Maximum Likelihood Amplitude Scale Estimation for Quantization-Based Watermarking in the Presence of Dither

Ivo D. Shterev and Reginald L. Lagendijk

Delft University of Technology, 2628 CD Delft, Netherlands
{i.shterev, r.l.lagendijk}@ewi.tudelft.nl

ABSTRACT

Quantization-based watermarking schemes comprise a class of watermarking schemes that achieves the channel capacity in terms of additive noise attacks.[1] The existence of good high dimensional lattices that can be efficiently implemented[2–4] and incorporated into watermarking structures, made quantization-based watermarking schemes of practical interest. Because of the structure of the lattices, watermarking schemes making use of them are vulnerable to non-additive operations, like amplitude scaling in combination with additive noise.

In this paper, we propose a secure Maximum Likelihood (ML) estimation technique for amplitude scaling factors using subtractive dither. The dither has mainly security purposes and is assumed to be known to the watermark encoder and decoder. We derive the probability density function (PDF) models of the watermarked and attacked data in the presence of subtractive dither. The derivation of these models follows the lines of,[5] where we derived the PDF models in the absence of dither. We derive conditions for the dither sequence statistics such that a given security level is achieved using the error probability of the watermarking system as objective function. Based on these conditions we are able to make approximations to the PDF models that are used in the ML estimation procedure. Finally, experiments are performed with real audio and speech signals showing the good performance of the proposed estimation technique under realistic conditions.

Keywords: watermarking, quantization, subtractive dither, probability of error, maximum likelihood estimation, statistics.

1. INTRODUCTION

Watermarking schemes based on quantization theory have recently emerged as a result of information theoretic analysis.[1,6] In terms of additive noise attacks, these schemes have proven to perform better than traditional spread spectrum watermarking because the used lattice codes achieve capacity for the AWGN channel. The existence of good lattices in high dimensions[2] that can be directly and efficiently implemented have made quantization-based schemes of practical interest. Another important feature of quantization-based watermarking schemes is that they can completely cancel the host signal interference, which makes them invariant to the host signal. A similar phenomenon exists in channel coding with side information at the encoder.[7,8]

Since lattices have a particular structure, watermarking schemes making use of them are vulnerable to non-additive attacks like amplitude scaling. Furthermore, operations like amplitude scaling, linear and non-linear filtering induce a large amount of distortion with respect to the mean-squared error, but do not cause significant perceptual degradations. Non-additive operations on watermarked signals are quite common in many applications. One example is audio play-out and capturing, where the watermarked signal is passed through an D-A converter, transmitted through an analog noisy channel, captured by a microphone, and converted back to a digital representation. Clearly the microphone will capture a less powerful and degraded watermarked signal, which has led us to model the noisy channel as an amplitude scaling operation followed by additive noise.

Furthermore, in channel coding with side information at the encoder using lattices, it is known that (see[9] for more information on lattice techniques) to achieve the capacity $0.5\log(1+SNR)$ over the AWGN channel, the decoder has to scale the received data with a factor α before lattice decoding. In the context of watermarking where the communication channel is not secure, an attacker may try to disturb the decoder by applying amplitude scaling with a factor β in addition to noise. To counter this scaling attack, the decoder has to know the scaling factor to apply the correct scaling $\frac{\alpha}{\beta}$ prior to lattice decoding.

Security, Steganography, and Watermarking of Multimedia Contents VII, edited by Edward J. Delp III, Ping Wah Wong, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 5681 © 2005 SPIE and IS&T · 0277-786X/05/$15

Several techniques are known in the literature for combating non-additive operations. One of the approaches is based on designing watermarking codes that are resilient to non-additive operations, such as modified trellis codes.[10,11] Another approach is based on estimating the non-additive operations and inverting them prior to watermark decoding.[12] However, most of the proposed techniques in the literature lack an underlying theoretical model and experimental validation with real signals.

In this paper we propose an ML procedure for estimating amplitude scaling factors, based on probabilistic models in the presence of subtractive dither. Moreover, in the watermarking literature, the dither has been mostly analyzed from a statistically independent quantization noise point of view, while the security aspects of the dither itself were absent. In this paper we also give conditions for the dither sequence statistics, such that a given level of security is achieved, using the probability of error of the watermarking system as an objective function. Adhering to the subtractive dither conditions we are able to make simplified approximations of the PDF models, on which the ML estimation procedure relies.

The paper is organized as follows. In Section 2 we formulate the problem mathematically, as well as the watermark encoder and decoder. In Section 3 we derive the PDF models in the presence of subtractive dither. In Section 4 we give conditions for the dither sequence such that an attacker without having knowledge of the dither is not able to decode the watermark. In Section 5 we give simplified approximations to the PDF models adhering to the subtractive dither conditions. A description of the estimation procedure is given in Section 6. Section 7 contains experimental results from real audio host signals, and Section 8 concludes the paper.

2. MATHEMATICAL FORMULATION

In this paper we focus on the most popular quantization-based watermarking scheme, namely Quantization Index Modulation with Distortion Compensation (QIM with DC). Throughout the paper, random variables are denoted by capital letters and their realizations by the respective small letters. The notation $X \sim f(x)$ indicates that the random variable X has a PDF $f(x)$.

Fig. 1 shows the watermark encoder, where $W \in \{0,1\}$ denotes the message bits that are embedded in the host data, $\widetilde{X}$ is the host signal itself with a variance $\sigma^2_{\widetilde{X}}$, X is the watermarked signal, D is the dither sequence with a variance σ^2_D, and U is the output of the quantizer. $Q(\cdot)$ denotes uniform quantization with step size Δ. The quantization noise is denoted by N_1 and has a variance $\sigma^2_{N_1}$. The heart of the watermark encoder is a quantizer, whose input-output characteristic is shown in Fig. 2.

The attack channel is shown in Fig. 3. It consists of the amplitude scale factor β and the noise $N_2 \sim \mathcal{N}(0, \sigma^2_{N_2})$. The coefficient $\alpha = \frac{\sigma^2_{N_1}}{\sigma^2_{N_1}+\sigma^2_{N_2}}$ in the encoder is known from.[8] In most applications, watermarking schemes operate in the small distortion case, i.e., $\sigma^2_{\widetilde{X}} \gg \sigma^2_{N_1}, \sigma^2_{N_2}$, so as to preserve the quality of the host signal. Therefore, high-resolution quantization theory is applicable, from which it follows that $N_1 \sim \mathcal{U}(0, \sigma^2_{N_1})$ and N_1 is statistically independent from the host signal $\widetilde{X}$. Moreover, the quantization noise is statistically independent of the host signal and uniformly distributed over the base quantization cell independently of the choice of the dither sequence, as long as D is statistically independent of the host signal. The total noise introduced by the watermark encoder is equal to N_1, which is shown as follows:

$$N_1 = \alpha\widetilde{X} - \big(X - (1-\alpha)\widetilde{X}\big) = \widetilde{X} - X. \tag{1}$$

The attacked (received) signal Y, which is an input to the watermark decoder, can be written in the following way:

$$Y = \beta X + N_2 = \beta\big(U - D + (1-\alpha)\widetilde{X}\big) + N_2. \tag{2}$$

Using the relation $\alpha\widetilde{X} = U - D + N_1$, we obtain the received data Y in terms of N_1, N_2, and the watermark-bearing signal U:

$$Y \;=\; \frac{\beta}{\alpha}\big(U - D + (1-\alpha)N_1\big) + N_2. \tag{3}$$

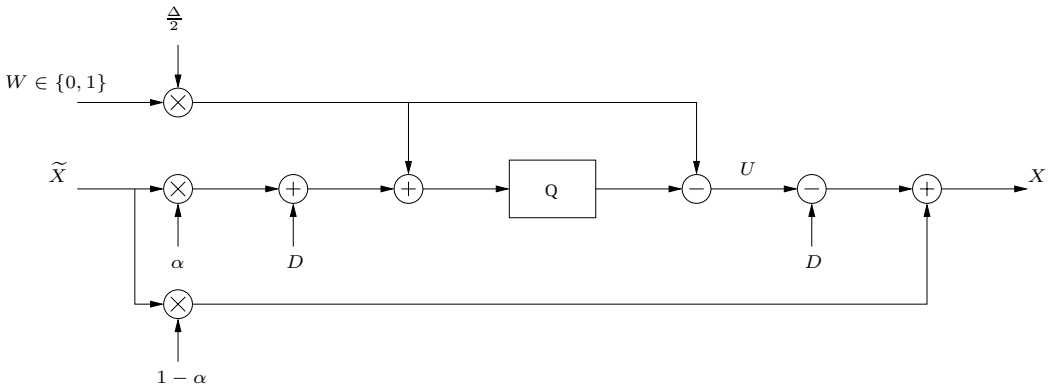

Figure 1. Watermark encoder.

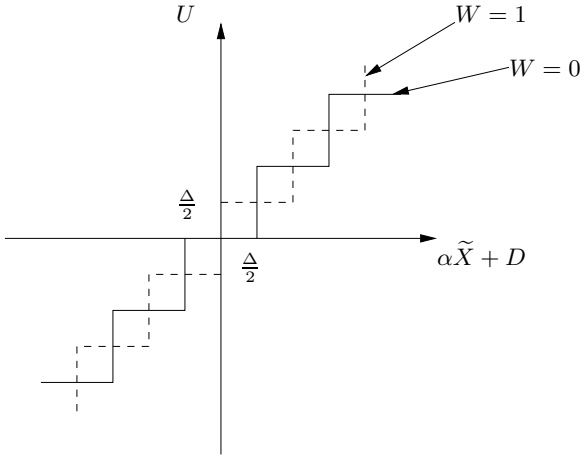

Figure 2. Quantizer input-output characteristics

From (3), we observe that the optimal countermeasure against the attacker's scale β is just scaling by $\frac{1}{\beta}$, which of course assumes that the decoder knows the value of β. The watermark decoding process is based on U and is depicted in Fig 4. To get an estimate that is as close to U as possible, the decoder first scales the received data by $\frac{\alpha}{\beta}$ and adds the dither D, thus obtaining:

$$\hat{U} = \frac{\alpha}{\beta}Y + D = U + (1-\alpha)N_1 + \frac{\alpha}{\beta}N_2. \tag{4}$$

The decoder then computes the absolute value of the quantization noise $\|\hat{U} - Q(\hat{U})\|$ and makes an estimate of the embedded watermark in the following way:

$$\hat{W} = \begin{cases} 0 & \text{if } \|\hat{U} - Q(\hat{U})\| \leq \frac{\Delta}{4} \\ 1 & \text{if } \|\hat{U} - Q(\hat{U})\| > \frac{\Delta}{4} \end{cases} \tag{5}$$

3. PDF MODELS

Since in the presence of subtractive dither the PDF of X will be perturbed by D, it is difficult to derive a useful exact mathematical expression for it. That is why we choose to manipulate X in a convenient way, having knowledge of D, so that we are able to mathematically describe the structure of the PDF of the resulting random variable.

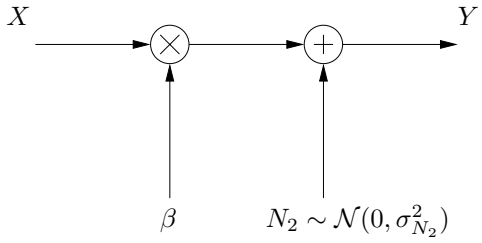

Figure 3. Attack Channel

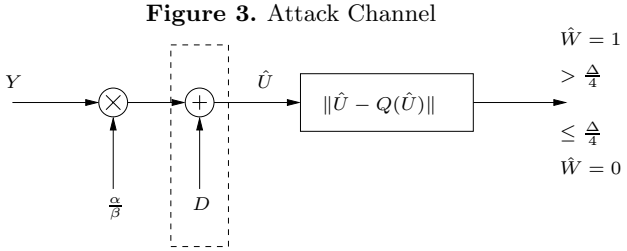

Figure 4. Watermark decoder

For the purpose of estimation and for simplicity, we will assume that only message $W = 0$ is embedded. Extension to the more general case of embedding zeros and ones is straightforward.

Referring to Fig. 1, let us assume that $\alpha\widetilde{X} + D$ belongs to the k-th quantization cell, i.e.:

$$\Delta\Big(k - \frac{1}{2}\Big) < \alpha\widetilde{X} + D < \Delta\Big(k + \frac{1}{2}\Big). \tag{6}$$

Multiplying by $\frac{1-\alpha}{\alpha}$ and adding $k\Delta$, we obtain:

$$\frac{\Delta}{\alpha}\Big(k - \frac{1-\alpha}{2}\Big) < (1-\alpha)\widetilde{X} + k\Delta + \frac{1-\alpha}{\alpha}D < \frac{\Delta}{\alpha}\Big(k + \frac{1-\alpha}{2}\Big). \tag{7}$$

Recognizing that $(1-\alpha)\widetilde{X} + k\Delta + \frac{1-\alpha}{\alpha}D = X + \frac{1}{\alpha}D$ (see Fig. 1), we can write the PDF of $X + \frac{1}{\alpha}D$ for a particular k as

$$f'_{X+\frac{1}{\alpha}D}(x) \quad = \quad f_{(1-\alpha)\widetilde{X}+k\Delta+\frac{1-\alpha}{\alpha}D}(x) I_{A_k|W=0}(x) \tag{8}$$

where $I_{A_{k|W=0}}(x)$ denotes the indicator function of the set $A_{k|W=0}$ defined as:

$$I_{A_{k|W=0}}(x) = \begin{cases} 1 & \text{if } x \in A_{k|W=0} \\ 0 & \text{if } x \notin A_{k|W=0} \end{cases} \tag{9}$$

and

$$A_{k|W=0} = \Big[\frac{\Delta}{\alpha}(k - \frac{1-\alpha}{2}), \frac{\Delta}{\alpha}(k + \frac{1-\alpha}{2})\Big]. \tag{10}$$

Generalizing for all k, we have

$$f_{X+\frac{1}{\alpha}D}(x) = \sum_{k=-\infty}^{+\infty} f_{(1-\alpha)\widetilde{X}+k\Delta+\frac{1-\alpha}{\alpha}D}(x) I_{A_k|W=0}(x). \tag{11}$$

Eq. (11) is the key expression for the estimation procedure in the presence of subtractive dither. We can see that although X is perturbed by the dither, if we add the term $\frac{1}{\alpha}D$ to the watermarked signal, we are able to obtain a signal that has a PDF with a structure.

Taking into account β and the additive noise N_2, we now have:

$$f_{Y+\frac{\beta}{\alpha}D}(x) \quad = \quad f_{N_2}(x) * f_{\beta X+\frac{\beta}{\alpha}D}(x), \tag{12}$$

where the convolution $*$ follows from the independence between N_2 and $\beta X + \frac{\beta}{\alpha}D$.

4. DESIGN OF THE DITHER SEQUENCE

Since the dither sequence has a security role in the watermarking system, we give conditions for the dither sequence statistics such that an attacker is not able to decode the watermark with an error probability* different than 0.5. These conditions will also allow for an approximation of the PDF models.

To derive the dither conditions, we first need to derive the error probability of the watermarking system, which is given by the following theorem.

Theorem 1: For the case where the dither sequence D is not known to the decoder, the error probability P_e of the watermarking system is given by the expression:

$$P_e = \sum_m \Pr[m\Delta - \frac{3\Delta}{4} \leq (1-\alpha)N_1 + \frac{\alpha}{\beta}N_2 - D \leq m\Delta - \frac{\Delta}{4}]. \quad (13)$$

Proof: The error probability P_e can be expressed as

$$\begin{aligned} P_e &= Pr[\hat{W}=1|W=0]Pr[W=0] + Pr[\hat{W}=0|W=1]Pr[W=1] \\ &= Pr[\hat{W}=1|W=0] \end{aligned} \quad (14)$$

where the last line follows from the fact that the encoder is a symmetric scheme of two quantizers, that the channel strategy is independent of the embedded message (i.e. $Pr(N_2|W) = Pr(N_2)$), and that $Pr(W=0) + Pr(W=1) = 1$. Therefore we can model the whole watermarking system, together with the attack channel, as a Binary Symmetric Channel (see Fig. 5).

From (14) and from Fig. 1, it is straightforward to show that the probability of error can be expressed as

$$\begin{aligned} P_e &= Pr[\hat{W}=1|W=0] \\ &= Pr[\|Q(\hat{U}-D) - (\hat{U}-D)\| \geq \frac{\Delta}{4}]. \end{aligned} \quad (15)$$

Using the following relation for scalar quantizers,

$$\|Q(\hat{U}-D) - (\hat{U}-D)\| = \|(\hat{U}-D+\frac{\Delta}{2}) \bmod \Delta - \frac{\Delta}{2}\|, \quad (16)$$

we have

$$P_e = Pr[\|(\hat{U}-D+\frac{\Delta}{2}) \bmod \Delta - \frac{\Delta}{2}\| \geq \frac{\Delta}{4}] \quad (17)$$

$$= Pr[(\hat{U}-D+\frac{\Delta}{2}) \bmod \Delta \leq \frac{\Delta}{4} \quad (18)$$

$$\bigcup \ (\hat{U}-D+\frac{\Delta}{2}) \bmod \Delta \geq \frac{3\Delta}{4}] \quad (19)$$

where $\bigcup$ denotes the union of two events.

Using (4) and taking into account that $U \in \Lambda$, the quantizer lattice, we can write

$$\begin{aligned} P_e &= Pr[((1-\alpha)N_1 + \frac{\alpha}{\beta}N_2 - D + \frac{\Delta}{2}) \bmod \Delta \leq \frac{\Delta}{4} \\ &\bigcup \ ((1-\alpha)N_1 + \frac{\alpha}{\beta}N_2 - D + \frac{\Delta}{2}) \bmod \Delta \geq \frac{3\Delta}{4}]. \end{aligned}$$

Using Number theory,[13] we can write that for any b, and any c such that $b > c > 0$, and any $a \neq mb$, where $m \in (-\infty, +\infty)$ is an integer, the solution to the inequalities

$$a \bmod b \geq c \quad (20)$$

$$a \bmod b \leq c \quad (21)$$

*Since we have one-dimensional, one-bit watermarking, the error probability and bit error probability are equal.

is

$$mb + c \leq a \leq (m+1)b \text{ and} \tag{22}$$

$$mb \leq a \leq mb + c \text{ , respectively} \tag{23}$$

Therefore, after simple arithmetics, we arrive at

$$P_e \quad = \quad \sum_m \Pr[m\Delta - \frac{3\Delta}{4} \leq (1-\alpha)N_1 + \frac{\alpha}{\beta}N_2 - D \leq m\Delta - \frac{\Delta}{4}]. \tag{24}$$

end of proof.

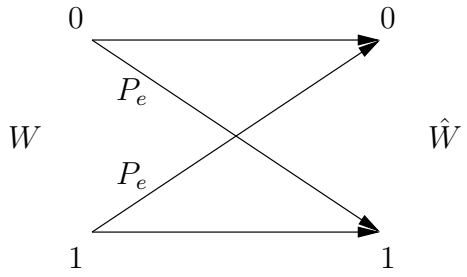

Figure 5. A representation of the watermarking system as a Binary Symmetric Channel with cross-over probability P_e.

An illustration of (24) for concrete values of the parameters $\sigma^2_{N_1}$, $\sigma^2_{N_2}$, σ^2_D, and β is shown in Fig. 6. The intervals $[m\Delta - \frac{3\Delta}{4}, m\Delta - \frac{\Delta}{4}]$ (black intervals in Fig. 6) cover half of the real axis.

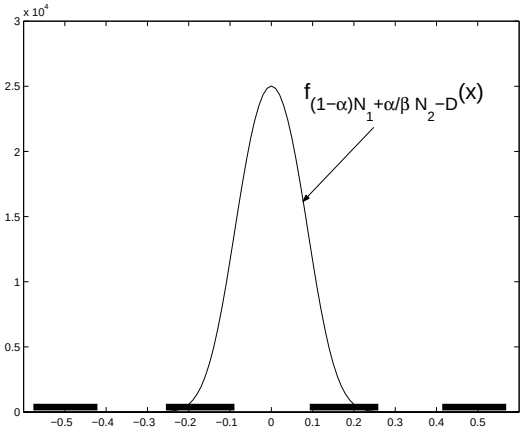

Figure 6. PDF of $(1-\alpha)N_1 + \frac{\alpha}{\beta}N_2 - D \sim f_{(1-\alpha)N_1+\frac{\alpha}{\beta}N_2-D}(x)$ for $N_1 \sim \mathcal{U}(0, 0.01)$, $N_2 \sim \mathcal{N}(0, 0.01)$, $D \sim \mathcal{U}(0, 0.01)$, and $\beta = 1$. The integral of $f_{(1-\alpha)N_1+\frac{\alpha}{\beta}N_2-D}(x)$ over the black intervals gives $P_e \approx 0.5$.

We would like to design the dither sequence statistics such that the error probability $Pe = 0.5$ for all choices of N_2. For this we will need to show the independence between N_1, N_2, and D. Since by assumption the noise N_2 is independent of N_1 and D, we will only need show the independence of N_1 and D. By definition this is equivalent to

$$f(N_1|D = d) \quad = \quad f(N_1). \tag{25}$$

From Fig. 1 and for particular k, we can express the quantization noise as

$$N_1 \quad = \quad -(\alpha\widetilde{X} + D) + k\Delta. \tag{26}$$

From (25) and (26), we can see that when the dither is known, the conditional PDF of the quantization noise is a function only of the PDF of $\alpha\widetilde{X}$, which we write as

$$f(N_1|D=d) \quad = \quad F\Big(f(\alpha\widetilde{X})\Big). \tag{27}$$

Since $\alpha\widetilde{X} = -D - N_1 + k\Delta$, we can write

$$f(N_1|D=d) = \sum_{k=-\infty}^{+\infty} f_{\alpha\widetilde{X}}(k\Delta - N_1 - d). \tag{28}$$

Following the same reasoning as in[14] with the roles of D and $\alpha\widetilde{X}$ interchanged, we can show that:

$$f(N_1|D=d) \quad = \quad f(N_1) \text{ , as } \alpha\widetilde{X} \sim \mathcal{U}(0, \sigma^2_{N_1}), \tag{29}$$

which is satisfied from the low-distortion case $\sigma^2 \gg \sigma^2_{N_1}, \sigma^2_{N_2}$, i.e., $\alpha\widetilde{X}$ is uniformly distributed over the base quantization interval.

If we exclude the terms N_1, and N_2 from P_e, we have

$$P_e^D \quad = \quad \sum_m \Pr[m\Delta - \frac{3\Delta}{4} \leq -D \leq m\Delta - \frac{\Delta}{4}], \tag{30}$$

which is actually the error probability when the attacker does not apply any additive noise.

It is easy to see that

$$P_e^D \quad \rightarrow \quad 0.5 \text{ , as } D \sim \mathcal{U}(0, \sigma^2_{N_1}). \tag{31}$$

This situation is illustrated in Fig. 7. Therefore, for security purposes, it is sufficient to choose a dither that is uniformly distributed over the interval $[-\frac{\Delta}{2}, +\frac{\Delta}{2}]$. Note that other distributions for D giving $P_e^D = 0.5$ also exist. We choose $f_D \sim \mathcal{U}(0, \sigma^2_{N_1})$ for simplicity reasons and because we want the power of D as small as possible.

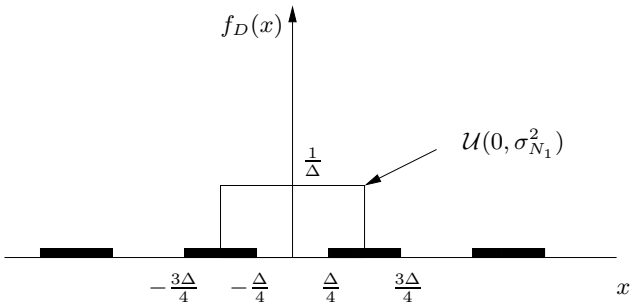

Figure 7. PDF of $D \sim \mathcal{U}(0, \sigma^2_{N_1})$ denoted as $f_D(x)$. The integral of $f_D(x)$ over the black intervals gives $P_{e|\sigma^2_{N_2}\rightarrow 0} = 0.5$.

Note that further increasing σ^2_D with respect to $\sigma^2_{N_1}$ will cause P_e^D to oscillate between the values $\frac{2}{5}$ and $\frac{2}{3}$, i.e., $P_e^D \in [\frac{2}{5}, \frac{2}{3}]$. An illustration of P_e^D as a function of σ^2_D is shown in Fig. 8. It can be seen that $P_e^D \rightarrow 0.5$ as $\frac{\sigma^2_D}{\sigma^2_{N_1}} \rightarrow \infty$, but oscillates for intermediate values. That is why we choose $D \sim \mathcal{U}(0, \sigma^2_{N_1})$ throughout the paper.

If we assume that the attacker applies small amount of noise with comparison to the watermark distortion, i.e. $\sigma^2_{N_2} \ll \sigma^2_{N_1} = \sigma^2_D$, then independently of $f_{N_2}(x)$, we have

$$\begin{aligned} P_{e|\sigma^2_{N_2} \ll \sigma^2_{N_1}} \quad &= \quad \sum_m \Pr[m\Delta - \frac{3\Delta}{4} \leq -D \leq m\Delta - \frac{\Delta}{4}] \\ &= \quad 0.5 \end{aligned} \tag{32}$$

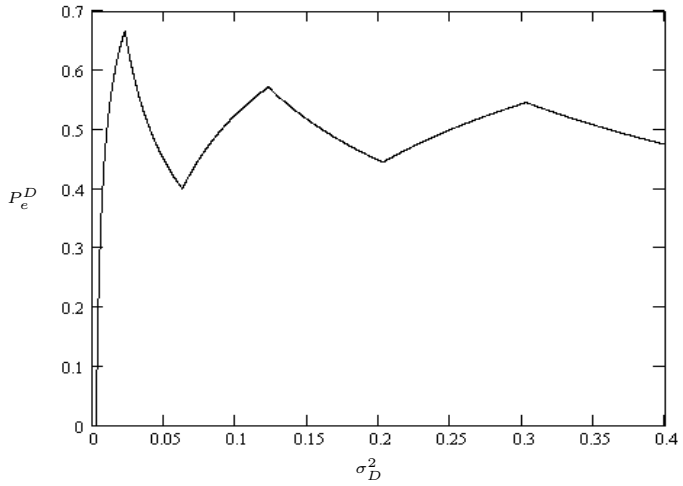

Figure 8. An illustration of P_e^D as a function of σ_D^2. The value for the watermark distortion is $\sigma_{N_1}^2 = 0.01$.

If we assume that the attacker applies a large amount of noise with comparison to the watermark distortion, i.e. $\sigma_{N_2}^2 \gg \sigma_{N_1}^2 = \sigma_D^2$, then independently of $f_{N_2}(x)$, we have

$$P_{e|\sigma_{N_2}^2 \gg \sigma_{N_1}^2} = \sum_m \Pr[m\Delta - \frac{3\Delta}{4} \leq N_1 - D \leq m\Delta - \frac{\Delta}{4}] \tag{33}$$

From the independence between N_1 and D, we can write $f_{N_1-D}(x) = f_{N_1}(x) * f_D(x)$. This situation is illustrated in Fig.9. We can see that integration of $f_{N_1-D}(x)$ over the black intervals will give $P_{e|\sigma_{N_2}^2 \gg \sigma_{N_1}^2} = 0.5$.

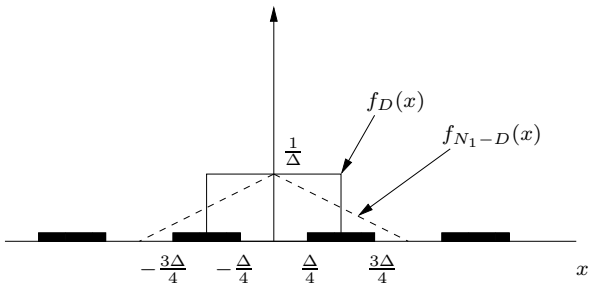

Figure 9. An illustration of $f_D(x)$ (solid line) and $f_{N_1-D}(x)$ (dashed line). The integral of $f_{N_1-D}(x)$ over the black intervals gives $P_{e|\sigma_{N_2}^2 \gg \sigma_{N_1}^2} = 0.5$.

Experimental curves for the probability of error P_e as a function of σ_D^2 for different values of $\sigma_{N_2}^2$ are shown in Fig.10. It can be seen that $P_e \to 0.5$ as $\sigma_D^2 \to \sigma_{N_1}^2$ independently of $\sigma_{N_2}^2$.

5. APPROXIMATION TO THE PDF MODELS

In this section, we will make simplified approximations of (11) obeying the subtractive dither conditions derived in the previous section. We can approximate $f_{X+\frac{1}{\alpha}D}(x)$ in the following way:

$$\begin{aligned} f_{X+\frac{1}{\alpha}D}(x) &= \sum_{k=-\infty}^{+\infty} f_{(1-\alpha)\widetilde{X}+k\Delta+\frac{1-\alpha}{\alpha}D}(x) I_{A_{k|W=0}}(x) \\ &\approx \sum_{k=-\infty}^{+\infty} f_{(1-\alpha)\widetilde{X}+k\Delta}(x) I_{A_{k|W=0}}(x) \end{aligned} \tag{34}$$

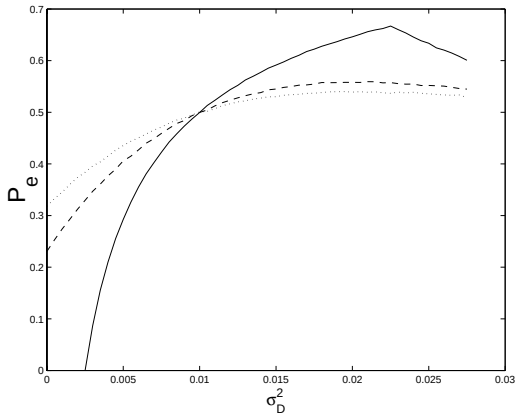

Figure 10. Experimental curves for P_e as a function of σ_D^2 for different values of $\sigma_{N_2}^2$. The solid curve is for $\sigma_{N_2}^2 = 0$, the dashed curve is for $\sigma_{N_2}^2 = 0.01$, and the dotted curve is for $\sigma_{N_2}^2 = 0.02$. Chosen settings are $\widetilde{X} \sim \mathcal{N}(0,1)$, $D \sim \mathcal{U}(0, \sigma_D^2)$, $N_2 \sim \mathcal{N}(0, \sigma_{N_2}^2)$, $\sigma_{N_1}^2 = 0.01$, and $\beta = 1$.

where the approximation follows from the small-distortion case $\sigma_{\widetilde{X}}^2 \gg \sigma_{N_1}^2 = \sigma_D^2$. Note that the output of the quantizer depends both on $\widetilde{X}$ and D, but since the variance of the first is assumed to be much larger, the term $k\Delta$ is present in the approximation together with $\widetilde{X}$. An illustration of $f_{X+\frac{1}{\alpha}D}(x)$, its approximation as given by (34), and $f_X(x)$ is given in Fig. 11. The difference between $f_{X+\frac{1}{\alpha}D}(x)$ and its approximation can hardly be recognized. We can also see the huge difference between $f_{X+\frac{1}{\alpha}D}(x)$ and $f_X(x)$.

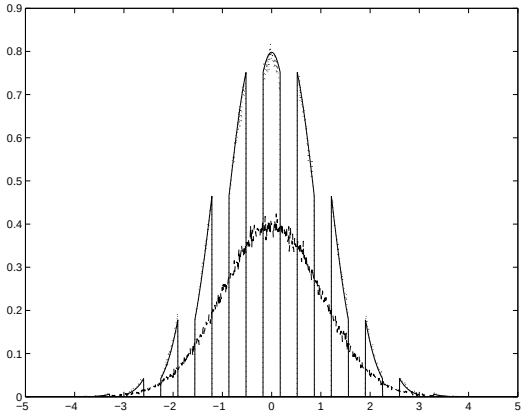

Figure 11. Comparison of $f_X(x)$ (dashed line), $f_{X+\frac{1}{\alpha}D}(x)$ (dotted line), and its approximation $\sum_{k=-\infty}^{+\infty} f_{(1-\alpha)\widetilde{X}+k\Delta}(x) I_{A_{k|W=0}}(x)$ (solid line). The difference between $f_{X+\frac{1}{\alpha}D}(x)$ and its approximation can hardly be recognized. Chosen settings are $\widetilde{X} \sim \mathcal{N}(0,1)$, $\sigma_{N_1}^2 = \sigma_{N_2}^2 = \sigma_D^2 = 0.01$, $\beta = 1$.

6. MAXIMUM LIKELIHOOD ESTIMATION

For the ML estimation approach we will assume that the host signal and attack channel noise are i.i.d. vector sources. We note though that such an assumption may result in a source of substantial loss in the case of real data (audio, video), exhibiting high correlation between the samples. The ML estimation of β is done based on the following relation:

$$f_{Y+\frac{\beta}{\alpha}D}(x) \quad = \quad f_{\beta X+\frac{\beta}{\alpha}D} * f_{N_2}(x) \tag{35}$$

By definition, the ML estimation $\hat{\beta}$ of the parameter β is given as:

$$\begin{aligned}\hat{\beta} &= \arg\max_{\beta} f_{Y_1+\frac{\beta}{\alpha}D_1, Y_2+\frac{\beta}{\alpha}D_2, \ldots, Y_n+\frac{\beta}{\alpha}D_n}(x) \\ &= \arg\max_{\beta} f_{Y_1+\frac{\beta}{\alpha}D_1}(x) f_{Y_2+\frac{\beta}{\alpha}D_2}(x) \ldots f_{Y_n+\frac{\beta}{\alpha}D_n}(x) \\ &= \arg\max_{\beta} \sum_i \log f_{Y_i+\frac{\beta}{\alpha}D_i}(x). \end{aligned} \tag{36}$$

Here the second line follows from the assumption that the received data consists of i.i.d. samples, and therefore the joint PDF can be written as a product of the marginal PDFs. The last line follows from the monotonicity of the logarithm.

The Maximum Likelihood Functional (MLF) is the expression $\sum_i \log f_{Y_i+\frac{\beta}{\alpha}D_i}(x)$. Experimental curves of the MLF for different values of β and $\frac{\sigma^2_{N_2}}{\sigma^2_{N_1}}$ are shown in Fig. 12. Since $f_{Y+\frac{1}{\alpha}D}(x)$ is not differentiable (due to the indicator function) it is difficult to find an analytical expression of $\hat{\beta}$. Therefore, we do a brute force search for the optimal value of β based on (36).

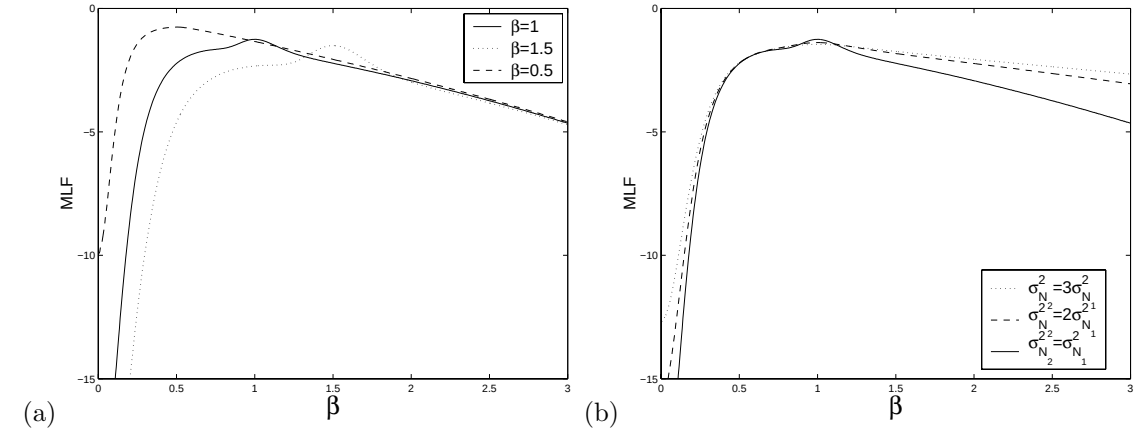

Figure 12. Graph of MLF for different values of $\hat{\beta}$ (a) and different values of $\frac{\sigma^2_{N_2}}{\sigma^2_{N_1}}$ (b). Chosen settings are $\widetilde{X} \sim \mathcal{N}(0,1)$, $D \sim \mathcal{U}(0, 0.01)$, $N_2 \sim \mathcal{N}(0, 0.01)$, and $\sigma^2_{N_1} = 0.01$.

7. EXPERIMENTS

In this section we describe experiments with real audio signals (audio and speech with sampling frequency 48kHz) carried out to test the estimation accuracy of the proposed techniques in terms of the ratio $\frac{\sigma^2_{N_2}}{\sigma^2_{N_1}}$, the parameter β, and the number of available signal samples s. In principle one aims at developing estimation techniques that require a small amount of data, so that they can be applied in situations where the estimating parameter slowly varies.

Experimental results in terms of $\frac{\sigma^2_{N_2}}{\sigma^2_{N_1}}$ and s are shown in Fig. 13. The assumed PDF model of the host signal at the estimator side is a zero-mean Laplacian PDF with variance equal to the variance of the sum of the variances of the host signal, watermark, and the noise in the attack channel, i.e., $\mathcal{L}(0, \sigma^2_{\widetilde{X}} + \sigma^2_{N_1} + \sigma^2_{N_2})$. This is a realistic assumption, because the decoder has access to the received data and can estimate its variance. Furthermore, in practice most audio signals have a PDF that resembles the Laplacian PDF. The loss in performance of the ML approach is due to the approximation in $f_{X+\frac{1}{\alpha}D}(x)$ and the fact that generally, ML estimation requires a large sample size.[15] In Fig. 14 we plot experimental results of $\beta - \hat{\beta}$ as a function of β for different audio signals. The assumed host signal PDF at the estimator side is $\mathcal{L}\big(0, \beta^2(\sigma^2_{\widetilde{X}} + \sigma^2_{N_1}) + \sigma^2_{N_2}\big)$. It can be seen that for small

values of β the estimation accuracy decreases which is due to the mismatch between the host signal PDF and the PDF model assumed at the estimator side. For large values of β there is also a mismatch, but it turns out to be insignificant for the ML estimation procedure.

The ML estimation procedure is computationally very expensive, because of the brute force searching for the optimal β. The paper[16] treats the problem of jointly estimating β and $\sigma^2_{N_2}$ by transforming the attack channel into one that is equivalent but computationally less expensive for the ML approach processing chain. However, this transform does not improve the estimation.

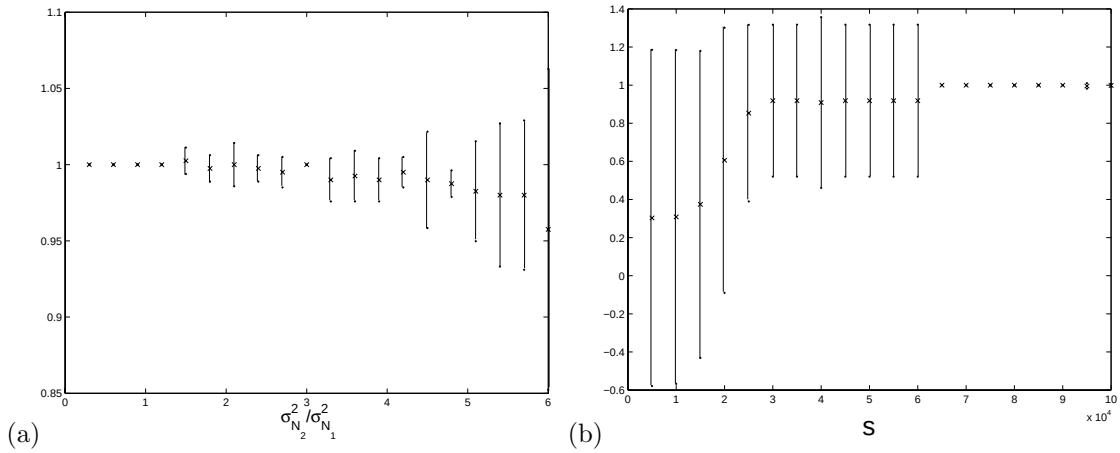

Figure 13. Graphs of $\hat{\beta}$ for real audio signals as a function of $\frac{\sigma^2_{N_2}}{\sigma^2_{N_1}}$ (a) and as a function of available signal samples s (b). The crosses represent the estimation mean, and the lines the estimation standard deviation in both directions. The chosen settings are $10\log\frac{\sigma^2_{\widetilde{X}}}{\sigma^2_{N_1}} = 30db$, and $\sigma^2_{N_2} = \sigma^2_{N_1}$ (b). The assumption for the estimator is $\widetilde{X} \sim \mathcal{L}(0, \sigma^2_{\widetilde{X}} + \sigma^2_{N_1} + \sigma^2_{N_2})$.

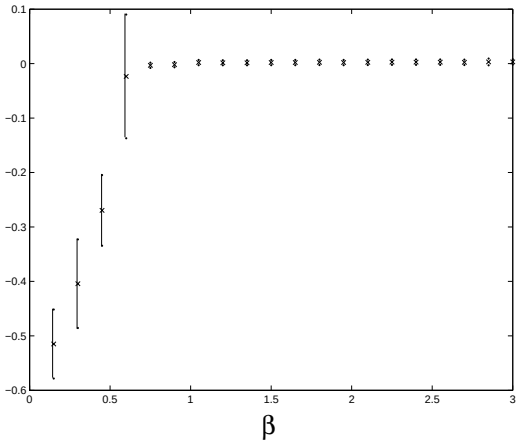

Figure 14. Graphs of $\beta - \hat{\beta}$ for real audio signals as a function of β. The crosses represent the mean, and the lines the standard deviation in both directions. The chosen settings are $10\log\frac{\sigma^2_{\widetilde{X}}}{\sigma^2_{N_1}} = 30db$, $\sigma^2_{N_1} = \sigma^2_{N_2}$. The assumption for the estimator is $\widetilde{X} \sim \mathcal{L}\big(0, \beta^2(\sigma^2_{\widetilde{X}} + \sigma^2_{N_1}) + \sigma^2_{N_2}\big)$.

8. CONCLUSIONS

We have presented a Maximum Likelihood estimation procedure for estimating amplitude scaling factors using subtractive dither in a quantization-based watermarking context. We gave sufficient conditions for the dither sequence such that a given level of security is achieved. The estimation approach performs well in terms of additive noise attacks and for a relatively wide range of values for the parameter β, under realistic assumptions. The disadvantage is the need for a relatively large amount of signal samples for estimating reliably β, which is mainly due to the approximations in incorporating the subtractive dither and to the nature of ML estimation. Another disadvantage is that the method is computationally expensive and currently not suitable for real-time applications.

REFERENCES

1. P. Moulin and A. O'Sullivan. Information-Theoretic Analysis of Information Hiding. *IEEE Transactions on Information Theory*, 49(3):563–593, March 2003.
2. J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups*. Springer-Verlag, 3 edition, 1999.
3. J. H. Conway and N. J. A. Sloane. Fast quantizing and decoding algorithms for lattice quantizers and codes. *IEEE Transactions on Information Theory*, 28(2):227–821, March 1982.
4. J. H. Conway and N. J. A. Sloane. A fast encoding method for lattice codes and quantizers. *IEEE Transactions on Information Theory*, 29(6):820–824, November 1983.
5. I. D. Shterev, R. L. Lagendijk, and R. Heusdens. Statistical Amplitude Scale Estimation for Quantization-based Watermarking. *SPIE Security, Steganography, and Watermarking of Multimedia Contents VI*, 5306, January 2004. CA, USA.
6. B. Chen and G. Wornell. Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding. *IEEE Transactions on Information Theory*, 47:1423–1443, May 2001.
7. S. I. Gel'fand and M. S. Pinsker. Coding for Channel with Random Parameters. *Problems of Control and Information Theory*, 9:19–31, 1980.
8. M. H. Costa. Writing on Dirty Paper. *IEEE Transactions on Information Theory*, 29(3):439–441, May 1983.
9. R. Zamir, S. Shamai (Shitz), and U. Erez. Nested Linear/Lattice Codes for Structured Multiterminal Binning. *IEEE Transactions on Information Theory*, 48(6):1250–1276, June 2002.
10. M. L. Miller, G. J. Doerr, and J. Cox. Dirty-Paper Trellis Codes For Watermarking. *IEEE International Conference On Image Processing*, 2:129–132, September 2002. Rochester, NY.
11. B. Bradley. Improvement to CDF Grounded Lattice Codes. *SPIE Security, Steganography, and Watermarking of Multimedia Contents VI*, 5306, January 2004. CA, USA.
12. J. J. Eggers, R. Bauml, and B. Girod. Estimation of Amplitude Modifications before SCS Watermark Detection. *SPIE Security and Watermarking of Multimedia Contents IV*, 4675:387–398, January 2002. San Jose, CA, USA.
13. K. Chandrasekharan. *Introduction to Analytic Number Theory*. Springer-Verlag, 1968.
14. L. Schuckman. Dither Signals and Their Effect on Quantization Noise. *IEEE Transactions on Communication Technology*, 12(4):162–165, December 1964.
15. H. V. Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, second edition, 1994.
16. R. L. Lagendijk and I. D. Shterev. Estimation of Attacker's Noise and Variance for QIM-DC Watermark Embedding. *IEEE International Conference on Image Processing*, October 2004. Singapore.