Proceedings of SPIE Vol. 4314:
Security and Watermarking
of Multimedia Contents III
San Jose, Ca., Jan. 2001.

Performance of a Practical Blind Watermarking Scheme

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

Jonathan K. Su
MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02420-9185, USA

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

ABSTRACT

In many blind watermarking proposals, the unwatermarked host data is viewed as unavoidable interference. Recently, however, it has been shown that blind watermarking corresponds to communication with side information (i.e., the host data) at the encoder. For a Gaussian host data and Gaussian channel, Costa showed that blind watermarking can theoretically eliminate all interference from the host data. Our previous work presented a practical blind watermarking scheme based on Costa's idea and called "scalar Costa scheme" (SCS). SCS watermarking was analyzed theoretically and initial experimental results were presented. This paper discusses further practical implications when implementing SCS. We focus on the following three topics: (A) high-rate watermarking, (B) low-rate watermarking, and (C) restrictions due to finite codeword lengths. For (A), coded modulation is applied for a rate of 1 watermark bit per host-data element, which is interesting for information-hiding applications. For (B), low rates can be achieved either by repeating watermark bits or by projecting them in a random direction in signal space (spread-transform SCS). We show that spread-transform SCS watermarking performs better than SCS watermarking with repetition coding. For (C), Gallager's random-coding exponent is used to analyze the influence of codeword length on SCS performance.

Keywords: blind watermarking, watermark capacity, channel coding, coded modulation, random coding exponent

1. INTRODUCTION

Blind digital watermarking is the art of communicating a message by embedding it into multimedia data (host data), and decoding it without access to the original, non-watermarked host data. Envisioned applications for such a method are copy control or ownership verification. A blind watermarking scheme must be designed such that the watermarked data has subjective quality close to that of the original host data 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. Although this technique allows for reliable communication even for strong attacks, blind detection of spread-spectrum watermarks suffers significantly from host data interference. In 1999, it was realized that the host data 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, which shows that, for Gaussian data and additive white Gaussian noise (AWGN) attacks, blind watermarking can perform as well as if the decoder had access to the original host data. Costa derived the capacity of blind watermarking facing an AWGN attack. Here, capacity means the maximal achievable rate $R =$ (number of watermark bits)/(number of host-data elements) for a given strength of the attack and any watermarking scheme, including any modulation and any coding.

Costa used a random codebook, which is not practical. We have previously presented a simplified practical blind watermarking scheme, called "scalar Costa scheme" (SCS).[1] The performance of SCS watermarking has been analyzed with respect to the maximal achievable rate R for a given strength of the attack and any coding scheme. SCS does not achieve capacity but is easy to implement, is host-data independent, and can perform significantly better than blind spread-spectrum watermarking. The achievable rate of SCS was analyzed for an AWGN attack,[1] as well as its performance after an optimized linear filtering and additive noise attack.[2] SCS watermarking and its achievable rate will be reviewed in Sec. 2.

This paper discusses further practical implications when implementing SCS. We focus on the following three items:

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

(A) For information hiding applications, high-rate watermarking might be possible. In Sec. 3, we compare the performance of SCS watermarking with three different coded modulation techniques to achieve a rate R of one watermark bit per host-data element. Measured bit-error results are related to those of Chou et al.[3] for the same watermark rate.

(B) Transmission of one watermark bit per host-data element is applicable for information hiding schemes, but for robust watermarking, where strong attacks must be considered, much lower watermark rates are more realistic. In Sec. 4 different methods for low-rate SCS watermarking are discussed.

(C) The achievable rates computed for SCS can be obtained only for codewords of infinite length. In practice, the codeword length is finite. With help of Gallager's random-coding exponent, we analyze the influence of codeword length on the performance of SCS in Sec. 5.

2. SCS WATERMARKING

We consider digital watermarking as a communication problem. The watermark encoder derives from the watermark message m 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. 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, and x_n,w_n,s_n,r_n, and k_n refer to their respective nth elements.

2.1. Watermarking as Communication with Side-Information at the Encoder

Fig. 1 depicts a block diagram of blind watermark communication, where the attacker introduces additive white Gaussian noise (AWGN) $\mathbf{v}$. The depicted scenario can be considered communication with side information about the host data at the encoder.[4]

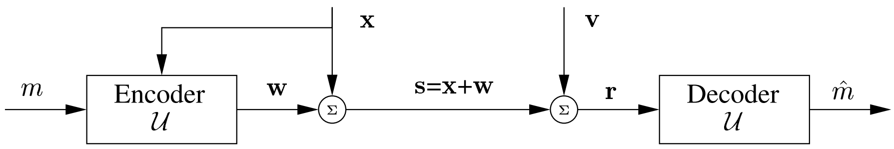


Figure 1. Watermark communication facing an AWGN attack.

Moulin and O'Sullivan[5] showed that for white Gaussian host data $\mathbf{x}$ and MSE distortion measurement, the Gaussian test channel (GTC) is the worst (or best, depending on perspective) possible attack in the sense that the rate of reliable communication is minimized for a constrained distortion of $\mathbf{r}$. The GTC attack combines scaling of the public data $\mathbf{s}$ by g (usually $g < 1$) and additive white Gaussian noise $\mathbf{z}$ of power σ_z^2. Watermark communication facing a GTC attack is depicted in Fig. 2. If $g \neq 0$, the receiver compensates for scaling by dividing $\mathbf{r}$ by g to produce $\mathbf{r}' = \mathbf{s} + \mathbf{z}/g$. Thus, the design of a watermark encoder and decoder in case of a GTC attack can be translated into the design for an *effective* AWGN attack with noise $\mathbf{v} = \mathbf{z}/g$. Note that the optimal scale factor g depends on the host-data power σ_x^2, or equivalently on the watermark-to-document power ratio $\text{WDR} = 10 \log_{10} \sigma_w^2/\sigma_x^2$ dB.

For the communication scenario depicted in Fig. 1, Costa[6] showed theoretically that for Gaussian host data of power σ_x^2, a watermark sequence of power σ_w^2, and AWGN of power σ_v^2 the capacity is $C = 0.5 \log_2(1 + \sigma_w^2/\sigma_v^2)$, independent of σ_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 $\mathbf{x}$. In this paper, we focus on the performance of communication systems for the scenario in Fig. 1, where the attack strength is completely characterized by the watermark-to-noise power ratio $\text{WNR} = 10 \log_{10} \sigma_w^2/\sigma_v^2$ dB.

2.2. Practical Communication Derived from Costa's Scheme

Costa's scheme involves a random codebook $\mathcal{U}$, which 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, we proposed replacing it by a structured codebook, in particular a product codebook of dithered uniform scalar quantizers, and called this scheme *SCS* (Scalar Costa Scheme).[1] In SCS, the watermark message m is encoded into a sequence of watermark letters $\mathbf{d}$, where the elements d_n belong

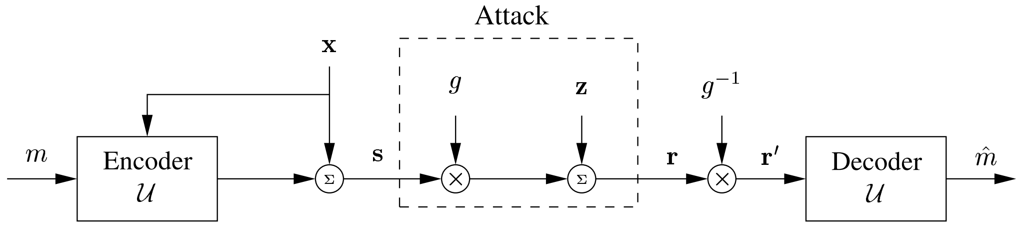


Figure 2. Watermark communication facing a GTC attack.

to a D-ary alphabet $\mathcal{D} = \{0, 1, \ldots, D-1\}$. D-ary signaling denotes SCS watermarking with an alphabet $\mathcal{D}$ of size $D = |\mathcal{D}|$. In many practical cases, binary SCS watermarking ($d_n \in \mathcal{D} = \{0, 1\}$) will be used. Each of the watermark letters is embedded into the corresponding host elements x_n. For example, x_n could be a signal sample or a frequency coefficient of multimedia data. The embedding rule for the nth element is given by

$$s_n = x_n + \alpha \left(\mathcal{Q}_\Delta \left\{ x_n - \Delta \left(\frac{d_n}{D} + k_n \right) \right\} + \Delta \left(\frac{d_n}{D} + k_n \right) - x_n \right), \tag{1}$$

where $\mathcal{Q}_\Delta \{\cdot\}$ denotes scalar uniform quantization with step size Δ. The key $\mathbf{k}$ is a pseudo-random sequence with $k_n \in (0, 1]$. Fig. 3 shows a block diagram of (1). This embedding scheme depends on two parameters: the quantizer step size Δ and the scale factor α. Both parameters can be jointly optimized to achieve a good trade-off between embedding distortion σ_w^2 and detection reliability for a given noise variance σ_v^2 of an AWGN attack. Optimal values for Δ and α must be computed numerically.[1] A good approximation is given by

$$\Delta_{\text{opt}} = \sqrt{12 \left(\sigma_w^2 + 2.71 \sigma_v^2 \right)}, \text{ and } \alpha_{\text{opt}} = \sqrt{\frac{\sigma_w^2}{\sigma_w^2 + 2.71 \sigma_v^2}}. \tag{2}$$

In case of the GTC attack with a certain constraint on the attack distortion, the parameters α and Δ are obtained from those for an equivalent effective AWGN attack with noise power σ_v^2.

At the decoder, the received data $\mathbf{r}$ is demodulated to obtain the data $\mathbf{y}$. The demodulation rule for the nth element is

$$y_n = \mathcal{Q}_\Delta \{ r_n - k_n \Delta \} + k_n \Delta - r_n. \tag{3}$$

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 $d_n = 1$.

The upper plot of Fig. 4 depicts one period of the PDF of the watermarked elements s_n conditioned on the transmitted watermark letter d_n, and $k_n = 0$ for binary SCS. The lower plot shows the respective PDFs of the demodulated received elements y_n after AWGN attack conditioned on the transmitted watermark letter d_n. The PDF $p_y(y_n|d_n)$ is derived numerically.[1] In case of using an incorrect key $\mathbf{k}$ at the receiver, the distribution of $p_y(y_n|d_n)$ will be uniform for any possible $\mathbf{r}$. This is indicated by the dotted line in the lower plot of Fig. 4.

2.3. Performance Limits of SCS Watermarking

A detailed analysis of the performance limits of SCS watermarking is given in our previous work.[1] Here, we summarize the most important results. Fig. 5 compares the achievable rates obtained for SCS watermarking with the capacity of the ideal Costa scheme. Obviously, SCS watermarking does not achieve capacity, but is not too far from an ideal scheme either. Further, the achievable rates of binary dither modulation (DM), proposed by Chen and Wornell,[4] and blind spread-spectrum watermarking (SS) are shown. Binary DM can be considered a special case of SCS watermarking with $\alpha = 1$ for all WNRs. Fig. 5 shows that DM performs poorly for negative WNRs, where the optimal value of α is significantly smaller than 1. Blind SS watermarking suffers from host-data interference, and its performance depends highly on the statistics of the host data. The depicted achievable rate of blind SS watermarking is for Gaussian host data with WDR $= -15$ dB. For weak to moderately strong attacks (i.e., WNRs greater than about -10 dB) SCS watermarking outperforms SS watermarking by far due to the host-data independent nature of SCS watermarking. However, the right plot in Fig. 5 also reveals that for very strong attacks (WNR < -15 dB), blind SS is more appropriate than SCS watermarking since here the attack distortion is more important than the host-data interference. Note that the ideal Costa scheme would outperform blind SS watermarking at all attack distortion levels.

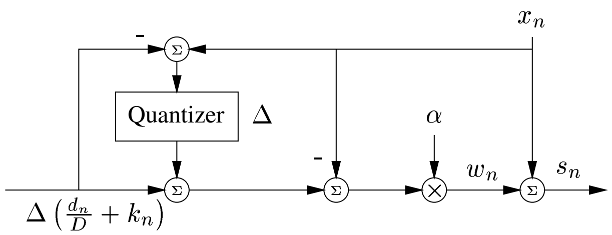


Figure 3. Watermark embedding using Costa's scheme with a scalar component codebook (SCS). The watermark letter $d_n \in \mathcal{D}$ is embedded after dithered uniform scalar quantization of x_n and the addition of the scaled quantization error as watermark w_n.

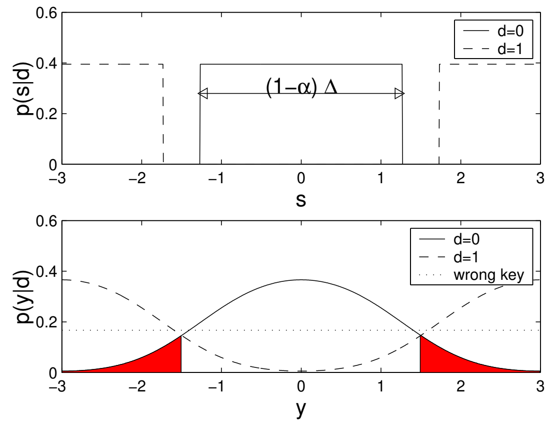


Figure 4. One period of the PDFs of the watermarked data s and the demodulated data r for binary SCS (σ_w^2=1, WNR $= 3\,$dB, $\Delta = 6$, $\alpha = 0.58$). The filled areas represent the probability of detection errors assuming $d = 0$ was transmitted.

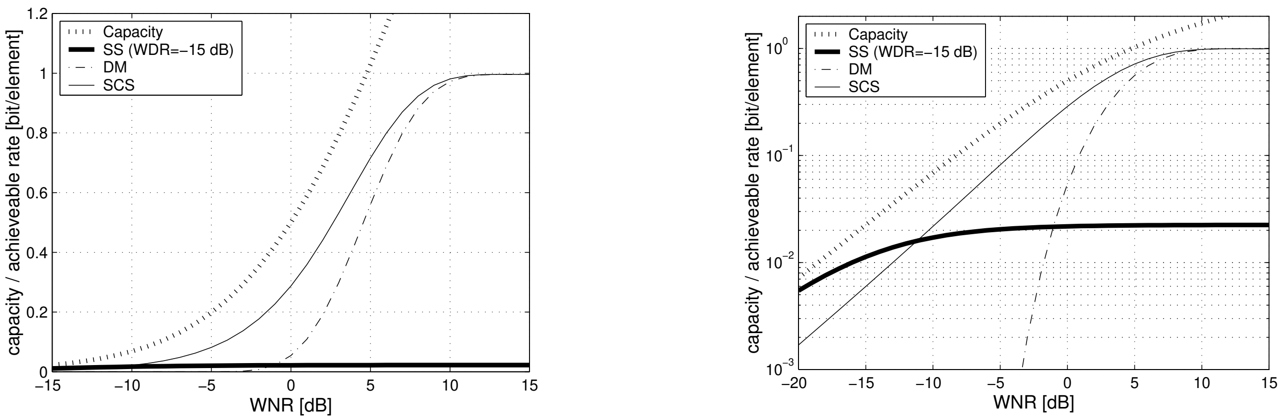


Figure 5. Capacity of blind watermarking facing an AWGN attack compared with the achievable rates of binary SCS, binary DM and blind spread-spectrum (SS) watermarking. The achievable rates are shown with linear (left) and logarithmical (right) scales.

3. HIGH-RATE SCS WATERMARKING

Here, SCS watermarking over an AWGN channel at rates $R > 0.5$ bit/element is considered. Although robust watermarking at these rates is unrealistic, related applications like information hiding might operate at such high rates. The threshold of 0.5 bit/element for the definition of "high-rate" SCS was chosen since for higher rates, the achievable rate of binary SCS is significantly lower than for D-ary signaling with $D > 2$, as shown in Fig. 6. We observe that the size of the alphabet $\mathcal{D}$ has a significant influence only for WNRs larger than about $\approx 4\,$dB, or equivalently $R > 0.6$ bit/element.

Coded modulation techniques are used to combine D-ary signaling with binary error-correction coding. Here, we investigate the performance of SCS at $R = 1$ bit/element for different coded modulation techniques. As shown in Fig. 6, for $R = 1$ bit/element, 3-ary signaling is as good as D-ary signaling with $D > 3$. However, 4-ary or 8-ary signaling is discussed here due to its efficient combination with binary coding techniques. Thus, the watermark letters d_n are from the alphabet $\mathcal{D} = \{0, 1, 2, 3\}$ or $\mathcal{D} = \{0, 1, 2, 3, 4, 5, 6, 7\}$. The letters are specified by the binary sequence $d_n = d_n^0 d_n^1$ or $d_n = d_n^0 d_n^1 d_n^2$, where $d_n^0, d_n^1, d_n^2 \in \{0, 1\}$, and d_n^0 is the least-significant bit.

A detailed discussion of coded modulation is beyond the scope of this paper. Our main goal is to demonstrate that, with $R = 1$, low bit-error rates ($p_b < 10^{-5}$) can be achieved within 1.6 dB of the maximal achievable rate of SCS. Fig. 7 shows

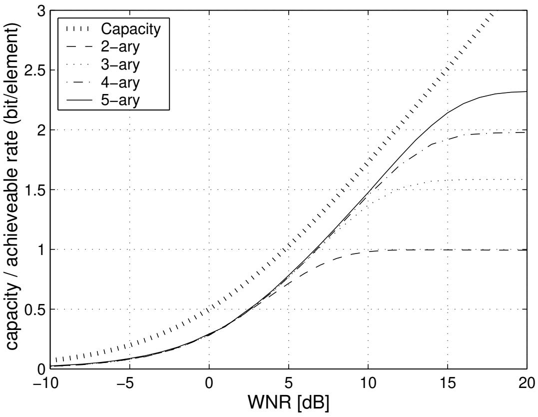


Figure 6. Achievable rate for SCS with D-ary signaling. Binary SCS is not appropriate for WNR $> 4\,\mathrm{dB}$ or equivalently rates > 0.6 bit/element.

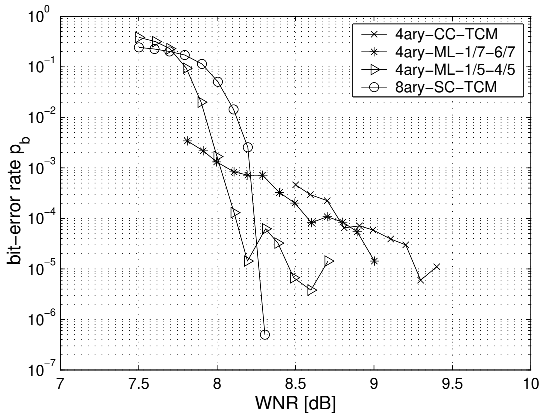
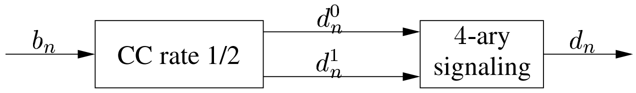
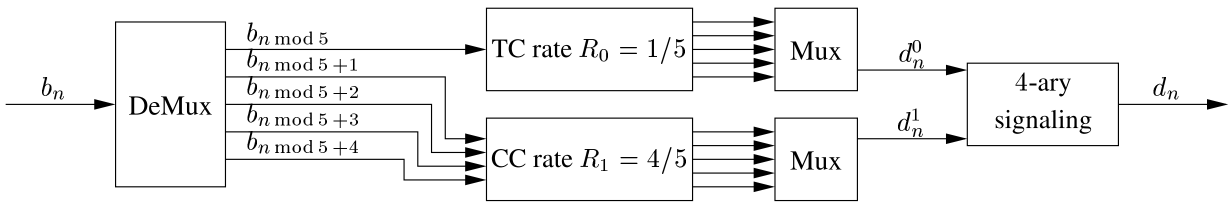


Figure 7. Measured bit-error rate p_b for rate 1 bit/element are shown for 4-ary convolutional coded trellis coded modulation (CC-TCM), 4-ary multilevel coded modulation with rate assignment of $R_0 = 1/7$ and $R_1 = 6/7$ (ML-1/7-6/7), or $R_0 = 1/5$ and $R_1 = 4/5$ (ML-1/5-4/5), and 8-ary serial concatenated trellis coded modulation (SC-TCM).

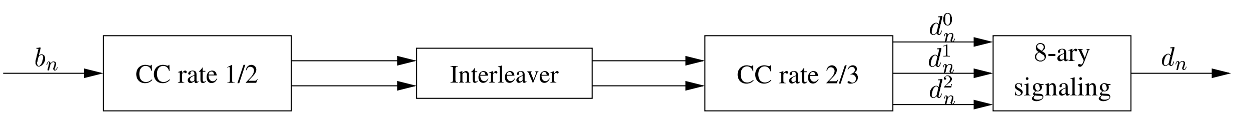
simulation results for *trellis coded modulation with convolutional codes* (CC-TCM), *multilevel coding* (ML) and *trellis coded modulation with serial concatenated codes and iterative decoding* (SC-TCM). A brief description of the encoding process for the different coded modulation schemes is given before the discussion of these simulation results. For the corresponding decoding processes the reader is referred to the literature.



(a) 4-ary-CC-TCM



(b) 4-ary-ML-1/5-4/5



(c) 8-ary-SC-TCM

Figure 8. Encoder for the applied coded modulation schemes 4-ary-CC-TCM, 4-ary-ML-1/5-4/5, and 8-ary-SC-TCM.

The mapping of a watermark message m onto a sequence of watermark letter d_n depends on the coded modulation technique. However, in all schemes considered, the message m is first mapped onto a binary sequence $\mathbf{b}$, with one element $b_n \in \{0, 1\}$ for each host-data element x_n. Fig. 8 depicts the block diagrams for the encoding of $\mathbf{b}$ into $\mathbf{d}$ for 4-ary-CC-TCM, 4-ary-ML-1/5-4/5, and 8-ary-SC-TCM. 4-ary-CC-TCM denotes the classical TCM proposed by Ungerboeck.[7] The information bits $\mathbf{b}$ are encoded with a rate 1/2 convolutional code so that for each data element x_n, one out of four possible watermark letters d_n is selected. A Viterbi decoder[8] with the conditional probabilities $p_y(y_n|d_n)$ as path weight can be used as decoder for 4-ary-CC-TCM.

Multilevel coding, originally proposed by Imai and Hirakawa,[9] is a combined coding and modulation method based on binary component codes for the least-significant bits $\mathbf{d}^0$ and the most-significant bits $\mathbf{d}^1$ of the 4-ary letters $\mathbf{d}$. An important issue in the design of multilevel codes is the choice of component codes and their code rates. Wachsmann et al.[10] proposed a technique for selecting the component code rates depending on the capacity of the equivalent binary input channels for communicating $\mathbf{d}^0$ and $\mathbf{d}^1$. Here, simulations for the rate assignment of $R_0 = 1/5$ and $R_1 = 4/5$ (ML-1/5-4/5), and $R_0 = 1/7$ and $R_1 = 6/7$ (ML-1/7-6/7) are presented. A turbo code[11] (TC) of rate R_0 is applied for level 0, and a convolutional code (CC) of rate R_1 is used for level 1. Fig. 8(b) shows the encoding process for ML-1/5-4/5. A multi-stage decoder[10] first decodes the level 0 code. Next, the level 1 code is decoded, where the decoding results for level 0 can be already exploited.

Fig. 8(c) depicts the encoder for SC-TCM. Here, 8-ary modulation is used, where the information bits $\mathbf{b}$ are first encoded with a non-systematic rate 1/2 convolutional code. Next, the encoded sequence is interleaved and further encoded with a systematic rate 2/3 convolutional code. The code design is equivalent to that by Vucetic et al.,[12] however, here the output of the second convolutional code is directly mapped on the 8-ary watermark letters $\mathbf{d}$. The decoder[12] iteratively decodes the inner and outer convolutional codes. The interleaver between both concatenated codes is important to break blocks of unreliably detected bits. Consequently, the interleaver length is an important parameter for the performance of SC-TCM. Here, we used an interleaver of length 20,000 for 10,000 information bits $\mathbf{b}$. The presented SC-TCM scheme is only one example for coded modulation techniques with concatenated codes and iterative decoding algorithms. Improved results may be achievable by using parallel concatenated codes and an optimized choice of component convolutional codes.

The simulation results in Fig. 7, measured for more than $1,000,000$ transmitted bits, demonstrate that bit-error rates around 10^{-5} are achieved by all considered coded modulation schemes for WNR > 9.3 dB. The best performance for $p_b < 10^{-5}$ was achieved by SC-TCM, with a minimal required WNR ≈ 8.3 dB. However, SC-TCM is also the most complex scheme with a block length of 10,000 information bits. The results for ML coded modulation are obtained for information bit blocks of length 1,000, and are less than a dB higher than ST-TCM at $p_b < 10^{-5}$. We observe that the rate assignment $R_0 = 1/5, R_1 = 4/5$ is superior at low error rates, and the rate assignment $R_0 = 1/7, R_1 = 6/7$ is more appropriate at high error rates. CC-TCM gives the worst results of all considered schemes. However, it is also the least complex scheme and thus might be an option for practical applications. Ideally, SCS with $R = 1$ is possible for WNR > 6.7 dB. Thus, the discussed coded modulation schemes come within $1.6 - 2.7$ dB of an optimal coding scheme.

Chou et al.[3] proposed a blind watermarking scheme where an optimization algorithm is used to design the codebook $\mathcal{U}$. They presented simulation results for different versions of their approach at bit-error rates of $p_b \leq 10^{-5}$ and a watermark rate $R = 1$ bit/element. Their best scheme operates at WNR ≥ 10 dB. Thus, SCS combined with coded modulation outperforms the approach of Chou et al.[3] by about $0.7 - 1.7$ dB.

4. LOW-RATE SCS WATERMARKING

In most watermarking applications, the distortion that can be introduced by an attacker into the watermarked data $\mathbf{s}$ will be at least as large as the watermark embedding distortion. For the scenario depicted in Fig. 1, this means that a WNR of about 0 dB or less must be considered. For these distortion levels, binary SCS watermarking is sufficient. Thus, the watermark message m, represented by a binary sequence $\mathbf{b}$, has to be encoded into a sequence $\mathbf{d}$ of binary watermark letters $d_n \in \{0, 1\}$. In order to achieve communication with low error rates, each bit of $\mathbf{b}$ has to be embedded redundantly into the host data $\mathbf{x}$. Here, we investigate different methods for the redundant embedding of $\mathbf{b}$, and compare their performance for an AWGN attack as depicted in Fig. 1.

4.1. SCS Watermarking with Repetition Coding

The simplest approach for the redundant embedding of the information bits $\mathbf{b}$ into the host data $\mathbf{x}$ is the repeated embedding of each bit. Let ρ denote the *repetition factor*, thus the sequence of watermark letters $\mathbf{d}$ is ρ times longer than $\mathbf{b}$. The watermark letters $\mathbf{d}$ are embedded, transmitted and demodulated as described in Sec. 2. However, instead of deciding for each demodulated value y_n what transmitted watermark letter d_n is most likely, the decoder can estimate directly the most likely transmitted watermark information bit b_k from ρ different demodulated values in $\mathbf{y}$. Without loss of generality, we assume that the kth information bit b_k has been embedded into the data elements $x_{\rho k}, x_{\rho k+1}, \ldots, x_{\rho k+\rho-1}$. For an AWGN attack, the demodulated values y_n are independent identically distributed, so the decoder computes the probability of a transmitted information bit $b_k = 1$ via

$$p\left(b_k=1\right) = \frac{\prod_{n=\rho k}^{\rho k+\rho-1} p_y\left(y_n|d_n=1\right)}{\prod_{n=\rho k}^{\rho k+\rho-1} p_y\left(y_n|d_n=1\right) + \prod_{n=\rho k}^{\rho k+\rho-1} p_y\left(y_n|d_n=0\right)}. \tag{4}$$

Finally, a hard estimate $\hat{b}_k$ of the kth information bit is obtained by

$$\hat{b}_k = \begin{cases} 1 & , p\left(b_k=1\right) > 0.5, \\ 0 & , p\left(b_k=1\right) \leq 0.5. \end{cases} \tag{5}$$

4.2. Spread-Transform SCS watermarking

A different approach to redundantly embed the information $\mathbf{b}$ into host data $\mathbf{x}$ is called spread-transform (ST) watermarking. ST watermarking was originally proposed by Chen and Wornell[13] to improve binary dither modulation (DM) watermarking. In ST watermarking, the watermark is not directly embedded into the host data $\mathbf{x}$, but into the projection $\mathbf{x}^{\mathrm{S}T}$ of $\mathbf{x}$ onto a random sequence $\mathbf{t}$. Let τ denote the spreading factor, meaning the number of host-data elements x_n belonging to one element $x_k^{\mathrm{S}T}$. For simplicity, we assume that τ is an integer value, although spread transforms with rational spreading factors can be implemented, too. Further, we assume that τ consecutive elements of $\mathbf{x}$ are transformed into one element of $x_k^{\mathrm{S}T}$. Thus, the spread transform can be computed by

$$x_k^{\mathrm{S}T} = \sum_{n=\tau k}^{\tau k+\tau-1} x_n t_n. \tag{6}$$

Now, any algorithm can be applied to embed a watermark into $\mathbf{x}^{\mathrm{S}T}$ to obtain $\mathbf{s}^{\mathrm{S}T}$. Note that proper normalization of the spreading vector $\mathbf{t}$ is assumed. The watermarked data $\mathbf{s}$ is computed by the inverse spread transform

$$s_n = x_n - x_k^{\mathrm{S}T} t_n + s_k^{\mathrm{S}T} t_n, \tag{7}$$

where $k = n \bmod \tau$. For watermark detection, the received data $\mathbf{r}$ has to be projected onto $\mathbf{t}$, too. Thus demodulation and decoding of the watermark information has to be performed on the transformed data $\mathbf{r}^{\mathrm{S}T}$, where

$$r_k^{\mathrm{S}T} = \sum_{n=\tau k}^{\tau k+\tau-1} r_n t_n. \tag{8}$$

The basic idea behind ST watermarking is that any component of the channel noise $\mathbf{v}$ being orthogonal to the spreading vector $\mathbf{t}$ does not impair watermark detection. Thus, an attacker, not knowing the exact spreading direction $\mathbf{t}$, has to introduce much larger distortions to impair a ST watermark as strong as a watermark embedded directly into $\mathbf{x}$. For an AWGN attack, the effective WNR_τ after ST with spreading factor τ is given by

$$\mathrm{WNR}_\tau = \mathrm{WNR}_1 + 10\log_{10}\tau. \tag{9}$$

Thus, doubling the spreading length τ gives an additional power advantage of 3 dB for the watermark in the ST domain.

4.3. Comparison of SCS with Repetition Coding and ST-SCS

The bit-error rate p_b for SCS watermarking with repetition coding or with ST-SCS after an AWGN attack have been measured for different WNRs. Fig. 9 shows simulation results for $\rho = 2, 4, 8$ and $\tau = 2, 4, 8$, where plots with linear and logarithmic axes for the error rate are provided. We observe that ST-SCS gives significantly lower error rates than SCS with repetition coding at the same watermarking rate, meaning $\tau = \rho$. The predicted WNR gain of 3 dB for the same detection reliability by doubling τ can be observed. However, the for SCS with repetition coding, the WNR gain is less than 3 dB when $\rho = 2$. At first glance, this result is surprising since repetition coding and bipolar transmission for conventional communication without side information at the encoder, both give a 3 dB advantage when $\rho = \tau = 2$. However, the observed effect can be explained

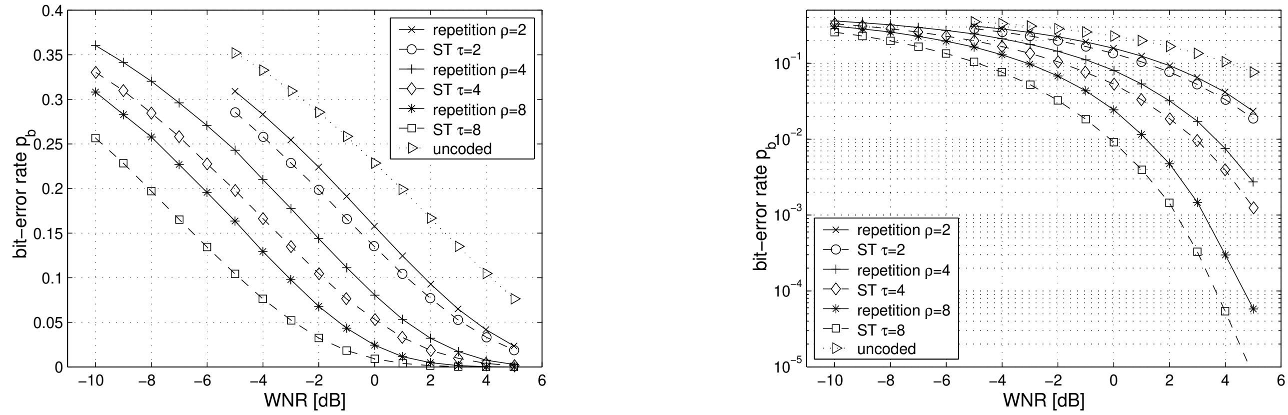


Figure 9. Measured bit-error rates p_b for ST-SCS communication and SCS watermarking with repetition coding. For identical watermarking rates (spreading factor τ = repetition factor ρ), ST-SCS gives lower error rates than SCS with repetition coding.

by examining the specific structure of the codebook $\mathcal{U}$ in SCS. The multiple representations of a single watermark letter d_n by several points in the signaling space leads to many nearest neighbors which can lead to detection errors. Fig. 10 shows a section of the two-dimensional PDFs of the received data $\mathbf{r}$ in the case of an information bit $b_k = 0$ for SCS with repetition coding with $\rho = 2$; bright areas indicate high probabilities. The key sequence $\mathbf{k}$ has been set to zero for illustration purposes. The circles and crosses depict the codebook entries corresponding to a transmitted watermark bit $b_k = 0$ and $b_k = 1$, respectively. Each circle is surrounded by four near-by crosses. Fig. 11 shows the corresponding two-dimensional PDFs in the case of ST-SCS with $\tau = 2$, where the spreading direction $\mathbf{t}$ was chosen to be the main diagonal. Obviously, any noise that is orthogonal to $\mathbf{t}$ does not affect the decision whether the transmitted bit was 0 or 1. Further, each circle is surrounded only by two crosses. Thus, the probability that AWGN pushes watermarked data into the area where a detection error occurs is lower for ST-SCS than for SCS with repetition coding.

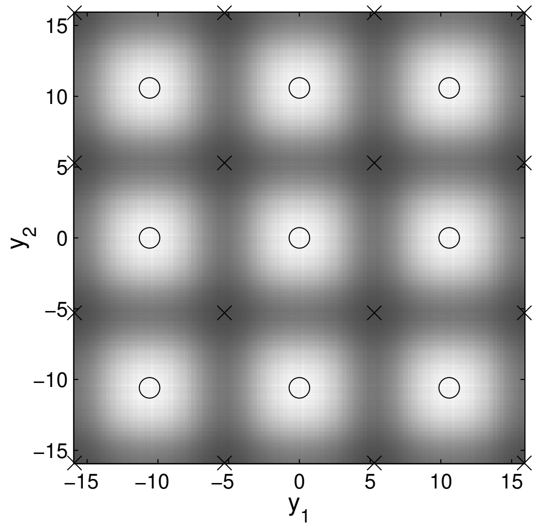


Figure 10. Detection statistics for SCS with repetition coding with $\rho = 2$.

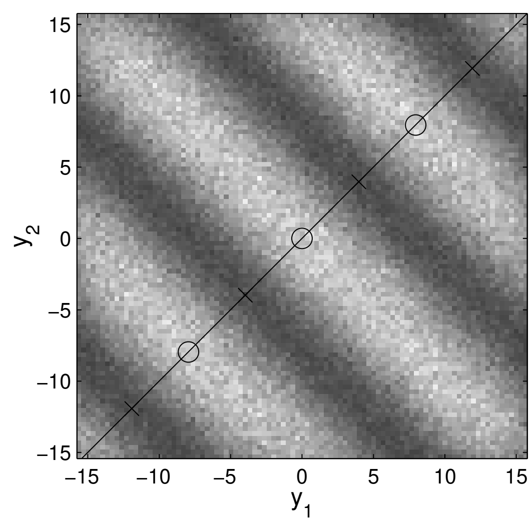


Figure 11. Detection statistics for ST-SCS with $\tau = 2$.

Attention: The advantage of ST-SCS over SCS with repetition coding is only possible if the spreading direction $\mathbf{t}$ is not known to an attacker. Otherwise, an attacker would place all the noise in the direction $\mathbf{t}$ and the WNR-advantage vanishes. Further, ST-SCS watermarking with large spreading factors τ might be impractical since perfect synchronization of the complete spreading vector $\mathbf{t}$ is necessary. In contrast, detection in the case of SCS with repetition coding is possible when only some of the watermarked data elements are synchronized. Another potential problem with large spreading factors τ is that the host-data power in the ST domain might become so low that the assumption that the host-data is approximately uniformly distributed in the range of one quantizer cell no longer holds; this assumption is used in quantization based watermarking schemes like SCS and DM. As a consequence, the power of the watermark can no longer be predicted by $\Delta^2/12$. However, this problem can be avoided by using a key sequence $\mathbf{k}$ (Sec. 2) that acts as a dither sequence that ensures a quantization noise power of $\Delta^2/12$.

4.4. Achievable Rate of ST Watermarking and Optimal Spreading Factor τ

ST-SCS watermarking should to be considered a different suboptimal approach to implementing a transmission scheme for channel coding with side-information at the encoder (Fig. 1). Thus, the achievable rate of ST-SCS might be larger than that of SCS. Note that ST-SCS can never perform worse than SCS since SCS is a special case of ST-SCS with $\tau = 1$. The performance improvement by ST-SCS has already been shown in previous work.[1] Here, we investigate the optimal spreading factor τ for attacks of differing noise power.

Let $C_\tau(\mathrm{WNR})$ denote the achievable rate of a specific watermarking scheme, e.g., an ideal Costa scheme, D-ary ST-SCS watermarking, or D-ary ST-DM, with spreading factor τ for a certain WNR. $C_1(\mathrm{WNR})$ is the achievable rate of the respective scheme without ST, e.g., D-ary SCS watermarking as shown for $D = 2, \dots, 5$ in Fig. 6. The performance of ST watermarking can be computed from that of the respective scheme without ST by

$$C_\tau(\mathrm{WNR}) = \frac{C_1(\mathrm{WNR}_\tau)}{\tau} = \frac{C_1(\mathrm{WNR} + 10\log_{10}\tau)}{\tau}. \tag{10}$$

The term $\log_{10} C_\tau(\mathrm{WNR})$ decreases linearly with an increasing value of $\log_{10}\tau$. Thus, ST watermarking can give a gain only if $\log_{10} C_1(\mathrm{WNR})$ has a slope which is steeper than one decade per $\Delta\mathrm{WNR} = 10\,\mathrm{dB}$. For an ideal Costa scheme, this slope is achieved only in the limit as $\mathrm{WNR} \to -\infty$; thus, the ST does not give a gain. However, for the suboptimal SCS and DM watermarking, there exists a critical $\mathrm{WNR}_{\mathrm{crit}}$ such that for all $\mathrm{WNR} < \mathrm{WNR}_{\mathrm{crit}}$, the slope of $\log_{10} C_\tau(\mathrm{WNR})$ is steeper than one decade per $\Delta\mathrm{WNR} = 10\,\mathrm{dB}$. Consequently, ST watermarking is useful for all $\mathrm{WNR} < \mathrm{WNR}_{\mathrm{crit}}$, and the optimal spreading factor is such that the effective $\mathrm{WNR}_\tau = \mathrm{WNR}_{\mathrm{crit}}$.

Fig. 12 shows the achievable rates of SCS, ST-SCS, DM, and ST-DM with logarithmic capacity axis. The curves clearly demonstrate that ST is advantageous if the slope of the logarithmic achievable rate curves is larger than one decade per $\Delta\mathrm{WNR} = 10\,\mathrm{dB}$. Since the achievable rates for SCS and DM watermarking are computed numerically, the corresponding critical $\mathrm{WNR}_{\mathrm{crit}}$ are also obtained numerically. We found that for SCS, $\mathrm{WNR}_{\mathrm{crit,SCS}} = 0.01\,\mathrm{dB}$, and for DM, $\mathrm{WNR}_{\mathrm{crit,DM}} = 5.81\,\mathrm{dB}$. Fig. 12 shows also that DM can be improved significantly for $\mathrm{WNR} < \mathrm{WNR}_{\mathrm{crit,DM}}$, where for SCS only a minor gain is accessible. Note that ST-DM performs worse than simple SCS for most practical WNRs. Also, there is a constant gain of about 1.8 dB for ST-SCS over ST-DM in the range of negative WNRs.

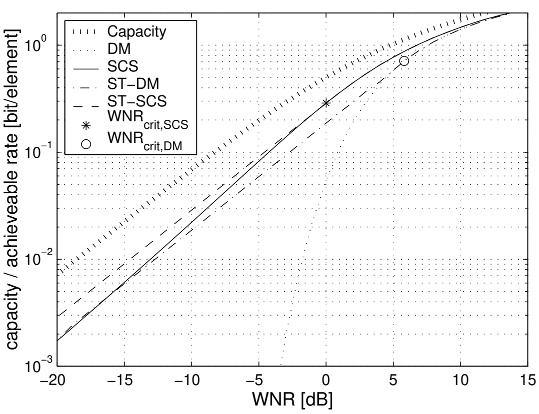


Figure 12. Performance improvement by spread-transform watermarking.

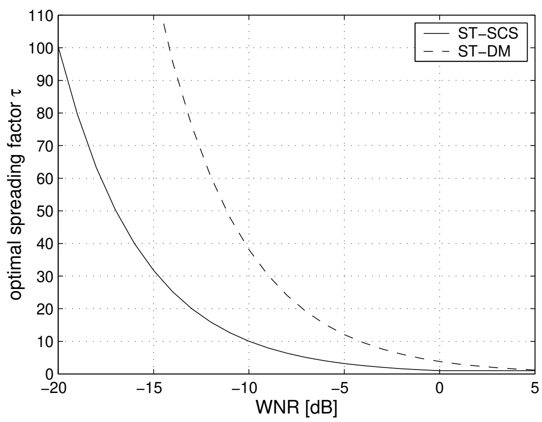


Figure 13. Optimal spreading factors τ for different WNRs (differently strong attacks).

Fig. 13 shows the optimal spreading factors τ for ST-SCS and ST-DM at different WNRs. We assume for simplicity that rational spreading factors τ are implementable. We observe that for each $\mathrm{WNR} < 0\,\mathrm{dB}$, the optimal spreading factor τ for ST-

DM is 3.8 times that of ST-SCS. Thus, ST-DM needs longer spreading vectors while giving poorer performance than ST-SCS. Large τ are not desirable in practice due to more complex synchronization and less host-data power in the ST domain.

4.5. SCS with State-of-the-Art Channel Coding

Repetition coding is known to be very inefficient. State-of-the-art error correction codes, e.g., turbo codes,[11] outperform repetition coding by far. Fig. 14 shows the measured bit-error rates for turbo coded SCS watermarking over an AWGN channel. Turbo codes with coding rates $R = 1/2$, $R = 1/3$, $R = 1/5$, and $R = 1/7$ and a random interleaver of length $N_i = 10,000$ were used. The shown bit-error rates reflect the typical behavior of turbo codes with random interleaving. The bit-error rate p_b decreases rapidly for a certain WNR, but does not decrease further than $p_b \approx 10^{-5}$, which is denoted as error floor. This error floor is mainly determined by codewords with low Hamming distance. We observe that the error floor increases for turbo coding at lower code rates R. This effect is again due to the multiple representation of SCS watermark letters in the signaling space. Note that the error floor of turbo codes can be reduced by an improved interleaver design.[14]

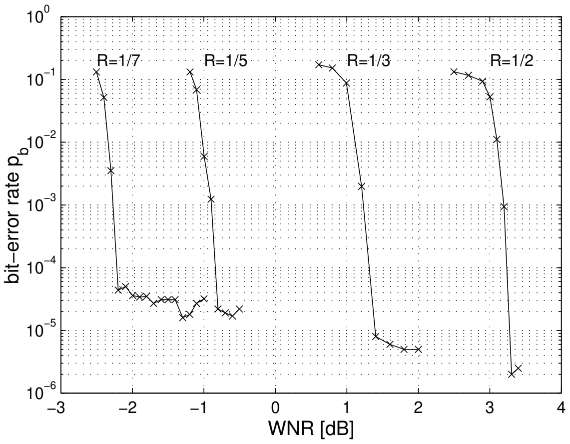


Figure 14. Turbo coded (TC) SCS with random interleaver of length 10,000 and code rates $R = 1/2$, $R = 1/3$, and $R = 1/5$, and $R = 1/7$.

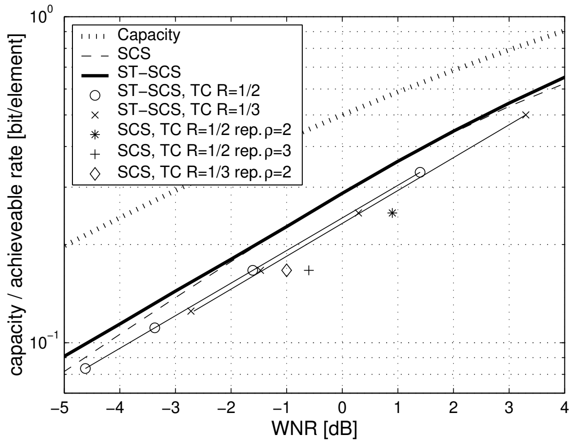


Figure 15. Reliable coded binary (ST-)SCS compared with theoretical limits. The measured points indicate the minimal WNR for that a specific coding technique achieves a bit-error rate $p_b < 10^{-5}$.

The minimal $\mathrm{WNR}_{\min}$ for that coded SCS watermarking gives bit-error rates of $p_b \approx 10^{-5}$ is shown in Fig. 15, where the actual rate R of a specific coded system and the maximal achievable rate are plotted logarithmically. We observe that turbo coded SCS performs indeed close to the maximal achievable rate of SCS watermarking. The coding results for $R = 1/2$ and $R = 1/3$ can be translated to lower rates R via ST watermarking, which is indicated by the straight lines. ST-SCS watermarking with a rate $R = 1/3$ turbo code seems to be a very good choice for low-rate watermarking if any desired ST length τ is applicable. Fig. 15 shows also that turbo coded SCS combined with repetition coding is less efficient than ST-SCS. Nevertheless, repetition coding might be useful in practice since it can be implemented in a very flexible way. Any received data element r_n with embedded watermark bit b_k increases the estimation reliability for $\hat{b}_k$, where for ST watermarking all data elements r_n required for computation of the projection r_k^{ST} must be available to the receiver.

5. SCS WATERMARKING PERFORMANCE FOR CONSTRAINED CODEWORD LENGTH

So far, the performance of SCS watermarking in the case of AWGN attacks has been evaluated either by the achievable rate R or the measured bit-error rates p_b for specific error-correction codes. Both evaluations are somewhat unsatisfying for practical watermarking applications. The achievable rate of SCS watermarking at a given WNR can be obtained for infinitely long codewords. However, at the very least, in practice the codeword length N_c is limited by the number of host-data elements to be watermarked. Further, simulation results for specific codes might be misleading since we cannot be sure that better performance cannot be achieved by some other code of identical codeword length N_c. To analyze the limits of a watermarking technology, we are interested in the achievable performance of *any* coding scheme with a constrained codeword length N_c. At present, we are not able to provide such a limit. However, with the help of Gallager's random coding exponent,[15] there at least exists a way to bound the *word error rate* p_w for SCS with an average random code of codeword length N_c.

Gallager's random coding exponent $E(R)$ is defined by

$$E(R) = \max_{0 \leq \eta \leq 1} \{E_0(\eta) - \eta R\}, \tag{11}$$

where for SCS watermarking

$$E_0(\eta) = -\log_2 \left\{ \int_y \left[\sum_{d \in \mathcal{D}} p\left(d\right) \left(p_y\left(y|d\right)\right)^{\frac{1}{1+\eta}} \right]^{1+\eta} \mathrm{d}y \right\} \tag{12}$$

with $p_y\left(y|d\right)$ as defined in Sec. 2.

Then[15] there exists a code of codeword length N_c such that $p_w \leq 2^{-N_c E(R)}$. $E(R)$ depends on the probabilities $p_y\left(y|d\right)$ and thus for AWGN attacks, on the WNR. Therefore, it is possible to compute the minimum $\mathrm{WNR}_{\min}$ for which a code with $p_w \leq 2^{-N_c E(R)}$ exists. Note that nothing is said about the converse, meaning there could be a code that fulfills the upper bound on p_w for WNRs lower than $\mathrm{WNR}_{\min}$. Nevertheless, in practice the random coding exponent provides a tight bound on the achievable performance of coding with a constrained codeword length N_c.

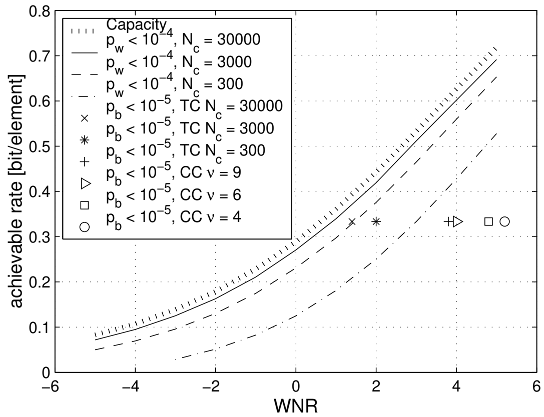


Figure 16. Achievable watermark rate for constrained codeword length N_c

If we fix N_c and p_w so that $E(R) = -\frac{1}{N}\log_2 p_w$, than the maximum rate R of SCS for a given WNR can be computed. Fig. 16 shows results for $p_w = 10^{-4}$ and three different codeword lengths $N_c = 30,000$, $N_c = 3,000$, and $N_c = 300$. For a fixed rate R, a significantly higher WNR is required to achieve reliable detection with short codewords than with long codewords. A loss of about 3 dB in $\mathrm{WNR}_{\min}$ for a certain rate R compared to the achievable rate of SCS must be tolerated when reducing the codeword length to $N_c = 300$. Equivalently, for a fixed WNR, a significant loss in achievable SCS rate must be tolerated. This effect is particularly strong for negative WNRs, where the capacity curve has a flat slope. For instance, at WNR $= 0$ dB, the rate of reliable SCS watermarking with codewords of length $N_c = 300$ is only 0.1 bit/element, rather than 0.3 bit/element as predicted by the theoretically computed achievable rate of SCS.

Fig. 16 depicts also $\mathrm{WNR}_{\min}$ for coded SCS watermarking with rate $R = 1/3$ achieving a *bit-error rate* $p_b < 10^{-5}$, where turbo codes (TC) and convolutional codes with different interleaver lengths N_i and memory lengths ν, respectively, are used. These experimental results show the same tendency as the theoretical results obtained with Gallager's random coding exponent. Turbo codes with short interleaver lengths, and thus short codeword lengths, perform significantly worse than those with long interleaver lengths. The rate $R = 1/3$ turbo code with interleaver length $N_i = 100$, which equals a codeword length of $N_c = 300$, does not perform significantly better than a simple convolutional code with memory $\nu = 9$. The performance of convolutional codes decrease with shorter memory of the encoder. This effect is related to the reduced performance of SCS watermarking for reduced codeword lengths N_c of block codes.

6. CONCLUSION

SCS watermarking is a practical blind watermarking scheme that is particularly good for weak to moderately severe attacks. In this paper, several practical aspects of SCS watermarking were discussed, where we focus on the performance of SCS watermarking facing an AWGN attack. The results for this attack can be translated directly into that for a GTC attack, which was shown to be an optimal attack for white Gaussian host data. We showed that for high-rate watermarking, e.g., useful in information-hiding applications, SCS combined with coded modulation achieves a rate of 1 bit/element at significantly lower WNR than the scheme proposed by Chou et al.[3] For low-rate watermarking, the performance of spread-transform (ST-)SCS watermarking and SCS watermarking with repetition coding was compared. ST-SCS watermarking turned out to be superior. This behavior results because, in SCS with repetition coding, the multiple representations of codewords for a single watermark message produce many nearest neighbors in the codeword space. For positive WNR, the spread transform does not give a gain over SCS without spread transform. For negative WNR the optimal spreading factor τ_{opt} is such that an effective WNR of about 0 dB is achieved in the spread-transform domain. Further, simulation results show that with turbo coding, performance close to the achievable rate of SCS watermarking can be obtained. Finally, the effect of limited codeword length N_c on the performance of SCS watermarking was analyzed theoretically with help of Gallager's random coding exponent, and practically with simulations for turbo codes with different interleaver sizes N_i and convolutional codes with different memory lengths ν. Particularly for negative WNRs, short codewords lead to a significant loss of achievable watermark rate.

7. ACKNOWLEDGEMENTS

The authors thank Marco Breiling, Robert Fischer, and Simon Hüttinger for their valuable support with respect to turbo coding, coded modulation, and choice of convolutional codes.

REFERENCES

1. 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.
2. J. J. Eggers, J. K. Su, and B. Girod, "Robustness of a Blind Image Watermarking Scheme," in *Proceedings of IEEE International Conference on Image Processing (ICIP 2000)*, (Vancouver, Canada), September 2000.
3. J. Chou, S. Pradhan, L. E. Ghaoui, and K. Ramchandran, "A robust optimization solution to the data hiding problem using distributed source coding principles," in *Proc. of SPIE Vol. 3974: Image and Video Communications and Processing 2000*, (San Jose, Ca, USA), January 2000.
4. B. Chen and G. Wornell, "Preprocessed and postprocessed quantization index modulation methods for digital watermarking," in *Proc. of SPIE Vol. 3971: Security and Watermarking of Multimedia Contents II*, pp. 48–59, (San Jose, Ca, USA), January 2000.
5. P. Moulin and J. A. O'Sullivan, "Information-Theoretic Analysis of Information Hiding." Preprint, September 1999.
6. M. H. M. Costa, "Writing on Dirty Paper," *IEEE Transactions on Information Theory* **29**, pp. 439–441, May 1983.
7. G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Transactions on Information Theory* **28**, pp. 55–67, January 1982.
8. J. G. Proakis, *Digital signal processing, principles algorithms and applications*, McGraw-Hill, New York, 2nd ed., 1989.
9. H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Transactions on Information Theory* **23**, pp. 371–377, May 1977.
10. U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel Codes: Theoretical Concepts and Practical Design Rules," *IEEE Transactions on Information Theory* **45**, pp. 1361–1391, July 1999.
11. C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding," *IEEE Transactions on Communications* **44**, pp. 284–287, October 1996.
12. B. Vucetic and J. Yuan, *Turbo Codes: Principles and Applications*, Kluwer Academic Press, Boston, Dordrecht, London, 2000.
13. B. Chen and G. W. Wornell, "Achievable performance of digital watermarking systems," in *Proceedings of the IEEE Intl. Conference on Multimedia Computing and Systems*, vol. 1, pp. 13–18, pp. 13–18, (Florence, Italy), June 1999.
14. S. Dolinar and D. Divsalar, "Weight distributions for turbo codes using random and nonrandom permutations," *TDA Progress Report* **42-122**, pp. 56–65, August 1995.
15. R. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, New York, 1968.