A NEW LOOK AT ML STEP-SIZE ESTIMATION FOR SCALAR COSTA SCHEME DATA HIDING

Gabriel Domínguez-Conde[#], Pedro Comesaña-Alfaro[#], and Fernando Pérez-González[#]*

[#]University of Vigo, EE Telecomunicación, Campus Universitario
[*] Gradiant (Galician Research and Development Center in Advanced Telecommunications)
36310 Vigo-SPAIN
{gdomin|pcomesan|fperez}@gts.uvigo.es

ABSTRACT

Watermarking schemes based on the Dirty Paper Coding (DPC) paradigm have been shown to achieve much higher rates than classical Spread-Spectrum methods. However, in practice, the latter continue to be used due to their higher security and robustness. In fact, the most prevalent DPC method, the so-called Scalar Costa Scheme (SCS), is prone to non-additive attacks, such as a simple gain which produces a desynchronization between the embedding and decoding codebooks thus severely affecting performance. Although some gain-robust modifications to the basic SCS exist, all have serious drawbacks. One alternative, which was somehow abandoned for its complexity, is to estimate the gain at the decoder, with the advantage of preserving the simplicity of SCS. In this paper we take a new look at the estimation problem and propose an affordable algorithm to perform Maximum Likelihood estimation of the channel gain, that is able to restore the original SCS performance. We also show and experimentally illustrate how our scheme can be effectively adapted to watermark decoding in filtered images.

Index Terms— Dirty paper coding, gain attack, image filtering, maximum likelihood, watermarking.

1. INTRODUCTION

The advantages of Dirty Paper Coding (DPC) techniques in watermarking have been widely recognized [1, 2, 3]. Specifically, DPC-based schemes can achieve the channel capacity for Additive White Gaussian Noise (AWGN) channels [4]. This good performance is obtained thanks to the host rejection property of DPC techniques, which rely on the quantization of the host signal by using a codebook (indexed by the embedded symbol) with multiple codewords (contrarily to Spread-Spectrum techniques, where the involved codebook contains a single codeword). Unfortunately, for typical DPC codebooks, such as the scalar ones used in the prevalent Scalar Costa Scheme (SCS), the multiplicity of codewords makes them sensitive to amplitude modifications of the codebook.

A simple but devastating special case is the fixed gain attack (a.k.a. linear valumetric attack), in which the channel simply multiplies the watermarked signal by a constant real number. Even such a simple channel has shown to have dramatic consequences on the decoding of SCS, yielding very large probabilities of decoding error. Due to its relevance, several approaches have been proposed in the literature to cope with this problem. They can be roughly classified into two categories:

- *Robust codebooks*: in this case the typical SCS codebooks [3] are replaced by codebooks implicitly robust against the gain attack [5, 6, 7]. While [7] proposes to use phase-based codebooks (contrarily to magnitude-based ones), in [6] the information is embedded by considering the maximum correlation between the host signal and a pseudo-randomly generated set of sequences, and in [5] a codebook that depends on the empirical statistics of the watermarked signal is used.
- *Gain equalization*: in this case standard codebooks are used, but the channel effect is equalized by estimating the gain value and dividing the received signal by the estimate [8, 9]. Balado *et al.* [8] developed a method based on uniform scalar quantizers and turbocodes that iteratively estimates the gain factor, compensates its effect, and decodes the embedded message. On the other hand, Shterev and Lagendijk [9] proposed an exhaustive-search-based implementation of the Maximum-Likelihood (ML) estimation of the gain factor; again, this value is used for equalizing the observations, and performing the decoding with the original codebook.

Despite these works, overcoming the gain attack is still an open problem, as embedding distortion is difficult to be controlled in phase quantization based techniques [7] and orthogonal dirty paper coding [6] (which is also more computationally demanding than SCS), the computational cost of [8, 9] is substantial, and the work in [5] requires a sample buffer to be filled before decoding can be performed in a robust way.

In this paper, we follow the equalization approach proposed in [3] and later developed in [8, 9] for which an estimate of the channel gain is needed. Our work is similar to [9] in that Maximum Likelihood (ML) estimation is performed, but instead of carrying out a computationally prohibitive exhaustive search, we propose a computationally efficient algorithm that considers a novel approximation of the likelihood function. The advantages of our new approach are shown on both scaled synthetic signals and filtered natural images.

The remaining of this paper is organized as follows: Sect. 2 overviews SCS, while Sect. 3 focuses on the description of the proposed gain estimation; then, this scheme is adapted to deal with filtered natural images in Sect. 4. Finally, experimental results are presented in Sect. 5, and conclusions are drawn in Sect. 6

1.1. Notation

Real random variables are denoted with capital letters (e.g., X) and their outcomes with lowercase letters (e.g., x). Similarly, L-

Research supported by the European Regional Development Fund (ERDF) and the Spanish Government under project COMONSENS (CONSOLIDER-INGENIO 2010 CSD2008-00010), and the Galician Regional Government under projects "Consolidation of Research Units" 2009/62, 2010/85.

dimensional real random vectors and their outcomes are denoted by bold letters (e.g., $\mathbf{X}$ and $\mathbf{x}$, respectively), and their jth component is indicated by a subindex (X_j and x_j). The probability density function (pdf) of random variable X is denoted by $f_X(x)$, its mean by $\mathrm{E}\{X\}$, and its standard deviation by σ_X. Superscript S is used for denoting variables/vectors in the spatial domain.

2. OVERVIEW OF SCS DATA HIDING

In the sequel we will focus on the binary implementation of SCS, i.e., the case where two scalar quantizers (corresponding to the embedded bit) are used. The binary vector $\mathbf{m}$ is embedded by modifying the host signal $\mathbf{x}$ (we assume $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \sigma_X^2 I_{L\times L})$) to the watermarked signal $\mathbf{y}$, which is given by

$$y_i = x_i + \alpha\left(\mathcal{Q}_\Delta\left(x_i - d_i - m_i\frac{\Delta}{2}\right) - \left(x_i - d_i - m_i\frac{\Delta}{2}\right)\right),$$

for all $1 \le i \le L$, where $\mathcal{Q}_\Delta(\cdot)$ stands for the uniform scalar quantizer with step-size Δ, $\alpha \in (0, 1]$ is the so-called distortion compensation parameter, and $\mathbf{d}$ is the dither vector, which is a secret-key-dependent realization of $\mathbf{D} \sim U([-\Delta/2, \Delta/2]^L)$. Therefore, the watermark signal is $\mathbf{w} \triangleq \mathbf{y} - \mathbf{x}$, and the mean embedding distortion $D_w \triangleq \mathrm{E}\left\{\mathbf{W}^2\right\}/L$ is generally constrained in terms of a minimum Document to Watermark Ratio (DWR) defined as σ_X^2/D_w.

Under the fixed gain attack the received signal $\mathbf{z}$ is defined as

$$\mathbf{z} = t_0\left(\mathbf{y} + \mathbf{n}_1\right) + \mathbf{n}_2, \tag{1}$$

where t_0 is a real gain factor, $\mathbf{N}_1 \sim \mathcal{N}(\mathbf{0}, \sigma_{N_1}^2 I_{L\times L})$, $\mathbf{N}_2 \sim \mathcal{N}(\mathbf{0}, \sigma_{N_2}^2 I_{L\times L})$, $\mathbf{N}_1$ and $\mathbf{N}_2$ are mutually independent and also independent of $\mathbf{Y}$. Be aware that in [8, 9] $\mathbf{z}$ corresponds to $\mathbf{z} = t_0\left(\mathbf{y} + \mathbf{n}_1\right)$; therefore, the model considered here is slightly more general (its usefulness in modeling practical situations will be shown in Sect. 4).

The most extended implementation of the decoder, estimates the ith embedded bit as

$$\hat{m}_i = \underset{m\in\{0,1\}}{\operatorname{argmin}} \left|\mathcal{Q}_\Delta\left(z_i - d_i - m\frac{\Delta}{2}\right) - \left(z_i - d_i - m\frac{\Delta}{2}\right)\right|.$$

However, if $t_0 \neq 1$ the embedding and decoding codebooks will be misaligned with the consequence of significantly increasing the decoding error probability [5]. This problem could be easily solved if the gain factor t_0 were known; as this is not the case, it must be estimated from the received samples. To this end, it is possible to take advantage of the structure of the watermarked signal distribution (which is induced by SCS embedding). The decoder can exploit this estimate, denoted by $\hat{t}_0(\mathbf{z})$, to equalize the received samples before decoding. Specifically,

$$\hat{m}_i = \underset{m\in\{0,1\}}{\operatorname{argmin}} \left|\mathcal{Q}_\Delta\left(\frac{z_i}{\hat{t}_0(\mathbf{z})} - d_i - m\frac{\Delta}{2}\right) - \left(\frac{z_i}{\hat{t}_0(\mathbf{z})} - d_i - m\frac{\Delta}{2}\right)\right|. \tag{2}$$

3. GAIN FACTOR ESTIMATION

Since *a priori* knowledge of t_0 is not available in general, we propose to obtain an approximation of the Maximum Likelihood (ML) estimate of t_0.[1] Due to the componentwise independence of $\mathbf{Z}$, the ML estimate is calculated as $\hat{t}_0(\mathbf{z}) = \operatorname{argmin}_t L(t, \mathbf{z})$, where $L(t, \mathbf{z}) \triangleq -2\sum_{i=1}^{L} \log f_{Z|T,K}(z_i|t, d_i)$, and the embedded bits are modeled by $\mathbf{M} \sim \mathrm{Binomial}(L, 1/2)$.

[1] A similar algorithm was proposed in [10] for the complex flat fading channel estimation problem.

Unfortunately, $L(t, \mathbf{z})$ is an involved function, so we propose to simplify the ML estimation by approximating the pdf of Z. Such approximation is based on the following assumptions: a) $\sigma_X^2 \gg \Delta^2/12$ (verified for a wide range of real applications) in order to use the flat-host assumption (see [11]); b) the scaled self-noise variance [12] is much smaller than the total channel noise variance, i.e., $(1-\alpha)^2 t_0^2 \Delta^2/12 \ll t_0^2\sigma_{N_1}^2 + \sigma_{N_2}^2$; c) the variance of the total noise (self-noise plus total channel noise) is larger than the second moment of the scaled quantization lattice used at the decoder, i.e., $(1-\alpha)^2 t_0^2\Delta^2/12 + t_0^2\sigma_{N_1}^2 + \sigma_{N_2}^2 > t_0^2\Delta^2/48$; and d) the variance of the total noise is much smaller than the variance of the scaled host, i.e., $(1-\alpha)^2 t_0^2\Delta^2/12 + t_0^2\sigma_{N_1}^2 + \sigma_{N_2}^2 \ll t_0^2\sigma_X^2$. Under these hypotheses, $f_{Z|T,K}(z|t, d)$ can be approximated as

$$f_{Z|T,K}(z|t,d) \approx \frac{e^{-\frac{z^2}{2\sigma_X^2 t^2}}}{\sqrt{2\pi\sigma_X^2 t^2}} \times \left(1 + 2e^{-\frac{2\pi^2\left(\sigma_{N_2}^2 + t^2\left(\sigma_{N_1}^2 + \frac{(1-\alpha)^2\Delta^2}{12}\right)\right)}{(\Delta/2)^2 t^2}} \cos\left(\frac{2\pi z}{\Delta/2t} - \frac{2\pi d}{\Delta/2}\right)\right).$$

Under the assumptions introduced above, the absolute value of the argument of the exponential is much larger than 1; consequently, since $\log(1+u) \approx u$ for $|u| \ll 1$, $L(t, \mathbf{z})$ can be approximated as

$$L(t,\mathbf{z}) \approx \frac{\|\mathbf{z}\|^2}{\sigma_X^2 t^2} + L\log(2\pi\sigma_X^2 t^2) - 4\sum_{i=1}^{L} e^{-\frac{2\pi^2\left(\sigma_{N_2}^2 + t^2\left(\sigma_{N_1}^2 + \frac{(1-\alpha)^2\Delta^2}{12}\right)\right)}{(\Delta/2)^2 t^2}} \cos\left(\frac{2\pi z_i}{\Delta/2t} - \frac{2\pi d_i}{\Delta/2}\right). \tag{3}$$

The approximate $L(t, \mathbf{z})$ has a large number of local minima, so standard off-the-shelf optimization algorithms can not be used; furthermore, brute-force minimization (as that proposed in [9]) is computationally prohibitive. We thus propose an *ad-hoc* optimization algorithm, which will be shown to be computationally efficient.

First, a search-interval for the absolute value of t_0 is obtained from the variance-based unbiased estimate of t_0^2. Specifically,

$$\hat{t}_{0,\mathrm{var}}^2(\mathbf{z}) = \frac{\frac{\|\mathbf{z}\|^2}{L} - \sigma_{N_2}^2}{\sigma_X^2 + \sigma_W^2 + \sigma_{N_1}^2}. \tag{4}$$

If L is large enough to use the Central Limit Theorem (CLT), then the distribution of $\hat{t}_0^2(\mathbf{z})_{\mathrm{VAR}}$ can be approximated by $\mathcal{N}(t_0^2, 2(t_0^2(\sigma_X^2 + \sigma_W^2 + \sigma_{N_1}^2) + \sigma_{N_2}^2)^2/(L(\sigma_X^2 + \sigma_W^2 + \sigma_{N_1}^2)^2))$, and t_0^2 will be within $[t_-^2, t_+^2]$ with large probability, where

$$t_\pm^2 \triangleq \max\left(\epsilon, \hat{t}_{0,\mathrm{var}}^2(\mathbf{z}) \pm K_2\sqrt{\frac{2(\hat{t}_0^2(\mathbf{z})_{\mathrm{var}}(\sigma_X^2 + \sigma_W^2 + \sigma_{N_1}^2) + \sigma_{N_2})^2}{L\left(\sigma_X^2 + \sigma_W^2 + \sigma_{N_1}^2\right)^2}}\right);$$

$\epsilon > 0$ guarantees that both t_-^2 and t_+^2 take positive values, and $K_2 \ge 0$ controls the probability with which $|t_0|$ lies in the interval $[t_-, t_+]$.

Once it is available, the search interval $[t_-, t_+]$ is sampled, producing a candidate set $\mathcal{T}^+$; this sampling must be fine enough to guarantee that if a sample is within the main lobe of the target function, then at least one of its neighbors in the sampling set will be also in the main lobe. This sampling criterion has a computational rationale: since the target function is convex in the main lobe, regular convex optimization algorithms can be locally applied on the points in the candidate set, and convergence to the global minimum will be ensured.

Specifically, the sampling criterion is based on the factor in (3) defining the lobes, i.e., the cosine function argument. Indeed, we consider the variance of $(\mathbf{z} - t\mathbf{d})\text{mod}(t\Delta/2)^2$ when t is in a neighborhood of t_0, and for $t = t_0$; the sampled points $t(l)$ are iteratively computed as $t(l+1) = \frac{t(l)\left(\alpha\frac{\Delta^2}{48}+\sigma_X^2+\frac{\Delta}{2\sqrt{12}}\nu\right)}{\sigma_X^2+\frac{\Delta^2(1-K_1)}{48}}$, where $\nu \triangleq \sqrt{\Delta^2/48\left((1-\alpha)^2+K_1(2\alpha-1)\right)+K_1\sigma_X^2}$, $t(1) = t_-$, and the iterative sampling stops when $t(l) \geq t_+$. Parameter K_1 is introduced to control the separation between consecutive points in $\mathcal{T}^+$ and, thus, the cardinality of such set; the larger K_1, the smaller $|\mathcal{T}^+|$ (less computational cost), but the more likely it will be that $\mathcal{T}^+$ misses the main lobe of the target function, with a consequent performance loss. Since t_0 can be negative, by symmetry we define $\mathcal{T} = \mathcal{T}^+ \cup -\mathcal{T}^+$.

The centroid used at embedding is estimated for each $t \in \mathcal{T}$; this is done by equalizing the received observation, i.e, $c_j = \mathcal{Q}_{\Delta/2}(z_j/t - d_j) + d_j$, $j = 1, \ldots, L$. Then, given $t \in \mathcal{T}$, the vector of centroids $\mathbf{c}$ is estimated, and from this choice the minimum mean square error gain factor, i.e., $t^* \triangleq \text{argmin}_t \|\mathbf{z} - t\mathbf{c}\|^2$, is computed. It is easy to show that $t^* = (\mathbf{z}^T\mathbf{c})/\|\mathbf{c}\|^2$. We will denote by $\mathcal{T}^*$ the set of local optimizers t^* thus obtained. Note that $|\mathcal{T}^*| \leq |\mathcal{T}|$. Since the sampling method guarantees that at least one $t \in \mathcal{T}$ belongs to the main lobe, the ML estimate of t_0 is finally approximated by $\hat{t}_0(\mathbf{z}) \approx \text{argmin}_{t\in\mathcal{T}^*} L(t, \mathbf{z})$.

4. ADAPTATION TO FILTERED IMAGES

An interesting application of the technique introduced in the previous section goes beyond a pure scaling and considers a watermarked image that is convolved with a linear filter. From the estimation result, the embedded bits must be reliably extracted. In this section we assume the embedding to be performed in the full-frame Discrete Cosine Transform (DCT)[3] domain, and the considered filters to be circularly symmetric; therefore, $\mathbf{x}$ will denote the coefficients in that domain of a gray level image $\mathbf{x}^S$ of size $N_r \times N_c$.

Typically, the energy of natural images is concentrated at the low frequencies, which are the most perceptually significant components. Therefore, an attacker could remove the high frequencies without a large semantic distortion; consequently, most robust watermarking schemes embed the messages at the low-middle frequencies, excluding the DC component (e.g., [14]).

After embedding, the full-frame Inverse DCT (IDCT) of $\mathbf{y}$ is calculated to obtain $\mathbf{y}^S$. The pixel values of the watermarked image are rounded to the nearest integer and clipped; this operation, which is modeled by the addition of $\mathbf{n}_1$ in (1), is denoted by rclip$(\cdot)$

$$\text{rclip}(y_i^s) = \begin{cases} \text{round}(y_i^s) & \text{if} \quad y_i^s \in [0, 2^q - 1] \\ 0 & \text{if} \quad y_i^s < 0 \\ 2^q - 1 & \text{if} \quad y_i^s > 2^q - 1, \end{cases}$$

where round$(\cdot)$ stands for the round function, and q denotes the pixel depth. Then, the watermarked image is filtered (and subsequently rounded and clipped) in the spatial domain, yielding $\mathbf{z}^S = (\mathbf{y}^S + \mathbf{n}_1^S) * \mathbf{h}^S + \mathbf{n}_2^S$, where $*$ denotes the convolution operation (we consider $\mathbf{z}^S$ to have the size of $\mathbf{y}^S$ and $\mathbf{n}_1^S$), $\mathbf{h}^S$ is an $N_r^h \times N_c^h$-sized spatial filter, and $\mathbf{n}_2^S$ models the rclip$(\cdot)$ operation after filtering.

Assuming $N_r \gg N_r^h$ and $N_c \gg N_c^h$, as customary, the filtering border effect is neglected in our analysis; the spatial domain filtering is approximated by a DCT domain frequency-dependent gain (although one must be aware that the filtering effect is not purely multiplicative). So, if one can estimate the gain factor corresponding to each frequency, then the SCS decoder in (2) may be used.

This gain estimate will be performed blockwise, relying on the assumption that the filter frequency response to be approximately constant within each block. Non-overlapped $N_B \times N_B$-sized blocks are used. If N_B were too large, then the frequency response could no longer be assumed constant within each block; on the other hand, if N_B were too small, then the estimate precision will be poor, due to the small number of samples.

We assume the AC full-frame DCT coefficients used for embedding to be i.i.d. zero-mean Gaussian distributed with known variance, and independent of the coefficients in other blocks. Furthermore, rclip$(\cdot)$ is modeled in the spatial domain by both $\mathbf{N}_1^S$ (rounding and clipping due to the pixel domain transformation of the watermarked image, before filtering) and $\mathbf{N}_2^S$ (rounding and clipping due to the pixel domain casting of the filtered image) following independent $U([-1/2, 1/2]^L)$ distributions. If $N_r \cdot N_c$ is large enough, the CLT can be applied, and $\mathbf{N}_1$ and $\mathbf{N}_2$ can be approximated to be i.i.d. zero-mean Gaussian distributed with variance $1/12$.

5. EXPERIMENTAL RESULTS

In this section we compare, by using synthetic signals, the performance of our proposed method with that of previous schemes in the literature; we also illustrate the application to filtered images. Throughout this section, the parameters of our method have been set to $K_1 = 10^{-3}$, $K_2 = 10$, and $\epsilon = 10^{-3}$. For the sake of comparison it will be useful to define the effective WNR as $\text{WNR}_e \triangleq t_0^2\sigma_W^2/(t_0^2\sigma_{N_1}^2 + \sigma_{N_2}^2)$, and $\alpha_{\text{Costa}} \triangleq \text{WNR}_e/(\text{WNR}_e + 1)$.

First, assuming that $t_0 > 0$, we compare the performance in terms of the Bit Error Rate (BER), of the scheme described in Sect. 3 with that of Balado *et al.* [8]. The results are shown in Fig. 1, where the *turbo-code* used in [8] is employed, i.e., a $1/15$ turbo code based on the *recursive systematic convolutional code* $\mathbf{g} = (31, 21, 25, 35, 23, 33, 27, 37)$ (octal coding) and interleaver size of 10^3 uncoded bits (yielding $L = 1.5 \cdot 10^4$) [15]. This coding is also considered for the results of the current approach shown in Fig. 1. It is worth noting that in order to reduce the complexity, our gain factor estimation algorithm does not explicitly exploit the code structure; in other words, for the results of the current method in Fig. 1 the code error correcting capabilities are employed solely for message decoding once the received signal is equalized by $\hat{t}_0(\mathbf{z})$. Hence, further improvements in the gain factor estimation would be afforded by exploiting the code underlying structure at the expense of a higher computational cost.

Fig. 1 shows that our scheme outperforms [8] for all the considered WNR_e's, except for $\text{WNR}_e \approx 1.76$ dB, where no decoding errors were found for either.[4] This is not surprising as this WNR_e corresponds to $t_0 = 1$. Indeed, the large sensitivity of [8] to gain attacks even slightly different from 1 is shown by the authors in their original paper; for the sake of numerical illustration, in Fig. 1 the gains corresponding to $\text{WNR}_e = 1$ and 2 dB are $t_0 \approx 0.850$ and $t_0 \approx 1.058$, respectively. Fig. 1 also shows the results obtained by initializing the scheme in [8] with the variance-based estimate introduced in (4); this initialization of Balado *et al.*'s method, newly proposed here, achieves the best results among all three methods for very small values of WNR_e (where the error in the variance-based estimate is very small), but it is clearly outperformed by the scheme described in Sect. 3 when larger WNR_e's are considered (corresponding to larger values of the estimate variance).

[2]The modulo operation is defined as $A \text{ mod } B \triangleq A - \mathcal{Q}_B(A)$.

[3]The definition proposed in [13] is used.

[4]Be aware that also no decoding errors were found for our method when $\text{WNR}_e = 1$ and 3 dB.

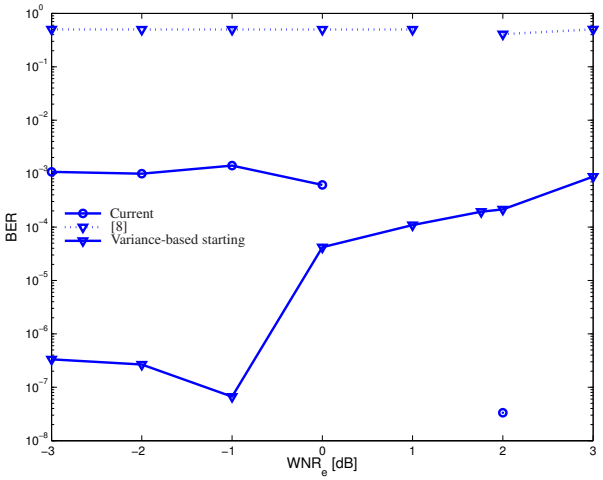

Fig. 1. BER as a function of WNR$_e$ for the method in [8], its variation when it is initialized by the variance-based estimate (4), and the current proposal. DWR $= 30$ dB, $L = 1.5 \cdot 10^4$, and $\alpha = \alpha_{\text{Costa}}$.

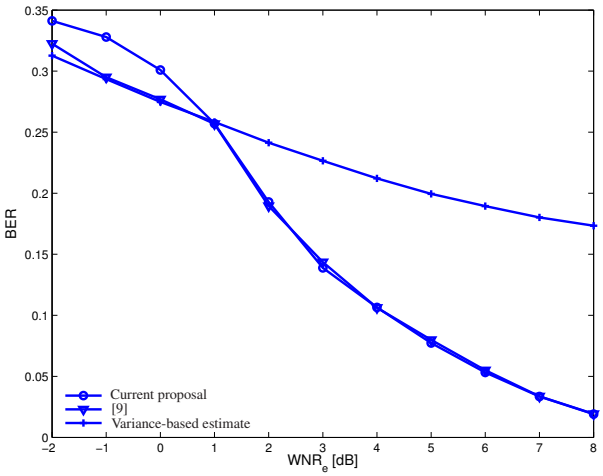

Fig. 2. BER as a function of WNR$_e$ for the method in [9], the variance-based estimate (4), and the current proposal. DWR $= 30$ dB, $L = 10^3$, and $\alpha = \alpha_{\text{Costa}}$.

Fig. 2 shows the BER as a function of WNR$_e$ for [9], the variance-based estimate in (4), and our proposal when channel coding is not used, and $t_0 > 0$. [9] is carried out by sampling finely enough a search interval. Special attention was paid to reducing its computational cost as much as possible (e.g., precomputing the pdfs depending on a quantized version of the dither).

Since [9] uses the exact received signal pdf and exhaustive search, it was expected to provide the best results, as it is indeed the case. Furthermore, and similarly to Fig. 1, the variance-based estimate outperforms our proposal for very low values of the WNR$_e$, as the structure on $f_{\mathbf{Y}}(\mathbf{y})$ induced by the watermark embedding is no longer observable; however, for larger WNR$_e$'s such structure is made evident, and our scheme clearly improves the results of the variance-based estimate. It is also interesting to note that Shterev and Lagendijk's method behaves almost exactly as the best result among the variance-based estimate and our proposal, showing that both schemes are good choices (depending on the WNR$_e$) to be used as alternatives to the method proposed in [9], with a dramatic reduction in the computational cost over the latter. Specifically, each Monte Carlo trial of [9] for WNR$_e = 6$ dB carried out in MatlabR2013b using a Core i5-2500 3.3GHZ 16 GB PC requires around 50 s, while our proposal approximately needs only 0.3 s.

Finally, Fig. 3 shows the results of the filtered-image-targeted adaptation proposed in Sect. 4 for a low-pass 5×5 spatial Gaussian filter with standard deviation 1, and a test set of 100 gray-converted 384×512-sized images pseudo-randomly selected from the UCID v2 image database [16]. For the reasons given in Sect. 4, only the first 10 zigzag-ordered DCT coefficient blocks of size 64×64 are used for hiding data.

Fig. 3 shows the BER averaged over the test images for the considered blocks, when $N_B = 16$, $\alpha = 1$, and the Peak Signal to Noise Ratio (PSNR), defined in this case as $255^2/\sigma_W^2$, is set to 40 dB. According to the shown results, the block BER approximately takes values between 10^{-1} and 10^{-2}, which illustrates that our scheme can be practically used in this demanding scenario. In addition, the BER seems to depend on the actual value of h (its estimate $\hat{h}$ is shown in this figure) as one would expect since σ_W^2, $\sigma_{N_1}^2$, and $\sigma_{N_2}^2$ are approximately constant for all the watermarked blocks and, thus, the WNR$_e$ only changes with h. These BER results are supported by the accuracy of the obtained estimates; specifically, in this example the mean square estimation error (MSE) of the gain factors takes values approximately around -30 dB in medium frequencies and less than -40 dB for low frequencies (where the energy of the images is concentrated).

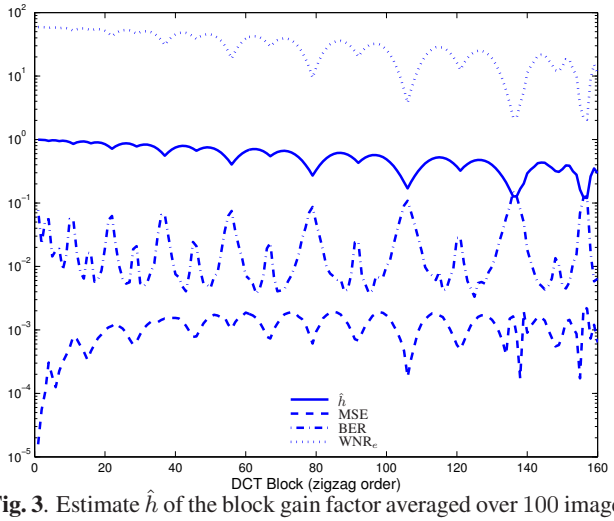

Fig. 3. Estimate $\hat{h}$ of the block gain factor averaged over 100 images, the mean square estimation error of the gain factor, averaged BER per block, and WNR$_e$ for each watermarked DCT block following the zigzag order. PSNR $= 40$ dB, $N_B = 16$, $\alpha = 1$, and Gaussian spatial filter of size 5×5 with standard deviation 1.

6. CONCLUSIONS

In this paper, a novel gain attack estimation algorithm based on ML is used to equalize the effect of this attack on SCS. Preserving the simplicity of SCS codebook, our proposal estimates the gain factor by taking advantage of the watermark signal pdf structure induced by the embedder; this estimate is used to equalize the received signal and decode the embedded message. The resulting method provides better results in terms of the BER than [8], partially due to the sensitivity of the latter to its starting point. On the other hand, the computational complexity of our scheme is substantially reduced in comparison with [9]; nevertheless, whenever the watermarked signal pdf structure arises, the BER achieved by the current proposal converges to that of [9], where the exact pdf of the received signal and exhaustive search are used. Finally, experiments with filtered real images ilustrate the usefulness of the proposed adaptation of our scheme for dealing with this real application scenario.

7. REFERENCES

[1] Max H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[2] Brian Chen and Gregory W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[3] Joachim Eggers and Bernd Girod, *Informed Watermarking*, Kluwer Academic Publishers, 2002.

[4] Uri Erez and Ram Zamir, "Achieving $\frac{1}{2}\log(1+\text{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293–2314, October 2004.

[5] Fernando Pérez-González, Carlos Mosquera, Mauro Barni, and Andrea Abrardo, "Rational dither modulation: A high-rate data-hiding method invariant to gain attacks," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3960–3975, October 2005.

[6] Andrea Abrardo and Mauro Barni, "Informed watermarking by means of orthogonal and quasi-orthogonal dirty paper coding," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 824–933, February 2005.

[7] Fabricio Ourique, Vinicius Licks, Ramiro Jordan, and Fernando Pérez-González, "Angle QIM: A novel watermark embedding scheme robust against amplitude scaling distortion," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2005, vol. II, pp. 797–780.

[8] Félix Balado, Kevin M. Whelan, Guénolé C. M. Silvestre, and Neil J. Hurley, "Joint iterative decoding and estimation for side-informed data hiding," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 4006–4019, October 2005.

[9] Ivo D. Shterev and Reginald L. Lagendijk, "Amplitude scale estimation for quantization-based watermarking," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4146–4155, November 2006.

[10] Gabriel Domínguez-Conde, Pedro Comesaña-Alfaro, and Fernando Pérez-González, "Flat fading channel estimation based on dirty paper coding," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2014, pp. 6529–6533.

[11] Luis Pérez-Freire, Pedro Comesaña, and Fernando Pérez-González, "Information-theoretic analysis of security in side-informed data hiding," in *Information Hiding*. Springer, 2005, pp. 131–145.

[12] Luis Pérez-Freire, Fernando Pérez-González, and Pedro Comesaña, "Secret dither estimation in lattice-quantization data hiding: a set membership approach," in *Electronic Imaging 2006*. International Society for Optics and Photonics, January 2006, pp. 60720W–1–60720W–12.

[13] A. K. Jain, *Fundamentals of Digital Image Processing*, pp. 150–153, Prentice Hall, 1988.

[14] Mauro Barni, Franco Bartolini, Alessia De Rosa, and Alessandro Piva, "Capacity of full frame DCT image watermarks," *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1450–1455, August 2000.

[15] Félix Balado, "Personal communication," February 2014.

[16] Gerald Schaefer and Michal Stich, "UCID - an uncompressed colour image database," in *SPIE Conference on Storage and Retrieval Methods and Applications for Multimedia*, January 2004, pp. 472–480.