Watermarking error exponents in the presence of noise: The case of the dual hypercone detector

Teddy Furon

Watermarking error exponents in the presence of noise: The case of the dual hypercone detector

Teddy Furon
Univ. Rennes, Inria, CNRS, IRISA
Rennes, France

ABSTRACT

The study of the error exponents of zero-bit watermarking is addressed in the article by Comesana, Merhav, and Barni, under the assumption that the detector relies solely on second order joint empirical statistics of the received signal and the watermark. This restriction leads to the well-known dual hypercone detector, whose score function is the absolute value of the normalized correlation. They derive the false negative error exponent and the optimum embedding rule. However, they only focus on high SNR regime, *i.e.* the noiseless scenario.

This paper extends this theoretical study to the noisy scenario. It introduces a new definition of watermarking robustness based on the false negative error exponent, derives this quantity for the dual hypercone detector, and shows that its performances is almost equal to Costa's lower bound.

KEYWORDS

Watermarking, Error exponents

ACM Reference Format:
Teddy Furon. 2019. Watermarking error exponents in the presence of noise: The case of the dual hypercone detector. In *IHMMSEC '19: ACM Workshop on Information Hiding and Multimedia Security, July 03–05, 2019, Paris, France.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

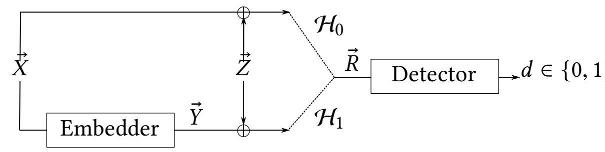

Figure 1: Zero-bit watermarking. The embedder hides a mark into the content. The detector checks for the presence of this mark.

1 INTRODUCTION

This paper is a theoretical study of the performances of a precise zero-bit watermarking scheme. It is theoretical because it considers an unrealistic model where the signals to be watermarked and the noise are assumed to be Gaussian distributed and infinitely long. It is specific to a given watermarking scheme as it focuses on the hypercone detector. This watermark detection scheme is important as [13] proves its optimality under some conditions.

Nevertheless, the performances of this scheme and its optimal watermark embedding are known only in the "high SNR regime" [3], *i.e.* when the attack noise power tends to zero. The main contribution of this paper is a follow-up extending paper [3] to any SNR regime. A shift of paradigm makes this extension tractable: Instead of optimising the performances of the scheme for a given noise power, the goal is to keep the performances acceptable over a maximum range of noise power. This shift brings a new viewpoint of this problem. Its gives birth to a new definition of watermarking robustness, which is a second contribution of the paper.

As a minor contribution, this paper revisits as well the noiseless setup (*i.e.* the limit of the high SNR regime). Failing detecting the watermark in the noiseless setup is equivalent to failing watermarking of a given host signal. The amount of watermark power is not big enough to make that host signal detectable. This observation eases the computation of the false negative error exponent thanks to a rolling-ball region filtering. It also has a nice connection with isoperimetric Gaussian inequality.

Section 2 introduces zero-bit watermarking and the theoretical setup. Section 3 lists the assumptions and the requirements specific to digital watermarking. The paper starts by revisiting in Sect. 4 the noiseless setup originally considered in [3], and extends this piece of theory to the noisy setup in Sect. 5. Section 6 proposes some upper and lower bounds adapting the rationale of M. Costa [4] to zero-bit watermarking. At last, a practical embedding strategy is deduced from this theoretical study in Sect. 7.

2 THE THEORETICAL SETUP

Zero-bit watermarking is different from multi-bit watermarking. While people usually knows what watermarking means, some get confused between the *detection* and the *decoding* of a watermark. In multi-bit watermarking, a first algorithm, so-called embedder, hides a message (possibly encoded in several bits) into a piece of content. A second algorithm analyses a piece of content and proceeds to a decoding. The decoding outputs the hidden message or the decision that the piece of content under scrutiny is indeed not watermarked.

In *zero-bit* watermarking, one is solely interested in distinguishing watermarked from non watermarked content. Therefore, the embedding does not hide any message, but just a mark. There is no modulation of a signal by the message to be transmitted since there is no message. Hence, the term *zero-bit* watermarking. In the same way, the second algorithm does not perform a decoding, but a *detection* of the presence or the absence of the mark (see Fig. 1).

IHMMSEC '19, July 03–05, 2019, Paris, France

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2.1 Notations

A feature vector in $\mathbb{R}^n$ is extracted from a piece of multimedia content. Vectors $\vec{x}$ and $\vec{r}$ denote respectively the extracted features from an original content, so-called the host, and from the content received by the detector. The embedder transforms $\vec{x}$ into $\vec{y}$ by adding a watermark signal: $\vec{y} = \vec{x} + \vec{w}(\vec{x})$. This vector depends on the host (for a side-informed watermarking scheme) and on a secret key (not indicated to keep notations simple).

We consider a power constraint watermark problem where the energy of the watermark per sample is limited.

$$\|\vec{w}(\vec{x})\|^2 \leq nP, \forall \vec{x} \in \mathbb{R}^n. \tag{1}$$

The Euclidean norm of vector $\vec{x} \in \mathbb{R}^n$ is denoted by $\|\vec{x}\|$.

The model of an attack is the addition of a noise vector $\vec{z}$, and the received vector extracted from the content under scrutiny is $\vec{r} = \vec{y} + \vec{z}$. At the detection side, two hypotheses are competing. The decision of the detector is denoted by d: $d = 1$ if the received content is deemed watermarked, $d = 0$ otherwise. There are two types of errors:

Under $\mathcal{H}_0$: The received vector has not been watermarked: $\vec{r} = \vec{x} + \vec{z}$. A *false positive* happens when $d = 1$ with probability $P_{\mathsf{fp}} := \mathbb{P}[d = 1|\mathcal{H}_0]$.

Under $\mathcal{H}_1$: The received vector has been watermarked: $\vec{r} = \vec{x}+\vec{w}(\vec{x})+\vec{z}$. A *false negative* happens if $d = 0$ with probability $P_{\mathsf{fn}} := \mathbb{P}[d = 0|\mathcal{H}_1]$.

To take a decision, we assume that the detector first computes a score from received vector $s(\vec{r})$ with $s(\cdot) : \mathbb{R}^n \to \mathbb{R}$. Then, it compares this score to a threshold τ: $d = 1$ if $s(\vec{r}) \geq \tau$ and $d = 0$ otherwise. This defines the region $\mathcal{W} \subset \mathbb{R}^n$ of the vectors deemed as watermarked:

$$\mathcal{W} := \{\vec{x} \in \mathbb{R}^n | s(\vec{x}) \geq \tau\}. \tag{2}$$

2.2 Theoretical setup

The theoretical setup assumes that the signals are instances of a white Gaussian distribution in $\mathbb{R}^n$. Denote by $\vec{X}$ the random host vector whose power is σ_X^2, $\vec{X} \sim \mathcal{N}(\vec{0}_n, \sigma_X^2.I_n)$, and $\vec{Z}$ the random noise vector of power σ_Z^2, $\vec{Z} \sim \mathcal{N}(\vec{0}_n, \sigma_Z^2.I_n)$. We assume that $\vec{Z}$ is independent of $\vec{X}$ and the secret key.

Computation of the performances $(P_{\mathsf{fp}}, P_{\mathsf{fn}})$ is difficult even under this simple setup. To facilitate comparison, the study focuses on the error exponents, *i.e.* the exponential decay rate of the error probabilities:

$$E_{\mathsf{fp}} := \lim_{n\to\infty} -\frac{1}{n} \log P_{\mathsf{fp}}, \quad E_{\mathsf{fn}} := \lim_{n\to\infty} -\frac{1}{n} \log P_{\mathsf{fn}}. \tag{3}$$

For the sake of simplicity, notations omit the fact that $(E_{\mathsf{fp}}, E_{\mathsf{fn}})$ depends on $(P, \sigma_X^2, \sigma_Z^2)$.

Note that, at then end of the paper, Section 7 deals with a more practical setup where no statistical model of the host is assumed.

3 REQUIREMENTS AND PRIOR ART

The prior art of zero-bit watermarking is mainly organised around the issue of obliviousness. When describing a watermarking scheme, assumptions about what the embedder and the detector know and do not know about the setup is critical. This matters when turning a theoretical watermarking scheme into a practical technique watermarking content.

3.1 Assumption on obliviousness

Multimedia contents have a wide diversity. Features extracted from these contents are certainly not white Gaussian distributed with a fixed power σ_X^2. This is the reason why the above setup is pure theory. As a small step towards being more realistic, content diversity may imply that σ_X^2 vary from one content to another. A watermarking scheme relying on the knowledge of the watermark detector about that parameter (to guarantee a given P_{fp}, for instance) is not applicable in practice.

This reasoning holds as well for P. The watermarking power usually depends on the masking properties of the host content. These properties are also very diverse from one content to another. For instance, the human eye is less sensitive to noise in textured areas than in flat regions. Therefore, one has to adapt the watermark power to the visual content of an image. The masking properties of the host content are usually analysed by a Human Visual System model at embedding. Yet, the watermark detector might receive a heavily distorted copy of the content preventing such analysis. In other words, the detector cannot know the value of P used at the embedding.

As for σ_Z^2, obliviousness is also a plus at the embedding side: The embedder may not know in advance the amount of noise power that the watermarked content will support.

To conclude, this paper integrates the specificities of watermarking in the theoretical setup by imposing the following obliviousness assumptions:

- The embedder is oblivious w.r.t. σ_Z^2.
- The detector is oblivious w.r.t $(\sigma_X^2, \sigma_Z^2, P)$.

3.2 Requirements

This section outlines the relevance of the concept of error exponent in practice where the length n is large enough. C. E. Shannon indeed motivated its use in his seminal work [15]. He warns that an error exponent a priori leads to inaccurate probability estimate: As $P_{\mathsf{fp}} = e^{-nE_{\mathsf{fp}}+o(n)}$, neglecting the term $o(n)$ may cause large multiplier uncertainty. Yet, for given E_{fp} and P_{fp}, Shannon outlines that $-\log(P_{\mathsf{fp}})/E_{\mathsf{fp}}$ sharply determines the necessary vector length. There is thus a trade-off between the exponent E_{fp} and the complexity of the scheme reflected by length n.

In practice, the main requirement is the probability of false positive P_{fp}. In many applications, its level is low and must be provably low. This means that one has to prove that the detector operates at a required low level. In the theoretical setup, operating a given exponent $E_{\mathsf{fp}} = E$ then determines the vector length n.

The false negative probability P_{fn} is usually less constrained than P_{fp}. In many applications, watermarking is a dissuasive weapon: P_{fn} should be small enough that attackers don't take the risk of pirating content. Indeed, $P_{\mathsf{fn}} \approx 1/2$ might be dissuasive enough. In the asymptotical setup, having $E_{\mathsf{fn}} = 0$ means that P_{fn} is not converging to zero exponentially fast. It might converge to zero more slowly or it might converge to another value.

Note that once the watermark detector operates at a fixed E_{fp}, say $E_{\text{fp}} = E$, E_{fn} depends on parameters $(\sigma_X^2, \sigma_Z^2, P, E)$. The above assumptions on obliviousness imply that there is no guarantee about E_{fn} at the detection side. However, just knowing that $E_{\text{fn}} > 0$ even if it is by a very small amount, would prove that the dissuasion is achieved. This motivates the following definition.

Definition 3.1. For a given setup $(\sigma_X^2, \sigma_Z^2, P, E)$, a watermarking scheme operating at $E_{\text{fp}} = E$ is deemed robust if $E_{\text{fn}} > 0$. We suppose that E_{fn} is always a decreasing function w.r.t. σ_Z^2. The robustness $R(\sigma_X^2, P, E)$ is the maximum noise power for which the watermarking scheme is robust. It is defined as

$$R(\sigma_X^2, P, E) := \sup\{\sigma_Z^2 | E_{\text{fn}}(\sigma_X^2, \sigma_Z^2, P, E) > 0\}. \tag{4}$$

For a given setup $(\sigma_X^2, \sigma_Z^2, P)$, the characteristic $E_{\text{fn}} = F(E_{\text{fp}})$ is a decreasing function, illustrating the trade-off between the false negative and false positive probabilities. Usually, this characteristic vanishes to zero at some point that we name the right endpoint.

Definition 3.2. The right endpoint of the characteristic is the biggest false positive error exponent for which the watermark is robust.

$$E_{\text{fp}}^R(\sigma_X^2, \sigma_Z^2, P) := \sup\{E | F(E) > 0\}. \tag{5}$$

3.3 Prior art

The issue of obliviousness w.r.t. (σ_X^2, σ_Z^2) at the detection side has been solved in two ways in the literature.

The first approach relies on Voronoï modulation (a.k.a. modulo channel) [16]. Lattices embedding have been widely studied for decoding hidden messages (often called Quanitzed Index Modulation) [2, 10, 12] but also in detecting zero-bit watermarking [9]. It uses a Euclidean lattice Λ and the corresponding modulo operator $(\vec{x} \mod \Lambda)$ to fold the space $\mathbb{R}^n$ onto the Voronoi cell of that lattice.

In a nutshell, the fine grain (a.k.a. high resolution) assumption states that if the typical scale of the lattice is small compared to $\sqrt{\sigma_X^2 + \sigma_Z^2}$, then $(\vec{X} + \vec{Z} \mod \Lambda)$ is uniformly distributed over the Voronoï cell of Λ. This can be also achieved thanks to a dithering signal which randomly shifts the lattice. In the end, the modulo operator succeeds to transform the unknown distribution of $\vec{X} + \vec{Z}$ (because the detector is oblivious w.r.t. (σ_X^2, σ_Z^2)) into a known distribution (uniformity over the Voronoï cell). This in turn allows to compute and guarantee probability P_{fp}.

The second approach uses a detection region $\mathcal{W}$ (2) which is a linear cone: If $\vec{x} \in \mathcal{W}$, then $\alpha\vec{x} \in \mathcal{W}$, $\forall \alpha > 0$. This provides an invariance to scaling. If the distribution of $\vec{X} + \vec{Z}$ is isotropic (as assumed in the theoretical setup), then $(\vec{X} + \vec{Z})/\|\vec{X} + \vec{Z}\|$ has a uniform distribution over the unit hypersphere. Again, this allows to compute and guarantee probability P_{fp}.

The well-known dual hypercone detection is an example of this second approach with a score function defined as

$$s(\vec{x}) = |\vec{x}^\top \vec{u}|/\|\vec{x}\|, \tag{6}$$

where $\vec{u} \in \mathbb{R}^n$, $\|\vec{u}\| = 1$, plays the role of a secret key. Threshold τ in (2) is defined as $\tau = \cos(\theta)$. Region $\mathcal{W}$ is then the circular dual hypercone of axis $\vec{u}$ and semi-angle $\theta \in [0, \pi/2]$.

This scheme has a long tradition in the history of digital watermarking. Since the seminal papers of I. Cox et al. [5, 6], normalized correlation has been used in a vast majority of papers [7] until side-information schemes were introduced [2, 6]. The argument of the seminal paper [5] was purely image processing oriented: normalizing the correlation is a way to be robust to contrast enhancement. Then some signal processing arguments defended this option [6, Sect. VI][1, Chap. 6, p. 237]: Decompose $\vec{R}$ as $(\vec{R}^\top \vec{u})\vec{u} + \vec{R}^\perp$ where $\vec{R}^\perp$ is the Euclidean projection of $\vec{R}$ onto the subspace orthogonal to $\vec{u}$. Under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, this projection has the same distribution $\mathcal{N}(\vec{0}_{n-1}, N.I_{n-1})$. Variance N is then estimated by $\|\vec{R}^\perp\|^2/n-1$ and used for comparing $\vec{R}^\top \vec{u}$ to the threshold $\tau = \sqrt{\hat{N}}\Phi^{-1}(1 - P_{\text{fp}})$. This indeed amounts to compare the ratio $\vec{R}^\top \vec{u}/\|\vec{R}^\perp\|$ to a threshold, say $1/\tan(\theta)$, or equivalently, to compare $s(\vec{R}) = \vec{R}^\top \vec{u}/\|\vec{R}\|$ to $\cos(\theta)$.

Ten years later, Merhav et al. show that this scheme is optimal from the information theoretical veiwpoint [13]. Here is a brief summary of results concerning the dual hypercone in the literature [3, 13]:

$$E_{\text{fp}} = -\log\sin(\theta), \tag{7}$$

$$\lim_{\sigma_Z^2 \to 0} E_{\text{fn}} = \begin{cases} 0 & \text{if } A < \cos(\theta) \\ S\left(\frac{A^2}{\cos^2(\theta)}\right) & \text{otherwise} \end{cases} \tag{8}$$

where

$$A := \sqrt{P/\sigma_X^2}, \tag{9}$$

$$S(x) := (x - 1 - \log(x))/2, \forall x \in \mathbb{R}_{>0}. \tag{10}$$

Note that function $S(\cdot)$ has a unique global minimum 0 at $x = 1$.

One can see that the characteristic $E_{\text{fn}} = F(E_{\text{fp}})$ is given by a parametric equation on θ. Usually, the watermarking power P is smaller than σ_X^2, so that $A < 1$. The right endpoint is then

$$E_{\text{fp}}^R = -1/2\log(1 - A^2). \tag{11}$$

Unfortunately, this characteristic is only known for $\sigma_Z^2 \to 0$. This is the reason why the authors of [3] speak about 'high SNR regime'. Since it is a zero-order expression for $\sigma_Z^2 \to 0$, this is indeed the characteristic in the noiseless scenario. Our main contribution provides new results in the noisy scenario, *i.e.* when $\sigma_Z^2 > 0$.

4 REVISITING THE NOISELESS SETUP

Before dealing with the noisy scenario, this section shows some hints about the noiseless scenario. Let us define the embeddable region as follows:

$$\mathcal{E}(P) := \{\vec{x} \in \mathbb{R}^n | \exists \vec{y} \in \mathcal{W} : \|\vec{x} - \vec{y}\|^2 < nP\}. \tag{12}$$

This is the set of vectors in $\mathbb{R}^n$ which can be successfully watermarked with a power budget P. This region is the filtering of $\mathcal{W}$ by a ball of radius $\sqrt{nP}$, a.k.a. the result of the rolling ball technique [14]: By rolling a ball of that radius over the boundary of $\mathcal{W}$, the center of that ball draws the boundary of region $\mathcal{E}(P)$.

The main idea of this section is to note that, under the noiseless scenario, a false negative happens at the detection side whenever the embedding fails watermarking a given signal. Therefore,

$$P_{\text{fn}} = \mathbb{P}(\vec{X} \notin \mathcal{E}(P)) = 1 - \mathbb{P}(\vec{X} \in \mathcal{E}(P)). \tag{13}$$

We are thus looking for a region $\mathcal{W}$ s.t. $\mathbb{P}(\vec{X} \in \mathcal{W}) = P_{\text{fp}}$ and which, once filtered by the rolling ball technique, gives the lowest P_{fn}, i.e. the biggest probability $\mathbb{P}(\vec{X} \in \mathcal{E}(P))$. This is an elegant way to theoretically study side-informed watermarking under the noiseless scenario because there is no need to specify anything about the embedding mechanism (*i.e.* function $\vec{w}(\cdot)$).

4.1 Lower bound

The Gaussian isoperimetric inequality [8] gives the worse possible region: For any region $\mathcal{W} \subset \mathbb{R}^n$ and $\vec{X} \sim \mathcal{N}(\vec{0}_n, \sigma_X^2 I_n)$ s.t. $\mathbb{P}(\vec{X} \in \mathcal{W}) = P_{\mathrm{fp}}$, we have

$$\mathbb{P}(\vec{X} \in \mathcal{E}(P)) \quad \geq \quad \Phi\left(\Phi^{-1}(P_{\mathrm{fp}}) + A\sqrt{n}\right), \tag{14}$$

$$P_{\mathrm{fn}} \quad \leq \quad 1 - \Phi\left(\Phi^{-1}(P_{\mathrm{fp}}) + A\sqrt{n}\right). \tag{15}$$

Function $\Phi(\cdot)$ is the cumulative density function of $\mathcal{N}(0,1)$. According to the Gaussian isoperimetric theorem, equality happens if and only if $\mathcal{W}$ is a half-space. Following the definition (2) of $\mathcal{W}$, this means that $s(\vec{x}) = \vec{x}^\top \vec{u}$. The upper bound (15) translates into the following lower bound for E_{fn}:

$$E_{\mathrm{fn}} \geq \left(\left|\frac{A}{\sqrt{2}} - \sqrt{E_{\mathrm{fp}}}\right|_+\right)^2. \tag{16}$$

In the same way, $E_{\mathrm{fp}}^R \geq A^2/2$.

This shows that a linear score function is indeed the worse choice in the noiseless setup, independently from the assumptions about obliviousness. Ironically, linear correlation was quite popular in the early ages of watermarking.

4.2 Dual hypercone

This section presents the methodology for calculating E_{fn} with the rolling ball technique over the dual hypercone. Suppose that $\vec{u} = (1, 0, \ldots, 0)^\top$, without loss of generality. The definition of the embeddable region writes as:

$$\mathcal{E}(P) = \left\{\vec{x} \in \mathbb{R}^n \middle| \sqrt{\sum_{i=2}^n x_i^2} \leq |x_1| \tan(\theta) + \sqrt{nP}\cos^{-1}(\theta)\right\}. \tag{17}$$

Define $V_1 = |X_1|/\sqrt{n}$ and $V_2 = \sqrt{\sum_{i=2}^n X_i^2/n}$. These are two χ_k random variables whose pdf is given by

$$f(v) = \frac{2\sqrt{n}}{2^{k/2}\Gamma(k/2)}\left(\frac{v\sqrt{n}}{\sigma_X}\right)^{(k-1)} e^{-n\frac{v^2}{2\sigma_X^2}} \tag{18}$$

with degree of freedom $k_1 = 1$ and $k_2 = n-1$ respectively. This change of variable yields a definition of set $\mathcal{E}(P)$ independent of n:

$$\mathcal{V} = \{(v_1, v_2) \in \mathbb{R}^2_{\geq 0} | v_2 \leq v_1 \tan(\theta) + \sqrt{P}\cos^{-1}(\theta)\}, \tag{19}$$

so that

$$P_{\mathrm{fn}} \quad = \quad \int_{\overline{\mathcal{V}}} f_{V_1}(V_1) f_{V_2}(v_2) dv_1 dv_2 \tag{20}$$

$$\quad = \quad K_n \int_{\overline{\mathcal{V}}} g(v_1, v_2) e^{-nh(v_1,v_2)} dv_1 dv_2, \tag{21}$$

for some functions $h(\cdot)$ and $g(\cdot)$, and a multiplicative constant K_n defined in the appendix. On one hand, we compute the 'exponent' of the multiplicative constant: $\kappa := \lim_{n\to\infty} -n^{-1}\log K_n$. On the other hand, the Laplace method states that, as $n \to \infty$, the integral is dominated by the value $e^{-nh^\star}$ where $h^\star$ is the minimum of function $h(\cdot)$ over $\overline{\mathcal{V}}$ (under some mild conditions). Then, the error exponent is given by:

$$E_{\mathrm{fn}} = \min_{\overline{\mathcal{V}}} h(v_1, v_2) + \kappa. \tag{22}$$

For the dual hypercone in the noiseless scenario, calculations lead to:

$$E_{\mathrm{fn}} = \min_{\overline{\mathcal{V}}} \frac{v_1^2}{2\sigma_X^2} + S\left(\frac{v_2^2}{\sigma_X^2}\right). \tag{23}$$

This function has a unique global minimum 0 at $v_1 = 0$ and $v_2 = \sigma_X$ (see (10)). This minimum lies in $\overline{\mathcal{V}}$ if $A \leq \cos(\theta)$. Otherwise, the solution of (22) lies on the boundary, *i.e.* $v_2 = v_1\tan(\theta) + A\cos^{-1}(\theta)$. This yields a univariate function in v_1 to be minimised, whose derivative takes only positive values. This shows that the minimum happens for the smallest value of v_1, *i.e.* $v_1 = 0$ so that $v_2 = A/\cos(\theta)$ (see Fig. 2). This rediscovers the results (8) of [3].

Interpretation: Probability P_{fn} is dominated by the probability that $\vec{X}$ lies around the closest point to the origin in $\mathcal{E}(P)$. If this minimum distance $\sqrt{nP}\cos(\theta)$ is lower than the typical module of $\vec{X}$, *i.e.* $\sqrt{n\sigma_X^2}$, then watermarking fails almost surely as $n \to \infty$ (by concentration) so that $E_{\mathrm{fn}} = 0$.

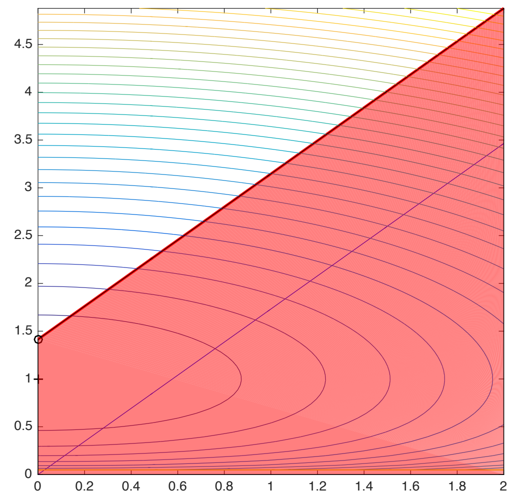

Figure 2: Laplace method for the dual hypercone with side information in the noiseless scenario. The red area is region $\mathcal{V}$ in the plane (v_1, v_2). E_{fn} is related the minimum of function (23) over the domain $\overline{\mathcal{V}}$. The level sets of function are depicted in colors. Here, the global minimum (black +) is not in $\overline{\mathcal{V}}$. The local minimum (black o) lies on the boundary.

5 THE NOISY SETUP

In the previous section, the rolling ball technique frees us to specify the way host vectors are watermarked. This section is now a little bit more specific for the hypercone detector. It is well known that $\vec{w}(\vec{x})$ must lie in the 2D subspace spanned by $\vec{u}$ and $\vec{x}$ [11, 13]. We work on the following basis of this subspace:

$$\vec{e}_1 = \vec{u}, \quad \vec{e}_2 = (\vec{x} - (\vec{x}^\top\vec{u})\vec{u})/\|(\vec{x} - (\vec{x}^\top\vec{u})\vec{u})\|. \tag{24}$$

The watermark signal is crafted as:

$$\vec{w}(\vec{x}) = \sqrt{n}(\tilde{w}_1\vec{e}_1 + \tilde{w}_2\vec{e}_2). \quad (25)$$

Any other embedding strategy wastes embedding energy in space directions not useful for detecting the watermark. For the moment, $(\tilde{w}_1, \tilde{w}_2)$ is not specified as a function of $\vec{x}$. The distortion constraint imposes that $\tilde{w}_1^2 + \tilde{w}_2^2 \leq P$.

The appendix A shows with the same kind of change of variables and the use of the Laplace method (as in Sect. 4) that:

$$E_{\mathsf{fn}} = \min_{\mathcal{F}} \frac{v_1^2}{2\sigma_X^2} + S\left(\frac{v_2^2}{\sigma_X^2}\right) + \frac{v_4^2 + v_5^2}{2\sigma_Z^2} + S\left(\frac{v_3^2}{\sigma_Z^2}\right), \quad (26)$$

with

$$\mathcal{F} = \{(v_1 + \tilde{w}_1 + v_5)^2 \tan^2(\theta) \leq (v_2 + \tilde{w}_2 + v_4)^2 + v_3^2\} \quad (27)$$

It is a priori not easy to solve this problem, but it is much simpler to see whether $E_{\mathsf{fn}} = 0$. This can only happen for $v_1 = v_4 = v_5 = 0$, $v_2 = \sigma_X$, and $v_3 = \sigma_Z$. Therefore, $E_{\mathsf{fn}} = 0$ if this point lies inside the feasible set $\mathcal{F}$. This holds if and only if

$$H(\tilde{w}_1, \tilde{w}_2) \leq \sigma_Z^2, \quad \text{with} \quad (28)$$

$$H(\tilde{w}_1, \tilde{w}_2) := \tilde{w}_1^2 \tan^2(\theta) - (\sigma_X + \tilde{w}_2)^2. \quad (29)$$

Interpretation: Asymptotically, the performance of the scheme is governed by the way the typical realization of a host signal is watermarked. This typical host is orthogonal to the axis of the hypercone ($v_1 = 0$) and has norm $\sqrt{n}\sigma_X$ (because $v_1^2 + v_2^2 = \sigma_X^2$). The typical realization of the noise has a norm $\sqrt{n}\sigma_Z$ (because $v_3^2 + v_4^2 + v_5^2 = \sigma_Z^2$), is orthogonal to the axis ($v_4 = 0$) and is orthogonal to the host ($v_5 = 0$). E_{fn} is null if this typical noise drives the watermarked signal outside the hypercone. The intersection of the hypercone with the plane $v_3 = \sigma_Z$ gives the hyperbola in $\mathbb{R}^2$:

$$C = \{(a, b) \in \mathbb{R}^2 | a^2 \tan^2(\theta) - b^2 = \sigma_Z^2\}. \quad (30)$$

5.1 Provably good embeddings

This subsection assumes that the watermark designer has chosen a given dimension n. The requirement on the false positive rate imposes $E_{\mathsf{fp}} \approx E := -n^{-1}\log(P_{\mathsf{fp}})$, which fixes the semi-angle θ by (7). What is the value of the robustness $R(\sigma_X^2, P, E)$?

We adopt now the point of view of the embedder. Our goal is to avoid such a null error exponent E_{fn} by carefully designing a watermark embedding $(\tilde{w}_1, \tilde{w}_2)$. In a 2D plane mapping point $(\tilde{w}_1, \tilde{w}_2)$, the embedding constraint $\tilde{w}_1^2 + \tilde{w}_2^2 \leq P$ defines a ball of radius $\sqrt{P}$ centered on $(0, 0)$ whereas (28) defines a region delimited by an hyperbola (equality in (28)) of center $(0, -\sigma_X)$. As $\sigma_Z \to 0$, the high-SNR regime tends to the noiseless scenario, and the hyperbola 'shrinks' towards its asymptotes: $\sigma_X + \tilde{w}_2 = \pm\tilde{w}_1 \tan\theta$. Figures 3, 4, and 5 shows the situation.

When P is small, the entire ball is contained 'inside' the hyperbola (*i.e.* in between the two branches of the hyperbola as depicted in Fig. 3): Whatever the embedding $(\tilde{w}_1, \tilde{w}_2)$, the false negative error exponent is zero. If P is big enough, the ball intersects the hyperbola and there are some embedding strategies $(\tilde{w}_1, \tilde{w}_2)$ which provide non zero error exponent (Fig. 5).

We are interested in the limit case when the ball has kissing points with the hyperbola (Fig. 4). The hyperbola is symmetric w.r.t. the axis $\{\tilde{w}_1 = 0\}$ and there are two kissing points (one on the

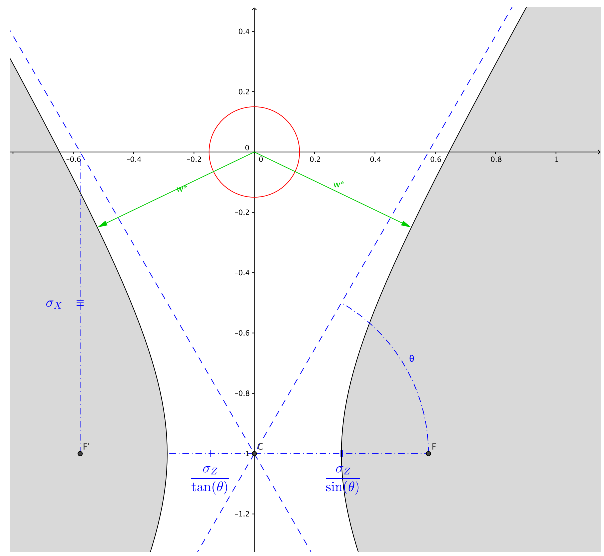

Figure 3: Conditions for $E_{\mathsf{fn}} \geq 0$ in the plane $(\tilde{w}_1, \tilde{w}_2)$: P is too small and $E_{\mathsf{fn}} = 0$. Setup: $\sigma_X = 1$, $\sigma_Z = 0.5$, and $\theta = \pi/3$. The distortion constraint defines the red circle of radius $\sqrt{P}$ centered on the origin $(0, 0)$; $E_{\mathsf{fn}} \geq 0$ defines the gray area outside the hyperbola of center $C = (0, -\sigma_X)$ and of asymptotes the dashed blue lines.

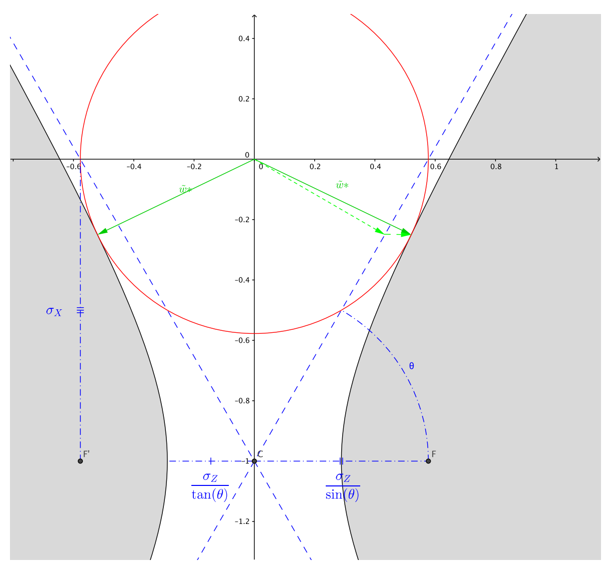

Figure 4: Conditions for $E_{\mathsf{fn}} \geq 0$ in the plane $(\tilde{w}_1, \tilde{w}_2)$: $P = \sigma_X^2 \cos^2\theta + \sigma_Z^2 \tan^{-2}\theta$. There are two kissing points given by (32) in green. Same setup as Fig. 3.

left hand side, the other on the right hand side of the hyperbola –Fig. 4). The system of equations provided by (28) (with equality) and $\tilde{w}_1^2 + \tilde{w}_2^2 = P$ implies that:

$$-(1+\tan^2(\theta)).\tilde{w}_2^2 - 2\sigma_X.\tilde{w}_2 + (P\tan^2(\theta) - \sigma_X^2 - \sigma_Z^2) = 0. \quad (31)$$

This polynomial of degree two in $\tilde{w}_2$ has a unique solution if and only if $P = P_0 + \sigma_Z^2 \tan^{-2}(\theta)$ with $P_0 := \sigma_X^2 \cos^2(\theta)$.

Consider the three following cases:

- if $P \leq P_0$ then the ball and the hyperbola never intersect for any σ_Z, including $\sigma_Z = 0$ (as in Fig. 3): $E_{\mathrm{fn}} = 0$ for any σ_Z. We rediscover result (8) from the noiseless scenario.
- if $P_0 < P \leq P_0 + \sigma_Z^2 \tan^{-2}(\theta)$, then $E_{\mathrm{fn}} = 0$ for that particular noise power σ_Z^2, but it might be strictly positive for a less harmful attack.
- if $P_0 + \sigma_Z^2 \tan^{-2}\theta < P$ (as in Fig. 5), then $E_{\mathrm{fn}} > 0$ for this noise power and, in this sense, the watermark is robust to that attack.

When we have exact equality $P = P_0 + \sigma_Z^2 \tan^{-2}\theta$ (as in Fig. 4), the two kissing points are given by (31):

$$(\tilde{w}_1^\star, \tilde{w}_2^\star) := \left(\pm\sqrt{P - \sigma_X^2 \cos^4\theta}, -\sigma_X \cos^2\theta\right). \quad (32)$$

A nice interpretation follows: if $P = P_0 + \delta P$ with $\delta P > 0$, then $\tilde{w}_1^{\star 2} = P_0 \sin^2\theta + \delta P$ and $\tilde{w}_2^{\star 2} = P_0 \cos^2\theta$. In words, the watermark signal first reaches the asymptotes of the hyperbola in order to guarantee $E_{\mathrm{fn}} > 0$ in the noiseless scenario. The 'shortest path' is to project $(0,0)$ on the asymptote by going along direction $(\sin\theta, -\cos\theta)$. This consumes the embedding power P_0. If it remains some extra embedding power $\delta P > 0$, the watermark signal carries on pushing the host signal *only* along the direction of the axis of the hypercone. This is depicted in Fig. 4.

The surprise is that this rediscovers the embedding shown to be 'optimum' in the noiseless scenario [3]. In the noisy scenario, this embedding can indeed be deemed as optimal as well with the following meaning: It is *not* the embedding that maximizes E_{fn} for a given σ_Z^2. This would certainly make $(\tilde{w}_1, \tilde{w}_2)$ a function of σ_Z^2, which violated the obliviousness of the embedder. On the contrary, (32) is independent of σ_Z^2. It is the embedding that makes $E_{\mathrm{fn}} > 0$ over the biggest noise power range $[0, R(\sigma_X^2, P, E)]$ with

$$R(\sigma_X^2, P, E) = \left|\frac{P}{e^{2E}-1} - \sigma_X^2 e^{-2E}\right|_+. \quad (33)$$

Interpretation: Again, we see that $R(\sigma_X^2, P, E) > 0$ if $P > P_0 = \sigma_X^2(1 - e^{-2E})$. When this is the case, the robustness is increasing with P, but decreasing with σ_X^2. The robustness is also a decreasing function of $E := (-\log P_{\mathrm{fp}})/n$: This scheme is extremely robustness for small E, *i.e.* very long signals. This complies with the well known rule of thumb in watermarking: The more spread, the more robust the watermark signal is. For a fixed P_{fp}, and long signals, we have:

$$R(\sigma_X^2, P, E) = \frac{nP}{2(-\log P_{\mathrm{fp}})} - \sigma_X^2 + o(n^{-1}). \quad (34)$$

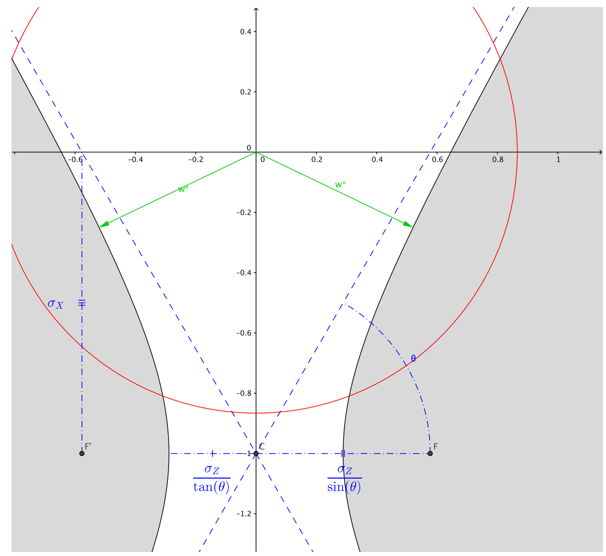

Figure 5: Conditions for $E_{\mathrm{fn}} \geq 0$ in the plane $(\tilde{w}_1, \tilde{w}_2)$: P is big enough so that there exist embeddings $(\tilde{w}_1, \tilde{w}_2)$ on the red circle inside the gray area. Same setup as Fig. 3.

5.2 Maximum E_{fp}

The previous section assumes that $E_{\mathrm{fp}} = E$, which fixes the semi-angle θ of the hypercone by (7), and establishes the expected robustness. This section provides another view: it assumes that $\sigma_Z^2 < R$ and looks for the biggest E_{fp} for which $E_{\mathrm{fn}} > 0$, *i.e.* E_{fp}^R. For a given P_{fp}, this gives a hint on the necessary vector length [15, Eq. (2) and below], $n > (-\log(P_{\mathrm{fp}}))/E_{\mathrm{fp}}^R$, to achieve a given robustness level.

This analysis is done for the watermarking strategy (32), whenever applicable (*i.e.* if $P > P_0$). Inequality (28) implies that error exponent E_{fn} is not null if:

$$\sigma_X^2 \sin^4\theta + (P + R - \sigma_X^2)\sin^2(\theta) - R \geq 0. \quad (35)$$

A special case is $R = 0$ and $P \geq \sigma_X^2$: The above inequality always holds which means that E_{fn} is not null for any the angle of the hypercone $\theta \in (0, \pi/2]$, and thus for any value of E_{fp}. This was shown in (8) and [3, 13].

Yet in the noisy scenario, (35) is a polynomial of degree two w.r.t. $\xi := \sin^2(\theta)$ which has two roots ξ_- and ξ_+ s.t. $\xi_- < 0 < \xi_+ < 1$ with:

$$\begin{aligned}\xi_+ &:= \frac{\sqrt{(P+R-\sigma_X^2)^2 + 4\sigma_X^2 R} - (P+R-\sigma_X^2)}{2\sigma_X^2} (\geq 0)\\ &= 1 - \frac{(P+R+\sigma_X^2) - \sqrt{(P+R+\sigma_X^2)^2 - 4P\sigma_X^2}}{2\sigma_X^2} (< 1)\end{aligned}$$

This polynomial takes positive values outside the interval $[\xi_-, \xi_+]$. This means that $E_{\mathrm{fn}} > 0$ if $\xi_+ < \sin^2\theta \leq 1$. This translates into the following right endpoint:

$$E_{\mathrm{fp}}^R = -\log\sqrt{\xi_+}. \quad (36)$$

Interpretation. : Again, as $R \to 0$ (*i.e.* in the noiseless setup), this E^R_{fp} tends to $+\infty$ if $P \geq \sigma_X^2$, or to $-1/2\log(1 - P/\sigma_X^2)$ if $P < \sigma_X^2$. Expression (36) is thus compliant with (11).

As a special case, let $R = \sigma_X^2$, *i.e.* the maximum noise power equals the power of the host signal. Then, the necessary vector length to achieve both the requirement under $\mathcal{H}_0$ (*i.e.* P_{fp}) and the target under $\mathcal{H}_1$ (*i.e.* $R = \sigma_X^2$) is (asymptotically):

$$n = \frac{2\log P_{\text{fp}}}{\log\left(\sqrt{1 + P^2/4\sigma_X^2} - P/2\sigma_X^2\right)} = \frac{4\sigma_X^2}{P}(-\log(P_{\text{fp}}) + o(1)). \quad (37)$$

Interpretation. : This formula is interesting in audio watermarking, where the rule of thumb is that $P/\sigma_X^2 = cst$, in the order of -20dB. This makes n in the order of some thousands.

6 COSTA'S BOUNDS

The previous section does not explicit the characteristic function $E_{\text{fn}} = F(E_{\text{fp}})$ but focuses on its feature E^R_{fn}. This is sufficient for assessing whether the watermark is robust according to definition 3.1 while operating at $E_{\text{fp}} = E$.

This section now compares the performances of the dual hypercone to bounds. It applies the idea of M. Costa in his famous paper [4][1]:

- **Non side-informed**: The lower bound is given by removing the dependence on $\vec{x}$ in the definition of the watermark signal. Now, the watermark signal is a fixed vector: $\vec{w} = \sqrt{nP}\vec{u}$. The received vector is $\vec{R} = \vec{w} + \vec{X} + \vec{Z}$.
- **Non blind**: The upper bound comes by giving the detector an advantage: it knows the value of $\vec{X}$. Removing this vector to received vector, we obtain $\vec{R} = \vec{w} + \vec{Z}$.

In both cases, the detector takes a decision based on $\vec{R} = \vec{w} + \vec{N}$, where $\vec{N}$ is a white Gaussian noise of variance $N = \sigma_X^2 + \sigma_Z^2$ (lower bound) or $N = \sigma_Z^2$ (upper bound).

We must not forget the specificities listed in Sect. 3.1. They make the detector oblivious to noise power N and watermark power P. This forbids the use of the Neyman-Pearson test $s(\vec{x}) = \vec{x}^\top\vec{u}$ because fixing the threshold fulfilling the requirement on the probability of false positive P_{fp} needs the knowledge of N. Again, we propose to use the scale invariant normalized correlation: $s(\vec{x}) = \vec{x}^\top\vec{u}/\|\vec{x}\|$. Note that this time the detection region is a single circular hypercone due to the absolute value operator missing.

C. E. Shannon already tackled the study of the probability P_{fn} in this case [15]. The error exponent is given by:

$$E_{\text{fn}} = \begin{cases} 0, & \text{if } 0 \leq \theta \leq \theta_0 \\ \frac{A^2}{2} - \frac{A}{2}G\cos\theta - \log(G\sin\theta) & \text{if } \theta_0 < \theta \leq \pi/2. \end{cases} \quad (38)$$

with $A := \sqrt{P/N}$, $\tan(\theta_0) = 1/A$, and $G = (A\cos(\theta)+\sqrt{A^2\cos^2(\theta)+4})/2$. It follows that:

$$E^R_{\text{fp}} = \frac{1}{2}\log\left(1 + \frac{P}{N}\right), \quad (39)$$

i.e. the capacity of a Gaussian channel. This is not a surprise. The semi-angle θ_0 defines the thinest cone for which the noise pushes the transmitted signal $\vec{w}$ outside with an exponentially vanishing probability P_{fn}. This probability represents the decoding error probability in a communication scenario. As for P_{fp}, it is the probability that a random vector falls into the hypercone under isotropic distribution. It equals the ratio of the solid angle of the hypercone and the one of the full hypersphere. E^R_{fp} can be though as representing, in logarithmic scale and per dimension, the number of hypercones with half angle θ_0 needed to fill the full hypersphere. In a communication scenario, this is the maximum number of messages (in logarithmic scale and per dimension) which can be reliably transmitted over this channel (*i.e.* with exponentially vanishing error probability).

[1] M. Costa used this idea for bounding the capacity of a side-informed communication channel. He never applied it to zero-bit watermarking, of course.

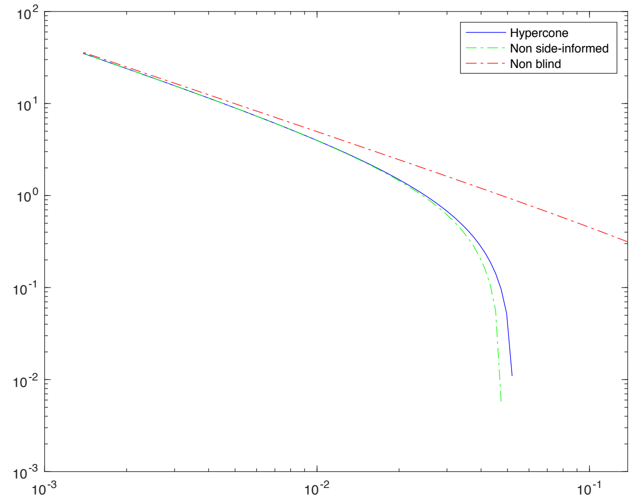

Figure 6: Robustness $R(\sigma_X^2, P, E)$ as a function of E for $\sigma_X^2 = 1$ and $P = 0.1$, *i.e.* function (33) **and bounds** (40)

The watermark is deemed robust if $E^R_{\text{fp}} > E$ for $N = \sigma_X^2 + \sigma_Z^2$ (lower bound) or $N = \sigma_Z^2$ (upper bound). The robustness, *i.e.* the value of σ_Z^2 making $E^R_{\text{fp}} = E$, is given by:

$$R(\sigma_X^2, P, E) = \begin{cases} \left|\frac{P}{e^{2E}-1} - \sigma_X^2\right|_+ & \text{non side-informed} \\ \frac{P}{e^{2E}-1} & \text{non blind} \end{cases} \quad (40)$$

Again by applying Shannon's argument (see Sect. 3.2), the necessary vector length for targeting a robustness equalling σ_X^2 is:

$$n = \begin{cases} \frac{2(-\log(P_{\text{fp}}))}{\log\left(1+P/2\sigma_X^2\right)} = \frac{4\sigma_X^2}{P}(-\log(P_{\text{fp}}) + o(1)) & \text{non side-informed} \\ \frac{2(-\log(P_{\text{fp}}))}{\log\left(1+P/\sigma_X^2\right)} = \frac{2\sigma_X^2}{P}(-\log(P_{\text{fp}}) + o(1)) & \text{non blind} \end{cases} \quad (41)$$

These expressions together with Figures 6 and 7 shows that the dual hypercone scheme has performances very close to the lower bound. When analysed under criteria and setups making sense in digital watermarking, this scheme is deceiving. Note that the authors of [3] claimed its optimality only in the noiseless scenario and under a so-called limited resources constraint.

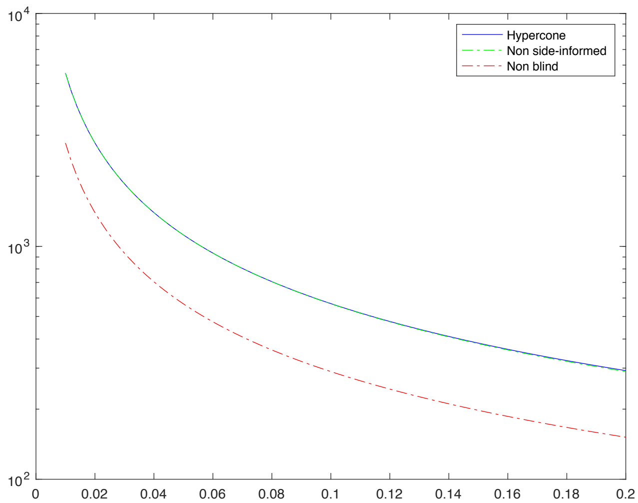

Figure 7: Necessary vector length n as a function of $P/\sigma_X^2 \in [0.01, 0.2]$, to achieve $P_{\mathrm{fp}} = 10^{-6}$ and $R(\sigma_X^2, P, E) = \sigma_X^2$, *i.e.* function (37) and bounds (41).

7 WATERMARKING REAL VECTORS

The above sections do not reveal how to watermark a host signal. They show that the error exponent E_{fn} is governed by the way the typical host vector is watermarked. Asymptotically, this typical host vector is such that $\vec{x}^\top\vec{u} = 0$ and $\|\vec{x}\|^2 = n\sigma_X^2$ because $(v_1, v_2) = (0, \sigma_X)$ in Sect. 5.1. In this case, the optimal watermark signal is given by (32). In practice, as n is not infinite, host vectors are different. How should they be watermarked?

The idea is to compute an error exponent $E_{\mathrm{fn}}(\vec{x})$ dedicated for that host vector. We no longer rely on a statistical model of the host (*i.e.* white Gaussian noise). It amounts to replace random variables (V_1, V_2) in appendix A.1 by their occurences $(v_1, v_2) = (\vec{x}^\top\vec{u}, \sqrt{\|\vec{x}\|^2 - \vec{x}^\top\vec{u}})/\sqrt{n}$. This modifies (42) to:

$$E_{\mathrm{fn}}(\vec{x}) = \min_{\mathcal{F}(\vec{x})} \frac{v_4^2 + v_5^2}{2\sigma_Z^2} + S\left(\frac{v_3^2}{\sigma_Z^2}\right), \tag{42}$$

while $\mathcal{F}(\vec{x})$ has the same definition (27). Cancelling $E_{\mathrm{fn}}(\vec{x})$ amounts to define a robustness level, whose maximisation defines the watermarking signal. This error exponent is null if and only if

$$R(\vec{x}, P, E) := \max_{\tilde{w}_1^2 + \tilde{w}^2 = P} (v_1 + \tilde{w}_1)^2 \tan^2(\theta) - (v_2 + \tilde{w}_2)^2 \le \sigma_Z^2. \tag{43}$$

Note that this robustness depends on E through the semi-angle θ (7). Having a priori $v_1 \neq 0$ prevents finding a close form. The optimal embedding has to be found numerically with a line search prototyping $(\tilde{w}_1, \tilde{w}_2) = \sqrt{P}(\cos\beta, \sin\beta)$ for $\beta \in [0, -\pi/2]$. Figure 8 shows that the difference with (32) is small. This means that (32) is a good approximation of the optimal embedding.

Note that this derivation is not rigorous: on one hand the host is not an infinite vector, one the other hand we compute error exponent, *i.e.* asymptotical quantities. Yet, it justifies an embedding proposed by M. Miller, I. Cox and J. Bloom nineteen years ago [11, Eq. 4], well before that the concept of error exponent was introduced in digital watermarking. This paper theoretically confirms the remarkable intuition of this research team.

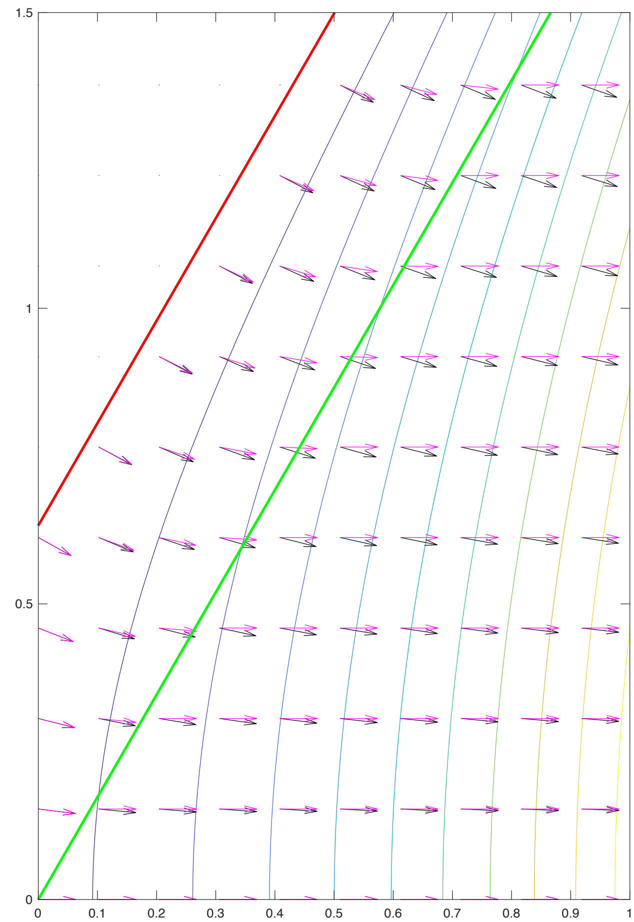

Figure 8: The watermark vector (w_1, w_2) is shown as a vector starting at point (v_1, v_2). They all have been scaled down to make the figure more visible. There is little difference between the numerical solution of (43) in black and the suboptimal embedding (32) in magenta. The green line represents the boundary of the hypercone. The red line is the boundary of $\mathcal{E}(P)$ (the rolling ball smoothing of the detection region). The color lines correspond to some level sets of $R(\vec{x}, P, E)$.

8 CONCLUSION

This paper completes the study of error exponents for the dual hypercone scheme. This scheme is important as it is one of the few schemes meeting the requirements on obliviousness at the embedding and detection sides. The paper extends the work of N. Merhav et al. to any SNR regime [3, 13]. It takes into account the specificities of digital watermarking. It introduces a new definition of robustness and the concept of necessary vector length to achieve false positive and robustness requirements.

In settings meaningful in digital watermarking, the asymptotical robustness is indeed not much bigger than the lower bound, *i.e.* the basic spread spectrum. This is in strong contrast with multi-bit watermarking, where M. Costa has shown that side-informed communication can perform as good as the upper bound. This is indeed due to the nature of zero-bit watermarking combined with the constraint of obliviousness. The open issue is whether there exist other detection regions granting obliviousness and closer to the upper bound. For instance, one can think of a union of thiner hypercones. Yet, their optimal number might be difficult to find.

A LAPLACE METHOD IN THE NOISY SCENARIO

A.1 Feasible set

Consider the basis $(\vec{e}_1, \dots, \vec{e}_n)$ of $\mathbb{R}^n$ where the first two vectors are defined in (24). We introduce the following random variables:

$$V_1 = (\vec{X}^\top \vec{e}_1)/\sqrt{n} \tag{44}$$
$$V_2 = (\vec{X}^\top \vec{e}_2)/\sqrt{n} \tag{45}$$
$$V_3 = \sqrt{\sum_{j=3}^{n} (\vec{Z}^\top \vec{e}_j)^2}/\sqrt{n} \tag{46}$$
$$V_4 = \vec{Z}^\top \vec{e}_2/\sqrt{n} \tag{47}$$
$$V_5 = \vec{Z}^\top \vec{e}_1/\sqrt{n} \tag{48}$$

V_1, V_4, and V_5 are Gaussian distributed while V_3 is a χ_{n-2} random variable scaled by $\sigma_Z/\sqrt{n}$. V_2 is a χ_{n-1} r.v. scaled by σ_X (by the definition of $\vec{e}_2$, $\vec{X}^\top\vec{e}_2$ is always positive). With this formulation, a false negative happens if:

$$\frac{(\vec{R}^\top \vec{u})^2}{\|\vec{R}\|^2} = \frac{(V_1 + \tilde{w}_1 + V_5)^2}{(V_1 + \tilde{w}_1 + V_5)^2 + (V_2 + \tilde{w}_2 + V_4)^2 + V_3^2} \leq \cos^2(\theta). \tag{49}$$

This means that the random vector $(V_1, \dots, V_5)$ lies in the domain $\mathcal{F} \subset \mathbb{R} \times \mathbb{R}^2_{\geq 0} \times \mathbb{R}^2$:

$$\mathcal{F} = \{(v_1 + \tilde{w}_1 + v_5)^2 \tan^2(\theta) \leq (v_2 + \tilde{w}_2 + v_4)^2 + v_3^2\} \tag{50}$$

A.2 Laplace potential functions

The random variables above defined are independent. The product of the their p.d.f. appears in the integral defining P_{fn}. When rewriting this integral in the form (21), K_n is thus the product of their multiplicative constants, whereas the potential function $h(\cdot)$ is the sum of their corresponding potential functions. This allows to deal case by case.

A.2.1 Gaussian distribution. The p.d.f. of V_1 for instance is $f_{V_1}(v) = \sqrt{n}e^{-v^2 n/2\sigma_X^2}/\sqrt{2\pi}\sigma_X$. This gives $K_n = \sqrt{n/2\pi}\sigma_X$ whose exponent is $\kappa = 0$, and the potential function $h(v) = -v^2/2\sigma_X^2$.

A.2.2 Chi distribution with fixed degree. The p.d.f. is given in (18). Its multiplicative constant is $K_n = 2(n/2)^{k/2}/\Gamma(k/2)\sigma$ whose exponent is $\kappa = 0$. The potential function is $h(v) = -v^2/2\sigma_X^2$.

A.2.3 Chi distribution with increasing degree. The p.d.f. of V_3 for instance is (18) where k is a function of n: $k = n-2$. Its multiplicative constant is $K_n = 2(n/2)^{n-2/2}/\Gamma(n-2/2)\sigma_Z$, whose exposant is $\kappa = -1/2$. The potential function is $h(v) = (v^2/\sigma_Z^2 - \log(v^2/\sigma_Z^2))/2$. In total, we get $S(v^2/\sigma_Z^2)$.

REFERENCES

[1] Mauro Barni and Franco Bartolini. 2004. *Watermarking Systems Engineering: Enabling Digital Assets Security and Other Applications* (1st ed.). Marcel Dekker.

[2] B. Chen and Gregory W. Wornell. 2001. Quantization index modulation: a class of provably good methods for digital watermarking and information embedding. *Information Theory, IEEE Transactions on* 47, 4 (May 2001), 1423–1443. https://doi.org/10.1109/18.923725

[3] P. Comesana, N. Merhav, and M. Barni. 2010. Asymptotically Optimum Universal Watermark Embedding and Detection in the High-SNR Regime. *Information Theory, IEEE Transactions on* 56, 6 (June 2010), 2804–2815. https://doi.org/10.1109/TIT.2010.2046223

[4] M.H.M. Costa. 1983. Writing on dirty paper (Corresp.). *Information Theory, IEEE Transactions on* 29, 3 (May 1983), 439–441. https://doi.org/10.1109/TIT.1983.1056659

[5] Ingemar J. Cox, Joe Kilian, F.T. Leighton, and T. Shamoon. 1997. Secure spread spectrum watermarking for multimedia. *Image Processing, IEEE Transactions on* 6, 12 (Dec 1997), 1673–1687. https://doi.org/10.1109/83.650120

[6] Ingemar J. Cox, M.L. Miller, and A.L. McKellips. 1999. Watermarking as communications with side information. *Proc. IEEE* 87, 7 (Jul 1999), 1127–1141. https://doi.org/10.1109/5.771068

[7] T. Furon and P. Bas. 2008. Broken arrows. *EURASIP Jounal on Information Security* 2008, ID 597040 (2008), doi:10.1155/2008/597040.

[8] Michel Ledoux. 1998. A Short Proof of the Gaussian Isoperimetric Inequality. In *High Dimensional Probability*, Ernst Eberlein, Marjorie Hahn, and Michel Talagrand (Eds.). Birkhäuser Basel, Basel, 229–232.

[9] Tie Liu and P. Moulin. 2003. Error exponents for one-bit watermarking. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, Vol. 3. III–65–8 vol.3. https://doi.org/10.1109/ICASSP.2003.1199108

[10] Tie Liu, P. Moulin, and R. Koetter. 2006. On error exponents of modulo lattice additive noise channels. *IEEE Transactions on Information Theory* 52, 2 (Feb 2006), 454–471. https://doi.org/10.1109/TIT.2005.862077

[11] M.L. Miller, Ingemar J. Cox, and J.A. Bloom. 2000. Informed embedding: exploiting image and detector information during watermark insertion. In *Image Processing, 2000. Proceedings. 2000 International Conference on*, Vol. 3. 1–4 vol.3. https://doi.org/10.1109/ICIP.2000.899260

[12] P. Moulin and R. Koetter. 2005. Data-Hiding Codes. *Proc. IEEE* 93, 12 (Dec 2005), 2083–2126. https://doi.org/10.1109/JPROC.2005.859599

[13] E. Sabbag and N. Merhav. 2006. Optimal Watermark Embedding and Detection Strategies Under Limited Detection Resources. In *Information Theory, 2006 IEEE International Symposium on.* 173–177. https://doi.org/10.1109/ISIT.2006.261759

[14] A. Seeger. 1997. Smoothing a nondifferentiable convex function: the technique of the rolling ball. *Revista de Matematicas Aplicadas* 18 (1997).

[15] C. E. Shannon. 1959. Probability of error for optimal codes in a Gaussian channel. *Bell System Tech. J.* 38 (1959), 611–656.

[16] Ram Zamir, Bobak Nazer, Yuval Kochman, and Ilai Bistritz. 2014. *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation and Multiuser Information Theory.* Cambridge University Press. https://doi.org/10.1017/CBO9781139045520