A constructive and unifying framework for zero-bit watermarking

Teddy Furon

Abstract

In the watermark detection scenario, also known as zero-bit watermarking, a watermark, carrying no hidden message, is inserted in content. The watermark detector checks for the presence of this particular weak signal in content. The article looks at this problem from a classical detection theory point of view, but with side information enabled at the embedding side. This means that the watermark signal is a function of the host content. Our study is twofold. The first step is to design the best embedding function for a given detection function, and the best detection function for a given embedding function. This yields two conditions, which are mixed into one 'fundamental' partial differential equation. It appears that many famous watermarking schemes are indeed solution to this 'fundamental' equation. This study thus gives birth to a constructive framework unifying solutions, so far perceived as very different.

Index Terms

Zero-bit watermarking, Pitman-Noether theorem, detection theory.

I. Introduction

In the past six years, side-informed embedding strategies have been shown to greatly improve watermark *decoding*. They exploit knowledge of the host signal during the construction of the watermark signal. The theory underlying these side-informed schemes was presented in the famous paper "Writing on Dirty paper" by M. Costa in 1983. Our work gives some theoretical aspect of the achievable performances when using side-information at the embedding side, as in Costa's correspondence, but for the watermark *detection* problem (a.k.a. zero-bit watermarking [1, Sect. 2.2.3]). This surprisingly received almost no study compared to the issue of watermark decoding, although it is perceived as a non trivial problem [2], [24]. Some other exceptions are works from M. Miller *et al.* (embedding cone) [3], JANIS [4] and watermark detection with distortion compensated dither modulation (DC-DM) schemes [5].

Manuscript submitted June 2006.

T. Furon is with the TEMICS project at INRIA. mail: teddy.furon@irisa.fr address: IRISA / TEMICS, Campus de Beaulieu, 35042 Rennes cedex. phone: +33 2 99 84 71 98. fax: +33 2 99 84 71 71.

A. Motivations from the application side

The trade-off between payload of the hidden message and robustness is a well known fact in watermarking. The main rationale for zero-bit watermarking is that maximum robustness that a watermarking primitive can inherently offer, is expected as the payload is reduced to the minimum. Here are two application scenarios where zero-bit watermarking might be sufficient, ie. it is not necessary to hide a message, but just the presence of a mark.

Some copy protection platforms [6] use watermarks as flags whose presence warns compliant devices that the piece of content they are dealing with, is a copyrighted material. Content access and copy protection are tackled by cryptographic primitives. Watermarking just prevents the 'analog hole' [7]–[9]. In other words, compliant devices expect three kinds of content: commercial content which are encrypted and watermarked, free content which are in the clear and not watermarked, and pirated content through the 'analog hole' which are in the clear but watermarked. Although most of DRM systems hide a message like a copy status, we have seen here that the presence of a mark is indeed sufficient.

Copyright protection is the most famous application of watermarking. However, hiding the name of the author in his Work is just a fact having no legal value. In Europe, the author first must be a member of an author society, then he registers his Work. The only legal proof is to give evidence that the suspicious image is indeed a version of a Work duly registered in an author society's database. Consequently, this is a yes/no question, which can be solved by detecting the presence or absence of a watermark previously embedded by an author society.

In these two applications, the presence of a watermark is not a secret, contrary to a steganographic scenario. The attacker obviously knows which content is watermarked. In the copy protection application, for instance, there is no point in attacking a personal video which is a free content, not protected neither by encryption nor by watermarking.

B. Motivations from the scientific side

Zero-bit watermarking is closely related to detection of weak signals in noisy environment: the watermark signal is embedded in a host signal, unknown to the detector. Its power is very weak compared the one of the host. Watermarkers resorted to classical elements of detection theory very early. This includes the use of Neyman-Pearson and Pitman-Noether theorems, calculus of efficacy, LMP tests (Locally Most Powerful) [10], and robust statistics [11].

The priority was at these times to design a better detector than the classical correlation, which is only optimal for white gaussian host signals. To name a few, this includes the works of teams such as Q. Cheng and T. Huang [12], A. Briassouli and M. Strinzis [13], M. Barni *et al.* [14]. They assume that the host signals are drawn from a known pdf (probability density function), and they apply the above-mentioned classical elements of detection theory. X. Huang and B. Zhang relax this implicit assumption considering that the 'real' pdf of the host belongs to a given family of distributions [15]. Their test is designed to fairly perform for the entire family. This allows to encompass attacks modifying the pdf within the family.

Another track is to see the host signal as a side information only available at the embedding. Side information brings huge improvements in watermark decoding. However, its use for zero-bit watermarking has received less interest. Pioneer works are mostly heuristic approaches [3], [16]. More recent works use the binning principle to achieve zero-bit watermarking [5], [17], although J. Eggers notices that SCS (Scalar Costa Scheme) is less efficient for zero-bit than for positive rate watermarking scheme [18, Sect. 3.6]. Indeed, Erez *et al.* prove the optimality of DC-DM based on Lattices (those whose Voronoi region asymptotically tends to an hypersphere) for strictly positive rate data hiding as far as an additive white gaussian noise attack is considered [19]. In the case of zero-rate watermarking, P. Moulin *et al.* reasonably conjecture that sparse Lattice DC-DM is optimal [20]. For zero-bit watermarking, Lattice DC-DM achieves high performances showing some host interference rejection [5]. However, there is a loss of efficacy compared to the private setup where the side information is also available at the detector.

At first glance, it would seem that the problem of watermark detection is simpler than the decoding of hidden symbols, because the decoder's output belongs to a message space which is bigger than the detector's range $\mathbb{B} = \{0, 1\}$. In other words, whereas watermark detection implies a simple binary hypotheses test, decoding of watermark is a complex multiple hypotheses test.

Yet, almost no theoretical limit, *ie.* an equivalent of Costa's result but for watermark detection, has been shown, except [21, Sect. 2] which only tackles the gaussian case. N. Merhav mentioned during the WaCha'05 workshop in Barcelona, that zero-bit watermarking is a hard problem whose optimal solution is not known for the moment [2]. Especially, up to now, there is no reason why the binning principle should be optimal, even if, as far as the author knows from the literature, it has the best performances against an AWGN attack. Yet, DC-DM schemes are known to be weak against scale gain attack.

II. Strategy and notation

Our goal is not to derive an accurate statistical model of the host signal as done in the above-mentioned prior works. On contrary, very basic assumptions (gaussian distribution or flat-host assumption) are in order, allowing us to stress the major role of side information at the embedding side. But, we consider that the binning principle is not the only way to enforce side information embedding. Our approach is indeed closer to the theory of weak signal detection.

A. Embedding side

The embedder transforms an original host signal $\mathbf{s}$ into a watermarked content $\mathbf{y} = \mathbf{f}(\mathbf{s}) = \mathbf{s} + \mathbf{x}$. The host signal or channel state $\mathbf{s}$ is a vector of n components of the original content, modeled as random variables. The notational key of the article is to decompose the watermark signal $\mathbf{x}$ as a unit power vector $\mathbf{w}$ and an amplitude θ.

$$\mathbf{f}(\mathbf{s}) = \mathbf{s} + \mathbf{x} = \mathbf{s} + \theta \mathbf{w}(\mathbf{s}). \tag{1}$$

$\mathbf{w}$ is a smooth function from $\mathbb{R}^n$ to $\mathbb{R}^n$, with the constraints $\mathrm{E}_{\mathbf{S}}\{\mathbf{w}(\mathbf{s})\} = 0$ and $\mathrm{E}_{\mathbf{S}}\{\|\mathbf{w}(\mathbf{s})\|^2\} = n$. It is a kind of direction pointing to an acceptance region of $\mathbb{R}^n$, towards which the host signal should be pushed. The scalar θ controls the gain or amplitude of the watermark signal. Theoretical frameworks often use a constant $\theta = \sqrt{P}$, where P is the fixed power of $\mathbf{x}$. Yet, in practice, host contents might support different watermark power depending on their individual masking property. This change might even occur within a content, such that we should resort to a vector $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)$ gathering positive and small gains affecting each sample. In this case, one might consider θ as the average gain.

Both parts of the watermark signal depends on the host content, either through side information, either for some perceptual reasons. Unfortunately, in blind schemes, side information is not made available at the detection side. Moreover, we wish to maintain a low detector's complexity; hence, no human visual or auditive system can recreate an estimate of θ. The only fact the detector knows is that the watermark amplitude θ is positive and small. We believe this model allows a great flexibility which eases practical implementations of watermarking schemes.

B. Detection side

Upon receipt of signal $\mathbf{r}$, the detector makes a binary decision: $d = 1$ ($d = 0$) means that, according to the detector, the piece of content under scrutiny is watermarked (resp. it has not been watermarked). There are two hypotheses: Under hypothesis $\mathcal{H}_0$, the detector receives an original content $\mathbf{r} = \mathbf{r}_0 = \mathbf{s}$ (see end of subsection I-A for justifications), whereas under hypothesis $\mathcal{H}_1$, the detector receives a watermarked and possibly attacked content $\mathbf{r} = \mathbf{r}_1$. Probability of false alarm P_{fa} and power of the test P_p are given by

$$P_{fa} = \Pr\{d = 1 | \mathcal{H}_0\} \quad ; \quad P_p = \Pr\{d = 1 | \mathcal{H}_1\}. \tag{2}$$

Once again, in zero-bit watermarking, no symbol is transmitted. Our problem is then fundamentally different from the communication of one bit because, under hypothesis $\mathcal{H}_0$, no processing is applied and $\mathbf{s}$, given by Nature, is directly sent to the detector.

We assume that the detector has the structure of a Neyman-Pearson test. First, it applies a detection function t mapping from $\mathbb{R}^n$ to $\mathbb{R}$. Then, this scalar is compared to a threshold τ: $d = 1$ if $t(\mathbf{r}) > \tau$, $d = 0$ else. The threshold is given by the constraint of a significance level α such that $P_{fa} = \mathrm{E}_D\{d | \mathcal{H}_0\} \leq \alpha$. Note that, for a given detection function, this threshold does not depend on what happens under hypothesis $\mathcal{H}_1$ (embedding function $\mathbf{w}$, watermark's amplitude θ). Moreover, we assume without loss of generality, that, under hypothesis $\mathcal{H}_0$, $t(\mathbf{r})$ is a centered random variable with unit variance:

$$\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r}) | \mathcal{H}_0\} = 0, \qquad \mathrm{Var}\{t(\mathbf{r}) | \mathcal{H}_0\} = 1. \tag{3}$$

If not the case, it is easy to built the test $\tilde{t}(\mathbf{r}) = (t(\mathbf{r}) - \mathrm{E}_{\mathbf{R}}\{t(\mathbf{r}) | \mathcal{H}_0\}) / \sqrt{\mathrm{Var}\{t(\mathbf{r}) | \mathcal{H}_0\}}$.

C. Pitman Noether efficacy

In this article, the tests are compared asymptotically for $n \rightarrow +\infty$. The Pitman-Noether theorem indicates that the best test has the higher efficacy η, whose general definition is given by [10, Sect. III.C.3]:

$$\bar{\eta} = \left(\lim_{n \rightarrow \infty} n^{-m\delta} \left. \frac{\partial^m}{\partial \theta^m} \mathrm{E}_{\mathbf{R}}\{t(\mathbf{r}) | \mathcal{H}_1\} \right|_{\theta = 0} \mathrm{Var}\{t(\mathbf{r}) | \mathcal{H}_0\}^{-1/2} \right)^{\frac{1}{m\delta}}, \tag{4}$$

where m is the first integer for which the m-th derivative of $\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r}) | \mathcal{H}_1\}$ is not null, and δ a positive scalar such that the limit is not null. In our problem, it is not unreasonable to assume $m = 1$ and $\delta = 1/2$ because the expectation of the detection function grows with $\sqrt{n}$ as $\mathrm{Var}\{t(\mathbf{r})|\mathcal{H}_0\}$ has been set to one for all n. This is at least true for well known watermarking schemes. We are not able to find a counter-example, ie. a watermarking scheme having a better growth rate than $\sqrt{n}$. Therefore, we restrict our analysis to $\delta = 1/2$.

The Pitman Noether theorem holds for composite one-sided hypothesis test. In Sect. II-A, motivations clearly show that our problem is not a simple hypothesis test ($\mathcal{H}_0 : \theta = 0$ versus $\mathcal{H}_1 : \theta = \sqrt{P}$ fixed), but a composite one-sided hypothesis test ($\mathcal{H}_0 : \theta = 0$ versus $\mathcal{H}_1 : \theta > 0$).

Last but not least, the proof of this theorem is based on an asymptotic study where the alternative hypothesis $\mathcal{H}_1$ has a vanishing parameter $\theta_n = kn^{-\delta}$, with k a positive constant. Important assumptions are the following regularity conditions:

$$\lim_{n\to\infty} \frac{\partial^m}{\partial\theta^m} \mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}\bigg|_{\theta=\theta_n} = \frac{\partial^m}{\partial\theta^m} \mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}\bigg|_{\theta=0} \quad \text{and} \quad \lim_{n\to\infty} \mathrm{Var}\{t(\mathbf{r}|\mathcal{H}_1)\} = \mathrm{Var}\{t(\mathbf{r})|\mathcal{H}_0\}, \tag{5}$$

and that $t(\mathbf{r}) - \mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})\}$ tends, as $n \to \infty$, to the normal distribution with mean 0 and variance 1, both under $\mathcal{H}_1$ and under $\mathcal{H}_0$.

We also define the efficiency per element (a.k.a. the differential detector SNR) in the same way as the efficacy but without the limit, such that in our case:

$$\eta = \frac{1}{n}\left[\frac{\partial}{\partial\theta}\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}\right]^2_{\theta=0}. \tag{6}$$

III. Detection of weak signal dependent on side information

The goal of this section is to give the expressions for the best detection and the best embedding functions. We mean ‘best’ in the sense of the Pitman Noether theorem, ie. such as they maximized the efficiency per element.

This section doesn’t consider any attack. Hence, the Pitman Noether theorem considers signals $r_0 = \mathbf{s}$ and $\mathbf{r}_1 = \mathbf{y} = \mathbf{s} + \theta_n \mathbf{w}(\mathbf{s})$, with $\mathrm{E}_{\mathbf{S}}\{\|\mathbf{w}(\mathbf{s})\|^2\} = n$ and $\theta_n = k/\sqrt{n}$, $k > 0$. It means that the proof of this theorem fixes the embedding distortion to $D_E = \theta_n^2 n = k^2$, but as n increases, the power of the watermarking signal vanishes.

A. Best detector for a given embedding function

In this subsection, embedding function $\mathbf{w}$ is fixed. A well known corollary of the Pitman Noether theorem [10, Sect. III.C.3] states that the Locally Most Powerful (LMP) test in $\theta = 0$ is asymptotically the best. A Cauchy-Schwarz inequality gives:

$$\begin{aligned} \frac{\partial}{\partial\theta}\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}\bigg|_{\theta=0} &= \int_{\mathbb{R}^n} t(\mathbf{r})\, \frac{\partial}{\partial\theta} p(\mathbf{r}|\mathcal{H}_1)\bigg|_{\theta=0} d\mathbf{r} && (7)\\ &\leq \sqrt{\int_{\mathbb{R}^n} t(\mathbf{r})^2 p(\mathbf{r}|\mathcal{H}_0) d\mathbf{r}} \sqrt{\int_{\mathbb{R}^n} p(\mathbf{r}|\mathcal{H}_0)\left(\frac{1}{p(\mathbf{r}|\mathcal{H}_0)}\, \frac{\partial}{\partial\theta} p(\mathbf{r}|\mathcal{H}_1)\bigg|_{\theta=0}\right)^2 d\mathbf{r}} && (8)\\ &= \sqrt{\int_{\mathbb{R}^n} p(\mathbf{r}|\mathcal{H}_0)\left(\frac{1}{p(\mathbf{r}|\mathcal{H}_0)}\, \frac{\partial}{\partial\theta} p(\mathbf{r}|\mathcal{H}_1)\bigg|_{\theta=0}\right)^2 d\mathbf{r}}, && (9) \end{aligned}$$

with equality for the LMP test:

$$t(\mathbf{r}) = k_t \frac{1}{p(\mathbf{r}|\mathcal{H}_0)} \left. \frac{\partial p(\mathbf{r}|\mathcal{H}_1)}{\partial \theta} \right|_{\theta=0}, \tag{10}$$

where k_t is a positive constant whose role is explained below. The use of the LMP with $\theta = 0$ is reinforced in practice by the fact the watermark power is very weak compared to the host power.

When there is no attack, $p(\mathbf{r}|\mathcal{H}_0) = p_{\mathbf{S}}(\mathbf{r})$ and $p(\mathbf{r}|\mathcal{H}_1) = p_{\mathbf{Y}}(\mathbf{r})$. We assume there exists $\bar{\theta} > 0$, such that function $\mathbf{f}(\mathbf{s})$ is invertible at least when $0 \leq \theta \leq \bar{\theta}$: $\mathbf{s} = \mathbf{f}^{-1}(\mathbf{y})$. This allows to write $p_{\mathbf{Y}}(\mathbf{r}) = p_{\mathbf{S}}(\mathbf{f}^{-1}(\mathbf{r}))|J_{\mathbf{f}^{-1}}(\mathbf{r})|$, with the last term being the determinant of the Jacobian matrix of $\mathbf{f}^{-1}$ taken at $(\mathbf{r}, \theta)$. Developing this last equation (see Appendix I), we finally get these expressions:

$$t(\mathbf{r}) = -k_t \frac{\nabla p_{\mathbf{S}}(\mathbf{r})^T}{p_{\mathbf{S}}(\mathbf{r})} \mathbf{w}(\mathbf{r}) - k_t \text{div}(\mathbf{w}(\mathbf{r})) \tag{11}$$

$$= -k_t \frac{\text{div}(p_{\mathbf{S}}(\mathbf{r})\mathbf{w}(\mathbf{r}))}{p_{\mathbf{S}}(\mathbf{r})}. \tag{12}$$

The first term of (11) corresponds to the classical non-linear correlation based LMP test [12]–[14], whereas the second term is not null whenever side information is enabled at the embedding side.

Let $\mathcal{B}_n(R)$ be the ball of radius R centered on $\mathbf{0}$, $\mathcal{S}_n(R)$ the associated hypersphere, and $E(R) = \int_{\mathcal{B}_n(R)} t(\mathbf{r}) p_{\mathbf{S}}(\mathbf{r}) d\mathbf{r}$. Then, thanks to the Gauss theorem, we have

$$|E(R)| = k_t \left| \int_{\mathcal{B}_n(R)} \text{div}(p_{\mathbf{S}}(\mathbf{r})\mathbf{w}(\mathbf{r})) d\mathbf{r} \right| \tag{13}$$

$$= k_t \left| \int_{\mathcal{S}_n(R)} p_{\mathbf{S}}(\mathbf{r})\mathbf{w}(\mathbf{r})^T \mathbf{e}(\mathbf{r}) d\mathbf{r} \right| \tag{14}$$

$$\leq k_t \int_{\mathcal{S}_n(R)} p_{\mathbf{S}}(\mathbf{r}) \|\mathbf{w}(\mathbf{r})\| d\mathbf{r}, \tag{15}$$

with $\mathbf{e}(\mathbf{r})$ the unit normal vector at position $\mathbf{r}$ on $\mathcal{S}_n(R)$. $E\{\|\mathbf{w}(\mathbf{r})\|^2\} < \infty$ implies that $\lim_{R \to +\infty} E(R) = 0$. Hence, the detection function given by (12) is centered under hypothesis $\mathcal{H}_0$, as required in II-B. The constant k_t enforces that Var$\{t(\mathbf{r})|\mathcal{H}_0\} = 1$:

$$k_t = \left(\int_{\mathbb{R}^n} \frac{1}{p(\mathbf{r}|\mathcal{H}_0)} \left[\frac{\partial p(\mathbf{r}|\mathcal{H}_1)}{\partial \theta} \right]^2_{\theta=0} d\mathbf{r} \right)^{-1/2}. \tag{16}$$

Finally, (6), (10) and (16) give the efficiency per element for such tests:

$$\eta = n^{-1} k_t^{-2} \tag{17}$$

B. Best embedding function for a given detection function

The detection function t being given (such that $t(\mathbf{r}_0)$ is a centered random variable with unit variance), we write:

$$\left. \frac{\partial}{\partial \theta} \mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\} \right|_{\theta=0} = \mathrm{E} \left\{ \left. \frac{\partial}{\partial \theta} t(\mathbf{s} + \theta \mathbf{w}(\mathbf{s})) \right|_{\theta=0} \right\} \tag{18}$$

$$= \mathrm{E}_{\mathbf{S}}\{\mathbf{w}(\mathbf{s})^T \nabla t(\mathbf{s})\}. \tag{19}$$

It appears that, for a given t, it is important to let $\mathbf{w}(\mathbf{s}) \propto \nabla t(\mathbf{s})$, $\forall \mathbf{s} \in \mathbb{R}^n$. The efficacy is then upper bounded by the following Cauchy-Schwarz inequality:

$$\eta = \frac{1}{n}\left(\int_{\mathbb{R}^n} p_{\mathbf{S}}(\mathbf{s})\|\mathbf{w}(\mathbf{s})\|\|\nabla t(\mathbf{s})\|d\mathbf{s}\right)^2 \leq \int_{\mathbb{R}^n} p_{\mathbf{S}}(\mathbf{s})\|\nabla t(\mathbf{s})\|^2 d\mathbf{s} \tag{20}$$

with equality when:

$$\mathbf{w}(\mathbf{s}) = k_w \nabla t(\mathbf{s}) \quad \forall \mathbf{s} \in \mathbb{R}^n, \tag{21}$$

where k_w is a normalizing constant to achieve $\mathrm{E}_{\mathbf{S}}\{\|\mathbf{w}(\mathbf{s})\|^2\} = n$:

$$k_w = \sqrt{n/\mathrm{E}_{\mathbf{S}}\{\|\nabla t(\mathbf{s})\|^2\}}. \tag{22}$$

(20) and (22) give the efficiency per element for such tests:

$$\eta = n k_w^{-2} = \mathrm{E}_{\mathbf{S}}\{\|\nabla t(\mathbf{s})\|^2\}. \tag{23}$$

C. *Synthesis*

For the moment, we know how to design the best embedding function for a given detection function, and how to design the best detection function for a given embedding function. This remembers the Lloyd-Max algorithm in quantization. However, dealing with closed form equations, we can insert (21) in (12) yielding a partial differential equation, that we loosely name ‘fundamental equation of zero-bit watermarking’:

$$p_{\mathbf{S}}(\mathbf{r})t(\mathbf{r}) + k_t k_w \mathrm{div}(p_{\mathbf{S}}(\mathbf{r})\nabla t(\mathbf{r})) = 0 \quad \forall \mathbf{r} \in \mathbb{R}^n. \tag{24}$$

Hence, the best couple of detection/embedding functions $\{t, \mathbf{w}\}$ is $\{t^\star, k_w \nabla t^\star\}$, with $t^\star$ a fundamental solution, ie. a solution of (24). Note that (17) and (23) are still valid. Therefore, it is possible to build a scheme of a given η (virtually, as high as possible), provided (24) admits a solution with $k_w k_t = \eta^{-1}$. The fundamental equation can also be written as:

$$\eta t(\mathbf{r}) + \frac{\nabla p_{\mathbf{S}}(\mathbf{r})^T}{p_{\mathbf{S}}(\mathbf{r})}\nabla t(\mathbf{r}) + \nabla^2 t(\mathbf{r}) = 0, \tag{25}$$

$\nabla^2 t(\mathbf{r})$ being the Laplacian of $t(\mathbf{r})$.

D. *A remarkable property of fundamental solutions*

A nice property induced by the fundamental equation is that a pair of its solutions with different efficiency are orthonormal for the scalar product $\langle ., . \rangle$ defined here for two functions g and h by:

$$\langle g, h \rangle = \mathrm{E}_{\mathbf{R}}\{g(\mathbf{r})h(\mathbf{r})|\mathcal{H}_0\}. \tag{26}$$

Denote $L[t] = \mathrm{div}(p_{\mathbf{S}}(\mathbf{r})\nabla t(\mathbf{r}))$. This differential operator is symmetric if $\int_{\mathbb{R}^n} t_i(\mathbf{r})L[t_j](\mathbf{r})d\mathbf{r} = \int_{\mathbb{R}^n} L[t_i](\mathbf{r})t_j(\mathbf{r})d\mathbf{r}$. In our case,

$$\int_{\mathbb{R}^n} t_i(\mathbf{r})L[t_j](\mathbf{r})d\mathbf{r} - \int_{\mathbb{R}^n} t_j(\mathbf{r})L[t_i](\mathbf{r})d\mathbf{r} = \int_{\mathbb{R}^n} \mathrm{div}(p_{\mathbf{S}}(\mathbf{r})(t_i(\mathbf{r})\nabla t_j(\mathbf{r}) - t_j(\mathbf{r})\nabla t_i(\mathbf{r})))d\mathbf{r}. \tag{27}$$

The symmetry is enabled for functions t_i, t_j if the last term, denoted by C, is zero. Let us write it as a limit:

$$\begin{aligned} C &= \int_{\mathbb{R}^n} \mathrm{div}(p_{\mathbf{S}}(\mathbf{r})(t_i(\mathbf{r})\nabla t_j(\mathbf{r}) - t_j(\mathbf{r})\nabla t_i(\mathbf{r})))d\mathbf{r} && (28)\\ &= \lim_{R\to\infty} \int_{\mathcal{B}_n(R)} \mathrm{div}(p_{\mathbf{S}}(\mathbf{r})(t_i(\mathbf{r})\nabla t_j(\mathbf{r}) - t_j(\mathbf{r})\nabla t_i(\mathbf{r})))d\mathbf{r} && (29)\\ &= \lim_{R\to\infty} \int_{\mathcal{S}_n(R)} p_{\mathbf{S}}(\mathbf{r})(t_i(\mathbf{r})\nabla t_j(\mathbf{r})^T\mathbf{e}(\mathbf{r}) - t_j(\mathbf{r})\nabla t_i(\mathbf{r})^T\mathbf{e}(\mathbf{r}))d\mathbf{r} && (30)\end{aligned}$$

The Gauss theorem gives the later equation. Assuming that the pdf of the host vanishes more quickly than the norm $\|t_i(\mathbf{r})\nabla t_j(\mathbf{r})\|$, we suppose in the sequel that the symmetry property is enabled for the solutions of the fundamental equation. Then, (24) in (27) gives

$$\begin{aligned} \int_{\mathbb{R}^n} t_i(\mathbf{r})L[t_j](\mathbf{r})d\mathbf{r} - \int_{\mathbb{R}^n} t_j(\mathbf{r})L[t_i](\mathbf{r})d\mathbf{r} &= -\int_{\mathbb{R}^n} t_i(\mathbf{r}).\eta_j p_{\mathbf{S}}(\mathbf{r})t_j(\mathbf{r})d\mathbf{r} + \int_{\mathbb{R}^n} t_j(\mathbf{r}).\eta_i p_{\mathbf{S}}(\mathbf{r})t_i(\mathbf{r})d\mathbf{r} && (31)\\ &= (\eta_i - \eta_j)\langle t_i, t_j\rangle = 0 && (32)\end{aligned}$$

The restriction to normalized detection functions and this last equation imply that $\langle t_i, t_j\rangle = \delta(j-i)$ where δ is the Kronecker delta function. Hence, the solutions of the fundamental equation with different efficiencies per element constitute a family of orthonormal functions (Subsection IV-B.1 even shows orthonormal functions sharing the same efficiency), if the symmetry property holds for all pairs of elements of this family.

IV. Some solutions of the fundamental equation of zero-bit watermarking

We are not able to find a general solution of the fundamental equation. However, in some cases, we show some examples of solution in this section.

A. *The scalar case*

To avoid multiplication of notation, we use the same letter to denote the scalar version of above-mentioned vectorial functions.

We suppose here that the host samples are i.i.d. such that $p_{\mathbf{S}}(\mathbf{s}) = \prod_{i=1}^n p_S(s_i)$. Moreover, our strategy is to maintain this statistical independence while embedding the watermark: $\mathbf{w}(\mathbf{s}) = (\epsilon_1 w(s_1), \cdots, \epsilon_n w(s_n))^T$, where $\boldsymbol{\epsilon}$ is a secret vector, with for instance, $\epsilon_i = \pm 1\,\forall i \in \{1, \cdots, n\}$. (11) shows that the detection function is indeed a sum $t(\mathbf{r}) = \sum_{i=1}^n \epsilon_i t(r_i)$; and (25) boils down to a scalar second-order ordinary differential equation with non constant coefficients:

$$\eta t(r) + \frac{p_S'(r)}{p_S(r)}t'(r) + t''(r) = 0. \tag{33}$$

1) Gaussian case: Assume that $s \sim \mathcal{N}(0, \sigma_x^2)$. (25) becomes even simpler: $\eta t(r) - rt'(r)/\sigma_x^2 + t''(r) = 0$. The solution is a linear combination of two 'independent' (*ie.* their Wronskian is not null) confluent hypergeometric functions of the first kind taken in $r^2/2$:

$$\begin{aligned} t^{(a)}(r) &= k_{t_1}.{}_1F_1\left(-\frac{\sigma_x^2\eta}{2}, \frac{1}{2}, \frac{r^2}{2\sigma_x^2}\right), && (34)\\ t^{(b)}(r) &= k_{t_2}.r.{}_1F_1\left(\frac{1-\sigma_x^2\eta}{2}, \frac{3}{2}, \frac{r^2}{2\sigma_x^2}\right). && (35)\end{aligned}$$

TABLE I

POLYNOMIAL SOLUTIONS OF THE SCALAR GAUSSIAN CASE $s \sim \mathcal{N}(0,1)$.

η	$w(s)$	$t(r)$	Var$\{t(r)\|\mathcal{H}_1\}$
1	1	r	1
2	s	$\frac{-1+r^2}{\sqrt{2}}$	$(1+\theta)^4$
3	$\frac{-1+s^2}{\sqrt{2}}$	$\frac{-3r+r^3}{\sqrt{6}}$	$1+66\theta^2+O(\theta^4)$
4	$\frac{-3s+s^3}{\sqrt{6}}$	$\frac{3-6r^2+r^4}{2\sqrt{6}}$	$1+12\sqrt{6}\theta+608\theta^2+O(\theta^3)$
5	$\frac{3-6s^2+s^4}{2\sqrt{6}}$	$\frac{15r-10r^3+r^5}{2\sqrt{30}}$	$1+5470\theta^2+O(\theta^4)$
6	$\frac{15s-10s^3+s^5}{2\sqrt{30}}$	$\frac{-15+45r^2-15r^4+r^6}{12\sqrt{5}}$	$1+40\sqrt{30}\theta+49122\theta^2+O(\theta^3)$
7	$\frac{-15+45s^2-15s^4+s^6}{12\sqrt{5}}$	$\frac{-105r+105r^3-21r^5+r^7}{12\sqrt{35}}$	$1+441392\theta^2+O(\theta^4)$

If $\sigma_x^2\eta$ is an even integer, $t^{(a)}$ is a polynomial function. If $\sigma_x^2\eta$ is an odd integer, $t^{(b)}$ is a polynomial function. Another way to see this is to recognize this later differential equation as the Hermite equation when η is a positive integer and $\sigma_x^2 = 1$. Therefore, if $\eta\sigma_x^2 = k \in \mathbb{N}$, $t_k(r) = \kappa_k H_k(r/\sigma_x)$, H_k being the Hermite polynomial of order k. This family of polynomials is known to be orthogonal with a weighting function[1] $\exp(-r^2/2)$. In our context, this is confirmed by (30), which reduces to the value of the integrand on the boundaries on an increasing interval of $\mathbb{R}$. The condition $C = 0$ is satisfied because $\lim_{r\to\infty} r^m \exp(-r^2/2\sigma_x^2) = 0, \forall m \in \mathbb{N}$. In the sequel, we call this set of fundamental solutions the 'polynomial family'.

Table IV-A.1 gives the expressions of the first elements of this family and their associated embedding function. The first line of this table is the well known direct spread spectrum scheme with a linear correlator, optimal detector in the Gaussian i.i.d. case. The second line is known as the proportional or multiplicative embedding, first proposed in [22, Sect. 4.2] for perceptual reason. A higher efficacy is another inherent advantage of proportional embedding. The remaining lines of this table generalize this idea to new schemes (as far as the author knows).

2) Uniform case: The classical 'flat-host' assumption used in DC-DM scheme studies states that the host pdf is a piecewise constant function. More precisely, we assume here the host pdf can be written as $p_S(s) = \sum_{i=-\infty}^{+\infty} P_i \Pi_i(s)$, with Π_i the indicator function of the elementary interval $[\frac{\pi}{\sqrt{\eta}}i, \frac{\pi}{\sqrt{\eta}}(i+1))$, and $\sum_{i=-\infty}^{+\infty} P_i = \sqrt{\eta}/\pi$. In this case, (25) defined almost everywhere[2], is a lot simpler: $\eta t(r) + t''(r) = 0$, whose obvious solution[3] is $t(r) = \sqrt{2}\cos(\sqrt{\eta}r)$ and hence, $w(s) = -\sqrt{2}\sin(\sqrt{\eta}s)$. Although these are not exactly the sawtooth embedding function of the scalar DC-DM (a.k.a. SCS), we find back at least periodic functions.

If the 'flat-host' assumption holds on the above partition of $\mathbb{R}$, then it also holds on the finer partition $\bigcup_{i=-\infty}^{+\infty}[\frac{\pi}{k\sqrt{\eta}}i, \frac{\pi}{k\sqrt{\eta}}(i+1))$, $k \in \mathbb{N}$. This gives birth to another fundamental solution $t_k(r) = \sqrt{2}\cos(k\sqrt{\eta}r)$, whose efficacy is k^2 greater. We call the sinusoidal family the set of fundamental solutions t_k indexed with integers. Once again, elements of

[1] This is the probabilists' definition of Hermite polynomials. However, these polynomials take different form according to the chosen standardization. For instance, $\kappa_k = 1/\sqrt{k!}$ when the coefficient of highest order of H_k is set to 1.

[2] Except on the boundaries due to discontinuities. This has little importance as the probability that the host signal is on a boundary is zero.

[3] The other solution $\{t(r) = \sqrt{2}\sin(\sqrt{\eta}r), w(s) = \sqrt{2}\cos(\sqrt{\eta}s)\}$ is valid on a shifted partition $\bigcup_i[\frac{\pi}{2\sqrt{\eta}}(2i-1), \frac{\pi}{2\sqrt{\eta}}(2i+1))$.

this family are orthonormal:

$$\langle t_k, t_\ell \rangle = \sum_i 2P_i \int_{i\frac{\pi}{\sqrt{\eta}}}^{(i+1)\frac{\pi}{\sqrt{\eta}}} \cos(k\sqrt{\eta}r)\cos(\ell\sqrt{\eta}r)dr = \delta(k-\ell). \tag{36}$$

B. The vector case

IV-A uses the cartesian system where the embedding processes in a sample wise manner. We generalize this idea to block based watermark schemes assuming there exists an integer p dividing n so that $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^p \cdots \times \mathbb{R}^p$ and that $p_{\mathbf{S}}(\mathbf{s}) = \prod_{i=1}^{n/p} p(s_{(i-1)p+1}, \cdots, s_{(i-1)p+p})$. If $t^{(p)}$ is a solution of the fundamental equation in $\mathbb{R}^p$ with a given efficacy, then $t^{(n)}(\mathbf{r}) = \sqrt{p/n}\sum_{i=1}^{n/p} t^{(p)}(r_{(i-1)p+1}, \cdots, r_{(i-1)p+p})$ is a solution in $\mathbb{R}^n$ yielding the same efficacy. This realizes a statistically independent embedding in the sense that the block of p watermark samples only depends on the same block of p host samples. The issue is now on finding solutions $t^{(p)}$. A usual technique is the separation of variables method in a specific orthogonal coordinate system [23].

1) Separation of variables: Classically, the separation of variables method considers a solution $t^{(p)}(\mathbf{r}) = \prod_{i=1}^{p} t_{\eta_i}(r_i)$, where each t_{η_i} have to satisfy (33) with their own efficacy η_i. The resulting efficacy of $t^{(p)}$ is then $\eta = \sum_{i=1}^{p} \eta_i$. For white gaussian hosts, this gives birth to an extension of the polynomial family which is indeed based on the multivariate Hermite polynomials, indexed by the n/p-uple $\mathbf{k} \in \mathbb{N}^p$: $H_{\mathbf{k}}(\mathbf{r}) = \prod_{i=1}^{n/p} H_{k_i}(r_i)$. Two different elements of this family are orthogonal for the scalar product (26), even having the same efficacy.

This extension of the polynomial family is illustrated in the following example. If $\mathbf{S} \sim \mathcal{N}(\mathbf{0}, \sigma_x^2 \mathbf{I}_n)$, then $\nabla p_{\mathbf{S}}(\mathbf{r}) = -p_{\mathbf{S}}(\mathbf{r})\mathbf{r}/\sigma_x^2$, and (24) becomes $\eta t(\mathbf{r}) - \mathbf{r}^T \nabla t(\mathbf{r})/\sigma_x^2 + \nabla^2 t(\mathbf{r}) = 0$. JANIS, a zero-bit watermarking scheme invented heuristically some years ago [4], [16], is a fundamental solution. Its detection function is the following one:

$$t(\mathbf{r}) = \sqrt{\frac{p}{n}} \sum_{i=1}^{n/p} \prod_{j=1}^{p} \frac{r_{(i-1)p+j}}{\sigma_x}. \tag{37}$$

Note that r_j appears only once in the detection function, $\forall j \in \{1, \cdots, n\}$. It is easy to see that $\mathbf{r}^T \nabla t(\mathbf{r}) = pt(\mathbf{r})$ and $\nabla^2 t(\mathbf{r}) = 0$. Thus, JANIS with order p is a solution to (24) provided that $\eta\sigma_x^2 = p$. This can be interpreted as follows: this is a block based watermarking scheme built on the p multivariate Hermite polynomial $H_{(1,\cdots,1)}$. This theoretical framework proves the optimality of the heuristic JANIS scheme.

Separation of variables can be done on another coordinate system. The following spherical coordinate system $(\rho, \theta_1, \cdots, \theta_{p-1})$ is adapted to isotropic host distributions, ie. $p_{\mathbf{S}}(\mathbf{s}) = f(\rho)$ with $\rho = \|\mathbf{s}\|$:

$$\begin{aligned}
r_1 &= \rho \sin\theta_{p-1} \sin\theta_{p-2} \cdots \sin\theta_2 \sin\theta_1 \\
r_2 &= \rho \sin\theta_{p-1} \sin\theta_{p-2} \cdots \sin\theta_1 \cos\theta_1 \\
r_3 &= \rho \sin\theta_{p-1} \sin\theta_{p-2} \cdots \cos\theta_2 \\
&\vdots \\
r_{p-1} &= \rho \sin\theta_{p-1} \cos\theta_{p-2} \\
r_p &= \rho \cos\theta_{p-1}.
\end{aligned}$$

For instance, we seek a function $t(\mathbf{r}) = t(\rho, \theta_{p-1}) = U(\rho)V(\theta_{p-1})$, which depends on two simple statistics $\rho = \sum_{i=1}^{p} r_i^2$ and $\theta_{p-1} = \arccos(\mathbf{r}^T\mathbf{e}_p/\|\mathbf{r}\|)$. $\mathbf{e}_p$ is a secret unit vector shared by the embedder and the detector taken as the p-th element of the canonical basis (ie. in the cartesian coordinate system). Separating variables in (25) yields two equations:

$$KV(\theta) + (p-2)\cot\theta V'(\theta) + V''(\theta) \quad = \quad 0 \tag{38}$$

$$(\eta\rho^2 - K)U(\rho) + \left((p-1)\rho + \rho^2\frac{f'(\rho)}{f(\rho)}\right)U'(\rho) + \rho^2 U''(\rho) \quad = \quad 0 \tag{39}$$

with $K \in \mathbb{R}$. The choice $U(\rho) = k_t\rho^2$ and $V(\theta) = p\cos^2\theta - 1$ is a solution provided $f'(\rho)/f(\rho) = -\rho/\sigma_x^2$ (white Gaussian host), $K = 2p$ and $\eta\sigma_x^2 = 2$. The detection function is then

$$t(\mathbf{r}) = k_t\left((\sqrt{p}\mathbf{r}^T\mathbf{e}_p)^2 - \|\mathbf{r}\|^2\right) = \frac{1}{\sigma_x^2\sqrt{2p(p-1)}}\left((p-1)r_p^2 - \sum_{i=1}^{p-1} r_i^2\right). \tag{40}$$

$t(\mathbf{r}) = \tau$ defines a p-dimensional two-sheet hyperboloid. This is closed to a two-sheet hypercone, acceptance region of the absolute normalized correlation, which is the optimum detection function based on such simple statistics for gaussian white host [24]. We agree here with N. Merhav and E. Sabbag that the acceptance region must be a two-sheet geometric form contrary to the well-known normalized correlation and its one-sheet hypercone [1]. Yet, neither the absolute normalized correlation nor the famous normalized correlation are fundamental solutions. Eq.(40) is however not unknown in the watermarking literature. This is the measure of robustness given in Cox *et al.* book [1, Eq.(5.13)].

Let us now invent a host such that

$$P(\mathbf{s} \in \mathcal{B}_p(R)) = \begin{cases} R/R_0 & \text{, if } R \leq R_0 \\ 1 & \text{, if } R > R_0. \end{cases}$$

This extension of the one dimension uniform distribution (in the sense that, in one dimension, a uniform distribution gives a linear cumulative distribution function over the interval $\mathcal{B}_1(R)$) implies that its isotropic pdf equals $f(\rho) = \rho^{1-p}/R_0$, if $0 < \rho < R_0$ (0, else). A solution in the form $t(\mathbf{r}) = U(\rho)$ must then satisfy $\eta U(\rho) + U''(\rho) = 0$, whose solutions are as follows:

$$t^{(a)}(\rho) \quad = \quad \sqrt{2\text{surf}(\mathcal{S}_p(1))}\cos(\sqrt{\eta}\rho) \quad \text{with } \sqrt{\eta}R_0 = 0\,[\pi], \tag{41}$$

$$t^{(b)}(\rho) \quad = \quad \sqrt{2\text{surf}(\mathcal{S}_p(1))}\sin(\sqrt{\eta}\rho) \quad \text{with } \sqrt{\eta}R_0 = 0\,[2\pi]. \tag{42}$$

surf$(\mathcal{S}_p(1))$ is the surface area of the p-hypersphere of unit radius: surf$(\mathcal{S}_p(1)) = 2\pi^{p/2}/\Gamma(p/2)$. This solution looks like the sphere hardening dither modulation scheme invented by F. Balado [25, Sect. 5].

2) Sparsity: Many possible coordinate systems allow a separation of variables [23], but their investigation is out of the scope of this paper. Preferably, we would like here to rediscover a famous principle in watermarking. Suppose we know a solution $t^\star$ to the scalar equation: $\eta^\star t^\star(x) + f(x)t^{\star\prime}(x) + t^{\star\prime\prime}(x) = 0$. We would like to extend this solution considering a solution in the form: $t = t^\star \circ g$, with $g : \mathbb{R}^n \to \mathbb{R}$ a derivable function. Gradient and laplacian have the following expressions:

$$\nabla t(\mathbf{r}) = t^{\star\prime}(g(\mathbf{r}))\nabla g(\mathbf{r}), \qquad \nabla^2 t(\mathbf{r}) = t^{\star\prime\prime}(g(\mathbf{r}))\|\nabla g(\mathbf{r})\|^2 + t^{\star\prime}(g(\mathbf{r}))\nabla^2 g(\mathbf{r}). \tag{43}$$

and the fundamental equation becomes:

$$t^{\star\prime}(g(\mathbf{r}))\left(-\frac{\eta}{\eta^\star}f(g(\mathbf{r}))+\frac{\nabla p_{\mathbf{S}}(\mathbf{r})^T}{p_{\mathbf{S}}(\mathbf{r})}\nabla g(\mathbf{r})+\nabla^2 g(\mathbf{r})\right)+t^{\star\prime\prime}(g(\mathbf{r}))\left(\|\nabla g(\mathbf{r})\|^2-\frac{\eta}{\eta^\star}\right)=0 \tag{44}$$

A linear form, ie. a projection $g(\mathbf{r})=\mathbf{r}^T\boldsymbol{\lambda}$, is a solution providing the following simplifications: $\nabla^2 g(\mathbf{r})=0$ and $\|\nabla g(\mathbf{r})\|=\|\boldsymbol{\lambda}\|$. Then, t is a fundamental solution with an efficiency per element $\eta=\eta^\star\|\boldsymbol{\lambda}\|^2$, provided we have:

$$\frac{\nabla p_{\mathbf{S}}(\mathbf{r})^T}{p_{\mathbf{S}}(\mathbf{r})}\boldsymbol{\lambda}=\|\boldsymbol{\lambda}\|^2 f(\mathbf{r}^T\boldsymbol{\lambda}). \tag{45}$$

For a white gaussian host, this implies that $f(x)=-x\|\boldsymbol{\lambda}\|^{-2}\sigma_x^{-2}$, which is the score (ie. $p'(x)/p(x)$) associated to $\mathcal{N}(0,\|\boldsymbol{\lambda}\|^2\sigma_x^2)$. Hence, the polynomial family is extended to the vector case with fundamental solutions of the form $t_k(\mathbf{r})=\kappa_k H_k(\mathbf{r}^T\boldsymbol{\lambda}/\|\boldsymbol{\lambda}\|\sigma_x)$ whose efficiency per element is $\eta=k/\sigma_x^2$.

For the flat assumption, f appears to be the null function. Hence, the sinusoidal family is extended to the vector case with fundamental solution of the form $t(\mathbf{r})=k_t\cos(\mathbf{r}^T\boldsymbol{\lambda})$ whose efficacy is $\eta=\|\boldsymbol{\lambda}\|^2$.

This kind of solutions illustrates the principle known as sparsity or time sharing [26, Sect. 5.2 and 8.2], where the watermark embedding is processed on the projection $\mathbf{r}^T\boldsymbol{\lambda}$. A typical implementation of this principle is the Spread Transform Dither Modulation [26, Sect. 5.2].

3) Space partitioning: Under the flat host assumption, (25) reduces to the well known Helmholtz equation: $\eta t(\mathbf{r})+\nabla^2 t(\mathbf{r})=0$. Suppose $t^\star$ is a solution, then the composition of this function by a translation operator yields another solution: $t_0(\mathbf{r})=t^\star(\mathbf{r}-\mathbf{r}_0)$. This property is due to the fact the score $\nabla p_{\mathbf{S}}(\mathbf{r})/p_{\mathbf{S}}(\mathbf{r})$ is invariant by translation since it is null, but this is also the case for a multivariate Laplace distribution. One can also mix different solutions defined over a specific region $\mathcal{C}_i\subset\mathbb{R}^p$: $t(\mathbf{r})=\sum_i t_i(\mathbf{r})\Pi_i(\mathbf{r})$, with $\Pi_i(.)$ the indicator function of region $\mathcal{C}_i$. Assume now, that regions $\{\mathcal{C}_i\}$ constitute a partition of $\mathbb{R}^p$ and that the host pdf is a piecewise constant function such that $p_{\mathbf{s}}(\mathbf{s})=\sum_i P_i\Pi_i(\mathbf{s})$. Then, the above mixture is a solution of the fundamental equation, except on the boundaries of contiguous regions where the gradients of $p_{\mathbf{S}}$ and t are a priori not defined.

An elegant way to set a partition is to define the regions as the Voronoi cells of a n-dimension lattice Λ: $\mathcal{C}_i=\mathcal{V}+\mathbf{c}_i$, $\mathbf{c}_i\in\Lambda$ and $\mathcal{V}$ the Voronoi cell centered on $\mathbf{0}$. With all these elements, we can write:

$$t(\mathbf{r})=\sum_i t_i(\mathbf{r})\Pi_i(\mathbf{r})=\sum_{\mathbf{c}_i\in\Lambda}t^\star(\mathbf{r}-\mathbf{c}_i)\Pi_i(\mathbf{r})=t^\star(\mathbf{r}-Q(\mathbf{r})), \tag{46}$$

with $Q(.)$ the quantization function mapping $\mathbb{R}^p$ onto Λ.

Under the flat host assumption, sparsity and space partitioning indeed give the same extension of the sinusoidal family: $t_{\mathbf{k}}(\mathbf{r})=\sqrt{2}\cos(\mathbf{r}^T\boldsymbol{\lambda}_{\mathbf{k}})$, when vector $\boldsymbol{\lambda}_{\mathbf{k}}$ is defined by $2\pi G^{-T}\mathbf{k}$, with G the generator matrix of lattice Λ and $\mathbf{k}\in\mathbb{N}^p$. $\mathbf{r}$ belonging to $\mathcal{C}_i$, means that $\mathbf{r}=\mathbf{c}_i+\tilde{\mathbf{r}}=G\mathbf{n}_i+\tilde{\mathbf{r}}$, with $\mathbf{n}_i\in\mathbb{Z}^p$ and $\tilde{\mathbf{r}}\in\mathcal{V}$. Thus, $t(\mathbf{r})=t(\tilde{\mathbf{r}})$ because $\mathbf{n}_i^T\mathbf{k}\in\mathbb{Z}$, $\forall(\mathbf{k},\mathbf{n}_i)\in\mathbb{N}^p\times\mathbb{Z}^p$. This gives $\eta=\|\boldsymbol{\lambda}_{\mathbf{k}}\|^2=4\pi^2\|G^{-T}\mathbf{k}\|^2$. Once again, this is not exactly the lattice quantizer based watermarking scheme, but at least we find back solutions which are periodic with respect to a lattice.

To conclude, the goal of this section is to show that several well-known watermarking schemes are indeed solutions of the fundamental equation, underlying the unifying character of this theoretical framework.

V. Attack noise

When there is an attack, the received signal under $\mathcal{H}_1$ is $\mathbf{r}_1 = \mathbf{a}(\mathbf{y})$. The attack channel $\mathbf{a}$ is defined through a conditional probability distribution $p_a(\mathbf{r}_1|\mathbf{y})$, whose associated attack power is $\sigma_a^2 = \int\int \|\mathbf{r}_1 - \mathbf{y}\|^2 p_a(\mathbf{r}_1|\mathbf{y}) p_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} d\mathbf{r}_1/n$. The parameters of the attack channel are unknown at the detection side. We would like to keep the detection as simple as possible so that the estimation of these parameters is not tractable in this strategy. The performance of the detector should degrade slowly with the strength of the attack, according to the definition of robust watermarking given in [27].

The Pitman Noether might then become useless because there is a disruption between the two hypotheses: $\mathcal{H}_1$ doesn't asymptotically converge to $\mathcal{H}_0$, in the sense that the regularity conditions (5) are violated due to the presence of the attack channel only under $\mathcal{H}_1$.

We present here two ways to tackle this problem, changing our framework in order to enforce the Pitman Noether theorem. A first idea is to restrict our analysis to a fixed WNR (watermark to noise power ratio): $\theta_n^2/\sigma_a^2 = g$. The received signal can be written as: $\mathbf{r}_1 = \mathbf{s} + \theta_n \mathbf{w}(\mathbf{s}) + \theta_n \sqrt{g}\tilde{\mathbf{z}}$, with $\mathrm{E}_{\mathbf{Z}}\{\|\tilde{\mathbf{z}}\|^2\} = n$. Therefore, the power of the difference signal $\mathbf{r}_1 - \mathbf{r}_0$ asymptotically vanishes with θ_n^2. The second idea considers attacks with fixed DNR (document -ie. host- to noise power ratio) where signals are corrupted by the same attack under both hypotheses as T. Liu and P. Moulin did [5]. Yet, the targeted applications as described in our introduction do not a priori motivate this possibility because the attack of unprotected content of $\mathcal{H}_0$ are clearly unlikely. We argue that a 'soft' attack on original pieces of content still produces regular content. The attack channel changes the value of the feature vectors, but it does not modify their inherent statistical structure.

Under both attack models, the fundamental equation appears to be statistically robust in the sense that it is not modified by the presence of the attack channel. However, this is only true for very particular conditions as described in the sequel.

A. *Fixed WNR attacks*

This subsection only shows that the fundamental equation remains unchanged when the watermarked signals goes through a fixed WNR AWGN attack channel.

1) Best embedding function for a given detection function: As usual, we write:

$$\begin{aligned}\left.\frac{\partial}{\partial\theta}\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}\right|_{\theta=0} &= \int\int \left.\frac{\partial}{\partial\theta} t(\mathbf{s}+\theta\mathbf{w}(\mathbf{s})+\theta\sqrt{g}\tilde{z})\right|_{\theta=0} p_{\mathbf{S}}(\mathbf{s}) p_{\tilde{\mathbf{Z}}}(\tilde{\mathbf{z}}) d\mathbf{s} d\tilde{\mathbf{z}} \quad (47)\\ &= \int \mathbf{w}(\mathbf{s})^T \nabla t(\mathbf{s}) p_{\mathbf{S}}(\mathbf{s}) d\mathbf{s} + \int\int \sqrt{g}\tilde{\mathbf{z}}^T \nabla t(\mathbf{s}) p_{\mathbf{S}}(\mathbf{s}) p_{\tilde{\mathbf{Z}}}(\tilde{\mathbf{z}}) d\mathbf{s} d\tilde{\mathbf{z}} \quad (48)\end{aligned}$$

We assume $\tilde{\mathbf{z}}$ is independent of $\mathbf{s}$ and centered, so that the second term is null. We find back the same best embedder as (21).

2) Best detection function for a given embedding function: The pdf of $\mathbf{r}_1 = \mathbf{y} + \sqrt{g}\theta\tilde{\mathbf{z}}$ is given by the following convolution:

$$p_{\mathbf{R}_1}(\mathbf{r}) = \int p_{\mathbf{Y}}(\mathbf{u}) p_{\sqrt{g}\theta\tilde{\mathbf{Z}}}(\mathbf{r}-\mathbf{u}) d\mathbf{u}, \quad (49)$$

whose derivative is composed of two terms:

$$\frac{\partial}{\partial\theta}p_{\mathbf{R}_1}(\mathbf{r})\bigg|_{\theta=0} = \int \frac{\partial}{\partial\theta}p_{\mathbf{Y}}(\mathbf{u})\bigg|_{\theta=0}\lim_{\theta\to 0}p_{\sqrt{g}\theta\tilde{\mathbf{Z}}}(\mathbf{r}-\mathbf{u})d\mathbf{u} + \int p_{\mathbf{S}}(\mathbf{u})\,\frac{\partial}{\partial\theta}p_{\sqrt{g}\theta\tilde{\mathbf{Z}}}(\mathbf{r}-\mathbf{u})\bigg|_{\theta=0}d\mathbf{u} \tag{50}$$

We assume that $\tilde{\mathbf{z}}$ is normal distributed. Then, $\lim_{\theta\to 0}p_{\sqrt{g}\theta\tilde{\mathbf{Z}}}(\mathbf{r}-\mathbf{u})$ is the Dirac distribution. Hence, the first term is, as detailed in Sect. III-A, $\partial/\partial\theta p_{\mathbf{Y}}(\mathbf{r})|_{\theta=0} = -\mathrm{div}(p_{\mathbf{S}}(\mathbf{r})\mathbf{w}(\mathbf{r}))$.

The second term is calculated being inspired by some proofs of the De Bruijn's identity (see [28, Th. 16.6.2]). It corresponds to the derivative of the pdf of $\mathbf{a}(\mathbf{s}) = \mathbf{s} + \sqrt{g}\theta\tilde{z}$ with respect to θ. In one hand, we have:

$$\frac{\partial}{\partial\theta}p_{\mathbf{a}(\mathbf{S})}(\mathbf{r}) = \int p_{\mathbf{S}}(\mathbf{u})\left(\frac{\|\mathbf{r}-\mathbf{u}\|^2}{g\theta^3} - \frac{n}{\theta}\right)p_{\sqrt{g}\theta\tilde{\mathbf{Z}}}(\mathbf{r}-\mathbf{u})d\mathbf{u}. \tag{51}$$

On the other hand, it appears that:

$$\nabla^2 p_{\mathbf{a}(\mathbf{S})}(\mathbf{r}) = \int p_{\mathbf{S}}(\mathbf{u})\left(\frac{\|\mathbf{r}-\mathbf{u}\|^2}{g^2\theta^4} - \frac{n}{g\theta^2}\right)p_{\sqrt{g}\theta\tilde{\mathbf{Z}}}(\mathbf{r}-\mathbf{u})d\mathbf{u} = \frac{1}{g\theta}\frac{\partial}{\partial\theta}p_{\mathbf{a}(\mathbf{S})}(\mathbf{r}). \tag{52}$$

Finally, the second term is null, because

$$\frac{\partial}{\partial\theta}p_{\mathbf{a}(\mathbf{S})}(\mathbf{r})\bigg|_{\theta=0} = \lim_{\theta\to 0} g\theta\nabla^2 p_{\mathbf{a}(\mathbf{S})}(\mathbf{r}) = 0, \tag{53}$$

and we find back the same best detection function as (12).

B. Fixed DNR attacks

The framework is changed so that the hypotheses are now: $\mathcal{H}_0 : \mathbf{r}_0 = \mathbf{a}(\mathbf{s})$ against $\mathcal{H}_1 : \mathbf{r}_1 = \mathbf{a}(\mathbf{s} + \theta\mathbf{w}(s))$. What are the impact of this new framework on the detection and embedding functions?

As already said, our analysis only holds for channel attacks conserving the statistical structure of the host signal. The restrictions are as follows. For gaussian host $\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$, the attack is an SAWGN channel: $\mathbf{a}(\mathbf{s}) = \gamma(\mathbf{s}+\mathbf{z})$, with $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma_z^2\mathbf{I}_n)$ independent of $\mathbf{s}$ and $\gamma = 1/\sqrt{1+\sigma_z^2}$. The attack is a Wiener filtering for this very simple case, which maintains $p(\mathbf{r}|\mathcal{H}_0)$ as a normal distribution. For the flat host assumption, the attack is an addition of an independent noise: $\mathbf{a}(\mathbf{s}) = \mathbf{s} + \mathbf{z}$. The new expression of $p(\mathbf{r}|\mathcal{H}_0)$ is given by a convolution, which renders the pdf under $\mathcal{H}_0$ even flatter and larger. Consequently, at the scale of the watermarking signal, $p(\mathbf{r}|\mathcal{H}_0)$ is still a piecewise constant function. The expression (11) of the best detection function given the embedding function is not modified when restricting to attack channels preserving $p(\mathbf{r}|\mathcal{H}_0)$.

This is not the case for the best embedding function given the detection function. For the class of attack channel considered in this paper, we can write $\mathbf{a}(\mathbf{s}) = \gamma(\mathbf{s}+\mathbf{z})$ with $\gamma = 1$ for the additive noise attack, and $\gamma = 1/\sqrt{1+\sigma_z^2}$ for the SAWGN attack. (18) is then modified as follows:

$$\begin{aligned}\frac{\partial}{\partial\theta}\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}\bigg|_{\theta=0}(\gamma,\sigma_z) &= \int\!\!\int \frac{\partial}{\partial\theta}t(\gamma(\mathbf{s}+\theta\mathbf{w}(\mathbf{s})+\mathbf{z}))\bigg|_{\theta=0}p_{\mathbf{Z}}(\mathbf{z})p_{\mathbf{S}}(\mathbf{s})d\mathbf{z}d\mathbf{s} &(54)\\ &= \int \gamma\mathbf{w}(\mathbf{s})^T\left(\int \nabla t(\gamma(\mathbf{s}+\mathbf{z}))p_{\mathbf{Z}}(\mathbf{z})d\mathbf{z}\right)p_{\mathbf{S}}(\mathbf{s})d\mathbf{s}. &(55)\end{aligned}$$

This last equation shows that the best strategy at the embedding side should set

$$\mathbf{w}(\mathbf{s}) \propto \mathrm{E}_Z\{\nabla t(\gamma(\mathbf{s}+\mathbf{z}))\}. \tag{56}$$

This implies that the embedder knows the attack channel parameters. This counter attack may not be realistic in general, and we keep our former strategy given by (21), so that

$$\eta(\gamma,\sigma_z) = \frac{\gamma^2\eta(1,0)}{n^2}(\mathrm{E}_{\mathbf{S}}\{\mathbf{w}(\mathbf{s})^T\mathrm{E}_{\mathbf{Z}}\{\mathbf{w}(\gamma(\mathbf{s}+\mathbf{z}))\}\})^2. \tag{57}$$

However, there are some cases where the counter attack (56) is surprisingly simple because it is indeed identical to the regular embedding strategy (21) whatever the parameters of the attack channel. This occurs when t is such that $\mathrm{E}_Z\{\nabla t(\gamma(\mathbf{s}+\mathbf{z}))\} = h(\gamma,\sigma_z)\nabla t(\mathbf{s})$. As a consequence, the fundamental equation (25) derived in the no attack case, remains valid under these particular attack cases. The efficiency per element is then equal to $\eta(\gamma,\sigma_z) = \gamma^2 h^2(\gamma,\sigma_z)\eta(1,0)$.

For the polynomial family, we rewrite the Wiener filtering denoting $\tilde{z} = \sigma_z^{-1}z$ distributed as $\mathcal{N}(0,1)$ and $\alpha = \arccos(\gamma)$. A less familiar identity of the Hermite polynomials allows to write:

$$t'_\ell(\gamma(s+z)) = \kappa_\ell \ell H_{\ell-1}(\cos(\alpha)s + \sin(\alpha)\tilde{z}) = \kappa_\ell \ell \sum_{k=0}^{\ell-1}\tbinom{\ell-1}{k}\cos^k(\alpha)\sin^{\ell-1-k}(\alpha)H_k(s)H_{\ell-1-k}(\tilde{z}) \tag{58}$$

$\mathrm{E}_Z\{t'_\ell(\gamma(s+z))\}$ reduces to $\mathrm{E}_{\tilde{Z}}\{t'_\ell(\gamma s + \sigma_z\gamma\tilde{z}))\} = \kappa_\ell\ell\gamma^{\ell-1}H_{\ell-1}(s) = \gamma^{\ell-1}t'_\ell(s)$ because $\mathrm{E}_{\tilde{Z}}\{H_k(\tilde{z})\} = \delta(k)$. Consequently, the polynomial family is a set of fundamental solutions for i.i.d. gaussian hosts and SAWGN attacks with Wiener filtering, whose efficiency per element is given by $\eta(\gamma,\sigma_z) = \ell\gamma^{2\ell}$. Wiener filtering means that $\gamma = (1+\sigma_z^2)^{-1/2}$. Two noticeable exemptions are t_1 and t_2, whose efficiency follows the same rule whatever the value of γ in the SAWGN channel. Last but not least: the higher the 'original' efficiency $\eta(1,0) = \ell$, the less robust is the scheme in the sense that $\eta(\gamma,\sigma_z)/\eta(1,0) = (1+\sigma_Z^2)^{-\eta(1,0)}$ decreases faster with the strength of the attack.

For the sinusoidal family, an additive noise leads to

$$\mathrm{E}_Z\{t'_\ell(s+z)\} = t'_\ell(s)\mathrm{E}_Z\{\cos(\ell\sqrt{\eta}z)\} - \ell\sqrt{2\eta}\cos(\ell\sqrt{\eta}s)\mathrm{E}_Z\{\sin(\ell\sqrt{\eta}z)\}. \tag{59}$$

The desired property is enable whenever the attack noise has an even pdf which sets the second term to zero. For instance, the AWGN channel gives $\mathrm{E}_Z\{t'_\ell(s+z)\} = t'_\ell(s)e^{-\ell\sqrt{\eta}\sigma_z^2/2}$. Consequently, the sinusoidal family is a set of fundamental solutions for flat hosts and additive symmetric noise attacks. For the AWGN channel attack, its efficiency is given by $\eta(1,\sigma_z) = \ell\sqrt{\eta}e^{-\ell\sqrt{\eta}\sigma_z^2}$. Once again, the higher the 'original' efficiency $\eta(1,0)$, the less robust is the scheme in the sense that $\eta(\gamma,\sigma_z)/\eta(1,0) = e^{-\eta(1,0)\sigma_z^2}$ decreases faster with the strength of the attack.

The same analysis also holds for the extension of the polynomial and sinusoidal family to the vector case. For instance, JANIS is a solution of the fundamental equation for i.i.d. gaussian hosts and SAWGN attack, such that $\mathrm{E}_Z\{\nabla t(\gamma(\mathbf{s}+\mathbf{z}))\} = \gamma^{p-1}\nabla t(\mathbf{s})$. The Wiener filtering restriction is not necessary as JANIS is based on first order Hermite polynomials. This gives the following efficiency per element $\eta(\gamma,\sigma_z) = p\gamma^{2p}$ which follows the same decreasing rule as the scalar polynomial family. The extended sinusoidal family follows the same rule: $\eta(1,\sigma_z)/\eta(1,0) = e^{-\eta(1,0)\sigma_z^2}$ with $\eta(1,0) = 4\pi^2\|G^{-T}\mathbf{k}\|^2$ as shown in Appendix III.

VI. Limitations and extensions

A. About the Pitman Noether theorem

The key element of this paper is the maximization of the efficiency per element, which plays a fundamental role in the Pitman Noether theorem. However, this theorem only holds if the tested statistic asymptotically converges in law to a normal random variable.

In Sect. IV-A, we propose to write the detection function as a sum of scalar functions: $t(\mathbf{r}) = \sum_{i=1}^{n} t(r_i)$. In our model, the $\{t(r_i)\}$ variables are independent and identically distributed with a bounded variance under both hypotheses. The central limit theorem can be applied to prove the asymptotic normality. However, the variance under $\mathcal{H}_1$ grows very fast with the efficacy, as shown in Appendix II and in Table IV-A.1. Consequently, the efficacy cannot tend to $+\infty$.

In Sect. IV-B, we propose the idea of decomposing the received signal in n/p blocks of p samples (p dividing n), so that the detection function is the sum of n/p i.i.d. random variables. The parameter p must be fixed to ensure the asymptotic normality. In all the block-based watermarking schemes rediscovered in this paper, (40) is the only one where $p = n$ still gives the asymptotic normality.

In any case, the central limit theorem shall not be used to estimate neither the threshold nor the power of the test in practice. The Berry-Esseen theorem shows that the rate of convergence to the normal distribution depends on the third moment of t, which we suspect to be fast increasing with the efficacy.

In our opinion, the usefulness of this framework is limited in practice. It nicely gives a unified theory gathering many known watermarking schemes. However, new fundamental solutions may not be adequate for practical implementations when host signal are not so long, or θ is not so small. A proper study requires a non asymptotic analysis which is out of the scope of this article.

B. Asymmetric tests

So far, the main idea of the paper is to take advantage of the knowledge of the host value $\mathbf{s}$ to boost the efficiency per element. This results in the increase of $\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\} = \theta\sqrt{n\eta} + O(\theta^2)$, while the variance $\mathrm{Var}\{t(\mathbf{r})|\mathcal{H}_1\}$ is maintained at the level of $\mathrm{Var}\{t(\mathbf{r})|\mathcal{H}_0\}$ at least to the first order. Asymptotically, the test has to make a clear cut between two gaussian distributions having the same variance. This is sometimes called a symmetric test. Let us focus more on the variance $\mathrm{Var}(t(\mathbf{r})|\mathcal{H}_1)$. As H. Malvar and D. Florencio did for zero-rate watermarking [29], we would like to control the value of $\mathrm{Var}(t(\mathbf{r})|\mathcal{H}_1)$, achieving so-called asymmetric tests[4].

The watermark signal is already dependent to the host through the vector $\mathbf{w}(\mathbf{s})$ which pushes the host towards a region in space where the detection function has a higher value, ie. hopefully the acceptance region. We add here another dependence which modulates the amplitude of this vector: host signals which are naturally far away from the acceptance region are more strongly pushed than those near the acceptance region. We write the watermark signal

[4] Be careful not to confuse with asymmetric watermarking where the detection key is different from the embedding private key.

$\mathbf{x}(\mathbf{s}) = \theta k_w(\mathbf{s})\mathbf{w}(\mathbf{s})$. For a fair comparison with the previous sections, the constraint reads: $\mathrm{E}_{\mathbf{S}}\{k_w(\mathbf{s})^2\|\mathbf{w}(\mathbf{s})\|^2\} = n$. The embedding strategy is not changed: $\mathbf{w}(\mathbf{s}) = \nabla t(\mathbf{s})$. Hence, we have:

$$\begin{aligned} n &= \mathrm{E}_{\mathbf{S}}\{k_w(\mathbf{s})^2\|\nabla t(\mathbf{s})\|^2\} & (60) \\ \frac{\partial}{\partial\theta}\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}\Big|_{\theta=0} &= \mathrm{E}_{\mathbf{S}}\{k_w(\mathbf{s})\|\nabla t(\mathbf{s})\|^2\} & (61) \\ \eta &= \frac{\mathrm{E}_{\mathbf{S}}\{k_w(\mathbf{s})\|\nabla t(\mathbf{s})\|^2\}^2}{\mathrm{E}_{\mathbf{S}}\{k_w(\mathbf{s})^2\|\nabla t(\mathbf{s})\|^2\}} & (62) \end{aligned}$$

Now, the goal is to choose function k_w such that it reduces the variance under $\mathcal{H}_1$:

$$\frac{\partial}{\partial\theta}\mathrm{Var}\{t(\mathbf{r})|\mathcal{H}_1\}\Big|_{\theta=0} = 2\mathrm{E}_{\mathbf{S}}\{t(\mathbf{s})\tilde{\nu}(\mathbf{s})\} \geq -2\mathrm{Var}\{\nu(\mathbf{s})\}, \quad (63)$$

where $\nu(\mathbf{s}) = k_w(\mathbf{s})\|\nabla t(\mathbf{s})\|^2$ such that its centered version is $\tilde{\nu}(s) = \nu(\mathbf{s}) - \frac{\partial}{\partial\theta}\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}\big|_{\theta=0}$. The Cauchy-Schwarz inequality gives $-2\mathrm{Var}\{\nu(\mathbf{s})\}$ as the lower bound, with equality when $\tilde{\nu}(\mathbf{s}) = -ct(\mathbf{s})$, c a positive constant. Hence, we achieve to reduce $\mathrm{Var}(t(\mathbf{r})|\mathcal{H}_1)$. However, this strategy consumes embedding distortion:

$$\begin{aligned} n &= \mathrm{E}_{\mathbf{S}}\{k_w(\mathbf{s})^2\|\nabla t(\mathbf{s})\|^2\} = \mathrm{E}_{\mathbf{S}}\{\nu(\mathbf{s})^2\|\nabla t(\mathbf{s})\|^{-2}\} \\ &= c^2\mathrm{E}_{\mathbf{S}}\{t(\mathbf{s})^2\|\nabla t(\mathbf{s})\|^{-2}\} + n\eta\mathrm{E}_{\mathbf{S}}\{\|\nabla t(\mathbf{s})\|^{-2}\} - 2c\sqrt{n\eta}\mathrm{E}_{\mathbf{S}}\{t(\mathbf{s})\|\nabla t(\mathbf{s})\|^{-2}\}. \quad (64) \end{aligned}$$

For the simple cases explored in this paper, we are able to find a bijection $\mathbf{s}' = h(\mathbf{s})$ such that $p_{\mathbf{S}}(\mathbf{s}')t(\mathbf{s}')\|\nabla t(\mathbf{s}')\|^{-2} = -p_{\mathbf{S}}(\mathbf{s})t(\mathbf{s})\|\nabla t(\mathbf{s})\|^{-2}$, which implies a third null term. Denote $a = \mathrm{E}_{\mathbf{S}}\{t(\mathbf{s})^2\|\nabla t(\mathbf{s})\|^{-2}\}$ and $b = \mathrm{E}_{\mathbf{S}}\{\|\nabla t(\mathbf{s})\|^{-2}\}$. (64) finally reads:

$$n = ac^2 + bn\eta. \quad (65)$$

A higher c decreases $\mathrm{Var}\{t(\mathbf{r})|\mathcal{H}_1\}$ (first order approximation) but also η due to the distortion constraint. In practice, this strategy brings a crucial issue. Assuming the tested statistic has a gaussian distribution under both hypotheses, a decrease of $\mathrm{Var}\{t(\mathbf{r})|\mathcal{H}_1\}$ yields a higher power of test only if $\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\}$ is greater than threshold $\tau > 0$. Now, if this is not the case (for instance, due to an attack), then the impact of this strategy is just the opposite. This phenomenon is not relevant in [29], as it tackles watermark decoding where threshold τ equals 0, the distributions under $\mathcal{H}_0$ (bit 1 has been hidden) and $\mathcal{H}_1$ (bit 0 has been hidden) being symmetric around this value.

VII. About DC-DM watermarking based on lattice quantization

Our theoretical framework doesn't succeed to find back well known DC-DM watermarking schemes based on lattice quantization. The detection function is usually defined as $t(\mathbf{r}) = k_t\|Q(\mathbf{r}) - \mathbf{r}\|^2$, and the embedding function $\mathbf{x}(\mathbf{s}) = \alpha(Q(\mathbf{s}) - \mathbf{s})$ complies with rule (21). The parameter α plays a crucial role in the trade-off between the embedding distortion and the inherent robustness of the scheme. Note that our point of view is very different as we suppose that the host signal is pushed in a direction given by $\mathbf{w}(\mathbf{s}) = k_t\nabla t(\mathbf{s}) = 2k_t(Q(\mathbf{s}) - \mathbf{s})$, but the watermark signal $\mathbf{x}(\mathbf{s}) = \theta\mathbf{w}(\mathbf{s})$ is not deterministic because the amplitude θ is not fixed.

A. *Efficacy without noise*

We consider a lattice Λ and a host whose pdf is a piecewise constant function over the partition induced by Λ: $\mathbb{R}^p = \bigcup_{\mathbf{c}_i \in \Lambda} (\mathcal{V} + \mathbf{c}_i)$. We study the detection function given by $t(\mathbf{r}) = k_t(\|Q(\mathbf{r}) - \mathbf{r}\|^2 - \mu)$, with Q the quantizer associated to Λ, and $\{k_t, \mu\}$ enforcing a centered unit variance tested statistic under $\mathcal{H}_0$:

$$\mu \;=\; \mathrm{vol}(\mathcal{V})^{-1} \int_{\mathcal{V}} \|\mathbf{r}\|^2 d\mathbf{r} = I(\Lambda, 2), \tag{66}$$

$$k_t \;=\; -\left(\mathrm{vol}(\mathcal{V})^{-1} \int_{\mathcal{V}} \|\mathbf{r}\|^4 d\mathbf{r} - \mu^2 \right)^{-\frac{1}{2}} = -(I(\Lambda, 4) - I(\Lambda, 2)^2)^{-\frac{1}{2}}. \tag{67}$$

$I(\Lambda, k)$ denotes the k-th normalized moment of $\mathcal{V}$, ie. $\mathrm{vol}(\mathcal{V})^{-1} \int_{\mathcal{V}} \|\mathbf{r}\|^k d\mathbf{r}$. The embedding function is $\mathbf{w}(\mathbf{s}) = 2k_w k_t(Q(\mathbf{r}) - \mathbf{r})$, ie. a vector pointing towards the nearest element of the lattice. Constant k_w is given by:

$$k_w = \frac{\sqrt{n}}{2k_t \sqrt{I(\Lambda, 2)}}. \tag{68}$$

Finally, (23) gives the following efficiency per element for the noiseless case:

$$\eta = \frac{4I(\Lambda, 2)}{I(\Lambda, 4) - I(\Lambda, 2)^2}. \tag{69}$$

For a positive scale factor $\beta < 1$ giving a finer partition induced by $\beta\Lambda$, we have a higher efficiency $\eta_{\beta\Lambda} = \beta^{-2}\eta_{\Lambda}$. Therefore, efficiencies should be compared for partitions with $\mathrm{vol}(\mathcal{V}) = 1$. Anyway, finding the optimal lattice giving the best efficiciency is out of the scope of this paper. As an example, for cubic lattice $\Lambda = \mathbb{Z}^p$, $\mathcal{V}$ is the centered hypercube $[-1/2, 1/2)^p$ and $\eta = 60$. For the two dimension hexagonal lattice A_2, whose associated generating matrix is $G = [2\,1; 0\,\sqrt{3}]/\sqrt{2\sqrt{3}}$ such that $\mathrm{vol}(\mathcal{V}) = 1$, we achieve a higher efficacy per element $\eta = 1800\sqrt{3}/43 \approx 72, 50$. Compared to the square lattice $\mathbb{Z}^2$, the 'more spherical' of the two lattices is the best, when no attack is considered. This is surprisingly different from the zero-rate case presented in [20, Sect. 3.3].

Increasing the integer p, there exist lattices with nearly spherical Voronoi cell. Assuming $\mathcal{V} = \mathcal{B}_p(R)$, the efficacy reads $\eta = (p + 4)(p + 2)R^{-2}$. Setting $R = \Gamma(p/2 + 1)^{1/p}/\sqrt{\pi}$ such that $\mathrm{vol}(\mathcal{V}) = 1$, and using Stirling's approximation, we achieve a linear efficiency per element: $\eta \approx 2\pi e p$. In view of Sect.VI-A, this issue is now whether we can increase parameter p, which is the size of the blocks. The tested statistic reads in term of the square norm of a quantization noise of a flat host, which is not asymptotically gaussian. Once again, we are facing the limitations of the Pitman Noether theorem: the block based watermarking must be done with a fixed p.

B. *Mixture of fundamental solutions*

We investigate the idea of mixing different fundamental solutions among a family. Suppose a family of orthonormal fundamental solutions $\{t_j\}$ with integer indices, and create the following detection function $t(\mathbf{r}) = \sum_{j=1}^{\Omega} \omega_j t_j(\mathbf{r})$. We have:

$$\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_0\} = \sum_{j=1}^{\Omega} \omega_j \mathrm{E}_{\mathbf{R}}\{t_j(\mathbf{r})|\mathcal{H}_0\} = 0, \qquad \mathrm{Var}\{t(\mathbf{r})|\mathcal{H}_0\} = \sum_{j=1}^{\Omega} \omega_j^2 = 1. \tag{70}$$

The last equation gives a constraint on the weights $\{\omega_j\}$.

The reader must be aware of two facts. First, we have chosen here to mix some detection functions, but we could also do the mixture on the embedding functions. Second, this mixture is a priori not a fundamental solution. Given this mixture, we select the best embedding function $\mathbf{w}(\mathbf{s}) = k_w \nabla t(\mathbf{s})$. However, it is a priori not true that the mixture is the best detection function knowing $\mathbf{w}(\mathbf{s})$. The mixture of detection functions implies a mixture of the associated embedding functions, $\mathbf{w}(\mathbf{s}) = \sum_{j=1}^{\Omega} \varpi_j \mathbf{w}_j(\mathbf{s})$, but with different weights:

$$\varpi_j = k_w \omega_j \sqrt{\eta_j(1,0)} \quad \text{and} \quad k_w = (\sum_{j=1}^{\Omega} \omega_j^2 \eta_j(1,0))^{-1/2}$$

(23) gives the efficacy when there is no attack:

$$\eta(1,0) = \sum_{j=1}^{\Omega} \omega_j^2 \eta_j(1,0). \tag{71}$$

(57) gives the following efficiency per element under attack,

$$\eta(\gamma,\sigma_z) = \left(\sum_{j=1}^{\Omega} \omega_j \varpi_j \sqrt{\eta_j(\gamma,\sigma_z)} \right)^2 = \frac{\left(\sum_{j=1}^{\Omega} \omega_j^2 \sqrt{\eta_j(1,0)\eta_j(\gamma,\sigma_z)} \right)^2}{\sum_{j=1}^{\Omega} \omega_j^2 \eta_j(1,0)}, \tag{72}$$

if we suppose that $\mathrm{E}_{\mathbf{S}}\{\mathbf{w}_j(\mathbf{s})\mathrm{E}_{\mathbf{Z}}\{\mathbf{w}_k(\gamma(\mathbf{s}+\mathbf{z}))\}\} = \delta(j-k)n/\gamma.\sqrt{\eta_j(\gamma,\sigma_z)/\eta_j(1,0)}$, ie. the functions stay orthogonal even under attack. This assumption considerably simplifies the expression of the efficiency. From Sect. V-B, we know this holds for the polynomial family ($\gamma = 1/\sqrt{1+\sigma_z^2}$), and for the sinusoidal family ($\gamma = 1$), because $\mathrm{E}_{\mathbf{Z}}\{\mathbf{w}_k(\gamma(\mathbf{s}+\mathbf{z}))\} \propto \mathbf{w}_k(\mathbf{s})$.

Mixture is a tool which renders the study of some watermarking schemes easier. For instance, the sinusoidal family weighted with $\omega_j = -(-1)^j 3\sqrt{10}/\pi^2/j^2$ gives the Fourier series decomposition of the SCS scheme:

$$t(s) = -\frac{6\sqrt{5}}{\pi^2} \sum_{j=1}^{\infty} \frac{(-1)^j}{j^2} \cos(j\sqrt{\eta}s) = \frac{\sqrt{5}}{2} - (s - Q(s))^2 \frac{6\sqrt{5}}{\Delta^2} \tag{73}$$

$$w(s) = k_w \frac{6\sqrt{5\eta}}{\pi^2} \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sin(j\sqrt{\eta}s) = -(s - Q(s)) \frac{\sqrt{12}}{\Delta} \tag{74}$$

with Q a quantizer whose step is $\Delta = 2\pi/\sqrt{\eta}$. The application of (72) gives the efficiency of SCS which is otherwise cumbersome to calculate with the direct expressions of t and w. Here, we simply have:

$$\eta_{SCS}(1,\sigma_z) = \frac{90\eta}{\pi^4}.\frac{(\sum_{j=1}^{\infty} j^{-2} e^{-j^2 \frac{\eta\sigma_z^2}{2}})^2}{\sum_{j=1}^{\infty} j^{-2}} = \frac{60}{\Delta^2} \left(1 + \frac{6\sigma_Z^2}{\Delta^2} - \frac{3}{\pi} \int_0^{\frac{2\pi\sigma_z^2}{\Delta^2}} \vartheta_3(0, e^{-\pi u}) du \right)^2, \tag{75}$$

where ϑ_3 is the third Jacobi theta function. Fig. (1) shows the efficacy of SCS with σ_z ranging from 0 to 1 for $\eta = 1$. When there is no attack, $\eta_{SCS}(1,0) = 60/\Delta^2 = 15\eta/\pi^2 \approx 1,52\eta$. A sinusoidal function starting from the same value is more robust in this range of noise as plotted in Fig. (1). This demonstrates the superiority of the sinusoidal family.

In the same way, the detection function based on lattice quantizer of Sect. VII-A can be decomposed through a Fourier series over lattice Λ, whose generator matrix is G:

$$t(\mathbf{r}) = I(\Lambda, 2) + \sqrt{2} \sum_{\mathbf{k} \in \mathbb{N}^p} \omega_{\mathbf{k}} \cos(2\pi \mathbf{r}^T G^{-T} \mathbf{k}), \tag{76}$$

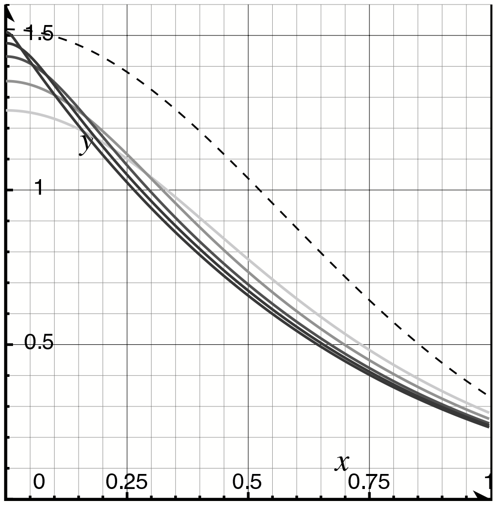

Fig. 1. Efficiency per element of the SCS scheme under AWGN attack against σ_z. The grey plots are the approximations by (75) for $j_{max} = \{3, 5, 10, 20, 100\}$. The dotted line is the efficiency of the sinusoidal solution with $\eta(1, 0) = 1, 52$.

with $\omega_{\mathbf{k}} = \sqrt{2}\text{vol}(\mathcal{V})^{-1} \int_{\mathcal{V}} \|\mathbf{r}\|^2 \cos(2\pi \mathbf{r}^T G^{-T} \mathbf{k}) d\mathbf{r}$. This decomposition in Fourier series may not be easy to obtain except for low dimension lattice.

VIII. Conclusion

Rewriting classical elements of detection theory with the assumption that the watermark signal depends on the host gives us the expression of the best embedding function knowing the detector. Coupling this result with the expression of the LMP test gives a partial differential equation we named ‘fundamental equation’ of zero-bit watermarking. Its main advantage is to offer a constructive theoretical framework unifying most of the watermarking schemes the community knows. Moreover, a side product is that the decomposition onto a family of orthogonal fundamental solutions provide an easier way to characterize the performance of DC-DM schemes.

IX. Acknowledgments

I would like to thank Pedro Comesana Alfaro for his numerous corrections and suggestions of improvement, and also Sandrine Le Squin and Julie Josse for having compared the results with their numerical simulations.

Appendix I
LMP test

For a given embedding function $\mathbf{w}$, we derive the Locally Most Powerful test, whose detection function is defined as:

$$t(\mathbf{r}) = \frac{k_t}{p_{\mathbf{S}}(\mathbf{r})} \left. \frac{\partial p(\mathbf{r}|\mathcal{H}_1)}{\partial \theta} \right|_{\theta=0}. \tag{77}$$

$\theta \to 0$ makes function $\mathbf{f}$ invertible: $\mathbf{s} = \mathbf{f}^{-1}(\mathbf{y})$, and $p(\mathbf{r}|\mathcal{H}_1) = p_{\mathbf{S}}(\mathbf{f}^{-1}(\mathbf{r}))|J_{\mathbf{f}^{-1}}(\mathbf{r},\theta)|$, with the last term being the Jacobian of $\mathbf{f}^{-1}$. Finally, the detection function is:

$$\begin{aligned} t(\mathbf{r}) &= \frac{k_t}{p_{\mathbf{S}}(\mathbf{r})}\left(\nabla p_{\mathbf{S}}(\mathbf{f}^{-1}(\mathbf{r}))^T \frac{\partial \mathbf{f}^{-1}}{\partial \theta}(\mathbf{r})|J_{\mathbf{f}^{-1}}(\mathbf{r},\theta)| + p_{\mathbf{S}}(\mathbf{f}^{-1}(\mathbf{r}))\frac{\partial |J_{\mathbf{f}^{-1}}(\mathbf{r},\theta)|}{\partial \theta}\right)_{\theta=0} &(78)\\ &= \frac{k_t}{p_{\mathbf{S}}(\mathbf{r})}(A(\mathbf{r}) + B(\mathbf{r})). &(79)\end{aligned}$$

Some simple equations are:

$$\begin{aligned} \mathbf{f}(\mathbf{s})|_{\theta=0} &= \mathbf{s}, &(80)\\ \mathbf{f}^{-1}(\mathbf{y})\big|_{\theta=0} &= \mathbf{y}, &(81)\\ \mathbf{f}^{-1}(\mathbf{y}) &= \mathbf{y} - \theta \mathbf{w}(\mathbf{f}^{-1}(\mathbf{y})). &(82)\end{aligned}$$

A. *Expression of* $A(\mathbf{r})$

Deriving this last expression gives:

$$\frac{\partial \mathbf{f}^{-1}}{\partial \theta}(\mathbf{y}) = -\mathbf{w}(\mathbf{f}^{-1}(\mathbf{y})) - \theta J_{\mathbf{w}}(\mathbf{f}^{-1}(\mathbf{y}))\frac{\partial \mathbf{f}^{-1}}{\partial \theta}(\mathbf{y}). \tag{83}$$

Hence,

$$\left.\frac{\partial \mathbf{f}^{-1}}{\partial \theta}(\mathbf{y})\right|_{\theta=0} = -\mathbf{w}(\mathbf{y}). \tag{84}$$

The elements of the Jacobian matrix are given by:

$$[J_{\mathbf{f}^{-1}}(\mathbf{y},\theta)](i,j) = \frac{\partial f_i^{-1}}{\partial y_j} = \delta(i-j) - \theta \nabla w_i(\mathbf{f}^{-1}(\mathbf{y}))^T J_{\mathbf{f}^{-1}}(\mathbf{y})\mathbf{e}_j. \tag{85}$$

The simplification taking $\theta = 0$ yields $|J_{\mathbf{f}^{-1}}(\mathbf{y},0)| = 1$, and the expression of A is as follows:

$$A(\mathbf{r}) = -\nabla p_{\mathbf{S}}(\mathbf{r})^T \mathbf{w}(\mathbf{r}). \tag{86}$$

B. *Expression of* $B(\mathbf{r})$

This term implies the derivative of the determinant of matrix $J_{\mathbf{f}^{-1}}(\mathbf{r},\theta)$ which is invertible as $\theta \to 0$:

$$\frac{\partial |J_{\mathbf{f}^{-1}}|}{\partial \theta}(\mathbf{r},\theta) = |J_{\mathbf{f}^{-1}}(\mathbf{r},\theta)|\text{tr}(J_{\mathbf{f}^{-1}}(\mathbf{r},\theta)^{-1}\frac{\partial J_{\mathbf{f}^{-1}}}{\partial \theta}(\mathbf{r},\theta)) \tag{87}$$

Taking $\theta = 0$ gives:

$$\frac{\partial |J_{\mathbf{f}^{-1}}|}{\partial \theta}(\mathbf{r},0) = \text{tr}\left(\frac{\partial J_{\mathbf{f}^{-1}}}{\partial \theta}(\mathbf{r},0)\right). \tag{88}$$

The derivative of (85) gives the elements of matrix $\frac{\partial J_{\mathbf{f}^{-1}}}{\partial \theta}(\mathbf{r},\theta)$:

$$\frac{\partial^2 f_i^{-1}}{\partial \theta \partial y_j}(\mathbf{r},\theta) = -\nabla w_i(\mathbf{f}^{-1}(\mathbf{r}))^T J_{\mathbf{f}^{-1}}(\mathbf{r},\theta)\mathbf{e}_j - \theta\frac{\partial}{\partial \theta}(\nabla w_i(\mathbf{f}^{-1}(\mathbf{r}))^T J_{\mathbf{f}^{-1}}(\mathbf{r},\theta)\mathbf{e}_j). \tag{89}$$

So that, these elements are equal to $-\nabla \mathbf{w}_i(\mathbf{r})^T \mathbf{e}_j$ when $\theta = 0$, and, finally, $B(\mathbf{r}) = -p_{\mathbf{S}}(\mathbf{r})\text{div}(\mathbf{w}(\mathbf{r}))$.

APPENDIX II

MACLAURIN SERIES OF $\mathrm{VAR}\{t(r)|\mathcal{H}_1\}$ WITHOUT ATTACK.

We make the Maclaurin series of $t(s+\theta w(s))^2$, and take the expectation:

$$\mathrm{E}_S\{t(s+\theta w(s))^2\} = 1 + 2\theta \mathrm{E}_S\{w(s)t'(s)t(s)\} + \theta^2 \mathrm{E}_S\{w(s)^2(t'(s)^2 + t(s)t''(s))\} + O(\theta^3). \qquad (90)$$

If t is an odd function, then t' and $w = k_w t'$ are even functions. The second term of the series is null. If t is an even function, the second term is not null as shown in Table IV-A.1.

A. First order term for even polynomial function

An even polynomial detection function means $t(s) = \kappa_k H_k(s)$, with k even and $\kappa_k = k!^{-1/2}$ (probabilists' definition). Then, $t'(s) = \kappa_k k H_{k-1}(s)$ and $w(s) = k_w \kappa_k k H_{k-1}(s) = \kappa_{k-1} H_{k-1}(s)$. Therefore, $\mathrm{E}_S\{w(s)t'(s)t(s)\} = \kappa_k^2 \kappa_{k-1} k \mathrm{E}_S\{H_k(s) H_{k-1}(s)^2\}$. A known formula of the square of Hermite polynomials is the following one:

$$H_{k-1}(s)^2 = \sum_{\ell=0}^{k-1} \binom{k-1}{\ell}^2 \ell! H_{2k-2-2\ell}(s) \qquad (91)$$

The orthogonality of the Hermite polynomial family allows us to conclude that:

$$\mathrm{E}_S\{w(s)t'(s)t(s)\} = \kappa_k^2 \kappa_{k-1} k \binom{k-1}{k/2-1}^2 (k/2-1)!k! = \frac{\sqrt{(k-1)!k!}}{(k/2-1)!(k/2!)^2}. \qquad (92)$$

The application of the Stirling approximation, when k is large, gives $\mathrm{E}_S\{w(s)t'(s)t(s)\} \approx \sqrt{2/e}(2\pi)^{-3/4} 2^{3k/2} k^{-1/4}$. The derivation of the second order term is tackled in the following section.

B. Second order term

In a similar way, we have:

$$w(s)^2 t'^2(s) = \frac{k}{(k-1)!^2} \left(\sum_{\ell=0}^{k-1} \binom{k-1}{\ell}^2 \ell! H_{2k-2-2\ell}(s) \right)^2, \qquad (93)$$

whose expectation, thanks to the orthogonality feature, simplifies to:

$$\mathrm{E}_S\{w(s)^2 t'^2(s)\} = \frac{k}{(k-1)!^2} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell}^4 \ell!^2 (2k-2-2\ell)! \qquad (94)$$

The second term is slightly different:

$$\begin{aligned} w(s)^2 t''(s) t(s) &= \frac{k(k-1)}{k!(k-1)!} H_{k-1}(s)^2 H_k(s) H_{k-2}(s), & (95)\\ &= \frac{k(k-1)}{k!(k-1)!} \left(\sum_{\ell=0}^{k-1} \binom{k-1}{\ell}^2 \ell! H_{2k-2-2\ell}(s) \right) \left(\sum_{\ell=0}^{k-2} \binom{k}{\ell}\binom{k-2}{\ell} \ell! H_{2k-2-2\ell}(s) \right), & (96)\\ &= \frac{k^2}{k!(k-1)!} \left(\sum_{\ell=0}^{k-1} \binom{k-1}{\ell}^2 \ell! H_{2k-2-2\ell}(s) \right) \left(\sum_{\ell=0}^{k-2} \frac{k-\ell-1}{k-\ell} \binom{k-1}{\ell}^2 \ell! H_{2k-2-2\ell}(s) \right), & (97) \end{aligned}$$

whose expectation is

$$\mathrm{E}_S\{w(s)^2 t''(s) t(s)\} = \frac{k}{(k-1)!^2} \sum_{\ell=0}^{k-2} \left(1 - \frac{1}{k-\ell}\right) \binom{k-1}{\ell}^4 \ell!^2 (2k-2-2\ell)!. \qquad (98)$$

C. Final expression

Withdrawing the square of $\mathrm{E}_{\mathbf{R}}\{t(\mathbf{r})|\mathcal{H}_1\} = \sqrt{k}\theta + O(\theta^2)$, we get:

$$\mathrm{Var} \quad \{t(\mathbf{r})|\mathcal{H}_1\} = \tag{99}$$

$$1 \quad +2\mathrm{mod}(k+1,2)\theta\frac{\sqrt{(k-1)!}k!}{(k/2-1)!(k/2)!^2} + \theta^2\frac{k}{(k-1)!^2}\sum_{\ell=0}^{k-2}\left(2-\frac{1}{k-\ell}\right)\tbinom{k-1}{\ell}^4\ell!^2(2k-2-2\ell)! + O(\theta^3) \tag{100}$$

Appendix III

Efficacy of the extended sinusoidal family under AWGN attack

We have $\nabla t(\mathbf{r}) = -\sqrt{2}\sin(\mathbf{r}^T\boldsymbol{\lambda}_\mathbf{k})\boldsymbol{\lambda}_\mathbf{k}$. Therefore:

$$\mathrm{E}_{\mathbf{Z}}\{\nabla t(\mathbf{r}+\mathbf{z})\} = -\sqrt{2}(\sin(\mathbf{r}^T\boldsymbol{\lambda}_\mathbf{k})\mathrm{E}_{\mathbf{Z}}\{\cos(\mathbf{z}^T\boldsymbol{\lambda}_\mathbf{k})\} + \cos(\mathbf{r}^T\boldsymbol{\lambda}_\mathbf{k})\mathrm{E}_{\mathbf{Z}}\{\sin(\mathbf{z}^T\boldsymbol{\lambda}_\mathbf{k})\})\boldsymbol{\lambda}_\mathbf{k} \tag{101}$$

The last term is null when the pdf of $\mathbf{Z}$ is odd (ie. $p_{\mathbf{Z}}(\mathbf{z}) = p_{\mathbf{Z}}(-\mathbf{z})$) because $\sin(\mathbf{z}^T\boldsymbol{\lambda}_\mathbf{k})$ is even. Thus, if $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \sigma_z^2\mathbf{I})$, then $\mathrm{E}_{\mathbf{Z}}\{\nabla t(\mathbf{r}+\mathbf{z})\} = h(1,\sigma_z)t(\mathbf{r})$, with

$$\begin{aligned}
h(1,\sigma_z) &= \mathrm{E}_{\mathbf{Z}}\{\cos(\mathbf{z}^T\boldsymbol{\lambda}_\mathbf{k})\} && (102)\\
&= \mathrm{E}_{\mathbf{Z}}\{\cos(\sum_{i=1}^{p} z_i\lambda_{k,i})\} && (103)\\
&= \mathrm{E}_{Z_1}\{\cos(z_1\lambda_{k,1})\}\mathrm{E}_{\mathbf{Z}}\{\cos(\sum_{i=2}^{p} z_i\lambda_{k,i})\} - \mathrm{E}_{Z_1}\{\sin(z_1\lambda_{k,1})\}\mathrm{E}_{\mathbf{Z}}\{\sin(\sum_{i=2}^{p} z_i\lambda_{k,i})\} && (104)\\
&= e^{-\lambda_{k,1}^2\sigma_z^2/2}\mathrm{E}_{\mathbf{Z}}\{\cos(\sum_{i=2}^{p} z_i\lambda_{k,i})\} && (105)
\end{aligned}$$

Repeating $p-1$ times the last two lines, we finally get:

$$h(1,\sigma_z) = e^{-\|\boldsymbol{\lambda}_\mathbf{k}\|^2\sigma_z^2/2} = e^{-\eta(1,0)\sigma_z^2/2} \tag{106}$$

Therefore: $\eta(1,\sigma_z) = \eta(1,0)e^{-\eta(1,0)\sigma_z^2}$.

References

[1] I. Cox, M. Miller, and J. Bloom, *Digital Watermarking*, Morgan Kaufmann Publisher, 2001.

[2] N. Merhav, "An information-theoretic view of watermarking embedding-detection and geometric attacks," presented at WaCha05, available at `www.ee.technion.ac.il/people/merhav/`, jun 2005.

[3] M. Miller, I. Cox, and J. Bloom, "Informed embedding: exploiting image and detector information during watermark insertion," in *Proc. of Int. Conf. on Image Processing*, Vancouver, Canada, September 2000, IEEE.

[4] J. Delhumeau, T. Furon, N. Hurley, and G. Silvestre, "Improved polynomial detectors for side-informed watermarking," in *Security and Watermarking of Multimedia Contents IV*, Santa Clara, Cal., USA, January 2003, SPIE Electronic Imaging, pp. 311–321.

[5] T. Liu and P. Moulin, "Error exponents for one-bit watermarking," in *Proc. of ICASSP*, Hong-Kong, apr 2003.

[6] J.P. Andreaux, A. Durand, T. Furon, and E. Diehl, "Copy protection system for digital home networks," *IEEE Signal Processing Magazine*, vol. 21, no. 2, pp. 100–108, March 2004, Special Issue on Digital Right Management.

[7] Wikipedia, "Analog hole," *Wikipedia, The Free Encyclopedia*, vol. `http://en.wikipedia.org/w/index.php?title=Analog_hole&oldid=38835021`, 2006.

[8] E. Diehl and T. Furon, "Closing the analog hole," in *Proc. IEEE Int. Conf. Consumer Electronics*, 2003, pp. 52–53.

[9] E. Lin, A. Eskicioĝlu, R. Lagendijk, and E. Delp, "Advances in digital video content protection," *Proc. of IEEE*, vol. 93, no. 1, pp. 171–183, jan 2005.

[10] H. Vincent Poor, *An introduction to signal detection and estimation*, vol. 2nd edition, Springer, 1994.

[11] P. Huber, *Robust statistics*, J. Wiley and Sons, 1991.

[12] Q. Cheng and T. Huang, "Robust optimum detection of transform domain multiplicative watermarks," *IEEE Trans. Sig. Processing*, vol. 51, no. 4, pp. 906–924, apr 2003.

[13] A. Briassouli and M. Strinzis, "Locally optimum nonlinearities for DCT watermarking detection," *IEEE Trans. Image Processing*, vol. 13, no. 12, pp. 1604–16017, dec 2004.

[14] M. Barni, F. Bartolini, A. de Rosa, and A. Piva, "Optimum decoding and detection of multiplicative watermarks," *IEEE Trans. Signal Processing*, vol. 51, no. 4, pp. 1118–1123, apr 2003.

[15] X. Huang and B. Zhang, "Robust detection of transform domain additive watermarks," in *Proc. of Int. Work. on Digital Watermarking*, M. Barni, Ed., Siena, Italy, sep 2005, vol. 3710 of *LNCS*, pp. 124–138, Springer.

[16] T. Furon, G. Silvestre, and N. Hurley, "JANIS: Just Another N-order side-Informed Scheme," in *Proc. of Int. Conf. on Image Processing ICIP'02*, Rochester, NY, USA, September 2002, vol. 2, pp. 153–156.

[17] L. Pérez-Freire, P. Comesa na, and F. Pérez-González, "Detection in quantization-based watermarking: performances and security issues," in *Security, Steganography, and Watermarking of multimedia contents VII*, E. Delp and P. W. Wong, Eds., San jose, CA, USA, jan 2005, vol. 5681 of *Proc. of SPIE-IS&T Electronic Imaging*, pp. 721–733.

[18] J. Eggers and B. Girod, *Informed Watermarking*, Kluwer Academic Publishers, 2002.

[19] U. Erez and R. Zamir, "Achieving $0.5 \log 1 + SNR$ on additive white gaussian noise channel with lattice encoding and decoding," *IEEE Tran. on IT*, pp. 2293–2314, oct 2004.

[20] P. Moulin, A. Goteti, and R. Koetter, "Optimal sparse-QIM codes for zero-rate blind watermarking," in *Proc. of ICASSP*, Montreal, may 2004.

[21] T. Furon, J. Josse, and S. Le Squin, "Some theoretical aspects of watermarking detection," in *Proc. Security, steganography and watermarking of multimedia content*, San Jose, CA, USA, jan 2006.

[22] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, December 1997.

[23] P. Moon and D. E. Spencer, "Theorems on separability in riemannian n-space," *Proc. Amer. Math. Soc.*, vol. 3, pp. 635–642, 1952.

[24] N. Merhav and E. Sabbag, "Optimal watermarking embedding and detection strategies under limited detection resources," *submitted to IEEE Trans. on Inf. Theory*, 2006.

[25] F. Balado, "New geometric analysis of spread-spectrum data hiding with repetition coding, with implications for side-informed schemes," in *Proc. of Int. Work. on Digital Watermarking*, M. Barni, Ed., Siena, Italy, sep 2005, vol. 3710 of *LNCS*, pp. 336–350, Springer-Verlag.

[26] P. Moulin and R. Koetter, "Data hiding codes," *Proceedings of the IEEE*, dec 2005.

[27] T. Kalker, "Considerations on watermarking security," in *Proc of the Fourth Workshop on Multimedia Signal Processing (MMSP)*, J.-L. Dugelay and K. Rose, Eds., Cannes, France, October 2001, IEEE, pp. 201–206.

[28] T. Cover and J. Thomas, *Elements of information theory*, Number ISBN-0-471-06259-6 in Wiley series in telecommunications. Wiley, 1991.

[29] H.S. Malvar and D.A.F. Florêncio, "Improved spread spectrum: A new modulation technique for robust watermarking," *IEEE Trans. on Signal Processing*, vol. 51, no. 4, pp. 868–905, April 2003, Special Issue on Signal Processing for Data Hiding in Digital Media and Secure Content Delivery & secure content delivery, IEEE Trans. on Signal Processing.