

# Calcul de la Probabilité d'Erreur pour une Chaîne de Tatouage Audionumérique à Bruit Non Gaussien

Sonia SAÏED<sup>1</sup>, Sonia LARBI<sup>1</sup>, Rached HAMZA<sup>2</sup>, Larbi BELHADJ SLAMA<sup>2</sup>, Mériem JAÏDANE<sup>1</sup>

<sup>1</sup>Unité Signaux et Systèmes

Ecole Nationale d'Ingénieurs de Tunis, BP 37, 1002 Tunis

<sup>2</sup>Laboratoire des Systèmes de Communications

Ecole Nationale d'Ingénieurs de Tunis, BP 37, 1002 Tunis

sonia.bouajina@yahoo.fr, sonia.larbi@enit.rnu.tn

{rached.hamza, larbi.bhslama}@supcom.rnu.tn, nej.jaidane@planet.tn

**Résumé** – La chaîne de tatouage audionumérique étudiée peut être vue comme une chaîne de communications numériques particulière du fait du caractère non gaussien du bruit additif d'une part et du faible rapport signal à bruit variable par fenêtre d'autre part.

L'objectif de cette étude est de mener un calcul théorique de la probabilité d'erreur que l'on est en mesure d'atteindre, en tenant compte de ces spécificités. Cette étude théorique est un préalable à un choix adéquat d'un système de codage canal dans un contexte de tatouage audio.

**Abstract** – The considered digital audio watermarking scheme can be viewed as a particular digital communications system, because of the non gaussian additive noise on the one hand, and the weakness of the variable signal to noise ratio on the other hand.

The goal of the presented study is to compute the theoretical error probability we can expect, taking into account the system particularities mentioned above. This theoretical study is a preamble for choosing an appropriate channel code for the considered digital audio watermarking system.

## 1 Introduction

Dans la chaîne de tatouage audionumérique étudiée (cf. figure 1) [1], une signature est insérée dans un signal audio de manière inaudible grâce à une mise en forme perceptuelle actualisée par fenêtre d'analyse du signal audio  $x(n)$ . De nombreuses études ont pour objectif l'amélioration de la détection du tatouage. Une des pistes, déjà explorée dans le cas du tatouage d'images, consiste à insérer un système de codage/décodage canal dans la chaîne de tatouage. Un choix adéquat d'un tel codeur et une évaluation préalable des performances de la chaîne de tatouage en terme de probabilité d'erreur nécessitent une meilleure connaissance de ses caractéristiques.

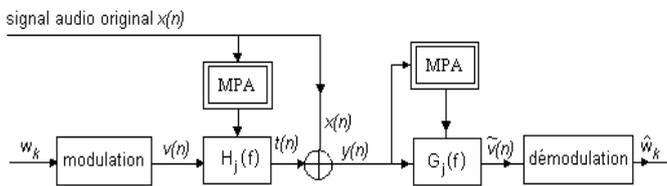


FIG. 1: Schéma de la chaîne de tatouage audio étudiée

Dans la deuxième section de cette étude, nous développons deux schémas équivalents de la chaîne de tatouage de la figure 1. Dans la troisième section, on se base sur ces schémas équivalents pour montrer que la chaîne considérée est équivalente à une chaîne de communications numériques non classique, où en particulier le rapport signal à bruit, variable par fenêtre d'analyse, est relativement faible et où le bruit additif est non gaussien. Ceci nous permet dans la quatrième section de prévoir

les performances attendues de cette chaîne en terme de probabilité d'erreur binaire, lorsque le bruit est soit gaussien soit laplacien par fenêtre.

## 2 Chaîne de tatouage et chaîne de communications numériques

### 2.1 Chaîne de tatouage étudiée

Le système de tatouage audionumérique considéré peut être vu comme un système de communications numériques [2, 3], où l'information  $v(n)$ , supposée blanche (construite à partir d'une signature binaire aléatoire  $w_k$ ) et centrée, est insérée dans le signal audionumérique  $x(n)$  de manière inaudible, grâce à un filtre de mise en forme spectrale  $H_j$  qui génère le signal de tatouage  $t(n)$  (cf. figure 1). Le module au carré de la réponse fréquentielle de  $H_j$  correspond à la courbe de masquage obtenue par le modèle psychoacoustique (MPA) pour chaque fenêtre d'analyse d'indice  $j$  de  $N$  échantillons de  $x(n)$ .

Différentes méthodes de construction de  $v(n)$  à partir des bits  $w_k$  sont possibles. Ici, nous avons choisi la modulation par un dictionnaire  $\{S_0, -S_0\}$ , où  $S_0$  est un vecteur aléatoire composé de  $\{\pm 1\}$  et de durée égale à  $N$  échantillons, afin d'unifier la durée symbole à la durée de la fenêtre d'analyse imposée par le modèle psychoacoustique.

Dans le système étudié, la détection se base sur un filtre de déconvolution de Wiener non causal  $G_j$ , qui à la fois « égalise »  $H_j$  et « réduit » le bruit  $x(n)$ . Notons que la détection de l'estimée de la signature  $\hat{w}_k$  de  $w_k$  est ensuite faite par démodulation avec le même dictionnaire qu'à l'émission [3].

## 2.2 Premier schéma équivalent

Le filtre de Wiener  $G_j$  basé sur la minimisation de l'erreur quadratique moyenne  $E\{[v(n) - \tilde{v}(n)]^2\}$  se décompose selon [3, 4] :

$$G_j(z) = \frac{1}{H_j(z)} F_j(z) \quad (1)$$

où  $H_j$  et  $F_j$  sont respectivement les filtres de mise en forme<sup>1</sup> et de réduction du bruit filtré  $b(n)$ , tel que  $x(z) = b(z)H_j(z)$ . La chaîne de tatouage est alors équivalente au schéma de la figure 2.

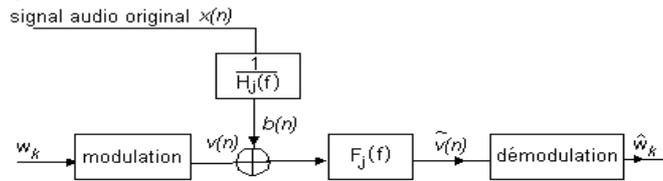


FIG. 2: Premier schéma équivalent de tatouage

## 2.3 Second schéma équivalent

La chaîne de tatouage de la figure 1 est également équivalente au schéma de la figure 3, où  $B(n)$  est la filtrée de  $x(n)$  par le filtre de Wiener  $G_j$  et  $v'(n)$  la filtrée de  $v(n)$  par  $F_j$ .

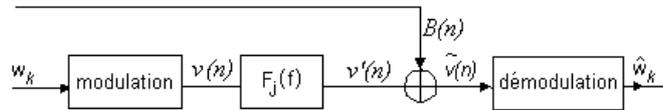


FIG. 3: Second schéma équivalent de tatouage

Pour illustrer le rôle de réduction de la puissance du bruit  $b(n)$  par le filtre  $F_j$ , nous avons superposé sur la figure 4, les densités spectrales de puissance de  $b(n)$  et de  $B(n)$  évaluées pour une fenêtre d'analyse d'un extrait de violon. La réduction de la puissance du bruit apportée par le filtre  $F_j$  est importante (un gain de 20 dB est atteint sur cet exemple). Par conséquent, nous adopterons dans la suite le schéma de équivalent de la figure 3 et on considérera  $B(n)$  comme étant le bruit additif.

1. Dans la relation (1), on a supposé que le filtre  $H_j$  estimé à la réception à partir de  $y(n)$  est identique à celui calculé à l'émission à partir de  $x(n)$ .

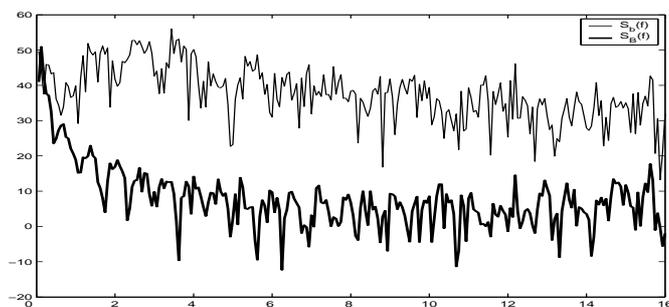
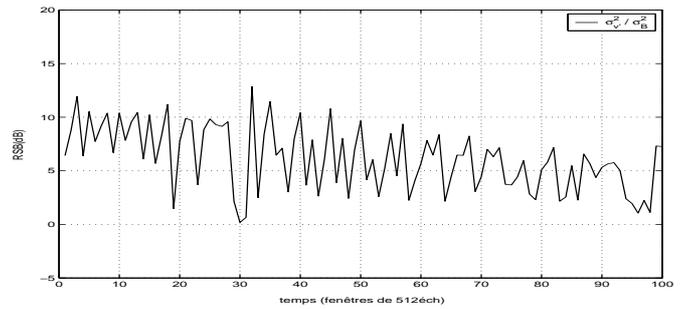
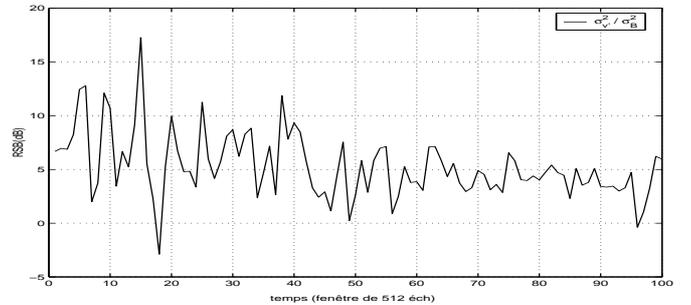


FIG. 4: Exemple de densités spectrales de puissance (dB) de  $b(n)$  et  $B(n)$  pour une fenêtre d'analyse du fichier  $f_2$  (extrait de violon)



(a) Extrait de jonasz



(b) Extrait de violon

FIG. 5:  $RSB_j$  par fenêtre

## 3 Caractéristiques du bruit additif

Dans les simulations, nous avons considéré deux fichiers audio ( $f_1$ : extrait de Michel Jonasz et  $f_2$  extrait instrumental de violon). Ces deux fichiers ont été pris de durée 2 secondes et échantillonnés à  $f_e = \frac{1}{T_e} = 32$  kHz. La taille des fenêtres du MPA est  $N = 512$  échantillons.

### 3.1 Evolution du rapport signal à bruit

Du fait de la réduction du bruit opérée par le filtre  $F_j$  et visualisée à l'aide de l'exemple de la figure 4, nous optons dans la suite de cette étude pour la définition suivante du rapport signal à bruit (cf. schéma de la figure 3):

$$RSB_j = \frac{E\{v'^2(n)\}}{E\{B^2(n)\}} \quad (2)$$

L'évolution du  $RSB_j$  défini plus haut sur des fenêtres de taille  $N = 512$  échantillons pour les deux fichiers  $f_1$  et  $f_2$  est présentée sur la figure 5. Le RSB varie pour le fichier  $f_1$  (figure 5.a) entre 0 et 12 dB autour d'une valeur moyenne de 6 dB et pour le fichier  $f_2$  (figure 5.b) entre -3 et 15 dB autour d'une valeur moyenne de 5 dB. Cette plage de variation sera considérée dans ce qui suit pour l'évaluation des performances de la chaîne de tatouage.

Par ailleurs, la figure 5 confirme que le RSB est relativement faible et variable au cours du temps par fenêtre, ce qui influera fortement sur la probabilité d'erreur que l'on est en mesure d'atteindre.

### 3.2 Bruit additif non gaussien

En première approximation, nous allons considérer que la densité de probabilité  $f_B$  du bruit  $B(n)$  évolue, par fenêtre  $j$ , entre la loi gaussienne et la loi laplacienne, l'hypothèse laplacienne étant couramment adoptée pour les signaux audio.

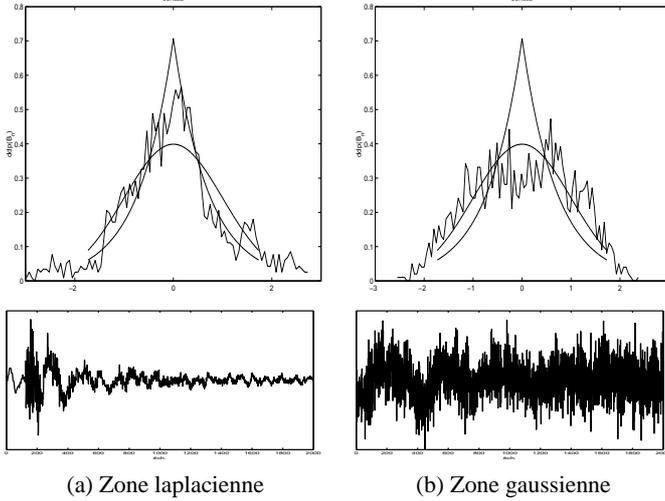


FIG. 6: Densité de probabilité du signal  $B(n)$  pour deux fenêtres différentes du fichier  $f_1$

Dans ce qui suit, les calculs seront menés sous l'hypothèse gaussienne et laplacienne. Pour justifier cette hypothèse, nous avons superposé sur les figures 6a et 6b, l'histogramme centré et normalisé en puissance, du bruit  $B(n)$  c'est à dire du signal  $x(n)$  filtré par Wiener, avec les ddp gaussienne et laplacienne. Ceci est fait pour une fenêtre d'analyse de 512 échantillons pour deux extraits différents du fichier  $f_1$ . Une étude plus précise [6], montre que la ddp de  $B(n)$  obéit à un modèle paramétré plus général, englobant le cas gaussien et laplacien (modèle gaussien généralisé).

## 4 Calcul de la probabilité d'erreur

La statistique du bruit additif équivalent  $B(n)$  variable par fenêtre, impose de mener le calcul de la probabilité d'erreur par fenêtre de  $N$  échantillons. Le signal  $v(n)$  est par fenêtre égal à l'un des vecteurs aléatoires  $\{S_0, -S_0\}$ . On peut donc considérer qu'il est construit à partir d'une séquence binaire  $\alpha(n)$  par codage en ligne NRZ<sup>2</sup> de durée  $T_e$ . L'introduction de  $\alpha(n)$  a uniquement pour rôle de se remettre dans le contexte du calcul classique de probabilité d'erreur relative à  $v(n)$ . Le schéma de tatouage peut alors être ramené à celui de la figure 7. Une détection dure de  $\tilde{v}(n)$  par rapport à un seuil de décision  $S$ , permet d'obtenir  $\hat{\alpha}(n)$ , estimée de  $\alpha(n)$ , équivalente à l'estimée de  $v(n)$ .

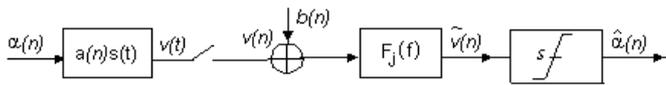


FIG. 7: Chaîne de tatouage avec codeur en ligne

Pour simplifier les calculs, on considère dans la suite l'émission d'un seul symbole binaire  $a(n_0)$  à l'instant  $n_0$ , on a alors :

$$v(n_0) = a(n_0)s(n_0T_e) \quad (3)$$

où  $s(t)$  est l'impulsion de mise en forme de durée  $T_e$ . D'après

la figure 7, nous avons :

$$\tilde{v}(n) = v(n) * f_j(n) + b(n) * f_j(n) \quad (4)$$

où  $f_j$  est la réponse impulsionnelle du filtre  $F_j$ . Or, vu la relation (3) et en négligeant les retards de transmission, on peut écrire :

$$\tilde{v}(n_0) = a(n_0)r(n_0) + B(n_0) \quad (5)$$

où  $r(n_0) = s(n_0T_e) * f_j(n_0)$ .

Comme  $v(n)$  correspond par fenêtre à un vecteur aléatoire de  $\{\pm 1\}$ , les bits  $\alpha(n)$  sont identiquement distribués sur  $\{0, 1\}$ . Il est alors aisé de montrer que la probabilité d'erreur est donnée par [5] :

$$P_e = \frac{1}{2} \Pr\{B(n_0) > S + r(n_0)\} + \frac{1}{2} \Pr\{B(n_0) < S - r(n_0)\} \quad (6)$$

Les ddp considérées étant symétriques, le seuil  $S$  optimal minimisant la probabilité d'erreur est nul. Nous obtenons ainsi la relation donnant la probabilité d'erreur pour une ddp du bruit  $B(n)$  quelconque :

$$P_e = \int_{r(n_0)}^{+\infty} f_B(x) dx \quad (7)$$

Dans le cas gaussien, on retrouve le résultat classique :

$$P_e^G = \frac{1}{2} \operatorname{erfc} \left( \frac{r(n_0)}{\sigma_B \sqrt{2}} \right) \quad (8)$$

Dans le cas laplacien, comme

$$f_B(x) = \frac{1}{\sqrt{2}\sigma_B} \exp \left( -\frac{\sqrt{2}}{\sigma} |x| \right)$$

et pour  $r(n_0)$  positif, la probabilité d'erreur correspondante s'écrit :

$$P_e^L = \frac{1}{2} \exp \left( \frac{-\sqrt{2}}{\sigma_B} r(n_0) \right) \quad (9)$$

$\sigma_B^2$  étant la variance du bruit  $B(n)$ .

On remarque d'après l'expression de la probabilité d'erreur de l'équation (9), que  $P_e$  est décroissante en fonction du rapport  $\rho = \frac{r(n_0)}{\sqrt{2}\sigma_B}$ . Minimiser la probabilité d'erreur revient donc à maximiser  $\rho$ .

Notons par ailleurs, que  $\rho^2$  est homogène au RSB : en effet,  $r(n_0)$  est fonction du signal issu du codeur en ligne et filtré par le filtre  $F_j$ , à savoir  $v'(n)$ ;  $\sigma_B$  représente l'écart type de l'échantillon du bruit  $B(n)$ .

Sur la figure 8, nous avons représenté graphiquement les probabilités d'erreur exprimées par les équations (8) et (9) en fonction du RSB dans la plage de variation déterminée plus haut par simulation et pour les fichiers audio  $f_1$  et  $f_2$ . Nous remarquons ainsi, d'après ces courbes, que jusqu'à 2 dB environ, les performances sont équivalentes pour un bruit gaussien et laplacien. Par contre, au delà de cette valeur, la dégradation de la probabilité d'erreur apportée par le caractère laplacien s'accroît de plus en plus pour les forts RSB. Par exemple, pour un rapport signal à bruit de 11 dB, les performances se dégradent de  $4.10^{-7}$  à  $3.10^{-4}$ .

Dans une situation de tatouage audionumérique, il y a deux paramètres qui interviennent : le rapport signal à bruit qui varie par fenêtre selon le signal audio considéré et selon le type de modulation adoptée dans le schéma de tatouage d'une part, et d'autre part, la ddp du bruit variable par fenêtre qui évolue entre la loi gaussienne et la loi laplacienne.

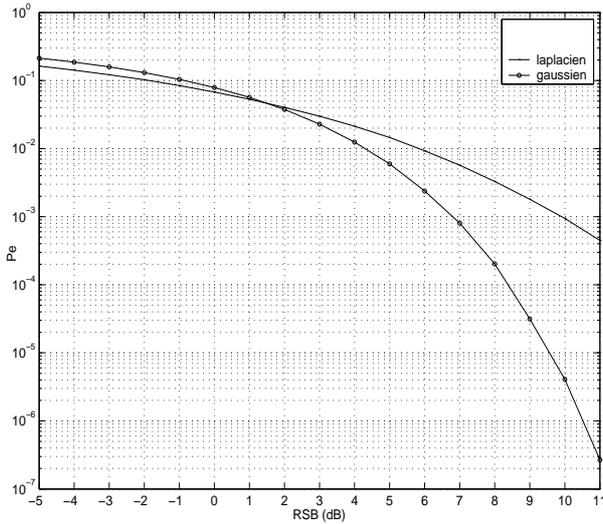


FIG. 8: Probabilité d'erreur pour une ddp laplacienne et gaussienne de  $B(n)$

## 5 Conclusion

Différents schémas équivalents d'une chaîne de tatouage audio numérique ont été proposés dans cette étude. Ces schémas ont permis de montrer qu'il s'agissait d'une chaîne de communications numériques non conventionnelle, où le rapport signal à bruit était relativement faible et où le bruit additif était non gaussien. L'accent a été mis sur ce caractère non gaussien (en particulier laplacien) et des premiers résultats que l'on est en mesure d'atteindre dans cette situation ont été trouvés.

Cette étude est un préalable à un choix adéquat de codeur canal dans un contexte de tatouage. Des résultats préliminaires ont déjà été obtenus [6], en généralisant le modèle du bruit (modèle gaussien généralisé) par fenêtre.

## Références

- [1] L. Boney, A.H. Tewfik et K.N. Hamdy. *Digital watermarks for audio signals*. IEEE Int. Conf. on Multimedia Computing Systems, Japan, June 1996.
- [2] L. de Campos Teixeira Gomez. *Tatouage de signaux audio*. Thèse de l'Université Paris 5, mai 2002.
- [3] S. Larbi, M. Jaïdane et N. Moreau. *Audiowatermark detection for all-pass pirate attack: hybrid blind equalization/Wiener deconvolution approach*. Eusipco2002, Toulouse.
- [4] M.H. Hayes. *Statistical Digital Signal Processing and Modeling*. Ed. Wiley, 1996.
- [5] A. Glavieux et M. Joindot. *Communications numériques*. Addison-Wesley, 1994.
- [6] S. Saïed. *Amélioration de la détection de tatouage audio numérique par codage canal. Choix d'une métrique adaptée au contexte non gaussien*. Rapport interne, Unité Signaux et Systèmes, ENIT, Tunis, Mars 2003.