



N° d'ordre : 04/2014/CED

## Université Hassan II– Casablanca

### Thèse de Doctorat

Préparée par :

**Mohammed KHALIL**

Spécialité :

**Informatique**

Sujet de la thèse :

## **Tatouage par modulations numériques pour l'enrichissement du contenu audio**

Thèse présentée et soutenue à Mohammedia le 27 septembre 2014  
devant le jury composé de :

<b>BOULMAKOUL Azedine</b>	<b>PES à la FST Mohammedia</b>	<b>Président</b>
<b>ABOUTAJDINE Driss</b>	<b>Directeur du CNRST</b>	<b>Rapporteur</b>
<b>BEKKHOUCHA Abdelkrim</b>	<b>PES à la FST Mohammedia</b>	<b>Rapporteur</b>
<b>GHOCHO Mounir</b>	<b>PES à l'université de Leeds, UK</b>	<b>Rapporteur</b>
<b>TAMTAOUI Ahmed</b>	<b>PES à l'INPT</b>	<b>Examineur</b>
<b>ADIB Abdellah</b>	<b>PES à la FST Mohammedia</b>	<b>Directeur de thèse</b>

**Etablissement** : Faculté des sciences et Techniques de Mohammedia

**CEDoc** : Sciences, Techniques, Ingénierie et Développement Durable

**Nom du laboratoire** : Laboratoire d'Informatique de Mohammedia

---

## Dédicaces

À la mémoire de mes **grands parents** que Dieu bénisse vos âmes

À ma **grand mère** qui m'accompagne par ses prières

À mes **chers parents** que Dieu vous procure longue vie

À mon **frère** à qui je souhaite beaucoup de réussite

À mes **tantes** et mes **oncles** pour leur encouragement

À mes **cousines** et mes **cousins** pour leur amour

À mes **amis** pour leur soutien

---

# Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude au professeur **Abdellah Adib**, pour m'avoir proposé ce sujet riche d'informations. J'aimerais bien le remercier pour sa disponibilité envers moi ainsi que pour ses conseils judicieux qui m'ont permis de mener à bien ce travail.

Je tiens à exprimer mes profonds remerciements au professeur **Driss Aboutajdine** pour l'honneur qu'il m'a fait en acceptant d'être rapporteur de ce travail et pour le temps qu'il m'a accordé pour m'écouter et discuter avec moi. Je tiens à remercier profondément le professeur **Abdelkrim Bekkhoucha** pour l'attention qu'il a manifestée à l'égard de cette recherche en s'engageant à être rapporteur. Je tiens à remercier sincèrement le professeur **Mounir Ghogho** pour l'intérêt qu'il a porté à ce travail ainsi que le temps qu'il a consacré pour la lecture de ce manuscrit.

Je tiens à remercier tout particulièrement le professeur **Azedine Boulmakoul** qui m'a fait l'honneur de présider le jury. Je voudrais remercier également le professeur **Ahmed Tantaoui** d'avoir examiner ce travail et être présent parmi les membre du jury.

Je souhaiterais aussi adresser mes remerciements au professeur **Mohamed Ramdani** de m'avoir accueilli au sein du laboratoire informatique.

Ces remerciements seraient incomplets si je n'en adressais pas à l'ensemble des membres du département informatique pour leur soutien moral ainsi que le cadre agréable dans lequel ils m'ont permis de préparer ce travail dans une ambiance chaleureuse et conviviale.

Enfin, j'adresse mes remerciements à tous ceux qui m'ont aidés de près ou de loin à réaliser ce travail.

---

# Résumé

Et si on vous donne la possibilité de cacher des informations supplémentaires dans un support numérique tel qu'une chanson, un extrait audio ou encore un enregistrement d'un discours sans que cela ne dégrade la qualité de ces derniers ? Puis, pourquoi ne pas exploiter ces informations pour des buts bien précis en relation avec une application de votre choix ? Cette idée aussi simple que séduisante est connue sous le nom du tatouage audio. Le principe de cette technique est d'insérer directement dans un signal audio, une marque voire une signature, contenant des informations particulières comme le titre d'une chanson ou le nom d'un artiste. Développé initialement pour la protection des droits d'auteurs, le tatouage est devenu un axe de recherche d'incessante expansion vu l'entendu de ses domaines d'applications.

Les travaux de cette thèse portent sur l'amélioration des performances des systèmes de tatouage afin d'enrichir le contenu audio par des informations diverses. Ces performances exigent des contraintes bien spécifiques. Tout d'abord, on doit insérer le tatouage de manière transparente. Ensuite, on doit pouvoir transférer le maximum d'informations possible dans le signal audio. Enfin, il faut récupérer les informations insérées avec le minimum d'erreurs possible même en présence de quelques perturbations. La motivation principale de cette thèse est l'exploration de l'apport des modulations numériques en tatouage audio. Outre les systèmes existants basés sur l'étalement de spectre, nous proposons des systèmes basés sur les modulations à porteuses sinusoïdales. Nous nous concentrons sur la modulation par déplacement de phase et celle à porteuses multiples. Ces systèmes permettent d'insérer le tatouage de manière fiable et robuste vis-à-vis des perturbations externes. En outre, nous présentons une nouvelle approche de tatouage multiple qui permet l'insertion et l'extraction de plusieurs tatouages afin d'augmenter le débit d'insertion.

Enfin, nous proposons un nouveau procédé qui va exploiter l'afficheur numérique de la radio pour transférer des informations supplémentaires. En effet, le signal radio avant sa diffusion sera tatoué automatiquement par des informations préparées au niveau de la station radio. Le récepteur radio doté d'un extracteur de tatouage permettra de récupérer les informations insérées et de les diffuser sur l'afficheur radio sous forme d'un texte lisible.

---

# Abstract

Have you ever tried to hide additional information in a digital format such as a song or an audio clip without any degradation of their quality? Then, why do not use this information for specific purposes in connection with an application of your choice? This simple idea is known as the audio watermarking. The principle of this technique is to embed directly into an audio signal a signature containing a specific information such as a song title or an artist name.

The purpose of this thesis focuses on improving the performance of watermarking systems to enrich the audio signal by various information. These performances require specific constraints. First, the watermark must be imperceptible. Then, we need to embed the maximum of information in the audio signal. Finally, we must recover the embedded information with minimum errors even in the presence of some disturbances. The main motivation of this thesis is to explore the contribution of digital modulations in audio watermarking. In addition to existing systems based on spread spectrum techniques, we present new systems based on carrier modulations; as phase shift keying and multi-carrier modulations. These systems permit to embed the watermark under the inaudibility constraint while ensuring a certain robustness against external disturbances. Furthermore, we present a new approach for audio watermarking that exploits multiple access techniques to increase the capacity of embedded information.

Finally, we propose a new service which will use the digital display on the radio to transfer additional information. At the emitter, the radio signal is watermarked by information prepared at the radio station. At the receiver, the radio equipment extracts the information from the watermarked signal and displays it in the form of a text or image.

---

# Table des matières

<b>Dedicaces</b>	<b>i</b>
<b>Remerciements</b>	<b>ii</b>
<b>Résumé</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>Table des figures</b>	<b>ix</b>
<b>Liste des tableaux</b>	<b>xii</b>
<b>Glossaire</b>	<b>xiii</b>
<b>Notations</b>	<b>xiv</b>
<b>1 Contexte général</b>	<b>1</b>
1.1 Enjeux du tatouage . . . . .	1
1.1.1 Un peu d'histoire . . . . .	1
1.1.2 Classes de tatouages . . . . .	2
1.1.3 Domaines d'applications . . . . .	3
1.1.4 Contraintes générales . . . . .	6
1.2 Le tatouage aux frontières de plusieurs disciplines . . . . .	7
1.2.1 Tatouage et cryptographie . . . . .	7
1.2.2 Tatouage et biométrie . . . . .	8
1.2.3 Tatouage et télécommunications . . . . .	9
1.2.4 Tatouage et compression . . . . .	10
1.3 État de l'art . . . . .	10
1.3.1 Classification des techniques de tatouage audio . . . . .	11
1.3.2 Domaine d'insertion du tatouage . . . . .	16
1.3.3 Techniques améliorées de tatouage par étalement de spectre . . . . .	17
1.3.4 Tatouage audio en tant que problème de communication numérique . . . . .	19
1.4 Protocole expérimental . . . . .	20
1.4.1 Signaux tests . . . . .	20
1.4.2 Inaudibilité . . . . .	21
1.4.3 Robustesse . . . . .	22

## TABLE DES MATIÈRES

---

1.4.4	Débit et coût . . . . .	23
1.5	Approche suivie dans cette thèse . . . . .	23
1.5.1	Applications envisagées . . . . .	23
1.5.2	Analyse critique de l'existant . . . . .	24
1.5.3	Contributions . . . . .	24
1.5.4	Organisation de la thèse . . . . .	25
<b>2</b>	<b>Tatouage audio comme problème de communication numérique</b>	<b>27</b>
2.1	Tatouage audio du point de vue de la communication numérique . . . . .	27
2.1.1	Modulation en bande de base . . . . .	28
2.1.2	Modulation par étalement de spectre . . . . .	29
2.1.3	Récepteur optimal . . . . .	31
2.2	Mise en forme perceptive . . . . .	34
2.2.1	Conditions psychoacoustiques . . . . .	34
2.2.2	Filtre de mise en forme . . . . .	35
2.2.3	Contrainte d'inaudibilité . . . . .	36
2.3	Récepteur aveugle : Égalisation et filtrage . . . . .	36
2.3.1	Annuleur d'interférences . . . . .	37
2.3.2	Filtre de Wiener . . . . .	38
2.3.3	Détection du tatouage . . . . .	39
2.4	Système de tatouage par étalement de spectre . . . . .	39
2.4.1	Étalement de spectre à séquence directe . . . . .	40
2.4.2	Étalement de spectre par saut de fréquence . . . . .	41
2.4.3	Comparaison entre les deux techniques . . . . .	42
<b>3</b>	<b>Tatouage audio par modulations à porteuses sinusoïdales</b>	<b>44</b>
3.1	Modulation . . . . .	44
3.1.1	Influence du type de modulation sur les performances du tatouage audio . . . . .	44
3.1.2	Modulations à bande étroite . . . . .	47
3.1.3	Comparaison entre les modulations à bande étroite . . . . .	49
3.2	Tatouage par modulation de phase . . . . .	50
3.2.1	Modulation de phase binaire . . . . .	50
3.2.2	Contrôle d'inaudibilité du tatouage PSK . . . . .	51
3.2.3	Impact de la fréquence porteuse sur les performances du tatouage . . . . .	53
3.3	Tatouage par modulation à porteuse sinusoïdale informée . . . . .	56
3.3.1	Algorithme adaptatif d'insertion du tatouage . . . . .	57
3.3.2	Algorithme d'extraction du tatouage . . . . .	59
3.3.3	Performances . . . . .	60
3.4	Tatouage par modulation multi-porteuses . . . . .	63
3.4.1	Motivations envers les modulations multi-porteuses . . . . .	63
3.4.2	Système de tatouage multi-porteuses . . . . .	65
3.4.3	Insertion d'un message texte . . . . .	72
<b>4</b>	<b>Approche multi-tatouage basée sur les techniques d'accès multiples</b>	<b>74</b>
4.1	Système multi-tatouage à accès multiples . . . . .	75

## TABLE DES MATIÈRES

---

4.1.1	Techniques d'accès multiples . . . . .	76
4.1.2	Système multi-tatouage proposé . . . . .	77
4.1.3	Évaluation des performances du système multi-tatouage . . . . .	83
4.2	Système multi-tatouage à base de séparation de sources . . . . .	88
4.2.1	Contrôle d'inaudibilité . . . . .	89
4.2.2	Séparation de sources . . . . .	89
4.2.3	De la détection aveugle vers la non-aveugle . . . . .	92
4.3	Tatouage multiple des signaux audio double canal . . . . .	93
4.3.1	Signaux audio double-canal . . . . .	93
4.3.2	Insertion des tatouages dans les signaux audio à deux canaux . . . . .	94
4.3.3	Extraction des tatouages à partir des signaux audio à deux canaux . . . . .	95
4.4	Discussions . . . . .	96
4.4.1	Performances des différents systèmes de tatouage proposés . . . . .	96
4.4.2	Insertion d'une image numérique . . . . .	99
4.4.3	Comparaison avec les systèmes de télécommunications . . . . .	103
<b>5</b>	<b>Vers un nouveau service radio</b>	<b>105</b>
5.1	La radio . . . . .	105
5.1.1	Importance . . . . .	105
5.1.2	Histoire de la radio . . . . .	106
5.1.3	Évolution . . . . .	107
5.2	Exploitation de l'afficheur radio . . . . .	108
5.2.1	Afficheur radio . . . . .	108
5.2.2	Système RDS . . . . .	109
5.2.3	Système DAB : diffusion audio numérique . . . . .	110
5.3	Transfert d'informations supplémentaires au contenu radio par tatouage numérique	111
5.3.1	Limitations des systèmes existants . . . . .	111
5.3.2	But du procédé proposé . . . . .	111
5.3.3	Principe du procédé . . . . .	112
5.4	Vers la finalisation du service radio . . . . .	113
5.4.1	Intérêt social . . . . .	113
5.4.2	Intérêt économique . . . . .	114
5.4.3	Verrous technologiques . . . . .	115
	<b>Conclusions et perspectives</b>	<b>116</b>
	<b>Annexes</b>	<b>120</b>
<b>A</b>	<b>Audition</b>	<b>120</b>
A.1	Système auditif humain . . . . .	120
A.1.1	Principe . . . . .	120
A.1.2	Mesure physique de l'intensité . . . . .	120
A.2	Psychoacoustique . . . . .	121
A.2.1	Seuil auditif . . . . .	121
A.2.2	Phénomène de masquage . . . . .	122

<b>B Filtrage</b>	<b>125</b>
B.1 Filtre de mise en forme . . . . .	125
B.2 Annuleur d'interférence . . . . .	127
B.3 Filtre de wiener . . . . .	128
<b>Bibliographie</b>	<b>137</b>
<b>Publications</b>	<b>138</b>

---

# Table des figures

1.1	Le nombre d'articles publiés chaque année dans le domaine de tatouage et de cryptographie. . . . .	2
1.2	Identification des informations émises par le pilote d'avion par le tatouage [57]. . . . .	4
1.3	Animation Réaliste par Tatouage audiovisuel à l'Usage des Sourds [10]. . . . .	5
1.4	Comparaison entre tatouage et cryptographie [118]. . . . .	7
1.5	Les différentes modalités de la biométrie [78]. . . . .	9
1.6	Chaîne de communication d'un système de tatouage [9]. . . . .	10
1.7	Comparaison entre le signal audio et le signal image [94]. . . . .	11
1.8	Schéma simple du tatouage QIM [68]. . . . .	12
1.9	Tatouage par échos [12]. . . . .	13
1.10	Système de tatouage comme une chaîne de communication. . . . .	19
1.11	Triangle magique. . . . .	20
2.1	Modèle synthétique de la chaîne de tatouage. . . . .	28
2.2	DSP du signal NRZ ainsi que le signal étalé pour $T_c = T_b/2$ , $T_c = T_b/4$ et $T_c = T_b/10$ . . . . .	30
2.3	Signal émis avant et après étalement de spectre dans les domaines temporel et fréquentiel. . . . .	31
2.4	Illustration du signal de tatouage $w(n)$ , du signal audio $x(n)$ et du signal tatoué $y(n)$ . . . . .	31
2.5	Performances théoriques et pratiques de la fiabilité de transmission du tatouage. . . . .	32
2.6	Fiabilité de transmission du tatouage en fonction du débit de transmission et du RSB. . . . .	33
2.7	Représentation temporelle du signal audio "Violoncello" et du tatouage avec différentes valeurs $\alpha$ . . . . .	34
2.8	Système de tatouage lorsque l'émetteur réalise la mise en forme perceptible du signal modulé en suivant un MPA. . . . .	35
2.9	Allure temporelle du signal audio et seuil de masquage. . . . .	36
2.10	Principe de la détection du tatouage par l'approximation $1/\hat{H}(z)$ de l'annuleur d'interférences et un démodulateur à corrélation. . . . .	37

TABLE DES FIGURES

---

2.11	Représentation de la réponse fréquentielle de l'annuleur d'interférences $G_{AI}(f)$ ainsi que le signal modulé avant et après égalisation. . . . .	37
2.12	Effet d'introduction du filtre de Wiener. . . . .	38
2.13	Système de tatouage en présence d'égalisation mixte par l'annuleur d'IES $G_{AI}(f)$ et par le filtre de Wiener $G_W(f)$ , suivis par le détecteur à corrélation. . . . .	39
2.14	La DSP du signal de tatouage par étalement de spectre à séquence directe. . . . .	40
2.15	La DSP du signal de tatouage par étalement de spectre à saut de fréquence. . . . .	41
2.16	Comparaison entre les deux modulations DSSS et FHSS dans un système de tatouage audio. . . . .	43
3.1	La DSP du signal de tatouage en utilisant la modulation à bande étroite et celle à étalement de spectre ainsi que le seuil de masquage. . . . .	46
3.2	Performances du système de tatouage audio en terme de fiabilité de détection : Comparaisons entre les modulations PSK, DSSS et FHSS. . . . .	51
3.3	Seuil auditif [92]. . . . .	54
3.4	Sélection préliminaire des fréquences porteuses qui assurent un compromis entre l'inaudibilité et la fiabilité de détection. . . . .	55
3.5	Processus d'insertion dans le système de tatouage informé basé sur la modulation PSK. . . . .	57
3.6	Processus d'extraction dans le système de tatouage informé basé sur la modulation PSK. . . . .	59
3.7	Performances du système de tatouage informé en termes d'inaudibilité et de fiabilité de détection en variant le seuil d'inaudibilité. . . . .	61
3.8	Comparaisons entre les performances des modulations numériques en tatouage audio dans le cas d'un canal sans et avec perturbations. . . . .	62
3.9	Système de tatouage multi-porteuses. . . . .	65
3.10	Seuil de masquage et DSP du signal de tatouage à étalement de spectre et multi-porteuses selon le nombre de sous porteuses. . . . .	67
3.11	Variation des performances du système de tatouage multi-porteuses en termes d'inaudibilité et de fiabilité de détection selon le nombre de sous-porteuses. . . . .	68
3.12	Performances du système de tatouage multi-porteuses selon trois types d'intervalles de gardes. . . . .	69
3.13	Robustesse du système de tatouage COFDM contre le BBAG. . . . .	70
3.14	Extrait du signal audio "Violin" avant et après l'opération de compression. . . . .	71
3.15	Robustesse du système de tatouage COFDM contre la compression. . . . .	72
4.1	Le système multi-tatouage audio proposé. . . . .	77
4.2	Comparaison entre la DSP du signal de tatouage en utilisant les techniques CDMA, FHMA et FDMA et le seuil de masquage. . . . .	79
4.3	Performances du système multi-tatouage selon le nombre de tatouages insérés. . . . .	82
4.4	Transparence du tatouage selon la fréquence porteuses $f_c$ . . . . .	84

---

## TABLE DES FIGURES

---

4.5	Comparaisons entres les techniques d'accès multiples dans le système multi-tatouage.	85
4.6	Fiabilité de détection des différentes techniques d'accès multiples selon le débit d'insertion. . . . .	86
4.7	Système multi-tatouage audio basé sur la séparation de sources. . . . .	88
4.8	Évaluation de la transparence des tatouages par le contrôle d'inaudibilité. . . . .	90
4.9	Représentation du signal audio et du signal audio estimé dans le domaine temporel.	92
4.10	Représentation dans le domaine temporel du signal de tatouage estimé par la méthode directe et indirecte. . . . .	92
4.11	Le système multi-tatouage pour les signaux audio à deux canaux. . . . .	93
4.12	Évaluation de qualité qui audio du signal tatoué double canal. . . . .	94
4.13	Fiabilité de détection selon le débit d'insertion. . . . .	96
4.14	Signal Audio : Saxophone. . . . .	99
4.15	Le tatouage sous forme d'image. . . . .	100
4.16	L'extraction de l'image binaire à partir du signal tatoué soumis à plusieurs perturbations. . . . .	101
4.17	L'extraction de l'image binaire à partir du signal tatoué soumis à plusieurs perturbations. . . . .	102
5.1	Stations radio au Maroc. . . . .	106
5.2	Exemples d'utilisation de l'afficheur radio. . . . .	108
5.3	Spectre d'une émission de radiodiffusion FM, avec stéréophonie et RDS [87]. . . . .	110
5.4	Exemple d'utilisation du service radio proposé. . . . .	113
A.1	Schéma de l'appareil auditif humain comprenant l'oreille externe [E], l'oreille moyenne [M] et l'oreille interne [I] [97]. . . . .	121
A.2	Diagramme de Fletcher : Courbes d'isotonie pour l'écoute binaurale [3] . . . . .	123
A.3	Seuil d'inaudibilité du son masqué en fonction de l'ordre et du temps d'apparition du son masqué et du son masquant. . . . .	123

---

# Liste des tableaux

1.1	Comparaisons entre cryptographie et tatouage. . . . .	8
1.2	Signaux audio de test appartenant à la base de données SQAM. . . . .	21
1.3	Échelle de dégradation à cinq notes et valeurs de l'ODG associées. . . . .	22
1.4	Perturbations utilisées pour évaluer les performances en terme de robustesse. . . . .	23
2.1	Le bruit en communication numérique et en tatouage audio. . . . .	33
2.2	Comparaison entre les modulations par étalement de spectre DSSS et FHSS. . . . .	42
3.1	Valeurs de l'ODG du signal tatoué $y(n)$ en utilisant les modulations DSSS, FHSS et PSK. . . . .	53
3.2	Normalisation de l'échelle des Barks et découpage en 25 bandes critiques. . . . .	54
3.3	Coût de complexité des différentes techniques de tatouage. . . . .	62
3.4	Extraction des messages textes par le système COFDM. . . . .	73
4.1	Le temps symbole $N_s$ selon le débit d'insertion $R$ . . . . .	83
4.2	Le nombre de bits par symboles $q$ selon le débit d'insertion $R$ . . . . .	83
4.3	Le nombre de tatouages $K$ selon le débit d'insertion $R$ . . . . .	83
4.4	Evaluation de la robustesse des techniques d'accès multiples vis-à-vis des perturbations externes à $R = 4\text{kb/s}$ . . . . .	86
4.5	Bande de fréquences utilisées pour les techniques FDMA et FHMA. . . . .	87
4.6	Comparaisons entre les système de tatouage audio en termes de débit, inaudibilité, fiabilité de détection et coût de calcul. . . . .	98
4.7	Comparaisons entre les système de tatouage audio en termes de robustesse. . . . .	98
4.8	Les valeurs PSNR t TEB des logos insérés. . . . .	100
5.1	Bref résumé de l'histoire de la radio . . . . .	106
A.1	Correspondance entre niveau de pression acoustique et pression. . . . .	122

---

# Glossaire

<b>ASK</b>	: Amplitude Shift Keying
<b>BBAG</b>	: Bruit Blanc Additif Gaussien
<b>CDMA</b>	: Code Division Multiple Access
<b>COFDM</b>	: Coded Orthogonal frequency-division multiplexing
<b>CP</b>	: Cyclic Prefix
<b>CS</b>	: Cyclic Suffi
<b>DCT</b>	: Discrete Cosine Transform.
<b>DFT</b>	: Discrete Fourier Transform.
<b>DSP</b>	: Densité Spectrale de Puissance
<b>DWT</b>	: Discrete Wavelet Transform.
<b>DSSS</b>	: Direct Sequence Spread Spectrum
<b>EMD</b>	: Empirical Mode Decomposition
<b>IES</b>	: Interférence Entre Symboles
<b>i.i.d</b>	: indépendants et identiquement distribués
<b>IMF</b>	: Intrinsic Mode Functions
<b>ISH</b>	: Interférence avec le Signal Hôte
<b>FDMA</b>	: Frequency Division Multiple Access
<b>FFT</b>	: Fast Fourier Transform.
<b>FHSS</b>	: Frequency Hopped Spread Spectrum
<b>FSK</b>	: Frequency Shift Keying
<b>LPC</b>	: Langage Parlé Complété
<b>LSB</b>	: Least Significant Bit
<b>MP</b>	: Multi-porteuse
<b>MPA</b>	: Modèle Psychoacoustique
<b>NRZ</b>	: Non Return to Zero
<b>RZ</b>	: Return to Zero
<b>ODG</b>	: Objective Difference Grade
<b>PEAQ</b>	: Perceptual Evaluation of Audio Quality
<b>PSK</b>	: Phase Shift Keying
<b>QIM</b>	: Quantization Index Modulation
<b>RSB</b>	: Rapport Signal à Bruit
<b>SDMI</b>	: Secure Digital Music Initiative
<b>SQAM</b>	: Sound Quality Assessment Material
<b>TEB</b>	: Taux d'Erreur Binaire
<b>UISA</b>	: Under-determined Independent Subspace Analysis
<b>ZP</b>	: Zero Padding

---

# Notations

$x(n)$	Signal audio original
$v(n)$	Signal modulé
$s(n)$	Signal en bande de base
$w(n)$	Signal de tatouage
$y(n)$	Signal audio tatoué
$c_m(n)$	Formes d'ondes constituant le dictionnaire de modulation
$c(n)$	Séquence d'étalement

---

$R$	Débit d'insertion ou débit de transmission
$N$	Durée totale du signal de tatouage
$N_s$	Temps symbole discret
$N_b$	Temps bit discret
$N_c$	Nombre de sous-porteuses
$N_{sym}$	Temps symbole MP
$N_G$	Longueur de l'intervalle de garde
$N_{symG}$	Temps symbole étendu
$T_s$	Temps symbole
$F_e$	Fréquence d'échantillonnage
$L_c$	Facteur d'étalement

---

$b_k$	Information binaire à tatouer
$a_k$	séquences de symboles
$s(n)$	Signal en bande base
$S_s(f)$	DSP du signal $s(n)$
$S_v(f)$	DSP du signal $v(n)$
$S_w(f)$	DSP du signal $w(n)$
$M_x(f)$	Seuil de masquage
$H(f)$	Filtre de mise en forme
$\hat{H}(f)$	Filtre de mise en forme estimé
$G_{AI}(f)$	Annuleur d'interférences
$G_W(f)$	Filtre de Wiener

---

$q$	Nombre de bits par symbole
$M$	Nombre de formes d'ondes
$A$	Amplitude du signal $v(n)$
$\varphi$	Phase du signal $v(n)$
$f_c$	Fréquence porteuse du signal $v(n)$

---

# Chapitre 1

## Contexte général

### 1.1 Enjeux du tatouage

#### 1.1.1 Un peu d'histoire

Bien que l'art de la fabrication du papier a été inventé en Chine il y a plus de mille ans, le tatouage papier n'est apparu en Italie qu'en 1282 [27]. Le but derrière l'utilisation du tatouage papier reste toujours inconnu et incertain, mais quelques sources historiques affirment qu'il a été utilisé probablement pour identifier le fabricant du papier. Au cours du 18<sup>ème</sup> siècle, le tatouage papier a été utilisé en Europe et en Amérique en tant que marque, pour enregistrer la date dans laquelle le document a été fabriqué et indiquer la taille des feuilles originales.

Deux cent ans plus tard, le tatouage électronique a vu le jour. En 1954, Emil Hembrooke de la société Muzak, a déposé un brevet d'invention sur le tatouage des oeuvres musicales [29]. Cette invention a permis d'identifier les oeuvres musicales et elle a constitué alors un moyen efficace pour empêcher le piratage des documents. Un certain nombre de travaux est apparu après, mais tous ces travaux restaient essentiellement heuristiques tant que le tatouage n'était pas reconnu autant que technologie distincte.

Cela a commencé à changer à la fin des années 1980, quand le terme "tatouage numérique" a été utilisé pour la première fois dans [29]. Dans les débuts des années 1990, le tatouage a commencé à intéresser les chercheurs et un certain nombre de documents publiés contenaient des descriptions de plusieurs algorithmes de tatouage numérique. La première conférence internationale qui comprenait le tatouage numérique dans ses principaux thèmes a eu lieu en 1996 [29]. Depuis lors, le tatouage numérique a connu un intérêt substantiel en tant que sujet de recherche. A cette époque, plusieurs organisations ont inclus le tatouage dans diverses normes internationales. Le CPTWG<sup>1</sup> a testé les systèmes de tatouage pour la protection des vidéos sur quelques disques DVD [27]. La SDMI<sup>2</sup> à son tour, a utilisé le tatouage comme un composant principal dans ses systèmes pour protéger la musique [27].

Au cours des 20 dernières années, le tatouage numérique a connu un intérêt considérable, et il a commencé à faire sa place dans quelques domaines d'applications. Des efforts énormes ont été fournis en vue de permettre à ces applications de perfectionner leurs besoins. Malheureusement, le nombre d'applications impliquant le tatouage numérique reste toujours restreint. En effet, les

---

1. Copy Protection Technical Working Group.  
2. Secure Digital Music Initiative.

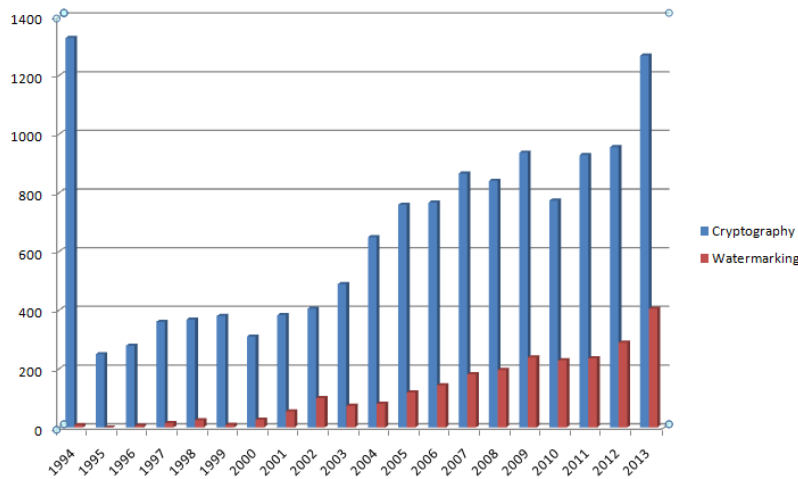


FIGURE 1.1 – Le nombre d’articles publiés chaque année dans le domaine de tatouage et de cryptographie.

chercheurs dans ce domaine prétendent que le tatouage vit son âge d’enfance [1]. De même, nous estimons qu’il passe maintenant à son âge d’adolescence où plusieurs compagnies investissent dans des projets de recherche mais n’arrivent pas encore à déployer de très bonnes solutions. Ceci est dû au manque d’algorithmes qui peuvent satisfaire les contraintes exigées puisque le tatouage numérique existe depuis moins de vingt ans. Si on le compare avec un domaine aussi connu comme la cryptographie dont les algorithmes sont déployés dans le monde entier pour la sécurité, on remarque que la cryptographie dépasse largement le tatouage en termes de nombre d’articles<sup>3</sup> comme c’est illustré dans la FIGURE 1.1. Néanmoins, il ne faut pas oublier que le tatouage est encore nouveau et le premier article est apparu en 1994 alors qu’il y avait 1328 articles en cryptographie. Aujourd’hui, les contributions dans le domaine du tatouage, que ce soit pour des images, des vidéo ou des signaux audio, sont de plus en plus nombreuses. En 2013 par exemple, 400 articles ont été publiés en tatouage numérique, et 1200 articles ont été publiés en cryptographie. Certes, le nombre de travaux en cryptographie est encore plus qu’en tatouage numérique mais on remarque qu’il y a un intérêt remarquable de la part des chercheurs en tatouage numérique.

### 1.1.2 Classes de tatouages

Le tatouage est souvent confondu avec deux termes techniques : l’enfouissement de données et la stéganographie [27]. Ces trois termes ont beaucoup de chevauchements et partagent de nombreuses notions. Néanmoins, ces termes présentent quelques points de différences, tant sur la conception que sur les contraintes d’applications.

- **L’enfouissement de données**<sup>4</sup> : est un terme général qui englobe un large éventail de problèmes au-delà de celui d’enfouir des messages dans le contenu. Il peut se référer soit

---

3. On a effectué la recherche des deux mots **watermarking** et **cryptography** dans la base de données Sciencedirect.

4. Datahiding.

à rendre l'information imperceptible (comme en tatouage) ou de maintenir le secret de l'existence de l'information (comme en stéganographie).

- **La stéganographie** : est l'art de cacher l'existence d'un message transmis, via un support (texte, image ou **son**) entre un émetteur et un récepteur. Elle étudie les mécanismes permettant de rendre la présence du message secrète et indétectable.
- **Le tatouage**<sup>5</sup> : est une technique qui permet d'incorporer une signature voire une marque portant une information particulière, généralement liée au signal de support.

### 1.1.3 Domaines d'applications

Le tatouage peut être utilisé dans plusieurs types d'applications qui visent deux contextes différents [27, 33, 46, 54] ; le premier visant la sécurité des documents numériques et le second visant la transmission de données.

#### 1.1.3.1 Contexte sécuritaire

Dans le contexte sécuritaire, le tatouage vise à adapter la marque insérée en fonction de l'action déroulée sur le document par le pirate. Dans ce cas, l'information intégrée à l'intérieur du document original doit être robuste vis-à-vis des différentes attaques intentionnelles du pirate [27, 33]. Parmi les applications visant la sécurité, on trouve :

- **L'identification du propriétaire** : un tatouage contenant des informations sur les droits d'auteur est intégré dans le signal multimédia original. Le tatouage, connu seulement de la part du titulaire des droits d'auteur, doit être très robuste et sécurisé, permettant au propriétaire de justifier la présence de ce tatouage dans le cas de litige de propriété.
- **Preuve d'appartenance** : il est encore possible d'utiliser le tatouage non seulement pour identifier la propriété du droit d'auteur, mais comme une preuve réelle de la propriété. Le problème se pose quand un adversaire essaie de remplacer la mention du droit d'auteur originale par une autre, puis prétend de posséder le droit d'auteur. Dans ce cas, au lieu d'avoir une preuve directe de la propriété en insérant, par exemple, une autre signature dans le document original, l'algorithme va plutôt essayer de prouver que le document de l'adversaire est dérivé de celui tatoué. Un tel algorithme fournit une preuve indirecte puisqu'il est plus probable que le véritable propriétaire possède le document litigieux, parce que le véritable propriétaire est celui qui dispose de la version originale à partir de laquelle les deux autres ont été créés.
- **L'authentification** : la signature est intégrée dans le contenu original, et elle est utilisée plus-tard pour vérifier si le contenu a été altéré ou non. Si on prend comme exemple l'enregistrement de la parole, il serait facile d'éliminer une partie d'une phrase enregistrée. Ceci peut changer complètement la compréhension de tout un discours. Ainsi, pour éviter ce traitement illégale, on insère une marque dans le document de telle sorte que si on enlève une partie de la phrase, une partie de la marque sera également enlevée et cela va empêcher sa détection correcte. Si la marque n'est pas détectée, on pourra conclure que le document a été modifié.

---

5. Watermarking.

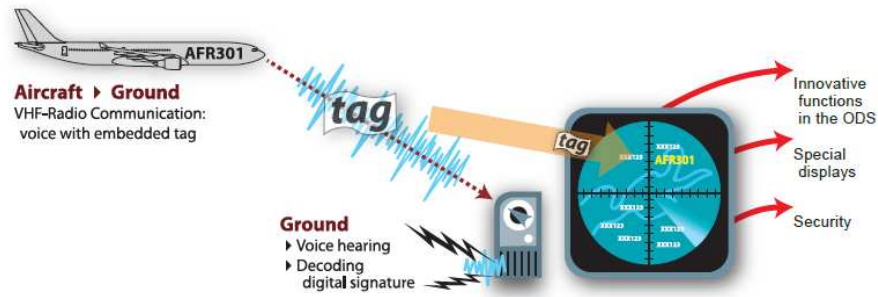


FIGURE 1.2 – Identification des informations émises par le pilote d’avion par le tatouage [57].

- **Le suivi des transactions** : connu aussi sous le nom d’empreintes digitales<sup>6</sup>. Le tatouage est utilisé ici pour tracer l’expéditeur d’un exemplaire de données multimédia. L’idée est d’utiliser une marque différente pour chaque copie du document. Dans ce cas, si on trouve une copie illégale dans le marché, on pourra facilement identifier la personne à qui on a vendu le contenu original.
- **Le contrôle de reproduction** : on cherche à éviter la génération et l’utilisation des produits illégalement copiés. Pour ce faire, l’information insérée serait utilisée pour bloquer les appareils électroniques qui servent à reproduire les copies illégales.

### 1.1.3.2 Contexte non sécuritaire

Le tatouage intégré dans ce type d’applications ne vise pas cette fois-ci la sécurité du tatouage mais s’intéresse plutôt à transférer des informations supplémentaires dans le contenu numérique [46, 54]. Bien que la robustesse contre les attaques intentionnelles n’est pas nécessaire, un certain degré de robustesse contre des perturbations licites comme la compression peut être souhaitable. Dans ce type d’applications, le tatouage doit contenir généralement une capacité d’informations élevée et doit être extrait en utilisant un algorithme de détection aveugle. Parmi les applications visant la transmission de données, on trouve :

- **Le contrôle de diffusion** : cette application, souvent associée au tatouage sécuritaire, peut être envisagée dans un contexte non sécuritaire dès lors que l’identifiant ne sert pas de preuve de propriété mais permet d’établir des statistiques sur l’utilisation du document. Dans les émissions de radio, les annonceurs veulent généralement s’assurer si leurs annonces et leurs publicités ont été correctement diffusées conformément au nombre de fois spécifié dans le contrat. Pour cela, on insère une marque dans chaque publicité. L’information identifie l’enregistrement publicitaire lors de son transfert dans un réseau de diffusion (radio, télévision ou Internet). Elle permet, par exemple, de savoir sur quelle radio le signal audio a été diffusé, combien de fois, et à quelle heure.
- **Contrôle du trafic aérien** : le tatouage peut être utilisé pour remplacer les identifications vocales des pilotes lors de leurs échanges avec les bases de contrôle aérien [47]. En effet, le contrôle aérien réglant le trafic s’effectue par l’intermédiaire d’une fréquence radio, partagée par tous les pilotes des avions en vol : plusieurs échanges entre les différents

---

6. Fingerprinting.

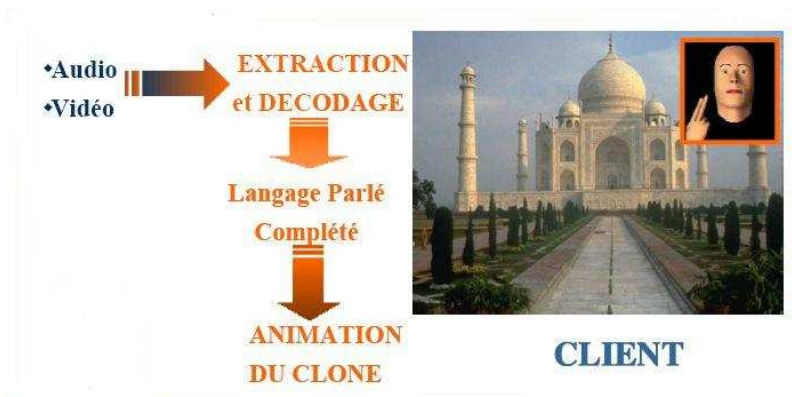


FIGURE 1.3 – Animation Réaliste par Tatouage audiovisuel à l’Usage des Sourds [10].

pilotes et la station de contrôle peuvent avoir lieu en même temps. Il est alors primordial que chaque pilote communique, à chaque échange, son identité. Au lieu d’effectuer cette identification vocalement, ceci pourrait être réalisée en tatouant la communication radio comme c’est illustré à la FIGURE 1.2.

- **Annotation de documents sonores** : cette application peut servir pour transférer une étiquette afin d’aider l’indexation des signaux dans les bases de données [110] : l’information cachée à destination de l’auditeur ou de l’administrateur de la base de données peut transférer des meta-données décrivant le contenu du signal ou des informations à destination d’une application cible. Par exemple, le message caché peut indiquer le nom de l’artiste, le lieu de l’enregistrement ou toute autre donnée relative au signal. Cette annotation peut être faite en studio ou en temps-réel lors d’un concert.
- **L’ajout d’informations publicitaires** : le signal audio peut cacher des informations annexes à diffuser vers un auditeur. Par exemple, l’information cachée peut contenir la référence d’un produit mis en vente par le télé-achat [108].
- **Augmentation de l’intelligibilité des émissions télévisées** : c’est une application qui fonctionne en temps réel dont l’objectif est de remplacer l’affichage du télétexte par l’insertion d’un clone dans les émissions télévisées (c.f. FIGURE 1.3). Ce dernier permettra aux sourds et malentendants d’améliorer leur compréhension envers une telle télédiffusion grâce au mouvement du visage et des mains qui reproduiront le Langage Parlé Complété<sup>7</sup> [10].
- **Séparation des extraits musicaux** : ayant à notre disposition des signaux sources musicaux, un ensemble d’informations portant certaines caractéristiques, peut être extrait à partir des signaux sources. Ces informations, représentant le tatouage, sont insérées de manière inaudible, dans le mélange des signaux audio sources. Grâce au tatouage, on peut extraire toutes les informations utiles à la séparation des signaux musicaux originaux [6].

---

7. Le Langage Parlé Complété (LPC) est un langage qui complète la lecture labiale à l’aide des gestes de la main.

#### 1.1.4 Contraintes générales

Les systèmes de tatouage peuvent être caractérisés par un certain nombre de contraintes [27, 33, 88]. L'importance de chaque contrainte dépend de l'application envisagée. La pertinence d'un système de tatouage audio visant une application particulière, peut être évaluée en fonction des propriétés suivantes :

- **L'inaudibilité** : l'introduction de la marque ne doit pas modifier la qualité auditive du fichier audio original et ne doit pas générer de distorsions perceptibles. L'idée consiste à rendre le contenu original et le contenu tatoué les plus identiques possible sur le plan perceptif. Il est courant d'utiliser la notion d'imperceptibilité ou de transparence pour faire référence à ce concept.
- **La fiabilité de détection** : l'information cachée doit être extraite, à partir du signal audio tatoué, avec le minimum d'erreurs possibles.
- **La robustesse** : elle caractérise la résistance de la marque insérée contre les opérations licites effectuées sur les signaux numériques (compression/décompression, conversion analogique-numérique, filtrage, changement de support, etc) et celles illicites (attaques malveillantes des pirates).
- **Le débit d'insertion** : il s'agit d'insérer une quantité d'information élevée dans un support tout en assurant les fonctionnalités envisagées.
- **La détection aveugle** : il serait relativement facile de retrouver une marque cachée dans un document si l'on dispose d'une copie de l'original. Une simple différence mathématique pourrait convenir. Une détection qui utilise le document original est connue sous le nom de détection non aveugle. Cependant, en pratique l'utilisation du document original est rarement possible. L'information cachée doit donc être détectée directement à partir du signal audio tatoué sans faire recours au signal audio original.
- **La sécurité** : le système de tatouage doit être sécurisé dans le sens que l'adversaire ne doit pas être en mesure de détecter la présence des données insérées, ou encore de les supprimer. Un utilisateur non autorisé ne devrait pas être capable d'extraire les données dans un laps de temps raisonnable, même s'il sait que le signal hôte contient un tatouage. Les exigences de sécurité varient en fonction de l'application envisagée et dans certains cas, les données sont cryptées avant le processus d'insertion dans le document original.
- **La synchronisation** : il s'agit de situer le début de la marque en concordance avec le début du signal à tatouer, ainsi même si un pirate effectue une attaque par décalage temporel, il ne pourra pas tromper le détecteur.
- **Le coût** : c'est une propriété particulière pour les applications qui fonctionnent en temps réel. Par exemple, dans le contrôle de diffusion, les processus d'insertion et de détection doivent être effectués en temps réel, tandis que dans les applications de protection du droit d'auteur, le temps n'est pas un facteur déterminant pour une mise en oeuvre pratique.



(a) Tatouage numérique.



(b) Cryptographie.

FIGURE 1.4 – Comparaison entre tatouage et cryptographie [118].

## 1.2 Le tatouage aux frontières de plusieurs disciplines

Entre cryptographie, biométrie, télécommunications et compression, le tatouage numérique présage la nécessaire pluridisciplinarité des savoirs. Le tatouage se distingue des autres techniques par trois points essentiels. Tout d'abord, il est imperceptible, de plus il reste toujours présent lorsqu'un document se convertit en un autre format de fichiers et enfin il subit les mêmes transformations que le document dans lequel il est inséré.

### 1.2.1 Tatouage et cryptographie

Récemment, la sécurité est devenue l'un des problèmes les plus difficiles pour la diffusion des nouvelles technologies de l'information. En effet, la distribution des copies illégales, au coût très bas et sans aucune dégradation de la qualité exige des solutions de sécurité de plus en plus importantes. La question qu'on peut se poser alors est : pourquoi ne pas se servir de la cryptographie à la place du tatouage numérique pour améliorer la sécurité des documents partagés sur internet. Pour répondre à cette question, il serait important de souligner la différence qui existe entre cryptographie et tatouage [37, 116].

La cryptographie est probablement la méthode la plus commune de la protection du contenu numérique. Le contenu est crypté avant la livraison, et une clef de déchiffrement est fournie uniquement à ceux qui ont acheté des exemplaires originaux. Le fichier crypté peut alors être mis à disposition via Internet, mais serait inutile à un pirate sans une clef appropriée. Malheureusement, un pirate peut effectivement acheter le produit, utiliser la clef de décryptage pour obtenir une copie non protégée du contenu, puis procéder à la distribution des copies illégales.

En d'autres termes, un fichier numérique (audio ou image) crypté est inutilisable sans le code de déchiffrement, alors qu'un fichier tatoué peut être utilisé sans fournir aucune information supplémentaire (c.f. FIGURE 1.4). Par ailleurs, le tatouage procure une protection permanente à son contenant ; la marque est toujours présente dans le fichier original, même dans les copies illégales. Ce n'est pas le cas d'un document crypté ; une fois déchiffré, le fichier n'est plus protégé [116].

De plus, la cryptographie ne permet pas de vérifier si le vrai message a été bien transmis ou non. En effet, si un pirate supprime une partie du message crypté, on ne pourra jamais récupérer

le message transmis même si on dispose de la clef de déchiffrement.

Enfin, les documents multimédias nécessitent des performances élevées, et leurs données peuvent être modifiées en raison d'erreurs de transmission, des taux de compression élevés, ou des opérations de mise à l'échelle. Par conséquent, il est difficile de définir une entrée appropriée des mécanismes cryptographiques. Si l'on applique des mécanismes cryptographiques directement aux documents multimédias, certains problèmes peuvent survenir ; le déchiffrement du bloc de texte chiffré peut échouer. Le tatouage numérique permet alors d'incruster directement dans le contenu numérique une information qui peut survivre vis-à-vis des différentes perturbations externes.

En résumé, on peut dire que le principe de la cryptographie est de rendre incompréhensible une communication dont on connaît l'existence, en employant une technique de chiffrement des informations échangées [116]. En contrepartie, le tatouage a pour objectif de cacher un message de manière imperceptible dans un support numérique [37]. Le tableau 1.1 illustre quelques exemples de comparaisons entre la cryptographie et le tatouage numérique.

<b>Cryptographie</b>	<b>Tatouage</b>
Écriture secrète : visible et incompréhensible.	Écriture discrète : imperceptible et compréhensible.
Le document crypté peut être vu par n'importe qui mais la cryptographie rend le message incompréhensible.	Le tatouage cache le message dans un autre média de telle sorte que personne ne peut le lire.
Le résultat final de la cryptographie est le document chiffré.	Le résultat final du tatouage est le support média tatoué.
Toute personne a le droit de détecter et de modifier le document chiffré.	Le message caché est imperceptible à tout le monde.
Le tatouage ne peut être utilisé pour augmenter la robustesse du système cryptographique.	Le tatouage peut être utilisé avec la cryptographie pour cacher un message crypté.

Tableau 1.1 – Comparaisons entre cryptographie et tatouage.

### 1.2.2 Tatouage et biométrie

Un autre terme très utilisé en sécurité et qui est souvent confondu avec le tatouage est celui de la biométrie [78]. Contrairement au tatouage qui est utilisé pour la protection des documents numériques, la biométrie est utilisée pour la sécurité des personnes, des biens ou des informations (c.f. FIGURE 1.5). En effet, la biométrie vise à vérifier l'identité d'une personne à l'aide d'une ou plusieurs modalités qui lui sont propres (voix, iris, empreintes digitales, visage, etc). Il existe deux classes de méthodes en biométrie ; à savoir : l'indentification et la vérification.

- **L'identification** : elle nécessite le stockage centralisé ou local des gabarits dans des bases de données et des algorithmes permettant de chercher un élément dans ces grandes bases de données. Elle permet la reconnaissance d'un individu produisant une annonce (ex. une signature, une phrase parlée...), parmi plusieurs individus.
- **La vérification** : elle peut se faire sans accès à de grandes bases de données, en préservant

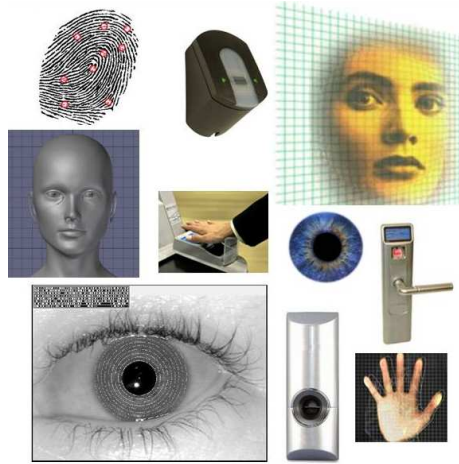


FIGURE 1.5 – Les différentes modalités de la biométrie [78].

l'anonymat (par exemple si on stocke les gabarits cryptés sur une carte à puce). Elle vise à vérifier si l'identité qu'une personne a annoncée est vraie ou fausse.

Le tatouage peut être utilisé comme un moyen fiable pour améliorer les systèmes biométriques. En effet, il peut servir à augmenter la performance de reconnaissance comme il peut servir à améliorer la sécurité :

- **Performance de reconnaissance** : le tatouage est inséré dans des échantillons de données pour être utilisé ensuite pour la reconnaissance.
- **Sécurité** : le but du tatouage ici n'est pas d'améliorer la reconnaissance des systèmes biométriques, mais d'employer un message particulier pour la protection du droit d'auteur afin de permettre la reconnaissance biométrique après l'extraction du tatouage.

### 1.2.3 Tatouage et télécommunications

L'objectif principal d'un système de télécommunications est de transmettre efficacement une information (un message binaire) d'un émetteur vers un récepteur [53]. Par exemple deux ordinateurs reliés par Internet : le message binaire, une suite de 0 et de 1 (représentant un texte ou une image), est converti par l'émetteur en un signal physique (analogique), le signal modulé, est transmis au récepteur, qui analyse le signal reçu (après l'avoir démodulé) pour retrouver le message binaire (numérique) initial. Or, au cours de la transmission du message, celui-ci peut être bruité; un bruit de canal peut s'ajouter au signal reçu (des bits seront erronés) et venir perturber la bonne récupération du message. Le tatouage peut être utilisé donc pour améliorer les systèmes de transmissions existants. En effet, il peut être utilisé pour [54] :

- Aider à récupérer le signal hôte transmis sur le canal de communication par la correction des erreurs [22]. L'information cachée est utilisée pour réduire le bruit lié à la transmission [79].
- Tester la qualité de services des transmissions multimédia [13, 17].

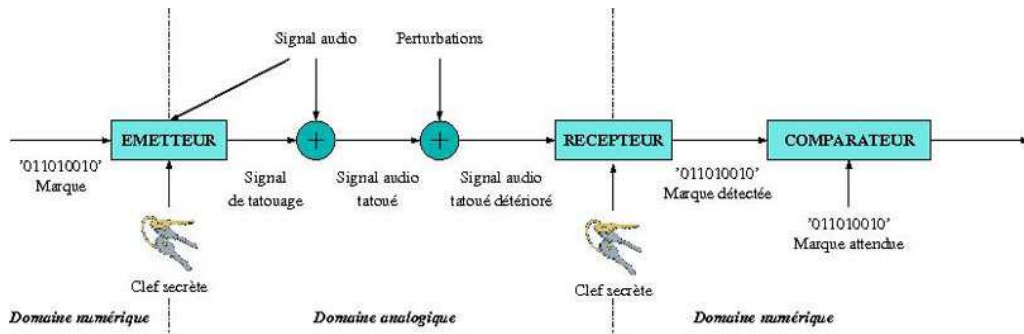


FIGURE 1.6 – Chaîne de communication d'un système de tatouage [9].

- Évaluer, de manière subjective, la qualité du signal de parole transmis sur la bande téléphonique [16].
- Élargir la bande passante du signal transmis dans un système de communication [45, 98].

D'un autre côté, les systèmes de tatouage additifs reposent eux aussi sur le même principe qu'un système de télécommunications numérique. En effet, au lieu d'exploiter le tatouage pour améliorer les performances en télécommunications, le système de tatouage peut être vu lui même comme une chaîne de communication particulière. Ce système permet la transmission d'une information supplémentaire dans un canal adjacent au signal audio destiné soit à l'auditeur soit à une application cible (c.f. FIGURE 1.6). Dans ce cas, le tatouage est le signal utile porteur d'informations tandis que le signal original représente le bruit perturbateur du canal.

### 1.2.4 Tatouage et compression

Plusieurs travaux de recherche en tatouage audio ont été proposés ces dernières années grâce aux algorithmes et méthodes déjà existants en compression audio. Le principe de la compression est basé sur l'exploitation des défauts du système auditif humain pour supprimer l'information redondante et non perceptible d'un signal audio [36]. En effet, notre système auditif néglige certaines parties de l'information contenue dans un signal audio. Le rôle de la compression audio est de réduire le débit du son numérique tout en préservant la qualité. Ces mêmes défauts sont exploités en tatouage audio pour ajouter une information inaudible et cacher un message, une marque ou une signature. Ceci est basé sur la théorie de la psychoacoustique qui consiste à modéliser le système auditif humain et la perception du son afin de déterminer les conditions pour lesquelles le tatouage inséré reste transparent pour l'auditeur.

## 1.3 État de l'art

Le tatouage numérique a été initialement appliqué aux images numériques. Depuis lors, plusieurs techniques en tatouage image ont été proposées [12, 23]. Toutefois, ces techniques ne peuvent pas être directement transposées aux signaux audio. En effet, il faut préciser que contrairement aux images qui sont essentiellement basses fréquences, les signaux audio possèdent une dynamique beaucoup plus forte et une occupation de la bande spectrale perceptible plus importante (c.f. FIGURE 1.7). Il est à noter que le tatouage des signaux audio est plus difficile

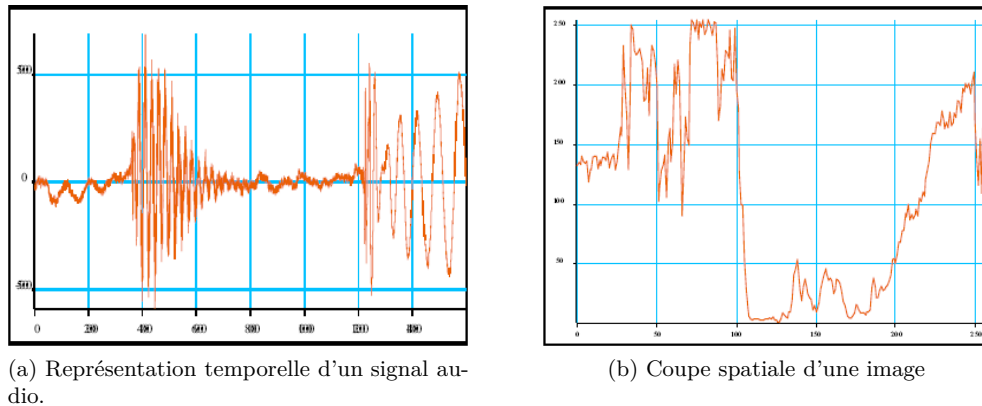


FIGURE 1.7 – Comparaison entre le signal audio et le signal image [94].

que le tatouage d'images ou des séquences vidéo en raison de la gamme dynamique plus large du système auditif humain par rapport au système visuel humain [12]. Par conséquent, la quantité de données qu'on peut intégrer de manière transparente dans un signal audio est considérablement inférieure à la quantité de données qui peut être incrustée dans des images numériques et des séquences vidéos.

### 1.3.1 Classification des techniques de tatouage audio

Les méthodes de tatouage sont nombreuses tant aux images numériques qu'aux signaux audio. On va se limiter ici seulement aux méthodes de tatouage audio. On va présenter les méthodes les plus importantes qui constituent la base de la plupart des algorithmes de tatouage existants.

#### 1.3.1.1 Codage par LSB

L'une des premières méthodes de tatouage des signaux audio est le codage du bit le moins significatif LSB<sup>8</sup> [115]. Le tatouage LSB consiste à remplacer les bits les plus faibles du signal hôte quantifié par l'information du tatouage. Le codeur LSB utilise une clef secrète pour choisir un sous-ensemble pseudo-aléatoire de tous les échantillons du signal audio. L'opération de substitution est effectuée sur des échantillons sélectionnés. Dans la phase de décodage, la même clef secrète est utilisée pour extraire les bits de tatouage à partir du signal reçu. Les principaux avantages de cette méthode de tatouage sont :

- Sa capacité d'insertion élevée. Par exemple, si on choisit seulement le bit le moins significatif du signal audio de qualité CD (de fréquence d'échantillonnage égale à 44.1 kHz et encodé sur 16 bits), on peut atteindre un débit d'insertion égale à 44 100 bits par seconde.
- Sa simplicité de mise en oeuvre, qui nécessite très peu de coût de calcul tant au codage qu'au décodage permettant ainsi l'insertion et l'extraction du tatouage en temps réel, même pour les dispositifs limités en puissance de calcul.

---

8. Least Significant Bit.

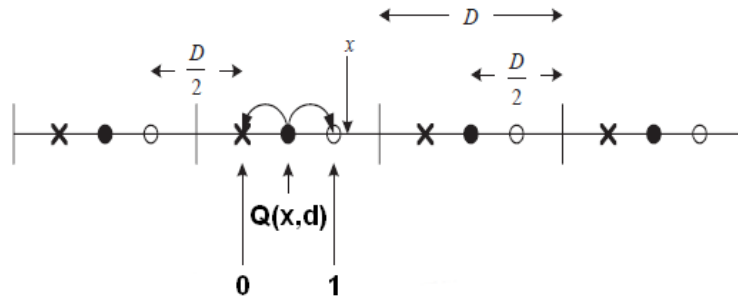


FIGURE 1.8 – Schéma simple du tatouage QIM [68].

Ces avantages sont contrebalancés par des inconvénients :

- Le tatouage LSB est très limité par la contrainte d'inaudibilité. La sélection aléatoire des échantillons, utilisés dans l'insertion du tatouage, introduit un bruit blanc gaussien de puissance faible. Ceci limite le nombre de bits de poids faible qui peuvent être imperceptiblement modifiés à cause de la sensibilité du système auditif humain au bruit blanc additif gaussien. Afin d'assurer la transparence du tatouage, seuls les 4 bits les moins significatifs des 16 bits par échantillons du signal original peuvent être utilisés dans le codage LSB.
- Le codage LSB est très fragile contre les perturbations. Une simple conversion analogique-numérique détruira complètement le tatouage inséré.

### 1.3.1.2 Tatouage par QIM

La Modulation de l'Indice de Quantification (QIM<sup>9</sup>) est une autre technique de tatouage qui cache l'information par la quantification des échantillons du signal audio  $x$  [68]. L'amplitude maximale en valeur absolue des échantillons du signal  $x$  ainsi que la résolution de quantification  $J$  détermine le pas de quantification  $D = 2 * \max(|x|)/(2^J)$ . Par la suite, chaque échantillon est quantifié avec ce pas de quantification. Une légère modification est alors faite à chaque échantillon quantifié selon le tatouage. Une simple implémentation du QIM peut se résumer comme suit [68].

Le tatouage par quantification consiste à modifier les niveaux d'amplitude du signal hôte d'entrée  $x$  par un quantificateur qui dépend du tatouage  $w=(0$  ou  $1)$  à cacher. L'échantillon tatoué  $y$  est représenté comme suit :

$$y = Q(x, d) + \frac{D}{4} \times (2 \times w - 1) \quad (1.1)$$

où  $Q(x, D)$  est la fonction de quantification et  $D$  est le pas de quantification. La fonction de quantification est définie comme suit :

$$Q(x, D) = \left[ \frac{x}{D} \right] \times D \quad (1.2)$$

où  $[x]$  est la fonction qui arrondit à l'entier le plus proche de  $x$ . Le concept du système de quantification dans l'équation (1.1) est illustré sur la FIGURE 1.8. L'échantillon  $x$  est d'abord

9. Quantization Index Modulation.

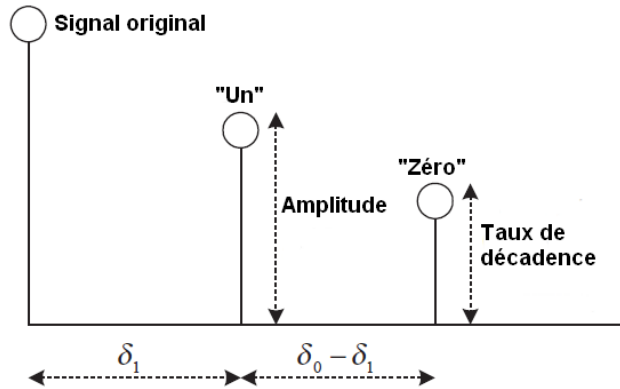


FIGURE 1.9 – Tatouage par échos [12].

quantifié par la fonction  $Q(x, D)$  noté par le cercle noir. Si le bit de tatouage est 1, alors la valeur  $D/4$  est ajoutée à la valeur de l'échantillon quantifié qui déplace l'échantillon jusqu'au rond blanc. Dans le cas contraire,  $D/4$  est soustraite de la valeur de l'échantillon quantifié, qui déplace l'échantillon vers la croix (x). Du côté du décodeur, la différence entre l'échantillon reçu et sa valeur quantifiée est calculée. Si elle est comprise entre  $[0, D/4]$ , alors le bit du tatouage extrait est "1". Si la différence se situe entre  $[-D/4, 0]$ , le bit de tatouage extrait est "0". Dans le cas contraire, le signal reçu n'est pas tatoué. Ceci peut être illustré par le système d'équations suivant :

$$\begin{cases} w = 1 & \text{si } 0 < y - Q(x, D) \leq \frac{D}{4} \\ w = 0 & \text{si } -\frac{D}{4} \leq y - Q(x, D) < 0 \end{cases} \quad (1.3)$$

Le tatouage QIM est très simple à mettre en oeuvre et il est robuste contre les attaques de type bruit additif à faible puissance. Tant que le bruit introduit au canal de transmission est inférieur à  $D/4$ , le détecteur peut toujours extraire correctement le tatouage. Cependant, si le bruit dépasse  $D/4$ , le tatouage devient difficile à être détecté avec précision. Il s'agit donc d'assurer un compromis entre la robustesse et la transparence du tatouage. Ainsi, un grand pas de quantification  $D$  augmentera la robustesse du tatouage avec un risque de créer une distorsion audible au signal audio.

### 1.3.1.3 Tatouage par échos

Les systèmes de tatouage par échos introduisent l'information cachée dans le signal hôte  $x(n)$  en ajoutant des échos pour construire le signal tatoué  $y(n)$  [12]. Les méthodes de tatouage par échos sont aveugles et ont été conçues spécialement pour les signaux audio<sup>10</sup>. Ces méthodes sont très robustes contre les attaques standards de tatouage et permettent un débit d'insertion de dizaines de bits par seconde.

Le tatouage par échos consiste à ajouter une résonance acoustique au signal hôte. Après avoir ajouté l'écho, le signal tatoué conserve les mêmes caractéristiques statistiques et perceptives. Le décalage (ou retard)  $\tau$  entre le signal original et le tatouage est suffisamment petit pour que

10. Ils ne sont pas conçus pour le tatouage image et vidéo.

l'écho soit perçue par le système auditif humain comme une résonance ajoutée.

$$y(n) = x(n) + \alpha x(n - \tau) \quad (1.4)$$

La FIGURE 1.9 présente les quatre principaux paramètres qui sont l'amplitude initiale, le taux de décadence, l'offset "un", et l'offset "zéro". Traiter un signal hôte par n'importe quel noyau se traduit par un signal codé. Le retard résultant entre le signal original et l'écho dépend du noyau utilisé ; si le noyau "un" est utilisé le retard  $\tau$  sera égale à  $\delta_1$  et si le noyau "zéro" est utilisé le retard  $\tau$  sera égale à  $\delta_0$ .

L'extraction du message inséré implique la détection du retard  $\tau$ . L'analyse cepstrale duplique l'impulsion cepstrale sur chaque  $\tau$  échantillons. L'amplitude des impulsions représentant les échos sont faibles par rapport au signal audio original. Ainsi, il faut calculer l'autocorrélation du cepstre du signal tatoué de telle sorte que le signal reçu

$$y(n) = x(n) + \alpha x(n - \tau) - \alpha x(n - \tau - \Delta) \quad (1.5)$$

puisse réduire les distortions du signal perçu et améliore sa robustesse. La valeur typique de  $\Delta$  est inférieure à trois ou quatre échantillons. Si l'auto-cepstre est supérieure à  $\delta_1$  plus qu'il l'est à  $\delta_0$ , le bit introduit est décodé comme un "1". L'augmentation de la robustesse de l'algorithme de tatouage nécessite des échos à haute énergie qui doivent être introduits, mais ceci dégrade la qualité du signal audio.

#### 1.3.1.4 Codage de phase

L'idée de base du tatouage par codage de phase est de diviser le signal audio original en blocs et intégrer le tatouage dans le spectre de phase du premier bloc [12]. L'inconvénient majeur de cette méthode de codage de phase c'est qu'elle est de faible puissance puisque seul le premier bloc est utilisé pour l'insertion du tatouage. De plus, le tatouage n'est pas dispersé sur l'ensemble des données disponibles, mais il est implicitement localisé et peut donc être facilement supprimé par un simple recadrage<sup>11</sup>. La détection par cette méthode est non aveugle, ce qui limite le nombre d'applications qui peuvent lui faire recours.

#### 1.3.1.5 Méthodes de Patchwork

La technique du patchwork a été présentée la première fois pour intégrer le tatouage dans les images numériques [12]. C'est une méthode statistique basée sur des tests d'hypothèses et qui s'appuie sur de grands ensembles de données. Le processus d'insertion du tatouage utilise une méthode pseudo-aléatoire pour insérer le tatouage dans le signal audio qui est extrait à l'aide d'indices numériques (comme la valeur moyenne), décrivant la distribution spécifique. La méthode est généralement appliquée dans un domaine Transformé (Fourier, ondelettes, etc) afin de diffuser le tatouage dans le domaine temporel et pour renforcer la robustesse contre les modifications de traitement du signal.

#### 1.3.1.6 Tatouage par étalement de spectre

Une des techniques les plus utilisées en tatouage audio est celle connue sous le nom d'étalement de spectre [11, 28, 69].

---

11. Cette attaque connue sous le nom du cropping, est une opération qui consiste à supprimer une partie du signal tatoué dans le but de l'améliorer ou de modifier son format.

Soit le vecteur  $x$  représentant le signal hôte original et le vecteur  $y$  représentant le signal tatoué après sa distorsion. Une clef secrète  $K$  est utilisée par une séquence pseudo-aléatoire  $u$  de moyenne nulle, et dont ses éléments sont soit  $+\rho_u$  ou  $-\rho_u$ . La séquence est multipliée par le message secret binaire  $m$ , qui est convertie en une suite de '1' et de '-1'. Ensuite, elle est ajoutée au signal audio original. Le signal tatoué  $y$  est noté par :

$$y = x + mu \quad (1.6)$$

La distorsion  $D$  causée par le tatouage est donnée donc par :

$$D = \|y - x\| = \|mu\| = \|u\| = \rho_u \quad (1.7)$$

où  $\| \cdot \|$  est la norme définie par  $\|x\| = \sqrt{\langle x, x \rangle}$  et  $\langle \cdot \rangle$  est le produit scalaire calculé par :

$$\langle x, u \rangle = \sum_{i=0}^{N-1} x_i u_i \quad (1.8)$$

avec  $N$  est la taille des vecteurs  $x$  et  $u$ . Le bruit  $n$  est considéré comme étant additif et le signal tatoué résultant s'écrit sous la forme suivante

$$\hat{y} = y + n \quad (1.9)$$

La corrélation normalisée est calculée entre le signal tatoué et la séquence pseudo-aléatoire  $u$  pour extraire le tatouage.

$$r = \frac{\langle \hat{y}, u \rangle}{\langle u, u \rangle} = \frac{\langle mu + x + n, u \rangle}{\rho_u^2} = m + x_u + n_u \quad (1.10)$$

où

$$x_u = \frac{\langle x, u \rangle}{\rho_u^2} \quad (1.11)$$

et

$$n_u = \frac{\langle n, u \rangle}{\rho_u^2} \quad (1.12)$$

Ensuite le tatouage extrait est donné par :

$$\hat{m} = \text{sign}(r) \quad (1.13)$$

La corrélation normalisée est considérée aussi comme un processus gaussien  $r \sim N(m_r, \rho_r^2)$  avec une moyenne

$$m_r = E(r) = m \rho_r^2 = \frac{\rho_x^2 + \rho_n^2}{N \rho_u^2} \quad (1.14)$$

Enfin, la probabilité d'erreur est donnée par l'équation

$$\begin{aligned} Pe &= Pr\{\hat{m} < 0 | m = 1\} \\ &= Pr\{\hat{m} > 0 | m = -1\} \\ &= \frac{1}{2} \text{erfc}\left(\frac{m_r}{\rho_r \sqrt{2}}\right) \\ &= \frac{1}{2} \text{erfc}\left(\sqrt{\frac{N \rho_u^2}{2(\rho_x^2 + \rho_n^2)}}\right) \end{aligned} \quad (1.15)$$

où  $\text{erfc}$  est la fonction d'erreur complémentaire.

### 1.3.2 Domaine d'insertion du tatouage

Le domaine d'insertion du tatouage a une grande importance dans la conception du système de tatouage. En effet, il a un grand impact sur les performances en termes d'inaudibilité, de fiabilité de détection et aussi de robustesse.

Si le domaine d'insertion est le domaine temporel, l'information est incorporée directement dans les amplitudes du signal audio. Bien que ce domaine offre un débit élevé, il ne permet pas une forte robustesse contre les perturbations externes [40]. Alors que si l'insertion se fait dans un domaine transformé, le tatouage doit être situé dans les coefficients invariants aux perturbations communes comme la compression, le rééchantillonnage, le filtrage passe-bas, et d'autres opérations de traitement du signal [39]. L'idée est que, après l'opération de transformation, un changement important dans le signal diminuerait considérablement la qualité subjective du signal audio tatoué.

#### 1.3.2.1 Domaine temporel

Le tatouage peut être ajouté au signal audio dans le domaine temporel [11, 31]. Ce domaine permet une insertion directe du tatouage dans le signal hôte en se basant sur des modèles psychoacoustiques qui assurent la contrainte d'inaudibilité.

#### 1.3.2.2 Domaine transformé

Ce domaine est souvent choisi pour son invariance vis-à-vis de certaines attaques désynchronisantes : en effet, la représentation fréquentielle est robuste aux modifications de l'échelle de temps couramment appliquées aux signaux audio. Le tatouage peut être ajouté au signal audio dans plusieurs domaines :

- **Domaine fréquentiel** : le tatouage peut être inséré dans les coefficients FFT<sup>12</sup> [62, 100, 104], dans le domaine des sous-bandes [69, 77, 105] ou encore dans le domaine DCT<sup>13</sup> [7]
- **Transformée en ondelettes** : l'insertion dans le domaine DWT<sup>14</sup> permet une robustesse élevée par rapport aux deux domaines DFT<sup>15</sup> et DCT [7].
- **Domaine cepstral** : ce domaine, très souvent considéré pour le tatouage des signaux de parole [75, 76], contient l'information la plus significative de la parole : il est du fait robuste à un ensemble de traitements sur les signaux dont les opérations désynchronisantes.
- **Domaine EMD**<sup>16</sup> : ce domaine d'insertion permet premièrement, d'augmenter la fiabilité de détection en insérant la même séquence binaire dans les composantes IMFs<sup>17</sup> hôtes, et d'augmenter ensuite le taux d'information binaire dans le signal audio en maintenant des séquences binaires différentes [48, 50].

---

12. Fast Fourier Transform.

13. Discrete Cosine Transform.

14. Discrete Wavelet Transform.

15. Discrete Fourier Transform.

16. Empirical Mode Decomposition.

17. Intrinsic Mode Functions.

### 1.3.3 Techniques améliorées de tatouage par étalement de spectre

La technique la plus utilisée en tatouage numérique est celle par étalement de spectre. Ceci est dû à plusieurs raisons :

- Insertion du signal de tatouage avec une puissance faible et donc une inaudibilité assurée.
- Robustesse contre les interférences entre symboles et donc une bonne fiabilité de détection.
- Sécurité et confidentialité du tatouage.
- Utilisation de toute la largeur de bande du signal audio.

Ainsi, on se focalise sur les méthodes de tatouage par étalement de spectre qui ont été améliorées au niveau du récepteur.

#### 1.3.3.1 Filtrage adapté

Le filtre adapté est généralement utilisé dans de tels cas pour détecter l'existence des séquences pseudo-aléatoires et de localiser précisément l'échantillon de départ de la séquence pseudo-aléatoire. Puisque la séquence pseudo-aléatoire ne dure qu'une très courte période de temps, elle peut être traitée comme une suite d'impulsions de bruit transitoire et détectée par un filtre dont la réponse impulsionnelle est adaptée à la séquence pseudo-aléatoire. Ceci peut être réalisé grâce au filtre adapté dont la réponse en fréquence est définie comme suit [112] :

$$H(f) = K \frac{PN^*(f)}{S_x(f)} \quad (1.16)$$

où  $K$  est le facteur d'échelle,  $PN^*(f)$  est le conjugué complexe du spectre de la séquence pseudo-aléatoire  $u$  et  $S_x(f)$  est la densité spectrale de puissance du signal hôte  $x$ . Dans les applications réelles, le signal hôte est très identique à un processus à moyenne nulle avec une variance  $\sigma_x^2$ . D'où, l'équation (1.16) devient

$$H(f) = \frac{K}{\sigma_x^2} PN^*(f) \quad (1.17)$$

avec une réponse impulsionnelle

$$h(n) = \frac{K}{\sigma_x^2} u(-n) \quad (1.18)$$

Le signal tatoué devient

$$y(n) = x(n) + \alpha u(n) \quad (1.19)$$

où  $\alpha$  est un facteur de contrôle d'amplitude. Ensuite, la sortie du filtre adapté est

$$\begin{aligned} o(n) &= \frac{K}{\sigma_x^2} u(-n) * (x(j) + \alpha u(j)) \\ &= \frac{K}{\sigma_x^2} u(-n) * x(j) + \frac{\alpha K}{\sigma_x^2} u(-n) * u(j) \\ &= d_{ux} + d_u \end{aligned} \quad (1.20)$$

Puisque le signal hôte et la séquence pseudo-aléatoire ne sont pas corrélés, le terme  $d_{ux}$  devient nul. La séquence pseudo-aléatoire est aussi orthogonale au signal hôte, et  $d_u$  atteint

ainsi son maximum, si la séquence est parfaitement alignée, sinon, elle sera nulle. Par conséquent, l'emplacement du début de la séquence pseudo-aléatoire détectée dans un bloc est exprimée par :

$$pn_{emp} = argmax_i(o(n)) \quad (1.21)$$

La détection par le filtre adapté est optimale pour un bruit additif blanc gaussien. Cependant, le signal audio hôte est généralement loin d'avoir des caractéristiques d'un bruit blanc gaussien. L'échantillon audio adjacent est généralement corrélé avec une grande variance, ce qui augmente l'erreur de détection des modulations par étalement de spectre.

### 1.3.3.2 Algorithme de lissage : Savitzky-Golay

Cvejic dans [30] a utilisé les filtres de lissage Savitzky-Golay à moindres carrés dans le domaine temporel pour lisser le bruit audio à large bande. Ainsi, la variance du signal audio hôte est fortement réduite. Par conséquent, le taux d'erreur de détection va diminuer.

Une autre manière de diminuer la variance du signal audio peut être faite grâce au processus de blanchiment, qui soustrait une moyenne mobile à partir du spectre de fréquence du signal reçu, juste avant la corrélation. Cette méthode, bien qu'elle supprime une partie de la corrélation dans le signal audio, elle supprime également une partie du tatouage puisque la marque insérée est étalée sur toute la bande de fréquence.

### 1.3.3.3 Filtrage cepstral

Kirovski et al dans [69] ont proposé une méthode de filtrage cepstral pour réduire la variance qui fonctionne comme suit :

- Le signal tatoué  $y$  est un vecteur MCLT<sup>18</sup>.
- Calcul du cepstre de l'amplitude du vecteur MCLT de  $y$  par la DCT ;  $Z = DCT(y)$ .
- Filtrer les  $K$  premiers coefficients cepstraux,  $z_i = 0, i = 1...K$ .
- Appliquer l'inverse de la DCT à  $Z$  et le signal résultant passe par le détecteur de corrélation.

L'idée derrière cette méthode est que les grandes variations dans le signal tatoué peuvent y provenir à partir des grandes variations dans le signal hôte  $x$ , considérant que l'amplitude du tatouage est faible. En éliminant les grandes variations en  $y$ , la variance du signal hôte  $x$  est fortement réduite donnant une meilleure détection.

### 1.3.3.4 Filtrage de l'enveloppe spectrale

Une autre manière pour améliorer la détection est de minimiser l'effet de corrélation du signal original. Au cours de la détection, le signal hôte est considéré comme du bruit qui dégrade les performances de détection en raison de sa forte énergie. Ce bruit est beaucoup plus fort en tatouage audio qu'en tatouage image puisque le système auditif humain est plus sensible que le système visuel humain. Afin d'éliminer de manière aveugle les effets de signal hôte pour

---

18. Modulated Complex Lapped Transform.

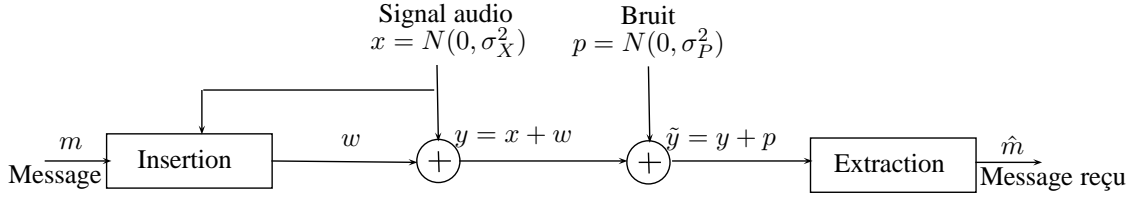


FIGURE 1.10 – Système de tatouage comme une chaîne de communication.

une bonne extraction du tatouage, Jung et al dans [63] ont proposé une méthode qu'ils ont appelé filtrage de l'enveloppe spectrale. Cette méthode supprime l'enveloppe spectrale du signal de réception et réduit la variance de bruit de la valeur de corrélation au niveau du détecteur. Étant donné que la majeure partie des caractéristiques spectrales de l'enveloppe du signal tatoué provienne du signal hôte, le filtrage de l'enveloppe spectrale permet de réduire les effets du signal hôte et d'améliorer la détection du tatouage. Les étapes de ce procédé sont les suivantes :

- Appliquer un filtre passe-bas LPF<sup>19</sup> au signal tatoué pour obtenir l'enveloppe spectrale du vecteur  $Y$ .

$$Y = LPF(\log(X)) \quad (1.22)$$

- Supprimer  $Y$  du  $\log(X)$ .
- Convertir les données filtrées en échelle linéaire.

$$\hat{X} = e^{(\log(X)-Y)} \quad (1.23)$$

Le procédé de filtrage de l'enveloppe spectrale améliore ainsi les performances de détection.

### 1.3.4 Tatouage audio en tant que problème de communication numérique

Le processus fondamental dans chaque système de tatouage peut être modélisé sous forme d'une chaîne de communication numérique où un message est transmis à partir de l'émetteur vers le récepteur [83]. Le procédé de tatouage peut être considéré comme la transmission d'un message de tatouage via un canal particulier à deux bruits additifs (c.f. FIGURE 1.10) ; le premier bruit est le signal audio et le second représente les perturbations externes appliquées au signal audio tatoué. Les deux bruits sont supposés être des processus aléatoires indépendants, notés  $X$  et  $P$ , de lois de probabilités respectives  $N(0, \sigma_X^2)$  et  $N(0, \sigma_P^2)$ . L'émetteur choisit alors le tatouage  $w$  le plus adapté pour représenter le message  $m$ . Ce choix doit assurer la contrainte d'inaudibilité qui est modélisée sous la forme d'un contrôle de puissance. En effet, la puissance du tatouage  $w$ , notée  $\sigma_w^2$ , doit être choisie inférieure ou égale à un seuil, noté  $\theta$  [9] :

$$\sigma_w^2 \leq \theta \quad (1.24)$$

En construisant le tatouage  $w$  indépendamment du signal audio  $x$ . Le signal  $x$  est alors considéré comme un bruit perturbateur au niveau de la détection. Le système de tatouage non-informé résultant est un système de communication standard de bruit BBAG. Ces performances

19. Low Pass Filter

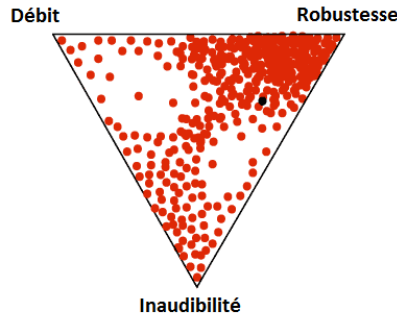


FIGURE 1.11 – Triangle magique.

en terme de capacité d'insertion  $C_{NI}$  sont données par [101] :

$$C_{NI} = \frac{1}{2} \log_2 \left( 1 + \frac{\sigma_w^2}{\sigma_X^2 + \sigma_P^2} \right) \quad (1.25)$$

## 1.4 Protocole expérimental

Les contraintes présentées précédemment sont générales au problème du tatouage numérique. Naturellement, elles dépendent du contexte envisagé. Dans le cas spécifique du tatouage audio considéré dans cette thèse, on se focalise sur le contexte non sécuritaire.

Nous cherchons à avoir le débit d'information le plus élevé possible tout en gardant dans l'esprit que le système doit prendre en compte les perturbations externes du canal. Les attaques des pirates destinées à enlever le message ne sont pas pris en compte car les signaux à transmettre ne sont pas des données sensibles. Enfin, une copie du signal original ne sera pas disponible pour réaliser la détection. Le système de détection sera donc développé pour permettre une détection aveugle.

Naturellement, un compromis existe entre les différentes contraintes (c.f. FIGURE 1.11). En effet, une augmentation importante de la robustesse entraînera souvent une augmentation de la distorsion du support (si l'on travaille à débit fixe). Ceci est dû au fait que pour garantir la robustesse, le système d'insertion devra augmenter la puissance du message inséré qui peut devenir perceptible. D'autre part, on peut aussi travailler à débit variable. Dans ce cas, les systèmes qui ont une robustesse importante peuvent utiliser un débit d'insertion faible afin de maintenir la distorsion sous un niveau minimal.

### 1.4.1 Signaux tests

L'insertion et la détection du tatouage dépendent fortement du type des signaux à tatouer [66]. Pour cela, l'évaluation est faite sur un sous-ensemble de fichiers audio de la base SQAM<sup>20</sup> échantillonnés à 44.1 kHz et quantifiés à 16 bits (c.f. Tableau 1.2).

---

20. Sound Quality Assessment Material, disponible à <http://soundexpert.org/sound-samples>

Type	Contenu
<b>Instruments simples</b>	1-Violin ; 2-Viola ; 3-Violoncello ; 4-Double-bass ; 5-Piccolo ; 6-Flute ; 7-Oboe ; 8-Cor anglais ; 9-Clarinet ; 10-Bass-clarinet ; 11-Bassoon ; 12-Contra-bassoon ; 13-Saxophone ; 14-Trumpet ; 15-Trombone ; 16-Horn ; 17-Tuba ; 18-Harp ; 19-Claves ; 20-Castanets ; 21-Side drum ; 22-Bass drum ; 23-Kettle-drums ; 24-Cymbal ; 25-Triangle ; 26-Gong ; 27-Tubular bells ; 28-Glockenspiel ; 29-Xylophone ; 30-Vibraphone ; 31-Marimba ; 32-Grand piano ; 33-Harpsichord ; 34-Celesta ; 35-Accordion ; 36-Organ.
<b>Vocal</b>	37-Soprano ; 38-Alto ; 39-Tenor ; 40-Bass ; 41-Quartet.
<b>Parole</b>	42-English Female speech, 43-English male speech 44-French Female speech, 45-French male speech 46-German Female speech, 47-German male speech
<b>Instruments solistes</b>	48-Haydn Trumpet ; 49-Handel Organ ; 50-Bach Organ ; 51-Guitar ; 52-Ravel Violin ; 53-Schubert Piano.
<b>Vocal et Orchestre</b>	54-Soprano Mozart ; 55-Soprano Spiritual ; 56-Soloists Verdi ; 57-Choir Orff.
<b>Orchestre</b>	58-Orchestra R. Strauss ; 59-Wind ensemble Stravinsky ; 60-Wind ensemble Mozart ; 61-Orchestra Baird.
<b>Musique Pop</b>	62-ABBA ; 63-Eddie Rabbitt.

Tableau 1.2 – Signaux audio de test appartenant à la base de données SQAM.

### 1.4.2 Inaudibilité

La première contrainte imposée au système de tatouage audio est la contrainte d'inaudibilité. Le signal tatoué doit être le plus identique possible au signal original d'un point de vue perceptif. Seul un ensemble de tests d'écoute, réalisés par un grand nombre d'auditeurs, peut établir de manière fiable la transparence du tatouage. Il s'avère malheureusement longs et coûteux surtout en phase de conception du système. Il convient alors de disposer d'une mesure objective de la différence perceptive de deux signaux, qui permet d'évaluer la performance du système pendant la phase de conception.

Pour mesurer la différence perceptive entre les signaux lors de la conception d'un système de tatouage sans mettre en oeuvre des tests d'écoute coûteux, plusieurs mesures objectives ont été proposées [80]. Bien que ces mesures de distorsion sont directement liées à la puissance du tatouage, elles sont souvent considérées comme insuffisantes pour mesurer la qualité du signal audio. Dans cette thèse, on va se servir d'une méthode de mesure objective basée sur l'algorithme

PEAQ<sup>21</sup> [64]. Cette méthode fournit une classe de différence objective (ODG<sup>22</sup>) révélant la qualité du signal tatoué de manière suffisamment fiable comme illustré au tableau 1.3.

ODG	Qualificatif de la dégradation
0	Imperceptible
-1	Perceptible mais non gênante
-2	Légèrement gênante
-3	Gênante
-4	Très gênante

Tableau 1.3 – Échelle de dégradation à cinq notes et valeurs de l'ODG associées.

### 1.4.3 Robustesse

Le Taux d'Erreur Binaire (TEB) est le critère d'évaluation des performances de la détection en communications numériques. Ce critère est défini comme étant le rapport du nombre de bits détectés erronés sur le nombre total de bits émis. Ce même critère sera utilisé ici pour évaluer et comparer les performances de la réception du tatouage. Les perturbations, auxquelles le système de tatouage dédié à l'augmentation du contenu doit être robuste, sont l'ensemble des opérations licites qui peuvent être appliquées au signal audio tatoué<sup>23</sup>. Les perturbations envisagées pour évaluer la robustesse, et qui sont adaptées à notre contexte applicatif de transmission des données utiles [107], sont donc illustrées au tableau 1.4 :

- **Opérations de filtrage** : utilisées pour augmenter ou supprimer certaines parties du spectre du signal audio.
- **Distorsions de la dynamique** : consistent à modifier la sonie du signal audio ; il s'agit de l'augmenter ou de la diminuer dans la plupart des cas.
- **Ajout du bruit** : lors d'un enregistrement d'un signal audio tatoué analogique dans un environnement bruité.
- **Les changements de format** : cette catégorie regroupe le rééchantillonnage ou l'interpolation des signaux à une fréquence d'échantillonnage différente du signal original, la modification du nombre de bits utilisés pour la quantification des échantillons et la compression du signal (compression MPEG par exemple).

---

21. Perceptual Evaluation of Audio Quality.

22. Objective Difference Grade.

23. comme nous l'avons déjà précisé, les attaques malveillantes ne seront prises en compte.

Type	Perturbations	Paramètres
Opérations de filtrage	Filtrage passe-bas	fréquence de coupure : 8kHz
	Filtrage passe-haut	fréquence de coupure : 2kHz
Distorsions de la dynamique	Modification de la sonie	Multiplier l'amplitude par 0.5
Bruit ajouté	BBAG	RSB : 20dB
Changement de format	Compression MPEG	Débit : 64kb/s, 96kb/s et 128kb/s
	Rééchantillonnage	De 44.1kHz à 32kHz
	Quantification	Résolution : 12Hz

Tableau 1.4 – Perturbations utilisées pour évaluer les performances en terme de robustesse.

#### 1.4.4 Débit et coût

L'une des contraintes les plus importantes d'un système de tatouage audio est le débit d'insertion, habituellement déterminé par les exigences spécifiques de l'application envisagée. Le débit binaire du tatouage est le nombre de bits insérés par unité de temps généralement calculé en bits par seconde.

Le temps de calcul est défini comme étant le rapport entre le temps nécessaire à la simulation et la durée du signal audio traité. Cette mesure est très dépendante du logiciel de programmation (ici MATLAB 2013) et du processeur utilisé (ici Intel(R) Core(TM) i5 CPU@ 1.80GHz avec 4GO de RAM). Elle sera donnée particulièrement à titre comparatif, puisque l'implémentation des systèmes n'a pas été optimisée.

### 1.5 Approche suivie dans cette thèse

#### 1.5.1 Applications envisagées

Dans cette thèse, on vise le contexte non sécuritaire du tatouage audio. Plus particulièrement, le domaine d'applications que nous envisageons est ce qu'on va appeler "**tatouage pour l'enrichissement du contenu audio**". Il s'agit de véhiculer des données supplémentaires à destination d'une application cible. Une application intéressante qui vise l'enrichissement du contenu audio est l'insertion dans un signal musical des méta-données contenant des informations sur le titre de la chanson, l'artiste, le nom de l'album ainsi que le genre de musique.

Dans ce contexte applicatif d'enrichissement du contenu, le tatouage doit satisfaire à un ensemble de contraintes spécifiques :

- Garantir la transparence du tatouage inséré.
- Obtenir un débit d'insertion le plus élevé possible.
- Détecter le tatouage avec le minimum d'erreurs possibles.

- Assurer une robustesse face à un ensemble de distorsions licites telles que le bruit blanc gaussien et la compression MPEG.

Ces enjeux spécifiques présentent la différence entre tatouage pour la protection des droits d'auteur et celui dédié à l'enrichissement du contenu auquel nous nous intéressons.

### 1.5.2 Analyse critique de l'existant

Étant donné le domaine d'applications envisagé ainsi que ses propres contraintes, la question qui se pose alors est : quel est le système de tatouage le plus adapté à notre contexte applicatif ? Parmi toutes les techniques de tatouage existantes dans la littérature [12], il y en a deux méthodes de tatouage qui sont les plus utilisées dans tous les travaux de recherche [2].

La première méthode est le tatouage par étalement de spectre [15, 69], qui consiste à utiliser un code indépendant de la séquence à envoyer pour étaler le signal à bande étroite en un signal dont la largeur de bande est plus large. La deuxième méthode est le tatouage par la méthode de quantification [68], qui consiste à modifier les niveaux d'amplitude du signal hôte par un quantificateur qui dépend du message à cacher. Les caractéristiques de ces deux méthodes de tatouage sont [68] :

- Les techniques d'étalement de spectre garantissent une grande robustesse contre un faible débit.
- Les méthodes de quantification assurent un haut débit de transmission contre une robustesse très limitée.

Il serait donc nécessaire de choisir la technique de tatouage la plus adaptée aux contraintes de l'application envisagée. Par exemple, dans le cas d'un tatouage de type sécuritaire, le premier critère à vérifier est la robustesse. Alors que pour le tatouage destiné à l'enrichissement du contenu, il exige un haut débit d'insertion mais dans ce type d'application, il n'existe pas d'attaques malveillantes.

### 1.5.3 Contributions

Le système de tatouage auquel on se réfère est présenté sur la FIGURE 1.6. Ce système se compose de trois grands blocs : émetteur, canal et récepteur. Après une analyse approfondie des travaux présentés en littérature, on a remarqué que la plupart des chercheurs se sont focalisés sur le canal et le récepteur tout en admettant la partie émettrice comme elle a été présentée<sup>24</sup> dans [15]. On propose donc de se focaliser sur la partie émettrice et améliorer ses performances. Mais dans quelle partie de l'émetteur faut-il intervenir ? Est-ce dans la partie codage ou la partie mise en forme ou au encore au niveau de la modulation.

Le codage n'a d'influence que sur la fiabilité de détection<sup>25</sup>. Alors que la mise en forme du signal de tatouage n'a un impact que sur l'inaudibilité. Le seul bloc qui reste au niveau de l'émetteur est la modulation. Ainsi, les travaux de cette thèse s'articulent autour de la modulation numérique pour améliorer les performances du système de tatouage audio. Le problème peut être abordé comme la définition d'une modulation appropriée pour obtenir les meilleures

---

24. Premier article apparu en tatouage audio.

25. Si on veut minimiser le TEB.

performances du système de tatouage. Le choix de la modulation parmi tous les autres blocs du système n'est pas arbitraire. En effet, la modulation numérique a un impact sur l'inaudibilité, la fiabilité de détection, la robustesse ainsi que le débit d'insertion. Ainsi, les axes traités dans cette thèse se font à plusieurs échelles :

- **Modulations à bande étroite** : garantir l'inaudibilité en adaptant les paramètres du signal modulé à ceux du filtre de mise en forme déjà conçu pour le cas d'utilisation de l'étalement de spectre.
- **Modulations multi porteuses** : améliorer la robustesse du tatouage en répartissant la puissance du signal de tatouage sur plusieurs porteuses différentes.
- **Modulations multi-tatouages** : insérer plusieurs tatouages en parallèle dans le même signal audio en s'inspirant du principe des méthodes d'accès multiples.
- **Tatouage non aveugle** : exploiter la connaissance à priori du signal audio récupéré au niveau du récepteur par la technique de séparation de source pour améliorer la fiabilité de détection.
- **Tatouage double canal** : augmenter le débit d'insertion en mixant plusieurs tatouages dans les deux canaux du signal audio.

#### 1.5.4 Organisation de la thèse

Étant donné le contexte et l'objectif de cette thèse, il paraît nécessaire de présenter la structure générale de ce manuscrit.

Le chapitre 2 consiste en la présentation des différentes entités qui constituent le système de tatouage du point de vue de communication numérique ; à savoir l'émetteur, le canal et le récepteur. Ensuite, les similitudes ainsi que les différences, entre un système de tatouage et celui de communication numérique permettront en particulier d'employer certains outils de la théorie des communications en tatouage et d'en dégager des techniques d'insertion, et plus précisément des techniques de modulations, que nous adapterons au contexte du tatouage. En particulier, nous présenterons les modulations par étalement de spectre qui sont les plus utilisées dans les systèmes de tatouage. On s'intéressera à deux types de modulations : modulation par étalement de spectre à séquences directe et celle par saut de fréquences.

Le chapitre 3 présente deux systèmes de tatouage basés sur les modulations sinusoïdales. Tout d'abord, on va commencer par mettre en valeur les avantages des modulations à bande étroite. On va montrer comment insérer de l'information dans la phase du signal de tatouage pour assurer un double gain en termes de fiabilité de détection et de transparence. Ensuite, on va étudier l'impact de la fréquence porteuse du signal modulé sur l'inaudibilité et la fiabilité de détection. Un système de tatouage informé est ensuite présenté, ayant comme objectif d'adapter les paramètres du signal modulé pour améliorer les performances en tatouage audio. Une dernière section est consacrée aux modulations multi-porteuses visant à améliorer la robustesse du tatouage contre les perturbations externes. On va montrer comment on peut contrôler la puissance du signal multi-porteuse afin d'assurer la contrainte d'inaudibilité. Enfin, on va évaluer la robustesse du système multi-porteuse face au bruit blanc et à la compression MPEG.

Le chapitre 4 s'intéresse à l'augmentation du débit d'insertion dans un système de tatouage audio. On va montrer comment le système de tatouage peut être modélisé comme un système de

communication multi-utilisateurs particulier. La nouvelle modulation utilisée dans ce système prendra le nom de modulation multi-tatouage et est inspirée des techniques d'accès multiples largement utilisées en télécommunications. Son principe est de permettre l'insertion simultanée de plusieurs tatouages dans le même signal audio. Trois techniques d'accès multiples seront discutées ; l'accès multiple par répartition en codes, l'accès multiple par saut de fréquences et l'accès multiple par répartition en fréquences. Ensuite, on va contrôler l'inaudibilité du tatouage par un facteur d'amplitude qui permettra l'insertion de plusieurs tatouages avec une qualité audio presque identique à celle du signal original. A la réception, on va proposer une nouvelle procédure d'extraction exploitant le concept de séparation de sources. Enfin, on vise à augmenter le débit d'insertion du système multi-tatouage dans le cas particulier d'un signal double canal.

Le chapitre 5 présente une idée originale qui consiste à utiliser le tatouage numérique des signaux audio dans la transmission radio. Notre but sera d'exploiter l'afficheur numérique de la radio comme un petit écran sur lequel on peut afficher des informations qui ont été insérées dans l'émission diffusée avant sa transmission. Afin d'atteindre ce but, notre application interviendra à deux niveaux : au niveau de la station radio où l'insertion est faite, puis au niveau des clients (auditeurs de la chaîne radio) où les phases d'extraction et d'affichage seront réalisées.

## Chapitre 2

# Tatouage audio comme problème de communication numérique

Dans ce chapitre nous allons étudier les éléments qui nous permettront d'aborder le tatouage par une approche héritée des télécommunications [53]. Cette analyse prend en compte les particularités liées aux études psychoacoustiques, notamment dans la construction du tatouage. En effet, l'objectif est de présenter la conception d'un système de tatouage audio, en exploitant l'analogie entre tatouage audio et communication numérique. Le système présenté est conçu sous forme d'un système de communication particulier, reprenant les deux étapes classiques qui sont :

- L'émetteur qui a pour rôle de construire le signal de tatouage  $w(n)$  qui va être inséré dans le signal audio hôte  $x(n)$ . Cette insertion du tatouage produira le signal audio tatoué  $y(n)$  dont la qualité doit être presque identique au signal original  $x(n)$ .
- Le récepteur qui a pour rôle d'extraire l'information cachée à partir seulement du signal tatoué  $y(n)$ .

Ce chapitre met également l'accent sur quelques différences entre le système de tatouage et celui de communication numérique. L'origine de cette différence se situe au niveau des conditions de transmission du signal de tatouage, qui sont généralement différentes du point de vue de la communication numérique.

### 2.1 Tatouage audio du point de vue de la communication numérique

Le système de tatouage peut être étudié comme un système de communication numérique très particulier. En effet, l'émetteur joue le rôle du système d'insertion de l'information dans le signal audio et le récepteur joue le rôle de la détection de cette information. Dans cette partie on propose de rappeler les outils de la théorie des communications numériques nécessaires à la conception et à l'évaluation des performances théoriques du système de tatouage.

La conception du système de tatouage se base principalement sur le fait que le signal de tatouage  $w(n)$  est ajouté directement au signal audio hôte  $x(n)$  dans le domaine temporel. Le

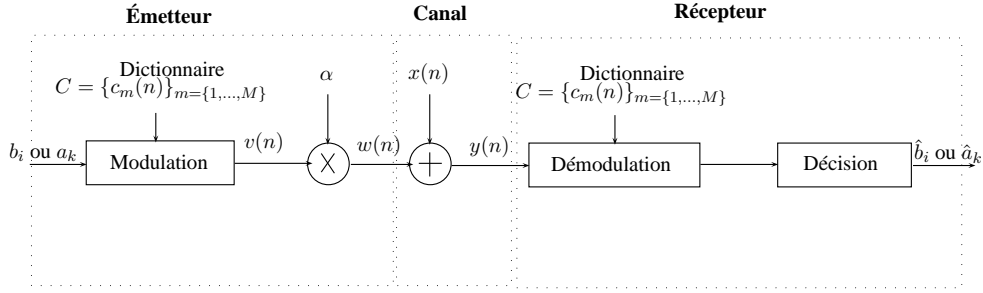


FIGURE 2.1 – Modèle synthétique de la chaîne de tatouage.

signal audio tatoué  $y(n)$  s'écrit donc :

$$y(n) = x(n) + w(n) \quad (2.1)$$

Ce formalisme établit le lien entre un système de tatouage et celui de communication numérique comme c'est illustré dans la FIGURE 2.1. Le signal de tatouage  $w(n)$  est le signal utile, porteur de l'information à transmettre, tandis que le signal audio  $x(n)$  est le bruit caractérisant le canal de transmission.

### 2.1.1 Modulation en bande de base

L'information de tatouage  $b_i$  est représentée initialement sous forme d'une suite de bits  $\{0,1\}$  indépendants et identiquement distribués. Le temps  $T_b$  pour transmettre un bit est lié à la fréquence d'échantillonnage  $F_e = \frac{1}{T_e}$  du signal audio par :

$$T_b = N_b T_e \quad (2.2)$$

où  $N_b$  est le nombre d'échantillons d'un seul bit d'information et  $T_e$  est la période d'échantillonnage. On regroupe la séquence  $b_i$  en  $q$ -uplets pour former une séquence de symboles  $a_k$  (chaque symbole  $a_k$  de durée  $T_s = qT_b$  est choisi parmi  $M = 2^q$  symboles). La séquence de symboles  $a_k$  est transmise ensuite avec un débit de transmission qui s'exprime par :

$$R = \frac{F_e}{N_b} = \frac{qT_e}{N_s} \quad (2.3)$$

où  $N_s$  est le temps symbole discret. La séquence de symboles  $a_k$  passe ensuite par une modulation en bande de base ou codage en ligne, pour se transformer en un signal continu mieux adapté au canal de communication. La séquence  $a_k$  est cadencée dans le temps en associant à chaque valeur  $a_k$  un signal en temps ; ce qui génère le signal :

$$a(n) = \sum_k a_k \delta(n - kN_s) \quad (2.4)$$

où  $\delta(n)$  est la fonction de Dirac. Le signal  $a(n)$  est ensuite convolué avec :

$$\psi_{(N_s)}(n) = \begin{cases} A & \text{si } 0 \leq n \leq N_s \\ 0 & \text{sinon} \end{cases} \quad (2.5)$$

Le signal résultant  $s(n)$  est un signal en bande de base :

$$s(n) = \sum_k a_k \psi(n - kN_s) \quad (2.6)$$

et la fonction de mise en forme  $\psi$  est choisie de façon à ce que la largeur de bande associée corresponde bien à celle qui est disponible dans le canal. En choisissant le codage NRZ<sup>1</sup>, la densité spectrale de puissance du signal modulé en bande de base  $s(n)$  s'écrit sous la forme suivante [44] :

$$S_s(f) = A^2 T_b \text{sinc}^2(f_c T_b) \quad (2.7)$$

où  $A$  est l'amplitude,  $f_c$  est la fréquence porteuse et *sinc* est le sinus cardinal. Ce type de modulation est souvent utilisé en communication numérique, dans les transmissions à courte distance et produit une probabilité d'erreur dans un canal BBAG<sup>2</sup> représentée sous la forme suivante [44] :

$$P_b = Q\left(\sqrt{\frac{2A^2 T_b}{N_0}}\right) \quad (2.8)$$

où  $N_0$  est la DSP du bruit blanc du canal et  $Q(x)$  est l'intégrale de Gauss<sup>3</sup>. Ce type de modulation a été testé dans les systèmes de tatouage audio et ne semble pas être applicable dans un objectif de débit et de fiabilité de transmission [19]. De plus, la puissance du signal  $s(n)$  est élevée autour de la fréquence porteuse  $f_c = 0$  ce qui peut générer des distorsions d'audibilité au signal audio tatoué  $y(n)$ .

### 2.1.2 Modulation par étalement de spectre

Pour réduire la densité spectrale de puissance  $S_s(f)$  du signal  $s(n)$ , les chercheurs ont proposé d'utiliser la modulation par étalement de spectre (DSSS<sup>4</sup>) [26, 35]. En effet, ce type de modulation utilise un code d'étalement indépendant de la séquence à envoyer pour étaler le signal en bande de base  $s(n)$  en un signal dont la largeur de bande est plus large. Un symbole  $a_k$  de durée  $N_s$  est construit à partir des séquences binaires d'étalement appelées 'Chips' générées aléatoirement<sup>5</sup> chacune de durée  $N_c$ .

La modulation par étalement de spectre consiste à réaliser une association bijective entre la marque numérique  $a_k$  et le signal modulé noté  $v(n)$ . Cette étape de modulation est caractérisée par la donnée de  $M$  formes d'ondes différentes de longueur  $N_s$  contenues dans le dictionnaire de modulation  $C = c_k(n)_{k \in [0, M-1]}$ . Le signal modulé  $v(n)$  est construit par les formes d'ondes  $c_k(n)$  selon la suite de symboles provenant de la source. Les éléments du dictionnaire  $C$  constituant le signal modulé  $v(n)$  sont représentés sous formes de codes pseudo-aléatoires. Le signal modulé résultant est un signal étalé sur toute la bande de fréquence du signal audio  $x(n)$ .

Bien que le signal étalé a la même quantité d'énergie et porte les mêmes informations que le signal à bande de base, il a une densité spectrale de puissance beaucoup plus faible en raison de sa largeur de bande étalée.

- 
1. Non Return to Zero.
  2. Bruit Blanc Additif Gaussien.
  3.  $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-u^2) du$
  4. Direct Sequence Spread Spectrum.
  5. des codes pseudo-aléatoires.
-

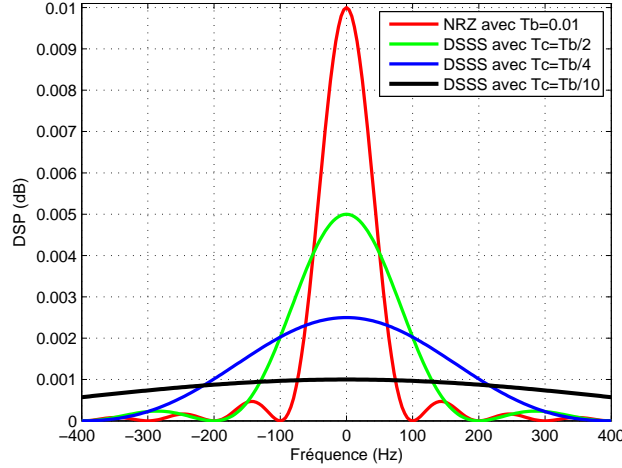


FIGURE 2.2 – DSP du signal NRZ ainsi que le signal étalé pour  $T_c = T_b/2$ ,  $T_c = T_b/4$  et  $T_c = T_b/10$ .

Le gain d'étalement, est donné par la relation existante entre le temps symbole  $N_s$  et la durée chip  $N_c$ , ou de façon équivalente, par la relation existante entre la largeur de bande d'un symbole  $B_s$  et la largeur de bande d'un chip  $B_c$ . Il s'écrit sous la forme :

$$L_c = \frac{N_s}{N_c} = \frac{B_c}{B_s} \quad (2.9)$$

où  $L_c$  est le facteur d'étalement. La FIGURE 2.2 compare la DSP  $S_s(f)$  du signal en bande de base  $s(n)$  avec la DSP  $S_v(f)$  du signal étalé  $v(n)$  en variant le facteur d'étalement  $L_c$ . On remarque que plus la valeur de  $L_c$  augmente, plus la puissance du signal  $s(n)$  est étalée sur toute la bande de fréquences. Le signal  $s(n)$  est utilisé pour moduler une séquence pseudo-aléatoire à large bande  $c(n)$  :

$$c(n) = \sum_{m=0}^{L_c-1} c_m \psi_{(N_c)}(n - mN_c), \quad c_m = \{-1, 1\} \quad (2.10)$$

Le signal résultant est obtenu en multipliant le signal  $s(n)$  par  $c(n)$ .

$$v(n) = \sum_{k=0}^{L-1} a_k c_m \psi_{(N_c)}(n - mN_c) \quad (2.11)$$

où  $L$  est le nombre de symboles émis. Dans ce cas, la DSP du signal modulé  $v(n)$  par étalement de spectre devient :

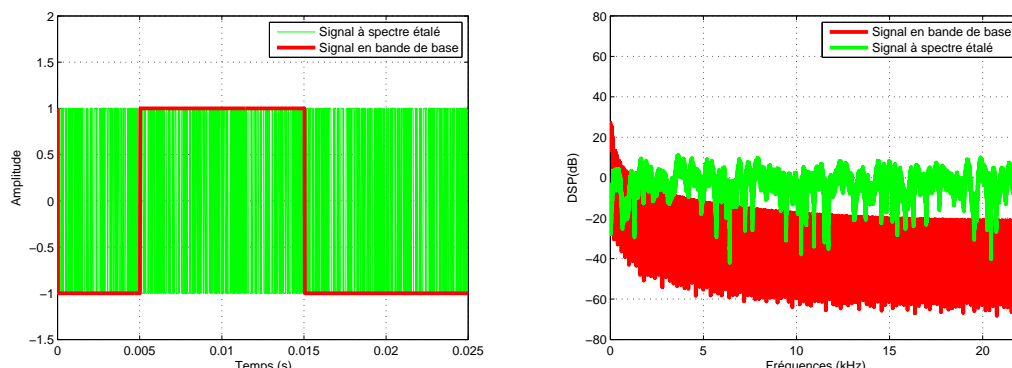
$$S_v(f) = A^2 T_c \text{sinc}^2(f_c T_c) \quad (2.12)$$

Le signal  $v(n)$  est le signal modulé par étalement de spectre et sa représentation temporelle et fréquentielle sont illustrées sur la FIGURE 2.3. Le signal modulé  $v(n)$  est ensuite multiplié par un facteur d'amplitude  $\alpha$  pour augmenter ou diminuer le RSB<sup>6</sup>. Le signal résultant  $w(n)$  appelé

---

6. Rapport Signal à Bruit.

*"Il faut se garder de trois fautes : parler sans y être invité, ce qui est impertinence ; ne pas parler quand on y est invité, ce qui est de la dissimulation ; parler sans observer les réactions de l'autre, ce qui est de l'aveuglement".*  
**Confucius**



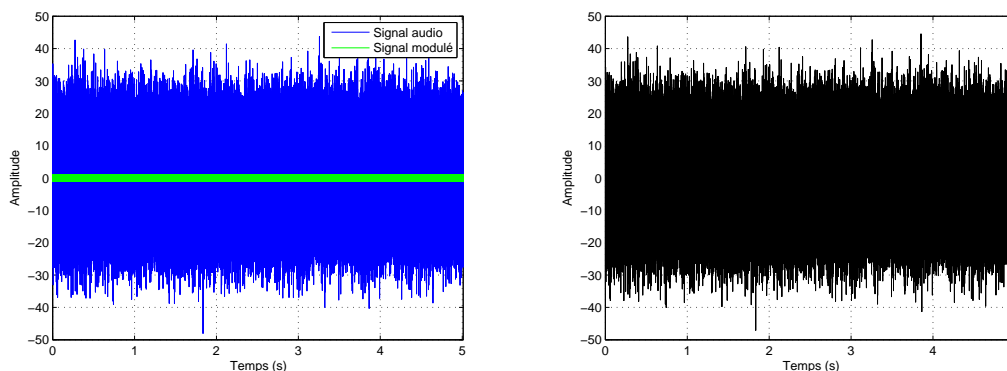
(a) Signal émis avant et après étalement de spectre dans le domaine temporel. (b) Signal émis avant et après étalement de spectre dans le domaine fréquentiel.

FIGURE 2.3 – Signal émis avant et après étalement de spectre dans les domaines temporel et fréquentiel.

signal de tatouage est ajouté au signal audio  $x(n)$ . Le signal tatoué  $y(n)$  s'écrit alors sous la forme :

$$\begin{aligned} y(n) &= x(n) + w(n) \\ &= x(n) + \alpha v(n) \end{aligned} \quad (2.13)$$

La FIGURE 2.4 illustre le processus d'insertion du signal de tatouage  $w(n)$  dans le signal audio  $x(n)$ .



(a) Représentation temporelle du signal modulé ainsi que le signal audio. (b) Représentation temporelle du signal tatoué.

FIGURE 2.4 – Illustration du signal de tatouage  $w(n)$ , du signal audio  $x(n)$  et du signal tatoué  $y(n)$ .

### 2.1.3 Récepteur optimal

Le récepteur optimal est celui qui permet de récupérer le tatouage inséré avec le minimum d'erreurs possible. Ce récepteur est complètement défini par la théorie de communication numé-

*"Les deux guerriers les plus puissants sont la patience et le temps. N'oublie pas que les grandes réalisations prennent du temps et qu'il n'y a pas de succès du jour au lendemain". Léon Tolstoï*

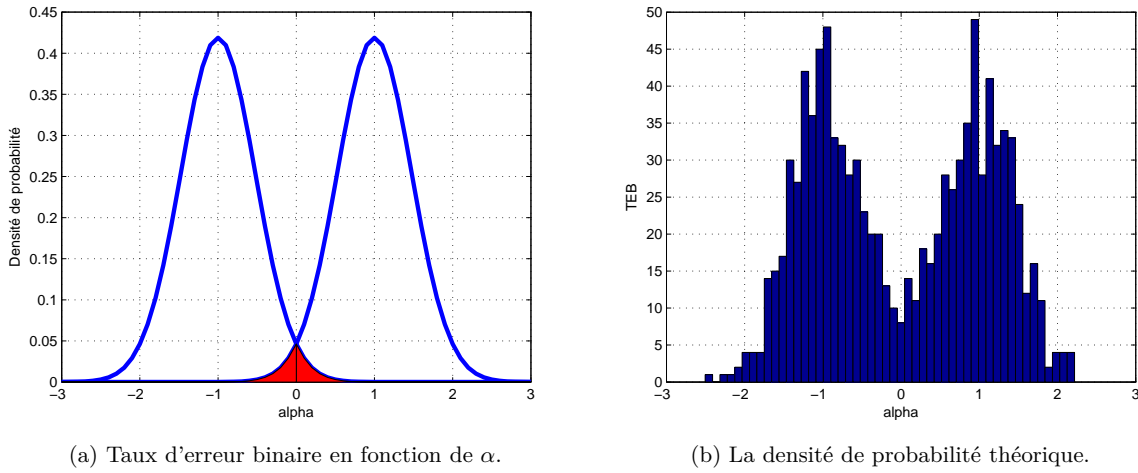


FIGURE 2.5 – Performances théoriques et pratiques de la fiabilité de transmission du tatouage.

rique [53] dans le cas d'un canal BBAG. Par conséquent, on suppose que le bruit audio  $x(n)$  est la réalisation d'un processus aléatoire  $X$  blanc gaussien de loi  $N(0, \sigma_X^2)$  et de variance  $\sigma_X^2$ .

Soit  $v$ ,  $x$  et  $y$  les vecteurs d'observations de  $v(n)$ ,  $x(n)$  et  $y(n)$  sur un horizon de  $N = LN_s$  échantillons. Le signal reçu  $y = x + \alpha v$  est la somme d'une composante déterministe  $\alpha v(n)$  et d'une composante aléatoire  $X$  dont est issue la réalisation  $x(n)$ . Le principe du récepteur optimal est donc de déterminer l'état du message numérique associé au signal  $v(n)$  qui minimise le nombre d'erreurs de transmission étant donné l'observation  $y(n)$ .

Le message numérique étant supposé être une suite de bits indépendants et identiquement distribués de  $q$ -uplets binaires. Dans ce cas particulier, la maximisation de la probabilité d'une décision correcte conduit à la règle du maximum de vraisemblance qui permet de définir la partition de l'espace d'observation. Cette règle énonce que l'état  $\hat{b}_i$  décidé est celui qui maximise la densité de probabilité suivante [53] :

$$\hat{b}_i = \underset{b_i \in \{0,1\}}{\operatorname{argmax}} (y|b_i) \quad (2.14)$$

L'expression du détecteur optimal nécessite donc d'explicitier la loi de probabilité du signal reçu  $y(n)$ . Cette loi dépend elle même de la nature du bruit audio, que nous avons dans cette section supposé BBAG de loi  $N(0, \sigma_X^2)$ . Ayant posé  $x(n)$  blanc, gaussien, centré et de variance  $\sigma_x^2$ , le vecteur reçu  $y$ , observation de  $y(n)$  sur  $N$  échantillons, peut alors être modélisé par un vecteur aléatoire  $Y$  dont les  $N$  composantes sont  $N$  variables aléatoires indépendantes. La densité de probabilité de  $Y$  conditionnellement à l'état du signal modulé  $v(n)$  est donnée par [53] :

$$P(y|b_i) = \frac{1}{\sqrt{(2\pi\sigma_x^2)^{N_b}}} \exp\left(-\frac{\|y - \alpha v\|^2}{2\sigma_x^2}\right) \quad (2.15)$$

où  $\|y - \alpha v\|^2 = \sum_{n=0}^{N_b-1} ((y(n) - \alpha v(n)))^2$ . La FIGURE 2.5 représente la densité de probabilité d'erreur donnée par l'équation (2.15) ainsi que l'histogramme du taux d'erreur binaire correspondant.

Pour évaluer les performances du système de tatouage présenté, on calcule le TEB en fonction de deux paramètres. Le premier paramètre est le débit en fixant le RSB à  $-20$ dB et le deuxième

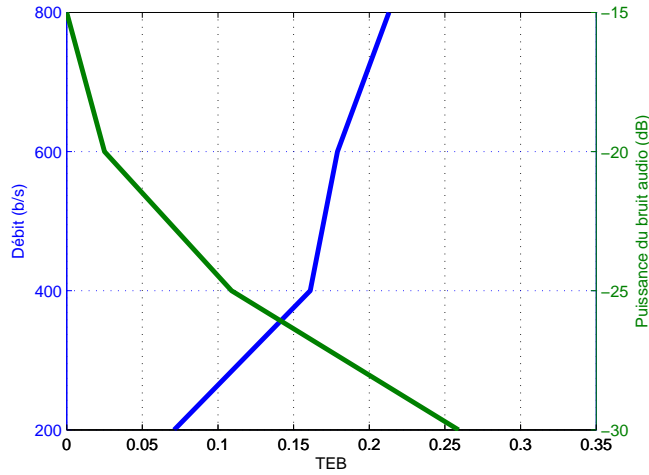


FIGURE 2.6 – Fiabilité de transmission du tatouage en fonction du débit de transmission et du RSB.

est le RSB en fixant le débit à 100b/s. La FIGURE 2.6 présente les résultats obtenus ; plus on augmente le débit ou le RSB, la fiabilité de transmission se dégrade. Par exemple, lorsque  $RSB = -20dB$ , on obtient un TEB de  $10^{-2}$  qui est acceptable. Malheureusement, ces résultats s'apparentent à un modèle théorique dans lequel le signal de tatouage serait le résultat d'une simple modulation et le bruit audio serait un bruit blanc additif gaussien. Cette modélisation est très éloignée des contraintes de transmission spécifiques au tatouage. Le récepteur optimal défini par la théorie des communications numériques doit donc être modifié pour prendre en compte les caractéristiques du bruit audio. Le tableau 2.1 présente une comparaison entre les caractéristiques du bruit en communication classique et celui en tatouage audio. On remarque que les caractéristiques des deux bruits sont parfaitement différentes ; ce qui rend les conditions de transmission du tatouage peu défavorables par rapport à la transmission en communication numérique.

Communication numérique	Tatouage audio
Blanc	Coloré
Stationnaire	Non stationnaire
Gaussien	Non Gaussien
Faible puissance	Très puissant
Non corrélé avec la réponse du canal	Corrélé avec la réponse du canal

Tableau 2.1 – Le bruit en communication numérique et en tatouage audio.

De plus, l'inaudibilité du tatouage ne peut être atteinte par une simple modulation paramétrée par un facteur d'amplitude  $\alpha$ . En effet, le signal  $v(n)$  modulé occupe toute la bande fréquentielle audio, mais sa puissance est incontrôlable. Dans la FIGURE 2.7, le signal audio considéré  $x(n)$  est un extrait d'un instrument de "Violoncello" dont les caractéristiques sont

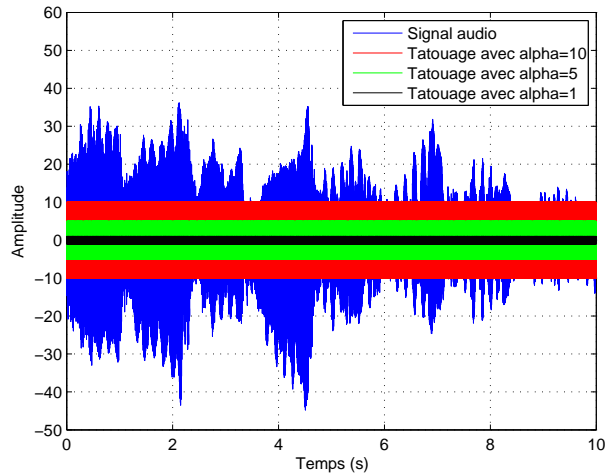


FIGURE 2.7 – Représentation temporelle du signal audio "Violoncello" et du tatouage avec différentes valeurs  $\alpha$ .

différentes du bruit blanc gaussien, et le signal de tatouage  $w(n)$  est représenté avec différentes valeur de  $\alpha$ . Si on augmente la valeur de  $\alpha$  on améliore la fiabilité de transmission mais le tatouage devient audible. D'un autre côté, si on diminue  $\alpha$  on peut garantir la transparence du tatouage mais sa détection devient difficile. Il serait alors nécessaire au signal modulé de passer par une étape de mise en forme qui permettra de lui procurer le niveau de puissance maximal sous la contrainte d'inaudibilité [35]. Ceci exige la prise en compte de la psychoacoustique, pour déterminer les conditions d'inaudibilité qu'il faut satisfaire pour modifier en conséquence la structure de l'émetteur.

## 2.2 Mise en forme perceptive

Le signal modulé  $w(n)$  ne peut pas être inséré directement dans le signal audio  $x(n)$ . En effet, la contrainte d'inaudibilité ne peut être garanti avec un simple contrôle d'amplitude par le paramètre  $\alpha$ . Pour assurer la transparence du tatouage inséré, on peut se servir de la psychoacoustique qui fournit des résultats permettant la mise en forme perceptive du signal de tatouage avant son insertion [35].

### 2.2.1 Conditions psychoacoustiques

La psychoacoustique s'attache à la modélisation du système auditif humain et la perception du son [120]. Ses résultats, essentiellement basés sur des expérimentations, mettent en évidence les défauts du système auditif, et plus particulièrement les phénomènes de masquage. Le masquage est le procédé par lequel un son, appelé son masqué est rendu inaudible en présence d'un autre, appelé son masquant. Ce phénomène, particulièrement exploité en compression audio [36], peut également être appliqué au tatouage pour déterminer les conditions d'inaudibilité que doit respecter le signal modulé  $w(n)$  pour qu'il soit transparent s'il est ajouté au signal audio  $x(n)$ . Ces conditions prennent la forme d'un seuil de masquage  $M_x(f)$  qui traduit la limite fréquen-

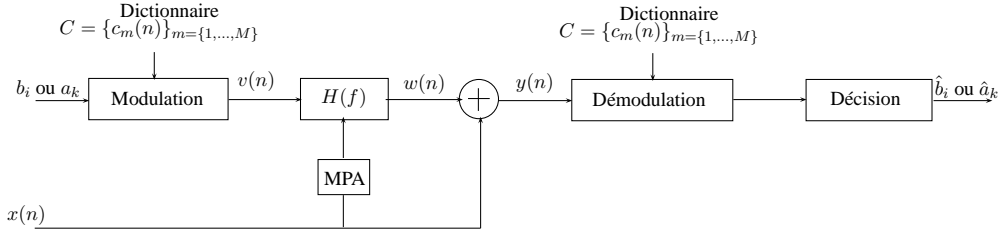


FIGURE 2.8 – Système de tatouage lorsque l'émetteur réalise la mise en forme perceptive du signal modulé en suivant un MPA.

tielle d'inaudibilité. Le signal  $w(n)$  ajouté à  $x(n)$  est transparent à condition que sa DSP soit inférieure au seuil de masquage quelle que soit la fréquence. Ce modèle psychoacoustique ne peut être appliqué que sur une fenêtre de durée  $N_{mpa}$  du signal audio, sur laquelle il peut être supposé stationnaire ; cette durée est de l'ordre de quelques millisecondes. L'analyse psychoacoustique doit donc être réévaluée à chaque fois que les propriétés statistiques du signal audio changent. Un résumé sur les études psychoacoustiques est présentée dans l'Annexe A.

## 2.2.2 Filtre de mise en forme

Assurer un compromis entre l'inaudibilité et la fiabilité de détection est le défi majeur dans la conception d'un tel système de tatouage audio. En effet, la puissance du signal de tatouage  $w(n)$  doit être choisie la plus élevée possible pour faciliter la détection correcte de l'information émise. D'autre part, cette puissance est limitée par le seuil de masquage puisque la DSP de  $w(n)$ , notée  $S_w(f)$ , ne doit pas excéder le seuil. La solution optimale à ce problème de maximisation sous contrainte consiste naturellement à choisir le signal  $w(n)$  de sorte que sa DSP coïncide avec le seuil de masquage  $M_x(f)$ . Ceci se traduit par la mise en forme spectrale du signal modulé  $v(n)$  par un filtre adaptatif, appelé filtre de mise en forme spectrale et noté  $H(f)$  (c.f. FIGURE 2.8). Ce filtre est conçu de sorte que la DSP du signal résultant du filtrage de  $v(n)$  par  $H(f)$ , appelé signal modulé mis en forme spectralement  $w(n)$ , coïncide avec le seuil de masquage. Le développement mathématique de ce filtre de mise en forme est présenté dans l'Annexe B.

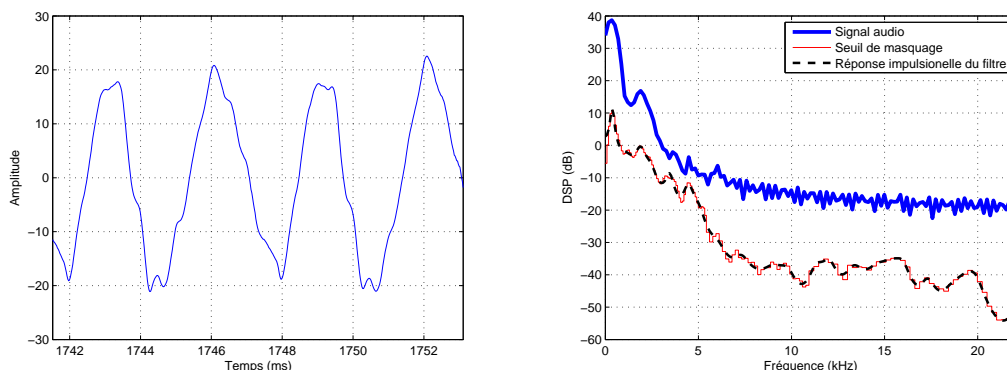
La propriété du masquage fréquentiel permet ainsi de déduire la DSP maximale que peut avoir un signal pour que l'on puisse l'insérer de façon imperceptible dans un signal audio hôte  $x(n)$ . Si  $M_x(f)$  est le seuil de masquage de  $x(n)$ , fourni par un modèle psychoacoustique (MPA) pour une fenêtre d'analyse de durée  $N_{mpa}T_e$  secondes, le signal  $y(n) = x(n) + w(n)$  sera perceptivement identique à  $x(n)$  si la DSP  $S_w(f)$  du signal  $w(n)$  vérifie :

$$S_w(f) \leq M_x(f) \quad (2.16)$$

Ainsi, la mise en forme perceptive, tout en étant incontournable au niveau de l'insertion du message, dans la mesure où elle lui procure le maximum de puissance imperceptible, est équivalente à un canal très particulier, dont les caractéristiques interviennent dans la conception du récepteur. La fonction de transfert  $H(z)$  du canal est purement récursive, stable et causale et s'écrit sous la forme suivante [74] :

$$H(z) = \frac{b_0}{1 + \sum_{i=1}^P a_i z^{-i}} \quad (2.17)$$

"La théorie, c'est quand on sait tout et que rien ne fonctionne. La pratique, c'est quand tout fonctionne et que personne ne sait pourquoi. Ici, nous avons réuni théorie et pratique; Rien ne fonctionne... et personne ne sait pourquoi!". **Albert Einstein**



(a) Trentième trame du signal audio dans le domaine temporel (b) DSP du signal audio  $x(n)$ , seuil de masquage et carré de la réponse fréquentielle de  $H(f)$ .

FIGURE 2.9 – Allure temporelle du signal audio et seuil de masquage.

Par ailleurs, le filtre  $H(z)$  est actualisé au rythme de calcul du seuil de masquage, ses coefficients varient tous les blocs de  $N_{mpa}$  échantillons, vu la non stationnarité du signal audio  $x(n)$ .  $H(z)$  est équivalent à un canal variant dans le temps, mais invariant par fenêtre d'analyse. La FIGURE 2.9(a) illustre un extrait du signal audio  $x(n)$  de  $N_{mpa} = 512$  échantillons où on peut remarquer que  $x(n)$  est localement stationnaire.

### 2.2.3 Contrainte d'inaudibilité

Si le filtre  $H(f)$  est correctement paramétré, choisir n'importe quel signal modulé de DSP constante et de puissance  $\sigma_v^2$  égale à 1 garantit l'inaudibilité du signal de tatouage  $w(n)$  en présence de  $x(n)$ . Cette condition peut être étendue au signal modulé de puissance inférieure à 1. En effet, dans ce cas, l'allure fréquentielle du signal de tatouage  $w(n)$  est toujours celle du seuil de masquage (c.f. FIGURE 2.9(b)). L'inaudibilité du tatouage est satisfaite si le signal modulé  $v(n)$  est choisi blanc et de puissance :

$$\sigma_v^2 \leq 1 \tag{2.18}$$

Cette condition sera par la suite appelée contrainte d'inaudibilité. Le signal modulé  $v(n)$  qui vérifie cette contrainte, est le résultat d'une modulation par étalement de spectre.

## 2.3 Récepteur aveugle : Égalisation et filtrage

Garantir la transmission fiable de l'information nécessite de prendre en compte la statistique du signal audio  $x(n)$ . Celle-ci est très différente de celle d'un canal BBAG étudié par la théorie. En effet, les signaux audio sont notamment connus pour être non-stationnaires et fortement colorés. Nous ferons l'hypothèse que le signal audio peut être modélisé par un processus Auto-Régressif (AR) d'ordre  $P$  globalement non stationnaire mais supposé localement stationnaire. Au regard de cette modélisation, le récepteur optimal défini pour un canal BBAG sera généralisé au cas d'un bruit AR. Ce récepteur devra également prendre en compte les choix spécifiques réalisés par l'émetteur pour construire le tatouage inaudible  $w(n)$ .

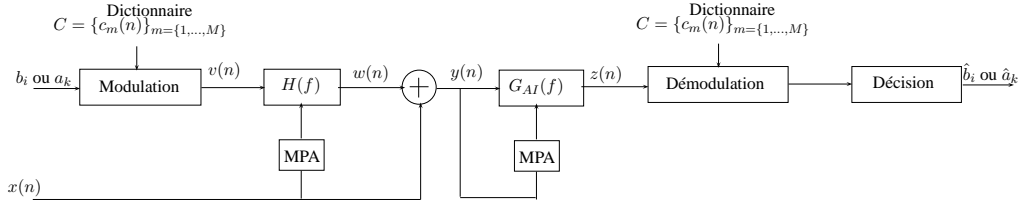
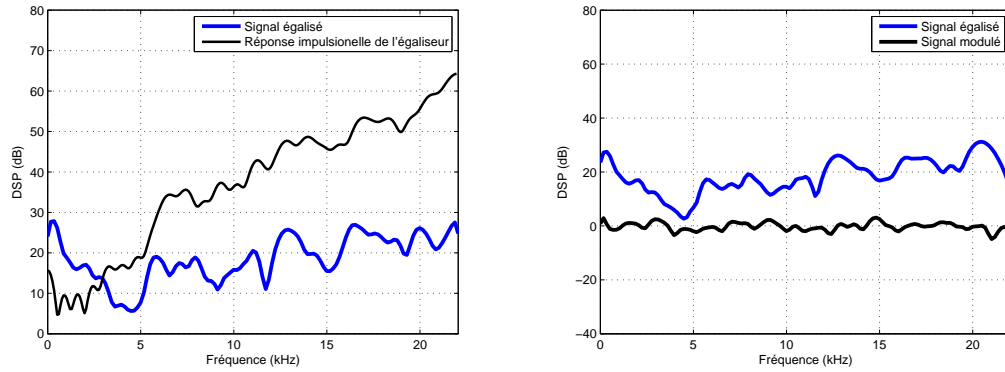


FIGURE 2.10 – Principe de la détection du tatouage par l'approximation  $1/\hat{H}(z)$  de l'anneur d'interférences et un démodulateur à corrélation.



(a) Réponse fréquentielle de l'anneur d'interférence et le signal égalisé  $z(n)$ .

(b) Signal égalisé et signal modulé.

FIGURE 2.11 – Représentation de la réponse fréquentielle de l'anneur d'interférences  $G_{AI}(f)$  ainsi que le signal modulé avant et après égalisation.

### 2.3.1 Anneur d'interférences

L'égalisation par l'anneur d'interférences est basée sur le critère du forçage à zéro [52]. Lorsque la réponse en fréquence du canal est connue au niveau du récepteur, cette technique consiste à annuler les interférences entre symboles en filtrant le signal reçu par un filtre de réponse fréquentielle égale à l'inverse de celle du canal. Ce principe est utilisé dans le contexte du tatouage, où la réponse fréquentielle du canal est approchée au récepteur en utilisant un modèle psychoacoustique [74].

Il est donc possible d'obtenir au niveau du récepteur une estimation  $\hat{H}(z)$  de  $H(z)$  : il suffit d'introduire un MPA au récepteur pour calculer le seuil de masquage  $M_y(f)$  et de procéder de la même manière que pour  $H(z)$  comme c'est illustré à la FIGURE 2.10.

Par la suite, on exprimera l'anneur d'interférences<sup>7</sup> d'ordre  $P$  par sa RI<sup>8</sup>  $g_{AI}(n)$  et sa fonction de transfert qu'on définit par :

$$G_{AI}(z) = \frac{1}{\hat{H}(z)} \quad (2.19)$$

7. Le développement mathématique de l'anneur d'interférence est présenté dans l'Annexe B.

8. Réponse impulsionnelle.

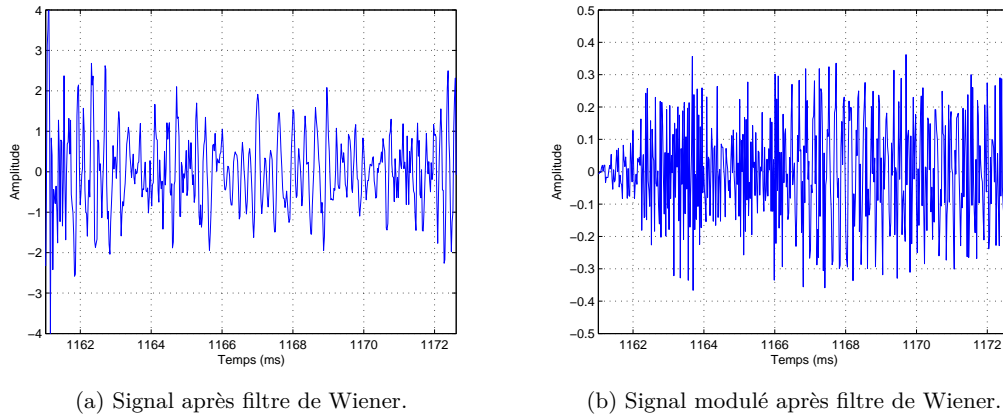


FIGURE 2.12 – Effet d'introduction du filtre de Wiener.

L'avantage de l'annuleur d'interférences utilisé dans ce contexte, est que non seulement il inverse la réponse du canal, mais décorrèle partiellement le bruit additif. Toujours sous l'hypothèse que  $\hat{H}(z)$  constitue une bonne approximation du canal  $H(z)$ , le signal à la sortie de l'annuleur d'interférences  $G_{AI}(z)$ , qu'on notera par la suite  $z(n)$ , s'écrit :

$$z(n) = (w(n) + x(n)) * g_{AI}(n) \approx v(n) + r(n) \quad (2.20)$$

où  $r(n)$  représente le signal audio égalisé. La DSP du signal égalisé  $z(n)$  est illustrée sur la FIGURE 2.11(a). Si on le compare avec le signal modulé  $v(n)$ , on remarque qu'ils sont différents comme le montre la FIGURE 2.11(b). Ceci montre que l'annuleur d'interférence  $G_{AI}(z)$  ne reconstruit pas le signal modulé avec précision. Afin d'améliorer les performances du détecteur à corrélation, il faudrait ajouter un autre filtre pour améliorer la détection du tatouage.

### 2.3.2 Filtre de Wiener

Le RSB à la sortie de l'égaliseur de forçage à zéro  $z(n)$  est très faible. En effet, la puissance du signal modulé  $v(n)$  est faible par rapport à celle du signal audio égalisé  $r(n)$ . Puisque les DSP de  $v(n)$  et  $r(n)$  sont différentes, il est possible de filtrer  $z(n)$  pour augmenter le RSB, en améliorant les composantes spectrales de  $z(n)$  dominées par  $v(n)$  et atténuer celles dominées par  $r(n)$ . Cela peut être fait de façon optimale en ajoutant le filtre de Wiener<sup>9</sup>  $G_W(z)$  à la sortie de l'égaliseur de forçage à zéro.

$$G_W(z) = \sum_{i=-p}^p g_W(i)z^{-i} \quad (2.21)$$

Le signal résultant  $\hat{v}(n)$  par le filtrage de Wiener  $G_W(z)$ , est très similaire au signal  $v(n)$  comme c'est illustré dans la FIGURE 2.12. Ses coefficients sont calculés de telle manière à minimiser la puissance du signal d'erreur  $e(n) = z(n) - v(n)$ . Cette condition est satisfaite si les coefficients  $g_W(i)$  du filtre sont la solution des équations de Wiener-Hopf [53] :

9. Le développement mathématique du filtre de Wiener est présenté dans l'Annexe B.

"Le succès ne consiste pas à ne jamais faire d'erreur mais à ne jamais faire la même erreur deux fois". **George Bernard Shaw**

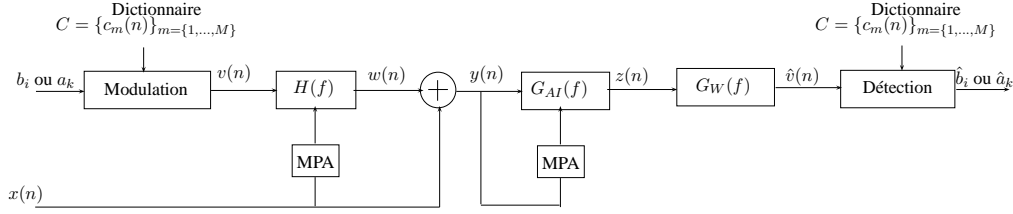


FIGURE 2.13 – Système de tatouage en présence d'égalisation mixte par l'annuleur d'IES  $G_{AI}(f)$  et par le filtre de Wiener  $G_W(f)$ , suivis par le détecteur à corrélation.

$$\begin{bmatrix} R_x(0) & R_x(1) & \cdots & R_x(2p) \\ R_x(1) & R_x(0) & \cdots & R_x(2p+1) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(-2p) & R_x(-2p+1) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} z_{-p} \\ z_{-p+1} \\ \vdots \\ z_p \end{bmatrix} = \begin{bmatrix} R_v(-p) \\ R_v(-p+1) \\ \vdots \\ R_v(p) \end{bmatrix} \quad (2.22)$$

où  $R_x(k)$  et  $R_v(k)$  sont les fonctions d'autocorrélation des signaux  $x(n)$  et  $v(n)$  respectivement.

### 2.3.3 Détection du tatouage

Dans la phase de détection, il faut effectuer une démodulation qui consiste en un processus d'étalement inverse. L'opération d'étalement inverse en réception est effectuée sur le signal reçu  $y(n)$  qui est la somme du signal audio  $x(n)$  et le tatouage  $w(n)$ . Comme l'énergie du signal modulé reste constante au cours de l'étalement, la densité de puissance du bruit audio devient plus faible que dans le cas sans étalement. Le processus de filtrage en réception élimine le bruit au delà de la bande étroite du signal. La puissance du bruit audio après le processus étalement-filtrage est donc divisée par  $N_c$ . Cela signifie une meilleure performance en terme de probabilité d'erreur. Le schéma final du système de tatouage à ce stade du développement est présenté sur la FIGURE 2.13.

## 2.4 Système de tatouage par étalement de spectre

En transmettant le message sur une bande passante beaucoup plus large que le minimum requis, les modulations par étalement de spectre offrent un avantage non négligeable pour la chaîne de tatouage. En effet, la redondance de l'information permet de protéger efficacement le système de transmission s'il est soumis à des interférences sévères [53]. De plus, l'aspect pseudo-aléatoire rend difficile la détection de l'information tatouée par un récepteur non autorisé [53]. Enfin, la DSP du signal modulé est faible du fait que la puissance transmise est étalée sur une large bande, et par conséquent, la quantité d'énergie par fréquence spécifique est très faible.

Dans cette partie, on s'intéresse à deux types de modulations par étalement de spectre : étalement de spectre par séquence directe (DSSS<sup>10</sup>) et étalement de spectre par saut de fréquence (FHSS<sup>11</sup>).

10. Direct Sequence Spread Spectrum.

11. Frequency Hopping Spread Spectrum.

"Un homme qui lit trop et qui fait trop peu d'efforts cérébraux prend vite des habitudes de paresse d'esprit".  
**Albert Einstein**

---

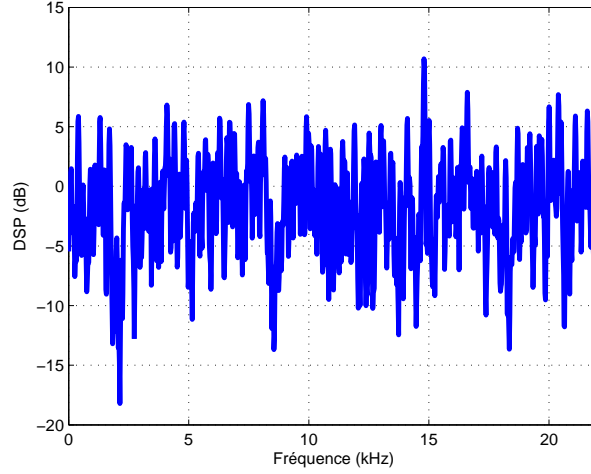


FIGURE 2.14 – La DSP du signal de tatouage par étalement de spectre à séquence directe.

### 2.4.1 Étalement de spectre à séquence directe

L'étalement de spectre à séquence directe a été initialement conçu pour des applications militaires [96]. En effet, l'étalement de spectre se définit comme un moyen de transmission dans lequel le signal occupe une bande passante supérieure au minimum nécessaire pour envoyer l'information. Le signal est réparti sur une bande de fréquence plus large, il est de faible puissance et moins sensible aux interférences [95]. L'étalement de la bande se fait par un code qui est indépendant des données. Ce code est sous forme d'une suite binaire et possède des propriétés similaires au bruit blanc. À la réception, le même code d'étalement est utilisé pour démoduler le signal reçu et donc de retrouver les données originales.

Le signal de tatouage étalé  $v(n)$  est obtenu en modulant la marque binaire  $b_m \in \{0, 1\}$  de taille  $L$  bits avec la forme d'onde étalée  $c(n)$  composée de  $N_b$  échantillons ( $c(0), c(1), \dots, c(N_b - 1)$ )

$$v(n) = \sum_{m=0}^{L-1} a_m c(n - mN_b) \quad (2.23)$$

où  $a_m$  est la séquence de symboles de  $\{-1, 1\}$  donnée par  $a_m = 2b_m - 1$ . Le signal de tatouage modulé  $v(n)$  est étalé sur toute la bande de fréquence disponible (cf. FIGURE. 2.14). Le signal tatoué est donné par :

$$\begin{aligned} y(n) &= v(n) + x(n) \\ &= c(n)a(n) + x(n) \end{aligned} \quad (2.24)$$

où  $a(n) = \sum_m a_m \psi_{N_b}(n - mN_b)$  et  $\psi_{N_b}$  est égale à 1 pour  $0 \leq n < N_b$  et zéro sinon. Ensuite, le signal tatoué  $y(n)$  passe par une étape de démodulation qui se compose d'un multiplicateur suivi d'un intégrateur et d'un dispositif de décision. Le multiplicateur est fourni avec un signal d'étalement  $c(n)$  produit localement qui est une copie exacte de celle utilisée dans l'émetteur. De plus, nous supposons que le récepteur fonctionne en synchronisation parfaite avec l'émetteur. La sortie du démodulateur est donc donnée par :

$$\begin{aligned} q(n) &= c(n)y(n) \\ &= c^2(n)a(n) + c(n)x(n) \end{aligned} \quad (2.25)$$

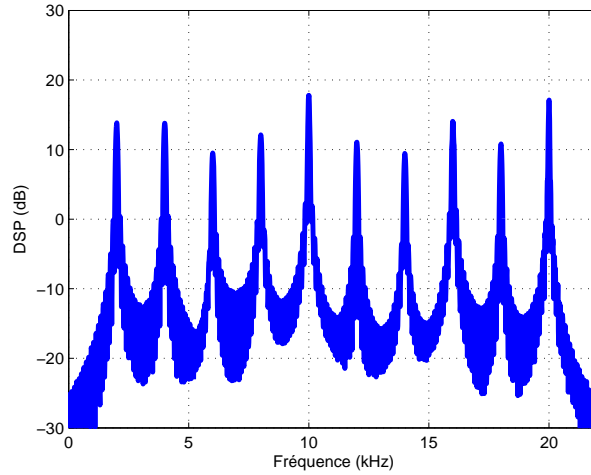


FIGURE 2.15 – La DSP du signal de tatouage par étalement de spectre à saut de fréquence.

La forme d'onde d'étalement  $c(n)$  oscille entre les niveaux  $+1$  et  $-1$ , cette alternance est perdue quand :

$$c^2(n) = 1 \quad \forall n \quad (2.26)$$

Par conséquent, on peut simplifier l'équation (2.25) en :

$$q(n) = a(n) + c(n)x(n) \quad (2.27)$$

D'après (2.27), on déduit que le signal  $a(n)$  est bien reproduit à la sortie du multiplicateur au niveau du récepteur, à l'exception de l'effet de l'interférence représenté par le terme additif  $c(n)x(n)$ . L'étalement de spectre affaiblit la puissance associée à l'expression  $c(n)x(n)$ . Maintenant, la composante de données  $a(n)$  est à bande étroite, alors que la composante  $c(n)x(n)$  est à bande étalée. Par conséquent, l'effet de l'interférence est réduit de manière significative au niveau du récepteur.

## 2.4.2 Étalement de spectre par saut de fréquence

Une autre technique d'étalement de spectre est la FHSS, qui ne vise pas cette fois à étaler le signal à bande étroite en un signal à large bande, mais consiste à diviser la largeur de bande disponible en des sous bandes (ou canaux) (cf. FIGURE. 2.15). Le signal modulé est ainsi réparti sur une large bande de fréquences selon une séquence de fréquences porteuses  $\{f_1, f_2, \dots, f_Q\}$  choisies de manière aléatoire<sup>12</sup>. Par ailleurs, il est envoyé avec un débit chip  $R_c$  défini par :

$$R_c = \max(R_h, R_s) \quad (2.28)$$

où  $R_h$  est le débit saut et  $R_s$  est le débit symbole.

Chaque canal de fréquence est défini comme étant une région spectrale qui inclut une seule fréquence porteuse. L'intervalle de temps entre les sauts de fréquence est appelé saut d'intervalle.

---

12. Cette séquence est appelé Hopping Code.

Techniques	DSSS	FHSS
<b>Caractéristiques</b>		
<b>Robustesse contre les IES</b>	Fort	Fort
<b>DSP</b>	Faible	élevé
<b>Débit</b>	Élevé	moyen
<b>Efficacité spectrale</b>	Grande	Petite
<b>Saut de fréquence</b>	Non	Oui
<b>Synchronisation</b>	Facile	Difficile
<b>Gain d'étalement</b>	Élevé	Aucun
<b>RSB après démodulation</b>	Élevé	Faible
<b>Coût</b>	Facile	Difficile
<b>Consommation de l'énergie</b>	Élevé	Faible

Tableau 2.2 – Comparaison entre les modulations par étalement de spectre DSSS et FHSS.

Sa durée est appelée le temps de saut et est noté  $T_h$ . La nouvelle bande du signal FHSS est  $W \geq QB$  avec  $B$  est la bande définie pour un seul saut de fréquence.

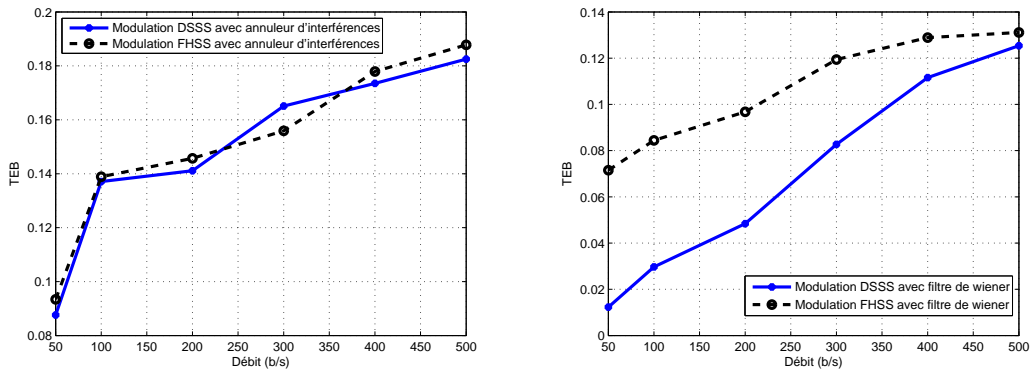
À l'émetteur, le signal modulé est construit à partir d'un synthétiseur de fréquences qui génère des fréquences aléatoires. Si le signal modulé a une forme de modulation de phase  $\phi(n)$ , alors le signal reçu du  $i^{\text{ème}}$  saut est :

$$v(n) = S \cos(2\pi f_j n + \phi(n) + \phi_i) \quad (i-1)N_h \leq n \leq iN_h \quad (2.29)$$

où  $S$  est la puissance moyenne,  $f_j$  est la fréquence porteuse du saut, et  $\phi_i$  est la phase aléatoire du  $i^{\text{ème}}$  saut. La séquence de saut de fréquence produite par le synthétiseur de récepteur est synchronisée avec la séquence produite par l'émetteur, mais elle est décalée d'une fréquence intermédiaire fixe, qui peut être nulle. Enfin, l'opération de détection consiste à supprimer la séquence de saut de fréquence à partir du signal reçu.

### 2.4.3 Comparaison entre les deux techniques

Les modulations DSSS et FHSS sont des techniques de modulations par étalement de spectre dont chacune a ses propres avantages et ses propres inconvénients [33] (c.f. Tableau. 2.2). On a évalué les modulations numériques DSSS et FHSS dans le système de tatouage audio. Tout d'abord, on a commencé par utiliser l'annuleur d'interférence seul au récepteur. La FIGURE 2.16(a) montre que la fiabilité de détection est faible et les deux modulations présentent des résultats équivalents. Par contre, lorsqu'on ajoute le filtre de Wiener, les performances du système se sont améliorées et la modulation DSSS présente une amélioration remarquable par rapport à la modulation FHSS avec un  $TEB = 10^{-2}$  pour un débit de 300b/s (c.f. FIGURE. 2.16(b)). En effet, la modulation FHSS produit plus d'interférences que le système DSSS et elle est très peu utilisée aujourd'hui en télécommunications numériques.



(a) Comparaison entre les deux modulations DSSS et FHSS en utilisant l'annuleur d'interférences . (b) Comparaison entre les deux modulations DSSS et FHSS en utilisant le filtre de Wiener .

FIGURE 2.16 – Comparaison entre les deux modulations DSSS et FHSS dans un système de tatouage audio.

## Conclusion

Dans ce chapitre, nous avons rappelé les outils nécessaires de la théorie des communications numériques qui nous ont permis d'établir le schéma du système de tatouage présenté dans la FIGURE 2.13. Ce système adapte les outils de la théorie de communication numérique au contexte du tatouage audio tout en respectant les spécificités et les contraintes de ce dernier. L'information à tatouer a été modulée en exploitant deux types de modulations par étalement de spectre à savoir : l'étalement de spectre par séquence directe et l'étalement de spectre par saut de fréquence. Cette phase de modulation a été suivie par une mise en forme qui permet d'ajouter le signal de tatouage  $w(n)$  au signal audio  $x(n)$  sous la contrainte d'inaudibilité. La réception est basée sur un annuleur d'interférence suivie d'un filtre de Wiener qui permet de récupérer au mieux l'information insérée avec le minimum d'erreurs possibles.

Malgré les similitudes qui existent entre les chaînes de tatouage et ceux de communication numérique, le système de tatouage présente aussi des caractéristiques différentes du modèle standard d'un système de communication numérique. En effet, la contrainte d'inaudibilité impose une puissance de tatouage faible par rapport au signal audio original. De plus, les caractéristiques du bruit audio sont différentes de celles du bruit en communication numérique qui est considéré blanc et gaussien. Ces particularités sont à l'origine des difficultés de la conception d'un système de tatouage performant.

L'objectif de cette thèse est d'étudier l'apport des techniques de modulations numériques et de les adapter au contexte du tatouage audio.

## Chapitre 3

# Tatouage audio par modulations à porteuses sinusoïdales

Les modulations par étalement de spectre sont les plus utilisées dans les systèmes de tatouage [33]; afin d'exploiter toute la bande du signal audio et insérer le tatouage avec une faible puissance. Ces propriétés garantissent d'une part la contrainte d'inaudibilité et assurent d'autre part une robustesse vis-à-vis des interférences entre symboles.

Outre les modulations par étalement de spectre, on se propose d'exploiter les modulations à porteuses sinusoïdales en tatouage audio. Ce chapitre s'intéresse à deux types de modulations; les modulations à bande étroite et les modulations multi-porteuses. Tout d'abord, on met l'accent sur quelques avantages que peuvent présenter les modulations à bande étroite. Ensuite, on montre qu'un choix judicieux des paramètres du signal de tatouage à bande étroite permet d'assurer un bon compromis entre l'inaudibilité et la fiabilité de détection. On présente également un système de tatouage informé basé sur les modulations à bande étroite qui exploite à priori la connaissance du signal audio pour choisir les formes d'ondes adéquates à l'insertion et à l'extraction du tatouage. Enfin, on intègre les modulations multi-porteuses permettant l'exploitation de toute la bande du signal audio pour effectuer l'opération d'insertion et d'extraction du tatouage.

### 3.1 Modulation

#### 3.1.1 Influence du type de modulation sur les performances du tatouage audio

En communication numérique [53], la modulation joue un rôle très important dans la conception d'un système de communication numérique. En effet, elle est l'élément responsable de la transmission du maximum de données binaires avec un taux d'erreur acceptable. Ainsi, le choix du type de modulation en communication numérique, se fait selon les critères suivants [53] :

- Débit de transmission (bits/seconde).
- Performance en terme d'erreur (Probabilité d'erreur).
- Puissance du signal émis (Rapport signal sur Bruit) pour atteindre une certaine performance.

De même, en tatouage audio, le choix du type de modulation se fait selon les trois critères suivants :

- Débit d'insertion (ou capacité d'information insérée).
- Fiabilité de détection (Taux d'erreur).
- Puissance du tatouage qui assure la contrainte d'inaudibilité.

Le problème semble donc très similaire à celui de la transmission des données numériques via un canal de communication. L'objectif est de pouvoir maximiser le débit d'insertion tout en minimisant le taux d'erreur binaire, même en présence du bruit perturbateur ou encore des distorsions rencontrées. En considérant le système de tatouage comme une chaîne de communication particulière, le tatouage peut bénéficier directement de 60 ans de la théorie de la communication numérique, après le travail pionnier de Claude Shannon [101]. Le problème peut être abordé comme la conception d'une modulation appropriée pour obtenir les meilleures performances du système de tatouage.

Le type de modulation le plus basique est celui à bande de base. Ce type de modulations est souvent utilisé en communication numérique, dans les transmissions à courte distance [44]. De plus, il a été déjà utilisé dans les systèmes de tatouage audio avec les codeurs NRZ<sup>1</sup> et RZ<sup>2</sup> (binaire et polaire) et ne semble pas être applicable dans un objectif d'inaudibilité et de fiabilité de détection [19]. En effet, si l'on choisit une modulation en bande de base, la puissance du signal modulé  $s(n)$  se situe dans les basses fréquences, typiquement autour de zéro. Si le signal audio est très puissant autour de cette fréquence, le tatouage ne sera pas transparent, et la contrainte d'inaudibilité ne sera pas satisfaite.

Néanmoins, on peut utiliser d'autres techniques de modulations numériques classiques ; connues sous le nom des modulations à bande étroite<sup>3</sup> afin d'insérer une information dans les hautes fréquences du signal audio. Ce type de modulations permet de translater le spectre du signal  $s(n)$  tout en préservant l'information supportée par le signal en bande de base. Le signal modulé en bande étroite s'écrit alors [53] :

$$\begin{aligned} v(n) &= \operatorname{Re}[s(n)\exp(j2\pi f_c n)] \\ &= A_c \cos(j2\pi f_c n + \varphi_c) \end{aligned} \quad (3.1)$$

où  $f_c$  est la fréquence porteuse,  $A_c$  est l'amplitude,  $\varphi_c$  représente la phase,  $\operatorname{Re}$  est la partie réelle et  $s(n) = A_c \exp(j\varphi_c)$ . L'information incorporée peut être affectée à :

- $A_c$  si on choisit une modulation par déplacement d'amplitude (ASK<sup>4</sup>) ;
- $\varphi_c$  si on choisit une modulation par déplacement de phase (PSK<sup>5</sup>) ;
- $f_c$  si on choisit une modulation par déplacement de fréquence (FSK<sup>6</sup>)

---

1. Non-Return-to-Zero.  
2. Return-to-Zero.  
3. Aussi appelé modulations sur ondes porteuse.  
4. Amplitude Shift Keying.  
5. Phase Shift Keying.  
6. Frequency Shift Keying.

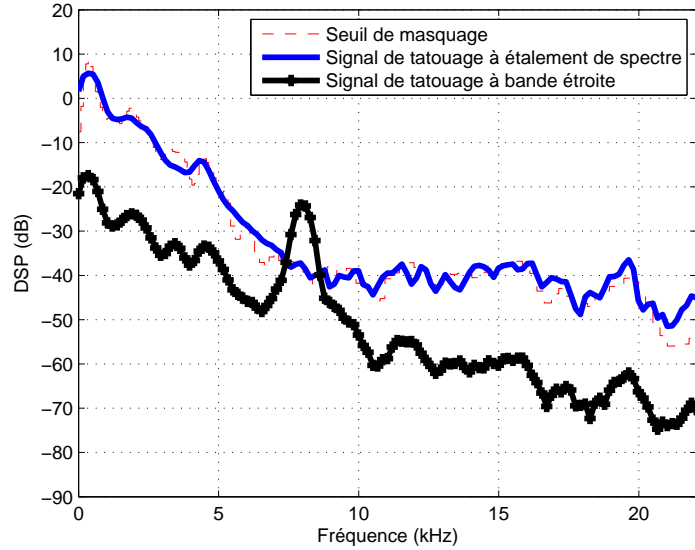


FIGURE 3.1 – La DSP du signal de tatouage en utilisant la modulation à bande étroite et celle à étalement de spectre ainsi que le seuil de masquage.

ou toute combinaison hybride entre ces paramètres. La DSP  $S_v(f)$  du signal modulé  $v(n)$  contient la somme de deux versions translatées en  $f - f_c$  et  $f + f_c$  de la DSP  $S_s(f)$  du signal modulé en bande de base [53] :

$$S_v(f) = \frac{1}{4}[S_s(f - f_c) + S_s(f + f_c)] \quad (3.2)$$

Si l'on choisit une modulation à bande étroite, la densité spectrale de puissance du signal modulé  $S_v(f)$  sera concentrée autour de  $f_c$ , ce qui pose deux problèmes : le choix de la fréquence porteuse  $f_c$  et la largeur de bande utilisée  $W_M$  par le spectre du signal modulé, qui est très faible devant la largeur de bande disponible  $W_A$  ; qui est celle du signal audio. Les chercheurs ont donc préféré de choisir une modulation par étalement de spectre pour quatre raisons principales :

- Insertion du signal de tatouage avec une puissance faible et donc une inaudibilité assurée.
- Robustesse contre les interférences entre symboles et donc une bonne fiabilité de détection.
- Sécurité et confidentialité du tatouage.
- Exploitation de toute la largeur de bande du signal audio.

En tatouage audio, les modulations à bande étroite ont acquis moins d'attention que les modulations à étalement de spectre. Ceci est dû à la DSP  $S_v(f)$  du signal à bande étroite  $v(n)$  concentrée autour de la fréquence porteuse  $f_c$  qui peut impacter la qualité audio du signal tatoué. La FIGURE 3.1 présente la DSP du signal de tatouage  $w(n)$  utilisant la modulation à bande étroite et celle à étalement de spectre ainsi que le seuil de masquage  $M_x(f)$ . Elle montre que la puissance du signal de tatouage par étalement de spectre est répartie sur toute la bande disponible du signal audio et suit la même allure du seuil de masquage. En revanche, la DSP du signal de tatouage

à bande étroite est concentrée autour de la fréquence porteuse  $f_c = 8775\text{Hz}$  alors qu'elle est faible ailleurs. Ainsi, il serait nécessaire de bien choisir la fréquence porteuse  $f_c$  qui permet de garantir d'une part l'inaudibilité et d'améliorer d'autre part la fiabilité de détection. Mais tout d'abord, dans quel paramètre du signal modulé  $v(n)$  doit-on insérer le tatouage pour assurer de telles performances ? Pour répondre à cette question, il faudrait évaluer les performances des modulations à bande étroite dans le système de tatouage audio.

### 3.1.2 Modulations à bande étroite

Dans cette partie, on envisage de traiter trois types de modulations à bande étroite : la modulation par déplacement d'amplitude, la modulation par déplacement de phase et la modulation par déplacement de fréquence.

#### 3.1.2.1 Modulation par déplacement d'amplitude

La modulation par déplacement d'amplitude (ASK) est probablement le premier type de modulation numérique appliqué en pratique. Elle a été utilisée grâce à sa simplicité pour la transmission de code Morse dans le télégraphe radio [89]. Un exemple simple de la modulation d'amplitude est la modulation binaire BASK<sup>7</sup> et plus précisément OOK<sup>8</sup>. Dans ce cas, un seul bit est transmis par période  $T_b$  (ou  $N_b$ ). La suite de bits '0' et '1' est représentée par les deux signaux :

$$\begin{cases} c_1(n) = 0 & b_i = 0 \\ c_2(n) = \sqrt{\frac{2E_b F_e}{N_b}} \cos(2\pi f_c n) & b_i = 1 \end{cases} \quad (3.3)$$

où  $E_b$  est l'énergie transmise par un seul bit et  $0 \leq n \leq N_b - 1$ .

Le fait d'augmenter le nombre de bits transmis par symbole permet d'augmenter l'efficacité spectrale<sup>9</sup> et le signal modulé va s'écrire dans ce cas sous la forme suivante :

$$c_i(n) = A_i \sqrt{\frac{2F_e}{N_s}} \cos(2\pi f_c n) \quad (3.4)$$

où  $i = 1, 2, \dots, M$ . Malheureusement, la probabilité d'erreur par symbole augmente aussi quand on augmente le nombre de bits par symbole. Ainsi, pour améliorer la fiabilité de transmission, il faudrait augmenter l'énergie émise par bit  $E_b$ . Par conséquent, ce type de modulation ne garantit pas un système fiable en transmission à cause de sa faible efficacité spectrale.

#### 3.1.2.2 Modulations par déplacement de phase

La modulation par déplacement de phase (PSK) désigne une famille de formes de modulations numériques qui ont toutes pour principe de véhiculer de l'information binaire via la phase d'un signal de référence (porteuse). Comme pour toute technique de modulation numérique, la phase

---

7. Binary Amplitude Shift Keying.

8. Tout ou rien.

9. L'efficacité spectrale est définie comme le flux binaire par Hz :

$$\rho = \frac{R_b}{B} \quad \text{bits/s/Hz}$$

en question ne peut prendre qu'un nombre fini de valeurs. Chacune de ces valeurs représente un nombre binaire unique, dont la taille (et donc la quantité d'information transmise) dépend du nombre de valeurs possibles pour la phase. Généralement, pour une modulation PSK donnée, les nombres binaires représentés sont tous de même taille. Ce type de modulation est utilisé dans les transmissions par satellite, la radiocommunication avec les mobiles et les faisceaux hertziens [89].

Parmi les formes de PSK les plus utilisées, il y a la modulation BPSK<sup>10</sup> dont les signaux  $c_1(n)$  et  $c_2(n)$  que vont représenter les symboles binaires '0' et '1', sont définis respectivement comme suit :

$$\begin{cases} c_1(n) = \sqrt{\frac{2E_b F_e}{N_b}} \cos(2\pi f_c n) & b_i = 0 \\ c_2(n) = -\sqrt{\frac{2E_b F_e}{N_b}} \cos(2\pi f_c n) & b_i = 1 \end{cases} \quad (3.5)$$

où  $f_c = \frac{n_c F_e}{N_b}$  est la fréquence de la porteuse,  $n_c$  est une constante,  $0 \leq n \leq N_b - 1$  et  $E_b$  est l'énergie du signal transmis par bit. Ces deux signaux sont antipodaux et possèdent la même fréquence et la même énergie. Ils ont été choisis du fait qu'ils ont un coefficient de corrélation égale à -1, ce qui donne une probabilité d'erreur minimale.

Les signaux  $\{c_i(n)\}$  pour  $i \in \{1, 2\}$  peuvent en général s'écrire comme combinaisons linéaires de fonctions appartenant à une base orthonormée obtenue par la méthode d'orthogonalisation de Gram Schmidt appliquée aux signaux  $\{c_i(n)\}$ . Les deux fonctions définissant cette base orthonormée s'écrivent [53] :

$$\begin{cases} \zeta 1(n) = \sqrt{\frac{2F_e}{N_b}} \cos(2\pi f_c n) & b_i = 0 \\ \zeta 2(n) = -\sqrt{\frac{2F_e}{N_b}} \cos(2\pi f_c n) & b_i = 1 \end{cases} \quad (3.6)$$

La propriété d'orthogonalité des vecteurs du dictionnaire permet de réduire la probabilité d'erreur lors de la détection par mesure de corrélation. Dans le cas général de la modulation PSK, le signal modulé s'écrit :

$$c_i(n) = \sqrt{\frac{2F_e}{N_s}} \cos(2\pi f_c n + \frac{2\pi i}{M}), \quad i = 0, 1, \dots, M - 1 \quad (3.7)$$

La motivation derrière l'utilisation de la modulation par déplacement de phase M-aire est d'augmenter l'efficacité spectrale. Mais le problème qui résulte lorsqu'on augmente le nombre de bits par symbole est que la puissance augmente aussi ; ce qui peut causer une perte en terme d'inaudibilité.

L'expression de la modulation PSK montre qu'il s'agit d'une modulation à enveloppe constante. Cette propriété lui garantit des performances acceptables, en termes d'inaudibilité et de fiabilité de transmission [53]. Il existe aussi d'autres variantes de la modulation PSK à l'exemple de la DPSK<sup>11</sup> et la OPSK<sup>12</sup> [44] qui n'ont pas été considérées dans ce travail ; mais qui peuvent améliorer les performances du système de tatouage.

---

10. Binary Phase Shift Keying.

11. DPSK (Differential-PSK.) : dans ce type de modulation, l'information est contenue non pas dans une valeur absolue de phase, mais dans le déphasage entre deux signaux successifs.

12. Les modulations décalées OPSK (Offset PSK) permettent de résoudre le problème des sauts de phases (rencontrés dans les modulations PSK.) qui font apparaître des discontinuités d'amplitude.

### 3.1.2.3 Modulation par déplacement de fréquence

La modulation par déplacement de fréquence (FSK) est un type de modulation par lequel le signal modulé varie entre des fréquences prédéterminées. Elle est probablement la première modulation numérique utilisée dans la communication pour l'industrie [44].

Dans le cas de la modulation par déplacement de fréquence binaire (BFSK<sup>13</sup>), on utilise deux fréquences différentes  $f_1$  et  $f_2 = f_1 + \Delta f$  afin de transmettre une séquence d'information binaire. Le dictionnaire de modulation est construit par les deux formes d'ondes :

$$\begin{cases} c_1(n) = \sqrt{\frac{2E_b F_e}{N_b}} \cos(2\pi f_1 n + \phi_1) & b_i = 0 \\ c_2(n) = \sqrt{\frac{2E_b F_e}{N_b}} \cos(2\pi f_2 n + \phi_2) & b_i = 1 \end{cases} \quad (3.8)$$

où  $\phi_1$  et  $\phi_2$  sont les phases initiales.

Lorsque  $\phi_1$  est différente de  $\phi_2$ , la forme d'onde est discontinue à chaque transition de bits. Ce premier type de la modulation FSK est discontinu et non cohérent. De plus, il est simple à réaliser mais son principal défaut est la grande bande passante dont il a besoin pour pouvoir transmettre les sauts de phase. Dans notre étude, on s'est intéressé à la modulation FSK cohérente caractérisée par des formes d'ondes de même phase initiale ( $\phi_1 = \phi_2$ ). Ce choix est basé sur les résultats découlant des télécommunications numériques qui ont montré que la modulation par déplacement de fréquence cohérente est plus performante que la modulation par déplacement de fréquence non cohérente [44]. En effet, ce type modulation est plus complexe à réaliser mais il requiert une bande passante plus étroite.

Dans le cas général, si on transmet un bloc de  $q = \log_2(M)$  bits par symbole. Les signaux du dictionnaire s'écrivent dans ce cas :

$$c_i(n) = \sqrt{\frac{2E_b F_e}{N_s}} \cos(2\pi f_c n + 2\pi i \Delta f n), \quad i = 0, 1, \dots, M-1, \quad 0 \leq n \leq N_s \quad (3.9)$$

$\Delta f = f_i - f_{i-1}$  est la différence de la fréquence instantanée et  $f_i = f_c + \Delta f$ .

Contrairement aux modulations ASK et PSK, les performances de la modulation FSK sont améliorées lorsqu'on augmente  $M$ . Cependant, l'augmentation de  $M$  entraîne aussi l'augmentation de l'occupation spectrale.

### 3.1.3 Comparaison entre les modulations à bande étroite

L'amélioration des performances des modulations à bande étroite revient à bien choisir la répartition des points sur le diagramme de constellation. En effet, il faut tenter d'éloigner les symboles au maximum pour diminuer la probabilité d'erreur et de les rapprocher de l'origine pour minimiser l'énergie nécessaire. L'augmentation ou la diminution du nombre de bits par symboles a fortement une influence sur les performances du système selon le type de modulation choisi.

La modulation BPSK est plus performante que les modulations BASK et BFSK en terme de fiabilité de transmission. D'un autre côté, la modulation BFSK occupe une bande plus large que les modulations BASK et BPSK [44].

---

13. Binary Frequency Shift Keying.

## 3.2 Tatouage par modulation de phase

Dans cette partie, on analyse la modulation par déplacement de phase afin d'améliorer les performances du système de tatouage audio. En effet, on vise à ajuster les différents paramètres de cette modulation pour les adapter au nouveau système de tatouage. Les paramètres de modulation qui seront traités sont :

- La phase dans laquelle on va insérer l'information.
- L'amplitude qui va assurer la contrainte d'inaudibilité.
- La fréquence porteuse adéquate pour l'insertion et l'extraction du tatouage.

### 3.2.1 Modulation de phase binaire

En utilisant les modulations par étalement de spectre, le signal de tatouage et le signal hôte sont fortement corrélés. Par conséquent, quelques chercheurs se sont intéressés à réduire cette corrélation pour améliorer la détection du tatouage [71]. Des études en tatouage ont déjà montré que la distribution de phase de manière uniforme diminue la corrélation entre le signal hôte et le signal de tatouage [70]. Ceci a été expérimenté sur une variété de signaux audio. De plus, le système auditif humain est généralement considéré comme insensible à la phase du son [85]. Par exemple, il est difficile pour les êtres humains de différencier entre deux sons complexes par la différence de phase dans les composantes hautes fréquences. Cependant, dans la gamme des basses fréquences, les différences de phase entre les composantes de fréquence influent certainement sur le timbre des sons complexes [91]. Même si une telle différence de phase relative est perceptible, il est connu que la phase absolue est imperceptible [106]. Ainsi, notre choix se porte sur la modulation PSK pour deux raisons principales :

- Le Système Auditif Humain (SAH) est insensible à la phase absolue du son, mais il est sensible à la phase relative [84]. Ainsi, nous transmettons l'information dans la phase absolue du signal de tatouage.
- La modulation PSK permet une séparation de phase maximale entre les points de constellations adjacents et donc une grande robustesse vis-a-vis des interférences entre symboles [53].

Les formes d'ondes du signal modulé PSK s'écrivent sous la forme :

$$c_i(n) = A \cos(2\pi f_c n + \varphi_i), \quad i = 0 \dots M - 1 \quad (3.10)$$

où  $A$  et  $f_c$  sont deux constantes et la phase  $\varphi_i$  est choisie parmi les  $M$  valeurs discrètes, données généralement par :

$$\varphi_i = \frac{2\pi i}{M}, \quad i = 0 \dots M - 1 \quad (3.11)$$

La question qui se pose maintenant est : pourquoi ne peut-on pas augmenter indéfiniment le débit en augmentant la taille de la constellation ? En effet, on peut effectivement augmenter le débit (en tant que vitesse de transmission) en augmentant la taille de constellation. Par contre, si l'on parle de débit en tant que nombre de bits par seconde arrivant sans erreur au niveau du

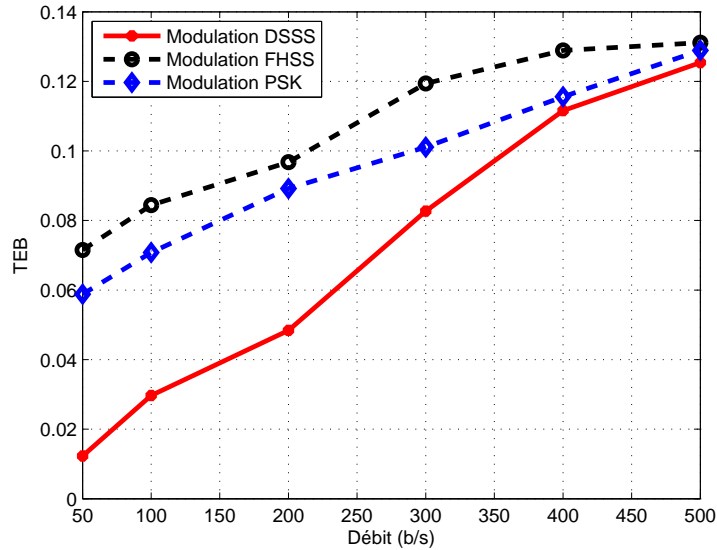


FIGURE 3.2 – Performances du système de tatouage audio en terme de fiabilité de détection : Comparaisons entre les modulations PSK, DSSS et FHSS.

récepteur, alors ceci n'est pas possible. En effet, plus la taille de la constellation augmente<sup>14</sup>, plus les symboles émis ont des valeurs de plus en plus proches. Il est alors difficile au niveau du récepteur de distinguer entre deux valeurs entachées d'erreurs dues au bruit. Les performances du système en seront alors affectées. Le débit a donc une limite déterminée par la capacité du canal.

D'autre part, plus la valeur de  $M$  est grande, plus la complexité de calcul augmente. En effet, si on transmet  $q$  bits par symboles, on aura besoin de  $2^q$  corrélateurs. Ainsi, on choisit  $M = 2$ . Ce choix découle de la performance élevée que représente la modulation BPSK, en terme de fiabilité de détection, par rapport aux modulations BASK et BFSK [53]. Par conséquent, le signal modulé  $v(n)$  sera composé de deux ondes porteuses différentes pour les deux phases ( $\varphi_0 = 0$  et  $\varphi_1 = \pi$ ) :

$$\begin{cases} c_0(n) = A \cos(2\pi f_c n) & \text{if } b_i = 0 \\ c_1(n) = -A \cos(2\pi f_c n) & \text{if } b_i = 1 \end{cases} \quad (3.12)$$

En utilisant la modulation PSK, on évalue maintenant les performances de cette technique en terme de fiabilité de détection en la comparant avec les modulations par étalement de spectre DSSS et FHSS. La FIGURE 3.2 montre le gain apporté par la technique DSSS par rapport à la modulation PSK. Cependant, la modulation PSK présente une nette amélioration par rapport à la technique FHSS. Mais est-ce que le tatouage PSK permet d'assurer la contrainte d'inaudibilité ?

### 3.2.2 Contrôle d'inaudibilité du tatouage PSK

La seule contrainte imposée au signal de tatouage  $w(n)$  et au signal modulé  $v(n)$  dont il est issu est la contrainte d'inaudibilité. D'où, il est nécessaire de faire appel à la psychoacoustique

14. à puissance fixée, la puissance est toujours normalisée pour des questions de coût de transmission.

pour établir les conditions d'une insertion inaudible et choisir l'allure des formes d'onde du dictionnaire à partir desquelles le signal  $v(n)$  est construit. En effet, le signal modulé  $v(n)$  ne peut pas être inséré directement dans le signal audio  $x(n)$  puisqu'il doit être inséré de manière transparente tout en rendant facile son extraction.

A cet effet, un filtre de mise en forme  $H(f)$  est conçu en bénéficiant des propriétés perceptives du système auditif humain afin d'augmenter le rapport signal sur bruit (RSB) du tatouage sous la contrainte d'inaudibilité. Ceci peut être réalisé par le filtrage du signal modulé  $v(n)$  par un filtre de mise en forme  $H(f)$  de sorte que le signal de tatouage résultant  $w(n)$  aura une DSP  $S_w(f)$  aussi proche que possible du seuil de masquage  $M_x(f)$ ,

$$\begin{aligned} M_x(f) &= S_w(f) \\ &= S_v(f) |H(f)|^2 \end{aligned} \quad (3.13)$$

Ainsi, la contrainte d'inaudibilité est assurée par la conception du filtre  $H(f)$ . Cette conception impose seulement de choisir un signal modulé  $v(n)$  qui vérifie l'inégalité suivante [9] :

$$\sigma_v^2 \leq 1 \quad (3.14)$$

Dans le cas où  $\sigma_v^2 = 1$ , on maximise la robustesse du tatouage sous la contrainte d'inaudibilité et le filtre de mise en forme  $H(f)$  devient :

$$|H(f)|^2 = M_x(f) \quad (3.15)$$

Si on utilise une modulation par étalement de spectre, le signal modulé  $v(n)$  est blanc et de puissance unité. Toutefois, en utilisant la modulation PSK, les points de la constellation choisis sont habituellement placés avec un espacement angulaire uniforme autour d'un cercle [53]. Par conséquent, les signaux qui constituent le signal modulé  $v(n)$  sont sinusoidaux avec une phase aléatoire [53, 70] :

$$c(n, \varphi) = A \cos(2\pi f_c n + \varphi) \quad (3.16)$$

où la phase  $\varphi$  est une variable aléatoire qui est uniformément distribuée sur l'intervalle  $[-\pi, \pi]$  avec une densité de probabilité [53] :

$$p(\varphi) = \frac{1}{2\pi} \quad \text{avec } \varphi \in [-\pi, \pi] \quad (3.17)$$

Afin d'utiliser le filtre  $H(f)$ , la puissance du signal modulé PSK doit satisfaire (3.14). D'après (3.16) et (3.17), on calcule les moments statistiques de la variable aléatoire  $c(n_k, \varphi)$  à un instant donné  $n_k$ .

$$\begin{aligned} E[c(n_k, \varphi)] &= E[c(n_k)] \\ &= \int_{-\infty}^{\infty} p(\varphi) A \cos(2\pi f_c n_k + \varphi) d\varphi \\ &= \frac{1}{2\pi} A \int_{-\pi}^{\pi} \cos(2\pi f_c n_k + \varphi) d\varphi \\ &= 0 \end{aligned} \quad (3.18)$$

et

$$\begin{aligned} E[c^2(n_k, \varphi)] &= E[c^2(n_k)] \\ &= \int_{-\pi}^{\pi} p(\varphi) A^2 \cos^2(2\pi f_c n_k + \varphi) d\varphi \\ &= \frac{A^2}{2} \end{aligned} \quad (3.19)$$

Par conséquent, le signal modulé  $v(n)$  est à moyenne nulle et de variance

$$\sigma_v^2 = \frac{A^2}{2} \quad (3.20)$$

Techniques	ODG
Modulation par étalement de spectre à séquence direct (DSSS)	-0.6729
Modulation par étalement de spectre à saut de fréquences (FHSS)	-0.4449
Modulation par déplacement de phase (PSK)	-0.2703

Tableau 3.1 – Valeurs de l’ODG du signal tatoué  $y(n)$  en utilisant les modulations DSSS, FHSS et PSK.

Pour garantir la contrainte d’inaudibilité (3.14), on doit choisir

$$A \leq \sqrt{2} \tag{3.21}$$

En choisissant  $A = \sqrt{2}$ , les ondes porteuses appartenant au dictionnaire  $C$  seront représentées par :

$$\begin{cases} c_0(n) = \sqrt{2} \cos(2\pi f_c n) \\ c_1(n) = -\sqrt{2} \cos(2\pi f_c n) \end{cases} \tag{3.22}$$

Ainsi, en utilisant les formes d’ondes présentées par (3.22), on garantit la transparence du tatouage. Le tableau 3.1 compare le degré d’inaudibilité des différentes modulations numériques. Tout d’abord, on remarque que les trois modulations permettent l’insertion du tatouage de manière transparente ( $ODG < -1$ ) mais assurent des degrés d’inaudibilité différentes en utilisant le même filtre de mise en forme  $H(f)$ . En effet, le filtre  $H(f)$  permet d’adapter la puissance du signal de tatouage  $w(n)$  selon le seuil de masquage qui est une version translatée vers le bas du spectre du signal audio. En effet, bien que le seuil de masquage soit théoriquement considéré comme la limite entre l’audible et l’inaudible, il ne représente qu’une approximation du comportement du système auditif humain. De plus, on remarque que la modulation PSK est celle qui permet l’insertion du tatouage avec le plus de transparence. Bien que la modulation par étalement de spectre est la plus utilisée dans la littérature, la séquence pseudo-aléatoire utilisée pour le tatouage peut être audible, même si sa puissance est faible [11, 61, 67].

### 3.2.3 Impact de la fréquence porteuse sur les performances du tatouage

Au cours des dernières années, les chercheurs ont fourni des efforts considérables pour comprendre les caractéristiques du système auditif humain [120] ; pour appliquer ces connaissances en compression audio en prime abord et par la suite en tatouage audio. La FIGURE 3.3 montre une courbe typique du seuil de perception absolu, où l’axe horizontal est la fréquence mesurée en hertz (Hz) et l’axe vertical représente le seuil absolu en décibels (dB). Comme on peut le voir, l’oreille humaine a tendance à être plus sensible à des fréquences dans la bande de 1 à 5 kHz, tandis que le seuil augmente rapidement à des fréquences très élevées et très basses. Il est donc clair qu’en insérant le tatouage dans la bande haute ou basse, la distorsion sera surtout inaudible et donc plus de transparence peut être atteinte.

La fréquence porteuse  $f_c$  est un paramètre important qui influe sur les performances du système de tatouage audio. En effet, la fréquence porteuse  $f_c$  doit satisfaire un compromis entre la transparence et la fiabilité de détection. Néanmoins, pour une valeur fixe de  $f_c$ , la qualité audio et la fiabilité de détection dépendent fortement du signal hôte  $x(n)$ . Ainsi, on ne peut pas choisir immédiatement une fréquence porteuse unique  $f_c$  adaptée à tous les signaux audio. Par conséquent, une première sélection des fréquences porteuses serait nécessaire pour faciliter le processus du tatouage audio.

"Dans chaque enfant il y a un artiste. Le problème est de savoir comment rester un artiste en grandissant".  
 Pablo Picasso

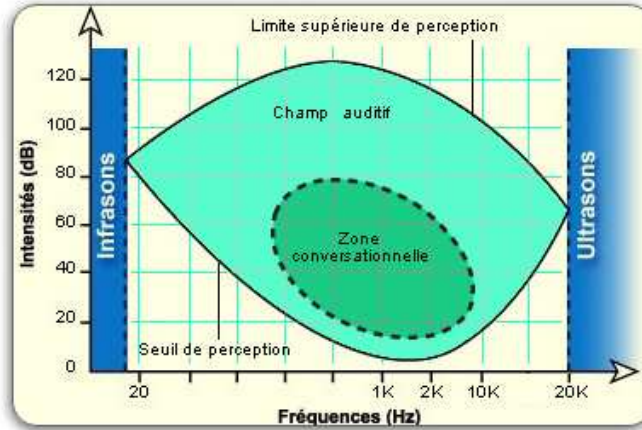


FIGURE 3.3 – Seuil auditif [92].

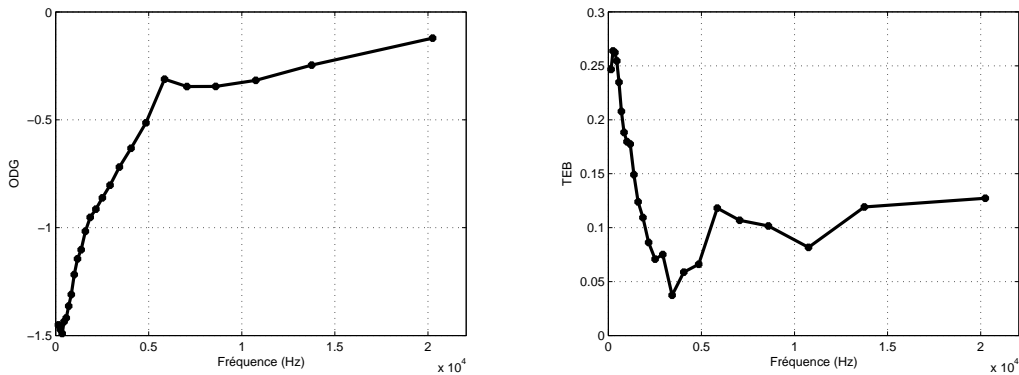
Le système auditif humain peut être modélisé comme un analyseur de fréquence constitué d'un ensemble de 25 filtres passe-bandes (appelé bandes critiques) qui couvrent la gamme audible humaine de 10Hz à 20kHz [120]. Ainsi, deux lignes spectrales (deux sons purs, par exemple) seront perçus comme une seule ligne, avec des puissances égales si leurs fréquences appartiennent au même filtre passe-bande [120]. Par conséquent, nous attribuons à chaque bande critique sa fréquence porteuse correspondante  $\{f_{c_i}\}_{i=\{1,\dots,25\}}$  comme c'est illustré sur le tableau 3.2.

Bande	Fréquence	Bande	Fréquence
1	50	14	2150
2	150	15	2500
3	250	16	2900
4	350	17	3400
5	450	18	4000
6	570	19	4800
7	700	20	5800
8	840	21	7000
9	1000	22	8500
10	1175	23	10500
11	1370	24	13500
12	1600	25	19500
13	1850		

Tableau 3.2 – Normalisation de l'échelle des Barks et découpage en 25 bandes critiques.

Pour la même séquence binaire  $b_i$ , la fréquence porteuse  $f_c$  des formes d'ondes sinusoïdales calculée en (3.22) est modifiée pour obtenir l'évolution des résultats en termes de transparence et de fiabilité de détection.

Afin d'évaluer la transparence du tatouage, on applique l'algorithme PEAQ. Cet algorithme fournit une valeur ODG caractérisant la qualité du signal audio tatoué  $y(n)$  par rapport au signal audio original  $x(n)$ . Cette valeur varie entre 0 (lorsque le tatouage est transparent) à



(a) Variation de l'ODG en fonction des fréquences porteuses  $f_c$ . (b) Variation du TEB en fonction des fréquences porteuses  $f_c$ .

FIGURE 3.4 – Sélection préliminaire des fréquences porteuses qui assurent un compromis entre l'inaudibilité et la fiabilité de détection.

–4 (dans le cas où la dégradation est très importante). La FIGURE 3.4(a) illustre l'évolution de la transparence du tatouage en fonction de la fréquence porteuse  $f_c$ . On remarque que le tatouage devient plus transparent dans les hautes fréquences. Ceci est dû à plusieurs raisons parmi lesquelles :

- Les études psychoacoustiques montrent que l'oreille humaine est moins sensible à la gamme des hautes fréquences [36].
- La gamme de fréquences auditive de 1 kHz à 5 kHz est extrêmement sensible à la membrane basilaire [84].
- La distribution du signal audio de l'énergie est concentrée autour des basses fréquences (au moins dans les signaux audio de test).
- Il est difficile pour les êtres humains de faire la différence entre deux sons complexes par la différence de phase aux composantes hautes fréquences. Cependant, dans la gamme des basses fréquences, les différences de phase entre les composantes de fréquence influent certainement sur le timbre des sons complexes [90].

De plus, l'évolution de la fiabilité de détection, selon la fréquence porteuse  $f_c$ , est représentée sur la FIGURE 3.4(b). On remarque que l'insertion du tatouage dans la bande de fréquence [2, 5] kHz permet l'extraction des données avec moins d'erreurs.

Les fréquences porteuses choisies influent énormément sur les performances du système de tatouage en termes d'inaudibilité et de fiabilité de détection [66]. Sur la FIGURE 3.4, on remarque que le tatouage se comporte différemment en fonction des trois bandes de fréquences :  $\Delta f_1 = [0, 2]$  kHz,  $\Delta f_2 = [2, 5]$  kHz et  $\Delta f_3 = [5, 22.5]$  kHz. Intuitivement, dans la bande de fréquences  $\Delta f_1$ , le tatouage est un peu perceptible et son extraction est très difficile. Toutefois, le tatouage devient moins perceptible dans la bande de fréquence  $\Delta f_2$  et sa détection est améliorée. Enfin, le tatouage est imperceptible dans la bande de fréquence  $\Delta f_3$  et peut être extrait avec un TEB acceptable. Il est donc clair que la bande de fréquence  $\Delta f_1$  n'est pas appropriée pour l'insertion du tatouage

en raison de sa faible performance en termes d'ODG et de TEB. Par conséquent, on se propose d'éliminer les fréquences porteuses appartenant à cette bande de fréquences et on sélectionne les  $K = 13$  fréquences appartenant à  $\Delta f_2$  et  $\Delta f_3$ . Les fréquences obtenues sont stockées dans  $\Delta f_{opt} = \{f_1, f_2, \dots, f_K\}$  et sont triées par ordre croissant selon leur TEB correspondant.

La manière d'insertion du tatouage doit être faite avec soin parcequ'une puissance élevée du tatouage aura un impact sur l'inaudibilité et la fiabilité de détection. En insérant le tatouage dans les hautes fréquences, la modulation PSK assure mieux l'inaudibilité mais entraîne une robustesse faible. D'un autre côté, les signaux audio sont généralement basses fréquences et donc l'insertion du tatouage dans cette bande de fréquence influence sur l'inaudibilité du tatouage et rend sa détection difficile. Enfin, l'insertion du tatouage dans les moyennes fréquences peut procurer au tatouage une robustesse élevée avec une transparence satisfaisante [93].

### 3.3 Tatouage par modulation à porteuse sinusoïdale informée

La conception d'un système de tatouage est un problème difficile vu que la contrainte d'inaudibilité influe sur la fiabilité de détection et vice versa [93]. Ainsi, un système de tatouage performant doit assurer un compromis entre ces deux contraintes. Pour réaliser un tel système, nous proposons une technique de tatouage audio informée qui va choisir automatiquement les paramètres d'insertion adaptés à chaque signal audio. En particulier, le système de tatouage informé doit non seulement garantir une bonne qualité du signal audio tatoué, mais doit aussi améliorer efficacement sa fiabilité de détection.

Le défi d'utilisation des modulations à bande étroite est de pouvoir choisir les paramètres convenables qui assurent l'imperceptibilité, la fiabilité de détection et la robustesse du tatouage. Ainsi, nous allons utiliser la phase  $\varphi$  des formes d'ondes porteuses pour insérer l'information du tatouage. L'amplitude  $A$  appropriée qui satisfait la contrainte d'inaudibilité a été calculée précédemment. Le choix de la fréquence porteuse  $f_c$ , qui assure un compromis entre la transparence et la fiabilité de détection du tatouage, dépend fortement du type du signal audio. Plusieurs méthodes exploitant la connaissance a priori du signal audio, pendant le processus d'insertion du tatouage ont été proposées [8, 72]. Ces techniques, connues sous le nom de tatouage informé, ont prouvé leur aptitude à améliorer les performances du système de tatouage. En effet, la capacité de transmission pouvait être nettement améliorée dès lors que la connaissance a priori du signal audio  $x(n)$  à l'émetteur est exploitée [25]. Cette capacité  $C_I$  est celle d'un canal BBAG soumis uniquement au bruit de perturbation  $P$  :

$$C_I = \frac{1}{2} \log_2 \left( 1 + \frac{\sigma_w^2}{\sigma_P^2} \right) \quad (3.23)$$

Dans cette partie, nous présentons un nouveau système informé qui permet de sélectionner la forme d'onde optimale qui assure le meilleur compromis mentionné ci-dessus. Dans le processus d'extraction, les données insérées seront récupérées à partir seulement du signal tatoué  $y(n)$ . Ensuite, les mêmes formes d'ondes sélectionnées dans le processus d'insertion seront utilisées dans l'étape de démodulation. L'objectif de ce système de tatouage informé consiste à rechercher le tatouage  $w(n)$  tel que le signal modulé  $v(n)$  dont il est issu, vérifie la contrainte d'inaudibilité à un certain degré tout en assurant une bonne détection au niveau du récepteur.

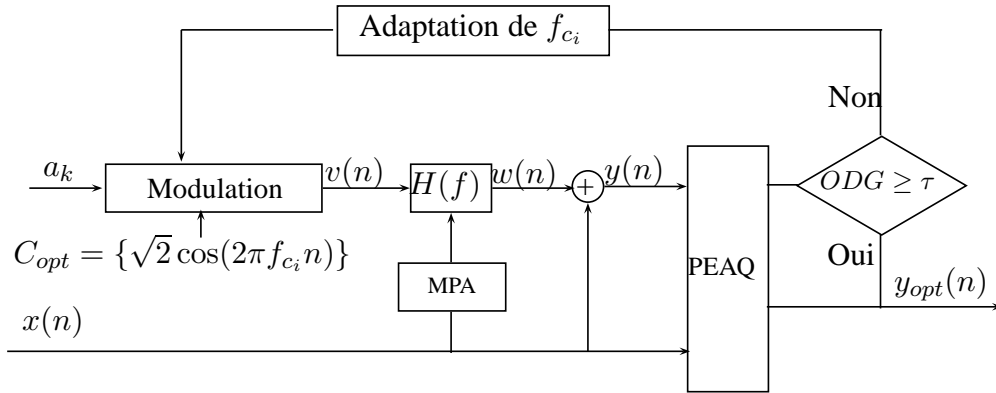


FIGURE 3.5 – Processus d’insertion dans le système de tatouage informé basé sur la modulation PSK.

### 3.3.1 Algorithme adaptatif d’insertion du tatouage

L’avantage d’utiliser des techniques de tatouage informées est de mettre à profit le signal audio  $x(n)$  pour établir une stratégie d’insertion adaptée, permettant de construire un tatouage inaudible et robuste aux perturbations du canal.

La transparence du tatouage est contrôlée par l’utilisation du filtre de mise en forme  $H(f)$  et la fréquence porteuse du signal modulé. Cette fréquence est choisie spécifiquement en fonction du signal audio pour garantir une transparence donnée, fixée par le concepteur du système. Ce choix est effectué par un module de contrôle adaptatif basé sur l’algorithme de mesure objective de la distorsion PEAQ qui évalue la transparence du tatouage.

Les conditions de détection correcte et robuste de l’information dépendent du signal audio dans lequel le tatouage sera inséré. Pour prendre en compte la connaissance de ce signal audio à l’émission, nous utilisons un dictionnaire  $C_{opt}$  contenant des formes d’ondes produisant les meilleures performances en terme de fiabilité de détection. Ces formes d’ondes se composent des fréquences porteuses appartenant au dictionnaire de fréquences  $\Delta f_{opt}$ . Ce dictionnaire permet d’estimer les signaux interférant dans le procédé de détection, et donc d’établir les conditions de la transmission correcte et robuste.

La FIGURE 3.5 illustre le système d’insertion du tatouage informé. Le dictionnaire  $C_{opt}$  est structuré par un ensemble de  $K$  formes d’ondes sinusoïdales qui correspondent aux fréquences porteuses stockées dans  $\Delta f_{opt}$ . Chaque forme d’onde peut être utilisée pour transmettre la séquence de symboles  $a_k$ .

Maintenant, nous visons à faire le choix de la forme d’onde optimale  $c_{opt}(n)$  du dictionnaire  $C_{opt}$  qui sera la plus susceptible d’être détectée tout en garantissant la contrainte d’inaudibilité. Cela est semblable à une technique de saut de fréquence particulière [5, 32, 42, 81], permettant de sélectionner la fréquence porteuse  $f_c$  appropriée à chaque signal audio  $x(n)$ . Le signal tatoué  $y(n)$  sera soumis par la suite à l’algorithme PEAQ qui mesure la qualité du son perçu. Cet algorithme permet de mesurer tout au long de la construction du signal audio tatoué  $y(n)$ , la valeur de l’ODG atteinte qui est comparée à un seuil  $\tau$  afin de garantir la transparence. Ce processus est répété jusqu’à l’obtention du signal de tatouage optimal  $y_{opt}(n)$ . Nous pouvons résumer l’algorithme proposé comme suit :

---

ALGORITHME 1 – Algorithme d’insertion informée du tatouage.

---

**Entrées : Réglage des paramètres**

- Séquence de symboles  $a_k$ .
- Débit du tatouage  $R$ .
- Signal audio  $x(n)$ .
- Seuil d’inaudibilité  $\tau$ .

**Sortie :** Signal tatoué  $y_{opt}(n)$ .

**Étape 1 : Processus préliminaire**

- Sélection des  $K$  fréquences porteuses comme expliqué dans la section 2.4 :

$$\Delta f_{opt} = [f_{c_1}, f_{c_2} \dots f_{c_K}]$$

- Construction du dictionnaire  $C$  calculé en 2.3 :

$$C_{opt} = \{-\sqrt{2} \cos(2\pi f_{c_i} n), \sqrt{2} \cos(2\pi f_{c_i} n)\}_{i=1, \dots, K}$$

**Étape 2 : Mise en forme du signal de tatouage**

- Faire une association bijective entre la séquence  $a_k$  et le signal modulé  $v(n)$ .
- Le signal de tatouage résultant est donné par :

$$w(n) = v(n) * h(n)$$

où  $h(n)$  est la réponse impulsionnelle du filtre de mise en forme :

$$H(z) = \frac{\beta_0}{1 + \sum_{i=1}^P \alpha_i z^{-i}}$$

**Étape 3 : Contrôle d’inaudibilité**

- Insertion du tatouage :  $y(n) = w(n) + x(n)$ .
  - $ODG = PEAQ(x(n), y(n))$ .
  - si  $ODG \geq \tau$  alors retourner  $y_{opt}(n)$ .
  - si  $ODG < \tau$  alors réguler  $f_{c_i}$  et retourner à **Étape 2**.
-

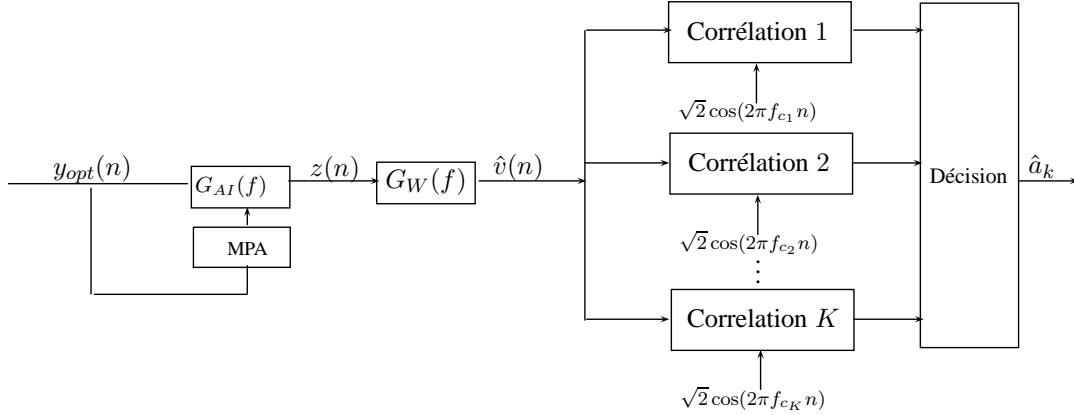


FIGURE 3.6 – Processus d'extraction dans le système de tatouage informé basé sur la modulation PSK.

### 3.3.2 Algorithme d'extraction du tatouage

A la réception, on doit pouvoir récupérer les informations du tatouage  $a_k$  à partir seulement du signal tatoué  $y_{opt}(n)$ , et ceci avec le minimum d'erreurs possible comme c'est illustré sur la FIGURE 3.6. Le système de tatouage est aveugle car le signal audio original est inconnu à l'étape de détection. En supposant que le récepteur est synchrone avec l'émetteur, l'égaliseur de forçage à zéro  $G_{AI}(f)$  est estimé à partir des propriétés de masquage du signal tatoué  $y_{opt}(n)$ . La réponse fréquentielle du filtre de mise en forme  $G_{AI}(f)$ , utilisée à l'émetteur, est obtenue à la réception par une analyse psychoacoustique du signal observé  $y_{opt}(n)$ . On suppose donc que le seuil de masquage de  $y_{opt}(n)$  est approximativement égal à celui de  $x(n)$ . Cet égaliseur permet d'annuler les interférences entre symboles en filtrant le signal tatoué par l'inverse du filtre de mise en forme [9].

L'inconvénient majeur de l'annuleur d'interférences est qu'il rehausse le bruit dans le cas où le canal présente une importante atténuation en fréquence [74]. Par ailleurs, le bruit filtré à la sortie de l'égaliseur est en général corrélé. Par conséquent, le détecteur à corrélation n'est pas dans un contexte fiable pour détecter les symboles émis  $\{a_k\}$ . Afin de rehausser les composantes spectrales de  $y_{opt}(n)$ , on le passe à travers un filtre de Wiener  $G_W(f)$  à la sortie de l'inverse du filtre de mise en forme  $G_{AI}(f)$  [74].

Enfin, la démodulation est effectuée de façon classique par  $K$  détecteurs de corrélation, où  $K$  est le nombre de formes d'ondes sélectionnées dans le processus d'insertion. Dans la phase de décision, il y a une seule forme d'onde du dictionnaire  $C_{opt}$  qui est la plus importante dans la détection des séquences  $b_i$ . Cette forme d'onde, notée  $\tilde{c}_{opt}(n)$ , est la seule forme d'onde qui maximise la corrélation avec le signal reçu  $z(n)$  parmi toutes les formes d'ondes du dictionnaire  $C_{opt}$ . Plus la corrélation entre le signal  $z(n)$  et  $\tilde{c}_{opt}(t)$  est élevée, plus la détection devient facile. Le processus d'extraction peut être résumé comme suit :

---

ALGORITHME 2 – Algorithme d'extraction du tatouage.

---

**Entrée :** Le signal tatoué  $y_{opt}(n)$ .

**Sortie :** La séquence estimée  $\hat{a}_k$ .

**Step 1 : Égalisation**

- Le signal tatoué  $y_{opt}(n)$  passe à travers l'égaliseur de forçage à zéro  $G_{AI}(f) = \hat{H}^{-1}(f)$ . Le signal résultant  $z(n)$  est donné par :

$$z(n) = \hat{v}(n) + r(n)$$

$$\text{où } r(n) = x(n) * \hat{h}^{-1}(n).$$

- On fait passer  $y(n)$  à travers le filtre de Wiener  $G_W(z)$  à la sortie de l'inverse du filtre de mise en forme  $\hat{H}^{-1}(f)$ .

**Step 2 : Choix de la forme d'onde qui sera détectée au récepteur**

- La forme d'onde  $\tilde{c}_{opt}(n)$  qui maximise la corrélation avec le signal reçu  $z(n)$  parmi les formes d'ondes du dictionnaire  $C$  est choisie de telle sorte que :

$$\tilde{c}_{opt}(n) = \operatorname{argmax}_{k=1,\dots,K} z(n)c_k(n)$$

$$\text{où } c_k(n) = \{\pm\sqrt{2} \cos(2\pi f_{c_k} n)\}_{k=1,\dots,K}.$$

**Step 3 : Décision**

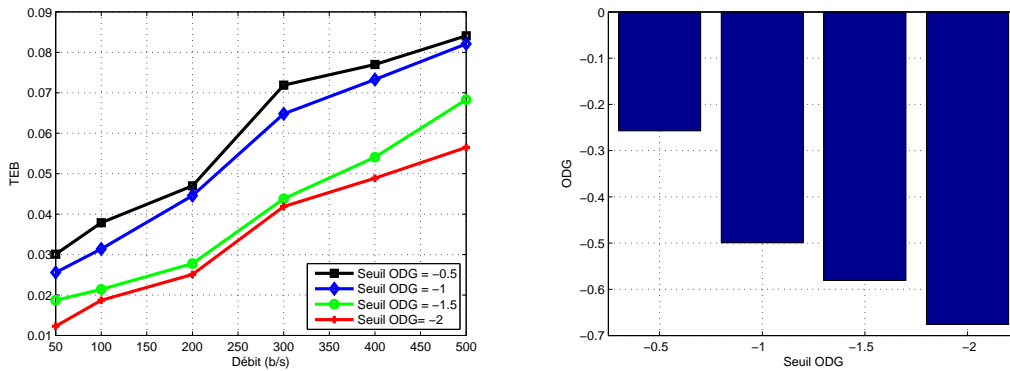
- L'étape de décision sélectionne les symboles estimés  $\hat{a}_k$  qui sont convertis par la suite à la séquence binaire estimée  $\hat{b}_i$ .

---

Puisque  $K$  corrélateurs sont nécessaires au récepteur, le décodeur pourrait être coûteux en terme de calcul quand  $K$  est grand. À cet effet, toutes les fréquences porteuses ont été sélectionnées à partir seulement de 25 bandes critiques, et par la suite en éliminant  $\Delta f_1$ , nous avons utilisé juste  $K = 13$  corrélateurs.

### 3.3.3 Performances

Les performances expérimentales du système de tatouage informé sont testés sur 63 signaux audio selon le protocole expérimental établi dans la section 1.4. Tout d'abord, on va évaluer l'efficacité du système de tatouage informé caractérisé par le contrôle adaptatif de l'insertion du tatouage, en terme d'inaudibilité. Ensuite, on va établir l'efficacité du système en terme de fiabilité de détection. Enfin, la robustesse du système sera évaluée face à quelques perturbations externes en la comparant avec les systèmes présentés précédemment. La performance des différents systèmes en terme de coût de calcul est également établie à la fin de cette section permettant de vérifier si le système peut fonctionner en temps réel ou en différé.



(a) Fiabilité de détection du tatouage selon le seuil d'inaudibilité  $\tau$ . (b) Transparence du tatouage selon le seuil d'inaudibilité  $\tau$ .

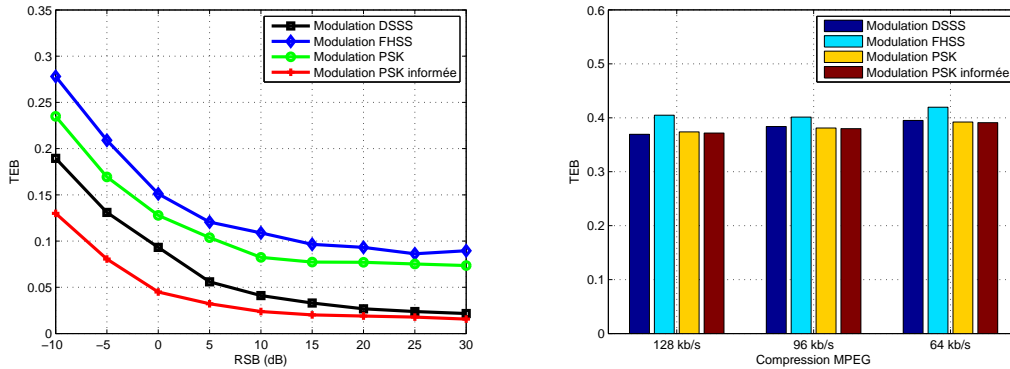
FIGURE 3.7 – Performances du système de tatouage informé en termes d'inaudibilité et de fiabilité de détection en variant le seuil d'inaudibilité.

Les performances du système de tatouage informé, en termes de fiabilité de détection, sont évaluées pour différentes valeurs du seuil d'inaudibilité  $\tau$ . Ce seuil permet de régler le niveau de transparence du tatouage selon les besoins du concepteur du système. On peut dire que la détection du tatouage est améliorée lorsque le seuil d'inaudibilité  $\tau$  diminue comme le montre la FIGURE 3.7(a). Ceci est normal du fait que l'extraction du tatouage avec un minimum d'erreurs est inversement proportionnel à la satisfaction de l'inaudibilité. D'un autre côté, en fixant le seuil d'inaudibilité  $\tau$ , on garantit une transparence inférieure au seuil fixé pour tous les signaux audio. De plus, la valeur moyenne de l'ODG obtenue est moins inférieure au seuil fixé  $\tau$ . La FIGURE 3.7(b) montre que le tatouage reste généralement imperceptible même lorsque le seuil d'inaudibilité est égale à  $\tau = -2$ .

Maintenant, on fixe le débit d'insertion  $R = 100\text{b/s}$  et on évalue la robustesse du système de tatouage informé vis-à-vis de deux types de perturbations : le bruit blanc gaussien et la compression MPEG.

D'abord, on ajoute un bruit blanc gaussien avec un RSB variant de  $-10$  à  $30\text{ dB}$ . Les TEBs obtenus sont illustrés sur la FIGURE 3.8(a). Comme on s'y attendait, le système de tatouage informé améliore la robustesse des techniques à bande étroite et à étalement de spectre non informées contre le bruit blanc gaussien. On remarque aussi que la technique informée présente une nette amélioration par rapport à la technique DSSS quand le tatouage est soumis à un bruit de forte puissance ( $RSB = -10\text{dB}$ ). Néanmoins, cette amélioration se dégrade quand la puissance du bruit s'affaiblit ( $RSB = 30\text{dB}$ ). Ceci est dû à la nature du bruit blanc qui est très proche des séquences pseudo-aléatoires générées par la modulation DSSS.

Le second type de perturbation auquel est soumis le tatouage, est la compression MPEG avec différents débits ( $128, 96, 64\text{kb/s}$ ). La FIGURE 3.8(b) illustre les résultats obtenus pour différentes techniques de tatouage. On remarque que les quatre techniques de tatouage présentent de faibles performances avec un TEB supérieur à  $3.10^{-1}$ . La compression MPEG reste toujours le type de perturbation qui détériore le plus le tatouage puisqu'il élimine une partie importante de l'information.



(a) Robustesse du système de tatouage informé contre le bruit blanc gaussien. (b) Robustesse du système de tatouage informé contre la compression MPEG.

FIGURE 3.8 – Comparaisons entre les performances des modulations numériques en tatouage audio dans le cas d'un canal sans et avec perturbations.

Enfin, le coût de calcul de chacune des méthodes de tatouage est présenté dans le tableau 3.3. Les techniques de tatouage considérées sont :

- Le système de tatouage par étalement de spectre à séquence directe (DSSS) présenté dans **2.4**.
- Le système de tatouage par étalement de spectre à saut de fréquences (FHSS) présenté dans **2.4**.
- Le système de tatouage par déplacement de phase (PSK) présenté dans **3.2**.
- Le système de tatouage par déplacement de phase informé (avec  $\tau = -1$  et  $\tau = -2$ ) présenté dans **3.3**.

Configurations	Coût
Modulation DSSS	2.1457
Modulation FHSS	3.1324
Modulation PSK	2.5324
Modulation PSK informée avec $\tau = -2$	5.3458
Modulation PSK informée avec $\tau = -1$	13.4565

Tableau 3.3 – Coût de complexité des différentes techniques de tatouage.

Le système de tatouage informé conduit à une augmentation du coût de calcul de manière significative. De plus, si on veut améliorer la transparence du tatouage et exiger un seuil d'inaudibilité de l'ordre de  $\tau < -1$ , ceci augmentera encore plus le temps de calcul.

Finalement, le système de tatouage informé, présenté dans cette section, améliore les performances des autres systèmes non informés en termes d'inaudibilité et de fiabilité de détection. Néanmoins, ce système présente des limitations en termes de coût de calcul et de robustesse contre la compression MPEG. Ainsi, il serait nécessaire de surmonter ces deux limitations en présentant un système moins coûteux en temps de calcul et plus robuste à la compression MPEG.

### 3.4 Tatouage par modulation multi-porteuses

Un des problèmes majeurs en tatouage audio, est l'extraction de l'information insérée avec le minimum d'erreurs possibles, même en présence des perturbations externes. Certaines perturbations comme la compression MPEG et le filtrage détériorent complètement le tatouage et rendent difficile sa détection. Ces perturbations peuvent être modélisées comme des distorsions d'évanouissement (ou fading) [71]. Outre la contrainte d'inaudibilité, le problème peut être considéré comme la conception d'un système de communication particulier, où le but est de réduire d'une part l'interférence avec le signal hôte et les interférences entre symboles et d'augmenter d'autre part la robustesse vis-à-vis des perturbations d'évanouissement.

Dans cette partie, on se propose d'adapter les modulations multi-porteuses, largement utilisées en télécommunications sans fil [51], au contexte du tatouage audio. Notre motivation principale est d'exploiter les avantages des modulations multi-porteuses afin d'améliorer les systèmes de tatouage basés sur les techniques sinusoïdales mono-porteuses et celles à étalement de spectre. Le nouveau système de tatouage multi-porteuses doit être en mesure d'augmenter la robustesse du tatouage inséré contre les perturbations d'évanouissement. De plus, la contrainte d'inaudibilité doit être assurée par l'adaptation des paramètres du tatouage qui seront conveables aux caractéristiques du modèle psychoacoustique.

#### 3.4.1 Motivations envers les modulations multi-porteuses

En tatouage audio, l'information insérée souffre de quelques effets, qui peuvent être caractérisés comme suit : l'Interférence Entre Symboles (IES) causé par les filtres d'émission et de réception, l'Interférence avec le Signal Hôte (ISH) provoqué par le signal audio original et l'évanouissement causé par certaines perturbations comme la compression MPEG.

##### 3.4.1.1 Un besoin accru envers les modulations multi-porteuses

En communication numérique haut débit, les transmissions sont limitées par des contraintes physiques [51] : le bruit qui est dû aux imperfections des systèmes et la déformation du signal au cours de la transmission. Elle impose une bonne séparation temporelle des informations émises afin qu'elles restent bien séparées à la réception.

Dans le cas d'un canal à trajets multiples, les techniques de modulations classiques sont très sensibles à l'interférence entre symboles [41]. L'intérêt des modulations multi-porteuses<sup>15</sup> (MP) est de placer l'information dans une fenêtre temps-fréquence telle que sa durée soit plus grande. En effet, l'idée est de transmettre des données numériques en les modulant sur un grand nombre de porteuses en même temps. Ainsi, les systèmes multi-porteuses transmettent les données par blocs : le flux original de données de débit  $R$  est multiplexé en  $N_c$  flux parallèles transmis

---

15. Multi-Carrier Modulation.

avec un débit de  $R/N_c$ . Il s'agit bien d'un multiplexage fréquentiel puisque les données sont transmises sur  $N_c$  canaux différents. Afin d'effectuer cette transmission, au lieu de transmettre les données en série comme le font les modulations mono-porteuses, les modulations multi-porteuses transmettent les données par bloc.

### 3.4.1.2 Robustesse contre l'ISH et l'IES

L'ISH est l'interférence du tatouage  $w(n)$  avec le signal hôte  $x(n)$ . Ce type de perturbations a un grand impact sur la fiabilité de détection. En effet, le signal audio  $x(n)$  est très puissant par rapport au signal de tatouage  $w(n)$ , ce qui rend la détection du message enfoui difficile au niveau du récepteur.

D'un autre côté, l'IES provient des différents filtres utilisés à l'émetteur  $H(f)$  et le récepteur  $\hat{H}^{-1}(f)$ . Quand la durée de la réponse impulsionnelle de ces filtres est supérieure à la durée de symbole  $N_s$ , les symboles  $a_k$  interfèrent entre eux. La détection du  $i^{\text{ème}}$  symbole est interféré avec les symboles qui précèdent (ou qui suivent) [53]. En effet, le signal estimé  $\hat{v}(n)$  peut s'écrire sous la forme suivante :

$$\hat{v}(n) = \sum_k a_k p(n - kN_s) + x(n) * \hat{h}^{-1}(n) \quad (3.24)$$

où  $p(n) = h(n) * \hat{h}^{-1}(n)$  et  $\hat{h}^{-1}(n)$  désigne la réponse impulsionnelle de l'égaliseur de forçage à zéro. Le signal estimé  $\hat{v}(n)$  est échantillonné au temps  $n_i = iN_s$  produisant :

$$\hat{v}(n_i) = a_i + \alpha \sum_{k=-\infty, k \neq i}^{\infty} a_k p[(i - k)N_s] + x(n_i) * \hat{h}^{-1}(n_i) \quad (3.25)$$

où le premier terme représente le  $i^{\text{ème}}$  symbole transmis, le second terme représente le bruit ajouté provenant de l'interférence entre symboles et le dernier terme restant symbolise l'interférence avec le signal hôte.

Ainsi, l'objectif est de minimiser les effets des IES et de l'ISH et de ce fait la détection des données avec le taux d'erreur le plus faible possible. Outre les modulations par étalement de spectre [9, 34, 73], les modulations multi-porteuses constituent une bonne solution pour réduire l'IES. En effet, la modulation multi-porteuses représente les bits du message original par une séquence de symboles qui seront ensuite convertis en des symboles MP. Dans ce cas, en augmentant le nombre de sous-porteuses ; la durée symbole MP devient élevée et le nombre de symboles interférant diminue. Cependant, l'ISH reste toujours présent même en utilisant les modulations multi-porteuses.

### 3.4.1.3 Robustesse contre l'évanouissement

Notre principale motivation, derrière l'utilisation des modulations multi-porteuses, est d'augmenter la robustesse du tatouage contre les bruits perturbateurs comme la compression. En effet, pour un même spectre de fréquence disponible, dans le cas des modulations multi-porteuses, l'information est répartie sur un grand nombre de porteuses modulées à un faible débit symbole. Ces sous-porteuses sont moins sensibles aux défauts du canal de communication. Un défaut majeur est celui de l'évanouissement ou encore la sélectivité en fréquence. Ces deux concepts expriment d'une part le fait que le RSB varie au cours du temps pouvant même s'annuler dans certains

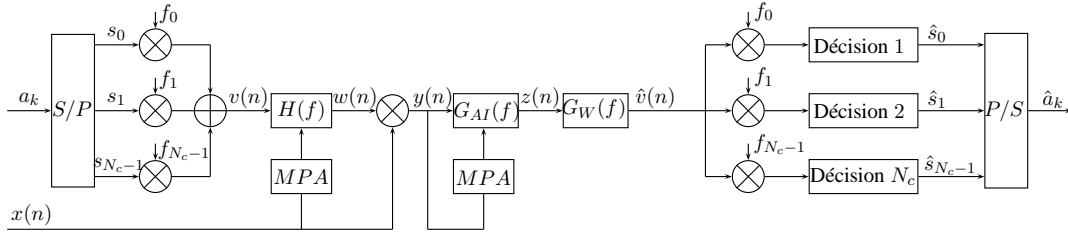


FIGURE 3.9 – Système de tatouage multi-porteuses.

cas et limitent d'autre part les performances en transmettant le signal dans des composantes fréquentielles qui sont atténuées différemment par le canal de propagation.

Les premiers travaux en tatouage ont considéré les effets des différentes perturbations comme l'ajout d'un bruit blanc gaussien stationnaire [15]. Cependant, certaines perturbations telles que la compression MPEG et le filtrage ont un effet différent de celui du BBAG qui détériore complètement le tatouage dans les composants associés au signal audio. Par exemple, le filtrage passe-haut supprime l'existence du tatouage inséré au niveau des coefficients basses fréquences. Par conséquent, ce type de perturbation ne peut être modélisé comme un BBAG, du fait de l'imprévisibilité de la variation du RSB du signal tatoué.

De même, nous affirmons comme dans [71], que de nombreuses perturbations y compris la compression et le filtrage sont plus convenablement modélisés comme des distorsions d'évanouissement. En tatouage audio, la compression MPEG peut être décrite comme un évanouissement sélectif en fréquence. Les approches de tatouage par étalement de spectre sont particulièrement vulnérables à ce type de perturbation ; aucune corrélation résiduelle entre le signal compressé et le tatouage ne peut aboutir à une détection fiable [82]. En outre, ces approches ne prennent pas en compte la non stationnarité du signal hôte. La mise en oeuvre des méthodes multi-porteuses dans le système de tatouage audio peut être une bonne solution pour fournir une grande robustesse contre ces perturbations.

### 3.4.2 Système de tatouage multi-porteuses

Les arguments présentés ci-dessus nous permettent de présenter un système de tatouage audio basé sur les modulations multi-porteuses présenté sur la FIGURE 3.9. L'objectif est de minimiser les interférences entre symboles et d'améliorer la robustesse contre l'évanouissement tout en satisfaisant la contrainte d'inaudibilité. Ainsi, on propose d'exploiter les modulations multi-porteuses afin d'améliorer les performances du système de tatouage audio. Nous montrons d'abord comment construire le signal modulé multi-porteuses  $v(n)$ . Ensuite, nous contrôlons la puissance du signal modulé MP en fonction du seuil de masquage afin d'assurer la contrainte d'inaudibilité. Enfin, nous visons à améliorer la robustesse du système en ajoutant des intervalles de gardes et des codes correcteurs d'erreurs.

#### 3.4.2.1 Modulation et démodulation multi-porteuses

Le principe des modulations multi-porteuses est de convertir une série de bits à haut débit en  $N_c$  symboles bas-débit. Chacun des  $N_c$  symboles est multiplié par différentes sous-porteuses. Puisque le débit symbole de chaque sous-porteuse est beaucoup moins faible que le débit de la suite des symboles des données initiales, l'IES diminue considérablement. En raison de la

conversion série/parrallèle (S/P), le temps discret pour  $N_c$  symboles est étendue à  $N_c N_s$ , ce qui constitue un symbole unique MC avec une longueur de  $N_{sym}$  (soit  $N_{sym} = N_c N_s$ ). L'enveloppe complexe du signal modulé MP s'écrit sous la forme :

$$v_{EC}(n) = \frac{1}{N_c} \sum_{i=0}^{N_c-1} s_i \exp(j2\pi f_i n) \quad 0 \leq n < N_{sym} \quad (3.26)$$

où  $s_i$  sont les  $N_c$  symboles parallèles et les signaux complexes  $\{\exp(j2\pi f_i n)\}_{i=0}^{N_c-1}$  représentent les différentes sous-porteuses du signal MP. Afin d'augmenter l'efficacité spectrale du système de tatouage, il faut choisir des signaux orthogonaux. En effet, les  $N_c$  fréquences des sous-porteuses sont localisées à  $f_k = \frac{kF_c}{N_{sym}} |_{k=\{0, \dots, N_c-1\}}$  afin d'atteindre l'orthogonalité entre les signaux des  $N_c$  sous-porteuses. Par conséquent, le signal résultant  $v(n)$  est appelé signal OFDM<sup>16</sup>. La spécificité de la technique OFDM vient du recouvrement mutuel des différentes sous-porteuses, d'une manière dite orthogonale. Cette orthogonalité permet une utilisation optimale des ressources spectrales et facilite l'implantation numérique.

### 3.4.2.2 Contrôle de puissance

La contrainte la plus importante dans un système de tatouage audio est la contrainte d'in-audibilité. Comme on l'a montré précédemment, un filtre de mise en forme  $H(f)$  est conçu selon un modèle psychoacoustique de telle manière que le signal modulé  $v(n)$  soit de puissance unité :

$$\sigma_v^2 = 1 \quad (3.27)$$

Afin d'exploiter le filtre de mise en forme  $H(f)$  dans le système de tatouage multi-porteuses, on doit adapter l'amplitude du signal MP tout en satisfaisant (3.27). Chaque sous-porteuse du symbole MP de durée  $N_{sym}$ , peut être considérée comme un signal de fréquence unique, multipliée par une fenêtre rectangulaire de longueur  $N_{sym}$ . Le signal modulé  $v(n)$  peut être écrit sous forme d'une somme des  $N_c$  sinusoides de phase aléatoires [53, 70] avec des points de constellations positionnés à un espacement angulaire uniforme autour d'un cercle [53],

$$v(n, \varphi) = \sum_{i=0}^{N_c-1} A \cos(2\pi f_i n + \varphi) \quad (3.28)$$

où  $A$  est l'amplitude de chaque signal de fréquence unique et la phase  $\varphi$  est une variable aléatoire uniformément distribuée sur l'intervalle  $[-\pi, \pi]$  avec une densité de probabilité définie par :

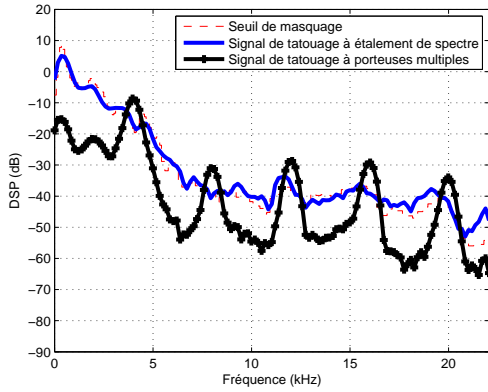
$$p(\varphi) = \frac{1}{2\pi} \quad \text{avec } \varphi \in [-\pi, \pi] \quad (3.29)$$

D'après (3.28) et (3.29), on calcule les moments statistiques de la variable aléatoire  $v(n_k, \varphi)$  à un moment donné  $n_k$ ,

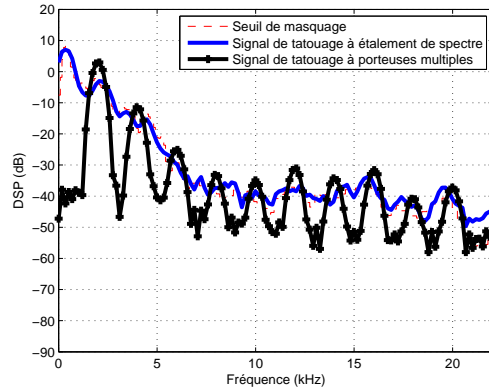
$$\begin{aligned} E[v(n_k, \varphi)] &= E[v(n_k)] \\ &= \int_{-\infty}^{\infty} p(\varphi) \sum_{i=0}^{N_c-1} A \cos(2\pi f_i n_k + \varphi) d\varphi \\ &= \frac{1}{2\pi} \sum_{i=0}^{N_c-1} A \int_{-\pi}^{\pi} \cos(2\pi f_i n_k + \varphi) d\varphi \\ &= 0 \end{aligned} \quad (3.30)$$

---

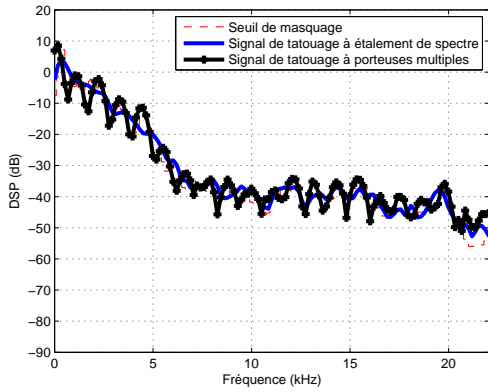
16. Orthogonal Frequency-Division Multiplexing.



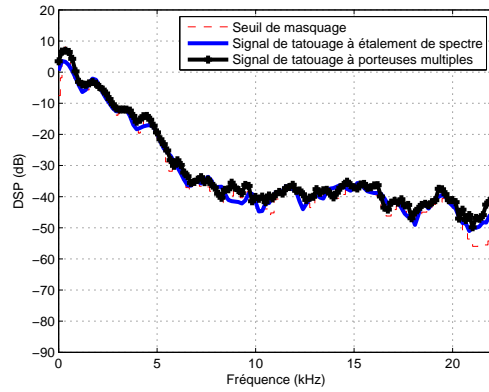
(a) Nombre de sous porteuses : 5



(b) Nombre de sous porteuses : 10



(c) Nombre de sous porteuses : 20



(d) Nombre de sous porteuses : 50

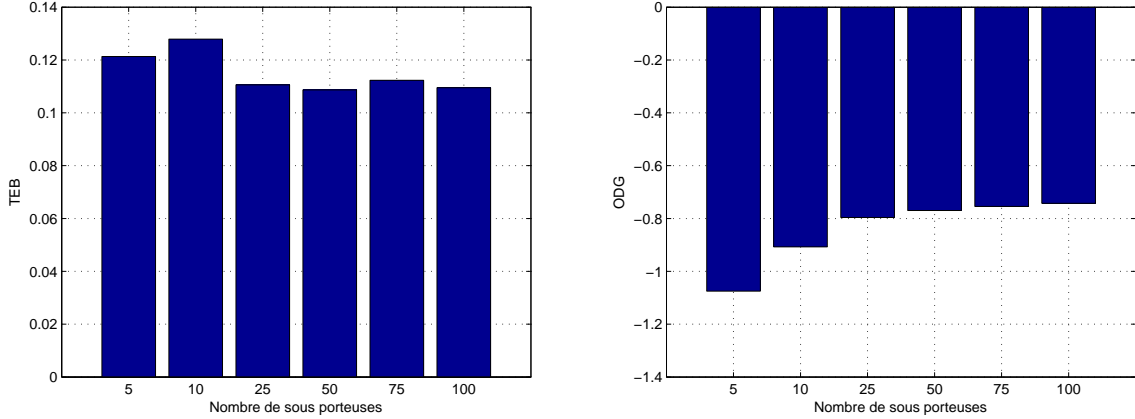
FIGURE 3.10 – Seuil de masquage et DSP du signal de tatouage à étalement de spectre et multi-porteuses selon le nombre de sous porteuses.

et

$$\begin{aligned}
 E[v^2(n_k, \varphi)] &= E[v^2(n_k)] \\
 &= \int_{-\pi}^{\pi} p(\varphi) \sum_{i=0}^{N_c-1} A^2 \cos^2(2\pi f_i n_k + \varphi) d\varphi \\
 &= N_c \frac{A^2}{2}
 \end{aligned} \tag{3.31}$$

Ainsi, le signal modulé  $v(n)$  est à moyenne nulle avec une variance  $\sigma_v^2 = N_c \frac{A^2}{2}$ . Afin de satisfaire la contrainte d'inéditabilité (3.27), on doit choisir  $A = \sqrt{2/N_c}$ . Ensuite, le signal modulé  $v(n)$  est passé à travers le filtre de mise en forme  $H(f)$  de sorte que le signal de tatouage résultant  $w(n)$  ait une DSP  $S_w(f)$  la plus proche possible du seuil de masquage  $M_x(f)$ . La FIGURE 3.10 illustre la DSP du signal de tatouage  $w(n)$  exploitant la modulation multi-porteuses et celle à étalement de spectre ainsi que le seuil de masquage. Cette figure montre que le signal de tatouage MP est considéré comme la somme de plusieurs signaux décalés en fréquence. Les différents symboles sont transmis en parallèle et plus on augmente le nombre de sous-porteuses  $N_c$ , plus le signal de tatouage suit l'allure du seuil de masquage.

La question qui se pose maintenant est comment choisir  $N_c$ ? Le choix de  $N_c$  dépend de



(a) Variation de l'ODG selon le nombre de sous-porteuses. (b) Variation du TEB selon le nombre de sous-porteuses.

FIGURE 3.11 – Variation des performances du système de tatouage multi-porteuses en termes d'inaudibilité et de fiabilité de détection selon le nombre de sous-porteuses.

plusieurs facteurs : le type du canal (variations rapides, longueur de la réponse impulsionnelle) et de la complexité que l'on veut bien accepter. La FIGURE 3.11 montre l'impact du nombre de sous-porteuses  $N_c$  sur les performances du système de tatouage. On remarque que plus on augmente le nombre de sous-porteuses plus l'ODG diminue. Cependant, le nombre de sous-porteuses n'influe pas sur la fiabilité de détection. On choisit donc 25 sous-porteuses du fait qu'à partir de la 25<sup>ème</sup> sous-porteuse l'ODG ne diminue pas trop.

### 3.4.2.3 Intervalle de garde

En prolongeant la durée du symbole  $N_c$  fois, l'IES diminue. Cependant, l'évanouissement reste toujours présent dans le système de tatouage multi-porteuses. Pour améliorer les performances du système MP, on ajoute un intervalle de garde entre deux symboles consécutifs.

L'intervalle de garde permet une grande amélioration des performances des modulations MP sur des canaux à évanouissement et à BBAG [53]. Par conséquent, l'ajout de l'intervalle de garde pour le système de tatouage MP, peut améliorer les performances. L'intervalle de garde peut être inséré de deux manières différentes. L'une est le remplissage de zéros (ZP<sup>17</sup>) qui remplit l'intervalle de garde par des zéros. L'autre est l'extension cyclique du symbole MP (pour une certaine continuité) avec préfixe cyclique (CP<sup>18</sup>) ou suffixe cyclique (CS<sup>19</sup>). Le CP étend le symbole MP en copiant les derniers échantillons du symbole MP avant. Les symboles MP étendus ont une durée de  $N_{symG} = N_{sym} + N_G$  ( $N_G$  désigne la longueur de l'intervalle de garde en terme d'échantillons). Comme la continuité de chaque sous-porteuse a été assurée par le CP, son orthogonalité avec tous les autres sous-porteuses est maintenue sur les  $N_{sym}$ . Notons que le débit symbole MP de données est réduit de  $N_{sym}/N_{symG} = N_{sym}/(N_{sym} + N_{GP})$  fois à cause de l'intervalle de garde.

La FIGURE 3.12 présente les performances du système de tatouage multi-porteuses avec

17. Zero Padding.

18. Cyclic Prefix.

19. Cyclic Suffix.

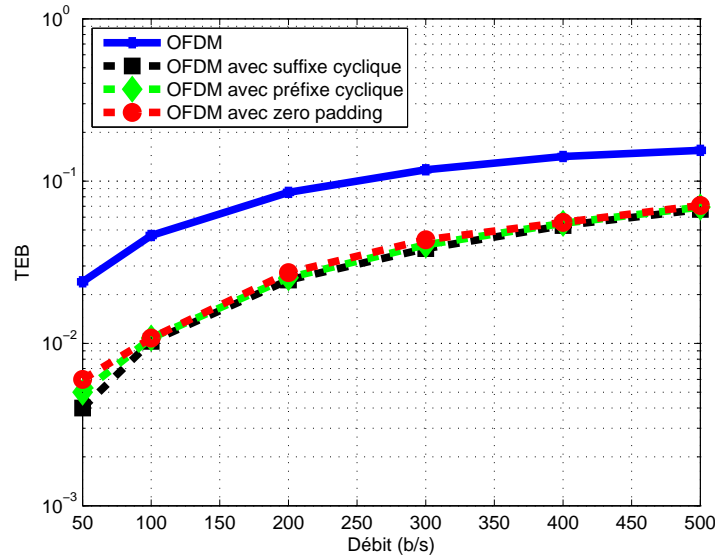


FIGURE 3.12 – Performances du système de tatouage multi-porteuses selon trois types d'intervalles de gardes.

l'ajout des intervalles de garde. Tout d'abord, on remarque que la fiabilité de détection est améliorée dans le système multi-porteuses. Ensuite, les différents intervalles de garde produisent des performances équivalentes avec une nette amélioration pour le CS. Ainsi le CS sera utilisé dans les tests suivants pour évaluer les performances du système de tatouage MP.

#### 3.4.2.4 COFDM

Le système de tatouage multi-porteuses est performant en terme de fiabilité de détection. Afin de rendre sa détection plus fiable, on peut ajouter des bits de correction d'erreurs dont les valeurs dépendent de celles des bits du signal qu'ils accompagnent [20, 114]. Les codes correcteurs d'erreurs sont divisés en deux catégories principales : les codes en blocs et les codes convolutifs [53]. Dans le schéma proposé, nous utilisons des codes en blocs linéaires en raison de la facilité de leur mise en oeuvre et leur performance acceptable.

En utilisant les codes en blocs linéaires, la séquence binaire  $b_i$  est divisée en des blocs de taille de  $k_1$  bits. Le nombre de bits dans le message source codé est augmenté en introduisant des bits de redondances ; ce qui augmente le nombre de bits en blocs à  $k_2$ . Les  $k_1$  bits d'informations ne sont pas modifiés et sont suivis par  $k_2 - k_1$  bits. Le décodage transforme la séquence binaire  $b_i$  en des mots de codes binaires plus longs  $a_m$ . En raison des différentes propriétés des codes en blocs linéaires existants, le codage utilisé fonctionne en codant les bits de tatouage  $b_i$  avec les codes cycliques<sup>20</sup> avant leur insertion dans le signal audio  $x(n)$  [53].

En ajoutant le codage cyclique au système OFDM, le système résultant est appelé COFDM<sup>21</sup>. Tout d'abord, ce système est évalué face au BBAG. En effet, les travaux précédents ont considéré l'effet de plusieurs perturbations équivalent à celui d'un BBAG [71]. La FIGURE 3.13 présente

---

20. Les codes cycliques sont des codes linéaires basés sur des polynômes sur les corps finis.

21. Coded OFDM.

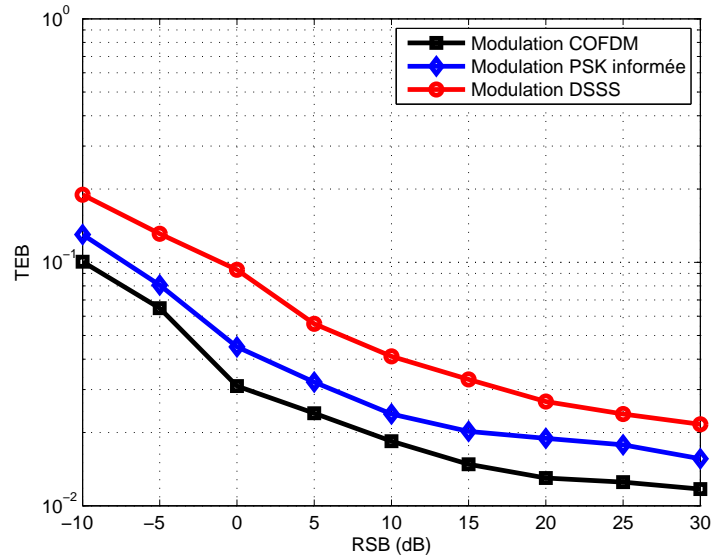


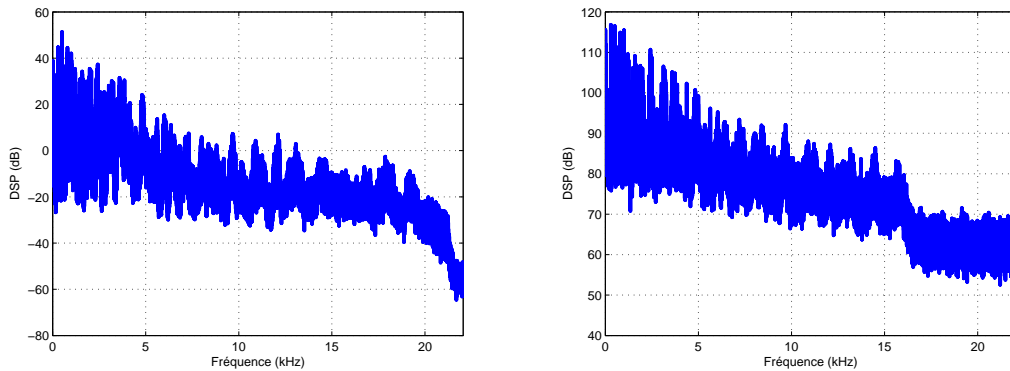
FIGURE 3.13 – Robustesse du système de tatouage COFDM contre le BBAG.

la robustesse du système COFDM contre le bruit blanc gaussien. Le débit d'insertion étant fixé à  $R = 100\text{b/s}$ , on remarque que le système COFDM est robuste vis-à-vis du BBAG. De plus, il présente une nette amélioration par rapport au système DSSS et au système PSK informé. Dans un tel système, le nombre de bits dans le tatouage est augmenté de manière contrôlée en introduisant des bits de redondance. En effet, si le débit binaire sans codage du canal est de  $R$ , le débit après codage de canal devient :

$$R_c = \frac{k_2}{k_1} R \quad (3.32)$$

L'opération du codage est caractérisée par un coefficient  $\frac{k_1}{k_2}$ , appelé rendement du code, qui témoigne de la redondance introduite<sup>22</sup>. Plus  $\frac{k_2}{k_1}$  est grand, plus il y a de redondance : par contre, le débit utile (bits d'information et non de redondance) peut alors devenir négligeable par rapport au débit total. Le rendement de notre système COFDM est de  $3/5$ ; ce qui conduit à une utilisation supplémentaire des échantillons audio pour insérer le tatouage.

En résumé, l'ajout de bits supplémentaires augmente la robustesse du tatouage mais réduit la capacité des informations intégrées pour une longueur de signal audio fixe. Par conséquent, il serait nécessaire d'assurer un compromis entre la robustesse et la capacité d'insertion. Mais ajouter de la redondance s'avère inutile si tous les bits redondants sont transmis sur la même porteuse du canal affecté d'un évanouissement ! Afin d'éviter ces inconvénients, les bits redondants sont transmis sur un grand nombre de porteuses. Comme l'affaiblissement du canal dépend de la fréquence, le message redondant peut passer sans déformation au moins sur quelques fréquences.



(a) Représentation du signal audio dans le domaine fréquentielle. (b) Représentation du signal audio compressé dans le domaine fréquentielle.

FIGURE 3.14 – Extrait du signal audio "Violin" avant et après l'opération de compression.

### 3.4.2.5 Robustesse face à la compression

Une des perturbations les plus destructives du tatouage est la compression MPEG. En effet, le codeur MPEG est assimilé à une opération de filtrage qui supprime les composantes fréquentielles du signal audio qui sont généralement moins sensibles à l'oreille humaine [36].

La FIGURE 3.14 illustre la DSP du signal audio avant et après la compression MPEG 96 kb/s. On remarque que les composantes fréquentielles supérieures à 16 kHz sont perdues. En effet, les hautes fréquences du signal audio sont en grande majorité inférieures au seuil de masquage et sont presque inaudibles. La compression peut donc être modélisée par un filtre passe-bas de fréquence de coupure  $F_c$  égale à 16 kHz dans notre cas.

Ce phénomène de compression est très similaire à celui d'un canal sélectif en fréquence qui ne se comporte pas identiquement suivant la fréquence du signal. Certaines fréquences peuvent être transmises ou encore atténuées plus que d'autres. Le signal est alors déformé lors de la transmission : les données sont dispersées dans le temps, pouvant mener à des interférences entre symboles [51].

Par conséquent, les performances du système de tatouage se dégradent lorsque le signal tatoué est soumis à la compression MPEG. En effet, pour certaines fréquences, les données sont très atténuées et perdues. Ces fréquences sont celles supérieures à 16 kHz. Ainsi, pour augmenter la robustesse du système COFDM, on insère le tatouage juste dans les fréquences inférieures à 16 kHz. De plus, pour assurer la contrainte d'inaudibilité, on doit éviter la bande de fréquence  $[0,6]$ kHz. Pour cela, la bande de fréquences choisie pour le système COFDM est  $[6,16]$ kHz. La nouvelle modulation est appelée COFDM adaptée.

---

22. un rendement de 1/2 équivaut à transmettre deux fois plus de bits

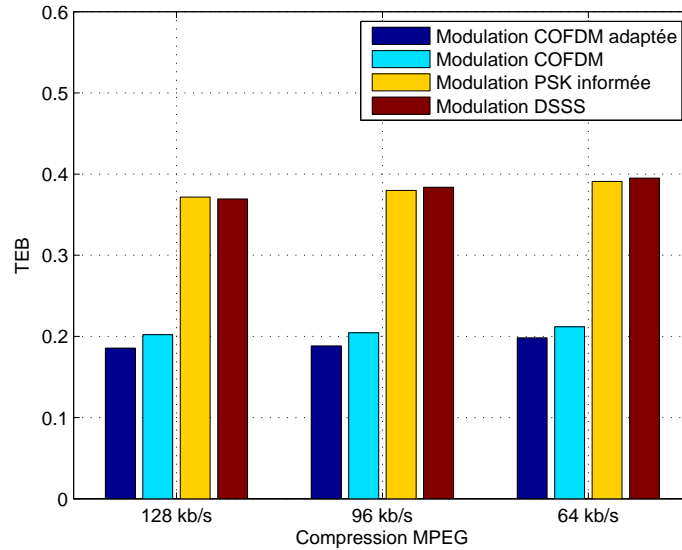


FIGURE 3.15 – Robustesse du système de tatouage COFDM contre la compression.

La FIGURE 3.15 montre le gain apporté par le système COFDM par rapport au système DSSS et le système PSK informé lorsque le signal tatoué est soumis à la compression MPEG. En effet, la compression MPEG quantifie les composantes spectrales de manière non-uniforme à différentes fréquences et élimine les plus élevées afin de préserver un niveau de fidélité perceptuelle [34]. Par conséquent, le système COFDM a été implémenté en choisissant uniquement les basses fréquences des sous-porteuses. Comme prévu, la robustesse du nouveau système est de nouveau améliorée grâce à l'insertion du tatouage dans les composantes basses fréquences du signal audio.

### 3.4.3 Insertion d'un message texte

Pour évaluer le système COFDM proposé à travers une application réelle, nous visons à insérer un message texte jouant le rôle du tatouage dans un signal audio. Le tatouage est inséré avec un débit de 100 b/s et le signal audio utilisé est le "Violin" dont la DSP est illustrée sur la FIGURE 3.14(a). Le message inséré est le texte suivant : **"Une personne qui n'a jamais commis d'erreurs n'a jamais tenter d'innover"**. L'objectif est de pouvoir récupérer le message inséré avec le minimum d'erreurs possible même en présence des perturbations externes. Le signal tatoué a subi deux types de perturbations, à savoir ; l'ajout du bruit blanc gaussien BBAG et la compression MPEG. Le message texte récupéré avec et sans perturbations est présenté sur le tableau 3.4. Quand le système n'est soumis à aucune perturbation, le message est détecté sans aucune erreur au niveau du récepteur. De même, lorsque le signal tatoué est soumis au BBAG avec un RSB de 20 dB, le tatouage est extrait parfaitement. Néanmoins, lorsqu'on diminue le RSB à 10 dB et ensuite à 0 dB, le message est récupéré avec des erreurs ; mais il reste compréhensible. Enfin, lorsqu'on applique la compression MPEG 96 kb/s au système COFDM, le message récupéré devient incompréhensible vu qu'il est détecté avec un TEB de l'ordre de  $10^{-1}$ .

Message texte récupéré	TEB	Perturbations
Une personne qui n'a jamais commis d'erreurs n'a jamais tenter d'innover	0	Sans
Une personng qui n'a jamais commis d'erreurs n'a jamais tenter d'innover	0.0020	BBAG 20dB
Une personne qui l'a jamais commis d'err'qrs n'a jaiais tentgr d'innmver	0.0139	BBAG 10dB
Une personle qui l'a jamais commis d'erreurs n'a jaiaiw tentgr d'mnnorep	0.0159	BBAG 0dB
Toe rersmlne qukanc"jamckq"aoooks d'cptcqrslbej akgmup gjteyy t'ohhipgv	0.1171	Compression

Tableau 3.4 – Extraction des messages textes par le système COFDM.

## Conclusion

L'analyse menée dans la première partie de ce chapitre a permis de mettre l'accent sur le rôle important que joue la modulation dans l'amélioration des performances du système de tatouage audio. Contrairement aux modulations par étalement de spectre utilisées dans les systèmes de tatouage présentés dans la littérature, nous avons intégré les modulations à bande étroite dans le système de tatouage. Après une étude comparative entre les modulations à bande étroite, il s'est avéré que la modulation par déplacement de phase est celle qui assure la meilleure performance en termes de fiabilité de détection.

Dans la deuxième partie du chapitre, nous avons commencé tout d'abord par justifier le choix de la modulation par déplacement de phase. Ensuite, nous avons adapté les paramètres du signal modulé au contexte du tatouage. En effet, nous avons calculé l'amplitude du signal modulé qui permet d'assurer la transparence du tatouage. La fréquence porteuse étant un paramètre essentiel, nous avons montré les bandes de fréquences qu'il faut choisir afin d'assurer un compromis entre l'inaudibilité et la fiabilité de détection.

Dans la troisième partie du chapitre, nous avons présenté un nouveau système de tatouage informé qui permet d'exploiter la connaissance du signal audio. Cette connaissance permet de bien choisir les formes d'ondes convenables pour améliorer les performances en tatouage audio. Ce nouveau système informé a présenté des résultats meilleurs que les systèmes d'étalement de spectre. L'inconvénient majeur de ce système est le coût de calcul qui ne permet pas son fonctionnement dans les applications à temps réel. De plus, le système informé est fragile face à la compression MPEG.

Afin de surmonter le problème du coût et celui de la robustesse, nous avons proposé un autre système multi-porteuse. Ce système présente une robustesse améliorée face aux perturbations externes y compris la compression MPEG. Ceci est dû au fait de l'ajout de l'intervalle de garde au signal multi-porteuse et l'introduction des codes correcteurs d'erreurs qui améliorent la fiabilité de détection. Enfin, nous avons calculé l'amplitude convenable qu'il faut choisir pour assurer la contrainte d'inaudibilité.

## Chapitre 4

# Approche multi-tatouage basée sur les techniques d'accès multiples

Actuellement, l'amélioration de la robustesse et du débit d'insertion présentent un défi à relever en tatouage audio. Toutefois, les études existantes montrent que le débit d'insertion influe sur la robustesse du tatouage [27, 33]. Malheureusement, la théorie d'information ne permet pas de calculer théoriquement la capacité des systèmes de tatouage, comme ce qui est déjà fait avec la capacité de Shannon<sup>1</sup> en communication numérique [101]. Ceci est dû au fait que la nature du bruit en tatouage audio est différente de celle du bruit en communication numérique [66].

Plusieurs travaux en tatouage se sont concentrés sur la recherche d'un système de tatouage robuste par étalement de spectre souvent au détriment du débit d'insertion. A l'inverse, nous proposons différentes techniques d'insertion du tatouage qui visent à augmenter en premier lieu le débit et à maximiser ensuite la robustesse vis-à-vis des perturbations externes. En effet, nous allons montrer comment le débit d'insertion ainsi que la robustesse pourront être améliorés tout en garantissant la contrainte d'inaudibilité. Nous allons présenter trois nouveaux systèmes de tatouage multiples permettant l'insertion de plusieurs tatouages en parallèle. Les systèmes de tatouage multiples seront analysés selon divers aspects, y compris non seulement le débit mais aussi la fiabilité de détection, la robustesse et l'inaudibilité. En d'autres termes, nous allons présenter une nouvelle approche de tatouage multiple qui permet d'insérer le maximum de bits d'informations de manière fiable tout en gardant une bonne qualité du signal audio tatoué. Ce chapitre présente trois systèmes de tatouage multiple qui ont pour but :

- L'augmentation du débit d'insertion en se basant sur les techniques d'accès multiples.
- La transformation du tatouage aveugle vers un autre non-aveugle grâce au concept de la séparation de sources.
- L'insertion et l'extraction de plusieurs tatouages dans des signaux audio à double canal.

---

1. La capacité de Shannon est le débit théorique maximal de transfert d'information sur un canal pour un certain niveau de bruit.

## 4.1 Système multi-tatouage à accès multiples

Les systèmes de tatouage basés sur la technique d'étalement de spectre sont les plus utilisés en tatouage audio, du fait qu'ils offrent une transparence et une robustesse au tatouage inséré. Toutefois, certaines questions fondamentales sur la base d'utilisation de l'étalement de spectre restent ouvertes. Par exemple, une tâche difficile pour la conception d'un système de tatouage par étalement de spectre est d'augmenter la quantité de données insérée, tout en fixant un degré d'inaudibilité et de robustesse. La robustesse est un attribut essentiel du spectre étalé en tatouage audio. Elle est considérablement détériorée si l'on essaie d'atteindre un taux d'insertion élevé en gardant les autres paramètres inchangés. En particulier, on peut inclure les systèmes de tatouage sinusoïdaux, présentés dans le chapitre précédent, dans la classe de tatouage par étalement de spectre. En effet, tous ces systèmes ne permettent l'insertion que d'un seul tatouage dans le même signal audio. Par conséquent, ils restent limités en terme de débit d'insertion.

De plus, la quantité d'informations qu'on peut cacher dans les signaux audio est relativement petite par rapport à celle qu'on peut insérer dans les images et les vidéos. Par conséquent, le tatouage multiple a été proposé pour insérer plusieurs tatouages dans le même signal audio [86]. Cette technique de tatouage est apparue récemment et elle a commencé à connaître un peu plus d'attention de la part des chercheurs ces dernières années [21, 58, 102, 113]. Néanmoins, les performances des techniques proposées restent toujours limitées en terme de débit d'insertion. Outre les interférences entre symboles provoqués au cours du processus d'insertion dans les systèmes mono-tatouage, un nouveau critère s'ajoute dans les systèmes multi-tatouage qui est celui des interférences entre les tatouages eux mêmes.

Dans cette partie, on se propose de présenter un nouveau système multi-tatouage basé sur les techniques d'accès multiples. En effet, les techniques d'accès multiples permettent à plusieurs utilisateurs de partager le même canal de communication [53]. Notre idée est d'exploiter ces techniques pour permettre à plusieurs tatouages de partager le même canal qui est le signal audio. En effet, le problème peut être vu comme la réalisation d'un système de communication multi-utilisateur particulier, où les différents utilisateurs représentent les différents tatouages qui accèdent simultanément au même canal audio. De cette manière, on peut exploiter quelques avantages des techniques d'accès multiples présentés en télécommunications afin d'augmenter la capacité d'insertion (la quantité d'information cachée) dans le nouveau système de tatouage audio.

Dans ce cas, deux facteurs limiteront le débit d'insertion : la fiabilité d'extraction et l'inaudibilité du tatouage. La technique d'accès multiple utilisée peut avoir un impact sur les performances du système multi-tatouage proposé. Par conséquent, nous discutons trois techniques d'accès multiples : l'accès multiple par répartition en code (CDMA<sup>2</sup>), l'accès multiple par saut de fréquences (FHMA<sup>3</sup>) et l'accès multiple par répartition de fréquences (FDMA<sup>4</sup>).

---

2. Code Division Multiple Access.  
3. Frequency Hopping Multiple Access.  
4. Frequency Division Multiple Access.

### 4.1.1 Techniques d'accès multiples

Les techniques d'accès multiples permettent à plusieurs utilisateurs de partager simultanément la bande passante disponible dans un canal de communication [96]. En exploitant l'analogie entre les systèmes de communications numériques et les systèmes de tatouage audio [26], on peut considérer les différents utilisateurs comme plusieurs tatouages qui partagent tous le même canal qui est le signal audio. Le nouveau système sera appelé système multi-tatouage et doit être en mesure d'éliminer les effets des interférences entre les différents tatouages qui limitent d'une part le débit d'insertion et dégradent d'autre part le taux d'erreur binaire. De plus, en augmentant le nombre de tatouages, les données insérées deviennent plus perceptibles. Pour cette raison, on se propose de présenter les trois techniques d'accès multiple; CDMA, FDMA et FHMA qui seront exploitées dans le nouveau système afin d'évaluer leurs impacts sur les performances du système de tatouage en fonction de leurs propres caractéristiques.

#### 4.1.1.1 Accès multiple par répartition en code

La technique CDMA<sup>5</sup> est considérée comme un élément essentiel dans les systèmes de communications mobiles de troisième génération [96]. La raison derrière l'utilisation de cette technique est la grande capacité qu'elle offre si on la compare avec la technique FDMA [51].

Le système CDMA est connu par son élimination des effets d'interférences d'accès multiple (IAM) qui limitent le débit et dégradent les performances de la fiabilité de transmission. Les interférences d'accès multiple représentent l'interférence entre les différents utilisateurs des séquences directes. Avec l'augmentation du nombre d'utilisateurs, les interférences d'accès multiple deviennent importantes et les performances du système CDMA se dégradent.

En exploitant le principe de la technique CDMA en tatouage audio, on peut considérer les utilisateurs comme plusieurs tatouages qui partagent simultanément la même bande de fréquences du canal audio. Chaque tatouage module directement ses informations par un code unique à un taux élevé qui est essentiellement non corrélé avec celui attribué à un autre tatouage.

#### 4.1.1.2 Accès multiple par saut de fréquences

L'accès multiple par saut de fréquences (FHMA) est une autre technique d'accès multiple qui permet à plusieurs utilisateurs d'accéder simultanément au canal de communication en allouant à chaque utilisateur une séquence de fréquences porteuses [53].

En tatouage audio, chaque message est modulé par plusieurs ondes porteuses sur différentes fréquences. Ces fréquences porteuses varient selon le nombre de canaux en utilisant une séquence de saut pseudo-aléatoire. Toutes les fréquences porteuses sont maintenues dans une table connue sous le nom de table de saut. Cette table permet le saut d'une fréquence à une autre dans un intervalle de temps bien déterminé. L'avantage d'utiliser cette technique est que chaque tatouage est réparti sur différents canaux. Ceci permet d'éviter le problème de bruit sur une fréquence particulière. Le saut de fréquences est réalisé au moyen d'un synthétiseur de fréquences, qui est à son tour produit par le générateur des codes pseudo-aléatoires. Au niveau du récepteur, les séquences pseudo-aléatoires doivent être les mêmes que celles utilisées à l'émetteur, sans aucun retard, car cela pourrait provoquer un taux d'erreurs élevé.

---

5. Cette technique aussi connue sous le nom de DS-SS : Direct Sequence-SS.

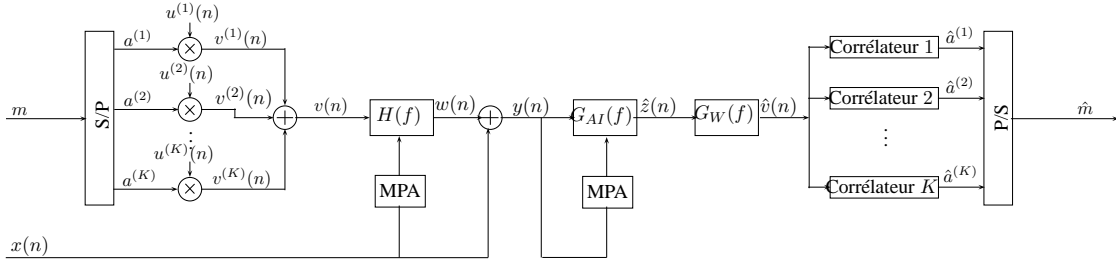


FIGURE 4.1 – Le système multi-tatouage audio proposé.

Bien que la technique FHMA n'offre aucun avantage contre le bruit blanc, le saut de fréquences permet aux signaux de sauter au delà des canaux de fréquences contenant des interférences ou un évanouissement sélectif en fréquence. Pour exploiter pleinement cet avantage contre les signaux perturbateurs à bande étroite, des canaux de fréquences disjoints sont nécessaires. Ces canaux disjoints peuvent être contigus ou peuvent avoir des régions spectrales inutilisées entre elles.

#### 4.1.1.3 Accès multiple par répartition en fréquence

Une des manières les plus simples pour minimiser les interférences entre les canaux adjacents consiste à utiliser des tatouages dont les spectres ne se chevauchent pas. Cette technique d'accès multiple connue sous le nom de FDMA, est couramment utilisée dans les systèmes cellulaires [51].

En tatouage numérique, le canal audio disponible est découpé en un certain nombre de canaux de fréquences qui sont utilisés par plusieurs tatouages. Si la modulation par déplacement de phase M-aire est utilisée pour insérer le tatouage avec un débit  $R$ , la durée de symbole de données est  $T_s = \frac{\log_2(M)}{R}$  et chaque tatouage occupe une bande passante [53] :

$$W = \frac{R}{\log_2(M)} \quad (4.1)$$

Dans les systèmes FDMA, la largeur de bande donnée est divisée en des segments de fréquences adjacents. Chaque segment permet au signal de communication associé de passer à travers un milieu de transmission avec un niveau acceptable.

#### 4.1.2 Système multi-tatouage proposé

Le schéma multi-tatouage proposé est présenté dans la FIGURE 4.1. Dans cette section, on va présenter les différentes étapes de construction du système multi-tatouage. On va inclure les trois techniques CDMA, FHMA et FDMA dans le nouveau système. L'impact du code d'étalement sur les performances du système multi-tatouage sera également établi.

##### 4.1.2.1 Processus d'insertion multi-tatouage

Le message original binaire  $m$  de taille  $N$  est représenté sous forme d'une suite de bits  $\{0, 1\}$  indépendants et identiquement distribués. On convertit le message binaire  $m$  en  $K$  symboles

parallèles  $a^{(k)} = [a_0^{(k)}, a_1^{(k)}, \dots, a_{\frac{N}{K}-1}^{(k)}]$ . Chaque symbole  $a_i^{(k)}$  de durée  $N_s = qN_b$  est choisi parmi  $M = 2^q$  symboles, où  $q$  représente le nombre de bits par symboles. Ensuite, le  $k^{\text{ème}}$  symbole  $a^{(k)}$  est converti en une forme d'onde appropriée selon la technique d'accès multiple utilisée [41, 53] :

$$u_i^{(k)}(n) = \begin{cases} \sum_{l=0}^{L-1} u_l^{(k)} p_{N_c}(n - lN_c) & 0 \leq n < N_c \quad \text{si } \nu = \{\text{CDMA}\} \\ \sqrt{2} \cos(2\pi f_u n) & 0 \leq n < N_h \quad \text{si } \nu = \{\text{FHMA}\} \\ \sqrt{2} \cos(2\pi(f_c + c^{(k)}\Delta f)n) & 0 \leq n < N_s \quad \text{si } \nu = \{\text{FDMA}\} \end{cases} \quad (4.2)$$

où  $L$  est la taille de la forme d'onde  $u_{CDMA}^{(k)}(n)$ ,  $N_c = \frac{N_s}{L}$  est la durée chip, l'impulsion rectangulaire  $p_{N_c}$  est égale à 1 pour  $0 \leq n < N_c$  et 0 ailleurs,  $u_l^{(k)}$  sont les séquences chips de la forme d'ondes  $u_{CDMA}^{(k)}(n)$ ,  $f_u$  est la fréquence porteuse associée au  $h^{\text{ème}}$  saut,  $N_h$  est la durée du saut,  $f_c$  est la fréquence porteuse,  $\Delta f$  est l'espacement des fréquences entre des données adjacentes  $a^{(k)}$  et  $c^{(k)}$  est le code FDMA affecté au  $k^{\text{ème}}$  symbole  $a^{(k)}$ . Ainsi, le signal  $v^{(k)}(n)$  affecté au  $k^{\text{ème}}$  tatouage  $a^{(k)}$  sera donné par

$$v^{(k)}(n) = a^{(k)} u_i^{(k)}(n), \quad 0 \leq n < N_s \text{ et } \nu \in \{\text{CDMA}, \text{FHMA}, \text{FDMA}\} \quad (4.3)$$

Le signal modulé résultant de l'ensemble des  $K$  tatouages est donné par :

$$v(n) = \sum_{k=1}^K v^{(k)}(n) \quad (4.4)$$

Le signal modulé  $v(n)$  ne peut pas être introduit directement dans le signal audio  $x(n)$  du fait que sa puissance  $\sigma_v^2$  est très faible par rapport à celle de  $x(n)$ . Afin d'augmenter la puissance du signal  $v(n)$ , tout en assurant la contrainte d'inaudibilité, et faciliter donc sa détection [9], le signal  $v(n)$  est soumis à un filtre de mise en forme  $H(f)$ .

$$w(n) = v(n) * h(n) \quad (4.5)$$

où  $*$  est l'opération de convolution et  $h(n)$  est la réponse impulsionnelle du filtre de mise en forme  $H(f)$ . Le filtre résultant est un filtre AR (auto-régressif) à réponse fréquentielle purement récursive dont la fonction de transfert  $H(z)$  s'écrit :

$$H(z) = \frac{b_0}{1 + \sum_{i=1}^P a_i z^{-i}} \quad (4.6)$$

Ce filtre transforme le signal modulé  $v(n)$  en un signal de tatouage  $w(n)$  qui assure au mieux la contrainte d'inaudibilité présentée par le seuil de masquage  $M_x(f)$ . La FIGURE 4.2 illustre la densité Spectrale de Puissance du signal de tatouage en utilisant les techniques CDMA, FHMA et FDMA ainsi que l'allure du seuil de masquage  $M_x(f)$ . On remarque que la DSP du signal de tatouage CDMA est répartie sur toute la bande de fréquences du signal audio et sa puissance augmente lorsqu'on incruste plus de tatouages. D'un autre côté, la DSP du signal de tatouage FDMA est très élevée autour des fréquences  $f_c + c^{(k)}\Delta f$  alors qu'elle est très faible dans les autres bandes de fréquences.

Une fois le filtre de mise en forme  $H(f)$  est conçu, le signal de tatouage  $w(n)$  est inséré dans le signal audio  $x(n)$  avec un nouveau débit d'insertion :

$$R = q \frac{KF_c}{N_s} \quad (4.7)$$

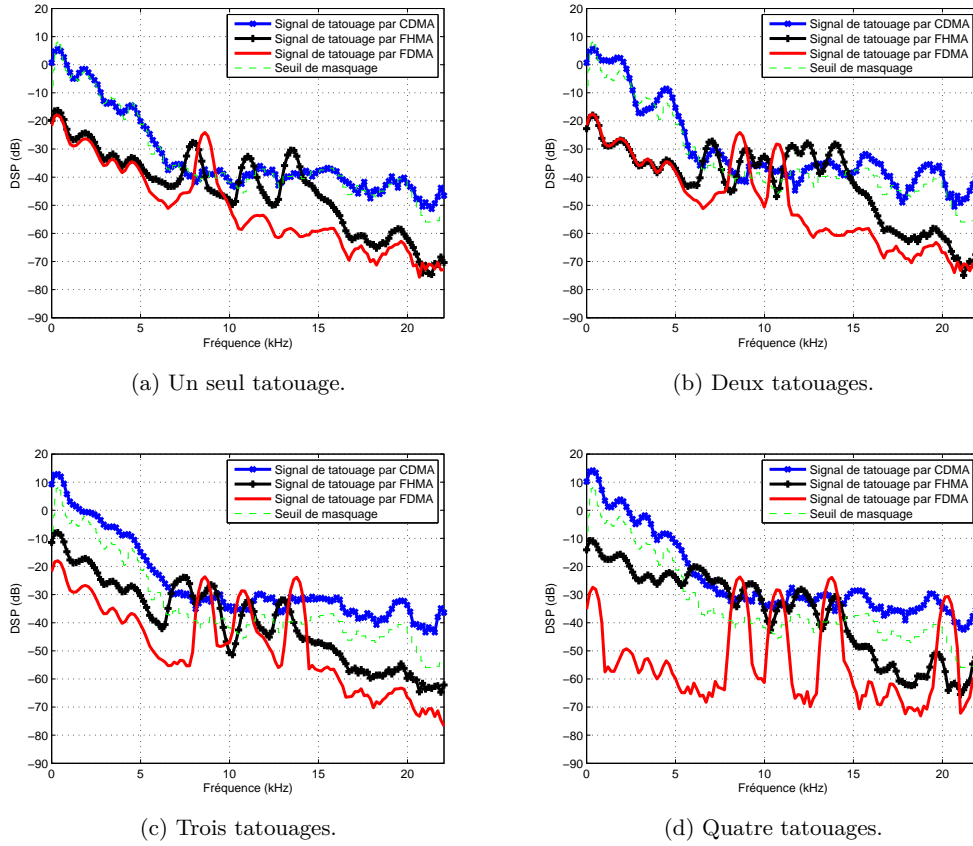


FIGURE 4.2 – Comparaison entre la DSP du signal de tatouage en utilisant les techniques CDMA, FHMA et FDMA et le seuil de masquage.

#### 4.1.2.2 Processus d'extraction multi-tatouage

Dans la phase d'extraction des tatouages, on doit estimer les différentes informations insérées à partir du signal tatoué seul. Pour cela, on fait passer le signal tatoué par l'estimé de l'inverse du filtre de mise en forme  $G_{AI}(f) = \hat{H}^{-1}(f)$  afin d'obtenir un signal  $y(n)$  tel que

$$y(n) = \hat{v}(n) + r(n) \quad (4.8)$$

où  $r(n) = x(n) * \hat{h}^{-1}(n)$  est le signal audio égalisé et  $\hat{v}(n)$  est le signal modulé estimé.

La puissance du signal  $\hat{v}(n)$  est faible par rapport à celle du signal  $r(n)$ . Afin de rehausser les composantes spectrales de  $y(n)$  dominées par celles de  $\hat{v}(n)$  et atténuer celles dominées par  $r(n)$ , on fait passer  $y(n)$  par le filtre de Wiener  $G_W(f)$  à la sortie de l'inverse du filtre de mise en forme  $G_{AI}$  [73]. Enfin, la phase de démodulation est réalisée par des corrélateurs où il est supposé qu'on dispose d'une copie des formes d'ondes  $u_i^{(k)}(n)$  utilisées à l'émetteur.

### 4.1.2.3 Codes d'étalements

Les systèmes de tatouage multiples sont plus difficiles à concevoir que les systèmes de tatouage simple<sup>6</sup>. En effet, les différents tatouages doivent être indépendants et décorrélés entre eux ; la modification d'un tatouage ne doit pas impacter les autres tatouages. Ainsi, il serait nécessaire de bien choisir les formes d'ondes du signal modulé. Ceci se traduit par un choix judicieux des codes d'étalements qui peuvent améliorer les performances en terme de fiabilité de détection.

Le choix du code d'étalement adéquat est important, car les propriétés d'auto-corrélation et la longueur du code fixe sont liées à la capacité du système. Les codes d'étalement CDMA peuvent être divisés en deux catégories : les codes orthogonaux et les codes non orthogonaux. Les séquences Walsh appartiennent à la première catégorie, tandis que les autres séquences (MLS<sup>7</sup> et Kasami) sont des séquences par registre à décalage [51]. D'un autre côté, Il existe deux types de codes dans la technique FHMA : l'étalement par saut de fréquences lent (SFH<sup>8</sup>) et l'étalement par saut de fréquences rapide (FFH<sup>9</sup>).

Lorsque les codes des tatouages sont orthogonaux, le signal de sortie du corrélateur dans le récepteur est égal à zéro à l'exception de la séquence souhaitée. Ainsi, nous étudierons l'impact des codes d'étalements sur les performances du système multi-tatouage.

### Codes MLS

Les types de codes les plus utilisés en tatouage par étalement de spectre sont les codes pseudo-aléatoires (ou MLS). Ces codes sont aussi connus sous le nom de séquences-m. Les séquences pseudo-aléatoires assurent les propriétés suivantes [51] :

- La fonction d'autocorrélation de ce type de codes est périodique et à valeurs binaires.
- Ils ont une corrélation très faible entre chaque deux versions de séquences décalées et une corrélation croisée faible entre chaque deux séquences.
- Dans chaque période de séquence MLS, le nombre des uns est toujours supérieur ou égale au nombre de zéros. Donc, il doit y avoir  $2^{m-1}$  uns et  $2^{m-1} - 1$  zéros dans une période entière de la séquence.

Les codes pseudo-aléatoires sont utilisés pour obtenir un signal à large bande. Leurs propriétés de bruit donnent une auto-corrélation étroite qui est bénéfique pour la synchronisation. La connexion des registres à décalage est déterminée par les coefficients  $h_i$ . Le polynôme générateur de la séquence  $h(p) = h_n p^n + h_{n-1} p^{n-1} + \dots + h_1 p + h_0$ , est un polynôme binaire de degré  $n$ . Les coefficients  $h_0 = h_n = 1$  et les  $h_i \in [0, 1]$ . Le signal  $a_i$  peut être écrit sous la forme :

$$\sum_{k=0}^{n-1} h_k a_{i-k} = a_{i-n} = h_0 a_i + h_1 a_{i-1} + \dots + h_{n-1} a_{i-n+1} \quad (4.9)$$

---

6. On veut dire par système de tatouage simple, celui qui permet l'insertion d'un seul tatouage.

7. Maximum Length Sequence.

8. Slow Frequency Hopping.

9. Fast Frequency Hopping.

## Codes Kasami

Les codes Kasami sont des séquences binaires très utilisées en raison de leur très faible corrélation croisée [51]. Ils sont principalement générés à partir d'une seule séquence-m de longueur maximale égale à  $2^n - 1$  tel que  $n$  est un nombre pair. Les séquences Kasami sont générées avec  $M = 2^{n/2}$  séquences binaires de période  $N = 2^{n-1}$ , où  $n$  est pair.

## Codes Walsh

Les codes Walsh sont utilisés pour atteindre une orthogonalité entre les différents tatouages. Le récepteur peut alors identifier le tatouage souhaité grâce à l'orthogonalité, sans que les autres messages n'interfèrent entre eux. Les codes Walsh sont obtenus par des rangées de mots de code de matrice carrée spéciale appelée matrice de Hadamard. La matrice contient une rangée de tous les zéros et les autres lignes ont chacune le même nombre des uns et des zéros [51].

Les codes Walsh-Hadamard sont générés par un ensemble de  $N = 2^n$  codes avec une taille  $N = 2^n$ . La matrice de Hadamard est définie comme suit :

$$H_{2N} = \begin{pmatrix} H_N & H_N \\ H_N & \overline{H_N} \end{pmatrix} \quad (4.10)$$

avec  $H_1 = [0]$  et  $N$  est une puissance de 2 qui désigne le complément binaire des bits dans la matrice et  $\overline{H_N}$  est le complément de  $H_N$ . En prenant comme exemple  $n = 8$ , on peut générer des codes Walsh codés sur 8 bits :

$$H_8 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \quad (4.11)$$

## Codes FFH

L'étalement par saut de fréquence rapide FFH est caractérisé par la variation des fréquences porteuses durant la transmission d'un seul symbole. Ainsi, le débit chip résultant peut être réécrit sous la forme :

$$R_c = R_h = K_2 R_s \quad (4.12)$$

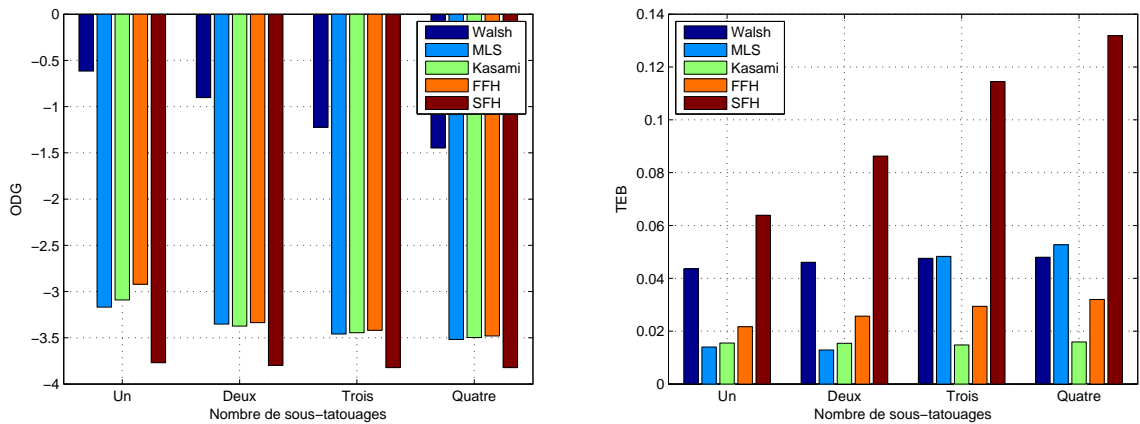
où  $K_2$  est le nombre de sauts par symbole,  $R_h$  est le débit saut et  $R_s$  est le débit symbole.

## Codes SFH

L'étalement par saut de fréquence lent SFH est caractérisé par la transmission de plusieurs symboles par saut (chaque symbole est un chip). D'où le débit chip résultant peut s'écrire sous la forme :

$$R_c = R_s = K_1 R_h \quad (4.13)$$

où  $K_1$  est le nombre de symboles par saut.



(a) Valeurs ODG pour différents codes d'étalements.

(b) Valeurs TEB pour différents codes d'étalements.

FIGURE 4.3 – Performances du système multi-tatouage selon le nombre de tatouages insérés.

#### 4.1.2.4 Impact des codes d'étalement

Les performances des codes d'étalement sont évaluées selon deux critères : l'inaudibilité et la fiabilité de détection des tatouages. Ces deux propriétés sont testées avec un débit d'insertion allant de  $R = 200\text{b/s}$  (quand on insère un seul tatouage) jusqu'à  $R = 800\text{b/s}$  (quand on insère quatre tatouages).

Tout d'abord, on compare les différents codes d'étalement en terme d'inaudibilité. La FIGURE 4.3(a) présente les valeurs ODG obtenues selon le nombre de tatouages insérés dans le signal audio. Elle montre que la qualité audio se dégrade quand la quantité d'informations insérée augmente. La qualité audio dépend aussi du code d'étalement utilisé dans la phase de modulation. On remarque que les codes Walsh sont ceux qui assurent le plus la transparence des tatouages. En effet, on peut insérer deux tatouages avec une très bonne qualité audio. Mais, en insérant plus de trois tatouages la qualité commence à se dégrader un peu.

Une fois qu'on a évalué la transparence des tatouages, il serait nécessaire de les récupérer à partir du signal tatoué [27]. La FIGURE 4.3(b) montre l'évolution de la fiabilité de détection dans le cas d'insertion d'un seul tatouage jusqu'à l'insertion de quatre tatouages. On remarque que lorsque le nombre de tatouages augmente, le TEB augmente aussi et la fiabilité de détection se dégrade. Ceci est dû aux interférences d'accès multiples produits par les différents tatouages, et aux interférences entre symboles produites dans chaque tatouage. Quand on insère un seul tatouage, tous les codes d'étalement permettent une fiabilité de détection acceptable. Cependant, quand on augmente le débit d'insertion en insérant plus de tatouages, les codes Kasami sont les meilleurs codes qui donnent un TEB minimum. Les codes MLS produisent également un bon TEB quand on insère un seul tatouage mais les performances se dégradent de manière significative lorsqu'on insère plus de deux tatouages.

### 4.1.3 Évaluation des performances du système multi-tatouage

#### 4.1.3.1 Paramétrisation du débit d'insertion

Le débit d'insertion est le nombre de bits insérés dans le même signal audio par unité de temps et il est mesuré en bits par seconde. Ainsi, le débit d'insertion peut augmenter en variant trois paramètres selon (4.7) :

- $q$  : Le nombre de bits par symbole.
- $N_s$  : le temps symbole discret.
- $K$  : le nombre de symboles parallèles  $a^{(k)}$  (ou tatouages  $v^{(k)}(n)$ ).

Par conséquent, il existe plusieurs scénarios pour augmenter le débit d'insertion  $R$  :

- On fixe  $q$  à 1 et  $K$  à 1 : Le tableau 4.1 présente le temps symbole discret  $N_s$  selon le débit d'insertion  $R$ .
- On fixe  $N_s$  à 441 et  $K$  à 1 : Le tableau 4.2 présente le nombre de bits par symbole  $q$  selon le débit d'insertion  $R$ .
- On fixe  $N_s$  à 441 et  $q$  à 1 : Le tableau 4.3 présente le nombre de tatouages  $K$  selon le débit d'insertion  $R$ .

R(b/s)	100	200	300	400	500	600	700	800	900	1000
$N_s$	441	221	147	110	88	74	63	55	49	44

Tableau 4.1 – Le temps symbole  $N_s$  selon le débit d'insertion  $R$ .

R(b/s)	100	200	300	400	500	600	700	800	900	1000
$q$	1	2	3	4	5	6	7	8	9	10

Tableau 4.2 – Le nombre de bits par symboles  $q$  selon le débit d'insertion  $R$ .

R(b/s)	100	200	300	400	500	600	700	800	900	1000
$K$	1	2	3	4	5	6	7	8	9	10

Tableau 4.3 – Le nombre de tatouages  $K$  selon le débit d'insertion  $R$ .

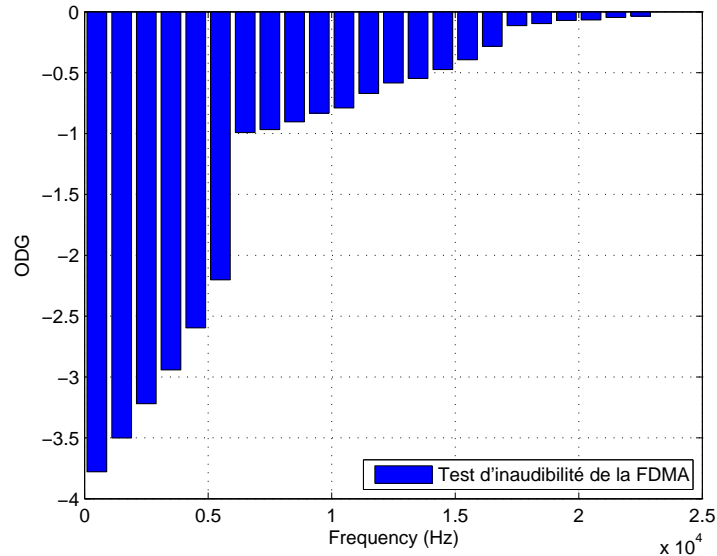


FIGURE 4.4 – Transparence du tatouage selon la fréquence porteuses  $f_c$ .

#### 4.1.3.2 Performances du système multi-tatouage

Les performances du système multi-tatouage est évalué selon l'inaudibilité, la fiabilité de détection et la robustesse; et ce en fonction du débit d'insertion. Tout d'abord, on évalue les performances du système multi-tatouage en comparant les trois techniques CDMA, FHMA et FDMA. En se basant sur l'étude comparative entre les codes d'étalement réalisée précédemment, on utilisera par la suite les codes Walsh dans la technique CDMA et les codes FFH dans la technique FHMA.

En comparant les techniques CDMA et FHMA (c.f. FIGURE 4.3), nous pouvons constater que la qualité audio se dégrade au fur et à mesure qu'on augmente le nombre de tatouages insérés. De plus la technique FHMA ne garantit pas du tout la contrainte d'inaudibilité. De même, en utilisant la technique FDMA, la qualité du signal audio va se dégrader encore plus. Ceci est dû au fait que la DSP du signal de tatouage FDMA est concentrée autour de quelques fréquences porteuses. Si le signal audio est très puissant autour des fréquences porteuses choisies par la technique FDMA, les tatouages insérés seront perceptibles. Ainsi, le choix des fréquences porteuses dans les techniques FDMA et FHMA auront un grand impact sur la transparence des tatouages. En effet, la FIGURE 4.4 montre que les hautes fréquences assurent mieux l'inaudibilité que les basses fréquences [38]. Par conséquent, on choisit juste les hautes fréquences du signal de tatouage FDMA et FHMA et on évalue l'imperceptibilité des tatouages.

La FIGURE 4.5(a) montre que la qualité audio du signal tatoué est améliorée grâce au choix judicieux des fréquences porteuses des techniques FHMA et FDMA. En effet, on remarque qu'on peut insérer :

- Dix tatouages en utilisant la technique FDMA.
- Cinq tatouages en utilisant la technique CDMA.

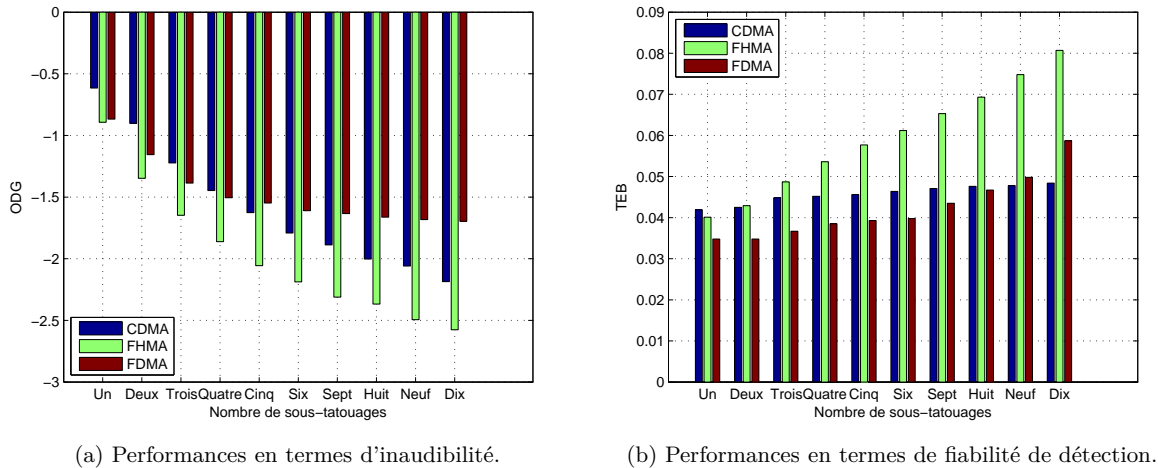


FIGURE 4.5 – Comparaisons entre les techniques d'accès multiples dans le système multi-tatouage.

- Trois tatouages en utilisant la technique FHMA.

avec le même degré d'inaudibilité  $ODG \simeq -1.6$ .

D'un autre côté, il serait important de comparer les performances de ces techniques en terme de fiabilité de détection. En augmentant le débit d'insertion  $R$ , la fiabilité de détection et la robustesse se dégradent. Mais il y a une grande différence dans la source d'erreurs. En effet, en diminuant la valeur de  $N_s$  ou en augmentant la valeur de  $q$ , la source d'erreurs provient des interférences entre symboles. Diminuer la valeur de  $N_s$  est équivalent à augmenter la valeur de  $q$ , ainsi on choisit  $q = 1$  dans tout le reste des simulations. Néanmoins, en augmentant la valeur de  $K$ , l'interférence d'accès multiples est introduit dans le cas d'utilisation du CDMA et l'interférence entre les canaux adjacents apparaît lorsqu'on utilise la FDMA. De plus, en diminuant la valeur de  $N_s$ , on améliore le débit d'insertion  $R$  sans impacter la contrainte d'inaudibilité. Par contre, quand on augmente la valeur de  $K$ , on perd en transparence.

Ainsi, on fixe  $N_s$  à 221, la FIGURE 4.5(b) montre l'évolution de la fiabilité de détection selon le nombre de tatouages insérés  $K$ . On remarque que la fiabilité de détection se dégrade quand le nombre de tatouages  $K$  augmente. La technique FDMA est celle qui présente le meilleur TEB jusqu'à l'insertion de huit tatouages. Ceci est dû au fait que les canaux adjacents se chevauchent entre eux produisant des interférences entre les tatouages. Cependant, la CDMA présente un TEB meilleur que la FDMA lorsqu'on insère plus de huit tatouages. Ceci est dû au fait que la CDMA permet d'éliminer les effets d'interférence d'accès multiple lorsque le nombre de tatouages devient grand.

Maintenant, on fixe la valeur de  $K$  à 4 et on augmente ensuite le débit d'insertion en diminuant la valeur de  $N_s$ . Même si le paramètre  $N_s$  n'affecte pas la qualité audio du signal tatoué, on est toujours limité par la fiabilité de détection. En effet, en diminuant  $N_s$ , les interférences entre symboles  $a_i^{(k)}$  augmentent et par conséquent la détection des tatouages devient plus difficile. Ainsi, la fiabilité de détection se dégrade quand  $N_s$  diminue et la technique FDMA est celle qui présente les meilleures performances en terme du TEB comme c'est illustré dans la FIGURE 4.6.

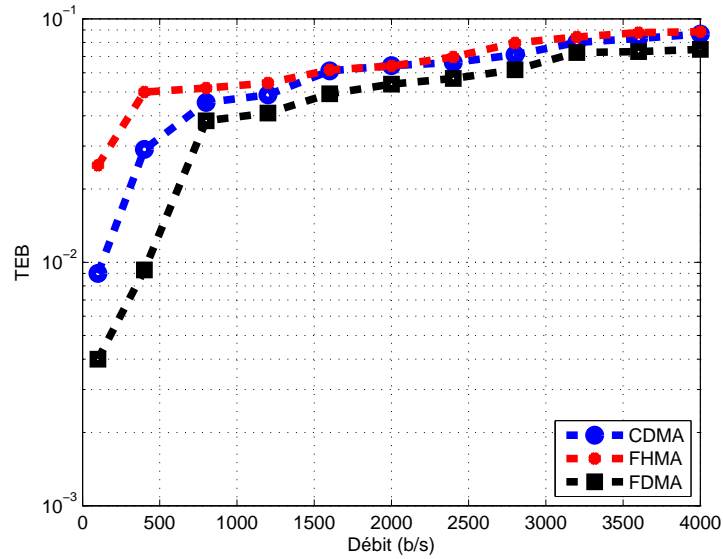


FIGURE 4.6 – Fiabilité de détection des différentes techniques d'accès multiples selon le débit d'insertion.

Perturbations	Performances (TEB)		
	CDMA	FHMA	FDMA
Filtre passe-bas	$2.064 \times 10^{-1}$	$2.298 \times 10^{-1}$	$1.731 \times 10^{-1}$
Filtre passe-haut	$2.235 \times 10^{-1}$	$1.041 \times 10^{-1}$	$5.81 \times 10^{-2}$
BBAG	$1.102 \times 10^{-1}$	$1.047 \times 10^{-1}$	$9.98 \times 10^{-2}$
Quantification	$1.067 \times 10^{-1}$	$9.88 \times 10^{-2}$	$9.45 \times 10^{-2}$
Amplification	$1.055 \times 10^{-1}$	$9.86 \times 10^{-2}$	$9.28 \times 10^{-2}$
Rééchantillonnage	$1.103 \times 10^{-1}$	$1.017 \times 10^{-1}$	$9.82 \times 10^{-2}$
Compression MPEG 128kb/s	$4.532 \times 10^{-1}$	$4.286 \times 10^{-1}$	$4.328 \times 10^{-1}$
Compression MPEG 96kb/s	$4.561 \times 10^{-1}$	$4.341 \times 10^{-1}$	$4.448 \times 10^{-1}$
Compression MPEG 64kb/s	$4.770 \times 10^{-1}$	$4.525 \times 10^{-1}$	$4.578 \times 10^{-1}$

Tableau 4.4 – Evaluation de la robustesse des techniques d'accès multiples vis-à-vis des perturbations externes à  $R = 4\text{kb/s}$ .

D'habitude, le signal tatoué  $y(n)$  est soumis à un ensemble de perturbations externes. Ainsi, on évalue le système multi-tatouage contre différents types de distorsions. La robustesse du système multi-tatouage contre les différentes perturbations est présentée dans le tableau 4.4. Les distorsions considérées produisent les valeurs du TEB lorsque le débit d'insertion est fixé à 4 kb/s. Ce tableau montre que les valeurs du TEB pour le tatouage extrait se dégrade, mais avec une fiabilité de détection acceptable sauf pour la compression et le filtrage passe-bas. À

partir de ces résultats, on constate que la technique FDMA est la plus robuste, contre toutes les perturbations considérées. D'un autre côté, on remarque que la CDMA est plus robuste que la FHMA contre certaines perturbations comme : le bruit additif, la quantification, l'amplification et le rééchantillonnage. Cependant, la FHMA permet une meilleure extraction du tatouage lorsque le signal tatoué est soumis à d'autres distorsions comme : le filtre passe-bas, le filtre passe-haut et la compression. Ceci est dû à la sélection adaptative des fréquences porteuses de la technique FHMA en fonction de chaque perturbation. Par exemple, nous avons inséré le tatouage  $w(n)$  dans les fréquences audio basses lorsque le signal tatoué a été soumis au filtre passe-bas et vice-versa pour le filtre passe-haut et d'autres perturbations. Une des distorsions les plus endommageantes est la compression MPEG. En utilisant la FDMA, on remarque que la robustesse du nouveau système est améliorée par rapport à la CDMA, mais reste toujours médiocre. En effet, il est connu que la compression MPEG quantifie les composantes spectrales de manière non uniforme à des fréquences différentes et élimine les fréquences les plus élevées, afin de conserver un niveau de fidélité perceptuelle [71]. Par conséquent, la technique FDMA est mise en oeuvre en choisissant uniquement les fréquences basses des sous-porteuses. Le tableau 4.5 illustre les bandes de fréquences des techniques FDMA et FHMA choisies pour chaque distorsion.

Perturbations	Bande de fréquences (kHz)
Filtre passe-bas	[6,8]
Filtre passe-haut	[8,22.05]
BBAG	[6,22.05]
Quantification linéaire	[6,22.05]
Amplification	[6,22.05]
Rééchantillonnage	[6,22.05]
Compression MPEG 128kb/s	[6,16]
Compression MPEG 96kb/s	[6,16]
Compression MPEG 64kb/s	[6,16]

Tableau 4.5 – Bande de fréquences utilisées pour les techniques FDMA et FHMA.

L'inaudibilité, la fiabilité de détection ainsi que la capacité d'insertion sont les trois critères que nous avons utilisés pour évaluer les performances du système multi-tatouage proposé. La technique FDMA a prouvé sa supériorité en tatouage audio par rapport à la CDMA, chose qui peut sembler en contradiction avec les résultats connus en télécommunications [53]. En effet, dans les communications sans fil, la technique CDMA donne des résultats<sup>10</sup> meilleurs que la FDMA en termes de TEB et de débit. Néanmoins, les différences entre les systèmes de communications et les systèmes de tatouage rendent les performances de la FDMA meilleures que celle de la CDMA. Ces différences sont les suivantes :

- **Inaudibilité** : le choix judicieux des fréquences porteuses permet au système FDMA d'insérer plus de tatouages que la CDMA tout en garantissant la contrainte d'inaudibilité.

---

10. Le système CDMA est capable d'offrir un débit huit fois plus élevé que le système FDMA.

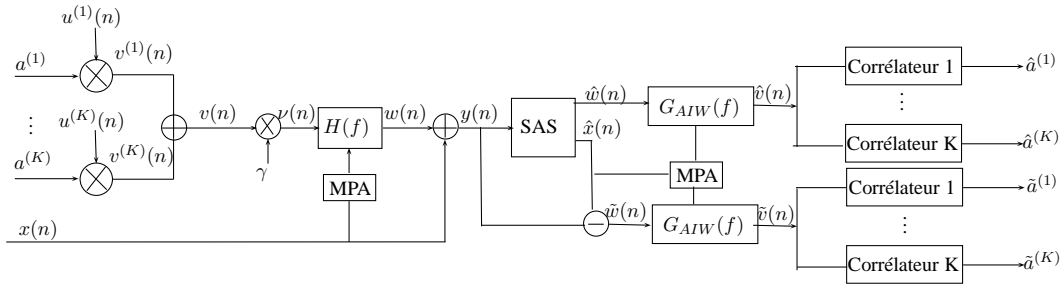


FIGURE 4.7 – Système multi-tatouage audio basé sur la séparation de sources.

- **Fiabilité de détection** : il est bien connu que la CDMA est la meilleure technique permettant à plusieurs utilisateurs de partager les canaux de communications avec moins d'interférences. Cela peut être clairement démontré quand un grand nombre d'utilisateurs partagent le même canal de communication. Dans ce cas, chaque utilisateur occupe toute la largeur de bande en même temps avec un code unique et les ressources du canal sont utilisées plus efficacement que la FDMA. Toutefois, en tatouage audio, le nombre de messages insérés est limité par la contrainte d'inaudibilité. Ainsi, nous avons montré que le maximum de tatouages qui peuvent être insérés est de dix. Par conséquent, en utilisant la FDMA, on peut introduire des intervalles de garde entre les bandes de fréquence pour réduire les interférences entre les sous-canaux.
- **Robustesse** : en utilisant la FDMA, et en choisissant des fréquences porteuses convenables à chaque perturbation, on améliore la robustesse du système multi-tatouage.

## 4.2 Système multi-tatouage à base de séparation de sources

Le système multi-tatouage présenté précédemment est performant et permet un débit d'insertion élevé. Cependant, ce débit est limité par la contrainte d'inaudibilité. On ne peut pas insérer autant de tatouages qu'on veut puisque la puissance du signal de tatouage augmente au fur et à mesure que le nombre de tatouage augmente. Pour cela, on se propose de contrôler l'inaudibilité du signal de tatouage avant son insertion. D'un autre côté, la faible puissance du signal de tatouage ainsi que sa non-stationnarité constituent un problème majeur aux approches classiques de filtrage au niveau de la détection [74]. Par conséquent, pour améliorer la robustesse des tatouages insérés vis-à-vis des perturbations externes et améliorer ainsi leur détection, les techniques de séparation de sources peuvent être un moyen efficace pour résoudre ce problème.

Malheureusement, la séparation de sources n'a pas connu une attention considérable en tatouage audio. Ainsi, nous limitons notre attention sur les techniques de séparation de sources aveugle afin d'améliorer l'efficacité de la détection du tatouage audio. Nous nous concentrons sur la conception d'un nouveau système qui considère le problème de détection du tatouage comme un problème de séparation de sources entre le signal audio et le signal de tatouage. En s'inspirant du concept de séparation de sources, on montrera également comment transiter de la détection aveugle vers celle non aveugle. Le système multi-tatouage audio proposé est illustré sur la FIGURE 4.7.

### 4.2.1 Contrôle d'inaudibilité

Pour garantir la contrainte d'inaudibilité, la puissance du signal de tatouage inséré est généralement contrôlée par un modèle psychoacoustique appliqué au signal audio  $x(n)$ . En effet, la conception du filtre de mise en forme  $H(f)$  impose de choisir un signal modulé  $\nu(n)$  dont la variance satisfait la contrainte suivante [9] :

$$\sigma_\nu^2 = 1 \quad (4.14)$$

Le modèle psychoacoustique est calculé à partir de la DSP  $S_x(f)$  du signal audio et peut être adapté au contexte du tatouage en utilisant un facteur d'échelle  $\gamma$  pour contrôler l'inaudibilité. Par conséquent, le signal modulé  $\nu(n)$  définie par l'équation (4.4) est multiplié par un facteur d'échelle  $\gamma$  avant de passer via le filtre de mise en forme  $H(f)$ . Enfin, le signal audio tatoué  $y(n)$  devient :

$$\begin{aligned} y(n) &= x(n) + w(n) \\ &= x(n) + \nu(n) * h(n) \\ &= x(n) + \gamma \nu(n) * h(n) \end{aligned} \quad (4.15)$$

où  $h(n)$  est la réponse impulsionnelle du filtre de mise en forme  $H(f)$ . Afin d'assurer la contrainte d'inaudibilité du signal de tatouage  $w(n)$ , on devrait ajuster le facteur d'échelle  $\gamma$  en calculant la variance du signal modulé  $\nu(n)$  :

$$\begin{aligned} Var(\nu(n)) &= Var(\gamma \nu(n)) \\ &= Var(\gamma \sum_{k=1}^K v^{(k)}(n)) \\ &= \gamma^2 Var(K v^{(k)}(n)) \\ &= \gamma^2 K^2 \end{aligned} \quad (4.16)$$

Pour satisfaire la contrainte d'inaudibilité (4.14), on doit choisir :

$$\gamma = \sqrt{\frac{1}{K}} \quad (4.17)$$

Dans ce cas, le nombre de tatouage  $K$  n'aura aucun impact sur la contrainte d'inaudibilité. La FIGURE 4.8 illustre les valeurs des ODGs obtenues par l'utilisation des techniques d'accès multiples après le contrôle d'inaudibilité. On remarque que la transparence des tatouages reste constante et elle est indépendante du nombre de tatouages insérés. De cette manière le débit d'insertion peut être amélioré autant de fois qu'on veut sans aucun impact sur la qualité du signal audio. Néanmoins, le débit du système multi-tatouage proposé reste toujours limité par l'interférence d'accès multiple qui se produit entre les différents tatouages.

### 4.2.2 Séparation de sources

Un des problèmes majeurs des systèmes multi-tatouage est la réalisation de la détection aveugle des différentes données insérées à partir du signal audio tatoué seul. Il serait donc nécessaire de concevoir un mécanisme d'extraction robuste qui permettra de récupérer toutes les informations cachées. La plupart des méthodes d'extraction de tatouage utilisent des techniques d'égalisation pour extraire les bits du tatouage [74]. Malgré la simplicité de leurs récepteurs, ils ont encore des limitations en terme de fiabilité de détection.

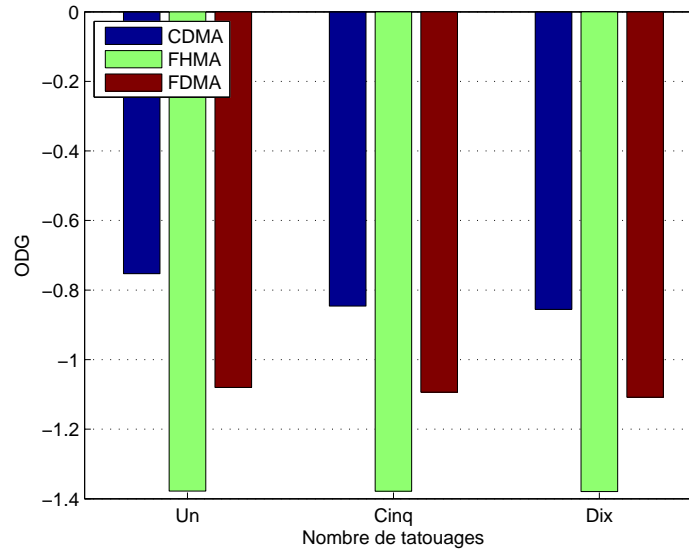


FIGURE 4.8 – Évaluation de la transparence des tatouages par le contrôle d’inaudibilité.

#### 4.2.2.1 Séparation de sources UISA

Parallèlement aux systèmes de tatouage audio standards, il est possible de modéliser le processus d’extraction du tatouage audio comme un modèle de séparation de sources aveugle [50]. Le signal tatoué  $y(n)$  est considéré comme un mélange linéaire entre le signal audio  $x(n)$  et le signal de tatouage  $w(n)$ . Néanmoins, la question qui se pose est : quelle méthode de séparation de sources doit-on choisir ? Les techniques classiques de séparation de sources (SAS) visent à estimer les  $\aleph$  signaux sources d’origine à partir des  $\hbar$  observations de leur mélange [4]. Dans notre cas, nous disposons de deux sources ( $x(n)$  et  $w(n)$ ) et une seule observation (le signal tatoué  $y(n)$ ). Quand on dispose de plus de sources que d’observations, le problème de séparation de sources devient beaucoup plus difficile. Pour surmonter cette difficulté, l’idée est d’utiliser la décomposition EMD<sup>11</sup> [59], afin d’exploiter un grand nombre d’IMFs<sup>12</sup> qui nous permettra de transiter d’un système sous-déterminé à un autre sur-déterminé<sup>13</sup> [49].

Comme les signaux audio sont souvent non-stationnaires, alors l’EMD est utilisé ici comme un outil approprié pour leur décomposition. En effet, l’EMD utilise un seul mélange pour extraire les composants IMFs [59]. Ainsi, le signal résultant mixé  $y(n)$  est décomposé par la technique EMD qui produit les signaux IMFs. Cela nous permet de passer d’une observation  $y(n)$  pour un ensemble d’observations qui sont les IMF du signal tatoué  $y(n)$ . Ensuite, la matrice temporelle résultante des IMFs est utilisée pour calculer un ensemble de vecteurs indépendants en appliquant la technique ISA<sup>14</sup> [18]. La technique résultante est connue sous le nom de UISA<sup>15</sup> ; l’algorithme de cette technique est le suivant [50] :

---

11. Empirical Mode Decomposition.

12. Intrinsic Mode Functions.

13. Il y a maintenant plus de capteurs au lieu d’un représenté par le signal tatoué.

14. Independent Subspace Analysis.

15. Under-determined Independent Subspace Analysis.

---

ALGORITHME 3 – Algorithme de séparation de sources UISA.

---

1. Initialement, le signal tatoué  $y(n)$  est convertit à des IMFs en appliquant la décomposition EMD, qui construit la matrice d'observation  $\mathbf{IMF}$  :

$$\mathbf{IMF} = [IMF_1(n), IMF_2(n), \dots, IMF_S(n)] \quad (4.18)$$

où  $S$  est le nombre des *IMFs*. La dimension de  $IMF_m(n)$  est  $1 \times L$  où  $L$  est la taille des données.

2. Appliquer l'analyse en composante principale en employant la SVD <sup>a</sup> [18] selon l'équation suivante :

$$\mathbf{IMF} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (4.19)$$

où  $\mathbf{U}$  and  $\mathbf{V}$  sont des matrices orthogonales avec des colonnes orthogonales de dimensions respectives  $L \times L$  et  $S \times S$ , et  $\mathbf{\Sigma}$  est la matrice qui contient  $S$  valeurs singulières.

3. Utiliser un ensemble réduit de  $\rho$  vecteurs de bases sélectionnées à partir de  $\mathbf{U}$ , i.e  $\mathbf{U}_{L \times \rho} = [\mathbf{u}_{L \times 1}, \mathbf{u}_{L \times 2}, \dots, \mathbf{u}_{L \times \rho}]$  en utilisant les premières  $\rho$  valeurs singulières définies par les informations contenues dans les valeurs singulières non nulles [111].
4. Appliquer l'analyse en composantes indépendantes pour exprimer le vecteur d'observation (ensemble de bases dérivées de la PCA)  $\mathbf{U}_{L \times \rho}$  comme le produit de la matrice de mélange  $\mathbf{Q}$  et du vecteur indépendant statistiquement  $\mathbf{s}$  :

$$\mathbf{U}_{L \times \rho} = \mathbf{Q}\mathbf{s} \quad (4.20)$$

où  $\mathbf{Q}$  est à  $\rho \times \hbar$  (pseudo-) matrice inversible avec des colonnes orthogonales,  $\mathbf{s}$  est un vecteur aléatoire de  $\hbar$  composantes de source.

5. Construire les sous-espaces indépendants utilisant les équations ci-dessous :

$$\mathbf{Y} = \mathbf{A}\mathbf{u} = \mathbf{A}\mathbf{Q}\mathbf{s} \quad (4.21)$$

$$\mathbf{Z} = \mathbf{v}\mathbf{\Sigma}\mathbf{A} \quad (4.22)$$

6. Construire la matrice des IMFs des sous espaces indépendants (sources) obtenues à partir de :

$$\mathbf{IMF}^j = y_j z_j^T \quad (4.23)$$

où  $y_j \in \mathbf{Y}$  et  $z_j \in \mathbf{Z}$ .

7. Implémenter un processus réversible de la décomposition EMD, de sorte que le signal temporel peut être calculé à partir de la simple addition des *IMFs*, pour chaque matrice  $IMF_s, IMF_j$ , d'un sous-espace indépendant.

8. Le regroupement des composants de sources résultants de l'algorithme k-means.

---

<sup>a</sup>. Singular Value Decomposition

"Tous les Hommes font la même erreur, de s'imaginer que bonheur veut dire que tous les vœux se réalisent. "  
*Léon Tolsto*

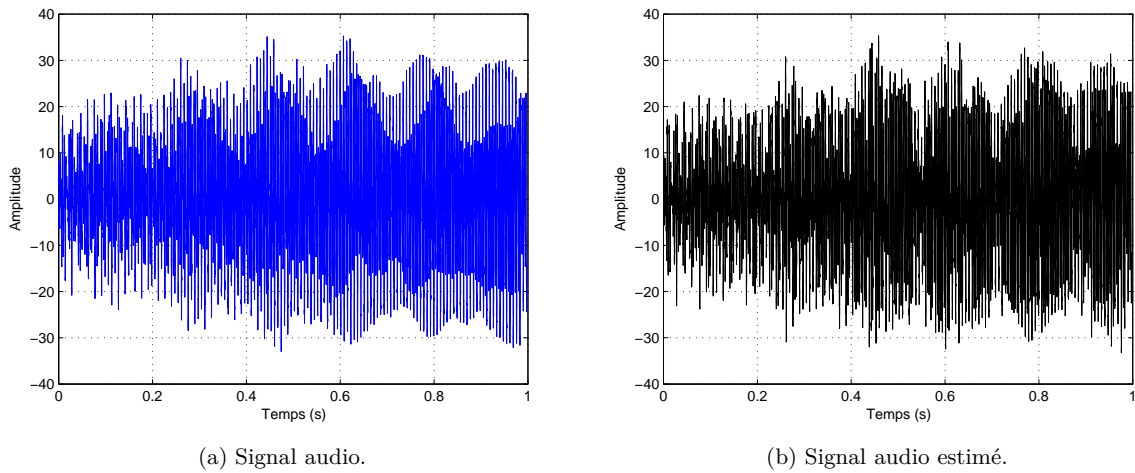


FIGURE 4.9 – Représentation du signal audio et du signal audio estimé dans le domaine temporel.

### 4.2.3 De la détection aveugle vers la non-aveugle

La séparation de sources aveugle permet l'extraction de deux signaux ; le tatouage estimé  $\hat{w}(n)$  et le signal audio estimé  $\hat{x}(n)$ . Un exemple d'extraction de ces deux signaux est illustré sur la FIGURE 4.9. Le fait de récupérer une estimation du signal audio par la technique UISA, peut transformer la détection aveugle du tatouage en une détection non aveugle. En effet, en soustrayant le signal audio  $\hat{x}(n)$  du signal tatoué  $y(n)$ , on obtient une autre estimation du signal de tatouage  $\tilde{w}(n)$ . La FIGURE 4.10 compare les deux signaux de tatouage estimés avec le signal de tatouage original en utilisant la méthode directe<sup>16</sup> et la méthode indirecte<sup>17</sup>. On remarque que la méthode indirecte dans cet exemple extrait mieux le signal de tatouage. Ensuite, les signaux  $\hat{w}(n)$  et  $\tilde{w}(n)$  sont soumis à un égaliseur de forçage à zéro  $G_{AI}(f)$  suivi du filtre de Wiener  $G_W(f)$ . L'ensemble de ces deux filtres est représenté par  $G_{HW}(f)$ .

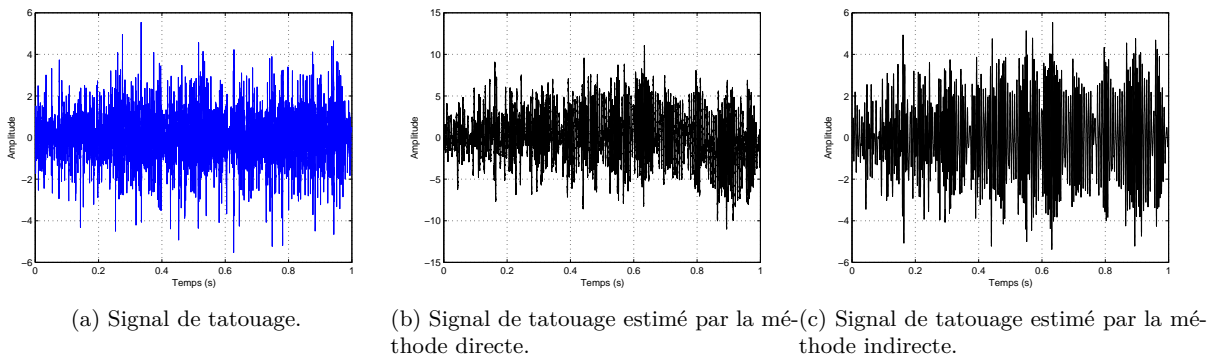


FIGURE 4.10 – Représentation dans le domaine temporel du signal de tatouage estimé par la méthode directe et indirecte.

16. Quand le tatouage est extrait directement à partir du signal tatoué.

17. Quand le tatouage est extrait à partir du signal audio estimé et celui tatoué.

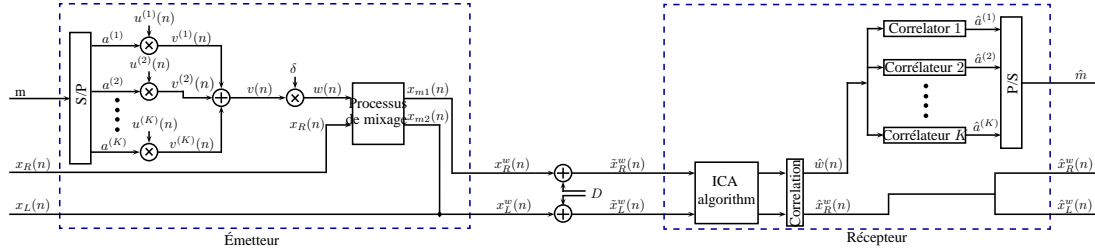


FIGURE 4.11 – Le système multi-tatouage pour les signaux audio à deux canaux.

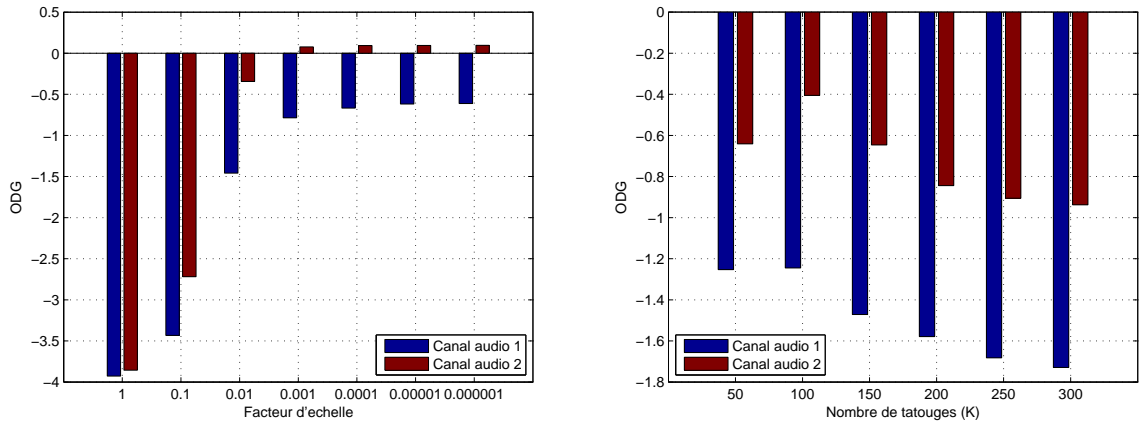
Maintenant que nous avons à notre disposition les signaux estimés  $\hat{v}(n)$  et  $\tilde{v}(n)$ , le processus de détection est effectué par les  $K$  corrélateurs, tout en supposant qu'on dispose d'une copie des formes d'ondes  $u^{(k)}(n)$  au niveau du récepteur. En effet, les signaux estimés  $\hat{v}(n)$  et  $\tilde{v}(n)$  sont soumis initialement à une étape de démodulation qui évalue la similarité entre les signaux estimés et les formes d'ondes  $u^{(k)}(n)$ . Enfin, une étape de décision détermine les messages binaires estimés  $\hat{a}_{k=1,\dots,K}^{(k)}$  et  $\tilde{a}_{k=1,\dots,K}^{(k)}$ . Les performances du système en terme de fiabilité d'extraction dépendent toujours de la technique de séparation utilisée au niveau du récepteur.

Notons que le processus de détection peut être amélioré en combinant les deux signaux de tatouage extraits  $\hat{w}(n)$  et  $\tilde{w}(n)$  en utilisant les techniques de fusion d'information [14]. Par exemple, on peut utiliser l'une des méthodes existantes [65] : les règles de combinaison, les réseaux de neurones, la machine à vecteurs de support, la théorie des probabilités, la théorie de l'évidence, la théorie des possibilités ou même la théorie des ensembles flous. Mais notre but ultime, dans ce travail, était de montrer que la méthode indirecte d'extraction peut mieux récupérer le tatouage inséré.

## 4.3 Tatouage multiple des signaux audio double canal

### 4.3.1 Signaux audio double-canal

Plusieurs signaux audio diffusés sur la radio ou encore en télévision ou même les CD de musique et les fichiers MP3, sont disponibles dans un format à deux canaux. On se propose donc, en utilisant l'algorithme ICA<sup>18</sup> comme outil d'extraction en considérant les deux canaux comme deux observations différentes<sup>19</sup>. Dans ce cas, nous surmontons le problème de la séparation de sources sous-déterminée rencontrée dans la partie précédente, en supposant que les canaux gauche et droite sont semblables. Néanmoins, l'ICA exige un nombre d'observations supérieur ou égal au nombre de sources indépendantes [24], ce qui signifie que nous pouvons insérer un seul signal de tatouage. Cette contrainte limite la capacité d'insertion du système multi-tatouage. Par conséquent, pour régler ce problème, on exploite la technique CDMA afin d'augmenter le débit d'insertion tout en maintenant le nombre de sources égal à deux (signal audio  $x(n)$  et signal de tatouage  $w(n)$ ). Le système multi-tatouage à double canal proposé est présenté dans la FIGURE 4.11.



(a) Valeurs ODG selon le facteur d'échelle  $\delta$ .

(b) Valeurs ODG selon le nombre de tatouages insérés.

FIGURE 4.12 – Évaluation de qualité qui audio du signal tatoué double canal.

### 4.3.2 Insertion des tatouages dans les signaux audio à deux canaux

Le processus d'insertion est effectué par la matrice de mélange  $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ . En considérant un des deux canaux audio  $x_j(n)$  ( $j = R$  ou  $L$  pour le canal droit et le canal gauche) et le signal tatoué  $w(n)$  comme deux signaux indépendants de même taille  $\frac{N}{K}N_s$ , la matrice  $\mathbf{A}$  est utilisée pour obtenir deux observations linéaires  $x_{m1}(n)$  et  $x_{m2}(n)$  comme suit :

$$\begin{aligned} \begin{bmatrix} a_{11}x_j(n) + a_{12}w(n) \\ a_{21}x_j(n) + a_{22}w(n) \end{bmatrix} &= \begin{bmatrix} x_{m1}(n) \\ x_{m2}(n) \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_j(n) \\ w(n) \end{bmatrix} \end{aligned} \quad (4.24)$$

Ensuite, les deux canaux du signal tatoué  $x_R^w(n)$  et  $x_L^w(n)$  seront obtenus en prenant :

$$\begin{aligned} x_R^w(n) &= x_{m1}(n) \\ x_L^w(n) &= x_{m2}(n) \end{aligned} \quad (4.25)$$

Puisque les deux canaux du signal audio original sont quasiment identiques<sup>20</sup>, la contrainte d'inaudibilité sera assurée même si un des canaux du signal tatoué est obtenu en le remplaçant simplement par l'un des mélanges (produit à la sortie du processus de mixage). De plus, les coefficients  $a_{ij}$  de la matrice sont choisis de manière à créer une légère différence entre  $x_{m1}(n)$  et  $x_{m2}(n)$ .

Pour évaluer l'impact de la puissance du tatouage sur la contrainte d'inaudibilité, on fixe le nombre de tatouages  $K = 1$  et on calcule ensuite les valeurs ODGs selon le facteur d'échelle  $\delta$ .

18. Independent Component Analysis.

19. Appelé également mélanges.

20. Contrairement aux signaux audio stéréophoniques où il y a une différence au niveaux d'amplitude et un retard de phase entre les deux canaux.

La FIGURE 4.12(a) montre que la qualité audio du signal tatoué s'améliore lorsque la valeur  $\delta$  diminue et les valeurs ODGs des deux canaux tatoués ne sont pas similaires. Ceci est dû à la faible différence entre les coefficients de la matrice de mélange ( $a_{11} = 0.9$ ,  $a_{12} = 0.1$  et  $a_{21} = 0.95$ ,  $a_{22} = 0.05$ ) utilisée pour construire les deux canaux du signal tatoué  $x_R^w(n)$  et  $x_L^w(n)$ . On fixe alors  $\delta = 10^{-3}$ .

On peut améliorer le débit d'insertion en variant trois paramètres à savoir : le nombre de bits par symboles  $q$ , le nombre de tatouages insérés  $K$  et le temps symbole discret  $N_s$ . Pour une valeur donnée de  $q$  qu'on fixe à 1, il faut augmenter la valeur de  $K$  et diminuer celle de  $N_s$  au minimum. Malheureusement, ces paramètres doivent être ajustés afin d'assurer un compromis entre l'inaudibilité et la fiabilité de détection des tatouages. Dans la FIGURE 4.12(b), on présente l'évolution des valeurs ODGs selon le nombre de tatouages insérés. On remarque qu'en augmentant la valeur de  $K$ , la qualité audio du signal tatoué se dégrade du fait qu'augmenter  $K$  est équivalent à augmenter la puissance du signal de tatouage.

### 4.3.3 Extraction des tatouages à partir des signaux audio à deux canaux

Pour extraire le signal de tatouage  $w(n)$ , on se propose d'utiliser l'algorithme ICA [24]. Considéré comme l'un des algorithmes les plus utilisés en séparation de sources, l'algorithme ICA vise à récupérer quelques signaux sources à partir de leurs mélanges linéaires instantanés. Cette approche est basée sur l'hypothèse que les signaux sources sont statistiquement indépendants [24], ce qui signifie que la densité de probabilité  $p(y_1, y_2, \dots, y_n)$  des variables aléatoires  $y_1, y_2, \dots, y_n$  peut se factoriser comme suit :

$$p(y_1, y_2, \dots, y_n) = p_1(y_1)p_2(y_2) \dots p_n(y_n) \quad (4.26)$$

où  $p_i(y_i)$  est la densité de probabilité marginale de  $y_i$ . Dans notre cas, cette hypothèse est assurément vérifiée puisque la connaissance des  $x_j(n)$  ne donne aucune information sur  $w(n)$  et vice versa. Ainsi, la récupération de la source (ou l'extraction du tatouage dans notre contexte) peut être obtenue par l'estimation d'une matrice de démixage  $\mathbf{B}$  qui transforme linéairement les signaux d'observation<sup>21</sup> et les transforment en des sources qui sont aussi indépendantes que possible [24] :

$$\begin{bmatrix} \hat{x}_j(n) \\ \hat{w}(n) \end{bmatrix} = \mathbf{B} \cdot \begin{bmatrix} \tilde{x}_R^w(n) \\ \tilde{x}_L^w(n) \end{bmatrix} \quad (4.27)$$

où  $\hat{x}_j(n)$  pour  $j = R$  ou  $L$  est une estimation du canal audio utilisée dans le processus de mixage et  $\hat{w}(n)$  est une estimation du signal de tatouage. Maintenant, pour récupérer le message binaire original  $m$ , on doit différencier entre  $\hat{x}_j(n)$  et  $\hat{w}(n)$  parce que le modèle ICA ne peut pas déterminer l'ordre des sources séparées [24]. Les deux sources estimées sont ensuite corrélées avec les codes d'étalement  $u^{(k)}(n)$ . La source estimée qui donnera la corrélation la plus élevée avec  $u^{(k)}(n)$  sera le signal de tatouage  $\hat{w}(n)$ . Enfin, la phase de démodulation est effectuée avec  $K$  détecteurs de corrélations en supposant la connaissance des codes d'étalement  $u^{(k)}(n)$  au niveau du récepteur.

Le paramètre  $K$  a un impact sur la fiabilité d'extraction, car il est la source de l'interférence d'accès multiples entre les symboles parallèles  $a^{(k)}$ . Par conséquent, on fixe  $K$  à 100 puis on augmente plus la capacité d'insertion en diminuant  $N_s$ . Même si  $N_s$  n'a aucune influence sur la qualité du signal audio tatoué, il a une certaine limite pour assurer une fiabilité de détection

---

21.  $\tilde{x}_R^w(n)$  and  $\tilde{x}_L^w(n)$  sont les deux canaux du signal tatoué après avoir subi la distorsion  $D$ .

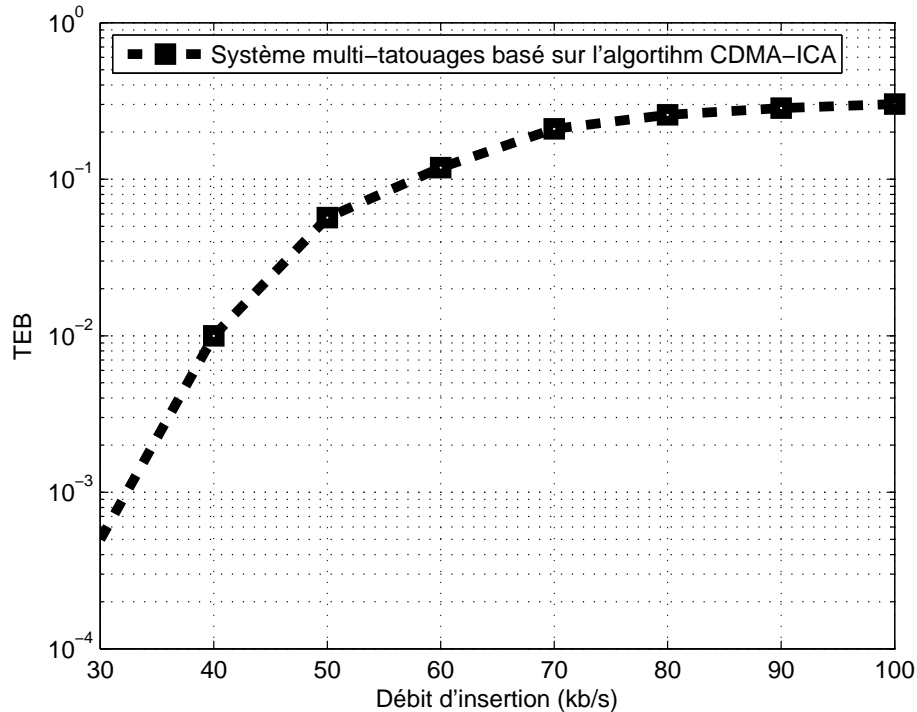


FIGURE 4.13 – Fiabilité de détection selon le débit d'insertion.

acceptable. En diminuant la valeur de  $N_s$ , l'interférence entre symboles produit entre les  $a_i^{(k)}$  augmente et par conséquent la détection des tatouages devient plus difficile. La FIGURE 4.13 montre qu'en tatouant les signaux audio double-canal, on atteint un débit d'insertion de 100kb/s mais avec moins de performances en termes de TEB. En effet, le TEB se dégrade lorsque  $N_s$  diminue jusqu'à atteindre un TEB égale à  $10^{-1}$  pour un débit de  $R \geq 60$ kb/s. Il faudrait alors diminuer le débit jusqu'à  $R < 60$ kb/s pour assurer une performance de l'ordre de  $10^{-2}$ .

## 4.4 Discussions

### 4.4.1 Performances des différents systèmes de tatouage proposés

Nous avons présenté dans cette thèse plusieurs systèmes de tatouage assurant chacun un degré de performances acceptable. Mais quel est le système qui présente les meilleures performances ? Pour répondre à cette question, il serait nécessaire de présenter une comparaison entre les performances de tous les systèmes de tatouage proposés. Chacun de ces systèmes présente des avantages et des inconvénients. En effet, il y a des systèmes qui assurent la transparence totale du tatouage mais présentent un débit d'insertion moyen. D'un autre côté, il y a des systèmes qui permettent l'insertion d'une grande quantité d'information mais qui sont fragiles vis-à-vis des perturbations externes.

Les systèmes de tatouage qu'on va comparer sont les suivants :

- **(P)** : Système de tatouage par étalement de spectre présenté dans le chapitre **2** et qui fait la base de plusieurs systèmes présentés dans l'état de l'art [8, 15, 28, 31, 50, 69, 73, 94, 103, 109, 117, 119].
- **(A)** : Système de tatouage par déplacement de phase informé présenté dans **3.3**.
- **(B)** : Système de tatouage multi-porteuses présenté dans **3.4**
- **(C)** : Système de tatouage multiple présenté dans **4.1**.
- **(D)** : Système de tatouage multiple basé sur la séparation de sources présenté dans **4.2**.
- **(E)** : Système de tatouage multiple pour les signaux à double canal présenté dans **4.3**.

Tous ces systèmes seront comparés selon les critères suivants :

- Inaudibilité.
- Débit d'insertion.
- Fiabilité de détection.
- Robustesse contre le bruit blanc gaussien et la compression.
- Coût de calcul.

Le tableau 4.6 illustre une première comparaison entre ces systèmes de tatouage présentés ci-dessus. Le système de tatouage audio de référence **(P)** offre un débit de 500b/s avec un TEB de l'ordre de  $10^{-1}$  tout en assurant l'inaudibilité et une exécution en temps réel. Le premier système proposé dans cette thèse **(A)** permet d'améliorer l'inaudibilité ainsi que la fiabilité de détection en insérant le tatouage avec le même débit du système **(P)**. L'inconvénient de ce système est qu'il est trop lent en calcul ce qui rend son utilisation difficile dans les applications fonctionnant en temps réel. Avec le même débit d'insertion, le système **(B)** améliore encore plus la fiabilité de détection des deux systèmes **(P)** et **(A)** et permet une exécution en temps réel tout en assurant le même degré d'inaudibilité du système **(P)**.

Les limitations communes des systèmes précédents est le débit d'insertion qui est moyen. Le système **(C)** permet d'augmenter le débit d'insertion jusqu'à 4kb/s mais en contre partie ne garantit pas la transparence totale du tatouage. Sa fiabilité de détection est acceptable et est équivalente à celle du système **(B)**. Ce système peut être exploité aussi dans les applications opérant en temps réel. En augmentant encore le débit jusqu'à 10kb/s tout en assurant la contrainte d'inaudibilité, le système **(D)** permet l'extraction du tatouage avec un TEB meilleur que celui offert par le système **(C)**. Malheureusement, ce système est lent en calcul à cause de la technique de séparation de sources UISA. Enfin, le système **(E)** offre un débit plus élevé avec une bonne fiabilité de détection, une transparence acceptable et une exécution en temps réel. L'inconvénient majeur de ce système est qu'il peut s'appliquer juste pour des signaux à double canal.

*"Il n'existe que deux choses infinies, l'univers et la bêtise humaine... mais pour l'univers, je n'ai pas de certitude absolue". Albert Einstein*

---

Systèmes de tatouage	Critères			
	Débit (b/s)	Inaudibilité (ODG)	Détection (TEB)	Temps réel
(P)	500	-0.7038	0.1254	Oui
(A)	500	-0.2573	0.0841	Non
(B)	500	-0.7425	0.0432	Oui
(C)	4000	-1.6958	0.0899	Oui
(D)	10000	-0.9208	0.0679	Non
(E)	40000	-1.124	0.0131	Oui

Tableau 4.6 – Comparaisons entre les système de tatouage audio en termes de débit, inaudibilité, fiabilité de détection et coût de calcul.

D'un autre côté, le tableau 4.7 compare les différents systèmes de tatouage en terme de robustesse. Les systèmes (P), (A) et (B) offre une robustesse un peu élevée contre le bruit blanc gaussien mais en se limitant à un débit de 100b/s. Un inconvénient majeur des systèmes proposés est leur fragilité contre la compression MPEG. En effet, aucun des systèmes n'est robuste contre ce type de perturbations sauf le système (B) à un degré près.

Systèmes de tatouage	Robustesse		
	Débit (b/s)	BBAG 10dB (TEB)	Compression 64kb/s (TEB)
(P)	100	0.0411	0.3951
(A)	100	0.0238	0.3911
(B)	100	0.0184	0.1982
(C)	4000	0.1145	0.4671
(D)	10000	0.1054	0.4125
(E)	40000	0.2414	0.4567

Tableau 4.7 – Comparaisons entre les système de tatouage audio en termes de robustesse.

Enfin, on peut dire que le système multi-tatouage (C) est celui qui assure le meilleur compromis entre toutes les contraintes imposées. Malheureusement, la seule contrainte qui est moins assurée par ce système est sa robustesse face à la compression MPEG. De plus, la transparence du tatouage dans ce système se dégrade lorsqu'on augmente la quantité d'information insérée. Par conséquent, on se propose de compléter les performances de ce système en ajoutant :

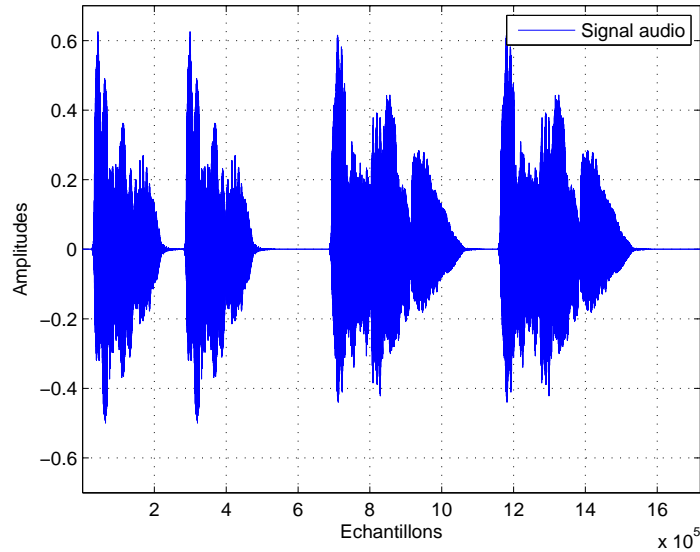


FIGURE 4.14 – Signal Audio : Saxophone.

- Les codes cycliques qui ont été utilisés dans le système **(B)** afin d'améliorer la robustesse.
- Le paramètre  $\gamma$  présenté dans le système **(D)** pour assurer la transparence du tatouage indépendamment de la quantité d'informations insérée.

Le système de tatouage résultant sera appelé système **(CBD)**.

#### 4.4.2 Insertion d'une image numérique

Pour mettre à l'épreuve le nouveau système multi-tatouage **(CBD)**, nous présentons une procédure qui vise une application réelle. L'application visée ne consiste pas, cette fois-ci, à insérer des messages textes mais vise à insérer une image numérique contenant un logo dans un signal audio. Le signal audio utilisé est le Saxophone illustré à la FIGURE 4.14. Le signal tatoué est soumis à un ensemble de perturbations présentées sous formes d'opérations de traitement du signal.

Le but est de récupérer l'image insérée à partir du signal tatoué perturbé avec le minimum d'erreurs possible. On se propose d'utiliser deux types d'images. La FIGURE 4.15(a) présente la première image qui est une image binaire de dimensions  $180 \times 100$ . La FIGURE 4.15(b) la deuxième image qui est une image en niveaux de gris de dimensions  $150 \times 100$ . Le nombre de tatouage est  $K = 10$  et le temps symbole est  $N_s = 44$ . Le débit d'insertion résultant  $R = 10\text{kb/s}$  et le tatouage inséré est imperceptible avec un  $ODG = -0.91$ .

La FIGURE 4.16 et la FIGURE 4.17 résument les résultats de la détection du tatouage inséré lorsqu'il est soumis aux perturbations externes. On remarque à partir des résultats obtenus que les images sont récupérées avec une qualité visuelle acceptable dans le cas des perturbations de type BBAG, quantification, amplification et rééchantillonnage. D'un autre coté, la qualité devient moyenne lorsque le signal tatoué est soumis à un filtrage passe-bas. Enfin, la qualité visuelle

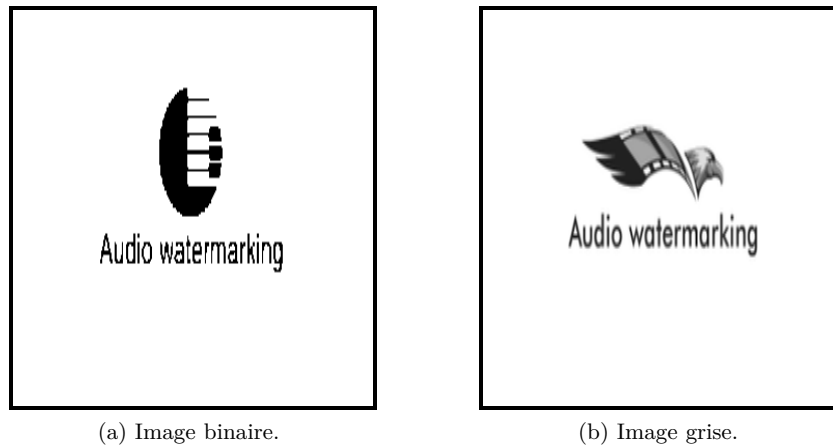


FIGURE 4.15 – Le tatouage sous forme d'image.

devient médiocre dans le cas de la compression MPEG. Le tableau 4.8 illustre les performances d'extraction pour chaque perturbation qui est mesurée par les valeurs TEB et PSNR<sup>22</sup>. Nous remarquons que les images binaires et celles en niveaux de gris sont extraites avec des valeurs de TEB similaires. Cependant, l'image en niveaux de gris a été extraite avec une meilleure qualité visuelle (PSNR) si on la compare avec l'image binaire.

Perturbations	Image binaire		Image grise	
	TEB	PSNR (dB)	TEB	PSNR (dB)
Filtre passe-bas	0.1199	4.5280	0.1191	36.5842
Filtre passe-haut	0.0486	6.4063	0.0404	38.2707
Bruit blanc	0.0188	12.2806	0.0112	37.4378
Quantification	0.0178	10.4531	0.0102	38.5560
Amplification	0.0169	5.5517	0.0191	36.2649
Rééchantillonnage	0.0183	5.4551	0.0204	36.2937
Compression MPEG 128kb/s	0.2017	1.5018	0.2119	35.3256
Compression MPEG 96kb/s	0.2284	-1.0068	0.2305	33.8008
Compression MPEG 64kb/s	0.2576	5.5517	0.2514	38.8699

Tableau 4.8 – Les valeurs PSNR t TEB des logos insérés.

---

22. PSNR (sigle de Peak Signal to Noise Ratio) est une mesure de distorsion utilisée en image numérique pour mesurer le degré de similarité entre deux images.



(a) Bruit blanc gaussien.

(b) Quantification.

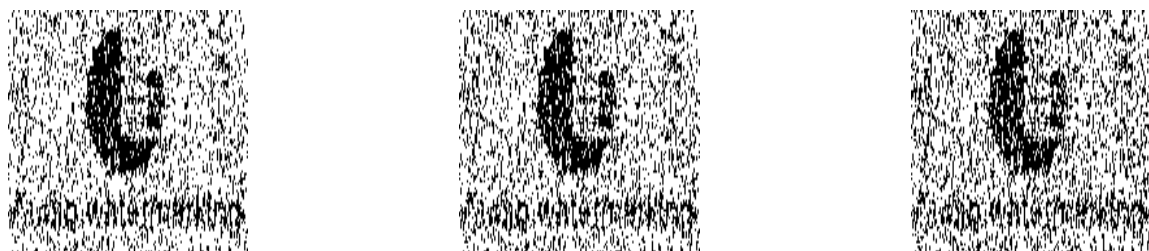
(c) Amplification.



(d) Rééchantillonnage.

(e) Filtre passe-haut.

(f) Filtre passe-bas.



(g) Compression 128kb/s.

(h) Compression 96 kb/s.

(i) Compression 64 kb/s.

FIGURE 4.16 – L'extraction de l'image binaire à partir du signal tatoué soumis à plusieurs perturbations.

*"Pour atteindre la vérité, il faut une fois dans la vie se défaire de toutes les opinions qu'on a reçues, et reconstruire de nouveau tout le système de ses connaissances". René Descartes*

---



(a) Bruit blanc gaussien.



(b) Quantification.



(c) Amplification.



(d) Rééchantillonnage.



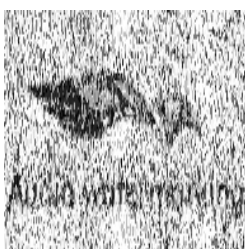
(e) Filtre passe-haut.



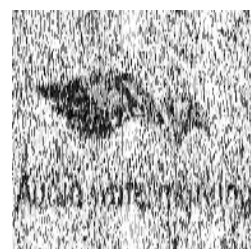
(f) Filtre passe-bas.



(g) Compression 128kb/s.



(h) Compression 96kb/s.



(i) Compression 64kb/s.

FIGURE 4.17 – L'extraction de l'image binaire à partir du signal tatoué soumis à plusieurs perturbations.

### 4.4.3 Comparaison avec les systèmes de télécommunications

Le système multi-tatouage (CBD) résultant offre un débit d'insertion élevé tout en garantissant la contrainte d'inaudibilité. De plus, il est robuste face à quelques perturbations externes. Ce nouveau système peut être exploité dans plusieurs applications en adaptant son utilisation selon la contrainte la plus exigée. Contrairement aux systèmes mono-tatouage présentés dans le chapitre 3 qui permettent l'insertion d'un seul message texte, le système multi-tatouage proposé peut permettre l'insertion de plusieurs messages textes en parallèle. De cette manière, au lieu de cacher une seule information, on peut incruster plusieurs messages contenant divers informations qui peuvent enrichir le contenu audio. D'un autre côté, si on veut augmenter la robustesse du tatouage inséré, on peut insérer plusieurs messages avec la même information afin de les utiliser toutes comme une aide à la décision et de la récupérer enfin avec le minimum d'erreurs possibles.

Comme les systèmes de tatouage audio présentent clairement des similitudes avec ceux de communications numériques [26], nous nous sommes inspirés des études déjà réalisées en télécommunication pour les adapter au contexte du tatouage audio. Néanmoins, les systèmes de tatouage présentent aussi des caractéristiques différentes du modèle standard d'un système de communication numérique, si on analyse les caractéristiques du bruit audio : coloré et très puissant. Par conséquent, les systèmes de tatouage audio présentent de faibles débits (plusieurs centaines voir des milliers de bits par seconde) et des taux d'erreur élevés (généralement de l'ordre de  $10^{-2}$ ), par opposition aux mégabits par seconde offerts par les systèmes de communication avec leur taux d'erreur faible (de l'ordre de  $10^{-6}$ ).

## Conclusion

Dans ce chapitre, nous nous sommes intéressés à la construction des systèmes multi-tatouage qui se basent essentiellement sur les techniques d'accès multiples connues en télécommunications. Ces systèmes ont pour objectif d'améliorer les débits d'insertion offerts par les systèmes de tatouage simples.

Tout d'abord, nous avons commencé par présenter le principe des techniques d'accès multiples. On s'est focalisé sur les techniques suivantes : CDMA, FHMA et FDMA. L'idée clef de l'utilisation de ces techniques d'accès multiples est la formulation du problème de tatouage en tant que problème de communication numérique. En communication numérique, les techniques d'accès multiples sont utilisées pour partager un canal de communication entre plusieurs utilisateurs. De même, nous avons proposé de partager simultanément le canal audio cette fois ci par plusieurs tatouages. La technique TDMA<sup>23</sup> n'a pas été prise en considération dans ce manuscrit car elle ne permet pas d'insérer simultanément les tatouages multiples dans les mêmes échantillons du signal audio mais permet plutôt d'insérer chaque information dans un ensemble d'échantillons audio indépendants. Ces systèmes multi-tatouage proposés étant limités par la contrainte d'inaudibilité, nous avons contrôlé la puissance du signal de tatouage afin de l'insérer de manière transparente. De plus, la détection du tatouage étant aveugle, nous avons proposé une nouvelle procédure basée sur la séparation de sources qui permet de récupérer le signal original afin de l'utiliser pour remonter au signal de tatouage par une simple soustraction avec le signal tatoué. En plus des performances des systèmes proposés qui ont été établies sur des signaux audio réels à un seul canal, nous avons présenté un autre système de tatouage qui est adapté aux signaux audio à double canal. Ce système a permis l'amélioration du débit d'in-

---

23. Time Division Multiple Access

*"Il suffit parfois d'aimer un peu moins ce qu'on aime pour éviter de le haïr". Jean Rostand*

---

sersion mais en contre-partie, il reste fragile vis-à-vis des perturbations externes. Enfin, nous avons présenté une comparaison entre tous les systèmes de tatouage présentés dans cette thèse en mettant l'accent sur les avantages et les inconvénients de chaque système.

## Chapitre 5

# Vers un nouveau service radio

La radio est l'un des médias les plus répandus comprenant des centaines de fournisseurs de programmes, des milliers d'émetteurs radio et des milliards de récepteurs dans le monde entier. Depuis le début de la diffusion radio dans les années 1920, le marché a été largement couvert par les services de radiodiffusion audio AM<sup>1</sup> et FM<sup>2</sup>. Dans ce chapitre, nous présentons un nouveau service qui consiste à utiliser le tatouage numérique des signaux audio pour la transmission radio. Notre but est d'exploiter l'afficheur numérique de la radio comme un petit écran sur lequel on peut afficher des informations qui ont été insérées dans l'émission diffusée avant sa transmission. Afin d'atteindre notre but, ce nouveau service interviendra à deux niveaux : au niveau de la station radio où l'insertion d'information est faite puis au niveau des clients (auditeurs de la chaîne radio) où les phases d'extraction et d'affichage d'information seront réalisées.

### 5.1 La radio

#### 5.1.1 Importance

Dans un monde de plus en plus dominé par les appareils de télécommunications, on peut se demander si la radio a encore sa place. Une récente étude de statistique<sup>3</sup> au Maroc indique que 61,3 % des marocains, de 11 ans et plus, écoutent quotidiennement la radio, soit 15 millions de personnes par semaine et près de 58,25 % le week-end ; soit 14,3 millions d'auditeurs. Le Marocain moyen écoute environ 3 heures (2h54) de radio par jour et près d'un tiers (27,6 %) écoutent la radio sur leur téléphone mobile. La radio, qui touche presque autant les urbains (60,6 %) que les ruraux (62,2 %), est principalement écoutée à la maison (72 %) et les Marocains l'écoutent généralement seuls (85 %). Que les auditeurs soient à la maison ou au travail, la radio reste toujours un élément essentiel et utilisée par plusieurs personnes dans le monde entier. La FIGURE 5.1 présente les chaînes radio diffusées sur le territoire Marocain.

---

1. Analog Modulation.

2. Frequency Modulation.

3. <http://www.lnt.ma/actualites/sondage-ipsos-radiometrie-medi1-et-2m-dans-le-top-5-les-privés-decrochent-38361.html>



FIGURE 5.1 – Stations radio au Maroc.

### 5.1.2 Histoire de la radio

La radio a été inventée après la découverte des ondes électromagnétiques et l'apparition du télégraphe<sup>4</sup>. Le physicien italien Guglielmo Marconi<sup>5</sup> est considéré comme le père des transmissions par radio ou TSF<sup>6</sup>, bien que ses inventions se sont inspirées des découvertes de nombreux prédécesseurs cités dans le tableau 5.1.

Année	Inventions et découvertes
1841	<b>Samuel Morse</b> invente le télégraphe
1876	<b>Alexander Graham bell</b> invente le téléphone <sup>7</sup> .
1888	<b>Hertz</b> met en évidence les ondes radio.
1889	<b>Tesla</b> réalise un générateur hautes fréquences (15 kHz).
1890	<b>Branly</b> découvre le principe de la radioconduction.
1891	<b>Tesla</b> brevette le système sans fil Tesla (radio télégraphe).
1893	<b>Popov</b> découvre le principe de l'antenne .
1895	<b>Marconi</b> expérimente les premières liaisons hertziennes.

Tableau 5.1 – Bref résumé de l'histoire de la radio

4. <http://www.histoire-pour-tous.fr/inventions/746-invention-de-la-radio.html>

5. Bien que Guglielmo Marconi ait d'abord été crédité de l'invention de la radio, et la plupart croient qu'il en est l'inventeur aujourd'hui encore, les brevets déposés par Nikola Tesla sur un système de radiodiffusion sans fils ont été utilisés par Marconi frauduleusement. La Cour suprême des États-unis a reconnu cet usage frauduleux des brevets de Tesla, le déclarant alors seul inventeur de la radiodiffusion sans fils. Tesla a démontré que les signaux radio sont juste une autre fréquence qui nécessitent un émetteur et un récepteur. Lors d'une présentation devant la National Electric Light Association, Tesla a démontré cette technologie. Bien que Tesla ait déposé deux brevets en 1897, en 1904, le US Patent Office a annulé sa décision, attribuant cette découverte à Marconi

6. Transmission sans fil.

En 1894, Guglielmo Marconi âgé de vingt ans, amorce des expériences dans le grenier de ses parents après avoir entendu parler des ondes hertziennes<sup>8</sup>. Il profite de plusieurs inventions et découvertes pour réaliser la première installation de radio : il utilise l'éclateur de Hertz comme émetteur, l'antenne de Popov, et le cohéreur de Branly comme récepteur. À force de ténacité, il augmente progressivement la portée des signaux qu'il émet jusqu'à atteindre une portée de plus de trois kilomètres en 1895.

Parti en Angleterre pour concrétiser son invention, il la perfectionne en utilisant un émetteur et un récepteur accordés sur une longueur d'onde particulière en faisant varier les caractéristiques électriques des antennes et des circuits. En décembre 1901 le premier signal radio transatlantique est reçu sur l'île de Terre-Neuve. Marconi a reçu le prix Nobel de physique en 1909 pour ses contributions au développement de la télégraphie sans fil. Ainsi, s'engage une exploitation commerciale de l'invention : la radio devient un outil incontournable de l'armée et un instrument de loisirs dans les foyers.

### 5.1.3 Évolution

La radiodiffusion est l'émission des signaux d'ondes électromagnétiques d'émetteurs de radios FM et les radios qui diffusent sur les ondes courtes, moyennes ou grandes destinées à être reçues directement par le public. Ce service peut comprendre plusieurs genres d'émission à l'exemple des émissions sonores, des émissions de télévision.

Par la suite, on s'intéressera qu'à la radiodiffusion pour la transmission des sons (la voix humaine et les signaux audio). Pour transmettre un signal radio, on a besoin de deux éléments essentiels :

- Une station de radio pour émettre des ondes électromagnétiques à l'aide d'un émetteur radio et d'une antenne.
- Un poste de radio ou récepteur radio pour recevoir les ondes radio et restituer les sons sur un haut-parleur.

Outre les transmissions radio analogiques, les systèmes de transmission de diffusion ont maintenant tendance à changer vers les transmissions numériques. En effet, la radio numérique couvre un spectre bien plus large depuis la fin des années 1990 puisque sa diffusion existe par satellite et par internet.

Contrairement à la radio analogique qui permet de transporter le signal électrique tel quel dans l'onde porteuse, la radio numérique numérise le signal et le compresse ensuite afin de le transmettre tout en optimisant la bande passante. Ce signal numérique peut être diffusé en direct ou encore en différé.

---

7. L'antériorité d'Antonio Meucci pour la l'invention du téléphone a été officiellement reconnue le 11 juin 2002 par la Chambre des représentants des États-Unis. Récemment le congrès de EEUU à octroyé la paternité du téléphone à Antonio Meucci.

8. [www.histoire-pour-tous.fr/inventions](http://www.histoire-pour-tous.fr/inventions)



(a) Post-CD.



(b) Montre numérique.



(c) Téléphones mobiles.



(d) Radio à afficheur numérique.

FIGURE 5.2 – Exemples d'utilisation de l'afficheur radio.

## 5.2 Exploitation de l'afficheur radio

### 5.2.1 Afficheur radio

La radio a pris une grande place dans notre vie quotidienne vue son rôle à éduquer, sensibiliser et informer ses auditeurs sur des sujets qui les concernent directement : santé, éducation, droits de l'homme, etc. Toutes ses activités sont assurées en exploitant seulement le sens de l'ouïe. Cependant, il y a un élément essentiel qui n'est pas bien exploité dans l'appareil radio : c'est l'afficheur. En effet, cet élément clef peut être exploité pour cibler les utilisateurs qui veulent lire des informations pertinentes en plus d'écouter la radio. L'afficheur numérique que l'on trouve dans plusieurs types d'appareils radio, tels que : le post-CD des voitures, les radios à afficheur numérique, les téléphones mobiles ainsi que les montres numériques ou les baladeurs Mp3 (c.f. FIGURE 5.2) :

- **Autoradio** : cet appareil est devenu un dispositif indispensable dans les voitures récentes. Ce composant contenant un afficheur numérique, peut être utilisé pour mettre le conducteur ainsi que les personnes qui l'accompagnent au courant des actualités tout en écoutant de la musique ou quelques informations sur la météo.
- **Poste de radio** : cet appareil que l'on retrouve presque dans toutes les maisons. Grâce à l'afficheur radio, l'utilisateur peut écouter la radio tout en restant au courant de toutes les nouvelles.
- **Téléphones mobiles** : tous les téléphones mobiles sont maintenant dotés de la radio et certains supportent même la technologie 3G qui permet de naviguer sur internet. Toutefois, dans certains endroits où le réseau 3G est faible ou avec des téléphones moins récents qui ne supportent pas la 3G, on a besoin de savoir ce qui se passe dans le monde. En exploitant l'afficheur radio, ce problème ne sera plus posé et on peut rester toujours informé.
- **Les montres numériques ou les baladeurs Mp3** : ce sont des dispositifs très légers que l'on peut porter partout. Équipés de la radio, ces appareils peuvent être utilisés pour

écouter des conseils médicaux lorsqu'on fait du sport et resté, en même temps, au courant de toutes les actualités.

L'idée d'exploiter l'afficheur radio n'est pas nouvelle puisqu'il existe plusieurs techniques permettant la transmission des données supplémentaires pour la diffusion du signal audio. Ces techniques sont présentes dans la radio analogique et la radio numérique.

### 5.2.2 Système RDS

La technique existante et la plus simple pour transmettre des services des données supplémentaires pour la diffusion est de les représenter sous forme de séquence de données à plusieurs tonalités au début de la transmission [56]. Dans ce cas, les techniques de modulation et de démodulation analogiques standards peuvent être exploitées pour créer un paquet de données de faible puissance qui sera transmis avant chaque message. En utilisant cette technique, il y a un bruit d'éclatement qui est nettement audible au début du message.

Un autre système connu sous le nom de 'Data-in-Voice' vise à diminuer l'audibilité de la séquence de tonalités multiples en réduisant la puissance de la bande spectrale [99]. En utilisant un filtre coupe-bande, le système supprime certaines parties du spectre vocal autour de 2 kHz. Afin de garantir complètement l'inaudibilité de la séquence de données, le système intègre un filtre au niveau des récepteurs.

Actuellement, le système de radiodiffusion le plus connu pour la transmission de données supplémentaires est le RDS<sup>9</sup> [60]. Ce service permet de transmettre des données numériques en parallèle des signaux audio de la radio FM<sup>10</sup> dans la gamme des très hautes fréquences (VHF) [55]. Notamment, le RDS permet l'écoute d'une station sans interruption lors d'un déplacement, en prenant en compte automatiquement le passage d'une fréquence à l'autre. Il fournit également une identification des stations par leurs noms, des signaux d'horloge, des messages textuels, des informations de commutation temporaire sur un canal de communication pendant l'émission d'un flash routier, et bien d'autres fonctionnalités supplémentaires. Un des problèmes techniques qui se pose lors de l'utilisation d'un tel système est l'utilisation d'une bande passante supplémentaire pour l'émetteur et le récepteur. En effet, les données numériques sont portées par la troisième harmonique du signal pilote à 19 kHz déjà présent dans le type de modulation de base utilisé en bande FM. Cette harmonique située autour de la fréquence sous porteuse 57 kHz, est injectée en tant que nouvelle composante du signal multiplexé de l'émission radio. Figurant dans la bande passante de l'émetteur et du récepteur, elle est complètement extérieure aux signaux audio, et donc n'introduit aucune distorsion audible comme c'est illustré sur la FIGURE 5.4.

D'autres standards de diffusion d'informations sur les stations de radio FM sont apparus après le RDS, notamment le DARC<sup>11</sup> et le DirectBand [55]. Ces deux standards offrent des débits bien supérieurs que celui du RDS, mais ils provoquent des interférences entre les stations.

On peut également citer le AMSS<sup>12</sup>; l'équivalent du RDS pour la radio AM. Le standard AMSS fournit des fonctionnalités équivalentes à la norme RDS. Cependant, l'AMSS permet la transmission des données avec un débit plus faible que celui du RDS [55].

---

9. Radio Data System.

10. La radio FM est destinée à être reçue plus précisément dans la bande 87.5-108 MHz (VHF - bande II) qui est utilisée.

11. Data Radio Channel.

12. Amplitude Modulation Signalling System.

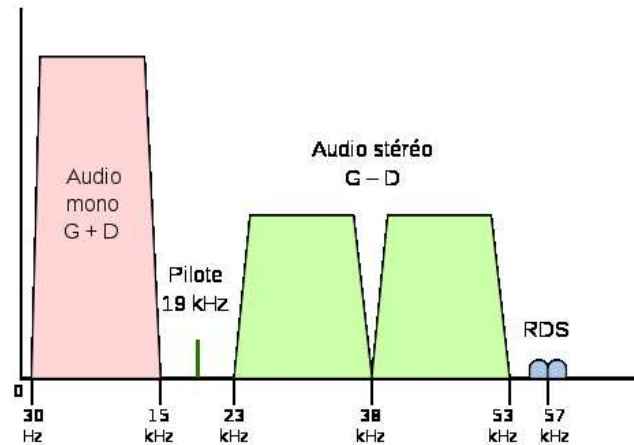


FIGURE 5.3 – Spectre d'une émission de radiodiffusion FM, avec stéréophonie et RDS [87].

### 5.2.3 Système DAB : diffusion audio numérique

Il existe plusieurs procédés de diffusion de la radio numérique. Le plus connu est le DAB<sup>13</sup> qui est un standard européen pour la radiodiffusion numérique qui fonctionne dans les bandes de fréquences VHF<sup>14</sup> (ondes ultra courtes) et dans les bandes UHF<sup>15</sup>.

L'idée principale de la création du DAB est que les ondes courtes, moyennes et longues peuvent offrir un certain nombre d'avantages par rapport aux autres systèmes de radiodiffusion<sup>16</sup> [55] :

- Couverture à très large échelle avec un seul émetteur.
- Qualité audio-numérique améliorée.
- Mise en Pause et Reprise du son.
- Les émetteurs d'une même stations peuvent être synchronisés sur la même fréquence.
- Recherche de stations sans connaître les fréquences pour chaque région.
- Données associées aux programmes : texte défilant, images, informations diverses, sites web.
- Transferts d'informations numériques diverses ; info-traffic, météo et vidéo.

Néanmoins, il y a toujours des problèmes, lors du remplacement d'une technologie existante par une autre [55] :

---

13. Digital Audio Broadcasting.

14. La bande des très hautes fréquences (very high frequency/VHF) est la partie du spectre radioélectrique s'étendant de 30 MHz à 300 MHz.

15. Bande des Ultra hautes fréquences (UHF) est la bande du spectre radioélectrique comprise entre 300 MHz et 3 000 MHz.

16. Satellite, VHF terrestre, etc.

- Absence de fréquences de transmission.
- Coût énorme pour le développement et l'investissement.
- Recherche de fournisseurs pour de nouveaux services non conventionnels (par exemple les services de données).
- Résolution du problème de la poule et de l'oeuf (qui sera le premier élément - le fournisseur de services ou le fabricant du récepteur ?).

## **5.3 Transfert d'informations supplémentaires au contenu radio par tatouage numérique**

### **5.3.1 Limitations des systèmes existants**

Le système DAB permet le transfert de données numériques supplémentaires à l'exemple du texte défilant et des images. Néanmoins, ce système présente une efficacité limitée en raison du manque de fréquences de transmission. De plus, le fait que tous les récepteurs doivent être équipés de cette technologie rend l'ensemble du système indésirable. En effet, même si la radio numérique existe dans de nombreux pays, la radio analogique continuera à être utilisée dans le monde entier et ce pour les nombreuses années à venir.

Pour un déploiement rapide, il serait intéressant alors d'éviter une modification majeure dans les équipements radios existants (des émetteurs-récepteurs supplémentaires, etc). De plus, en tenant compte de la conjecture actuelle de la bande passante et des fréquences pour la radio, des canaux supplémentaires de transmission numérique en bande passante exigeants ne sont pas réalisables dans un avenir proche. Il serait donc nécessaire de proposer un procédé qui permettra de remédier aux inconvénients précités.

### **5.3.2 But du procédé proposé**

Dans le cas présent, le tatouage est utilisé pour transmettre de l'information supplémentaire au contenu audio diffusé sur le canal radio. Le contenu audio fait référence à tout contenu se présentant sous forme d'un signal analogique ou numérique provenant d'un signal audio en diffusion sur un canal de radiodiffusion en temps réel. Grâce au tatouage, on peut insérer des méta-données décrivant le contenu tatoué avant sa transmission (tel que la description d'élément du contenu) qui peuvent être utilisées ultérieurement afin d'assurer un service à valeur ajoutée, ou bien encore des informations annexes (tel qu'un service de type télétexte ou des sous titres). Bien que la robustesse contre les attaques intentionnelles n'est pas nécessaire dans ce contexte, un certain degré de robustesse contre les perturbations licites comme la compression peut être souhaitable.

Le procédé de tatouage doit permettre un débit d'insertion élevé pour transférer différents types d'informations. Lors de la communication radio, les données supplémentaires doivent être intégrées dans le contenu audio. Ces données ne doivent pas perturber la communication et doivent toujours être transmises de façon fiable. Les contraintes propres à ce nouveau service seront donc un débit d'insertion (la quantité d'information insérée dans le contenu audio) élevé tout en garantissant la transparence du tatouage. De plus, le tatouage doit maintenir une certaine

robustesse contre les différentes perturbations licites subies lors de la transmission du signal radio tatoué dans le canal de diffusion.

La qualité du son d'un canal particulier (FM ou DAB) est importante pour déterminer le degré d'inaudibilité assuré par un tel système d'insertion de tatouage. Par exemple, le tatouage doit être incrusté de telle manière qu'il soit moins audible dans les systèmes DAB que dans les systèmes FM. Ceci est dû au fait que les systèmes analogiques FM assurent une qualité intrinsèque plus faible que les systèmes numériques DAB. Le procédé peut se caractériser dans ce cas par un contrôle de puissance du signal de tatouage n'introduisant aucune perturbation audible.

Ainsi, le but principal du service proposé est la transmission d'un certain type de données en tant qu'information supplémentaire au contenu audio original. En particulier, le contenu audio doit comprendre des informations qui ajoutent des fonctionnalités supplémentaires au niveau des récepteurs équipés en conséquence. De plus, l'ajout de l'aspect visuel à la radio doit être plus flexible et plus adapté aux normes radio existantes. Enfin, le procédé doit être à la fois simple à mettre en oeuvre tout en assurant des fonctionnalités plus performantes que les systèmes actuels.

### 5.3.3 Principe du procédé

Le procédé proposé est basé sur le concept du tatouage numérique qui est considéré comme une solution possible répondant à l'objectif visé tout en respectant les restrictions pré-citées. En particulier, le tatouage peut être utilisé pour fournir un service similaire aux systèmes RDS et DAB mais qui ne dépend pas d'une bande de transmission particulière. Le fait que l'information supplémentaire se situe à l'intérieur de la gamme de fréquences audible permet au signal de tatouage d'être transmis sur n'importe quel canal de transmission (analogique ou numérique) et peut tout simplement être détecté par l'ajout d'un extracteur de tatouage au niveau du récepteur.

Mais la question qui se pose est : où insérer le tatouage ? La réponse à cette question nécessite de savoir le programme de chaque station radio. En effet, le tatouage doit s'insérer dans la séquence audio directement avant sa transmission. La station radio doit disposer de l'application qui permet d'insérer le tatouage dans l'extrait audio qui sera diffusé. Puis, l'utilisateur doit disposer de l'extracteur du tatouage dans son équipement radio.

Le procédé permet d'exploiter l'afficheur radio pour fusionner l'effet visuel à celui de l'écoute. Pour réaliser ceci, un message texte portant des informations précises sera inséré en superposition à des documents audio classiques. Les étapes de réalisation sont illustrées sur la FIGURE 5.4. Le diffuseur de contenu (station radio par exemple) insère à l'intérieur de l'extrait audio un message portant des informations particulières. Pour réaliser ceci deux étapes sont nécessaires :

- La génération de la séquence binaire à partir du message texte. Cette tâche implique l'ajout d'un convertisseur qui permet de convertir un texte en un code binaire et vice versa. Ce convertisseur supporte bien les accents et les caractères spéciaux.
- L'insertion des paramètres binaires du message texte dans le flux audio. Pour que cette information ne dépende pas du format de diffusion (analogique, numérique) elle doit être insérée directement dans le contenu audio. Les outils de télécommunications et du tatouage constitueront le support technique pour la réalisation de cette tâche. C'est sur cette dernière tâche, à savoir, l'insertion de l'information sur le flux audio, que porte les systèmes de tatouages présentés dans cette thèse.

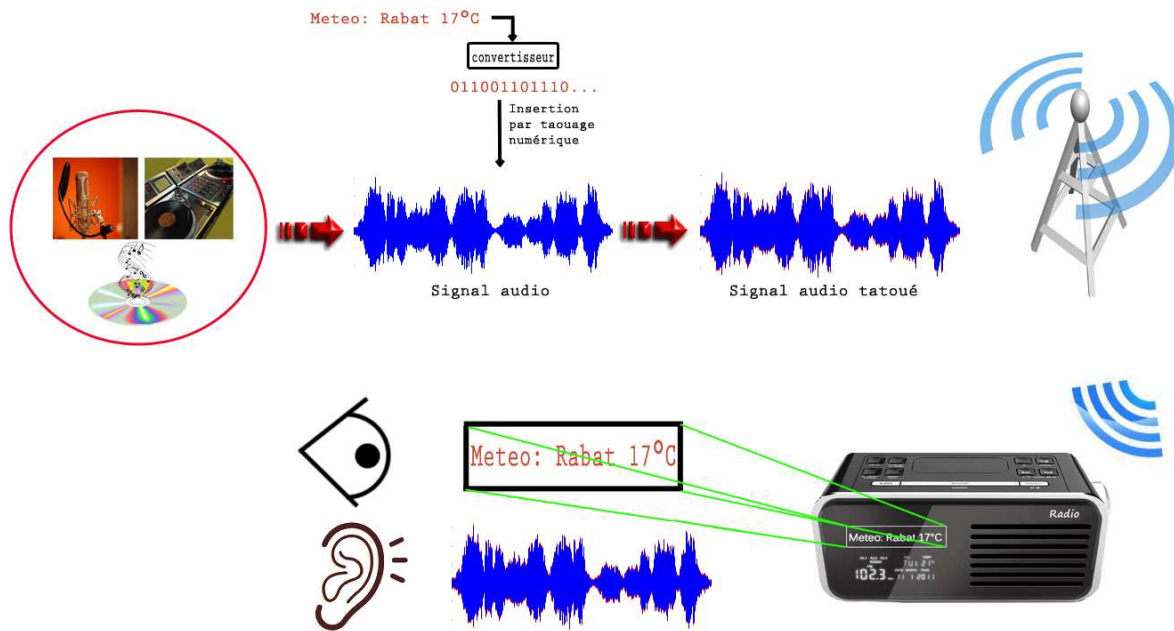


FIGURE 5.4 – Exemple d'utilisation du service radio proposé.

La deuxième partie se situe directement chez l'utilisateur, qui, mené d'un décodeur, restitue le message inséré. Cette partie est aussi décomposée en deux étapes :

- La détection et le décodage des paramètres du message binaire à partir du flux audio.
- La restitution du message texte sur l'afficheur radio. C'est la diffusion de ce message texte qui permet de rendre le programme diffusé plus riche en informations pour l'utilisateur. Il est intéressant de préciser que l'insertion d'une information annexe par tatouage permet à l'utilisateur de choisir la présence ou non du message texte.

## 5.4 Vers la finalisation du service radio

### 5.4.1 Intérêt social

Toutes les informations ou méta-données qui seront visualisées sur l'afficheur radio seront incluses dans le signal radio sous forme d'un tatouage. Ce dernier, tout en étant transparent peut contenir plusieurs types d'informations qu'on peut les classer en deux catégories :

- Flux constant lié au contenu.
- Flux variable portant des informations générales.

#### 5.4.1.1 Flux constant lié au contenu

Les informations liées au contenu sont des informations constantes qui peuvent, par exemple, être une description de l'extrait audio diffusé. Selon l'application et le profil de l'utilisateur, plusieurs niveaux de contenu d'informations peuvent être définis. Les utilisateurs ordinaires seraient intéressés par des informations générales sur une piste musicale, telles que le titre, le compositeur, l'étiquette et l'année d'édition. Les musiciens pourraient vouloir savoir quels instruments ont été joués, tandis que les ingénieurs du son pourraient être intéressés par des informations sur le processus d'enregistrement.

#### 5.4.1.2 Flux variable portant sur des informations générales

Le tatouage peut être également utilisé comme un moyen pour transmettre tout type d'information, pas nécessairement liée au contenu audio. Les stations de radio, par exemple, peuvent tatouer leurs extraits audio juste avant leur transmission, en intégrant du texte dedans. Par la suite, ce texte sera visualisé, par l'utilisateur, sur les afficheurs de leurs récepteurs radios. Diverses informations peuvent être transmises :

- Actualités et nouvelles.
- Prévisions météorologiques.
- Cotations boursières
- Nom de l'animateur radio.
- Titre de l'émission radio.
- Score d'un match de foot-ball.
- Énigmes et questions présentées dans une émission.
- etc.

#### 5.4.2 Intérêt économique

La radio est un média gratuit pour ses consommateurs finaux. En effet, sa principale source de revenus vient du fait qu'elle vend ses espaces publicitaires à des annonceurs. La publicité à la radio existe depuis quasiment sa création et a très rapidement financé les radios et a permis de développer leurs offres auprès des auditeurs. Au fil des années, différentes stratégies ont été développées pour tenter de présenter une publicité. La radio reste toujours un support publicitaire efficace. Ceci est dû au fait qu'elle est :

- **Omniprésente** : la radio est le média le plus proche de l'acte d'achat.
- **Vivante** : la radio est le média le plus souple et le plus réactif.
- **Économique** : la radio permet la répétition des messages à moindre coût avec un ticket d'entrée très attractif. Côté prix, même s'il n'y a pas de grille fixe, les professionnels

estiment<sup>17</sup> le prix moyen d'un affichage à 12.500 DH par mois. Le message radio coûte en moyenne 4.000 DH. Alors que pour une page dans la presse, il faut compter 30.000 DH.

La publicité radio est une partie importante de la programmation, occupant plusieurs minutes de chaque émission. Ceci est évident puisque la radio ne dispose que de la voix pour véhiculer le message. Pour cela, elle utilise très souvent des messages basés sur une promotion prix et répétés massivement pour atteindre le niveau de mémorisation souhaité. De plus, lorsqu'il y a trop de publicité, l'auditeur peut changer de station, et s'il y a moins d'auditeurs, donc il y a moins de performances et la demande des annonceurs se fait moindre.

Grâce au tatouage, on peut transmettre la publicité dans l'afficheur radio sans que ce dernier occupe aucun temps de la station radio. Un tel service aidera la station radio à avoir plus de ressources financières, sans aucun impact sur le nombre des auditeurs de la chaîne.

### 5.4.3 Verrous technologiques

Le service présenté dans ce chapitre offre un enjeu applicatif de grande envergure tout en soulevant des problèmes scientifiques extrêmement motivant. Les systèmes présentés dans cette thèse ne permettent pas d'accomplir le service proposé. Ainsi, les tâches et les enjeux scientifiques clés sont les suivants :

- L'information ajoutée doit être aussi importante que possible de manière à ce que son apparition sur l'afficheur soit aussi réaliste que possible. Ainsi, il faut concevoir des algorithmes de tatouage haute capacité.
- Le décodage de l'information dissimulée doit être possible quelque soit la chaîne de diffusion. Ainsi, le message doit être lisible après une compression du signal, une distorsion due à une réception de faible qualité ou encore un transcodage du flux audio.
- Les techniques de tatouage mises en oeuvre doivent également être de complexité très faible afin de pouvoir être utilisées pour un décodage en temps-réel. La complexité doit être inférieure à la complexité nécessaire au décodage de la source audio
- Les conditions de la transmission radio peuvent être considérées comme de nouveaux types de bruit qui entache le signal tatoué. Il serait nécessaire de développer de nouvelles techniques robustes aux perturbations liées à la transmission radio.
- La synchronisation est un problème majeur dans quelques applications du tatouage audio. Dans l'application proposée, il est connu que la transmission radio nécessite une bonne synchronisation. Par conséquent, il serait nécessaire de développer de nouvelles techniques de synchronisation.

## Conclusion

L'exploitation de l'afficheur radio numérique permettra d'ajouter un nouveau service audiovisuel à l'intention de ses clients qui consiste à adjoindre à des extraits audio un ensemble d'informations indélébiles et imperceptibles qui seront utiles à l'utilisateur.

---

17. <http://www.leconomiste.com/article/889229-investissements-publicitaires-br-radio-le-nouveau-support-fetich>

Les outils du développement pour réaliser l'application peuvent être considérés comme à mi-parcours. En effet, on peut dire que les méthodes nécessaires pour compléter ce service sont réalisées dans le cas idéal. Cependant, il nous reste à tester nos méthodes dans le cas réel, en ajoutant les conditions de la transmission radio. Ceci nécessitera l'étude des types de bruits auxquels va être soumis le tatouage. Et si nécessaire, proposer de nouvelles techniques plus robustes aux perturbations externes liées au nouveau service proposé.

Ce nouveau service et l'ensemble des techniques développées peuvent être encore améliorés en insérant plusieurs informations (de différents genres : paroles, actualités, météo) dans le même extrait audio. Dans ce cas, en utilisant l'extracteur du tatouage qui sera disponible chez le client, le choix du type d'informations pourra se faire selon les besoin de l'utilisateur. Avec les nouveaux algorithmes proposés dans les chapitre **3** et **4**, nous pouvons aussi incruster les paroles d'une chanson dans les CD audio. Une fois la musique se lance, les paroles s'affichent sur la radio qui peuvent être visualisées dans différentes langues.

# Conclusions et perspectives

## Conclusion

Le travail mené dans cette thèse a permis d'évaluer le potentiel de l'utilisation des modulations numériques dans le cadre de la mise en oeuvre d'un processus d'insertion et d'extraction du tatouage pour enrichir le contenu audio. Le problème a été abordé comme la définition d'une modulation appropriée pour obtenir de meilleures performances du système de tatouage en termes d'inaudibilité, de débit, de fiabilité de détection et de robustesse. Cette étude a fait donc intervenir deux disciplines : le traitement du signal audio et la modulation en communications numériques.

L'objectif de cette thèse est d'améliorer les performances du système de tatouage selon les critères cités ci-dessus. Ainsi, nous avons commencé par étudier les systèmes de tatouage les plus utilisés dans la littérature. Le fait que la plupart de ces systèmes sont basés sur la modulation par étalement de spectre nous a poussé à se demander sur quel levier pourrait reposer son large utilisation plus que les autres types de modulations comme celles à bande étroite.

L'objectif de recherche de cette thèse s'est porté sur la conception des techniques de modulations numériques dans lesquelles le débit et la robustesse sont maximisés tout en garantissant la contrainte de l'inaudibilité. Ainsi, les contributions majeures de cette thèse ont consisté en l'apport des modulations numériques en tatouage audio. Les techniques de modulations développées dans ce manuscrit sont les suivantes :

- Modulations à bande étroite pour une bonne fiabilité de détection. Ce premier travail a permis de :
  - Mettre en valeur les performances des modulations à bande étroite.
  - Insérer de l'information dans la phase du signal de tatouage pour assurer un double gain en termes de fiabilité de détection et de transparence.
  - Calculer l'amplitude du signal modulé par déplacement de phase pour assurer la transparence du tatouage.
  - Étudier l'impact de la fréquence porteuse du signal modulé PSK sur l'inaudibilité et la fiabilité de détection.
  - Proposer un nouveau système de tatouage informé dont les paramètres s'adaptent selon le signal audio afin d'assurer une meilleure performance en terme de fiabilité de détection.
- Modulations multi-porteuse pour améliorer la robustesse vis-à-vis des perturbations externes. Ce deuxième travail a permis de :

- Contrôler la puissance du signal modulé à porteuses multiples pour assurer la contrainte d'inaudibilité.
- Améliorer la robustesse du système de tatouage face aux différentes perturbations externes y compris la compression MPEG.
- Modulations multi-tatouage pour augmenter le débit d'insertion. Ce troisième travail a permis de :
  - Modéliser le système de tatouage comme un système de communication multi-utilisateurs particulier, où les différents utilisateurs jouent le rôle de plusieurs tatouages qui partagent le même canal audio.
  - Exploiter les trois techniques d'accès multiples (CDMA, FHMA et FDMA) dans le nouveau système multi-tatouage.
  - Contrôler l'inaudibilité du tatouage par un facteur d'amplitude qui permet l'insertion de plusieurs tatouages avec une qualité audio presque identique à celle du signal original.
  - Proposer une nouvelle procédure d'extraction exploitant le concept de séparation de sources.
  - Augmenter le débit d'insertion du système multi-tatouage dans le cas particulier d'un signal double canal.

## Perspectives

Ce travail pourrait être poursuivi sur plusieurs points essentiels afin d'améliorer les performances du système de tatouage :

- Les systèmes de tatouage basés sur les modulations multi-porteuses sont robustes vis-à-vis des perturbations externes et ceux basés sur les techniques d'accès multiples offrent un débit d'insertion élevé. Pour bénéficier des avantages présentés par les deux systèmes, une première perspective consiste à proposer un nouveau système de tatouage basé sur les techniques MC-CDMA<sup>18</sup> utilisées dans la téléphonie mobile de quatrième génération afin d'améliorer d'une part la robustesse et d'autre part augmenter le débit d'insertion.
- Une deuxième perspective consiste à insérer les tatouages dans le domaine Transformé (DWT, MDCT) qui fournit généralement une meilleure qualité de perception et de robustesse contre les distorsions au prix de l'augmentation de la complexité de calcul.
- Le bruit audio étant particulièrement différent du bruit blanc gaussien, une troisième perspective consiste à modéliser quelques signaux audio par un modèle théorique qui permettra de calculer la capacité de Shannon dans une chaîne de tatouage audio.
- Le système de tatouage basé sur la séparation de sources présenté dans 4.2 a permis l'extraction du tatouage de deux manières différentes mais en un temps de calcul un peu lent. Une troisième perspective consiste à concevoir de nouvelles techniques de séparation de

---

18. Multi-Carrier CDMA

*"Je sais pourquoi tant de gens aiment couper du bois. C'est une activité où l'on voit tout de suite le résultat".*  
**Albert Einstein**

---

sources adaptées au contexte du tatouage qui fonctionnent en temps réel. Il serait également intéressant d'intégrer les méthodes de fusion pour améliorer la fiabilité de détection.

- Une quatrième perspective consiste à proposer de nouveaux schémas de tatouage pour les signaux stéréo et les signaux multi-canal.
- Une dernière perspective consiste à proposer de nouvelles méthodes robustes aux perturbations de transmission du signal radio.

---

# Annexe A

## Audition

### A.1 Système auditif humain

Le système auditif humain s'étend sur une plage de fréquences audibles allant de 20Hz à 20 kHz. De plus, les capacités de discrimination aussi bien en fréquence qu'en intensité sont excellentes. Ceci est dû au faible nombre de cellules sensorielles dont dispose la cochlée<sup>1</sup> en comparaison avec les cellules photosensibles de l'oeil<sup>2</sup>.

#### A.1.1 Principe

L'appareil auditif est divisé en trois éléments de nature et de fonction différentes mais complémentaires [43] :

- L'oreille externe joue le rôle acoustique pour l'audition ; elle reçoit les vibrations acoustiques aériennes et les transmet à l'oreille moyenne.
- L'oreille moyenne permet de transmettre les vibrations aériennes vers l'oreille interne. Elle est l'élément essentiel de la transmission sonore.
- L'oreille interne est l'organe responsable de la transduction du signal acoustique en message nerveux.

La structure de l'appareil auditif est représentée par la Figure A.1. Tout d'abord, les vibrations acoustiques sont captées par l'oreille externe, ensuite transmises par l'oreille moyenne au milieu liquidien de la cochlée. À l'intérieur de la cochlée, les vibrations provoquent la mise en mouvement des différentes membranes qui la constituent. Par conséquent, les fibres nerveuses s'activent et transmettent alors un message électrique vers le cortex cérébral.

#### A.1.2 Mesure physique de l'intensité

Le système auditif humain est sensible aux variations de la pression atmosphérique provoquées par une onde acoustique. En acoustique, pour définir le niveau de pression, on utilise le

---

1. Organe de l'audition de l'oreille interne contenant 3500 cellules ciliées internes.  
2. Des millions de cellules.

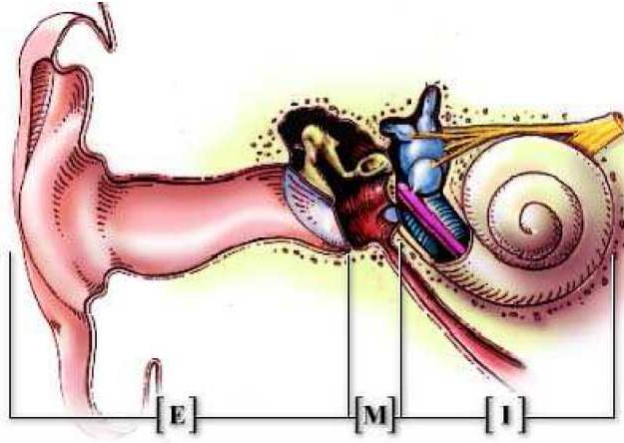


FIGURE A.1 – Schéma de l'appareil auditif humain comprenant l'oreille externe [E], l'oreille moyenne [M] et l'oreille interne [I] [97].

décibel ou (dB SPL<sup>3</sup>). Il est calculé à partir des valeurs de références en pression,  $P_0$ , ou en intensité,  $I_0$ , telles que

$$P_0 = 20\mu Pa \tag{A.1}$$

$$I_0 = 10^{-12} W/m^2 \tag{A.2}$$

La pression acoustique de référence  $P_0$  est choisie comme étant approximativement la plus petite valeur de pression acoustique audible.  $I_0$  est l'intensité correspondante. Le niveau de pression acoustique en décibels,  $L_{dB}$ , est alors défini en fonction de l'intensité acoustique  $I$ , ou de la pression efficace  $P$ , par la relation :

$$\begin{aligned} L_{dB} &= 10 \log \frac{I}{I_0} \\ &= 20 \log \frac{P}{P_0} \end{aligned} \tag{A.3}$$

Le tableau A.1 indique quelques ordres de grandeur de niveau de dB SPL pour des exemples de bruits correspondants.

## A.2 Psychoacoustique

La psychoacoustique s'attache à la modélisation du système auditif humain et la perception du son [120]. Ses résultats, essentiellement basés sur des expérimentations, mettent en évidence les défauts du système auditif, et plus particulièrement les phénomènes de masquage.

### A.2.1 Seuil auditif

Les seuils auditifs indiquent le niveau sonore minimal requis pour détecter un stimulus sonore en fonction de sa fréquence. Ainsi, on constate, par exemple, que le seuil de détection d'un son

---

3. Sound Pressure Level

---

dB SPL	Pression ( $\mu\text{Pa}$ )	Exemple de situation
130 - 140 dB	$63 \cdot 10^6 - 20 \cdot 10^7$	Seuil intolérable
120 dB	$20 \cdot 10^6$	Réacteur d'avion à 10 m
110 dB	$63 \cdot 10^5$	Atelier de chaudronnerie
100 dB	$20 \cdot 10^5$	Marteau-piqueur à 2 m
90 dB	$63 \cdot 10^4$	Atelier de tissage
80 dB	$20 \cdot 10^4$	Rue bruyante
60 dB	$20 \cdot 10^3$	Conversation vive
50 dB	$63 \cdot 10^2$	Musique douce
40 dB	$20 \cdot 10^2$	Conversation normale
30 dB	630	Résidence tranquille
20 dB	200	Studio d'enregistrement
10 dB	63	Laboratoire d'acoustique, minimum extérieur
5 dB	36	Chambre anéchoïque
0 dB	20	Niveau de référence

Tableau A.1 – Correspondance entre niveau de pression acoustique et pression.

grave est plus haut que celui d'un son aigu. L'échelle des phones indique des sensations d'intensité sonore équivalentes le long de la plage des fréquences audibles. Elle est naturellement illustrée par les courbes isophoniques ou courbes d'isophonie.

La Figure A.2 représente les courbes isophoniques mesurées en champ libre pour une écoute binaurale. Leur allure est légèrement différente dans la plage de fréquences 1-8 kHz du fait des diffractions sur la tête et au niveau de l'oreille externe. Les courbes d'isophonie représentent des limites du champ sonore de l'audition. Les courbes hautes (110 à 120 dB) indiquent le seuil de douleur et la courbe la plus basse correspond au seuil d'audition qui peut être approchée par la fonction non linéaire :

$$S_a(f) = 3,64f^{-0.8} - 6,5 \exp(-0,6(f - 3,3)^2) + 10^{-3}f^4 \quad (dB) \quad (A.4)$$

## A.2.2 Phénomène de masquage

Le masquage est le phénomène par lequel un son, le son masqué, peut être rendu inaudible par la présence d'un second, le son masquant. Deux types de masquage peuvent être distingués comme c'est illustré Figure A.3 :

- Le masquage temporel, lorsque les deux sons se produisent à intervalle de temps rapproché mais disjoint
- Le masquage fréquentiel ou simultané, lorsque les deux sons se produisent simultanément.

### A.2.2.1 Masquage temporel

Le masquage temporel peut être de deux natures, en fonction de l'ordre d'apparition des sons masquant et masqué. Si le son masqué est antérieur au son masquant, on parle de pré-masquage ;

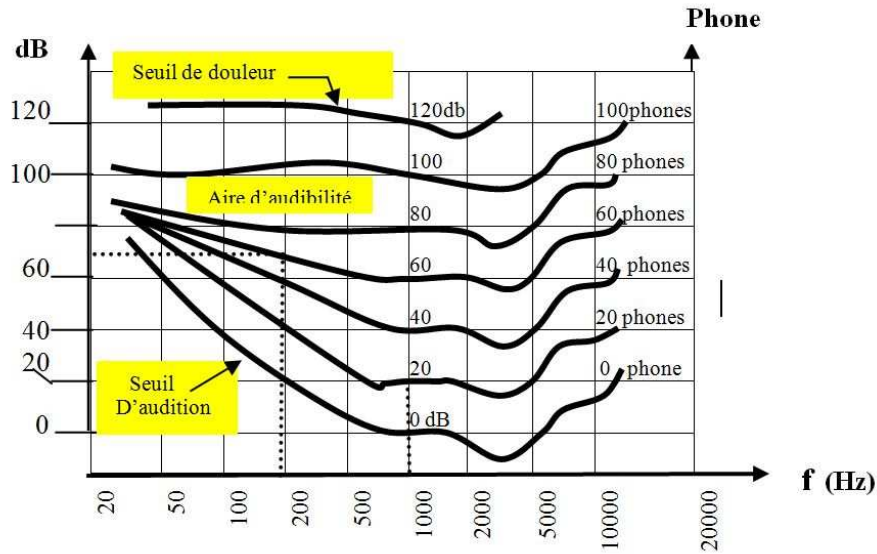


FIGURE A.2 – Diagramme de Fletcher : Courbes d'isophonie pour l'écoute binaurale [3]

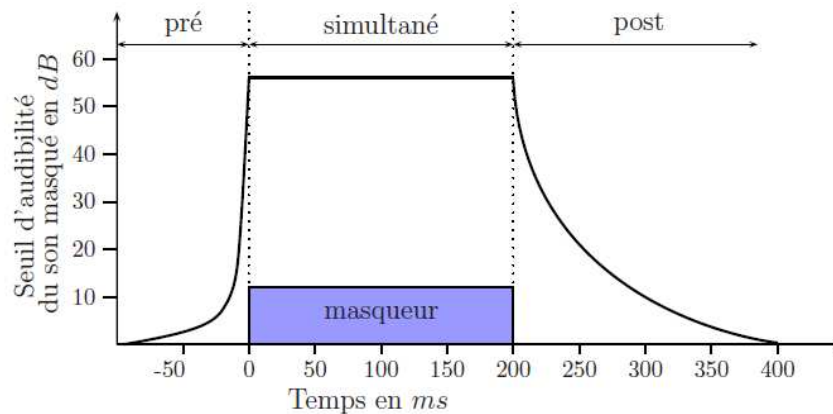


FIGURE A.3 – Seuil d'inaudibilité du son masqué en fonction de l'ordre et du temps d'apparition du son masqué et du son masquant.

sinon, de post-masquage. Un son n'est masqué que si son intensité est inférieure à un certain seuil. Ce seuil est fonction de l'écart<sup>4</sup> temporel entre les deux sons, de la fréquence des deux sons, et de l'intensité et de la durée du son masquant.

#### A.2.2.2 Masquage fréquentiel

Le masquage fréquentiel dépend des spectres d'amplitude du son masquant et du son masqué. Là encore, le son est masqué si son intensité est inférieure à un seuil, fonction de l'intensité

4. Cet écart est en fait l'écart d'apparition entre les sons dans le cas du pré-masquage et de l'écart entre la disparition du masquant et de l'apparition du masqué dans le cas du post-masquage.

---

du masquant, des fréquences respectives des sons masqué et masquant et de leurs natures (sinusoïdales ou bruitées). Ce phénomène peut être expliqué d'un point de vue anatomique. La présence du son masquant de forte intensité, provoque la vibration de la membrane basilaire. Cette vibration est relativement forte au point spécifique de la membrane correspondant à la fréquence du masquant et atténuée dans les zones avoisinantes. Une seconde vibration, liée au son masqué, ne peut donc être détectée que si elle dépasse la vibration (même atténuée) provoquée par le son masquant.

---

## Annexe B

# Filtrage

### B.1 Filtre de mise en forme

Le filtre de mise en forme (perceptive ou spectrale)  $H(f)$  est un filtre dont la réponse fréquentielle à une excitation par un signal d'entrée (signal modulé  $v(n)$ ) s'approche au mieux de la courbe du seuil de masquage  $M_x(f)$ . A l'issue du filtrage, le signal résultant  $w(n)$  utilisera au mieux la puissance disponible. En effet, cette opération permet d'augmenter au maximum le rapport des puissances tatouage à signal audio ( $\sigma_w^2/\sigma_x^2$ ) sous la contrainte d'inaudibilité [9, 74] :

$$S_w(f) \leq M_x(f) \quad (\text{B.1})$$

Le signal audio  $x(n)$  étant localement stationnaire, le modèle psychoacoustique (MPA) détermine le seuil de masquage sur des fenêtres d'analyses de longueurs  $N_{mpa}$  échantillons. Le filtre doit donc être réactualisé pour chaque fenêtre d'analyse. En effet, on désire avoir :

$$\begin{aligned} S_w(f) &= M_x(f) \\ &= S_v(f) |H(f)|^2 \end{aligned} \quad (\text{B.2})$$

Si on utilise une technique de modulation par étalement de spectre, le signal modulé  $v(n)$  est blanc et de puissance unité ( $\sigma_v^2 = 1$ ). On aura alors :

$$|H(f)|^2 = M_x(f) \quad (\text{B.3})$$

Le calcul des coefficients optimaux du filtre de mise en forme peut être le résultat d'une méthode de synthèse de filtre quelconque. Il s'agit de construire un filtre dont le module de la réponse en fréquence satisfait la relation (B.3). Une façon de le faire consiste à modéliser le signal  $w(n)$  par un processus autorégressif (AR). Un processus AR est modélisé par la réaction d'un filtre linéaire tout-pôles d'ordre  $P$ , stable et causal, excité par un signal  $v(n)$ . La modélisation d'un tel processus s'écrit sous la forme suivante :

$$w(n) + \sum_{i=1}^P a_i w(n-i) = b_0 v(n) \quad (\text{B.4})$$

où les  $a_i$  sont les coefficients et  $b_0$  le gain du modèle AR. La fonction d'autocorrélation  $R_w(k)$  du signal  $w(n)$  est obtenue à partir de (B.4) par :

$$E[w(n)w(n-k)] = E[b_0 v(n)w(n-k) - \sum_{i=1}^P a_i w(n-i)w(n-k)], \quad (\text{B.5})$$

---

avec  $k = 1, \dots, P$ , ce qui est équivalent à :

$$\sum_{i=1}^P a_i E[w(n-i)w(n-k)] = b_0 E[v(n)w(n-k)], \quad (\text{B.6})$$

avec  $a_0 = 1$  et  $k = 1, \dots, P$ .

Le terme  $E[v(n)w(n-k)]$  est nul pour  $k > 0$ , car les valeurs passées du signal  $w(n)$  ne dépendent pas de la valeur de  $v(n)$  à l'instant  $n$ , puisque  $v(n)$  est blanc :

$$\sum_{i=0}^P a_i R_w(k-i) = 0, \quad k > 0; \quad (\text{B.7})$$

d'où

$$-\sum_{i=1}^P a_i R_w(k-i) = R_w(k), \quad k > 0; \quad (\text{B.8})$$

En formulant (B.8) sous forme d'un système d'équations linéaires pour  $k = 1, \dots, P$  on obtient les équations de Yule-Walker :

$$\text{III} - \begin{bmatrix} R_w(0) & R_w(1) & \cdots & R_w(P-1) \\ R_w(1) & R_w(0) & \cdots & R_w(P-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_w(P-1) & R_w(P-2) & \cdots & R_w(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} R_w(1) \\ R_w(2) \\ \vdots \\ R_w(P) \end{bmatrix} \quad (\text{B.9})$$

ce qui est équivalent à la notation matricielle :

$$-R_w a = r_w \quad \text{donc} \quad a = -R_w^{-1} r_w \quad (\text{B.10})$$

où  $R_w$  est la matrice d'autocorrélation d'ordre  $P$  de  $w(n)$  qui est inversible car elle est définie positive. Le facteur de gain  $b_0$  est déterminé à partir de la relation (B.6) pour  $k = 0$  :

$$\begin{aligned} \sum_{i=0}^P a_i R_w(i) &= b_0 E[v(n)w(n)] \\ &= b_0 E[v(n)(b_0 v(n) - \sum_{i=1}^P a_i w(n-i))] \\ &= b_0^2 E[v(n)^2] - b_0 \sum_{i=1}^P a_i E[v(n)w(n-i)] \end{aligned} \quad (\text{B.11})$$

Comme  $E[v(n)w(n-i)] = 0$  pour  $i > 0$ , on a :

$$b_0 = \sqrt{\sum_{i=0}^P a_i R_w(i)} \quad (\text{B.12})$$

où  $R_w(k)$  est obtenue par transformée de Fourier inverse de la contrainte (B.2) :

$$R_w(k) = \int_{-\frac{1}{2}}^{\frac{1}{2}} M_x(f) \exp(j2\pi f k) df \quad (\text{B.13})$$

La matrice d'autocorrélation  $R_w$  est symétrique, de Toeplitz et définie positive : les coefficients  $a_i$ ,  $i = 0, \dots, P$  définissent un polynôme d'ordre  $P$  dont toutes les racines sont de module inférieur à l'unité. Ainsi, en appliquant la transformée en  $z$  à l'équation (B.4), on obtient la fonction de transfert du filtre  $H(z)$  recherché :

$$H(z) = \frac{b_0}{1 + \sum_{i=1}^P a_i z^{-i}} \quad (\text{B.14})$$

qui est tout-pôle stable et causale.

---

---

## B.2 Annuleur d'interférence

Le signal de sortie  $y(n)$  s'écrit [9, 74] :

$$y(n) = \sum_i h(i)v(n-i) + x(n) \quad (\text{B.15})$$

Soit l'égaliseur de fonction de transfert  $G_{AI}(z)$ , causal et de RI finie  $g_{AI}(n)$  de support  $M$ , dont la sortie  $\hat{v}(n)$  constitue une estimation du signal modulé  $v(n)$ . L'ensemble canal-égaliseur a pour RI  $c(n)$  donnée par :

$$c(n) = \sum_i h(i)g_{AI}(n-i) \quad (\text{B.16})$$

Le signal estimé  $\hat{v}(n)$  s'exprime par :

$$\begin{aligned} v(n) &= \sum_i c(i)v(n-i) + \sum_{k=0}^M g_{AI}(k)v(n-k) \\ &= c(0)v(n) + \sum_{i \neq 0} c(i)v(n-i) + \sum_{k=0}^M g_{AI}(k)x(n-k) \end{aligned} \quad (\text{B.17})$$

L'égaliseur dont la RI vérifie :

$$\sum_i h(i)g_{AI}(n-i) = 1, \quad n = 0 \quad (\text{B.18})$$

$$\sum_i h(i)g_{AI}(n-i) = 0, \quad n \neq 0 \quad (\text{B.19})$$

est appelé annuleur d'interférence ou égaliseur par forçage à zéro.

On peut déduire de (B.16) et (B.19) que l'échantillon estimé est égale à l'échantillon émis perturbé par le bruit filtré :

$$\hat{v}(n) = v(n) + \sum_{k=0}^M g_{AI}(k)x(n-k) \quad (\text{B.20})$$

La transformée en  $z$  de la condition (B.16) donne la fonction de transfert de l'annuleur d'interférences :

$$G_{AI}(z) = \frac{1}{H(z)} \quad (\text{B.21})$$

Sous l'hypothèse de canal connu, l'annuleur d'interférences est le filtre inverse du canal. Si celui-ci n'est pas à minimum de phase (i.e. il possède des zéros à l'extérieur du cercle unité), l'égaliseur  $G(z)$  donné par (B.25) n'est pas causal.

L'inconvénient majeur de l'annuleur d'interférences est qu'il rehausse le bruit dans le cas où le canal présente une importante atténuation en fréquence. Par ailleurs, le bruit filtré à la sortie de l'égaliseur est en général corrélé. Par conséquent, le détecteur à corrélation n'est pas dans un contexte optimal pour détecter les symboles émis  $\{s_k\}$ . La relation (B.25) implique par ailleurs que l'ordre de  $G_{AI}(z)$  est égal à celui de  $H(z)$  ( $M = P$ ).

---

### B.3 Filtre de wiener

Pour déterminer la RI  $g_W(n)$  du filtre de Wiener, on doit résoudre les équations de Wiener-Hopf [74] :

$$\sum_{l=-\infty}^{\infty} g_W(l)R_y(k-l) = R_{vy}(k), \quad \forall k \in \mathbb{Z} \quad (\text{B.22})$$

où  $R_y(k) = E[y(n)y(n-k)]$  et  $R_{vy}(k) = E[v(n)y(n-k)]$  sont respectivement les fonctions d'auto-corrélation de  $y(n)$  et d'inter-corrélation de  $v(n)$  et de  $y(n)$ . Étant donné que  $y(n) = h(n) * v(n) + x(n)$ , on a :

$$\begin{aligned} \sum_{l=-\infty}^{+\infty} g_W(l)R_y(k-l) &= E[v(n) \sum_i h(i)v(n-k-i)] \\ &= \sum_i h(i)R_v(k+i), \quad \forall k \in \mathbb{Z} \end{aligned} \quad (\text{B.23})$$

La résolution de ce système nécessite des approximations en terme de troncature des RI, ainsi que les hypothèses suivantes :

- La réponse fréquentielle du filtre de mise en forme  $H(z)$  utilisé à l'émetteur est obtenue à la réception par analyse psychoacoustique du signal observé  $y(n)$ . On suppose donc que le seuil de masquage de  $y(n)$  est approximativement égal à celui de  $x(n)$ . On notera  $\hat{h}(n)$  la RI du filtre de mise en forme  $\hat{H}(z)$  estimé au récepteur. L'équation (B.23) s'exprime alors par :

$$\sum_{l=-\infty}^{+\infty} g_W(l)R_y(k-l) = \sum_i \hat{h}(i)R_v(k+i), \quad \forall k \in \mathbb{Z} \quad (\text{B.24})$$

- La RI  $\hat{h}(n)$  est causale mais de longueur infinie, car le filtre  $\hat{H}(z)$  est récursif. Pour calculer  $g_W(n)$ ,  $\hat{h}$  est tronquée à un ordre fini  $P_h$ .
- La RI  $g_W(n)$  du filtre de Wiener est infinie et non causale. Pour une réalisation transverse de  $G_W(z)$ ,  $g_W(n)$  est tronquée à un nombre fini de coefficients. On appellera respectivement  $M_c$  et  $M_{nc}$  les longueurs des parties causale et anticausale de  $g$ .
- L'estimation de la fonction d'autocorrélation  $R_v(k)$  de  $v(n)$  peut se faire sur une longue observation. En effet, on suppose que le récepteur possède des connaissances sur les statistiques des symboles émis  $s_k$ , ainsi que l'alphabet d'émission  $A$ . Le récepteur dispose également du dictionnaire d'émission  $D$ . Il suffit alors de générer une séquence aléatoire aussi longue que l'on veut à partir de l'alphabet d'émission et de la moduler à partir du dictionnaire. On obtient ainsi un signal, dont les statistiques sont identiques à celle de  $v(n)$ . Une bonne estimation de la fonction d'autocorrélation  $R_v(k)$  est donc possible au récepteur.
- Par contre, l'estimation de  $R_y$  ne peut se faire qu'au rythme des fenêtres d'analyse imposé par le modèle psychoacoustique, donc tous les  $N_{mpa} = 512$  échantillons. L'estimation de  $R_y(k)$  est donc beaucoup moins bonne que celle de  $R_v(k)$ , d'autant plus que le signal  $y(n)$  est fortement non stationnaire.

Une fois les deux fonctions d'autocorrélation  $R_v(k)$  et  $R_y(k)$  et la RI  $\hat{h}(n)$  estimées, on résout

---

le système :

$$\sum_{l=-M_{nc}}^{M_{nc}} g_W(l)R_y(k-l) = \sum_{i=0}^{P_h} \hat{h}(i)R_v(k+i), \quad \text{pour } k = -M_{nc}, \dots, M_{nc} \quad (\text{B.25})$$

et on obtient

$$- \begin{bmatrix} g_W(-M_{nc}) \\ \vdots \\ g_W(0) \\ \vdots \\ g_W(M_{nc}) \end{bmatrix} = R_y^{-1} R_v \begin{bmatrix} \hat{h}(0) \\ \vdots \\ \hat{h}(P_h) \end{bmatrix} \quad (\text{B.26})$$

où les matrices d'autocorrélation  $R_y$  et  $R_v$  sont données par :

$$R_y = \begin{bmatrix} R_{\hat{y}}(0) & \cdots & R_{\hat{y}}(-M_{nc} - M_c) \\ \vdots & & \vdots \\ R_{\hat{y}}(M_{nc}) & \cdots & R_{\hat{y}}(-M_c) \\ \vdots & & \vdots \\ R_{\hat{y}}(M_{nc} + M_c) & \cdots & R_{\hat{y}}(0) \end{bmatrix} \quad (\text{B.27})$$

$$R_v = \begin{bmatrix} R_v(-M_{nc}) & \cdots & R_v(-M_{nc} + P_h) \\ \vdots & & \vdots \\ R_v(0) & \cdots & R_v(P_h) \\ \vdots & & \vdots \\ R_v(M_c) & \cdots & R_v(M_c + P_h) \end{bmatrix} \quad (\text{B.28})$$

La réalisation transverse décrite par le système (4.10) du filtre  $G_W(z)$  de l'équation (4.2) n'est plus optimale du fait de la troncature des RI  $g_W(n)$  et  $\hat{h}(n)$  d'une part, et du fait de l'approximation de  $h(n)$  par  $\hat{h}(n)$  d'autre part.

---

# Bibliographie

- [1] What is the future for watermarking? (part i). *IEEE Signal Processing Magazine*, 20(5) :55–60, 2003.
- [2] What is the future for watermarking? (part ii). *IEEE Signal Processing Magazine*, 20(6) :53–59, 2003.
- [3] *Intensité sonore et niveau sonore*. <http://sciences-physiques.ac-montpellier.fr/spip.php?article69>, 2010.
- [4] Abdellah Adib. Séparation aveugle de sources : Fonctions de contrastes et systèmes à référence. *Doctorat d'état spécialité Physique Option Traitement du signal.*, 2004.
- [5] A. Anastasijevic and D. Coja. Frequency hopping method for audio watermarking. *Telfor Journal, Vol. 4, No. 2*, 2012.
- [6] R. Badeau, L. Girin, A. Liuktus, and G. Richard. Procédé et dispositif de formation d'un signal mixé numérique audio, procédé et dispositif de séparation de signaux, et signal correspondant., April 19 2012. WO Patent App. PCT/EP2011/067,730.
- [7] N. Baranwal and K. Datta. Comparative study of spread spectrum based audio watermarking techniques. *International Conference on Recent Trends in Information Technology*, 2011.
- [8] C. Baras, N. Moreau, and P. Dymarski. Controlling the inaudibility and maximizing the robustness in an audio annotation watermarking system. *IEEE Transactions on Audio, Speech and Language Processing*, 14(5) :1772–1782, 2006.
- [9] Cléo Baras. *Tatouage informé de signaux audio numériques*. Thèse présentée pour obtenir le grade de docteur de l'Ecole Nationale Supérieure des Télécommunications, Décembre 2005.
- [10] P. Bas, J. Liénard, J. M. Chassery, D. Beautemps, G. Bailly, P. Bas, J. Lienard, J m Chassery, D. Beautemps, and G. Bailly. Artus : Animation réaliste par tatouage audiovisuel à l'usage des sourds.
- [11] P. Bassia, I. Pitas, and N. Nikolaidis. Robust audio watermarking in the time domain. *Trans. Multi.*, 3(2) :232–241, June 2001.
- [12] W. Bender, D. Gruhl, N. Morimoto, and A. Lu. Techniques for data hiding. *IBM Syst. J.*, 35(3-4) :313–336, September 1996.

- 
- [13] F. Benedetto, G. Giunta, and A. Neri. Qos assessment of 3g video-phone calls by tracing watermarking exploiting the new colour space 'y'st'. *IET Communications*, 1(4) :696–704, 2007.
- [14] I. Bloch. Information fusion in signal and image processing : Major probabilistic and non-probabilistic numerical approaches. *Digital signal and image processing series*, 2008.
- [15] L. Boney, A. H. Tewfik, and K. N. Hamdy. Digital watermarks for audio signals. In *International Conference on Multimedia Computing and Systems*, pages 473–480, 1996.
- [16] L. Cai, R. Tu, J. Zhao, and Y. Mao. Speech quality evaluation : A new application of digital watermarking. *IEEE T. Instrumentation and Measurement*, 56(1) :45–55, 2007.
- [17] P. Campisi, M. Carli, G. Giunta, and A. Neri. Blind quality assessment system for multimedia communications using tracing watermarking. *IEEE Transactions on Signal Processing*, 51(4) :996–1002, 2003.
- [18] M.A. Casey and A. Westner. Separation of mixed audio sources by independent subspace analysis. In *ICMC 2000*, pages 154–161, Berlin, Germany, 2000.
- [19] C.Baras. *Etude de la mise en forme de l'information binaire dans un système de tatouage audio*. Master's thesis, ENST, Août, 2002.
- [20] F. Chaabane, M. Charfeddine, and C. Ben Amar. The impact of error correcting coding in audio watermarking. *International Conference on Next Generation Networks and Services (NGNS)*, 2011.
- [21] R. Chamlawi, A. Khan, and I. Usman. Authentication and recovery of images using multiple watermarks. *Computers & Electrical Engineering*, 36(3) :578–584, 2010.
- [22] M. Chen, Y. He, and L.R. Lagendijk. A fragile watermark error detection scheme for wireless video communications. *IEEE Transactions on Multimedia*, 7(2) :201–211, 2005.
- [23] I. Christine and J. Edward. Digital watermarking : Algorithms and applications, 2001.
- [24] Pierre Comon. Independent component analysis, a new concept? *Signal Process.*, 36(3) :287–314, April 1994.
- [25] M. Costa. Writing on dirty paper (corresp.). *IEEE Trans. Inf. Theor.*, 29(3) :439–441, September 2006.
- [26] I. Cox, M. L. Miller, and A. L. Mckellips. Watermarking as communications with side information. In *Proceedings of the IEEE*, pages 1127–1141.
- [27] I. Cox, M.L. Miller, and A.J. Bloom. *Digital watermarking*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002.
- [28] I J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image processing*, 6(12) :1673–1687, 1997.
- [29] I.J. Cox and L.M Miller. The first 50 years of electronic watermarking. *EURASIP J. Adv. Sig. Proc.*, 2002(2) :126–132, 2002.
- [30] N. Cvejic. *Algorithms for Audio Watermarking and Steganography*. Acta Universitatis Ouluensis : Technica. Oulun yliopisto, 2004.
-

- 
- [31] N. Cvejic, A. Keskinarkaus, and T. Seppanen. Audio watermarking using m-sequences and temporal masking. *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, 2001.
- [32] N. Cvejic and T. Seppänen. Spread spectrum audio watermarking using frequency hopping and attack characterization. *Signal Process.*, 84(1) :207–213, January 2004.
- [33] N. Cvejic and T. Seppanen. *Digital Audio Watermarking Techniques and Technologies : Applications and Benchmarks*. IGI Global, Hershey, PA, USA, 2007.
- [34] N. Cvejic, D. Tujkovic, and T. Seppanen. Increasing robustness of an audio watermark using turbo codes. In *ICME*, pages 217–220. IEEE, 2003.
- [35] Leandro de C. T. Gomes. Tatouage de signaux audio. *Doctoral Thesis, Université Paris V, Paris, France*, 2002.
- [36] O. Derrien, S. Larbi, M. Guimaraes, and N. Moreau. Le codeur mpeg-2 aac expliqué aux traiteurs de signaux. *Annales des Télécommunications*, 55(9-10) :442–461, 2000.
- [37] J. Dittmann, P. Wohlmacher, and K. Nahrstedt. Using cryptographic and watermarking algorithms. *IEEE MultiMedia*, 8(4) :54–65, October 2001.
- [38] E. Erçelebi and L. Batakçı. Audio watermarking scheme based on embedding strategy in low frequency components with a binary image. *Digit. Signal Process.*, 19(2) :265–277, March 2009.
- [39] M. Fallahpour and D. Megias. Dwt-based high capacity audio watermarking. *IEICE Trans Fundamentals*, E93-A(1), 2010.
- [40] M. Fallahpour and D. Megias. High capacity audio watermarking using the high frequency band of the wavelet domain. *Multimedia Tools Appl.*, 52(2-3) :485–498, 2011.
- [41] K. Fazel and S. Kaiser. *Multi-Carrier and Spread Spectrum Systems*. John Wiley & Sons, Inc., New York, NY, USA, 2003.
- [42] W. Feng, W. Xiaolin, F. Gui, and C. Canhui. A novel blind watermarking scheme based on watermark hopping technique. *8th International Conference on Signal Processing*, 2006.
- [43] Thomas FILLON. *Traitement Numérique du Signal Acoustique pour une Aide aux Malentendants*. Thèse présentée pour obtenir le grade de docteur de l’Ecole Nationale Supérieure des Télécommunications, 14 décembre 2004.
- [44] F.Xiong. *Digital Modulation Techniques*. Library of Congress Cataloging, 2000.
- [45] B. Geiser, P. Jax, and P. Vary. Artificial bandwidth extension of speech supported by watermark-transmitted side information. In *INTERSPEECH*, pages 1497–1500. ISCA, 2005.
- [46] L. Gomes, P. Cano, E. Gómez, M. Bonnet, and E. Batlle. Audio watermarking and fingerprinting for which applications? *Journal of New Music Research*, 32, 2003.
- [47] M. Hagsmüller, H. Horst, A. Kröpfl, and G. Kubin. Speech watermarking for air traffic control. In Speech European Association for Signal and Image Processing, editors, *Proceedings of 12th European Signal Processing Conference*, volume n/a, pages 1653 – 1656, Sep 2004.
-

- 
- [48] N. El Hamdouni, A. Adib, S. Larbi, and M. Turki. Hybrid embedding strategy for a blind audio watermarking system using EMD and ISA techniques. In *International Symposium on Control, Communications and Signal Processing*, 2010.
- [49] Nawal El Hamdouni. Technique hybride emd/sas pour le tatouage audio : Une approche potentielle pour augmenter le débit de transmission. *Thèse de doctorat, à la faculté des sciences de rabat, Maroc*, 2012.
- [50] N. Hamdounil, A. Adib, S. Larbi, and M. Turki. A blind digital audio watermarking scheme based on emd and uisa techniques. *Multimedia Tools Appl.*, 64(3) :809–829, 2013.
- [51] L. Hanzo, M. Manster, B. J. Choi, and Thomas Keller. *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting [Hardcover]*. John Wiley & Sons, Inc., New York, NY, USA, 2003.
- [52] Simon Haykin. *Adaptive Filter Theory (3rd Ed.)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.
- [53] Simon Haykin. *Communication Systems*. Wiley Publishing, 5th edition, 2009.
- [54] X. He. *Signal Processing, Perceptual Coding and Watermarking of Digital Audio : Advanced Technologies and Models*. Premier Reference Source. Igi Global, 2011.
- [55] W. Hoeg and T. Lauterbach. *Digital Audio Broadcasting : Principles and Applications of Digital Radio, 2nd Edition*. Wiley, 2003.
- [56] K. Hofbauer, H. Hering, and G. Kubin. Speech watermarking for the vhf radio channel. In *n/a*, volume n/a, pages 0 – 0, Dec 2005.
- [57] Konrad Hofbauer. Speech watermarking and air traffic control. *Doctoral Thesis*, 2009.
- [58] H. Huang, S. Chu, J. Pan, C. Huang, and B. Liao. Tabu search based multi-watermarks embedding algorithm with multiple description coding. *Inf. Sci.*, 181(16) :3379–3396, 2011.
- [59] N. E. Huang, Z. Shen, S. R. Long, M. L. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and hilbert spectrum for nonlinear and non-stationary time series analysis. In *R. Soc. London A*, volume 454, pages 903–995, 1998.
- [60] R. Hung and M. Kendall. Rds encoder for fm transmitter, June 12 2008. US Patent App. 11/839,169.
- [61] M. Ikeda, K. Takeda, and F. Itakura. Audio data hiding by use of band-limited random sequences. *IEEE International Conference on Acoustics, Speech, and Signal Processing*.
- [62] M. Ikeda, K. Takeda, and F. Itakura. Audio data hiding by use of band-limited random sequences. *International Conference on Acoustics, Speech, and Signal Processing*, 1999.
- [63] S. Jung, J. Seok, and J. Hong. An improved detection technique for spread spectrum audio watermarking with a spectral envelope filter. *ETRI journal*, 25(1) :52–54.
- [64] P. Kabal. An examination and interpretation of itu-r bs.1387 : Perceptual evaluation of audio quality. *Department of Electrical and Computer Engineering McGill University*, 2002.
-

- 
- [65] B. Khaleghi, A. Khamis, F. Karray, and N. Razavi. Multisensor data fusion : A review of the state-of-the-art. *Inf. Fusion*, 14(1) :28–44, 2013.
- [66] Mohammed Khalil. *Influence du type de modulation sur un système de tatouage audio*. PFE Mater Faculté des sciences de Rabat, 2010.
- [67] H. Kim. Stochastic model based audio watermark and whitening filter for improved detection. *IEEE International Conference on Speech, and Signal Processing*.
- [68] J. Kim. Audio watermarking techniques. *Proceedings of the Pacific Rim Workshop on Digital Steganography*, 2003.
- [69] D. Kirovski and H. Malvar. Spread-spectrum watermarking of audio signals. *IEEE Transactions on Signal Processing*, 51(4) :1020–1033, 2003.
- [70] S. Kumar Krishna and S. Thippur. Increased watermark-to-host correlation of uniform random phase watermarks in audio signals. *Signal Process.*, 87(1) :61–67, January 2007.
- [71] D. Kundur and D. Hatzinakos. Diversity and attack characterization for improved robust watermarking. *IEEE Transactions on Signal Processing*, 49(10) :2383–2396, 2001.
- [72] M. Miller L., G. DoÁnrr J., and J.I. Cox. Applying informed coding and embedding to design a robust high-capacity watermark. *IEEE Transactions on Image Processing*, 13(6) :792–807, 2004.
- [73] S. Larbi, M. Jaïdane, and N. Moreau. A new wiener filtering based detection scheme for time domain perceptual audio watermarking. In *ICASSP (5)*, pages 949–952, 2004.
- [74] Sonia Larbi. Structures d’égalisation en tatouage audionumérique. *Thèse de doctorat, Ecole Nationale d’Ingénieurs de Tunis (ENIT) et Télécom Paris (ENST)*,, 2005.
- [75] S. Lee and Y. Ho. Digital audio watermarking in the cepstrum domain. *IEEE Trans. Consumer Electronics*, 46(3) :744–750, 2000.
- [76] X. Li and H. Yu. Transparent and robust audio data hiding in cepstrum domain. In *IEEE International Conference on Multimedia and Expo (I)*, pages 397–, 2000.
- [77] X. Li and H. Yu. Transparent and robust audio data hiding in subband domain. In *ITCC*, pages 74–81. IEEE Computer Society, 2000.
- [78] Z. Li. *Encyclopedia of Biometrics*. Springer Publishing Company, Incorporated, 1st edition, 2009.
- [79] C. Lin, D. Sow, and S. Chang. Using self-authentication-and-recovery images for error concealment in wireless environments. In *SPIE ITCOM/OptiComm*, volume 4518, Denver, CO, August 2001.
- [80] Y. Lin and H. Abdulla. Perceptual evaluation of audio watermarking using objective quality measures. In *ICASSP*, pages 1745–1748. IEEE, 2008.
- [81] S. Mahdavi, S. Seydnejad, and S. Saryazdi. A robust audio watermarking algorithm based on frequency hopping. *18th Iranian Conference on Electrical Engineering (ICEE)*, 2010.
- [82] H. Malik, A. Khokhar, and R. Ansari. Robust audio watermarking using frequency selective spread spectrum theory. In *ICASSP (5)*, pages 385–388, 2004.
-

- 
- [83] M. Miller, G. DoÁñrr, and I. Cox. Applying informed coding and embedding to design a robust high-capacity watermark. *IEEE Transactions on Image Processing*, 13(6) :792–807, 2004.
- [84] B. C. J. Moore. *An Introduction to the Psychology of Hearing*. New York : Academic, 4th edition 1997.
- [85] Brian C. J. Moore. *An Introduction to the Psychology of Hearing, Fifth Edition*. Academic Press, April 2003.
- [86] H. Murata, A. Ogihara, M. iwata, and A. Shiozaki. Multichannel audio watermarking method suitable for multiple watermarks. *International Symposium on Communications and Information Technologies (ISCIT)*, 2008.
- [87] Arthur Murray. Rds vs directband fm spectrum. 2009.
- [88] C. Neubauer and J. Herre. Advanced watermarking and its applications 5176. *109th Convention.*, 2000.
- [89] H. Nguyen and E. Shwedyk. *A First Course in Digital Communications*. Cambridge University Press, 2009.
- [90] K. Ozawa, Y. Suzuki, and T. Sone. Monaural phase effect on timbre of two signals, vol. 93, no 2 ,pp. 1007 - 1011. *J. Acoust. Soc. Amer*, 1993.
- [91] K. Ozawa, Y. Suzuki, and T. Sone. Monaural phase effects on timbre of twoÁÑtone signals. *J. Acoust. Soc. Amer*, 93(2) :1007–1011, 1993.
- [92] P.Blanc. *Centre Auditif de la Riviera -communication active*.
- [93] H. Peng and J. Wang. Optimal audio watermarking scheme using genetic optimization. *Annales des Telecommunications*, 66(5-6) :307–318, 2011.
- [94] Alejandro Lobo Guerrero Perez. Etude de techniques de tatouage audio pour la transmission de données. *Thèse de doctorat, Institut National Polytechnique de Grenoble (INPG), Grenoble, France*, 2004.
- [95] L. Pickholtz, L. Schilling, B. Milstein, and S. Member. Theory of spread spectrum communications, a tutorial. *IEEE Transactions on Communications*, 30 :855–884, 1982.
- [96] R. L. Pickholtz, D. L. Schilling, and B.Milstein. Spread spectrum for mobile communications. *IEEE Transactions on Vehicular Technology*, 40(2), May 1991.
- [97] R. Pujol. *Promenade autour de la cochlée*. 2004.
- [98] A. Sagi and D. Malah. Bandwidth extension of telephone speech aided by data embedding. *EURASIP J. Adv. Sig. Proc.*, 2007.
- [99] M. Sajatovic, J. Prinz, and A. Kroepfl. Increasing the safety of the atc voice communications by using in-band messaging. In *Digital Avionics Systems Conference*, 2003.
- [100] J. Seok and J. Hong. Audio watermarking for copyright protection of digital audio data. *Electronics Letters*, 37(1) :60–61, 2001.
-

- 
- [101] Claude Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27 :379–423, 623–656, July, October 1948.
- [102] H. Shen and B. Chen. From single watermark to dual watermark : A new approach for image watermarking. *Computers and Electrical Engineering*, 38(5) :1310–1324, 2012.
- [103] S. Shokri, M. Ismail, N. Zainal, and A. Shokri. Error probability in spread spectrum (ss) audio watermarking. *IEEE International Conference on Space Science and Communication*, 2013.
- [104] D. Swanson, M. Zhu, A. Tewfik, and L. Boney. Robust audio watermarking using perceptual masking. *Signal Processing*, 66(3) :337–355, 1998.
- [105] R. Tachibana, S. Shimizu, S. Kobayashi, and T. Nakamura. An audio watermarking method using a two-dimensional pseudo-random array. *Signal Processing*, 82(10) :1455–1469, 2002.
- [106] A. Takahashi, R. Nishimura, and Y. Suzuki. Multiple watermarks for stereo audio signals using phase-modulation techniques. *IEEE Transactions on Signal Processing*, 53 :806–815, 2005.
- [107] Information Society Technologies. Audio benchmarking tools and steganalysis. *European Network of Excellence in Cryptology*, 2006.
- [108] J F Tilki and A A Beex. Encoding a hidden digital signature onto an audio signal using psychoacoustics masking. *109th Convention.*, 1996.
- [109] N. Trivedi, T. Kumar, and M. Chandra. An algorithmic digital audio watermarking in perceptual domain using direct sequence spread spectrum. *International Conference on Communication Systems and Network Technologies*, 2014.
- [110] G. Tzanetakis. Music information retrieval : theory and applications. In Wen Gao, Yong Rui, Alan Hanjalic, Changsheng Xu, Eckehard G. Steinbach, Abdulmotaleb El-Saddik, and Michelle X. Zhou, editors, *ACM Multimedia*, pages 915–916. ACM, 2009.
- [111] C. Uhle, C. Dittmar, and T. Sporer. Extraction of drum tracks from polyphonic music using independent subspace analysis. In *Int. Symp. Independent Compon. Anal. Blind Signal Separation (ICA2003)*, pages 843–848, 2003.
- [112] V. Vaseghi. *Advanced Digital Signal Processing and Noise Reduction*. John Wiley & Sons, 2006.
- [113] J. Wang and S. Lian. On the hybrid multi-watermarking. *Signal Processing*, 92(4) :893–904, 2012.
- [114] H. H. Yee and F. S. Wei. Audio watermarking with error-correcting code. *TENCON, Singapore*, 2009.
- [115] C.H. Yeh and S.J.Kuo. Digital watermarking through quasi m-arrays. *IEEE Workshop on Signal Processing Systems*, pages 456 – 461.
- [116] B. Zaidan, A. Zaidan, A. Al-Frajatand, and H. Jalab. On the differences between hiding information and cryptography techniques : An overview. *Journal of Applied Sciences.*, 10(15) :1650–1655.
-

- 
- [117] P. Zhang, Y. Li, Y. Fan, J. Jiang, X. Ma, and Q. Hao. Robust audio watermarking based on frequency-domain spread spectrum using cazac sequence. *International Conference on Advanced Computational Intelligence (ICACI)*, 2013.
- [118] J. Zhao, E. Koch, and C. Luo. In business today and tomorrow. *Communications of the ACM*, 1998.
- [119] X. Zhao, Y. Guo, J. Liu, and Y. Yan. A spread spectrum audio watermarking system with high perceptual quality. In Dongfeng Yuan, Maoyong Cao, Cheng-Xiang Wang, and Hua Huang, editors, *CMC*, pages 266–269. IEEE Computer Society, 2011.
- [120] E. Zwicker, R. Feldtkeller, and C. Sorin. *Psychoacoustique : l'oreille, receptrer d'information*. Collection technique et scientifique des télécommunications. Masson, 1981 (73-Chambéry, Paris, New York, Barcelone).

---

# Publications

## Prix

Troisième Prix R&D Maroc : "Transport d'informations par Tatouage audio dans l'afficheur radio", Huitième édition du Concours National de l'innovation, de la recherche-développement et de technologies, 2013.

## Brevet national

**Mohammed Khalil** and Abdellah Adib "Transport d'informations par tatouage audio sur l'afficheur radio ", No 364652, OMPIC, Maroc (2014).

## Book chapter

**Mohammed Khalil**, Nawal El Hamdouni and Abdellah Adib " Improved Watermark Extraction Exploiting Undetermined Source Separation Methods" Image and Signal Processing, LNCS, vol 7340, ISBN 978-3-642-31253-3, ISSN 0302-9743, (2012).

## Publications

**Mohammed Khalil** and Abdellah Adib " Robust audio watermarking based on multi-carrier modulation " **submitted** to International Journal of Computational Vision and Robotics, ISSN : 1752-9131 (**Under review** from December, 31, 2013).

**Mohammed Khalil** and Abdellah Adib " Multiple AudioWatermarking System based on CDMA " International Journal of Information and Communication Technology, ISSN : 1741-8070 (**Accepted** 2014).

**Mohammed Khalil** and Abdellah Adib " Audio Watermarking with High Embedding Capacity Based on Multiple Access Techniques" Digital Signal Processing (Elsevier), ISSN : 1051-2004 (**Accepted** 2014).

---

**Mohammed Khalil** and Abdellah Adib " Informed audio watermarking based on adaptive carrier modulation " Multimedia tools and applications Springer, ISSN : 1380-7501, (2014).

**Mohammed Khalil** and Abdellah Adib " Improving audio watermarking performance by carrier modulations" International Journal of Computing and Information Technology, Vol.3 No.2, pages 123-136, ISSN : 0974-696X, (2011).

### **Conférences internationales**

**Mohammed Khalil** and Abdellah Adib "Audio watermarking based on PSK modulation and FDMA technique", accepted in International Symposium on signal, Image, Video and Communications (ISIVC), Marrakech, 2014.

**Mohammed Khalil** and Abdellah Adib "Audio watermarking system based on Frequency Hopping modulation", International Conference on Multimedia Computing and Systems (ICMCS), Marrakech, 2014.

**Mohammed Khalil** and Abdellah Adib " Embedding and Extracting Multiple Watermarks in Audio Signals using CDMA", International Conference on Intelligent Systems : Theories and Applications (SITA), Rabat, 2014

**Mohammed Khalil** and Abdellah Adib "Tatouage audio multiple basé sur les techniques d'accès multiples", Congrès Méditerranéen des Télécommunications (CMT), Mohammedia, 2014.

**Mohammed Khalil**, Adil El BAHY and Abdellah Adib " A high capacity blind watermarking for two-channel audio signals based on CDMA-ICA ", European Signal Processing Conference (EUSIPCO), Marrakech, 2013.

**Mohammed Khalil** and Abdellah Adib " Impact of Spreading Codes on CDMA Based Audio Watermarking System ", International Conference on Telecommunications (ICT), Casablanca, 2013.

**Mohammed Khalil** and Abdellah Adib " Audio watermarking with high embedding capacity based on multiple access techniques ", TAIMA, Hammamet, Tunisie, 2013.

**Mohammed Khalil** and Abdellah Adib "Une nouvelle technique de tatouage audio par déplacement de phase des signaux à porteuses multiples", GRETSI Brest, France, 2013.

**Mohammed Khalil**, Nawal El Hamdouni and Abdellah Adib " Increasing the information Capacity and Improving the detection reliability in audio watermarking system", International Symposium on Communications, Control and Signal Processing (ISCCSP), Rome, Italie, 2012.

---

**Mohammed Khalil** and Abdellah Adib " Tatouage audio par étalement à saut de fréquence", International Conference on Intelligent Systems : Theories and Applications (SITA), Mohammedia, 2012.

### Conférences nationales

**Mohammed Khalil** and Abdellah Adib " Increasing Audio Watermarking Capacity Based on CDMA Technique", Workshop on Information Technologies and Communication (WOTIC), Casablanca, 2011.

**Mohammed Khalil** and Abdellah Adib " Performance Evaluation of Bandpass Modulations in Digital Audio Watermarking", Workshop on Codes, Cryptography and Communication Systems (WCCCS), Rabat, 2011.

### Autres

Adil el Bahi and **Mohammed Khalil** and Abdellah Adib " A Dual Audio Watermarking System Based on Spread Spectrum and QIM Technique", International Conference on Intelligent Systems : Theories and Applications (SITA), Rabat, 2013.