS, en remplaçant les espérances mathématiques par des estimés instantanés, on obtient alors:

$$\frac{\partial J_m(k)}{\partial \Gamma_m} \approx -2[f_m(k)b_{m-1}(k-1) + f_{m-1}(k)b_m(k)] \quad (\text{A.3})$$

Nous obtenons l'équation de mise à jour pour les paramètres Γ_m

$$\Gamma_m(k+1) = \Gamma_m(k) - \frac{1}{2}\mu_m(k)\frac{\partial J_m(k)}{\partial \Gamma_m} \quad (\text{A.4})$$

En substituant l'équation (A.3) dans (A.4), on obtient l'équation de mise à jour des coefficients de réflexion.

$$\Gamma_m(k+1) = \Gamma_m(k) + \mu_m(k)[f_m(k)b_{m-1}(k-1) + f_{m-1}(k)b_m(k)] \quad (\text{A.5})$$

Dans un but de normalisation du pas d'adaptation, la valeur suivante de μ_m est adoptée :

$$\mu_m(k) = \frac{\mu}{\varepsilon_{m-1}(k)} \quad (\text{A.6})$$

Où μ représente le pas d'adaptation ordinaire, comme le cas de l'algorithme LMS et $\varepsilon_{m-1}(k)$ représente l'énergie totale des erreurs de prédiction, exprimée par :

$$\varepsilon_{m-1}(k) = E[f_{m-1}^2(k) + b_{m-1}^2(k-1)] \quad (\text{A.7})$$

L'expression itérative de $\varepsilon_{m-1}(k)$ est donnée par :

$$\varepsilon_{m-1}(k) = \varepsilon_{m-1}(k-1) + f_{m-1}^2(k) + b_{m-1}^2(k-1) \quad (\text{A.8})$$

Ou bien, en introduisant une constante positive α comprise entre ($0 < \alpha < 1$), par exemple ($\alpha=0,01$).

$$\varepsilon_{m-1}(k) = \alpha \varepsilon_{m-1}(k-1) + (1-\alpha)(f_{m-1}^2(k) + b_{m-1}^2(k-1)) \quad (\text{A.9})$$

Pour l'obtention des coefficients du filtre transverse d'une manière itérative, ce critère d'optimisation est adopté :

$$J_m(k) = E[e_m^2(k)] \quad (\text{A.10})$$

Avec :

$$e_m(k) = d(k) - y(k) \quad (\text{A.11})$$

Les coefficients g_m du filtre transverse peuvent être adaptés de la manière suivante :

$$g_m(k+1) = g_m(k) - \frac{1}{2} \mu_{m,g}(k) \frac{\partial J_m(k)}{\partial g_m} \quad (\text{A.12})$$

On obtient la mise à jour des coefficients g_m de la même manière que pour l'algorithme LMS (équation (2.17)):

$$g_m(k+1) = g_m(k) + \mu_{m,g}(k) e_m(k) b_m(k) \quad (\text{A.13})$$

Le pas d'adaptation est exprimé par :

$$\mu_{m,g}(k) = \frac{\mu}{\varepsilon_{m,b}(k)} \quad (\text{A.14})$$

Où $\varepsilon_{m,g}(k)$ représente l'énergie totale des erreurs de prédiction arrière, exprimée par :

$$\varepsilon_{m,g}(k) = E[b_m^2(k)] \quad (\text{A.15})$$

Ou encore évaluée de manière itérative comme suit:

$$\varepsilon_{m,g}(k) = (1 - \alpha) \varepsilon_{m,g}(k-1) + b_{m-1}^2(k) \quad (\text{A.16})$$

Les coefficients de réflexion de la structure en treillis et les coefficients de la structure transverse sont adaptés simultanément par les équations (A.5) et (A.13).

Notons que la convergence de l'algorithme GAL est moins dépendante du conditionnement de la matrice de corrélation du signal d'entrée.

Annexe B : Les effets de distorsions et code MATLAB

B.1 Les effets de distorsions des canaux

Dans cette partie de simulation on présente les effets de distorsions des canaux à noter : le bruit, l'interférence inter symboles (ISI) et la non linéarité.

Les figures B.1 et B.2 montrent les effets de distorsion engendrés par l'ISI, la non linéarité et le bruit pour les canaux Ch₁ et Ch₂ respectivement. La partie (a) présente le signal transmis, la partie (b) présente le signal transmis perturbé par l'effet de l'ISI, la partie (c) montre les effets de l'ISI et la non linéarité, la partie (d) montre l'effet total du canal. L'effet du canal est très sévère, certaines zones du signal sont amplifiées, l'amplitude des distorsions dépassent celle du signal, d'autre sont atténuées à proximité du zéro, ces distorsions peuvent engendrer des erreurs de décision. On s'aperçoit également que les effets de distorsions augmentent pour le canal à phase non minimale Ch₂. Au récepteur le filtre égaliseur permet de corriger ces distorsions et de restaurer le signal original.

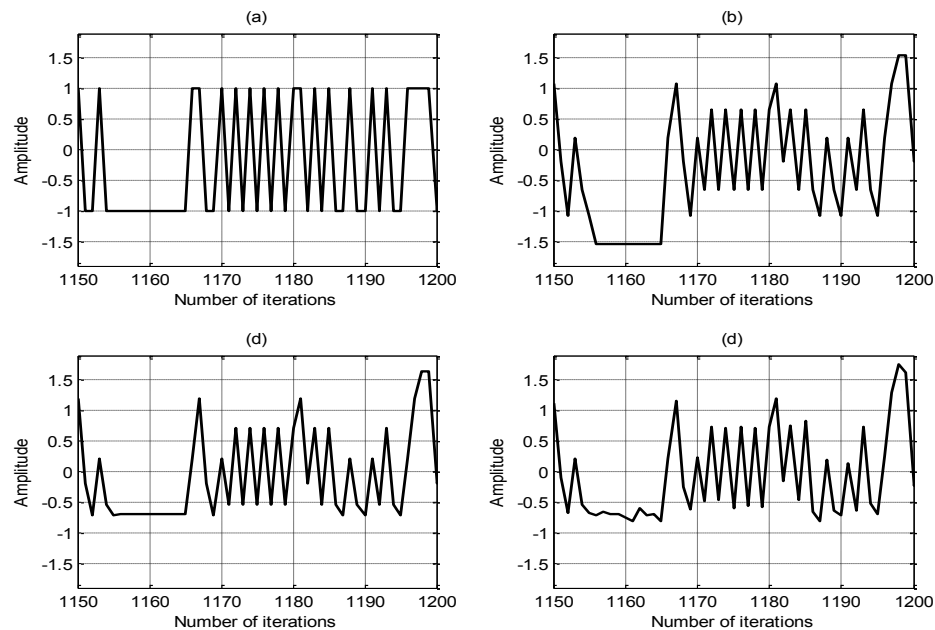


Figure B.1 Les effets de distorsions du canal Ch₁: (a) Canal idéal (b) Effet de l'ISI (c) Effets de l'ISI+NL (d) effets de l'ISI + NL+bruit, SNR 20dB

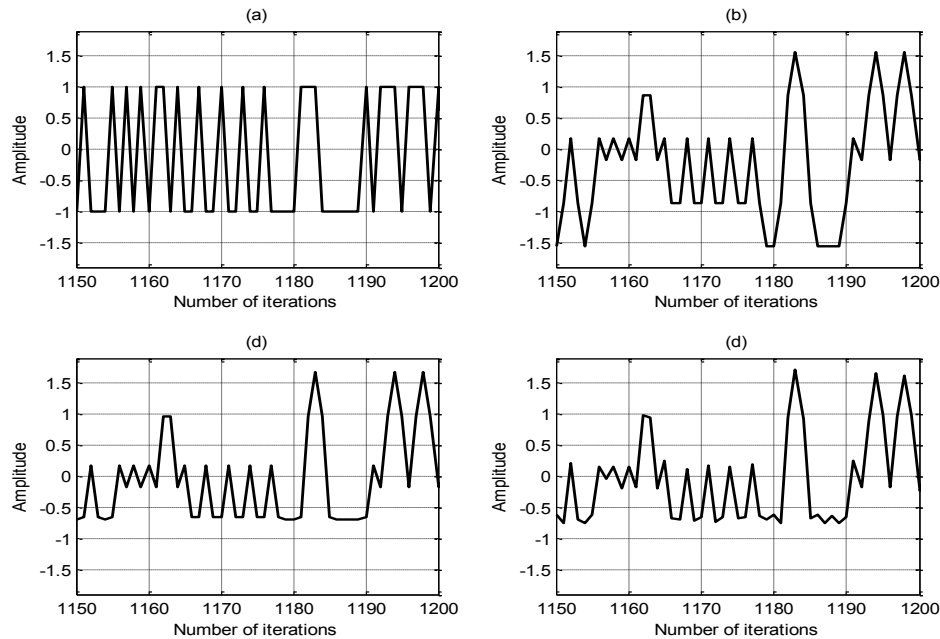


Figure B.2 Les effets de distorsion du canal Ch_2 : (a) Canal idéal (b) Effet de l'ISI (c), Effets de l'ISI+NL (d) effets de l'ISI + NL+bruit, SNR 20dB

B.2 Code MATLAB des réponses en fréquences des canaux

```
%caractéristiques des canaux utilisés en simulations
h = [.3482 .8704 .3482];% non minimum phase
r = roots(h)
N=4;w = -pi:pi/256:pi;
%*****PLAN DE Z*****
figure; subplot(224);
zplane(h), grid;hold on ;
xlabel(' Réel ');ylabel('Imaginaire '),title('Zéros du canal')
%*****Répnse impulsionnelle du canal*****
subplot(222)
y=impz(h,1),hold on,stem(y);line([0 N],[0 0]);
xlabel('n');ylabel('Amplitude');title('Réponse impulsinnelle'),
grid;
%*****Réponse fréquentielle du canal*****
h=freqz(h,1,w/pi,1); g=20*log10(abs(h));
subplot(221), plot(w/pi,g);grid;
xlabel('fréquence normalisée');ylabel('Amplitude (dB)'),
title('Réponse d\'amplitude'),hold on;
teta=unwrap(angle(h));
subplot(223),plot(w/pi,teta),grid;
xlabel('fréquence normalisée');ylabel('Phase (rad)')title('Réponse de
phase'),hold on;
```

Annexe C : Caractéristiques du canal et du filtre égaliseur linéaire

Dans cette partie, nous présentons les caractéristiques du canal, du filtre égaliseur et du système composé par le couple (canal, filtre égaliseur). Le filtre égaliseur considéré est un filtre linéaire transverse, le canal de communication considéré est défini par l'équation (4.3).

C.1 Réponses impulsionnelles

La figure C.1 montre les réponses impulsionnelles du canal, de l'égaliseur et du système (canal, filtre égaliseur). La partie (a) représente les coefficients de la réponse impulsionnelle du canal qui sont au nombre de trois coefficients, la partie (b) représente les coefficients du filtre égaliseur qui sont au nombre de sept. Les coefficients de la réponse du système (canal, filtre égaliseur) sont représentés en (c). Ils sont tous nuls sauf celui à l'origine qui est mis à 1. Nous remarquons que le système (canal, filtre égaliseur) se comporte comme un canal idéal.

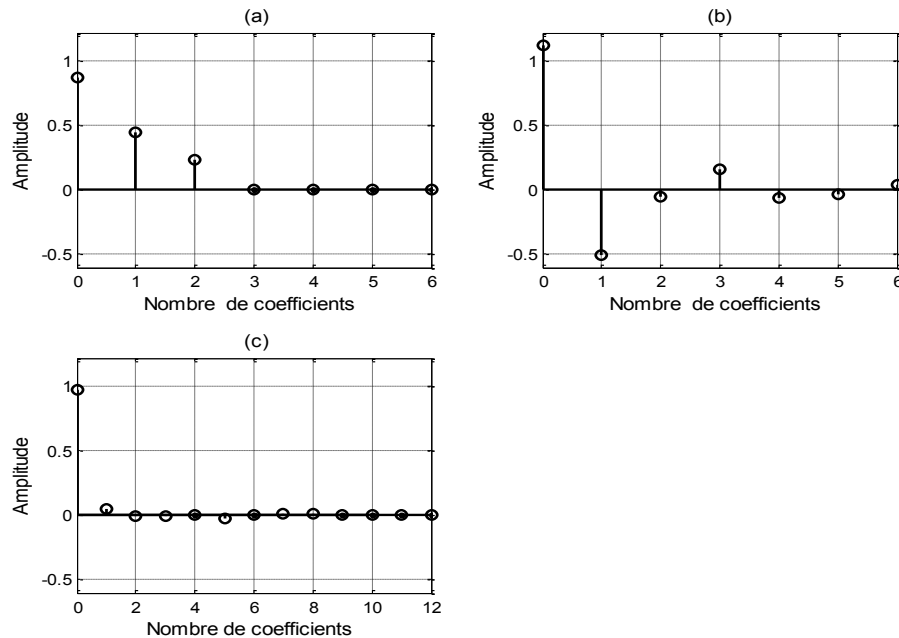


Figure C.1 Réponses impulsionnelles: (a) canal (b) filtre égaliseur
(c) système (canal, filtre égaliseur)

C.2 Réponses fréquentielles

La figure C.2 montre les réponses d'amplitudes des fonctions de transferts en fonction de la fréquence, du canal, du filtre égaliseur et du système (canal, filtre égaliseur). La partie (a) représente la réponse du canal. La réponse du filtre égaliseur dont les coefficients sont adaptés à travers l'algorithme LMS à inverser la réponse du canal (b). La partie (c) représente la réponse du système (canal, filtre égaliseur) qui se comporte comme un canal idéal.

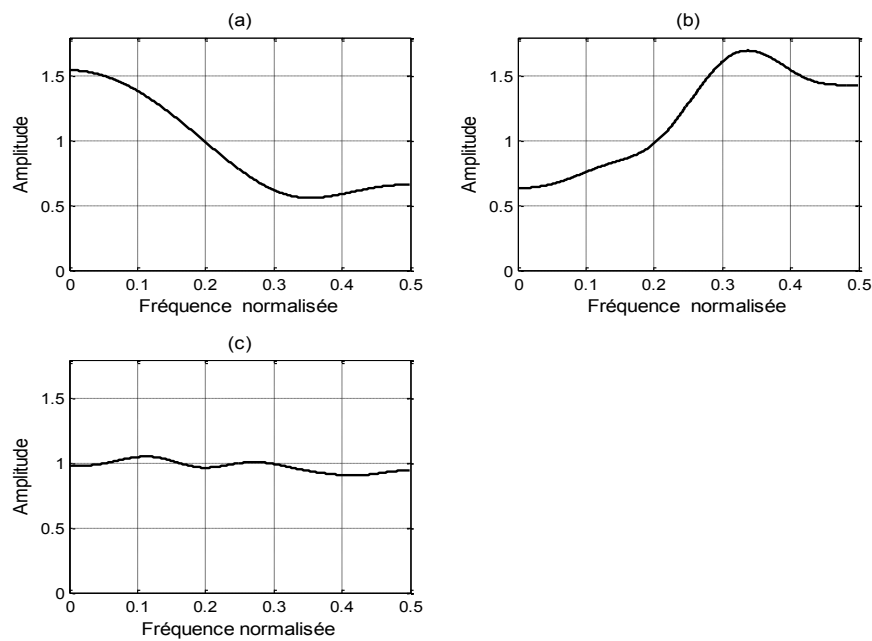


Figure C.2 Les réponses fréquentielles: (a) canal (b) filtre égaliseur
(c) système (canal, filtre égaliseur)

Bibliographie

- [1] J.G. Proakis, M. Salehi, *Digital Communications*, 5th Ed, New York: McGraw Hill, 2008.
- [2] S. Jalali, “ Wireless Channel Equalization in Digital Communication Systems,” Doctor of Philosophy in Claremont graduate University, California 2012.
- [3] Y. S. Yameogo, “ Etudes de nouvelles techniques d’estimation et d’égalisation de canal adaptées au système SC-FDMA , ” Thèse de doctorat université de rennes, 2011.
- [4] S. Qureshi, “Adaptive equalization,” Proc. IEEE, vol. 73, pp.1349–1387, 1985.
- [5] C.H.tseng, E.J.Powers “Non linear channel equalization in digital satellite systems” IEEE, pp.1639-1643, 1993.
- [6] E. Biglieri, S. Barberis, and M. Catena, “Analysis and compensation of nonlinearities in digital transmission systems,” IEEE J. Select. Areas Communication, vol.6, pp. 42-51,1988.
- [7] F. Jiuchao, Signal reconstruction with application to chaos-based communications, Doctor of philosophy Hong Kong polytechnic University, China, 2002.
- [8] Y. Liu, W. B. Mikhael, “ A Fast-converging Adaptive FIR Technique for Channel Equalization,” IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS), 2012.
- [9] H. Zhang, Y. Shi, A. Mehr et H. Huang, “Robust FIR equalization for time-varying communication channels with intermittent observations via an LMI approach, ” Signal Processing, vol. 91, pp.1651–1658, 2011.
- [10] M.U. Otaru, A. Zerguine, L. Cheded, “ Channel equalization using simplified least mean-fourth algorithm,” Digital Signal Processing, vol.21, pp. 447–465, 2011.
- [11] A.M. Grigoryan, E. R. Dougherty, S. S.Agaian, “ Optimal Wiener and homomorphic filtration : Review, ” Signal Processing, vol.121, pp.111–138, 2016.
- [12] J. Fickers, A. Ghazisaeidi, M. Salsi, G. Charlet, P. Emplit, and F. Horlin, "Decision-feedback equalization of bandwidth-constrained n-wdm coherent optical communication systems," Journal of Lightwave Technology, vol. 31, no. 10, pp. 1529-1537, 2013.
- [13] A.Youcef, “ Amélioration des performances du traitement d’interférences dans le domaine fréquentiel pour les communications acoustiques sous-marines,” Thèse de Doctorat, Université de Télécom Bretagne, 2013.

- [14] Olivier Sebe, “Déconvolution aveugle et sismologie : estimation de la source sismique et des effets de site,” Docteur de l’Université Joseph Fourier Grenoble , 2004.
- [15] M. F. Mosleh, A. H. Nakkash “Combination of LMS and RLS Adaptive Equalizer for Selective Fading Channel,” European Journal of Scientific Research Vol.43 No.1, pp.127-137, 2010.
- [16] A. Zerguine, A. Shafi, and M. Bettayeb, “Multilayer perceptron-based DFE with lattice structure,” IEEE Trans. Neural Netw, vol. 12, pp.532–545, May 2001.
- [17] K. Burse, R. N. Yadav, and S. C. Shrivastava, “Channel Equalization Using Neural Networks: A Review,” IEEE Transactions on Systems, Man, and Cybernetics-part c: Applications and reviews, vol. 40, no. 3, pp.352-357, 2010.
- [18] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," Neural Networks, vol. 2, pp. 359-366, 1989.
- [19] P.Sunita, K. M. Pradyumna, et P. Siba, “ A new training scheme for neural networks and application in non-linear channel equalization” , Applied Soft Computing , vol.27, no.1, pp. 47–52, 2015.
- [20] X.Lyu, W.Feng, R. Shi, Y. Pei, and N.Ge, “Artificial neural network-based nonlinear channel equalization: A soft-output perspective”, 22nd International Conference on Telecommunications (ICT), Pages: 243 – 248, 2015.
- [21] B. Amgothu, and G.Kalaichelvi, “Artificial Neural Network based interference mitigation through nonlinear channel equalization,” 3rd International Conference on Signal Processing, Communication and Networking (ICSCN), Pages: 1-6, 2015.
- [22] P.Corrall, O.Ludwig, and A.C. Lima, ‘Time-varying channel neural equalization using Gauss-Newton algorithm’, electronics letters, Vol. 46 No. 15, pp. 1055 – 1056, 2010.
- [23] S.Baloch, J.A.Baloch, M.A. Unar, “Channel Equalization Using Multilayer Perceptron Networks”, Research Journal of Engineering & Technology, Vol.31, No. 3, pp.469-474, 2012.
- [24] Z. Zerdoumi, D.Chicouche, and D. Benatia, “Neural Networks Based Equalizer for Signal Restoration in Digital Communication Channels,” International Letters of Chemistry, Physics and Astronomy, Vol. 55, pp 191-204, 2015.
- [25] B.Mulgrew, “Applying radial basis function networks,” IEEE Signal Process. Mag., pp. 50–65, Mar. 1996.
- [26] S. Chen, Y. Wu, and B. Luk, “ combined genetic algorithm optimisation and regularized orthogonal Lest squares learning for radial basis function networks”, IEEE Transactions on Neural Networks, vol.10, no.5, pp.1239–1245, 1999.
- [27] J. Lee, and R. Sankar, “Theoretical derivation of minimum mean square error of RBF based equalizer,” Signal Processing, vol.87, pp.1613–1625, 2007.
- [28] J. Choi, C. Lima and S. Haykin, “ Kalman Filter-Trained Recurrent Neural Equalizers for Time-Varying Channels, ” IEEE Trans on communications , vol 53, n° 3, 2005.
- [29] W. Chagra, F. Bouani, R.B. Abdenmour, M. Ksouri, and G. Favier, “ Equalization with decision delay estimation using recurrent neural networks”, Advances in Engineering Software, Vol.36, Issue 7, pp. 442-447, 2005.
- [30] H. Zhao, X.P. Zeng, Z. Y. He, W. D. Jin and T. R. Li, “Complex-Valued Pipelined Decision Feedback Recurrent Neural Network for Non-Linear Channel Equalization,” Electronics Letters, Vol. 6, No. 9, pp. 1082-1096, 2012.

- [31] S.K.Sharma and P. Chandra, “ An Adaptive Sigmoidal Activation Function Cascading Neural Networks,” 6th International Conference In Soft Computing Models in Industrial and Environmental Applications, SOCO 2011, published in, *Advances in Intelligent and Soft Computing*, vol 87. Springer, Berlin, Heidelberg .
- [32] V. E. Ismailov, “On the approximation by neural networks with bounded number of neurons in hidden layers,” *J. Math. Anal. Appl.* Vol.417, pp.963–969, 2014.
- [33] I.Castelli, E. Trentin, “ Combination of supervised and unsupervised learning for training the activation functions of neural networks, ” *Pattern Recognition Letters*, vol. 37 ,178–191, 2014.
- [34] F. Cao, D. Wang, H. Zhu ,Y. Wang , “ An iterative learning algorithm for feed forward neural networks with random weights, ” *Information Sciences*, Vol. 328 , pp.546–557, 2016.
- [35] J. Schmidhuber, “Deep learning in neural networks: An overview”, *Neural Networks*, vol.61, pp. 85–117, 2015.
- [36] Shitong Wang, Fu-LaiChung, JunWang JunWu, “ A fast learning method for feed forward neural networks, ”*Neurocomputing* vol. 149, pp.295–307, 2015.
- [37] A.T. Al-Awami, A. Zerguine, L. Cheded, A. Zidouri, W. Saif, “ A new modified particle swarm optimization algorithm for adaptive equalization,” *Digital Signal Processing* , vol.21,pp. 195–207, 2011.
- [38] S. Panda, P. K. Mohapatra, S. P. Panigrahi, “A new training scheme for neural networks and applicationin non-linear channel equalization, ” *Applied Soft Computing*, vol.27, PP.47–52, 2015.
- [39] J. Nayak, B. Naik, H.S. Behera, “A novel nature inspired firefly algorithm with higher order neural network: Performance analysis, ” *Engineering Science and Technology, an International Journal* Vol.19, N°. 1, pp. 197–211, 2016.
- [40] S. Panda, A. Sarangi, S. P. Panigrahi, “A new training strategy for neural network using shuffled-leaping algorithm and application to channel equalization, ”*Int. J. Electron. Commun. (AEÜ)* , vol .68, pp.1031–1036, 2014.
- [41] J.A. Bullinaria , K. AlYahya, “Artificial Bee Colony training of neural networks: comparison with back-propagation, ” *Memetic Computing*, Vol. 6, N°.3, pp 171–182, 2014.
- [42] S. Haykin, *Neural Networks and Learning Machines*, 3rd Ed, Pearson Prentice Hall, New Jersey 2009.
- [43] E.M. Corwin, A.M. Logar, W.J.E. Oldham, “An iterative method for training multilayer networks with threshold functions, ” *IEEE Trans. Neural Networks*, vol.5, N°.3 ,pp. 507–508,1999.
- [44] M. Negnevitsky, “ Multi-Layer Neural Networks with Improved Learning Algorithms, ” *proceedings of the IEEE conf on Digital Imaging Computing: Techniques and Applications (DICTA)*,2005.
- [45] G. Thimm, P. Moerland, E. Fiesler, “The interchangeability of learning rate and gain in backpropagation neural networks, ” *Neural Comput*, vol.8, N°.2, pp. 451–460, 1996.
- [46] A.H.Norhamreeza, , M.N. Nazri, G. Rozaida, and M .S. MNajib, “ Improvements of Back Propagation Algorithm Performance by Adaptively Changing Gain, Momentum and Learning Rate,” *International Journal on New Computer Architectures and Their Applications (IJNCAA)*, vol.1,N°.4,pp.866-878, 2011.

- [47] S. Abid, F. Fnaiech, and M. Najim, "A Fast Feedforward Training Algorithm Using a Modified Form of the Standard Backpropagation Algorithm," IEEE transactions on neural networks, vol. 12, no. 2, 2001.
- [48] M.A.W. Saduf, "Comparative Study of Back Propagation Learning Algorithms for Neural Networks," International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, N°12, pp.1151-1156, 2013.
- [49] A.D. Almási, S. Woźniak, V. Cristea, Y. Leblebici, T. Engbersen, "Review of advances in neural networks: Neural design technology stack," Neuro-computing, vol.174, pp. 31–41, 2016.
- [50] C.T. Chen, W.D. Chang, "A feed-forward neural network with function shape auto-tuning," Neural Networks, vol. 9, N°4, pp. 627–641, 1996.
- [51] L.N. Castro, L.A. Ramirez, F. Gomide, and F. J. V. Zuben, "Hybrid Tuning of Activation Functions in Feed forward Neural Networks," In Proceedings, International Joint Conference on Neural Networks (IJCNN '99), Vol. 6, pp. 4263 – 4268, 1999.
- [52] S. Xu, M. Zhang, "A novel adaptive activation function," In Proceedings International Joint Conference on Neural Networks (IJCNN '01), vol. 4, pp. 2779-2782, 2001.
- [53] G. Daqi, Y. Genxing, "Influences of variable scales and activation functions on the performances of multilayer feedforward neural networks," Pattern Recognition, vol. 36, pp. 869 – 878, 2003.
- [54] C. C. Yu, Y.C. Tang, B.D. Liu, "An adaptive activation function for multilayer feed forward neural networks", IEEE Region 10th Conference on Computers, Communications, Control and Power Engineering, TENCON '02 Proceedings, Vol. 1, pp. 645 – 650, 2002.
- [55] P. Chandra, Y. Singh, "A case for the self-adaptation of activation functions in FFANNs," Neurocomputing, vol. 56, pp. 447 – 454, 2004.
- [56] P. Chandra, Y. Singh, "An activation function adapting training algorithm for sigmoidal feedforward networks," Neurocomputing, vol. 61, pp. 429-437, 2004.
- [57] Z. Zerdoumi, D. Chicouche, D. Benatia, "Multilayer Perceptron Based Equalizer with an Improved Back Propagation Algorithm for Nonlinear Channels", International Journal of Mobile Computing and Multimedia Communications Vol. 7, N°3, pp.16-31, July-September 2016.
- [58] Z. Zerdoumi, D. Chicouche, D. Benatia, "An improved back propagation algorithm for training neural network-based equaliser for signal restoration in digital communication channels", Int. J. Mobile Network Design and Innovation, Vol. 6, No. 4, pp.236-244, 2016.
- [59] S. Hayin, *Communication systems*, 4th Ed, New York: Wiley & sons, 2001.
- [60] A. Glavieux, M. Joindot, *Communications numériques : Introduction*, Paris, Masson, 1996.
- [61] J. Guillet, "Caractérisation et modélisation spatio-temporelles du canal de propagation radioélectrique dans le contexte MIMO," Thèse de Doctorat, Institut National des Sciences Appliquées de Rennes, 2004.
- [62] O. Berder, "Optimisation et stratégies d'allocation de puissance des systèmes de transmission multi-antennes," Thèse de Doctorat, Université de Bretagne Occidentale, 2002.

- [63] S. Daumont, “ Techniques de démodulation aveugle en interception de signaux MIMO,” Thèse de doctorat, Université de Rennes 1, 2009.
- [64] E. Biglieri, G. Caire, G. Taricco, “ Coding for the fading channel: a survey Signal Processing, ” vol. 80 , pp. 1135-1148, 2000.
- [65] G.L. Stuber, *Principe of mobile communication*, 2^{ed} ed, Kluwer Academic Publishers, Boston, 2001.
- [66] M. Bellanger, *Traitement numérique du signal théorie et pratique*, 8^{eme} éd. Dunod, Paris, 2006.
- [67] Paulo S.R. Diniz, *Adaptive Filtering Algorithms and Practical Implementation*, 3rd Edition, Springer, Science and Business Media, 2008.
- [68] A. José, J. Apolinario , L.N. Sergio , *QRD-RLS Adaptive Filtering*, J.A Apolinario JR. (ed.), Springer, Science and Business Media, LLC 2009.
- [69] F. Michaut, M. Bellanger, *Filtrage adaptatif : théorie et applications*, Volume 1- *théorie et algorithmes*, version 6, Lavoisier, 26 Juin 2005.
- [70] Ibrahim Kacha, “Egaliseurs multicanaux aveugles rapides et robustes aux erreurs de surestimation de l’ordre du canal,” Thèse de Doctorat, Ecole Nationale Polytechnique, 2007.
- [71] M. ZIDANE, “ Contribution à l’identification et égalisation aveugles des canaux de transmission pour des systèmes de 4^{eme} génération,” MC-CDMA, Université Sultan Moulay Slimane Béni Mellal, Maroc 2015.
- [72] I. S. Misra, *wireless communications and Networks 3G and Beyond* ,2^{and} edition, Mc Graw Hill Education(India), 2013.
- [73] S.w. Kim, Y.c.Park and D.H.Youn, “A variable step-size gradient adaptive lattice algorithm for multiple sinusoidal interference cancelation”, EURASIP Journal on Advances in Signal Processing, pp 1-10, vol. 106, 2013.
- [74] G. Dreyfus et al, *Réseaux de neurones : Méthodologie et application*, Eyrolles, Paris 2002.
- [75] S. Haykin, *Neural Networks and Learning Machines*, 3rd Ed, Pearson Prentice Hall, New Jersey 2009.
- [76] Y. H. Hu , J-N, Hwang, *Handbook of neural network signal processing*, by CRC PRESS LLC Florida, USA, 2002.
- [77] R. Murugadoss, M. Ramakrishnan, “ Universal approximation using probabilistic neural networks with sigmoid activation functions, “ International Conference on Advances in Engineering and Technology Research (ICAETR), pp. 1- 4 , 2014.
- [78] Y. Singh, P. Chandra, “A class sigmoidal activation functions for FFANNs, ” Journal of Economic Dynamics & Control, vol. 28, pp. 183–187, 2003.
- [79] M. Ibnkahl, “Application of neural networks to digital communication”, Signal processing, vol. 80, pp 1185-1215, 2000.
- [80] T. Lin, B. Horne, P. Tino, C. Giles, “ Learning long-term dependencies in NARX recurrent neural networks, ” IEEE Transactions on Neural Networks, vol.7 N°6, pp.1329–1338, 1996.
- [81] Tsungnan Lin, Bill G. Horne, C. Lee Giles, “ How embedded memory in recurrent neural network architectures helps learning long-term temporal dependencies, ” Neural Networks Vol. 11, N°5, pp.861–868, 1998.

- [82] Z. Zerdoumi, D.Chicouche, and D. Benatia " Adaptive Decision Feedback Equalizer Based Neural Network for Nonlinear Channels, " 3rd International conference on systems and control (ICSC'13), IEEE conf, 28-31 October 2013, Hotel Hilton, Algiers, Algeria.
- [83] Z. Zerdoumi, D. Chicouche, and K. Rouabah " Channel Equalization Using Multilayer Perceptron Structure," 5th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications SETIT, March 22-26, 2009, TUNISIA.
- [84] S. SinghSodhi, P. Chandra, " Bi-modal derivative activation function for sigmoidal feed forward networks, " Neurocomputing vol. 143, pp. 182–196, 2014.
- [85] W.Ouch, N. Jankowski, "Survey of neural network transfer functions," Neural Computing Surveys, vol. 2, pp. 163-212, 1999.
- [86] D. Costarelli, R. Spigler, " Multivariate neural network operators with sigmoidal activation functions, " Neural Networks, vol. 48, pp. 72-77, 2013.
- [87] Z. Chena, F. Caob, " The approximation operators with sigmoidal functions," Computers and Mathematics with Applications , vol.58, pp. 758-765, 2009.
- [88] Q. Liang and J.M. Mendel, " Equalization of Nonlinear Time-Varying Channels Using Type-2 Fuzzy Adaptive Filters," IEEE Transactions On Fuzzy Systems, Vol. 8, No. 5, pp. 551- 563, 2000.
- [89] S. V. Vaerenbergh, M. Lázaro-Gredilla, I.Santamaría, " Kernel Recursive Least-Squares Tracker for Time-Varying Regression," IEEE Transactions On Neural Networks And Learning Systems, Vol. 23, No. 8, 2012.
- [90] J. Choi, M. Bouchard, and T. H. Yeap, " Decision Feedback Recurrent Neural Equalization With Fast Convergence Rate, " IEEE Transactions On Neural Networks, Vol. 16, No. 3, pp. 699- 708, 2005 .