Practical Watermarking scheme based on Wide Spread Spectrum and Game Theory

Stéphane Pateux * , Gaëtan Le Guelvouit

IRISA / INRIA-Rennes
Campus de Beaulieu
35042 Rennes Cedex, FRANCE.

Abstract

In this paper, we consider the implementation of robust watermarking scheme for non i.i.d. Gaussian signals and distortion based on perceptual metrics. We consider this problem as a communication problem and formulated it as a game between an attacker and an embedder in order to establish its theoretical performance. We first show that known parallel Gaussian channels technique does not lead to valid practical implementation, and then propose a new scheme based on Wide Spread Spectrum and Side Information. Theoretical performances of this scheme are established and shown to be very close to the upper bound on capacity defined by Parallel Gaussian channels. Practical implementation of this scheme is then presented and influence of the different parameters on performance is discussed. Finally, experimental results for image watermarking are presented and validate the proposed scheme.

Key words: Watermarking, Information theory, Game theory, Channel Coding with Side information

1 Introduction

A lot of effort has been dedicated over the last years for designing practical watermarking systems. The approaches were often viewing the media content as noise from the watermark detection perspective, hence regarding watermarking as a form of wide spread spectrum communication (WSS) with various forms of distortion measures (MSE or weighted MSE) and of channel characterizations [1], [2], [3]. The authors in [4] suggest to take into account the

* Corresponding author: Stephane.Pateux@irisa.fr

perceptual properties of the content and to embed in perceptually significant frequency components. Other approaches based on WSS and exploiting the perceptual sensitivity of the host data can also be found in [5], [6], [7]. However these techniques are based on empirical assumptions.

Attacks have often been modeled as the addition of White Gaussian noise (AWGN) [8], [9], and more recently as linear filtering plus white or colored additive noise [10], [11]. It has been shown in [12] and in [13] that expressing the problem of watermarking as a problem of communication with side information leads to optimal performances. Costa [14] has indeed shown that in the context of attacks modeled by AWGN, the capacity is not dependent on the cover signal. However the solution proposed, known as the *Ideal Costa Scheme* (ICS), requires very large codebooks, hence is not realistic. Different approaches have then been proposed to reach performances of Costa's scheme using structured codebooks; Scalar Costa Scheme (SCS) [15], syndrome based coding [16] or more recently trellis with multiple paths [17]. Dithered quantization techniques [12], [18] may also be seen as techniques exploiting side information. Most of these schemes are defined for i.i.d. signals which is not a valid assumption for usual considered signals. Techniques based on parallel-Gaussian channels [19],[20] have then been proposed to deal with such non i.i.d. Gaussian signals. However practical implementation of Parallel-Gaussian necessitates to know the original signal [21] and does not lead to a valid implementation (see discussion in section 2).

This paper deals with the robust data hiding problem, assuming a blind (the extraction system has no knowledge of the host signal) and symmetric (same private key for embedding and extraction) system. In this paper we consider this problem as a communication problem: one seeks the maximum hiding capacity (or rate of reliable transmission) over any hiding and attack strategies. The rate obviously depends on the perceptual distortion levels considered admissible and on the watermark channel (or attack scenarios) characterization. We especially present a technique based on Wide Spread Spectrum and Side Information facing Scaling and Additive White Gaussian Noise (SAWGN) optimized by considering Game Theory formalism in order to define performance limits. Practical implementation as well as efficiency of the proposed technique are further presented.

This paper is organized as follows. In section 2, general consideration about watermarking of non i.i.d. signals is first presented as well as limitation of the previously proposed techniques. In section 3, we then present an optimized watermarking scheme based on Wide Spread Spectrum and Side Information and discuss about its practical implementation. In section 4, experimental results are shown for image watermarking. Finally section 5 concludes this work.

2 Watermarking of non i.i.d. signals

Most of the techniques proposed in watermarking are assuming i.i.d. signals, however this assumption is rarely valid. For example, when performing embedding in a transform domain, coefficients are generally not i.i.d. (e.g. for images, low frequency coefficients have higher energy than high frequency coefficients). In [10], author then showed that in order to resist to filtering attacks, power spectrum of the watermark should be proportional to the power spectrum of the host signal, what they called the PSC condition. In [11],[22], optimizations of watermarking techniques based on wide spread spectrum have been proposed for non i.i.d. Gaussian signals considering Scaling and Additive White Gaussian Noise. While exploiting statistical properties of the host signal, those techniques are still not optimal since they do not exploit the realization of the host signal.

In [19],[20], theoretical analysis of watermarking for non i.i.d. signals have been carried. Capacity bounds have been derived by considering a game between an attacker and the embedder. First for i.i.d. Gaussian signals, capacity can be expressed as:

$$C = \frac{1}{2}\log_2\left[1 + \frac{D_1}{D_2 - D_1}(1 - \frac{D_2}{\sigma_X^2})\right] \quad (1)$$

Where D_1, D_2 corresponds respectively to the embedding distortion and to the attack distortion [1]; σ_X^2 corresponds to the variance of the host signal X. Optimal strategies for the embedder and the attacker take the following forms:

$$\begin{cases} Y &= \gamma_1(X + W) \\ Y' &= \gamma_2(Y + \delta) \end{cases} \quad (2)$$

with

$$\begin{cases} \gamma_1 &= \frac{\sigma_X^2 - D_1}{\sigma_X^2} \\ \gamma_2 &= \frac{\sigma_X^2 - D_2}{\sigma_X^2 - D_1} \\ \sigma_\delta^2 &= (D_2 - D_1)\frac{\sigma_X^2 - D_1}{\sigma_X^2 - D_2} \end{cases} \quad (3)$$

where X corresponds to the host signal, W to the watermark (that is defined taking into account Side Information), Y to the watermarked signal, Y' to

[1] In [20], capacity formulation differs since author considered the attacker using a measure of distortion between the attacked signal and the watermarked one - noted as type X constraints in [19]. As discussed in [19], distortion between the attacked signal and the original one - type S constraints, is more suited.

the attacked signal and δ is a Gaussian noise. Eqns. (2) show that considering simple additive Gaussian noise attack is not sufficient and that Scaling and Additive White Gaussian Noise attacks should rather be considered. It can be shown that γ_1 factor corresponds to the multiplicative factor γ that would have been used when using Wiener filtering to reduce the impact the added noise considered being W [2]. Embedding can then be considered as classical side information technique followed by Wiener filtering. Further, it can also be shown that γ_2 scaling factor has the effect of performing Wiener filtering when considering the addition of noises W and δ [3].

When considering non i.i.d. Gaussian signals, Parallel Gaussian channels are introduced and global capacity is estimated as the sum of the capacities over these different channels. If K channels are considered, and if we note $d_{1k}, d_{2k}, \sigma_k^2, r_k$ respectively the embedding distortion, the attack distortion, the variance of the host signal, and the ratio of occurrence for the k^{th} channel, the global capacity is defined as:

$$C = \max_{d_1} \min_{d_2} \sum_{k=1}^{K} r_k \Gamma(\sigma_k^2, d_{1k}, d_{2k}) \tag{4}$$

where Γ expresses the capacity of a given channel depending on its characteristics (see Eqn. (1)). The max-min operation represents the game between the attacker and the embedder for given constraints of embedding distortion D_1 and attack distortion D_2:

$$\begin{cases} \sum_{k=1}^{K} r_k d_{1k} \leq D_1 \\ \sum_{k=1}^{K} r_k d_{2k} \leq D_2 \end{cases} \tag{5}$$

Thus following this result, practical embedding should be performed on separate channels as proposed in [21]. However this technique needs to define the set of parallel i.i.d. channels. This definition of the channels may be disturbed when signal is being attack, and actually, work in [21] relies on the knowledge of the host signal in order to retrieve the different channels and their properties.

Moreover when considering practical implementation, embedding on separate channels does not guaranty to retrieve all the embedded information. Effec-

[2] Wiener multiplicative factor is defined as $\gamma = \frac{\sigma_X^2}{\sigma_X^2+\sigma_W^2}$. Distortion after Wiener filtering will be $D_1 = \frac{\sigma_X^2 \sigma_W^2}{\sigma_X^2+\sigma_W^2}$. Expressing σ_W^2 in function of D_1 and substituting it in the expression of γ leads to $\gamma = \gamma_1$.

[3] To this extent, we write $Y' = (\gamma_1.\gamma_2) \times [X + W + \frac{\delta}{\gamma_1}]$. Then it is easy to verify that $(\gamma_1.\gamma_2)$ corresponds to the multiplicative factor of the Wiener filter when considering X subject to the noise $(W + \frac{\delta}{\gamma_1})$.

tively, embedding strategy have been defined in order to respond to the optimal strategy of the attacker, but what happens when the attacker does not perform this "optimal" attack? As an example, instead of spreading its attack distortions d_{2k} according to the solution of the game defined in Eqn. (4), let us consider the case where the attacker decides to put more distortion on some channels and less on others. Then on the channels that have higher distortion, embedded information may not be fully retrieved. Channels being less distorted will not allow to retrieve the lost information due to the separate embedding/extraction on the channels! This example shows that when considering separate embedding/extraction on each channel, we do not have in fact a Nash equilibrium.

One way to deal with this problem is to exploit information from all the channels. That is, considering embedding and extraction globally on all channels. In line with Side Information technique, we will consider that a given message is associated with a set of code-vectors. Extraction will then consist in looking for the closest code-vector to the degraded watermarked signal $\underline{\mathbf{y}'}$ among all the possible ones (in terms of probabilistic distance). Considering SAWGN attacks, we have for the i^{th} coefficient:

$$\begin{cases} y_i = x_i + w_i = \sigma_{w_i}\beta u_i + v_i \\ y'_i = \gamma_i y_i + \delta_i \end{cases} \tag{6}$$

Where notation $\sigma_{w_i}\beta u_i + v_i$ is inspired from the side information interpretation proposed in appendix A; u_i represents the code-vector associated to the embedded message, v_i represents the remaining noise of the host signal (Gaussian noise independent of code-vector $\underline{\mathbf{u}}$), and σ_{w_i} is introduced in order to locally adapt the strength of the watermark to robustness and perceptual distortion. When searching for the embedded message, hypothesis testing between two code-vectors u_0 and u_1 will be performed. That is looking for the maximum value among $p(\underline{\mathbf{y}'}|u_0)$ and $p(\underline{\mathbf{y}'}|u_1)$:

$$\begin{array}{lll} & p(\underline{\mathbf{y}'}|u_0) & <> p(\underline{\mathbf{y}'}|u_1) \\ \Longleftrightarrow & \Pi_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp[-\frac{(y'_i - \gamma_i \sigma_{w_i}\beta u_{0,i})^2}{2\sigma_i^2}] & <> \Pi_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp[-\frac{(y'_i - \gamma_i \sigma_{w_i}\beta u_{1,i})^2}{2\sigma_i^2}] \\ \overset{\log(.)}{\Longleftrightarrow} & -\sum_{i=1}^{n} \frac{(y'_i - \gamma_i \sigma_{w_i}\beta u_{0,i})^2}{2\sigma_i^2} & <> -\sum_{i=1}^{n} \frac{(y'_i - \gamma_i \sigma_{w_i}\beta u_{1,i})^2}{2\sigma_i^2} \\ \Longleftrightarrow & \sum_{i=1}^{n} \frac{\gamma_i \sigma_{w_i}}{2\sigma_i^2} y'_i . u_{0,i} & <> \sum_{i=1}^{n} \frac{\gamma_i \sigma_{w_i}}{2\sigma_i^2} y'_i . u_{1,i} \end{array} \tag{7}$$

where $\sigma_i^2 = \sigma_{\delta_i}^2 + \gamma_i^2 \sigma_{v_i}^2$. Last line is obtained assuming that all code-vectors have the same energy. Extraction is thus performed by maximization of a weighted cross product between observations and code-vectors.

In order to simplify the extraction process, we then propose to use structured code-books based on the concatenation of Error Correcting Codes and Wide

Spread Spectrum. Since Wide Spread Spectrum techniques can be seen as a special case of linear transform and due to its linear form, maximization of the weighted cross product of Eqn. (7) will be performed on the components of the associated linear demodulation of WSS.

We will now consider the optimization and implementation of such watermarking schemes when facing SAWGN attacks.

Remark

Watermarking based on Wide Spread Spectrum and Side Information is similar to previously proposed scheme such as: Spread Transform schemes (ST-DM [12] and ST-SCS [23,24]) or Quantized Index Modulation in a projection domain [25]. Developments made in this paper extend those results for non-i.i.d. signals and study theoretical performance of such schemes.

3 Watermarking of non i.i.d. Gaussian signals based on Wide Spread Spectrum and Side Information

In [26], we have formulated the optimization of watermarking of non i.i.d. Gaussian signals based on WSS and Side Information facing SAWGN attacks as the solution of a game between an attacker and the embedder. We first recall the main steps of this result and then discuss of the practical implementation of such a scheme.

3.1 Wide Spread Spectrum watermarking technique with Side Information

Let us consider a non i.i.d signal $\underline{\mathbf{x}}$ modeled by a set of random variables $X^n = \{X_1, X_2, \ldots X_n\}$ with $X_i \sim \mathcal{N}(0, \sigma_{X_i}^2)$, and a message to embed through vector $\underline{\mathbf{b}}$ of size m of 0-mean and variance $E[b_j^2] = 1$ [4]. Wide spread spectrum uses a set of quasi-orthogonal vector to represent the message. To embed m symbols in a signal of length n, a $n \times m$ random matrix $\underline{\mathbf{G}}$ is generated. The embedding stage can be written as:

$$y_i = x_i + w_i = x_i + \frac{\sigma_{W_i}}{\sqrt{\sum_{j=1}^{m} (G_{i,j})^2}} \sum_{j=1}^{m} b_j G_{i,j}. \quad (8)$$

The watermark is then also non i.i.d and modeled by $W^n = \{W_1, W_2, \ldots W_n\}$

[4] Definition of $\underline{\mathbf{b}}$ is discussed in section 3.3.

with $W_i \sim \mathcal{N}(0, \sigma^2_{W_i})$. We further consider a perceptual metric for distortion. This distortion is considered as a weighted distortion with perceptual factors φ_i [5] . The embedding distortion can be written as:

$$\begin{aligned} D_{xy} &= E\left[\sum_{i=1}^{n} \varphi_i^2 \left(y_i - x_i\right)^2\right] \\ &= \sum_{i=1}^{n} \varphi_i^2 \sigma^2_{W_i}. \end{aligned} \tag{9}$$

Considering SAWGN attacks, the received signal is expressed as

$$y'_i = \gamma_i y_i + \delta_i, \tag{10}$$

where γ_i is a scaling factor and $\underline{\delta}$ is a non i.i.d. noise signal $\sim \mathcal{N}(0, \sigma^2_{\delta_i})$. The distortion introduced by this attack can be quantified with

$$\begin{aligned} D_{xy'} &= E\left[\sum_{i=1}^{n} \varphi_i^2 \left(y'_i - x_i\right)^2\right] \\ &= \sum_{i=1}^{n} \varphi_i^2 \left(\sigma^2_{X_i} \left(1 - \gamma_i\right)^2 + \gamma_i^2 \sigma^2_{W_i} + \sigma^2_{\delta_i}\right). \end{aligned} \tag{11}$$

Further we have shown in [26], that is beneficial to perform Wiener filtering at embedding [6] . We then rather consider embedding distortion after Wiener filtering; that is:

$$D_{xy} = \sum_{i=1}^{n} \varphi_i^2 \frac{\sigma^2_{X_i} \sigma^2_{W_i}}{\sigma^2_{X_i} + \sigma^2_{W_i}} \tag{12}$$

Considering general Eqn. (7), extraction is performed by searching the closest code-vector to the vector $\widehat{b_j}$ obtained after linear demodulation of the WSS. Considering using side information technique, the optimal demodulation can be expressed as:

$$\widehat{b_j} \propto \sum_{i=1}^{n} \frac{\gamma_i \sigma_{W_i}}{\sigma^2_{\delta_i}} y'_i G_{ij} \tag{13}$$

[5] Those perceptual factor depends on the kind of signal that is treated. Watson weighting factor [27] may be used for example for images.

[6] In fact, strategy of the attacker can be shown to be noise addition followed by Wiener filtering. When no attack noise is added, this filtering lowers the distortion without degrading performances. Using such filtering at embedding is then beneficial for the embedder in order to lower its embedding distortion.

$\widehat{b_j}$ being i.i.d. Gaussian variables, global performance is then defined through the signal to noise ratio E_b/N_0 defined as

$$\frac{E_b}{N_0} = \frac{E\left[\widehat{b_j}^2\right]}{\sigma^2_{\widehat{b_j}}} = \sum_{i=1}^{n} \frac{\gamma_i^2 \sigma_{W_i}^2}{\sigma_{\delta_i}^2}. \tag{14}$$

The max-min game resolution used to estimate theoretical performance of this scheme is performed in two steps. First, the attacker tries to find the optimal attack defined by the optimal parameters $\gamma_i^\star$ and $\sigma_{\delta_i}^\star$. This is done by a Lagrangian optimization [7] :

$$\left(\underline{\gamma}^\star, \underline{\sigma_\delta}^\star\right) = \arg\min_{\underline{\gamma}, \underline{\sigma_\delta}} \left\{ J_\lambda = \frac{E_b}{N_0} + \lambda\left[D_{xy'} - D_{xy'}^{\max}\right]\right\}, \tag{15}$$

where λ is a Lagrangian multiplier introduced in order to respect constraint on the attack distortion. The second part of the game is focused on the embedder strategy: he must find the optimal parameters σ_{W_i} in order to maximize the performance of the extractor. This is done with a Lagrangian approach:

$$\underline{\sigma_W}^\star = \arg\max_{\underline{\sigma_W}} \left\{ J_\chi = J_\lambda - \chi\left[D_{xy} - D_{xy}^{\max}\right]\right\}, \tag{16}$$

where χ is a Lagrangian multiplier introduced in order to respect the constraint on the embedding distortion. This maximization leads to the final optimal embedding parameters given by

$$\sigma_{W_i} = \frac{\varphi_i^2(\lambda - \chi)\sigma_{X_i}^2 - 1 + \sqrt{\left(\varphi_i^2(\lambda-\chi)\sigma_{X_i}^2 - 1\right)^2 + 4\varphi_i^2\lambda\sigma_{X_i}^2}}{2\varphi_i\sqrt{\lambda}} \tag{17}$$

In practical scenarios (λ, χ) parameters are defined to fulfill application constraints among capacity, embedding distortion or maximal allowable attack distortion. Additional definition such as $\gamma_i, \sigma_{\delta_i}, \frac{E_b}{N_0}$ values can be found in [26].

Remark

In Parallel Gaussian channels of Moulin [21], game formulation is similar. However the difference lies in the metrics characterizing the performance of the system. Global extraction in our scheme leads to performance measure defined by Eqn. (14) which later defines the capacity of the resulting Gaussian channel $C = \frac{1}{2}\log_2[1 + \frac{E_b}{N_0}]$. While parallel Gaussian game uses performance

[7] See [26] for details.

measure defined as the sum of the capacities of the different channels (thus assuming possible separate treatment on each channel):

$$\sum_{i=1}^{n} \frac{1}{2} \log_2[1 + \frac{\gamma_i^2 \sigma_{W_i}^2}{\sigma_{\delta_i}^2}] \tag{18}$$

(a).

(b).

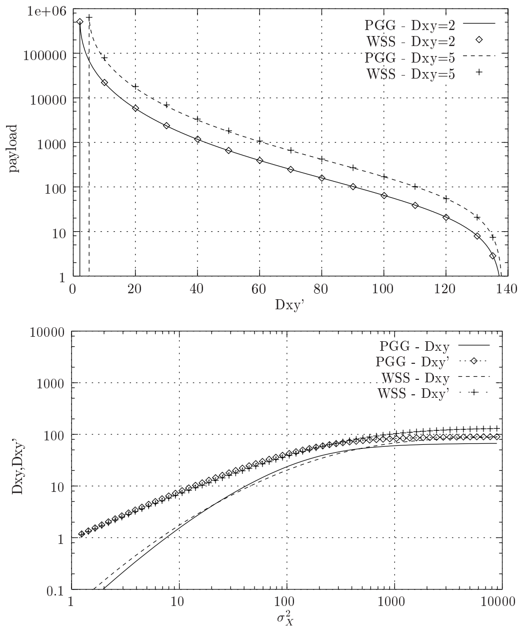


Fig. 1. Comparison between parallel Gaussian Channels and WSS with side Information. (a) capacity comparison for embedding distortions of 2 and 5. (b) embedding and attack distortions on the channels for global embedding and attack distortion of 10 and 20. Host signal is image Lenna after 3 levels DWT.

Fig. 1 shows a comparison between parallel Gaussian channels technique and our proposed approach. Capacity obtained with our proposed scheme is very close to the upper-bound defined by parallel Gaussian channels. On this figure are also presented the embedding and attack strategies (in terms of distortion) on each "channels". It can be observed that strategies for allocating distortions are very similar.

3.2 Recall of Costa's approach

Before presenting practical implementation for our proposed scheme, we first recall Costa's approach for channel coding with side information and it's direct application to watermarking of i.i.d. Gaussian signals. Further, in appendix A, a geometrical interpretation of Costa's embedding scheme is provided. In Costa's approach, we consider an i.i.d. Gaussian host signal $\underline{\mathbf{x}}$ modeled by $X \sim \mathcal{N}(0, Q)$. In the case of additive watermarking, the marked signal is

$\underline{\mathbf{y}} = \underline{\mathbf{x}} + \underline{\mathbf{w}}$. In order to control the embedding strength, the watermark signal must verify the bounded power constraint:

$$\frac{1}{n}\sum_{i=1}^{n} w_i^2 \leq P. \tag{19}$$

The $\underline{\mathbf{y}}$ signal may be attacked. This is modeled by an Additive White Gaussian Noise $\underline{\delta}$ whose mean is equal to zero and whose variance is N. The received signal is then $\underline{\mathbf{y}'} = \underline{\mathbf{x}} + \underline{\mathbf{w}} + \underline{\delta}$.

Costa has shown in [14] that the capacity of the transmission scheme described previously is given by

$$C = \frac{1}{2}\log_2\left[1 + \frac{P}{N}\right]. \tag{20}$$

This capacity can be reached with the introduction of a known signal $\underline{\mathbf{u}}$ modeled by $U \sim \mathcal{N}(0, P + \alpha^2 Q)$ so that $\underline{\mathbf{u}} = \underline{\mathbf{w}} + \alpha\underline{\mathbf{x}}$. The capacity can then be written as [28]:

$$\begin{aligned} C &= \max_{\alpha} R(\alpha) \\ &= \max_{\alpha} \left\{I(U;Y) - I(U;X)\right\}, \end{aligned} \tag{21}$$

where the random variable Y models the signal $\underline{\mathbf{y}}$. The maximum of the previous equation is reached for the value $\alpha = \frac{P}{P+N}$ leading back to Eqn. (20). Costa also proposed a constructive coding scheme [8] . It is based on a codebook $\mathcal{U}$ of $2^{n(I(U;Y)-\varepsilon)}$ elements, whose code-vectors are drawn according to the law $\mathcal{N}(0, (P + \alpha^2 Q)I)$. The term ε is chosen to be very small as $n \to \infty$. Each message that may be embedded is associated with $2^{nI(U;X)}$ code-vectors, i.e. the codebook is partitioned into $2^{n(C-\varepsilon)}$ bins $\mathcal{U}_r$, the index r corresponding to the r^{th} message. The code-vector $\underline{\mathbf{u}}^\star$ used for embedding is the closest one to $\underline{\mathbf{x}}$, leading to joint typical variables $(U, \alpha X)$ (i.e. $E\left[(U - \alpha X)^{\mathrm{T}} X\right] = 0$). The watermark is then defined as $\underline{\mathbf{w}} = \underline{\mathbf{u}}^\star - \alpha\underline{\mathbf{x}}$.

Watermarking embedding is performed in two steps. First the closest code-vector among $\mathcal{U}_M$ is searched [9] . Second $\underline{\mathbf{w}}$ is set to go towards this code-vector. Given the received message $\underline{\mathbf{y}'} = \underline{\mathbf{x}} + \underline{\mathbf{w}} + \underline{\delta}$, the extractor searches for the closest code-vector $\widehat{\underline{\mathbf{u}}}$ to $\underline{\mathbf{y}'}$.

Due to the fact that the ICS is based on random large codebooks, its implementation is not realistic since it requires to make an exhaustive search on all code-vectors. Practical, but suboptimal, approaches have been proposed for

[8] Known as ICS for Ideal Costa Scheme.

[9] It should be noted here that the norm of the code-vectors $\underline{\mathbf{u}}$ can take any constant value in the Gaussian case since the closest code-vector will always be the same.

i.i.d. Gaussian cover signals and AWGN channels based on structured codebooks [18,15,12,29]. All these scheme rely on the observation that codebooks provided by Error Correcting Codes will provide efficient codebooks for watermarking.

Since Side Information schemes do consider codebook larger than the set of messages, we can consider without loss of generality that a code-vector can be indexed with $(nC + nI(U;X))$ bits. The first nC bits identifying the message, while the last $nI(U;X)$ bits identify the code-vector into $\mathcal{U}_M$. Using ECC with fast decoding technique such as for example convolution codes or turbo-codes, it is then quite easy to retrieve the closest code-vector by setting an a priori on the first bits of the code-vectors [10] . Other dirty paper codes such as proposed in [16] and [17] may also be used for this purpose.

3.3 Practical Side Information embedding technique

Costa's embedding scheme is assuming i.i.d. Gaussian signals subject to Additive White Noise. When considering non i.i.d. signals subjected to SAWGN attacks with perceptual metrics, Costa's approach can not be directly applied. However using our proposed scheme presented in section 3.1, Costa's scheme can be used in the subspace defined by the linear watermark estimation (after WSS demodulation defined by Eqn. (13), we have an i.i.d. Gaussian channel [11]). We will now consider practical implementations of Side Information. To simplify the notations, $(\underline{\mathbf{x}}, \underline{\mathbf{w}}, \underline{\mathbf{y}}, \underline{\mathbf{y}}', \underline{\delta})$ will now be considered as the i.i.d. observations considered in the linear space generated by watermark estimation defined by Eqn. (13) with the optimal attack parameters [12] . It can be noted that when considering this optimal attack, WSS demodulation can be expressed as [13]

$$\widehat{b_j} \propto \sum_{i=1}^{n} \varphi_i y'_i G_{ij} \tag{22}$$

---

[10] When using trellis ECC, experimental results show that it is better not to put the $nI(U;X)$ bits at the end, but rather to spread it among other useful bits.
[11] This channel is also facing scaling operations due to the attacks. However since it is an i.i.d. Gaussian channel, and that code-vectors used have constant norm, scaling factors do not impact on performances when using ECC decoding with soft inputs $\widehat{b_j}$.
[12] In this linear transformation, formulation of γ_i, $\sigma^2_{W_i}$, $\sigma^2_{X_i}$ and $\sigma^2_{\delta_i}$ are used to express the different distortion constraints as signal power constraints similarly to Costa's formulation.
[13] This formulation is also valid for all optimal attacks performed on the system with lower distortion $D_{xy'} \leq D^{\max}_{xy'}$. For other attacks, channel state estimation has to be performed in order to estimate $(\gamma_i, \sigma^2_{\delta_i})$ parameters.

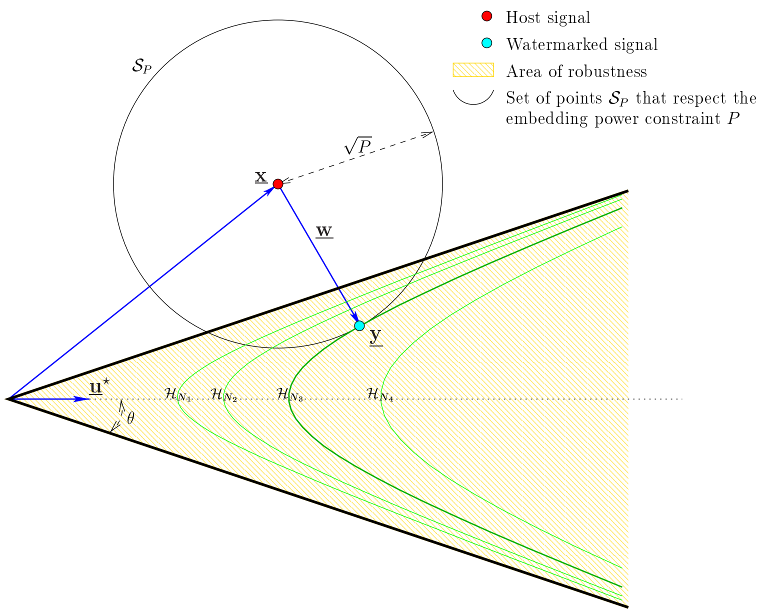


Fig. 2. Search method to get the best robustness with a fixed bounded embedding energy P.

In order to go towards the code-vector, Costa's scheme defined a watermark signal as the difference between the appropriately scaled code-vector and the host signal. This technique corresponds to the limit case where host signal is the farther away [14] . We give here an other technique which permits to get better results [15] . Further this technique turns out to be similar to the one previously proposed by Cox and al in [13] for detection technique.

Considering a maximum embedding distortion P, the watermark signal $\underline{\mathbf{w}}$ must be chosen efficiently to have a robust scheme against the noise addition of power N. This search for the best $\underline{\mathbf{w}}$ is illustrated by Fig. 2. As explained previously, the closest code-vector $\underline{\mathbf{u}}^\star$ is chosen. It defines a conical area of robustness: if the received signal $\underline{\mathbf{y}'}$ is inside this cone, the good code-vector can be extracted. The set of possible watermarked signals $\underline{\mathbf{y}}$ that respect the embedding power constraint P define an hyper-sphere $\mathcal{S}_P$ whose center is $\underline{\mathbf{x}}$. All the points of this sphere inside the cone are potential candidates for watermarked signal. However they don't all have the same robustness.

Considering the addition of a noise of power N, $\underline{\mathbf{y}}$ shouldn't go outside the cone. The set of points that may fall on the limit of the cone when subject to a noise of power N defines an hyperboloid inside the conical area. Given N,

[14] Costa's demonstration relies on this property which is statistically true as $n \to \infty$.
[15] Better results are effectively obtained, but since the limit case as a probability of occurrence that tends to 1, no significant improvements have to be expected.

they are defined by:

$$\mathcal{H}_N = \left\{ \underline{\mathbf{y}} \mid N = \left(\underline{\mathbf{y}} \cdot \underline{\mathbf{u}}\right)^2 \left(1 + \tan^2\theta\right) - |\underline{\mathbf{y}}|^2 \right\}. \tag{23}$$

Figure 2 shows examples of such hyperboloids. The optimal watermarked signal is then defined as the point of the hyper-sphere which have the highest robustness. Visually it corresponds to the tangent point between $\mathcal{S}_P$ and the hyperboloid $\mathcal{C}_N$ that maximize N.

In this embedding scheme cone aperture has to be defined. Considering that codebook $\mathcal{U}$ can be considered as a codebook for channel coding [16], we have $\tan\theta = \sqrt{\frac{N(P+N)}{P(P+Q+N)}}$ where θ is the half angle of the cone.

At extraction, the code-vector $\underline{\mathbf{u}}$ that is closest to $\underline{\mathbf{y}'}$ is simply searched using the ECC decoding technique. The message associated to this code-vector is then the one considered.

Remark

In these Side Information schemes, it is very important to use ECC decoding techniques based on soft inputs. These soft inputs allow to not normalize observations $\underline{\mathbf{y}}$, $\underline{\mathbf{y}'}$ and code-vectors $\underline{\mathbf{u}}$ when searching for the closest code-vector. This feature is further extremely important when considering SAWGN attacks otherwise scaling factor estimation would have to be performed such as proposed in [30] [17].

3.4 Subspaces dimension selection

Wide spread spectrum provides a way to embed m bits in a host signal of length n. As said earlier this could be interpreted as a kind of Spread Transform defining a linear embedding subspace. The embedding process applies only on $m = \varepsilon \times n$ components, with $\varepsilon \in]0;1]$. If we use the notations introduced in Sec. 3.2, the embedding distortion on this subspace is then P/ε

[16] See appendix A, where code-vectors of $\mathcal{U}$ are spread over an hyper-sphere. Code-vector βU should resist to noise addition of power $\sigma_V^2 + N$. The half angle θ of the hyper-cone associated to such a code-vector is thus defined as $\tan^2\theta = \frac{\sigma_V^2+N}{E[(\beta U)^2]} = \frac{N(P+N)}{P(P+Q+N)}$.

[17] Further, in order to work properly these estimation techniques necessitate that the attacker does not add noise to its scaling factors. In our proposed scheme, noise on the scaling factor does not impact and just acts as an additive noise similar to $\underline{\delta}$ noise.

while the global bounded power constraint from Eqn. (19) is still respected in the full space. Since this subspace is not known to the attacker, the noise $\underline{\delta}$ is spread equitably over all the components. The signal to noise ratio then becomes $P/\varepsilon N$. This leads to a new capacity definition:

$$C_\varepsilon = \frac{1}{2}\log_2\left[1 + \frac{P}{\varepsilon N}\right]. \tag{24}$$

in the subspace. For the whole signal, this gives $C' = \varepsilon C_\varepsilon$. When considering low signal to noise ratio, the capacity from Eqn. (20) can be approximated by $C \simeq \frac{P}{2N\log 2}$ and $C_\varepsilon \simeq \frac{P}{2\varepsilon N \log 2}$. From Eqn. (24), we can deduce:

$$\begin{aligned} C_\varepsilon &= \frac{\varepsilon}{2}\log_2\left[1 + \frac{P}{\varepsilon N}\right] && (25)\\ &= \frac{\varepsilon}{2}\log_2\left[1 + 2r\log 2\right] \text{ with } r \simeq C_\varepsilon && (26)\\ &\simeq C \times f(r) \text{ with } f(r) = \frac{1}{2r}\log_2\left[1 + 2r\log 2\right]. && (27) \end{aligned}$$

The function $f(r)$ then represents the ratio between the capacity limit defined by Costa and the achievable capacity using a linear subspace embedding technique. The term r can be interpreted as the rate between useful bits and inserted bits in the subspace [18] . Fig. 3 shows the variation of f. This figure shows that in order to get the highest performance, rate of the ECC should be the lowest. The maximal capacity can only be obtained when $r \to 0$, i.e. $\varepsilon \to 1$ (i.e. subspace represents the whole space), that is dimension of the subspace should be the largest. If we use $r = 1/3$, such is the case when using ECC with rate close to 1/3, about 85% of the maximal theoretic capacity can be achieved. This demonstrates that our proposed approach is close to optimal solution even when using ECC rates around 1/2 or 1/3. Further this allows to use subspaces with low number of dimension without significant loss in performance.

3.5 On the design of Side Information code-books

In Side Information techniques, it is necessary to define a set of code-vectors $\mathcal{U}$ of size $2^{nI(U;Y)}$ which is split into 2^{nC} sets $\mathcal{U}_M$ of code-vectors associated to the different existing messages; each sets $\mathcal{U}_M$ having $2^{nI(U;X)}$ elements. Ideal

[18] If only bits related to the message were considered, this could be interpreted as the rate of the ECC. However additional bits due to $I(U;X)$ have to be taken into account.

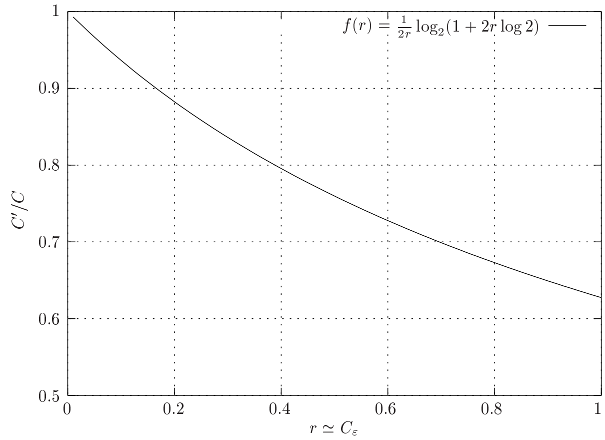


Fig. 3. Achievable capacity using subspaces.

value of these parameters are given according to Costa's paper as:

$$\begin{cases} I(U;Y) &= \frac{1}{2}\log_2(1+\frac{P(P+Q+N)}{N(P+N)}) \\ I(U;X) &= \frac{1}{2}\log_2(1+\frac{PQ}{(P+N)^2}) \\ C &= I(U;Y)-I(U;X) = \frac{1}{2}\log_2(1+\frac{P}{N}) \end{cases} \quad (28)$$

When considering WSS with Side Information, those values depends on the size of the subspace used. If $\varepsilon = \frac{m}{n}$ represents the ratio between the subspace and the full space, (P, Q, N) energy terms change to $(P' = \frac{P}{\varepsilon}, Q' = Q, N' = N)$, and capacity in the full space becomes $C = \varepsilon C_\varepsilon = \frac{\varepsilon}{2}\log_2(1+\frac{P'}{N'})$.

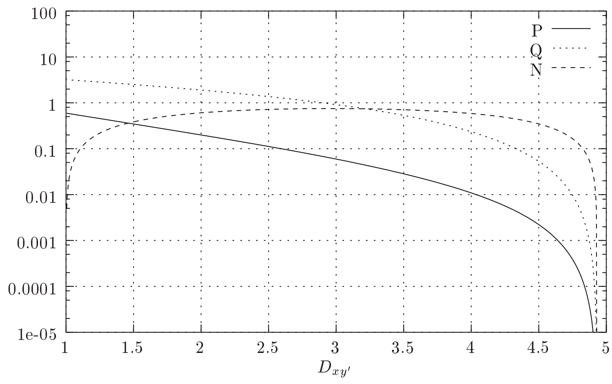
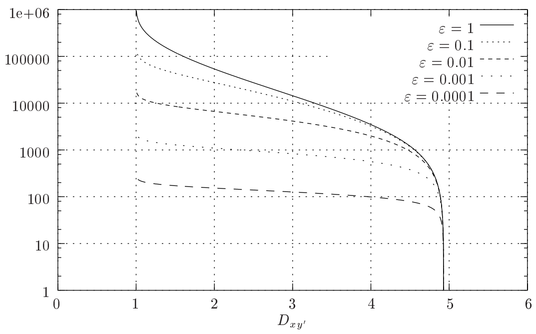
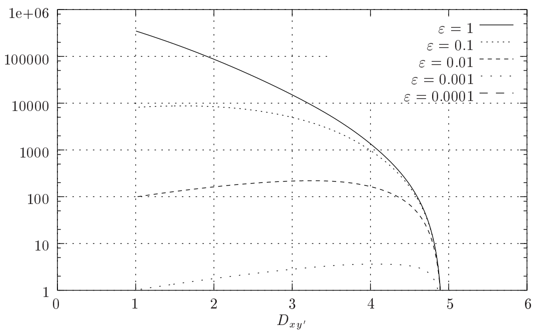


Fig. 4. Evolution of energies P, Q, N when using the estimator $\widehat{b_j} = \sum_i \varphi_i y'_i G_{ij}$. Image Lena, 512x512, embedding on 3 wavelet levels, use of perceptual factor $\varphi_i = \frac{1}{\sqrt{1+\sigma_{X_i}}}$, embedding distortion is set to 1.

Figure 4 shows the values of these different energies. Energies P and Q of respectively, the watermark and the host signal diminish since scaling factors tends to lower their response. Noise energy N first increases then decreases to zero (on the extreme case where all sites have been nullified, it is no more necessary to add noise). Figure 5 shows the impact of subspace dimension on capacities and additional bits necessary for side information scheme. For high distortions (corresponding to low payloads ~ 100 bits) we can observe that $I(U;X) \ll C$ and $n\varepsilon I(U;X)$ gets close to one or less. This means that the specificity of Side Information to provide several code-vectors for one message is not necessary in those situations. Thus for low payloads traditional WSS technique enriched by embedding technique of section 3.3 may be used without loss of performance. In other situations, number of additional bits is a fraction of the message length. Thus ECC technique will work with lengths that are of same order than the one of the message. Use of fast decoding techniques such is the case for convolution codes or turbo-codes then renders feasible such Side Information watermarking scheme [19] .

(a). 

(b). 

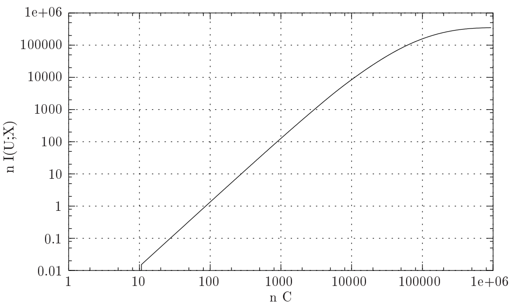
(c). 

Fig. 5. Impact of embedding subspace relative dimension ε on capacities $n\varepsilon C_\varepsilon$ (a) and additional bits $n\varepsilon I(U;X)$ (b). (c) show the evolution of $nI(U;X)$ function of nC for full space embedding. Image Lena, 512x512, embedding on 3 wavelet levels, use of perceptual factor $\varphi_i = \frac{1}{\sqrt{1+\sigma_{X_i}}}$, embedding distortion is set to 1. Results expressed in bits for the whole image.

[19] Complexity of those decoders is linear and is very low even for message of length around 1000 (for turbo-codes convergence is generally observed after a few iterations).

4 Experimental results

We now consider the application of the previous results to image watermarking. Message is embedded in the coefficients resulting from a wavelet decomposition of an image. A three level decomposition of the image has been used, and embedding is performed in all the subbands but the low frequency band. A perceptual factor [20] defined by $\varphi_i = (1+\sigma_{X_i})^{-\frac{1}{2}}$ is considered. Performances are measured using the signal to noise ratio E_b/N_0 of $\underline{\mathbf{b}}$. Eqn. (24) may then be used in order to express associated capacity.

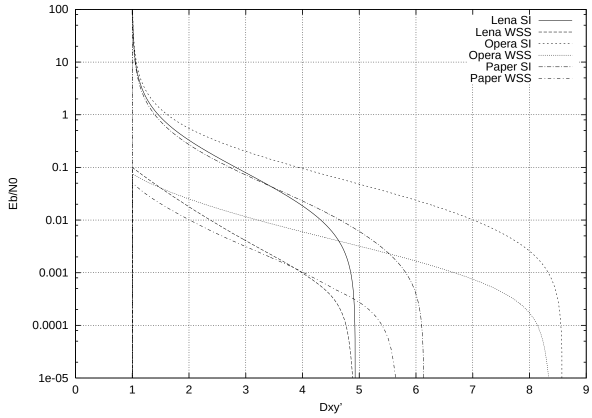


Fig. 6. Estimated signal to noise ratio estimated for Images *Lenna*, *Opera* and *Paper* using WSS with or without Side Information. Embedding distortion is set to 1.

Fig. 6 presents performance estimations for various images of our scheme using WSS with side information (SI) and compares it to results obtained in [22] using optimized WSS without side information (WSS). As expected, side informed schemes outperforms non informed schemes. Signal to noise ratio being increased by a factor 10 for medium attacks, capacity is nearly increased by a factor 10 [21]. It can also be observed that capacity is dependent on the image since variance of the host signal has an impact when considering SAWGN attacks.

In order to test our scheme against usual attacks, we then consider using Stirmark benchmark [31]. A 64 bits length message is embedded using half rate Error Correcting Code; embedding distortion is set to $D_{xy} = 1$. (λ, χ) parameters are tuned in order to ensure $\frac{E_b}{N_0} = 1$ with highest attack distortion $D_{xy'}$. Tab. 1 reports results for the non geometrical attacks of Stirmark. Once again, this demonstrates the importance of side information compared to non informed scheme of [22]. Further, since payload was low, no additional bits have been added, significant improvements have mainly been obtained thanks

[20] This is an adaptation of the metrics used in the context of JPEG 2000 compression.

[21] $\frac{1}{2}\log_2[1+\frac{E_b}{N_o}] \simeq \frac{1}{2\log 2}\frac{E_b}{N_o}$ for low signal to noise ratio.

	Side information	Spread spectrum
No attack	3028.29	130.55
2×2 median filtering	23.97	29.72
3×3 median filtering	90.43	68.24
90% quality JPEG compression	255.75	78.60
10% quality JPEG compression	11.74	7.60
3×3 Gaussian filtering	120.57	59.03
3×3 sharpening	420.26	142.94
FMLR	34.44	21.02

Table 1
Stirmark benchmark results for non-geometric attacks applied on image *Lena* (512×512 gray-scale image, three levels DWT). Embedding distortion is set to 1.

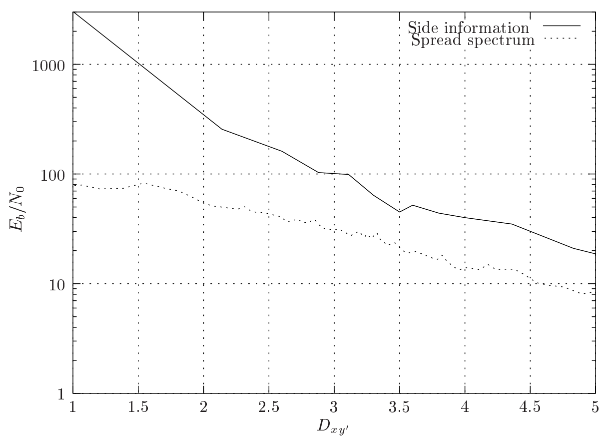


Fig. 7. Performance against JPEG compression for *Lena* (512×512 gray-scale image, three levels DWT). The psycho-visual factor used is $\varphi_i = (1+\sigma_{X_i})^{-\frac{1}{2}}$.

to the embedding technique described in section 3.3. E_b/N_0 measures may also be used in order to define the probability of bit error. $\frac{1}{2}erfc(\sqrt{\frac{E_b/N_0 d_{\min}}{2}})$ is a good estimation for this error probability where $d_{\min}$ represents the minimal distance of the ECC used. For results obtained with Stirmark benchmark, measures of error probability are always below 10^{-20}. For all of these tests, message was thus extracted without any errors. Fig. 7 shows the robustness of the presented scheme against JPEG compression. The watermarked image is compressed from 95 % to 10 % quality. The solid line represent the proposed solution, while the dashed one corresponds to WSS embedding scheme of [22]. At any of these levels, message is extracted without any errors.

5 Conclusion

In this paper we have studied the implementation of a practical watermarking scheme for non i.i.d. Gaussian signals and perceptual metrics for distortion. We have first shown that theoretical approach based on parallel Gaussian channels should not be perform with embedding/extraction on separate channels. We then reformulated the watermarking problem considering global embedding/extraction based on WSS and Side Information. Theoretical performances of this scheme has been established by considering a game between an attacker and the embedder for SAWGN attacks. This watermarking scheme leads to a practical implementation of Side Information scheme with performance very close to the upper-bound defined by parallel Gaussian Channels. Application to image watermarking has been validated by successfully resisting to all non geometrical Stirmark attacks.

A Geometrical interpretation of watermarking with Side Information

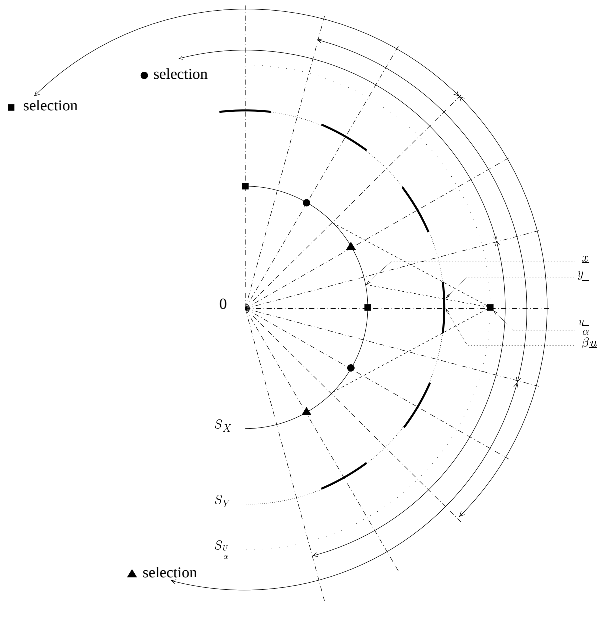


Fig. A.1. Geometrical interpretation of Costa's embedding scheme

Fig. A.1 gives a geometrical interpretation of Costa's embedding scheme. For visual rendering, this figure is in 2D but all informations drawn should be considered being in an n-dimensional space. Host signal $\underline{\mathbf{x}}$ lies on hyper-sphere S_X of square radius Q. Codebook $\mathcal{U}$ is created with code-vectors of energy

$P+\alpha^2 Q$ with $\alpha = \frac{P}{P+N}$ [22]. This codebook is split into sets $\mathcal{U}_M$ of size $2^{n.I(U;X)}$ associated with each 2^{nC} possible messages [23]. Cones drawn on Fig. A.1 show the areas containing the points that are the closest to a given code-vector. Several set of cones are considered. First, the set of small cones when considering all code-vectors. Second the sets when only code-vectors associated to a given message are considered (represented by angular sectors for various symbol selection).

The closest code-vector $\underline{\mathbf{u}}$ in $\mathcal{U}_M$ to signal $\underline{\mathbf{x}}$ is retrieved, and the embedding is done by letting $\underline{\mathbf{y}} = \underline{\mathbf{x}} + \alpha(\frac{\underline{\mathbf{u}}}{\alpha} - \underline{\mathbf{x}})$ (see on figure A.1 for such an example). Watermarking of the host signal $\underline{\mathbf{x}}$ leads to points that lay on the hyper-sphere S_Y. Bold arcs on this hyper-sphere show the different reachable values of $\underline{\mathbf{y}}$. As observed on this figure, this technique allows to move any signal $\underline{\mathbf{x}}$ in a cone associated to the corresponding code-vector. This can be demonstrated as follows. First resulting watermarked signal Y can be written as:

$$\underline{\mathbf{y}} = \beta \underline{\mathbf{u}} + \underline{\mathbf{v}} \tag{A.1}$$

with $\underline{\mathbf{v}}$ being a Gaussian noise V independent from U. It can be easily shown that we have:

$$\begin{cases} \beta &= \frac{P+\alpha Q}{P+\alpha^2 Q} \\ \sigma_V^2 &= \frac{(1-\alpha)^2 PQ}{P+\alpha^2 Q} \end{cases} \tag{A.2}$$

We thus have:

$$\begin{cases} E[(\beta U)^2] &= P\frac{(P+Q+N)^2}{(P+N)^2+PQ} \\ \sigma_V^2 &= N\frac{NQ}{(P+N)^2+PQ} \end{cases} \tag{A.3}$$

According to sphere packing theorem, when $n \to \infty$, we can put $2^{\frac{n}{2}\log_2(1+\frac{P}{N})}$ non overlapping spheres of square radius N centered on the hyper-sphere of square radius P. Since

$$\begin{cases} I(U;Y) &= \frac{1}{2}\log_2(1+\frac{P(P+Q+N)}{N(P+N)}) \\ \frac{P(P+Q+N)}{N(P+N)} &= \frac{P\frac{(P+Q+N)^2}{(P+N)^2+PQ}}{N\frac{NQ}{(P+N)^2+PQ}+N} = \frac{E[(\beta U)^2]}{\sigma_V^2+N} \end{cases} \tag{A.4}$$

we can then have $2^{nI(U;Y)}$ non intersecting spheres of square radius (σ_V^2+N) centered on an hyper-sphere of square radius $E[(\beta U)^2]$.

[22] These code-vectors lay on the hyper-sphere $S_{\frac{U}{\alpha}}$

[23] They are represented as square, disk and triangle symbols on the hyper-spheres. Each symbol is associated with a different message.

Costa's scheme is then similar to considering code-vectors βU subject to two noise: V the noise due to the host signal, and $\underline{\delta}$ the attack noise. Watermarked contents are thus already noisy and lay and the hyper-sphere S_Y (see figure A.1).

From these observations, watermarking with side information is similar to the problem of a Gaussian channel subject to additive noise (V and $\underline{\delta}$) although part of this noise is already present.

Remark

When looking at figure A.1, we can observe that adding random noise is not necessarily the best strategy for an attacker. By reducing the amplitude of the watermark signal, he can lower the distortion while using lower noise in order to get out the decoding cone. This remark just emphasizes the role and importance of SAWGN attacks.

References

[1] K. Matsui, K. Tanaka, Video-steganography: how to secretly embed a signature in a picture, Journal of the Interactive Multimedia Association Intellectual Property Project 1 (1) (1994) 187–205.

[2] J. R. Smith, B. O. Comiskey, Modulation and information hiding in images, in: Proc. Int. Workshop on Information Hiding, Vol. 1174, Cambridge, UK, 1996, pp. 207–226.
URL `citeseer.nj.nec.com/smith96modulation.html`

[3] F. Hartung, B. Girod, Watermarking of uncompressed and compressed video, IEEE Trans. Signal Proc. 66 (3) (1998) 283–302.

[4] I. J. Cox, J. Kilian, T. Leightom, T. Shamoon, Secure spread spectrum watermarking for multimedia, IEEE Trans. Image Proc. 6 (12) (1997) 1673–1687.

[5] J. Ruanaidh, W. Dowling, F. Boland, Phase watermarking of digital images, in: Proc. Int. Conf. on Image Processing, Vol. 3, Lausanne, Switzerland, 1996, pp. 239–242.

[6] M. D. Swanson, B. Zhu, A. H. Tewfik, L. Boney, Robust audio watermarking using perceptual masking, IEEE Trans. Signal Proc.: Special Issue on Copyright Protection and Control 66 (3) (1998) 337–355.

[7] C. I. Podilchuk, W. Zeng, Image-adaptive watermarking using visual models, IEEE Journal on Special Areas in Communications 16 (4) (1998) 525–539.

[8] S. Servetto, C. I. Podilchuk, K. Ramchandran, Capacity issues in digital image watermarking, in: Proc. Int. Conf. on Image Processing, Vol. 1, Chicago, IL, 1998, pp. 445–449.

[9] P. Moulin, J. A. O'Sullivan, Information-theoretic analysis of information hiding, IEEE Trans. Info. Thy .

[10] J. K. Su, J. J. Eggers, B. Girod, Analysis of digital watermarks subjected to optimum linear filtering and additive noise, IEEE Trans. Signal Proc.: Special Issue on Information Theoretic Issues in Digital Watermarking 81 (6).
URL `http://www.stanford.edu/ bgirod/pdfs/SignalProc2001.pdf`

[11] P. Moulin, A. Ivanovic, The watermark selection game, in: Proc. Conf. on Info. Sciences and Systems, 2001.
URL `http://www.ifp.uiuc.edu/ moulin/Papers/verify-ciss01.ps`

[12] B. Chen, G. W. Wornell, Quantization index modulation: a class of provably good methods for digital watermarking and information embedding, IEEE Trans. Info. Thy 47 (4) (2001) 1423–1443.

[13] I. J. Cox, M. L. Miller, A. L. McKellips, Watermarking as communications with side information, Proc. IEEE 87 (7) (1999) 1127–1141.

[14] M. H. M. Costa, Writing on dirty paper, IEEE Trans. Info. Thy 29 (3) (1983) 439–441.

[15] J. J. Eggers, J. K. Su, B. Girod, A blind watermarking scheme based on structured codebooks, in: Proc. IEE Colloq.: Secure Images & Image Authentification, London, UK, 2000.

[16] J. Chou, S. S. Pradhan, L. E. Ghaoui, K. Ramchandran, A robust optimization solution to the data hiding problem using distributed source coding principles, in: Proc. SPIE Image & Video Communications & Processing, Vol. 3974, 2000.

[17] M. L. Miller, G. J. Doerr, I. J. Cox, Dirty-paper trellis codes for watermarking, in: Proc. Int. Conf. on Image Processing, Vol. 2, Rochester, USA, 2002, pp. 129–132.

[18] M. Ramkumar, A. Akansu, A capacity estimate for data hiding in internet multimedia, in: Symp. on Content Security and Data Hiding in Digital Media, Newark, NJ, 1999.

[19] P. Moulin, M. K. Mihcak, The data-hiding capacity of image sources, IEEE Trans. Image Proc. Submitted.

[20] A. S. Cohen, A. Lapidoth, The gaussian watermarking game, to appear in IEEE Trans. Inform. Theory .

[21] M. K. Mihçak, P. Moulin, Information embedding codes matched to locally stationnary gaussian models, in: Proc. Int. Conf. on Image Processing, Vol. 2, Rochester, USA, 2002, pp. 137–140.

[22] G. Le Guelvouit, S. Pateux, C. Guillemot, Information-theoretic resolution of perceptual wss watermarking of non i.i.d gaussian signals, in: Proc. Eur. Signal Processing Conf., Vol. 1, Toulouse, France, 2002, pp. 454–457.

[23] J. J. Eggers, J. K. Su, B. Girod, Performance of a practical blind watermarking scheme, in: SPIE. (Ed.), Electronic Imaging 2001, San Jose, CA, 2001.

[24] J. J. Eggers, R. Bäuml, B. Girod, Digital watermarking facing attacks by amplitude scaling and additive white noise, in: 4th Int. ITG Conf. on Source and Channel Coding, 2002.

[25] F. Perez-Gonzalez, F. Balado, Inmproving data hiding performance by using quantization in a projected domain, in: Proc. Int. Conf. on Multimedia and Expo, Lausanne, Switzerland, 2002.

[26] G. Le Guelvouit, S. Pateux, C. Guillemot, Perceptual watermarking of non i.i.d. signals based on wide spread spectrum using side information, in: Proc. Int. Conf. on Image Processing, Vol. 3, Rochester, USA, 2002, pp. 477–480.

[27] A. B. Watson, DCT quantization matrices visually optimized for individual images, Proc. SPIE 1913 (1993) 202–216.

[28] S. Gel'fand, M. Pinsker, Coding for channel with random parameters, Problems of Control and Information Theory 9 (1) (1980) 19–31.

[29] J. C. Chou, K. Ramchandran, Turbo-coded trellis-based constructions for data hiding, in: Proc. SPIE Security & Watermarking of Multimedia Contents, Vol. 4675, 2002.

[30] J. J. Eggers, R. Bäuml, B. Girod, Estimation of amplitude modifications before scs watermark detection, in: Proceedings of SPIE: Electronic Imaging 2002, Security and Watermarking of Multimedia Contents IV, Vol. 4675, San Jose, CA, USA, 2002.

[31] F. A. P. Petitcolas, R. J. Anderson, Evaluation of copyright marking systems, in: Proc. Int. Conf. Multimedia Systems, Vol. 1, Florence, Italy, 1999, pp. 574–579.
URL
`http://www.cl.cam.ac.uk/ fapp2/papers/ieeemm99-evaluation.pdf`