Proceedings of SPIE Vol. 4675
Security and Watermarking
of Multimedia Contents IV
San Jose, Ca., Jan. 2002

# Estimation Of Amplitude Modifications before SCS Watermark Detection

Joachim J. Eggers and R. Bäuml
Telecommunications Laboratory
University of Erlangen-Nuremberg
Cauerstrasse 7/NT, 91058 Erlangen, Germany

Bernd Girod
Information Systems Laboratory
Stanford University
Stanford, CA 94305-9510, USA

**ABSTRACT**

New blind digital watermarking schemes that are optimized for additive white Gaussian noise (AWGN) attacks have been developed by several research groups within the last two years. Currently, the most efficient schemes, e.g., the scalar Costa scheme (SCS), involve scalar quantization of the host signal during watermarking embedding and watermark reception. Reliable watermark reception for these schemes is vulnerable to amplitude modification of the attacked host signal. In this paper, a method for the estimation of possible amplitude modifications before SCS watermark detection is proposed. The estimation is based on a securely embedded SCS pilot watermark. We focus on linear amplitude modifications, but investigate also the extension to nonlinear amplitude modifications. Further, the superiority of our proposal over an estimation method based on a spread-spectrum pilot watermark is demonstrated.

**Keywords:** blind digital watermarking, channel estimation

## 1. INTRODUCTION

Blind digital watermarking is the art of communicating a message by embedding it into multimedia data (host signal), and decoding it without access to the original, non-watermarked host signal. Envisioned applications for such a method are copy control or ownership verification. A blind watermarking scheme must be designed such that the watermarked signal has subjective quality close to that of the original host signal and that the decoder can correctly decode the embedded message after any attack that does not destroy the commercial value of the multimedia data.

Early blind watermarking schemes were built on the principle of spread spectrum (SS). Although this technique allows for reliable communication even for strong attacks, blind detection of spread-spectrum watermarks suffers significantly from host signal interference. In 1999, several researchers[1–3] realized that the host signal can be considered as side information at the watermark encoder, and thus improved blind watermarking schemes can be designed. A key paper in this field is the work by Costa,[4] which shows that for additive white Gaussian noise (AWGN) attacks blind watermarking can perform as well as if the decoder had access to the original host signal. We[5–7] developed a simplified practical watermarking scheme based on Costa's ideas, called "scalar Costa scheme" (SCS), which performs over a large range of attack strengths significantly better than blind spread-spectrum watermarking.

So far, the performance of SCS and related schemes has been mainly analyzed for AWGN attacks. However, in practical watermarking applications, the attack is not constrained to AWGN attacks. One particularly interesting class of extended attacks is (non-)linear amplitude modification. This class of attacks includes simple scaling of the watermarked signal, e.g. contrast reduction for image data, or the addition of a constant DC value. A typical example for non-linear amplitude modification is gamma-correction for image data. Blind spread-spectrum watermarking schemes are typically believed to survive such attacks without significant losses. However, quantization based watermarking schemes, like SCS, are vulnerable against such amplitude modifications. The SCS watermark decoder needs to estimate amplitude modifications for reliable watermark detection.

In this paper, we present a scheme for estimating linear amplitude modifications and simple parametrized non-linear amplitude modifications, e.g. gamma-correction, based on a securely embedded pilot sequence. Note that in watermarking application the secure embedding of pilot sequences is essential, since, otherwise, an attacker could simply focus on removing the embedded pilot sequence. Thus, we propose to embed a pilot sequence via secure SCS watermarking. The pilot sequence is known to the watermark receiver and thus can be exploited to estimate any amplitude modifications. In particular, we propose an estimation algorithm based on a Fourier analysis of the histograms of different parts of the received pilot samples.

SCS watermarking is briefly reviewed in Sec. 2 and the influence of amplitude modifications is highlighted. Our new algorithm for estimating linear amplitude modifications is derived and investigated in Sec. 3. In Sec. 4, we demonstrate the superiority of our new approach over an estimation based on SS pilot sequences. The extension of this work to simple non-linear amplitude modifications is outlined in Sec. 5.

Further author information: Send correspondence to J. Eggers. Email: eggers@LNT.de

## 2. SCS WATERMARKING AND AWGN AND AMPLITUDE SCALING ATTACK

We consider digital watermarking as a communication problem. Here, we assume that the watermark message is encoded into a sequence of watermark letters $\mathbf{d}$ of length $L_x$. The elements $d_n$ belong to a $D$-ary alphabet $\mathcal{D} = \{0, 1, \ldots, D-1\}$ of size $D = |\mathcal{D}|$. In many practical cases, binary watermark letters ($d_n \in \mathcal{D} = \{0, 1\}$) are used. The watermark encoder derives from the encoded watermark message $\mathbf{d}$ and the host data $\mathbf{x}$ an appropriate watermark sequence $\mathbf{w}$, which is added to the host data to produce the watermarked data $\mathbf{s}$. $\mathbf{w}$ must be chosen such that the distortion between $\mathbf{x}$ and $\mathbf{s}$ is negligible. Next, an attacker might modify the watermarked data $\mathbf{s}$ into data $\mathbf{r}$ to impair watermark communication. The attack is only constrained with respect to the distortion between $\mathbf{x}$ and $\mathbf{r}$. Finally, the decoder must be able to detect the watermark message from the received data $\mathbf{r}$. In *blind* watermarking schemes, the host data $\mathbf{x}$ are not available to the decoder but can be considered side information to the encoder. The codebook used by the watermark encoder and decoder is randomized dependent on a key $\mathbf{k}$ to achieve secrecy of watermark communication. Here, $\mathbf{x}$,$\mathbf{w}$,$\mathbf{s}$,$\mathbf{r}$, and $\mathbf{k}$ are vectors of identical length $L_x$, and $x_n$,$w_n$,$s_n$,$r_n$, and $k_n$ refer to their respective $n$th elements. Random variables are in Sans Serif fonts, e.g., $\mathsf{x}$ for a random variable describing the host signal.

Fig. 1 depicts a block diagram of blind watermark communication, where the attacker scales the watermarked data $\mathbf{s}$ by $g$ (usually $g < 1$) and introduces additive white Gaussian noise (AWGN) $\mathbf{v}$, with $\mathsf{v} \sim \mathcal{N}(r_{\text{offset}}, \sigma_v^2)$, that is

$$\mathbf{r} = g\mathbf{s} + \mathbf{v} = g(\mathbf{x} + \mathbf{w}) + \mathbf{v}. \tag{1}$$

Ideally, the receiver knows $g$ and $r_{\text{offset}}$ and thus compensates for the DC offset by subtracting $r_{\text{offset}}$ and compensates for scaling by division by $g$ (if $g \neq 0$). In this paper, we characterize the attack strength by the effective watermark-to-noise power ratio $\text{WNR} = 10 \log_{10}(g^2 \sigma_w^2 / \sigma_v^2)\,\text{dB}$.

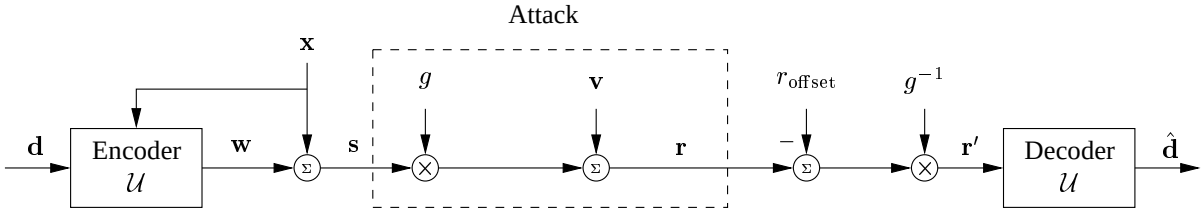


**Figure 1.** Watermark communication facing an attack by amplitude scaling and AWGN with mean $r_{\text{offset}}$.

For the communication scenario depicted in Fig. 1, Costa[4] showed theoretically that for Gaussian host data the watermark capacity is independent of the host data variance $\sigma_x^2$. The result is surprising since it shows that the host data $\mathbf{x}$ need not be considered as interference at the decoder although the decoder does not know $\mathbf{x}$. Costa's scheme involves a random codebook $\mathcal{U}$ that is available at the encoder and decoder. Unfortunately, for good performance $\mathcal{U}$ must be so large that neither storing it nor searching it is practical. Thus, several research groups proposed suboptimal but practical versions of Costa's scheme that are based on dithered uniform scalar quantization[1,2,5,8] where the $L_x$-dimensional codebook $\mathcal{U}$ is constructed by a concatenation of one-dimensional (scalar) codebooks. The derivation and realization of these schemes differ only slightly. We derived the scalar Costa scheme (SCS) and presented a detailed capacity analysis[5] and experimental results[7] for SCS watermarking. SCS watermarking will be considered throughout this paper, although any of the other proposals based on dithered scalar quantization could be used with minor modifications as well.

In SCS, each of the watermark letters $d_n$ is embedded into the corresponding host elements $x_n$. The encrypted scalar component codebook used in SCS is given by

$$\mathcal{U}_n^1(k_n) = \left\{ u_n = \left( l_n + \frac{d_n}{D} + k_n \right) \alpha\Delta \,\middle|\, d_n \in \mathcal{D}, l_n \in \mathbb{Z} \right\}, \tag{2}$$

where $\alpha$ and $\Delta$ are codebook parameters that are discussed below. $\mathcal{U}_n^1(k_n)$ can be described by the reconstruction points of $D$ scalar uniform quantizers which are shifted against each other by $d_n/D$. The given watermark letter $d_n$ selects one of these quantizers. The SCS embedding rule for the $n$th element is given by

$$s_n = x_n + \alpha \left( \mathcal{Q}_\Delta \left\{ x_n - \Delta \left( \frac{d_n}{D} + k_n \right) \right\} + \Delta \left( \frac{d_n}{D} + k_n \right) - x_n \right), \tag{3}$$

where $\mathcal{Q}_\Delta\{\cdot\}$ denotes scalar uniform quantization with step size $\Delta$. The key $\mathbf{k}$ is a pseudo-random sequence with $k_n \in [0,1)$. The SCS embedding scheme depends on two parameters: the quantizer step size $\Delta$ and the scale factor $\alpha$. Both parameters can be jointly optimized to achieve a good trade-off between embedding distortion $\sigma_w^2$ and detection reliability for a given noise variance $\sigma_v^2$ of an AWGN attack. Optimal values for $\Delta$ and $\alpha$ must be computed numerically.[5]

At the receiver, after compensation for $g$ and $r_{\text{offset}}$, the extraction rule for the $n$th element is

$$y_n = \mathcal{Q}_\Delta\{r'_n - k_n\Delta\} + k_n\Delta - r'_n. \tag{4}$$

For binary SCS, $|y_n| \leq \Delta/2$, where $y_n$ should be close to zero if $d_n = 0$ was sent, and close to $\pm\Delta/2$ for $d_n = 1$. If no compensation for $g$ and $r_{\text{offset}}$ is applied, the proper codebook for SCS watermark reception is

$$\hat{\mathcal{U}}_n^1(k_n) = \left\{ u_n = \left(l_n + \frac{d_n}{D} + k_n\right)\alpha\Delta_r + r_{\text{offset}} \,\middle|\, d_n \in \mathcal{D}, l_n \in \mathbb{Z} \right\}. \tag{5}$$

Here, $\Delta_r = g\Delta$ is the scaled quantizer step size which has to be used for SCS detection.

Fig. 2 illustrates the effect of the considered amplitude scaling and AWGN attack on the PDF of the received data elements $r_n$ for binary SCS watermarking. For better clarity, we assume a flat distribution of the host signal elements $x_n$ over a number of quantizer step sizes $\Delta$. The upper plot of Fig. 2 depicts several periods of the PDF of the watermarked elements $s_n$ conditioned on the transmitted watermark letter $d_n$, and $k_n = 0$. The lower plot shows the respective PDFs of the extracted received elements $y_n$ conditioned on the transmitted watermark letter $d_n$, where the attack is amplitude scaling by $g > 1$ and AWGN with nonzero mean $r_{\text{offset}}$. The large crosses and circles in both plots indicate the codebook entries of $\mathcal{U}_n^1(0)$ for $d_n = 0$ and $d_n = 1$. We observe that $\mathcal{U}_n^1(0)$ is no longer appropriate for SCS watermark reception if $g \neq 1$ and $r_{\text{offset}} \neq 0$.

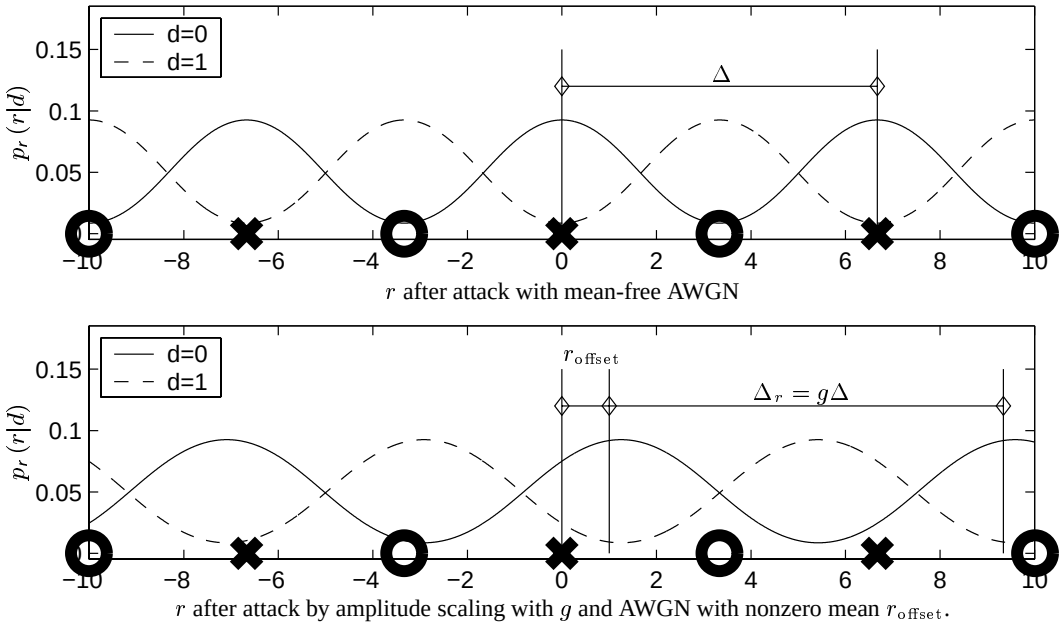


**Figure 2.** PDFs for SCS watermarked data before and after amplitude scaling and AWGN attack

## 3. ESTIMATION BASED ON SCS PILOT SEQUENCES

In the previous section, it has been assumed that the watermark receiver has perfect knowledge of the scale factor $g$ and a possible DC offset $r_{\text{offset}}$ in the added noise $\mathbf{v}$. Here, we propose a technique for estimating the attack channel parameters $g$ and $r_{\text{offset}}$ with the aid of a securely embedded pilot sequence $\mathbf{d}_{\text{pilot}} = \mathbf{0}$ of length $L_{\text{pilot}}$. Note that estimation of $\Delta_r = g\Delta$ is sufficient to enable SCS watermark reception. $g$ can be derived when $\Delta$ is known to the receiver.

The key idea behind our method for the estimation of $\Delta_r$ and $r_{\text{offset}}$ is to analyze the histograms of the received samples $r_{n,\text{pilot}}$, where $\mathbf{r}_{\text{pilot}} = (r_{0,\text{pilot}} \dots r_{n,\text{pilot}} \dots r_{L_{\text{pilot}}-1,\text{pilot}})$ is the sequence of received samples with embedded pilot symbols $\mathbf{d}_{\text{pilot}} = \mathbf{0}$. The suffix "pilot" is suppressed subsequently since only pilot samples are considered here. Note that for security, the pilot is embedded dependent on a secure random key sequence $\mathbf{k}$. Thus, without knowing $\mathbf{k}$, no structure in the watermarked signal is visible. However, the key-dependent embedding of the pilot sequence is also a problem for the estimation of $\Delta_r$ and $r_{\text{offset}}$. For SCS embedding, the key sequence is scaled with the embedding quantizer step size $\Delta_r$, but the proper quantizer step size $\Delta_r$ for reception still has to be found. Therefore, instead of exploiting the key sequence $\mathbf{k}$ directly, the histograms of the samples $r_n$ with key $k_n \in \mathbb{K}_m$ are analyzed separately, where

$$\mathbb{K}_m = \left\{ k \,\middle|\, \frac{m}{M} \leq k < \frac{m+1}{M} \right\} \quad \text{for } m \in \{0, 1, \dots, M-1\} \text{ and } M > 1. \tag{6}$$

Here, $M$ denotes the number of different ranges considered for the key values. The $M$ conditional histograms will show local maxima with a relative distance of $\Delta_r$. The absolute position of these maxima gives an estimate of $r_{\text{offset}}$.

### 3.1. Model for the Conditional PDFs of Received Pilot Elements

Let $p_r(r)$ denote the PDF of the received signal samples $r_n$. Here, IID signals are considered so that the sample index $n$ can be neglected in the statistical analysis. It can be assumed that $p_r(r)$ reflects more or less the host signal PDF ($p_r(r) \approx p_x(x)$) if the embedding distortion is small, the host signal PDF $p_x(x)$ is sufficiently smooth, and a key with $k_n \in [0, 1)$ is chosen.

First, a model for the conditional PDF $p_r(r|k \in \mathbb{K}_m)$ of the received signal $r_n$ for which $k_n \in \mathbb{K}_m$ is provided. The model is motivated by the observation that each PDF $p_r(r|k \in \mathbb{K}_m)$ shows local maxima with a distance of $\Delta_r$ and that $p_r(r|k \in \mathbb{K}_m)$ is a valid conditional PDF. An exact characterization of $p_r(r|k \in \mathbb{K}_m)$ is not necessary for our purpose. A sufficiently accurate model is given by

$$p_r(r|k \in \mathbb{K}_m) = p_r(r) \left( 1 + \gamma \cos\left( 2\pi f_0 r - \Phi_0 - \frac{2\pi}{M}\left( m + \frac{1}{2} \right) \right) \right), \tag{7}$$

where $\gamma$ is an appropriate constant with $0 < \gamma < 1$. The model parameters $f_0$ and $\Phi_0$ are directly related to the unknown parameters $\Delta_r$ and $r_{\text{offset}}$. $f_0$ determines the distance between two local maxima, and $\Phi_0$ determines their absolute position. The exact relationship is given by

$$f_0 = \frac{1}{\Delta_r} \quad \text{and} \quad \Phi_0 = \frac{2\pi}{\Delta_r} r_{\text{offset}} = 2\pi f_0 r_{\text{offset}}. \tag{8}$$

Fig. 3 depicts an example for the given model. The local maxima of the conditional PDFs $p_r(r|k \in \mathbb{K}_m)$ with a relative distance of $\Delta_r = 10$ are clearly visible. Further, it can be verified that the given model for the conditional PDFs fulfills the property

$$\sum_{m=0}^{M-1} p_r(r|k \in \mathbb{K}_m)\, p(k \in \mathbb{K}_m) = p_r(r). \tag{9}$$

### 3.2. Parameter Estimation Based on Fourier Analysis

The parameters $f_0$ and $\Phi_0$ of the model given in (7) have to be computed from the given conditional PDFs $p_r(r|k \in \mathbb{K}_m)$ and the given unconditional PDF $p_r(r)$. Fourier analysis is appropriate for this task since $f_0$ and $\Phi_0$ are the frequency and a constant phase contribution of the cosine term in (7).

For the $m$th conditional PDF, the normalized spectrum $A_m(f)$ is defined as

$$\begin{aligned} A_m(f) &= \mathcal{F}\left\{ \frac{p_r(r|k \in \mathbb{K}_m)}{p_r(r)} - 1 \right\} = \mathcal{F}\left\{ \frac{p_r(r|k \in \mathbb{K}_m) - p_r(r)}{p_r(r)} \right\} = \mathcal{F}\left\{ \gamma \cos\left( 2\pi f_0 r - \Phi_0 - \frac{2\pi}{M}\left( m + \frac{1}{2} \right) \right) \right\} \\ &= \frac{\gamma}{2} \left[ \mathrm{e}^{j\left(-\Phi_0 - \frac{2\pi}{M}\left(m+\frac{1}{2}\right)\right)} \delta(f_0 - f) + \mathrm{e}^{-j\left(-\Phi_0 - \frac{2\pi}{M}\left(m+\frac{1}{2}\right)\right)} \delta(f_0 + f) \right]. \end{aligned} \tag{10}$$

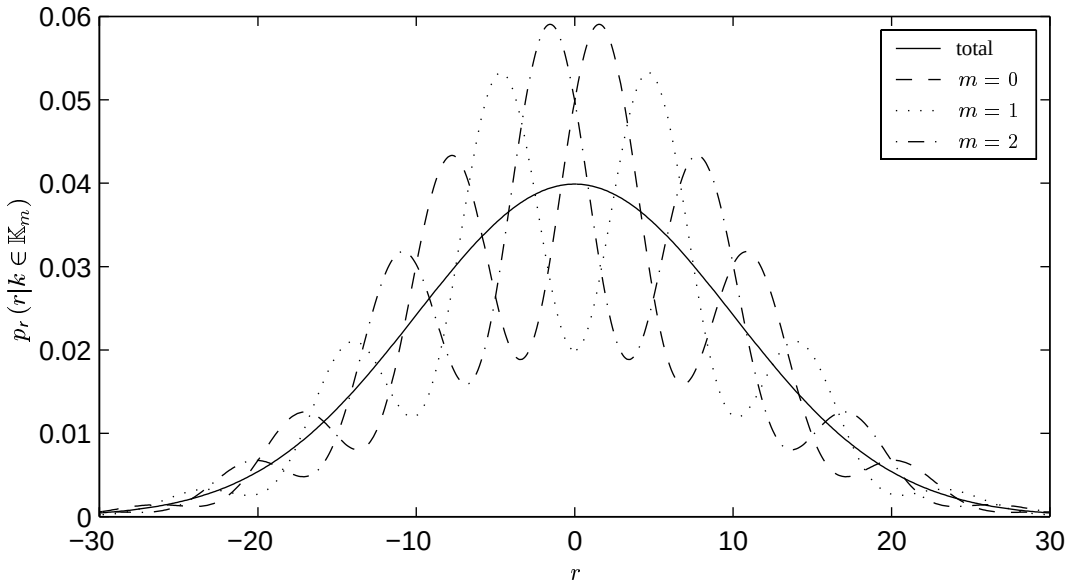


**Figure 3.** Total and conditional PDFs of the received pilot sequence. $M = 3$ different ranges for the key are distinguished. The example is for a Gaussian distribution of $r_n$, and for the parameters $\Delta_r = 10$ and $r_{\text{offset}} = 0$.

All $M$ spectra can be combined in an elegant way due to the systematically different phase at $A_m(f_0)$ and $A_m(-f_0)$. The $M$ spectra $A_m(f)$ are multiplied by $\mathrm{e}^{-j\frac{2\pi}{M}m}$ prior to their summation to an overall spectrum $A(f)$, that is

$$\begin{aligned} A(f) &= \sum_{m=0}^{M-1} A_m(f)\,\mathrm{e}^{j\frac{2\pi}{M}m} = \frac{\gamma}{2}\,\mathrm{e}^{-j\left(\Phi_0+\frac{\pi}{M}\right)}\delta\left(f_0-f\right)\underbrace{\sum_{m=0}^{M-1} 1}_{=M} + \frac{\gamma}{2}\,\mathrm{e}^{j\left(\Phi_0+\frac{\pi}{M}\right)}\delta\left(f_0+f\right)\underbrace{\sum_{m=0}^{M-1}\mathrm{e}^{j\frac{4\pi}{M}m}}_{=0} \\ &= \frac{\gamma M}{2}\,\mathrm{e}^{-j\left(\Phi_0+\frac{\pi}{M}\right)}\delta\left(f_0-f\right). \end{aligned} \tag{11}$$

Thus, for the model given in (7), $|A(f)|$ has only one peak, which is located exactly at the frequency $f_0$. Further, $\Phi_0 = -\arg\{A(f_0)\} - \frac{\pi}{M}$. Note that the multiplication by $\mathrm{e}^{-j\frac{2\pi}{M}m}$ is superior to a multiplication by $\mathrm{e}^{-j\frac{2\pi}{M}mf}$ which would correspond to a shift of the different conditional PDFs by $\frac{\Delta_r}{M}$. In the latter case, the spectrum $|A(f)|$ would have another peak at $f = -f_0$ which increases the required sampling interval for the numerical computation of the conditional PDFs.

### 3.3. Implementation Based on Histograms of Received Pilot Elements

The exact PDFs of the received signal do not fit exactly to the model given in (7). Further, in practice, the PDFs $p_r\left(r|k \in \mathbb{K}_m\right)$ and $p_r\left(r\right)$ can be only estimated from the $L_{\text{pilot}}$ pilot samples $\mathbf{r}$. This estimation is obtained from histograms with $L_{\text{bin}}$ bins that cover the total range of all received samples. Note that removal of outliers is useful in practice. Based on these histograms, $A_m(f)$ is computed at $L_{\text{DFT}} \geq L_{\text{bin}}$ discrete frequencies via a length-$L_{\text{DFT}}$ DFT. Here, a single peak in the spectrum $A(f)$ cannot be exptected due to estimation errors and the inaccuracy of the model (7). Nevertheless, for $L_{\text{pilot}}$ sufficiently large, a dominating peak should occur at $f_0$. Details of the outlined implementation are briefly described below.

First, the histogram $\hat{p}_r\left[\kappa\right]$ of all received pilot symbols $\mathbf{r}_{\text{pilot}}$ is computed, where $\kappa \in \{0, 1, \ldots, L_{\text{bin}} - 1\}$ is the bin index. The width $\varsigma$ of the histogram bins is computed by

$$\varsigma = \frac{r_{\max} - r_{\min}}{L_{\text{bin}}}, \tag{12}$$

where

$$r_{\min} = \min_{n\in\{0,1,\ldots,L_{\text{pilot}}-1\}} r_{n,\text{pilot}}, \tag{13}$$

$$r_{\max} = \max_{n\in\{0,1,\ldots,L_{\text{pilot}}-1\}} r_{n,\text{pilot}}. \tag{14}$$

Thus, the $\kappa$th bin covers the range $[r_{\min} + \kappa\varsigma, r_{\min} + (\kappa+1)\varsigma)$.

Next, the conditional histograms $\hat{p}_{r,m}[\kappa]$ are computed, where the index $m$ indicates the considered range of key values $\mathbb{K}_m$. The bins for these conditional histograms are identical to those used for the computation of $\hat{p}_r[\kappa]$.

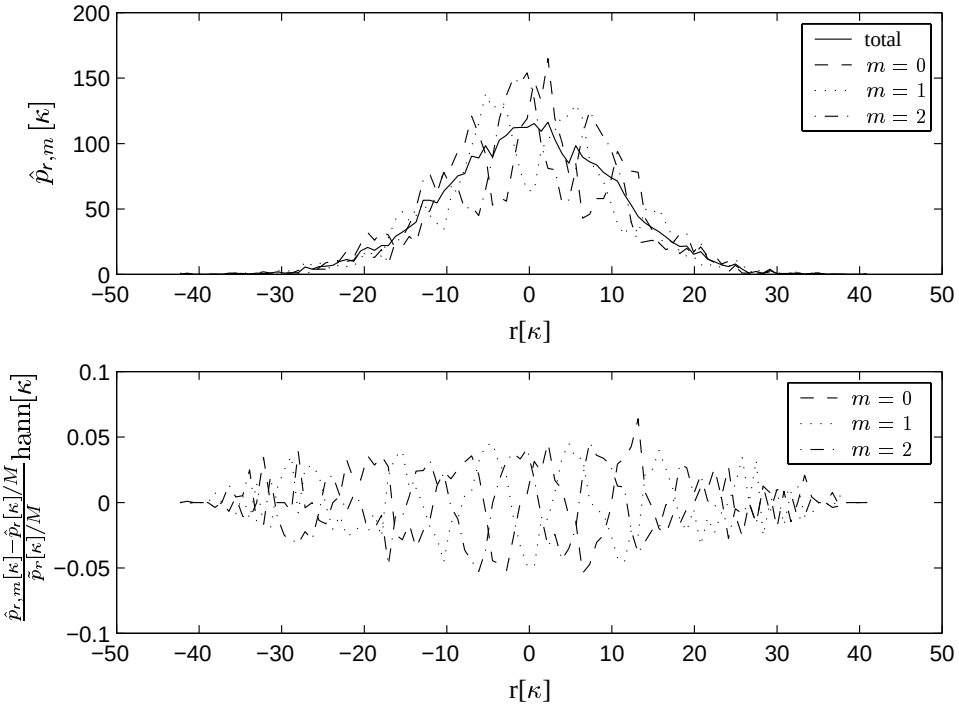


**Figure 4.** The upper plots shows the scaled total histogram $\hat{p}_r[\kappa]/M$ and the conditional histograms $\hat{p}_{r,m}[\kappa]$ of the received pilot sequence for $M = 3$. The lower plot shows the normalized and windowed conditional PDFs that are input to the DFT analysis.

Normalization of the conditional histograms $\hat{p}_{r,m}[\kappa]$ by division by the total histogram $\hat{p}_r[\kappa]$ is critical since empty bins in the total histogram can occur. In such a case, the corresponding bins of all conditional histograms are also empty. Thus, no useful information can be obtained from such a bin. Therefore, the modified total histogram

$$\tilde{p}_r[\kappa] = \begin{cases} 1 & \text{if } \hat{p}_r[\kappa] = 0, \\ \hat{p}_r[\kappa] & \text{else}, \end{cases} \tag{15}$$

is defined, which will be used for normalizing the conditional histograms. Further, bins that are not empty but contain only a few samples provide little useful information as well. For Gaussian distributed $r_n$, but also for many other typical signal distributions, empty and almost empty bins occur mainly at the tails of the histograms. Therefore, it is useful to weight the normalized histograms, e.g., with a von Hann window $\mathrm{hann}[\kappa] = 0.5\left(1 - \cos(\frac{2\pi(\kappa)}{L_{\mathrm{bin}}-1})\right)$. Note that an improved window might be available if a priori information about the distribution of $r_n$ exists. Fig. 4 depicts example histograms before and after normalization. For better illustration, a long pilot sequence $L_{\mathrm{pilot}} = 10000$ has been used.

Next, in analogy to (10), the discrete spectra

$$A_m[l] = \mathrm{DFT}_{L_{\mathrm{DFT}}}\left\{\frac{\hat{p}_{r,m}[\kappa] - \hat{p}_r[\kappa]/M}{\tilde{p}_r[\kappa]/M}\mathrm{hann}[\kappa]\right\} \tag{16}$$

are computed. Note that the histograms can be considered discretized PDFs, where the sampling frequency is $f_A = \frac{1}{\varsigma}$. The spectra of such discrete sequences are periodic and usually parameterized by the normalized frequency $\Omega = 2\pi\frac{f}{f_A} = 2\pi\varsigma f$. Sampling these periodic spectra at $L_{\mathrm{DFT}}$ equidistant frequencies $\Omega_l = \frac{2\pi}{L_{\mathrm{DFT}}}l$, for $l \in \{0, \ldots, L_{\mathrm{DFT}} - 1\}$, defines the discrete spectra $A_m[l]$.

Finally, all $M$ spectra $A_m[l]$ are combined to obtain $A[l]$ corresponding to (11), that is

$$A[l] = \sum_{m=0}^{M-1} A_m[l]\, \mathrm{e}^{j\frac{2\pi}{M}m}. \tag{17}$$

From $A[l]$, the frequency index $l_0$ with maximum $|A[l]|$ is determined, thus

$$l_0 = \arg \max_{l \in \{0,\ldots,L_{\mathrm{DFT}}-1\}} |A[l]|, \tag{18}$$

and the desired decoder parameters $\Delta_r$ and $r_{\mathrm{offset}}$ can be estimated by

$$\hat{\Delta}_r = \frac{\varsigma L_{\mathrm{DFT}}}{l_0}, \tag{19}$$

$$\hat{\Phi}_0 = -\arg\{A[l_0]\} - \frac{\pi}{M}, \tag{20}$$

$$\hat{r}_{\mathrm{offset}} = \mathrm{rem}(r_{\min} + \frac{\varsigma}{2} + \hat{\Phi}_0 \frac{\hat{\Delta}_r}{2\pi}, \hat{\Delta}_r). \tag{21}$$

Here, $\mathrm{rem}(a, b)$ denotes the remainder of the division $a/b$. Note that the offset $\hat{r}_{\mathrm{offset}}$ has to be computed relative to the center of the first bin.

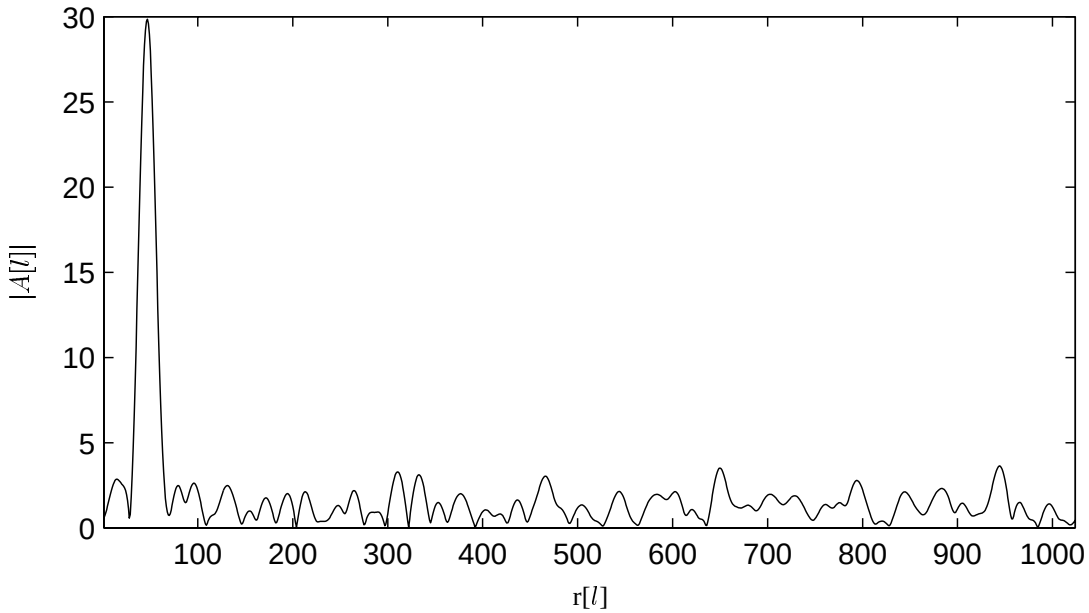


**Figure 5.** DFT spectrum $|A[l]|$ obtained from the normalized histograms shown in Fig. 4

Fig. 5 shows $|A[l]|$ for the normalized histograms in Fig. 4, where $L_{\mathrm{DFT}} = 1024$. The dominating peak at $l_0 = \frac{\varsigma L_{\mathrm{DFT}}}{\hat{\Delta}_r}$ is clearly visible. Note that for shorter pilot sequences, larger DFT components have to be expected for all indices $l \neq l_0$.

### 3.4. Estimation performance for different $L_{\mathrm{pilot}}$

The described algorithm for the estimation of $\Delta_r$ and $r_{\mathrm{offset}}$ is dependent on the following set of parameters:

| | | |
|---|---|---|
| $L_{\mathrm{pilot}}$ | : | length of pilot sequence |
| $L_{\mathrm{bin}}$ | : | number of bins used for the histograms |
| $M$ | : | number of different intervals for the key |
| $L_{\mathrm{DFT}}$ | : | DFT length |

The estimation accuracy also depends on the document-to-watermark power ratio ($\mathrm{DWR} = 10\log_{10}\sigma_x^2/\sigma_w^2\,\mathrm{dB}$), and on the watermark-to-noise power ratio (WNR). In this paper, the estimation performance for different pilot length $L_{\mathrm{pilot}}$ is discussed for $\mathrm{WNR} = -10\ \mathrm{dB}, \ldots, 5\ \mathrm{dB}$. This range for WNR covers the most interesting range of attack strengths for that SCS watermarking might be useful. The DWR has been fixed to $\mathrm{DWR} = 20\ \mathrm{dB}$ and the remaining parameters are $L_{\mathrm{bin}} = 50$, $L_{\mathrm{DFT}} = 1024$, and $M = 5$. Experimental results that support this choice of parameters are given in.[9]

The influence of the number $L_{\mathrm{pilot}}$ of received pilot elements is studied experimentally. For simplicity, $g = 1$ and no offset has been considered so that the estimator should ideally find $\hat{\Delta}_r = \Delta$ and $\hat{r}_{\mathrm{offset}} = 0$. For the evaluation of the estimation performance, three different figures of merit have been used:

$$\text{relative error of } \hat{\Delta}_r: \quad \delta_{\Delta_r} = \frac{\sqrt{\mathrm{E}\left\{(\hat{\Delta}_r - \Delta)^2\right\}}}{\Delta}, \tag{22}$$

$$\text{relative error of } r_{\mathrm{offset}}: \quad \delta_{r_{\mathrm{offset}}} = \frac{\sqrt{\mathrm{E}\left\{(\hat{r}_{\mathrm{offset}} - 0)^2\right\}}}{\Delta}, \tag{23}$$

$$\text{relative increase of bit-error probability}: \quad \delta_{p_{\mathrm{b}}} = \frac{\mathrm{E}\left\{\hat{p}_{\mathrm{b}} - p_{\mathrm{b}}\right\}}{p_{\mathrm{b}}}. \tag{24}$$

$\delta_{\Delta_r}$ and $\delta_{r_{\mathrm{offset}}}$ effectively measure the root of the mean squared estimation error relative to the exact step size $\Delta$. These figures of merit have been chosen since not only the variance of estimation errors is important, but also a possible biased estimate. The relative increase of the bit-error probability $p_{\mathrm{b}}$ for uncoded binary SCS reception with estimated $\Delta_r$ and $r_{\mathrm{offset}}$ is given by $\delta_{p_{\mathrm{b}}}$. It is sufficient to measure the expected difference of the bit-error probability since imperfect estimates $\Delta_r$ and $r_{\mathrm{offset}}$ can only increase the bit-error probability on average. $p_{\mathrm{b}}$ of uncoded binary SCS is relatively high for the considered WNRs. However, many new parameters would have to be introduced for simulations with coded SCS communication, which would make a fair comparison more difficult. Further, the increase of $p_{\mathrm{b}}$ can be considered a good indicator for the effect of estimation errors on coded communication. The free parameters can be optimized only for a certain range of different WNRs where here the focus is on $\mathrm{WNR} = -5\ \mathrm{dB}$ to $\mathrm{WNR} = 0\ \mathrm{dB}$. In particular the relative increase of the uncoded error probability ($\delta_{p_{\mathrm{b}}}$) shows a local minimum for a certain WNR, since for large negative WNRs, the estimation accuracy is decreased due to the strong noise, and for high WNRs, the absolute decoding error is so low that any decoding error increases the relative decoding error significantly.

In general, it is desired to make the pilot sequence as short as possible, however, very short pilot sequences lead to an inaccurate PDF estimation, and thus to incorrect estimations of $\Delta_r$ and $r_{\mathrm{offset}}$. Fig. 6 shows the estimation performance for $L_{\mathrm{pilot}} = 250, 500, 1000$, and 2000. Fig. 6.(a) depicts $\delta_{\Delta_r}$ which describes the relative estimation error of $\Delta_r$. For $L_{\mathrm{pilot}} = 2000$, $\delta_{\Delta_r}$ decreases monotonically with increasing WNR, and is lower than 1% for $\mathrm{WNR} > -3\ \mathrm{dB}$. Shorter pilot sequences lead to an increased relative estimation error. However, for some WNR, robust estimation is no longer possible at all. Lowering the WNR further introduces so much noise into the PDF estimation that the largest component of the spectrum $|A[l]|$ appears at any random index $0 < l < L_{\mathrm{DFT}} - 1 = 1023$. For $L_{\mathrm{pilot}} = 250$, this effect occurs for $\mathrm{WNR} < -1\ \mathrm{dB}$. For $L_{\mathrm{pilot}} = 500$, a minimum WNR of about -5 dB is required. Fig. 6.(b) depicts $\delta_{r_{\mathrm{offset}}}$ which follows in general the behavior of $\delta_{\Delta_r}$. The resulting relative increase of the uncoded error rate $\delta_{p_{\mathrm{b}}}$ is shown with linear and logarithmic axes in Fig. 6.(c) and Fig. 6.(d), respectively. $\delta_{p_{\mathrm{b}}}$ increases monotonically with decreasing pilot length $L_{\mathrm{pilot}}$. Further, it can be observed again that for some low WNR the estimation algorithm starts to fail completely. Nevertheless, it is quite promising that even for $L_{\mathrm{pilot}} = 500$, $\delta_{p_{\mathrm{b}}}$ is lower than 2% for all $\mathrm{WNR} \geq -5\ \mathrm{dB}$.

## 4. ESTIMATION BASED ON SS PILOT SEQUENCES

In the previous section, an estimation of the SCS receiver parameter $\Delta_r$ based on a known SCS watermark has been proposed. However, it is also possible to estimate the scale factor $g$, and thus $\Delta_r = g\Delta$, with help of an additive spread-spectrum (SS) pilot watermark. Here, we present an analysis of the estimation accuracy $\delta_{\Delta_r}$, as defined in Sec. 3.4, when using SS pilot watermarks and compare the result with those for SCS pilot watermarks.

We consider again the attack channel defined in (1). However, now, we assume that $\mathbf{w}$ is a pseudo-noise sequence of length $L_w = L_{\mathrm{pilot}}$ with zero mean ($\sum_{n=1}^{L_{\mathrm{pilot}}} w_n = 0$) and power $\sigma_w^2 = \frac{1}{L_{\mathrm{pilot}}}\sum_{n=1}^{L_{\mathrm{pilot}}} w_n^2$. Throughout this analysis, an IID host signal $\mathbf{w}$ and additive noise signal $\mathbf{v}$ is assumed so that $x_n = x$ and $v_n = v$, respectively. $\mathbf{w}$ is known to the watermark receiver so that $g$ can be estimated from $\mathbf{r}$ based on the correlation $\hat{c}$ between $\mathbf{r}$ and $\mathbf{w}$, that is

$$\hat{c} = \frac{1}{L_{\mathrm{pilot}}}\sum_{n=1}^{L_{\mathrm{pilot}}} r_n w_n. \tag{25}$$

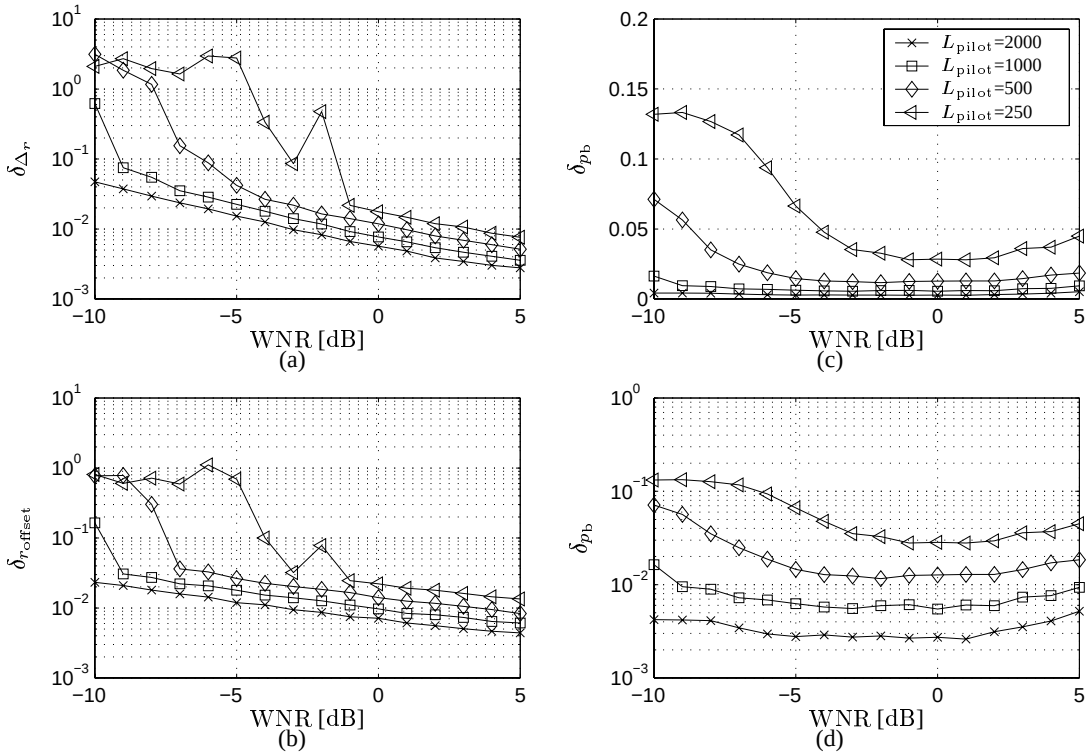


**Figure 6.** Estimation performance for different pilot lengths $L_{\text{pilot}}$ ($\text{DWR} = 20\,\text{dB}$, $L_{\text{bin}} = 50$, $M = 5$).

The unbiased estimate $\hat{g}$ of $g$ derived from $\hat{c}$ is derived as follows:

$$\mathrm{E}\{r_n\} = \mathrm{E}\{g(x_n + w_n) + v_n\} = g\mathrm{E}\{x\} + \mathrm{E}\{v\} + gw_n \tag{26}$$

$$\begin{aligned}\mathrm{E}\{\hat{c}\} &= \frac{1}{L_{\text{pilot}}}\sum_{n=1}^{L_{\text{pilot}}}\mathrm{E}\{r_n\}w_n = \frac{1}{L_{\text{pilot}}}\sum_{n=1}^{L_{\text{pilot}}}(g\mathrm{E}\{x\} + \mathrm{E}\{v\})w_n + \frac{g}{L_{\text{pilot}}}\sum_{n=1}^{L_{\text{pilot}}}w_n^2 \\ &= \frac{g\mathrm{E}\{x\} + \mathrm{E}\{v\}}{L_{\text{pilot}}}\underbrace{\sum_{n=1}^{L_{\text{pilot}}}w_n}_{=0} + \frac{g}{L_{\text{pilot}}}\sum_{n=1}^{L_{\text{pilot}}}w_n^2 = \frac{g}{L_{\text{pilot}}}\sum_{n=1}^{L_{\text{pilot}}}w_n^2 = g\sigma_w^2\end{aligned} \tag{27}$$

$$g = \frac{\mathrm{E}\{\hat{c}\}}{\sigma_w^2} \tag{28}$$

$$\hat{g} = \frac{\hat{c}}{\sigma_w^2} \tag{29}$$

(29) describes the estimation rule for $\hat{g}$ using the SS pilot watermark $\mathbf{w}$ that is known to the receiver. Next, the variance of $\hat{g}$ dependent on the pilot length $L_{\text{pilot}}$ is derived. For simplicity, we assume that the host signal $\mathbf{x}$ and the attack noise $\mathbf{v}$ are mean-free ($\mathrm{E}\{x\} = 0$, and $\mathrm{E}\{v\} = 0$) so that the variance of $\mathbf{x}$ and $\mathbf{w}$ is given by $\sigma_x^2 = \mathrm{E}\{x^2\}$ and $\sigma_v^2 = \mathrm{E}\{v^2\}$, respectively. The derivation of the variance $\mathrm{Var}\{\hat{g}\}$ is tedious but not difficult so that only the main steps are presented here:

$$\mathrm{E}\{\hat{c}^2\} = (g\sigma_w^2)^2 + \frac{g^2\sigma_x^2 + \sigma_v^2}{L_{\text{pilot}}}\sigma_w^2 \tag{30}$$

$$\mathrm{Var}\{\hat{c}\} = \mathrm{E}\{\hat{c}^2\} - \mathrm{E}\{\hat{c}\}^2 = \frac{g^2\sigma_x^2 + \sigma_v^2}{L_{\text{pilot}}}\sigma_w^2 \tag{31}$$

$$\mathrm{Var}\{\hat{g}\} = \mathrm{Var}\left\{\frac{\hat{c}}{\sigma_w^2}\right\} = \frac{\mathrm{Var}\{\hat{c}\}}{(\sigma_w^2)^2} = \frac{g^2\sigma_x^2 + \sigma_v^2}{L_{\text{pilot}}\sigma_w^2} = \frac{g^2\sigma_x^2/\sigma_w^2 + \sigma_v^2/\sigma_w^2}{L_{\text{pilot}}}. \tag{32}$$

We observe that Var $\{\hat{g}\}$ depends on the WNR via $\sigma_x^2/\sigma_w^2$ and on the DWR via $\sigma_v^2/\sigma_w^2$. The term $\sigma_x^2/\sigma_w^2$ dominates for realistic DWRs about 20 dB and WNR $> -10$dB. Further, we observe that Var $\{\hat{g}\}$ decreases with increasing pilot length $L_{\text{pilot}}$.

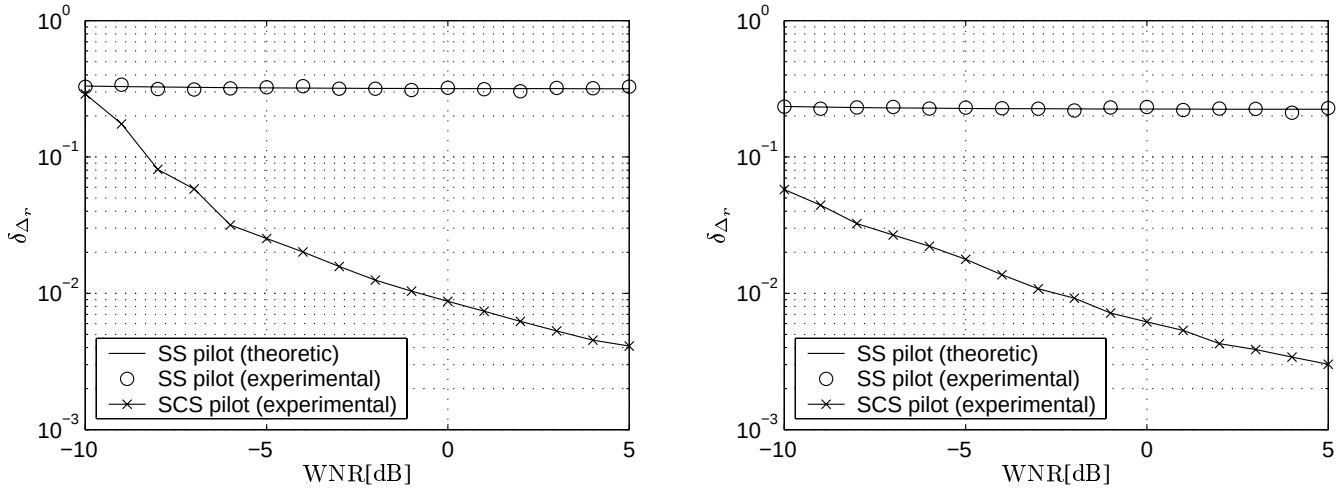


**Figure 7.** Estimation performance for $L_{\text{pilot}}$=1000 (left) and $L_{\text{pilot}}$=2000 (right). The performance for SS pilot watermarks and SCS pilot watermarks is compared (DWR $= 20$ dB, $M = 3$, $L_{\text{bin}} = 50$). The experimental results are averaged over 1000 simulations.

Fig. 7 compares the achieved estimation accuracy using SS pilot watermarks and SCS pilot watermarks for $L_{\text{pilot}} = 1000$ and $L_{\text{pilot}} = 2000$. Note that the estimation accuracy $\delta_{\Delta_r}$ for SS pilot watermarks can be computed theoretically from Var $\{\hat{g}\}$ via

$$\delta_{\Delta_r} = \frac{\sqrt{\mathrm{E}\left\{(\Delta_r - g\Delta)^2\right\}}}{g\Delta} = \frac{\sqrt{\mathrm{E}\left\{(\hat{g}\Delta - g\Delta)^2\right\}}}{g\Delta} = \frac{\sqrt{\mathrm{E}\left\{(\hat{g} - g)^2\right\}}}{g} = \frac{\sqrt{\mathrm{Var}\left\{\hat{g}\right\}}}{g}. \tag{33}$$

The results shown in Fig. 7 clearly demonstrate the superiority of the estimation algorithm based on SCS pilot watermarks. The advantage of the SCS pilot watermarks stems from the reduced influence of host-signal interference on the estimation accuracy.

## 5. NONLINEAR AMPLITUDE MODIFICATION

So far, attacks by linear amplitude modifications have been investigated where the estimation of $\Delta_r$ and $r_{\text{offset}}$ is sufficient for reliable SCS watermark reception after such attacks. However, general nonlinear modifications of the signal amplitude are much more difficult to handle due to the increased number of free parameters for the attack. Further, an objective signal quality evaluation appropriate for nonlinear amplitude modification is often difficult to find. In this paper, we do not consider general nonlinear amplitude modification, however, we demonstrate how the estimation algorithm for linear amplitude modifications described in Sec. 3 can be exploited to invert nonlinear amplitude modifications that are parameterized by one scalar parameter. For demonstration purpose, nonlinear amplitude modification by gamma correction is chosen, which is typical for image processing. First, an overview of the corresponding channel model is given. Next, an outline of our proposed estimation algorithm and some experimental results are presented.

### 5.1. Nonlinear Amplitude Modification by Gamma Correction

Gamma correction is the nonlinear mapping of image pixel intensities $i_n$ that is important for compensating the characteristics of image intensity reproduction of different camera and monitor devices.[10] Here, gamma correction of 8-bit grayscale image data is considered as an example for nonlinear amplitude modification, where the gamma correction is defined by

$$i'_n = 255 \left(\frac{i_n}{255}\right)^{1/\gamma}. \tag{34}$$

Note that $\gamma = 1$ results into an identity mapping. Fig. 8 shows the gamma correction for $\gamma \in \{0.5, 0.6, \ldots, 2\}$, which covers slightly more than the common range of values for $\gamma$.

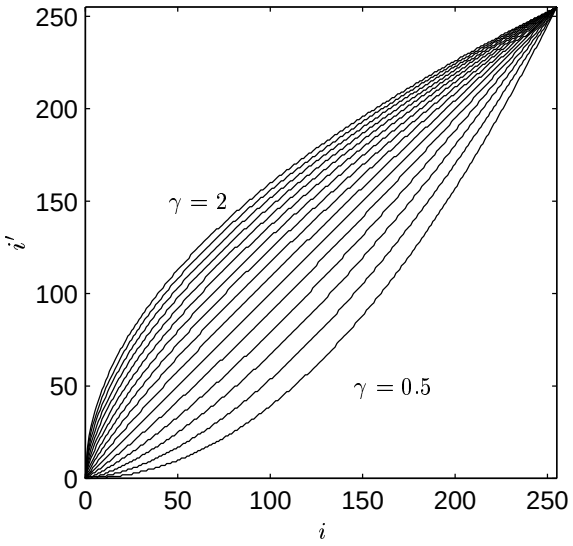


**Figure 8.** Gamma correction for grayscale 8-bit pixel values $i$ and $\gamma \in \{0.5, 0.6, \ldots, 2\}$

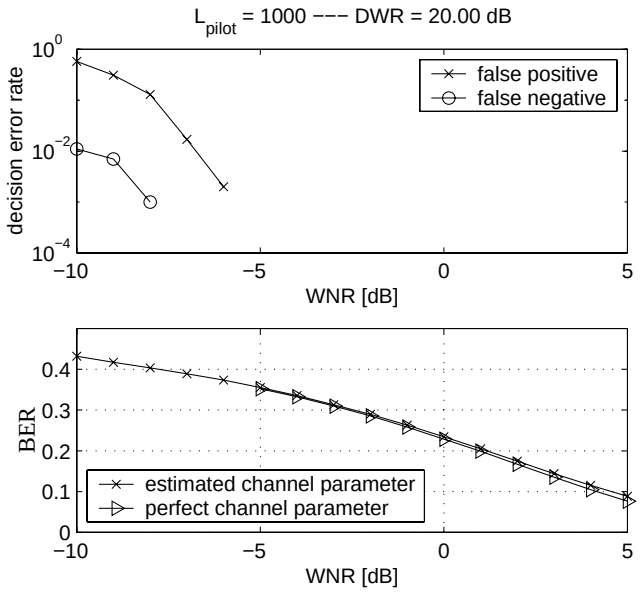


**Figure 9.** Watermark receiver performance after SAWGN attack and gamma correction with $\gamma = 1.2$. Upper plot: SCS watermark detection from $L_{\text{pilot}} = 1000$ data elements. Lower plot: uncoded binary SCS watermark bit-error rate (BER).

It is possible to combine a gamma correction attack with amplitude scaling and additive noise, where the order of the different operations has an influence on the effect of the attack. Here, we do not investigate all possible combinations of the different attack operations, however, we restrict the discussion to the channel model shown in Fig. 10. We propose to invert the effects of the attack before common SCS watermark decoding as much as possible with help of the estimates $\hat{\gamma}$, $\hat{g}$, and $\hat{r}_{\text{offset}}$. Thus, the problem is reduced to obtain the estimates $\hat{\gamma}$, $\hat{g}$, and $\hat{r}_{\text{offset}}$ from the host signal samples with an embedded SCS pilot watermark.

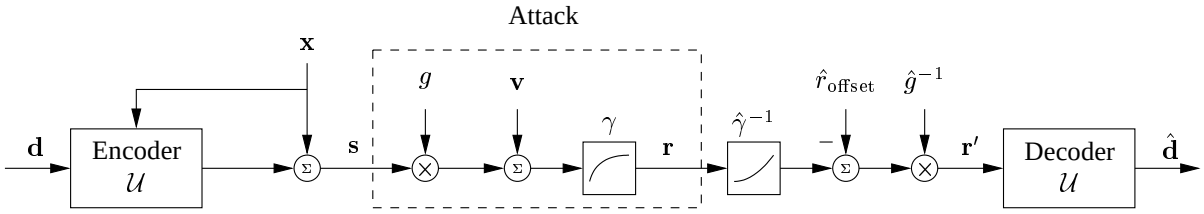


**Figure 10.** Watermark communication facing an attack by amplitude scaling and AWGN with mean $r_{\text{offset}}$ and additional nonlinear amplitude modification by gamma correction.

### 5.2. Estimation of Gamma, Gain, and DC Offset for SCS Watermark Reception

Due to space constraints, only a brief outline of our estimation algorithm for $\hat{\gamma}$, $\hat{g}$, and $\hat{r}_{\text{offset}}$ is given. The basic approach is a combination of the estimation of $g$ and $r_{\text{offset}}$ using the algorithm described in Sec. 3, SCS watermark detection from data elements with the embedded SCS pilot watermark, and a search over all gamma corrections with $\gamma \in \{0.5, 0.6, \ldots, 2\}$.

SCS watermark detection is described in more detail in our previous publications.[11,9] Note that watermark detection refers to the decision whether the received data $\mathbf{r}$ is not watermarked with key $K$ ($H_0$) or is watermarked with key $K$ ($H_1$). We formulate watermark detection as an hypothesis test with the hypotheses $H_0$ and $H_1$. Let $P_r$, with $P_r \in [0, 1]$, denote the reliability that the received data elements $\mathbf{r}$ are watermarked using the key sequence $\mathbf{k}$ derived from the key $K$. $P_r > 0.5$ indicates an embedded SCS watermark ($H_1$).

We propose to estimate $\gamma$, $g$, and $r_{\mathrm{offset}}$ using the following procedure:

1. Inverse gamma correction using all $\gamma \in \{0.5, 0.6, \ldots, 2\}$ is applied to the received data elements $\mathbf{r}_{\mathrm{pilot}}$.
2. For each $\gamma$, the estimation algorithm described in Sec. 3 delivers the maximum histogram frequency component $|A[l_0]|(\gamma)$. Note that all histograms are computed using $L_{\mathrm{bin}} = 256$ bins covering the entire 8-bit range [0,255].
3. For each $\gamma$, the SCS watermark detection reliability $P_r(\gamma)$ is computed.
4. The estimate $\hat{\gamma}$ is equal to $\gamma$ with maximum $|A[l_0]|(\gamma)$ and $P_r(\gamma) > 0.5$. The maximum $|A[l_0]|(\gamma)$ for all $\gamma$ is considered if $P_r(\gamma) \leq 0.5$ for each $\gamma$.
5. $\hat{g}$ and $\hat{r}_{\mathrm{offset}}$ are computed from $A[l_0](\hat{\gamma})$.

Fig. 9 shows some experimental results when using the outlined estimation of $\hat{\gamma}$, $\hat{g}$, and $\hat{r}_{\mathrm{offset}}$ and the attack involves gamma correction with $\gamma = 1.2$. The upper plot depicts the false positive and false negative decision error rate for SCS watermark detection using the estimates $\hat{\gamma}$, $\hat{g}$, and $\hat{r}_{\mathrm{offset}}$. Note that the result is biased in the direction of more false positive errors due to the search over several values for $\gamma$. We are basically looking for some $\hat{\gamma}$ that allows SCS watermark detection and thus produce more false positive errors. In practice, compensation of this bias within the hypothesis test should be applied or SCS watermark detection should be applied for received data elements that are not used within the estimation of $\hat{\gamma}$. We observe that no decision errors are measured for $\mathrm{WNR} > -5\mathrm{dB}$ within 1000 experiments. The lower plot of Fig. 9 shows the bit-error rate (BER) for uncoded binary SCS watermarks. The BER for reception with perfect knowledge of the channel parameters $\gamma$, $g$, and $r_{\mathrm{offset}}$ and the BER for reception using the corresponding estimates are compared. We observe that the BER increases only very slightly for SCS watermark reception using the estimated channel parameters.

## 6. CONCLUSIONS

A method for estimating (non-)linear amplitude modifications of watermarked data based on securely embedded SCS (scalar Costa scheme) pilot watermarks is proposed. The basic algorithm exploits the periodic structure in the histograms of received SCS watermarked data that becomes visible by exploiting the watermark key sequence. The performance of the proposed algorithm is described in detail for linear amplitude modifications. Experimental results show that accurate estimation of the channel parameters requires 500 to 2000 pilot elements depending on the attack noise. It is also demonstrated that the estimation based on SCS pilot watermarks is superior to an alternative approach using spread spectrum watermarks. Finally, the extension of the estimation algorithm to parameterized nonlinear amplitude modifications is outlined.

## REFERENCES


1. B. Chen and G. W. Wornell, "Provably robust digital watermarking," in *Proceedings of SPIE: Multimedia Systems and Applications II (part of Photonics East '99)*, vol. 3845, pp. 43–54, (Boston, MA, USA), September 1999.
2. M. Ramkumar, *Data Hiding in Multimedia: Theory and Applications*. PhD thesis, Dep. of Electrical and Computer Engineering, New Jersey Institute of Technology, Kearny, NJ, USA, November 1999.
3. I. J. Cox, M. L. Miller, and A. L. McKellips, "Watermarking as communications with side information," *Proceedings of the IEEE, Special Issue on Identification and Protection of Multimedia Information* **87**, pp. 1127–1141, July 1999.
4. M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory* **29**, pp. 439–441, May 1983.
5. J. J. Eggers, J. K. Su, and B. Girod, "A blind watermarking scheme based on structured codebooks," in *Secure Images and Image Authentication, Proc. IEE Colloquium*, pp. 4/1–4/6, (London, UK), April 2000.
6. J. J. Eggers, J. K. Su, and B. Girod, "Robustness of a blind image watermarking scheme," in *Proceedings of the IEEE Intl. Conference on Image Processing 2000 (ICIP 2000)*, (Vancouver, Canada), September 2000.
7. J. J. Eggers, J. K. Su, and B. Girod, "Performance of a practical blind watermarking scheme," in *Proc. of SPIE Vol. 4314: Security and Watermarking of Multimedia Contents III*, (San Jose, Ca, USA), January 2001.
8. M. Kesal, M. K. Mihçak, R. Kötter, and P. Moulin, "Iterative decoding of digital watermarks," in *Proc. 2nd Symp. on Turbo Codes and Related Topics*, (Brest, France), September 2000.
9. J. J. Eggers, *Information Embedding and Digital Watermarking as Communication with Side Information*. PhD thesis, Lehrstuhl für Nachrichtentechnik I, Universität Erlangen-Nürnberg, Erlangen, Germany, November 2001. preprint.
10. C. Poynton, "Frequently asked questions about gamma," tech. rep., available at http://www.inforamp.net/~poynton, 1998.
11. J. J. Eggers, R. Bäuml, R. Tzschoppe, and J. Huber, "Applications of information hiding and digital watermarking," in *ECDL Workshop on Generalized Docuements*, (Darmstadt, Germany), September 2001.