AN EFFICIENT LOW BIT-RATE INFORMATION EMBEDDING COSTA BASED SCHEME USING A PERCEPTUAL MODEL

Claude Delpha, Brice Djeumou Touko, Abdellatif Zaidi, Pierre Duhamel

{delpha, djeumou, zaidi, pduhamel}@lss.supelec.fr
Laboratoire Signaux et Systèmes, Supelec/CNRS/Univ. P-Sud, 3 rue J. Curie - 91190 Gif/Yvette

ABSTRACT

In this paper, we propose an audio watermarking scheme based on the Scalar Costa Scheme specifically calibrated with a perceptual model allowing to increase the embedding power. For our study, this scheme is designed to be efficient for low bit-rate embedding with sufficient robustness to channel degradations. We present here the main characteristics of our scheme and the way for introducing perceptual models in such a Costa based watermarking scheme without introducing any noticeable artifacts. An evaluation of the robustness of the embedding system is also theoretically discussed by the way of the main channel attack : additive noise from low to high level. We illustrate its relevance in practice using Monte Carlo simulations.

Index Terms— Multimedia systems, Security, Copyright protection, Digital Watermarking, Channel coding

1. INTRODUCTION

Digital technologies development has risen a challenging problem: ownership on multimedia data. Information embedding techniques, specially digital watermarking, have been proposed as potential solutions. In digital watermarking context, a signal, generally weak (watermark: proof of ownership), is transmitted within another signal, generally stronger (host signal: multimedia data). Three conflicting requirements have to be calibrated in evaluating watermarking schemes performances: the *embedding rate*, the *robustness* to legal or malicious degradations (attacks) and the *imperceptibility* of the watermark. In this study, we concentrate on audio watermarking for a secured diffusion of music on mobile phones: we relax the embedding rate constraint (few information to be embedded) and tune the process to obtain good imperceptibility and wide robustness to attacks.

As Scalar Costa Scheme (SCS) [1], [2] is one of the most promising solutions that have been thoroughly studied in recent years, we have focused our work on this kind of scheme. In such schemes, watermarking can be considered as the transmission of a message $\mathbf{m}$ (watermark) into a host signal $\mathbf{x}$ trough a channel (noisy host data) which can be corrupted by an additive noise $\mathbf{v}$ (Fig.1).

With such blind watermarking schemes with side information at the encoder no perceptual model can be directly used due to the few information at the decoder. In these cases the watermark embedding power is then limited to keep the imperceptibility, inducing limitations for the system robustness. Thus, several kind of attacks can widely affect the watermark detection. We show that, in the case of high level of additive noise induced for example by an Advance Audio Coding (AAC) compression, such perceptual model is necessary to increase the watermark power and allow a good detection. Few studies have been done combining Costa and perceptual models mainly for image watermarking cases [3], [4]. Solanki, in [4], proposes two approaches of this problem in the case of JPEG compression: one based on entropy thresholding and the other one based on selecting non zero DCT coefficients. Both approaches gives good results without embedding rates constraint and theoretical study. We propose here a specific way to use particular perceptual model in audio applications and describe the theoretical solution for this study specifically for low bit-rate embedding.

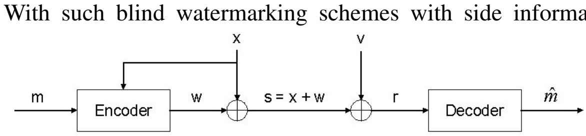

Fig. 1. Watermarking scheme with side information at the encoder.

In this paper, we first describe in Section 2 the encoding and decoding processes designed with different channel coding techniques applied on top of egger's SCS to perform minimum robustness to mid level Additive White Gaussian Noise (AWGN) in low rate embedding cases. We also show the robustness limitations without psycho-acoustic model. In Section 3, we modify the scheme and show how it is possible to introduce a perceptual model and manage with a particular detection process to improve the performances. For both, theoretical analysis and Monte Carlo simulations are given. Finally, concluding remarks are stated in Section 4.

2. ORIGINAL WATERMARKING SCHEME

For our study based on the *Costa's scheme* (Fig.1), if $\mathbf{x}$ is a $\mathbf{L}_x$-sample size of the host signal with power P and m a message to be embedded into the host signal through the embedded signal $\mathbf{w}$, we are able to obtain a watermarked signal $\mathbf{s} = \mathbf{x} + \mathbf{w}$ where $\mathbf{w}$ is zero-mean and power constrained ($\sigma_w^2 \leq P$). When $\mathbf{s}$ is attacked and if this attack is represented by an additive noise $\mathbf{v}$ of power N. The decoder receives the attacked data $\mathbf{r} = \mathbf{s} + \mathbf{v}$. Considering the host signal $\mathbf{x}$ Gaussian and the attack an AWGN, *Costa* have proposed an encoding process [1] that theoretically achieves the AWGN capacity $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$, which is the capacity of a communication scheme with side information available at both the encoder and the decoder. However, Costa's theoretical scheme is not feasible in practice. Based on dithered scalar quantizers, Eggers [2] proposed an efficient suboptimal scheme, referred as SCS. In the case of Gaussian noise attacks, different channel coding techniques have been applied but separately on the SCS.

For this work, we combine the use of several coding methods. We rely on Spread Transform (ST) coding to enforce the embedded signal. This ST coding is applied in a different way with comparison to that usually considered in, for example, [2] and [5]. In our implementation, since ST coding inevitably reduces the power of the host signal in the transform domain, the high resolution

The authors would like to thank the RNRT project SDMO (Secured Diffusion of Music on mObiles) for funding.

assumption is respected by limiting the spreading factor. These limitations are partially compensated by using a slight variation of ST: Repetition Coding (RC). This is inline with the use of RC as means of strengthening the signal at high noise levels in conventional communication. A third layer of channel coding is added to further strengthening the embedded signal: Turbo Coding (TC). TC is well suited for the application considered in this work, due to its convenience for transmission over channels where fading-like attacks may occur. For our study, we consider specific encoding and decoding functions (Fig.2) to obtain a performed robust and low rate embedding scheme. With the specific encoding and decoding functions, we provide a performed theoretical analysis of the error probability in the cases of AWGN attacks.

We denote the unspreading function by $f(\cdot)$ and the spreading

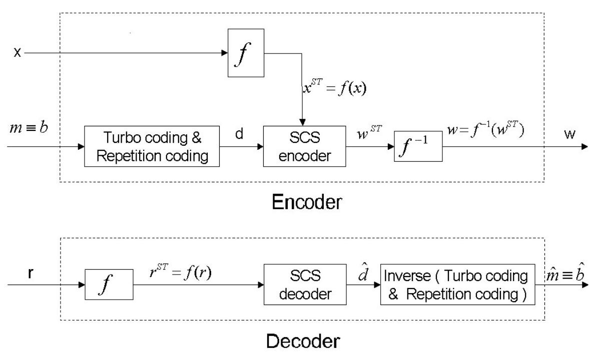

Fig. 2. Encoder and decoder scheme against AWGN attacks.

function by $f^{-1}(\cdot)$. For this work, the unspreading function is based on Spread Transform (ST), the binary representation of the message $\mathbf{b}$ is not embedded directly into the host data $\mathbf{x}$ but into its projection $\mathbf{x}^{\mathrm{ST}}$ onto a zero-mean random sequence $\mathbf{t}$ such that $x_j^{\mathrm{ST}} = \sum_{i=(j-1)\tau+1}^{j\tau} x_i \mathbf{t}(i)$. To come back in the temporal domain after obtaining the watermark in the ST domain, $\mathbf{w}^{\mathrm{ST}}$ is spread onto another zero-mean random sequence $\mathbf{t}'$ such that $w_i = w^{\mathrm{ST}}_{\lfloor \frac{i}{\tau} \rfloor} \mathbf{t}'(i)$. The spread transform reduces the effective noise power in the ST domain where the decoding is done. For example, for an AWGN attack, the power of its noise is divided by the spreading factor τ in the ST domain. By the same way, in the ST domain, the host data power Q is also divided by τ. In our schemes, to obtain the expected gains, spreading sequence $\mathbf{t}'$ and unspreading sequence $\mathbf{t}$ are normalized as follows: $\hat{\sigma}^2_{\mathbf{t}'} = 1$ and $\hat{\sigma}^2_{\mathbf{t}} = \left(\frac{1}{\tau}\right)^2$. We now provide a theoretical analysis of the error probability of this scheme before analyzing simulation results.

2.1. Error probability analysis

To find the expression for the error probability of this watermarking scheme, at first, we don't take into account the TC in the theoretical analysis. Some assumptions has been made before going through calculation: High Resolution ($Q \gg (P, N)$ or the Probability Density Function (PDF) of the host data is uniform on at least two quantizer steps) and decoding by hard decision. We suppose $\mathbf{d}$ to be an Independent Identically Distributed (IID) sequence so the error probability is independent from the bit sent (0 or 1). Then we take $d_k = 0$ and we will suppose that the *k*th host data element x_k is known ($x_k = x_0$). We denote $p_u(u)$ as the PDF of a random variable u, $\mathbf{p}_e$ for the error probability of the system without any channel coding, $\mathbf{p}_e^{ST}$ for the error probability with spread transform, $\mathbf{p}_e^{ST,r}$ for the error probability with spread transform and repetition coding which is equal to the error probability $\mathbf{p}_e^{th}$ of our partial watermarking scheme. The expression for $\mathbf{p}_e$ corresponds to the case where no channel coding is applied and the embedding rate is $1bit/element$. We denote $\mathbf{p}_e^k$ the error probability corresponding to the kth host data element $x_k = x_0$. At the encoding stage, the quantization error of the *k*th host data element x_k is

$$q_k = \mathcal{Q}_\Delta \left\{ x_k - \Delta \frac{d_n}{D} \right\} - \left(x_k - \Delta \frac{d_k}{D} \right) = \mathcal{Q}_\Delta \left\{ x_k \right\} - x_k$$

where $\mathcal{Q}_\Delta$ denotes the uniform scalar quantifier of step Δ. We have $s_k = x_k + w_k = x_k + \alpha q_k$

At the decoding stage, $r_k = s_k + v_k$ and

$p_r(r|d_k = 0, s_k = s_0) = p_s(r|d_k = 0, s_k = s_0) \otimes p_v(r)$

$p_s(r|d_k = 0, s_k = s_0)$ is a dirac at s_0 then $p_r(r|d_k = 0, s_k = s_0)$ is a Gaussian centered at $s_k = s_0$ with N as variance. We have

$$\begin{aligned} q_k^r &= \mathcal{Q}_\Delta \left\{ r_k \right\} - (r_k) \\ &= \mathcal{Q}_\Delta \left\{ x_k + \alpha q_k + v_k \right\} - (x_k + \alpha q_k + v_k) \end{aligned}$$

On the other hand $x_k + q_k = l \cdot \Delta$, hence

$$q_k^r = \mathcal{Q}_\Delta \left\{ (\alpha - 1) q_k + v_k \right\} - ((\alpha - 1) q_k + v_k) \quad (1)$$

The decision $\hat{d}_k$ on d_k is made from the quantization error q_k^r of the kth received data sample r_k. As $d_k = 0$ and a decoding by hard decision is applied, the good decision is taken only if $-\frac{\Delta}{4} \leq q_k^r \leq \frac{\Delta}{4}$. The decoder acts like it quantizes the noise v shifted forward in $(\alpha - 1) q_k$. The expression of $\mathbf{p}_e^k$ is

$$\begin{aligned} \mathbf{p}_e^k &= \sum_{i \geq 0} \left[p_v \left(v \in \left[-b_i^{max} + \delta_k, \ -b_i^{min} + \delta_k \right] \right) \right] \\ &+ \sum_{i \geq 0} \left[p_v \left(v \in \left[b_i^{min} + \delta_k, \ b_i^{max} + \delta_k \right] \right) \right] \end{aligned} \quad (2)$$

where $b_i^{min} = i\Delta + \frac{\Delta}{4}$, $b_i^{max} = i\Delta + \frac{\Delta}{4} + \frac{\Delta}{2}$ and $\delta_k = (1 - \alpha) q_k$ standing for the offset introduced by the SCS encoder factor α. $p_v(v)$ is the Gaussian PDF of the zero-mean random variable v with variance N. To obtain $\mathbf{p}_e$, we calculate the average of $\mathbf{p}_e^k$ over all possible values of x_k. As $\mathbf{x} - \mathbf{q} - \mathbf{w}$ is a Markov chain, the average of $\mathbf{p}_e^k$ over x is equivalent to the average of $\mathbf{p}_e^k$ over q. $\mathbf{p}_e$ is then

$$\begin{aligned} \mathbf{p}_e &= \mathrm{E}_{\mathrm{p}_q(q)} \left[\mathbf{p}_e^k \left(q, \alpha, \Delta \right) \right] \\ &= \int_{-\Delta/2}^{+\Delta/2} \frac{1}{\Delta} \mathbf{p}_e^k \left(q, \alpha, \Delta \right) dq \end{aligned} \quad (3)$$

where $\mathrm{p}_q(q) = \frac{1}{\Delta}$ is the uniform PDF of the quantization error since the High Resolution assumption is verified. In our simulations, $\mathbf{p}_e$ is approximated by

$$\hat{\mathbf{p}}_e = \lim_{N \to +\infty} \frac{1}{N} \sum_{k=0}^{N} \mathbf{p}_e^k (q_k, \alpha, \Delta) \quad (4)$$

where $q_k = -\frac{\Delta}{2} + k \frac{\Delta}{N}$.

α and Δ take the most convenient values for a good channel capacity [2]: $\alpha = \alpha_{SCS} = \sqrt{\frac{P}{P + 2.71N}}$ and $\Delta = \Delta_{SCS} = \sqrt{12\left(P + 2.71N\right)}$. By including the channel codings, we obtain the following results for the error probability:

Spread Transform: The noise power is reduced by the spread factor τ in the ST domain but this noise is still Gaussian ($N_{ST} = \frac{N}{\tau}$). So the error probability expression is unchanged but new α and Δ values has to be used: $\alpha' = \alpha_{SCS} = \sqrt{\frac{P}{P + 2.71\frac{N}{\tau}}}$, $\Delta' = \Delta_{SCS} = \sqrt{12\left(P + 2.71\frac{N}{\tau}\right)}$. It is important to observe that the rate is divided by the spreading factor τ. We obtain

$$\mathbf{p}_e^{ST} = \mathbf{p}_e \left(\alpha', \Delta' \right) \quad (5)$$

Repetition coding: If φ is the repetition factor, each bit of $\mathbf{b}$ is repeated φ times. The global good decision is taken on the kth bit b_k of $\mathbf{b}$ if an absolute majority of good decisions is taken on the φ repetition bits. The error probability is then

$$\mathbf{p}_e^{ST,r} = \sum_{i=0}^{k} \binom{\varphi}{i} \left[\mathbf{p}_e^{ST}\right]^{\varphi-i} \left[1-\mathbf{p}_e^{ST}\right]^{i} \tag{6}$$

with $k = \left\lfloor \frac{\varphi}{2} \right\rfloor$. Finally, we have: $\mathbf{p}_e^{th} = \mathbf{p}_e^{ST,r}$.

2.2. Performance analysis and discussion

We implement our watermarking scheme and also the theoretical model that we obtained previously, and for both, we compute the Bit Error Rate (BER) for different Watermark to Noise Ratio (WNR). For these two implementations we first do not include TC which will allow to increase the robustness of our scheme. As the watermark may not be perceptible in the considered audio data, we first determine the Document to Watermark Ratio (DWR) lower limit value to ensure this imperceptibility condition. For simulations, the main settings are: embedding rate R=100 bits/second of signal, sampling frequency of the host signal is 22050 Hz, spreading factor $\tau = 10$, DWR=45dB and WNR taken from -15 to -7 dB. We obtain on Fig.3 the BER evolution for both empirical and theoretical studies : the BER values decreases when the noise level is low.
As shown on Fig.3, theoretical BER ($\mathbf{p}_e^{th}$) matches with empirical BER ($\mathbf{p}_e^{emp}$). This illustrates the relevancy of this theoretical model. It could then be used to preview an upper bound for the BER of our watermarking system since the partial scheme do not include TC.
We now compute the BER of the watermarking scheme including

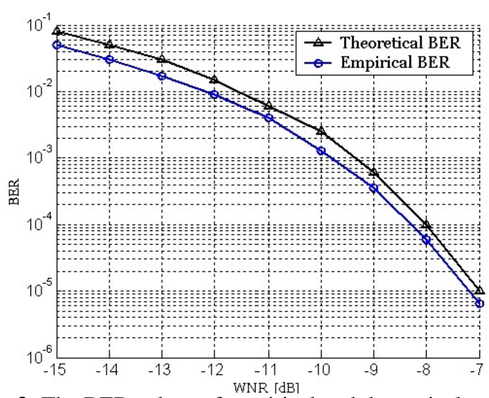

Fig. 3. The BER values of empirical and theoretical model

all channel codings : ST, TC, and RC. The settings still unchanged except the rate and the spreading factor which are now varying. The results are summarized in Fig.4. Acceptable BER values are obtained. These BER values are lowered with the reduction of the embedding rate or with the increase of τ. Within the framework of our project, the rate value needed is $\frac{32}{15}$bps. It corresponds to the rate that widely allows the embedding of a reference in a music song of average length. As this rate is widely under the value that has been used for our computations, best robustness performance can be expected. Nevertheless, we can note that the BER increase when the WNR decrease : for high levels of noise (WNR $< -\mathbf{11}$dB) more detection errors can be obtained (BER $\approx$3.10^{-1}). This can give problems, for example, in the case of compression of the considered audio signal which can introduce high levels of additive noise. By using the scheme described in the previous subsection the observed WNR value is around $-\mathbf{31}$dB for the chosen compression technique AAC. So acceptable performance cannot be expected even if we significantly reduce the embedding rate.

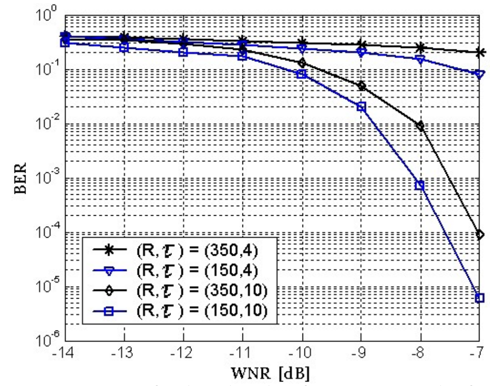

Fig. 4. BER vs WNR for low levels of AWGN attacks for two watermarking rates (R) and spreading factor (τ) combinations

3. EMBEDDING SCHEME WITH PERCEPTUAL MODEL

Due to the lack of robustness of our initial scheme with the presence of high additive noise levels, it is then necessary to modify the embedding scheme to increase the watermark embedding strength. This change must allow us to shift up the WNR value in the area where acceptable performance is reachable (close to $-\mathbf{10}$dB). To solve this problem, we develop a simple perceptual model that allows us to increase the watermark power (DWR) in the developed SCS based encoder without any noticeable artifact on the watermarked data $\mathbf{s}$. With this method, the watermark is subjected to a simple psychoacoustic model behaving as a filtering function which allows its spectral density to look like the host signal's. This solution allows us to take the advantage of the masking phenomenon. The filter used has a frequency response which looks like the host signal spectral density. A gain of $\mathbf{20}$dB can be achieved on the DWR and the WNR without any perceptible artifact. We are then able to scale down our DWR value from $\mathbf{45}$dB to $\mathbf{25}$dB. It is the threshold value for which the watermark still unnoticeable in the audio considered signal. The coefficients of this filter must be normalized such that the watermark power stays unchanged after the filtering. Fig.5 summarizes the new encoder scheme.
The harmonic response of the introduced filter $\mathbf{H}(f)$ is obtained

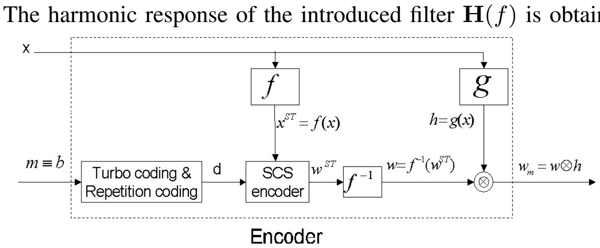

Fig. 5. The encoder scheme against high levels of AWGN attacks

from the host signal spectral density $\mathbf{X}(f)$ by the following expression: $\mathbf{H}(f) = \frac{|\mathbf{X}(f)|^{\beta}}{\text{var}(|\mathbf{X}(f)|^{\beta})}$ where β is the parameter which commands the level of filtering. In the temporal domain, the filter, supposed to be a N-order filter, is represented by $\mathbf{h}(z) = \sum_{i=0}^{N} h_i z^{-i}$. Since the watermark is an IID sequence and the filtering must not modify the watermark power, the following relation is checked: $\sigma_{w_m}^2 = P \sum_{i=0}^{N} h_i^2$ with $\sum_{i=0}^{N} h_i^2 = 1$. The filter coefficients depend on the parameter β of the filtering. Particularly, h_0 decreases as β increases. It's observed that the case $h_0 = 1$ corresponds to previous case (no filtering) described in section2. To optimize the encoding step, the host signal is divided in windows and and the message is embedded only in the windows which power is over a minimal threshold. The masking phenomenon is then improved.

3.1. Error probability analysis with perceptual model

As we know, the AWGN is more harmful than a compression noise at equivalent power. For our study, the compression noise is approximated like Gaussian data and this approximation gives an upper bound of the error probability for this kind of attacks. As for the previous study, the theoretical analysis is drawn up without considering TC. The same assumptions are supposed to be verified. We use the same notations and $\mathbf{n}$ denotes the compression noise. The present case differs from the previous because two noises are taken in account here.

At the encoding stage, $q_k = \mathcal{Q}_\Delta\{x_k\} - x_k$ and $w_k = \alpha q_k$. After the filtering, $w_m^k = \mathbf{h}^T \mathbf{w} = \alpha \sum_{i=0}^{N} h_i q_{k-i}$. The watermarked sample is $s_k = x_k + w_m^k$.

At the decoding stage, $r_k = s_k + n_k$. We have: $q_k^r = \mathcal{Q}_\Delta\{r_k\} - (r_k)$. On the other hand $x_k + q_k = l\Delta$. Hence

$$q_k^r = \mathcal{Q}_\Delta\left\{(\alpha-1)\,q_k + \alpha\sum_{i=0}^{N} h_i q_{k-i} - \alpha q_k + n_k\right\} - \left((\alpha-1)\,q_k + \alpha\sum_{i=0}^{N} h_i q_{k-i} - \alpha q_k + n_k\right)$$

If $\mathbf{z}$ denotes the noise due to the filtering under power constraint σ_z^2, its elements follow the expression: $z_k = \alpha \sum_{i=0}^{N} h_i q_{k-i} - \alpha q_k$ and, as $\mathbf{w}$ is an IID sequence, $\sigma_z^2 = 2\,(1-h_0)\,P$ and the relation $\sigma_z^2 \leq 2P$ always holds. The filtering brings much noise than watermark for $h_0 < 0.5$. For the level of the filtering, we have to find β values such that $h_0 > 0.5$. We hence implemented our scheme and verified that for $\beta = \frac{2}{\ln 10}$, on all data which have been watermarked, the noise power of the filtering is under the watermark power.

The filtering noise $\mathbf{z}$ is particular because, in the ST domain we do not benefit from the gain due to the spread factor τ as for the AWGN. This is because $\mathbf{z}$ is a combination of elements of the watermark sequence. If we suppose that the filtering noise and the encoding noise are independent, the effective WNR is given by: $\text{WNR} = 10\log\left(\frac{P}{\sigma_n^2+\sigma_z^2}\right)$. We take the watermark to compression noise ratio as $\text{WCR} = 10\log\left(\frac{P}{\sigma_n^2}\right)$ and the watermark to filtering noise ratio $\text{WFR} = 10\log\left(\frac{P}{\sigma_z^2}\right)$ (we have $\text{WFR}_{\text{ST}} \approx \text{WFR}$). The observed WCR values are around -31dB and the minimum value of the WFR is -3dB. The filtering noise $\mathbf{z}$ can be disregarded in the calculation of the SCS parameters for τ values which give $\text{WCR}_{\text{ST}} \ll \text{WFR}$. This implies $\text{WNR}_{\text{ST}} = \text{WCR}_{\text{ST}}$. For τ values which gives $\text{WCR}_{\text{ST}} \approx \text{WFR}$, $\mathbf{z}$ and $\mathbf{n}$ must be taken in account. For other values of τ, $\text{WNR}_{\text{ST}} = \text{WFR}$ and the system looses its efficiency because we no longer benefit from the spread transform but we loose in RC. $\mathbf{z}$ and $\mathbf{n}$ are approximated by Gaussian. For an upper bound of the error probability, eq. (2), (3), (4), (5) and (6) still hold with new parameter values obtained from the characteristics of this scheme.

3.2. Performance analysis and discussion with perceptual model

With these modifications in the encoding function including a perceptual model we have for the DWR=$\mathbf{25}$dB, the WNR value is around $-\mathbf{11}$dB. It's then possible to have settings which will give good performances. For simulations, in the calculation of α and Δ values, σ_z^2 is unknown. It's estimated by encoding the original host signal alone with the AAC model and then calculate the noise power resulting. It's a good approximation since $Q \gg P$. Fig.6 gives the BER for two different watermarking embedding rate: 1.5 bits per second of signal and 0.75 bit per second of signal. The sampling frequency of the host signal is unchanged.

Good performances are obtained especially for the rate 0.75 bps.

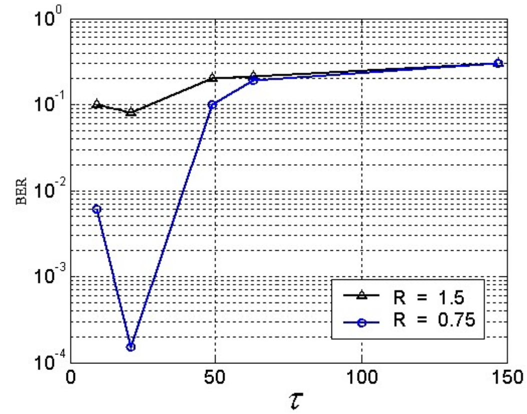

Fig. 6. BER vs spreading factor (τ) in the case of high noise levels for two watermarking rates (R)

The optimal value of the parameter τ is found. It corresponds to $\text{WCR}_{\text{ST}} \approx$ WFR. Over this optimal value of τ, we loose in repetition coding without getting an additive gain on the WNR_{ST}. That's why performances fall down.

4. CONCLUSION

The audio watermarking scheme developed here is based on the SCS and designed to have high performance specially in a low bit-rate information embedding applications. For this work the watermark power is chosen under the constraint of inaudibility in the audio signal. We have shown that by using such a scheme with properly designed and calibrated encoder and decoder functions, combining several types of channel coding methods (spread spectrum, turbo coding and repetition coding) and a perceptual model, good robustness results to AWGN (low or high level) are obtained. For the highest level of noise for example induced by AAC Compression, we have shown how it is necessary and possible to calibrate a perceptual model to be used in Costa based schemes. A gain of 20dB on WNR is obtained with this model allowing to increase the watermark embedding power and obtain a good detection. This watermarking scheme, specifically tuned to watermark audio data, is efficient and enough robust for low and high AWGN levels. For future works, its robustness can be evaluated for other types of attacks.

5. REFERENCES

[1] M.H.M. Costa. "Writing on dirty paper". *IEEE Trans. on Information Theory*, 29(3), May 1983.

[2] J. Eggers and B. Girod. "Informed watermarking". Kluwer Academic Publishers, 2002.

[3] Q. Li, G. Doërr and I. J. Cox. "Spread transform dither modulation using a perceptual model". *IEEE MMSP*, Oct. 2006.

[4] K. Solanki, N. Jacobsen, S. Chandrasekaran, U. Madhow, and B.S. Manjunath. "High-volume data hiding in images: introducing perceptual criteria into quantization based embedding". *IEEE Int. Conf. Proc. ICASSP 02*, 4, May 2002.

[5] B. Chen and G. W. Wornell. "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding". *IEEE Trans. on Information Theory*,47(4), May 2001.