submitted to
4th Intl. ITG Conference
on Source and Channel Coding
Berlin, Jan. 28-30, 2002

A Channel Model for Watermarks Subject to Desynchronization Attacks

Robert Bäuml, Joachim J. Eggers and Johannes Huber
Telecommunications Laboratory
University of Erlangen-Nuremberg
Cauerstr. 7/NT, 91058 Erlangen, Germany
{baeuml,eggers,huber}@LNT.de

***Abstract* —**

Digital watermarking has been proposed for the copyright protection of multimedia documents, e.g., natural images, audio, video. One important unsolved problem in the field of digital watermarking is the synchronization of the watermark receiver after D/A-A/D conversion or local desynchronization attacks, e.g. StirMark's random bending attack. In this paper, we present a channel model for watermark reception with non-perfect synchronization. Under the given channel model the capacity of SCS watermarking and imperfectly synchronized reception is derived. One important result is that the capacity of SCS watermarking depends on the document power as long perfect synchronization is not achieved. For realistic document-to-watermark power ratios, synchronization errors up to 10 % of the sampling interval can be accepted without too significant losses off the SCS watermark capacity.

I. Introduction

Digital watermarking is the art of communicating information, "the watermark message," by embedding it into digital multimedia documents, called "host documents" or "host signals," to produce "marked signals." The embedded *watermark* should be reliably decodable even after further processing of the marked data, which is also denoted as *attack* against the embedded watermark. Such processing can be simple D/A-A/D conversion of the document, but also a malicious attempt to impair watermark reception. Digital watermarking has gained a lot of attention in the recent years for it's potential in several areas like proof of ownership and copyright enforcement. For instance, the embedded watermark can provide information about the copyright holder of a document or indicate the copy-state of the digital content.

The research community has come up with a vast variety of watermarking algorithms for different types of multimedia data, e.g., natural images, audio, video. Depending on the data type, watermark embedding is implemented in the spatial or time domain, or in transform domains like the DFT/DCT-spectrum or a wavelet domain. The constraints of a certain application for digital watermarking must be taken into account during the design of a watermarking scheme. One important aspect is the availability of the original document at the watermark receiver. In many applications, the original document cannot be used during watermark reception, which is denoted as *blind watermark reception* or more generally *blind watermarking*. Blind watermarking is considered throughout this paper.

We consider digital watermarking as a communication problem, where the watermark communication channel is characterized by possible attacks against the embedded watermark. A complete characterization of the watermark channel is currently not available, though theoretical analyses of specific attack scenarios have been published within the last two years, e.g., [5, 10, 13, 9]. One specifically interesting attack is the addition of white Gaussian noise (AWGN), since the analysis of extended attack scenarios can often be based on the analysis of the AWGN attack. Further, the AWGN attack can be applied easily so that each watermarking scheme should show good robustness at least against this type of attack. The design of watermarking schemes facing AWGN attacks and the resulting watermark capacity is reviewed in Section II. In particular, the *Scalar Costa Scheme* (SCS) watermarking [6, 7] is described, which is currently the most powerful practical blind watermarking scheme in terms of watermark capacity in the case of AWGN attacks.

The goal of this paper is to extend the characterization of the watermark attack channel with respect to *desynchronization attacks*. During the analysis of AWGN attacks, it is assumed that the watermark receiver can look for the watermark information exactly at the same position where it has been embedded, which is denoted as *perfectly synchronized reception*. However, in real-world scenarios, this assumption does not hold necessarily. It is even possible that an attacker intentionally modifies the watermarked document in order to desynchronize the watermark receiver. Note that, for simplicity, the term "synchronization" is used here in a quite general way, although, in a strict sense, synchronization is only relevant for time depending data. A more detailed description of possible desynchronization attacks and the state-of-the-art in solving the synchronization problem for watermark receivers is given in Section III. Next, we present a model for im-

perfectly synchronized watermark reception in Section IV. Based on this model, we analyze in Section V the watermark capacity of SCS watermarking depending on the synchronization accuracy. Section VI concludes the most important results and gives an outlook on future research on desynchronization attacks.

II. Blind Watermarking Facing AWGN Attacks

We consider digital watermarking a communications problem which can be described as follows: The encoder derives from the *watermark message* m and the host signal $\mathbf{x}$ an appropriate watermark signal $\mathbf{w}$ which is added to the host signal to produce the watermarked signal $\mathbf{s}$. $\mathbf{w}$ must be chosen such that the distortion between $\mathbf{x}$ and $\mathbf{s}$ is negligible. Next, the watermarked signal $\mathbf{s}$ might be processed, which gives a signal $\mathbf{r}$. Such processing potentially impairs watermark communication and thus is denoted as an *attack* against the embedded digital watermark. In general, attacks against digital watermarks are only constrained with respect to the distortion between $\mathbf{x}$ and $\mathbf{r}$. Finally, the receiver must be able to decode the watermark message from the received (attacked) signal $\mathbf{r}$. Both, encoding and decoding, depend on a key sequence $\mathbf{k}$, which ensures that only authorized parties can embed, decode, and modify the embedded watermark message m. Fig. 1 depicts the described blind watermark communication scenario, where an attack by an additive white Gaussian noise (AWGN) signal $v \sim \mathcal{N}(0, \sigma_v^2,)$ is assumed. Further, the analysis is constrained to independent identically distributed (IID) Gaussian original signals $x \sim \mathcal{N}(0, \sigma_x^2)$. In this paper, $\mathbf{x}$,$\mathbf{w}$,$\mathbf{s}$,$\mathbf{r}$ $\mathbf{v}$ and $\mathbf{k}$ are vectors, and $x[n]$,$w[n]$,$s[n]$,$r[n]$,$v[n]$ and $k[n]$ refer to their respective nth elements.

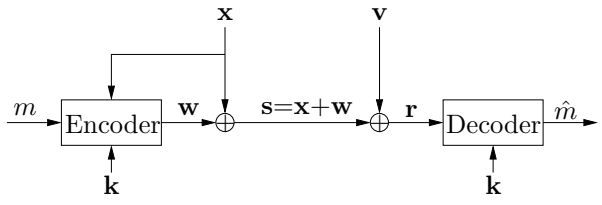

Fig. 1: Blind watermarking facing an AWGN attack.

Up to now, the most popular watermark embedding technique is based on the addition of a watermark signal $\mathbf{w}$ which is chosen independently from the host signal $\mathbf{x}$. Here, we assume a Gaussian watermark signal with $v \sim \mathcal{N}(0, \sigma_w^2,)$. This watermarking technique is also denoted as *spread-spectrum* (SS) watermarking, a term derived from spread-spectrum communication, although used in a slightly different way. For blind SS watermark reception, the unknown host signal $\mathbf{x}$ is considered as unavoidable interference. The watermark capacity of SS watermarking for Gaussian host signals and AWGN attacks is $C = 0.5 \log_2(1 + \sigma_w^2/(\sigma_x^2 + \sigma_v^2))$ bit/sample, which can be easily derived from the capacity of an AWGN channel [3]. Unfortunately, in realistic watermarking scenarios we have $\sigma_x^2 \gg \sigma_w^2$ to ensure imperceptibility of the watermark signal. Thus, the capacity of SS watermarking is limited by huge host-signal interference.

In 1999, it has been realized that blind watermarking can be considered communication with side information at the encoder [4, 1], which is obvious from the block diagram in Fig. 1. Costa [2] showed theoretically that for a Gaussian host signal of power σ_x^2, a watermark signal of power σ_w^2, and AWGN of power σ_v^2 the maximum rate of reliable communication (capacity) is $C = 0.5 \log_2(1 + \sigma_w^2/\sigma_v^2)$ bit/sample, independent of σ_x^2. The result is surprising since it shows that the host signal $\mathbf{x}$ need not be considered as interference at the decoder although the decoder does not know $\mathbf{x}$.

Costa's ideal scheme involves a *random* codebook which must be available at the encoder and the decoder. Unfortunately, for good performance the codebook must be so large that neither storing it nor searching it is practical. Thus, for practical application, the random codebook is replaced by a structured codebook, in particular a product codebook of dithered uniform scalar quantizers. The such simplified scheme is denoted as *Scalar Costa Scheme* (SCS) [6, 7]. This paper focusses on SCS watermarking, so that a brief review of the basic principle is given in the following.

For SCS watermarking, the watermark message m is encoded into a sequence of watermark letters $\mathbf{d}$, where $d[n] \in \mathcal{D} = \{0, 1\}$ in case of binary SCS. Each of the watermark letters is embedded into the corresponding host elements $x[n]$. The embedding rule for the nth element is given by

$$\begin{aligned} a[n] &= \Delta\left(\frac{d[n]}{2} + k[n]\right) \\ s[n] &= x[n] + \alpha\left(\mathcal{Q}_\Delta\left\{x[n] - a[n]\right\}\right) \\ &+ \alpha\left(a[n] - x[n]\right), \end{aligned} \tag{1}$$

where $\mathcal{Q}_\Delta\{\cdot\}$ denotes scalar uniform quantization with step size Δ. The key $\mathbf{k}$ is a pseudo-random sequence with $k[n] \in (0, 1]$. 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 and detection reliability for a given noise variance of an AWGN attack. Optimal values for Δ and α are given in [6].

Watermark decoding from the received signal $\mathbf{r}$ is based on the pre-processed received signal $\mathbf{y}$. The extraction rule for the nth element is

$$y[n] = \mathcal{Q}_\Delta\left\{r[n] - k[n]\Delta\right\} + k[n]\Delta - r[n], \tag{2}$$

where $|y[n]| \leq \Delta/2$. $y[n]$ should be close to zero if $d[n] = 0$ was sent, and close to $\pm\Delta/2$ for $d[n] = 1$.

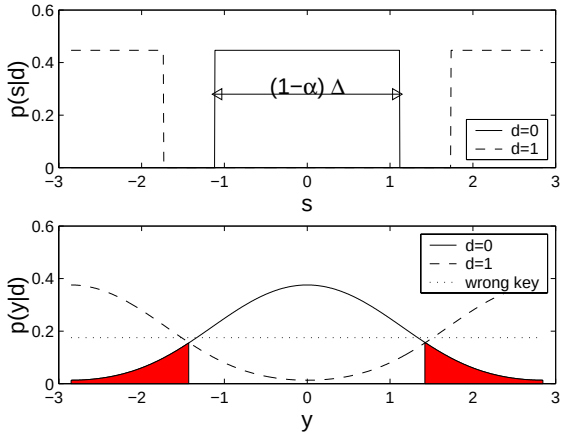

Fig. 2: One period of the PDFs of the transmitted and the received signal for binary SCS (σ_w^2=1, WNR = 2 dB, $\Delta = 5.7$, $\alpha = 0.61$). The filled areas represent the probability of detection errors assuming $d = 0$ was sent. The dotted line in the lower plot depicts the PDF when detecting with a wrong key $\mathbf{k}$.

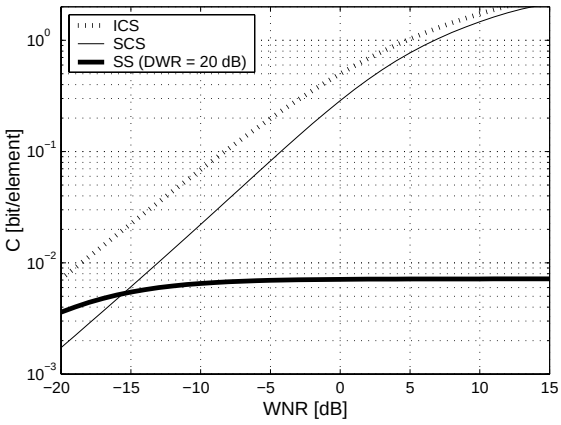

Fig. 3: Watermark capacities in case of AWGN attacks for the Ideal Costa Scheme (ICS), the Scalar Costa Scheme (SCS), and spread-spectrum (SS) watermarking

The basic properties of binary SCS watermarking can be demonstrated by the probability density function (PDF) of the transmitted signal $\mathbf{s}$ and the PDF of the extracted signal $\mathbf{y}$. Note that conditioning on the key sequence $\mathbf{k}$ is assumed in the following. The upper plot of Fig. 2 depicts one period of the PDF of the transmitted samples s conditioned on the sent watermark letter d. For $d = 0$, the transmitted value s is concentrated around integer multiples of Δ. Contrary, for $d = 1$, s is concentrated around $\Delta/2$ plus integer multiples of Δ. The lower plot shows the PDF of the extracted samples y after AWGN attack conditioned on the sent watermark letter d. $p_y\left(y[n]|d[n]\right)$ is computed numerically as described in [6]. We observe that the PDFs of y in case of $d = 0$ and in case of $d = 1$ can be still distinguished. Note that in the case of an incorrect key $\mathbf{k}$ at the receiver, the distribution of $p_y\left(y[n]|d[n]\right)$ will be uniform for any possible $\mathbf{r}$. This is indicated by the dotted line in the lower plot of Fig. 2.

Fig. 3 shows the watermark capacities for the mentioned watermarking schemes, namely the ideal Costa scheme (ICS), SCS, and SS and for AWGN attacks with varying *Watermark-to-Noise power Ratio* (WNR) of WNR $= 10\log_{10}\sigma_w^2/\sigma_v^2$ [dB]. Only SS watermarking suffers from host-signal interference, which limits the achievable capacity. The shown capacity of SS watermarking is for the realistic *Document-to-Watermark power Ratio* (DWR) of DWR $= 10\log_{10}\sigma_x^2/\sigma_w^2 = 20$ dB. SCS watermarking does not achieve the capacity of an ideal Costa scheme, but comes close to that for a large range of practically relevant WNRs. In particular, SCS watermarking achieves significantly larger watermark capacities than blind SS watermarking for WNR > -15 dB.

III. Desynchronization Attacks

The analysis of desynchronization attacks against digital watermarks and the development of efficient counter-attacks is still one of the most demanding problems in the field of digital watermarking. In this section, we illustrate the problem of desynchronization attacks in the case of watermarked image data, and give a brief overview of the state-of-the-art in this research area.

Desynchronization attacks have been a problem for a considerable time, especially in the field of image watermarking. Early desynchronization attacks consisted of rather simple global affine transformations. Robustness against such global desynchronization attacks can be achieved by watermark embedding into transform invariant domains. For instance, watermark embedding in the log-polar domain enables robustness against rotation, translation and scaling of the watermarked image [8]. Further, global affine transformations can be estimated relative easily due to the small number of free attack parameters. The estimation of these parameters is usually based on a known embedded synchronization pattern, where the estimation accuracy increases with the image size.

One of the most popular software tools for attacks on image watermarks is the StirMark package [12], which offers a wide range of different attacks to render watermark extraction hard to impossible. One of the most effective attacks within StirMark is the random bending attack which exloits the inability of the human eye to detect small local geometric distortions. For this attack, a smooth transformation of the sampling grid is applied which desynchronizes a simple watermark detector. Thus, preprocessing prior to standard watermark detection is required to enable watermark detection. An example transformation is presented in Fig. 4, where the transformation is obvious from the prior knowledge

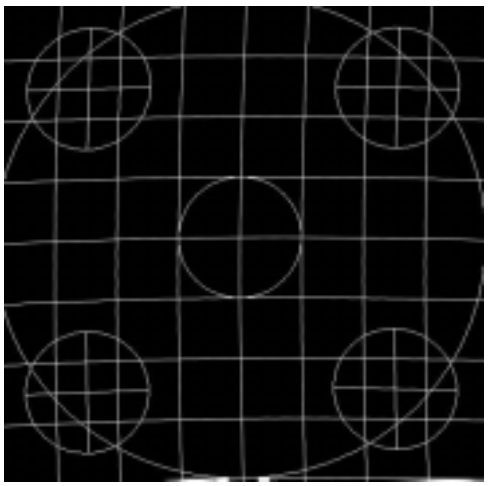

Fig. 4: StirMark applied to a regular grid

of the viewer about the image. The same transform imposed on a natural image yields negligible visual artefacts, while synchronization of standard watermark detectors is destroyed. The human visual perception is a complex process which takes a lot of prior knowledge about the image into account. Thus, it is difficult to find an objective quality measure which reflects the subjective quality loss due to the random bending attack. Note that the popular mean-squared error (MSE) distortion measure is not appropriate within the context of desynchronization attacks.

Counter-attacks against the StirMark random bending attack have been investigated mainly for non-blind watermarking, e.g, [11], where the knowledge of the original image can be exploited to achieve synchronization of the watermark detector. A promising approach for blind watermarking is based on a model for the local transformations, e.g., local affine transformations, where a synchronization pattern is used to estimate the model parameter. As for global transform models, the synchronization accuracy increases with the number of pixels available for the parameter estimation. In practice, it is highly unlikely that the original sampling grid can be reconstructed perfectly. Therefore, we investigate in this paper the influence of inaccurately synchronized watermark detection. We assume that resynchronization has been performed on the received data so that only a jitter in the sampling grid remains as the effective distortion. All effects are viewed in the coordinate domain, where warping effects can be handeled easiest.

It has to be noted that desynchronization attacks are also applicable to other media, e.g. audio data, though the specific attack model may need to be adapted to the given media type. For instance, the amount of subjectively acceptable local modifications of the sampling grid may differ significantly between image data and audio data.

Synchronization is also a major issue in communications, especially in wireless communication, and has been solved satisfactory for current applications. Unfortunately, the methods developed in these fields cannot be easily transferred to the synchronization problem in digital watermarking. Typical synchronization problems have been solved for proper models of specific transmission channels. Such models are still missing for desynchronization attacks against digital watermarks. One major problem is that the attacker has many degrees of freedom to implement desynchronization attacks and at the same time has malicious intent.

IV. A Channel Model for Desynchronization Attacks

In this section, a channel model for imperfectly synchronized watermark detection is developed. We assume that coarse resynchronization as been applied, e.g., based on the estimation of parameters of local transforms using an embedded synchronization pattern. The artefacts of imperfect resynchronization are similar across all resynchronization methods in the sense that the estimated sampling grid generally contains a certain deviation from the original sampling grid. For simplicity, one-dimensional signals are considered subsequently. The extension to multi-dimensional signals, e.g., image or video data, is straigt forward. The developed model gives insights into the principle limits of watermark detection after desynchronization attacks.

Let $s[n] = x[n] + w[n]$ denote the discrete watermarked signal. This signal corresponds to the critically sampled continuous signal $s(t)$, which is bandlimited to $f_G = 2/T$, where T denotes the width of one sampling interval. Then, $\hat{s}[n] = s(nT + T_\Delta)$ denotes the resampled signal, where an offset of T_Δ in the sampling grid has been introduced. Assuming ideal interpolation, $\hat{s}[n]$ can be computed from $s[n]$ with

$$\hat{s}[n] = \sum_{\nu=-\infty}^{\infty} s((n+\nu)T) \cdot \mathrm{sinc}\left(\nu T + T_\Delta\right), \qquad (3)$$

where $\mathrm{sinc}\,(x) = sin(\pi x)/\pi x$. Further signal distortions due to attack operations are described by an additive noise source $v[n]$ with power σ_v^2, so that the received attacked signal is given by

$$r[n] = \hat{s}[n] + v[n]. \qquad (4)$$

The described channel model is depicted in Fig. 5.

Note that considering only a constant sampling offset by T_Δ is not very restrictive. The model can be extended without difficulties to a sampling offset $T_\Delta[n]$ so that $\hat{s}[n] = s(nT + T_\Delta[n])$. However, in this paper we focus on a constant offset T_Δ which gives already important insights concerning the required resynchronization accuracy for watermark detection.

Next, the nth received signal sample $r[n]$ is decomposed into a component derived from the nth

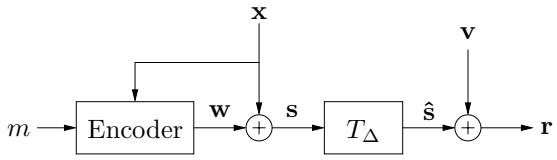

Fig. 5: Channel model for a desynchronization attack

watermarked sample $s[n]$ and additional contributions from samples $s[n+\nu]$, with $\nu \neq 0$, which gives

$$\begin{aligned} r[n] &= \sum_{\nu=-\infty}^{+\infty} s((n+\nu)T) \cdot \operatorname{sinc}(\nu T + T_\Delta) + v[n] \\ &= \underbrace{(w[n] + x[n]) \cdot \operatorname{sinc}(T_\Delta)}_{\hat{s}_e[n]} + \\ &+ \underbrace{\sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{+\infty} s((n+\nu)T) \cdot \operatorname{sinc}(\nu T + T_\Delta)}_{\hat{s}_z[n]} \\ &+ v[n] \end{aligned} \tag{5}$$

$\hat{s}_e[n]$ corresponds to the information bearing signal component, and $\hat{s}_z[n]$ describes *Inter-Symbol-Interference* (ISI). In common communication scenarios without side-information at the encoder, ISI by $s[n+\nu]$, for $\nu \neq 0$, can in principle be utilized to improve the detection performance. However, when exploiting side-information at the encoder, as in Costa's scheme or its practical version SCS, little is known about possible exploitations of ISI. Thus, we assume that ISI is unavoidable interference for SCS watermark detection. Further, we assume in the following that the watermarked signal is white and Gaussian distributed with a power of $\sigma_s^2 = \sigma_x^2 + \sigma_w^2$.

As we can derive from $\hat{s}_e[n]$, the signal bearing component in our model, containing $w[n]$, is attenuated by $\operatorname{sinc}(T_\Delta)$. Thus, we can determine the power $\sigma_{\hat{w}}^2$ of the attenuated watermark $\hat{w}[n]$ after the warping operation by

$$W(T_\Delta) = \sigma_{\hat{w}}^2 = \sigma_w^2 \cdot \operatorname{sinc}(T_\Delta)^2 .$$

In turn, the resulting noise power $N(T_\Delta)$ contains now the ISI term from $\hat{s}_z[n]$ and the AWGN $v[n]$:

$$N(T_\Delta) = \sigma_s^2 \cdot (1 - \operatorname{sinc}(T_\Delta)^2) + \sigma_v^2.$$

In the following, we investigate our channel model for $\sigma_v^2 = 0$ and $\sigma_v^2 = \sigma_w^2$. Fig. 6 shows the resulting effective watermark-to-noise power ratio $10 \cdot \log_{10}(W/N)$ for DWR = 15 dB, 20 dB, and 25 dB. We observe that the power of the AWGN $v[n]$ does not play a dominant role if the relative sampling offset T_Δ/T is larger than about 0.1 to 0.3, depending on the DWR. Further, a significant influence of the DWR appears. This result is a consequence of the assumption that the entire signal $\hat{s}_z[n]$ is unavoidable noise.

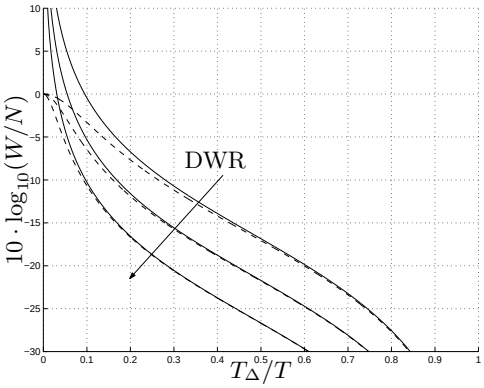

Fig. 6: W/N of $\hat{s}[n]$ depending on the sampling deviation T_Δ. The depicted results are for $\sigma_v^2 = 0$ (–) and $\sigma_v^2 = \sigma_w^2$ (- -) and for DWR = 15 dB, 20 dB, 25 dB.

Another interesting value is the *Attack to Interference Power Ratio* (AIR), which is defined as

$$AIR = \frac{\sigma_v^2}{\sigma_s^2 \cdot (1 - \operatorname{sinc}(T_\Delta)^2)}, \tag{6}$$

and represents the influence of the noise produced by the sampling jitter compared to the AWGN $\mathbf{v}$. As we can derive from Fig. 7, very little deviation from the exact sampling grid produces relatively large disturbance. In the best shown case, where the DWR is 15 dB, a deviation of 10% from the exact grid leads to additional noise from ISI which is as strong as the AWGN. In cases of larger DWRs, the influence of ISI increases and thus even more accurate resynchronization is required.

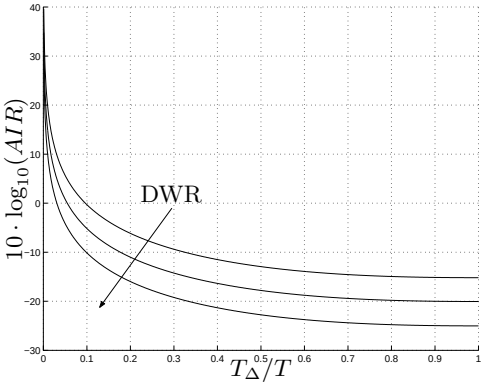

Fig. 7: Attack-to-Interference ratio depending on the sampling deviation T_Δ. The result is depicted for DWR = 15 dB, 20 dB, 25 dB.

V. Watermark Capacity for Imperfectly Synchronized Reception

In Section II, SCS watermarking has been introduces as a powerful blind watermarking technology. Significant gains over state-of-the-art SS water-

marking are predicted due to the host-signal independence of blind SCS watermarking. However, the described channel model for imperfectly synchronized watermark detection shows that the strength of ISI interference is strongly dependent on the host signal, in particular on the DWR. In SCS, the side-information about the host signal $\mathbf{x}$ at the encoder is exploited in a quite simple way. That is, the watermark sample $w[n]$ is chosen such that interference from $x[n]$ during blind watermark detection vanishes or is negligible at least. The influence of samples $x[n+\nu]$, for $\nu \neq 0$, which contribute strongest to the total ISI, is not considered during SCS watermark embedding. Thus, the performance of SCS in case of desynchronization attacks is no longer host signal independent. As soon as there is a desynchronization attack and this attack cannot be reversed perfectly, SCS suffers from host signal interference similar to SS watermarking. Here, the capacity of SCS watermarking after AWGN and desynchronization attacks is derived using the model described in Section IV. This allows us to investigate the remaining advantage of SCS over SS watermarking,

We assume that the ISI $\hat{s}_z[n]$ has a Gaussian distribution, which is reasonable for a white and Gaussian host signal $\mathbf{x}$. Then, the capacity of SCS watermarking after AWGN and desynchronization attacks can be obtained from the the capacity of SCS watermarking facing a simple AWGN attacks using the effective watermark-to-noise ratio $10 \cdot \log_{10}(W/N)$ as derived in Section IV. Fig. 8 shows the resulting capacity curves for $\sigma_v^2 = 0$ and $\sigma_v^2 = \sigma_w^2$, and for three different DWRs. As already suggested by the analysis in Section IV, a relative deviation T_Δ/T of more than 10 % results in a significant reduction of the SCS watermark capacity. ISI dominates the influence of AWGN for large T_Δ/T.

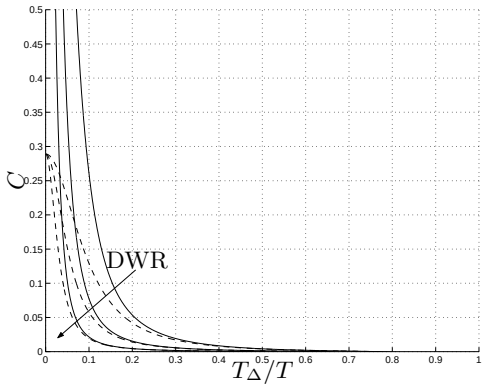

Fig. 8: Capacity of SCS in bit/sample under the influence of Inter-Symbol-Interference due to a sampling-grid deviation. The depicted results are for $\sigma_v^2 = 0$ (–) and $\sigma_v^2 = \sigma_w^2$ (- -) and for DWR = 15 dB, 20 dB, 25 dB.

Next, the influence of variable AWGN power σ_v^2 for differently fixed grid shifts T_Δ and DWR = 20 dB is investigated. Fig. 9 shows the capacity of SCS watermarking after AWGN attacks with WNR$=-20,\ldots 5$ dB and grid shifts $T_\Delta/T = 0, 0.05, 0.1$, and 0.2. We observe that ISI is less important for strong AWGN attacks since in these cases the AWGN dominates ISI for larger grid shifts. The SCS watermark capacity is still reasonably high, even under moderate synchronization errors up to $T_\Delta = T/10$. The comparison to SS watermarking with perfect synchronization shows that only for very strong AWGN attacks (WNRs below -15 dB) SCS watermarking with imperfect synchronization ($T_\Delta/T \approx 0.1$) performs worse than SS watermarking.

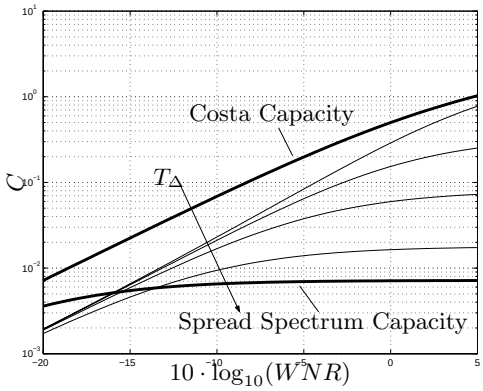

Fig. 9: Capacity of SCS watermarking under the influence of a sampling deviation T_Δ and AWGN with different WNR. The depicted results are for DWR = 20 dB and T_Δ/T=0, 0.05, 0.1, 0.2, as indicated by the arrow. The upper broad curve represents the channel capacity for an ideal Costa watermarking algorithm with perfect synchronization; the lower broad curve is the capacity for SS watermarking with perfect synchronization.

VI. Conclusions and Future Research

Robust watermark detection after desynchronization attacks is still an important open problem in the field of digital watermarking. In this paper, a channel model for imperfectly synchronized watermark detection has been investigated. The focus of our analysis is on blind scalar Costa scheme (SCS) watermarking, which is for perfectly synchronized detection independent from the host signal statistics and thus outperforms the popular spread-spectrum (SS) watermarking by far. We observed that SCS suffers from inter-symbol-interference (ISI) in case of imperfectly synchronized watermark detection. This is especially true for large document-to-watermark power ratios (DWRs), where ISI dominates other attack distortions. Thus, the property of SCS being host signal independent is no more true under desynchronization attacks. We investi-

gated the SCS watermark capacities after AWGN attacks and imperfect resynchronizion. One important result is that, for realistics DWRs, a synchronization error up to 10 % of the sampling interval is acceptable. For such accurate resynchronization, SCS watermarking performs for weak to medium-strong attacks still significantly better than SS watermarking. Nevertheless, our analysis highlights the fact that very exact resynchronization plays a major role for this watermarking method to keep up a reasonable watermark capacity.

The presented model for desynchronization attacks is very simple, but gives already significant insights into the synchronization problem of digital watermarking. In future, the model has to be extended to include random synchronization errors. Further, amplitude synchronization before SCS watermark detection has to be investigated as well.

VII. 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)*, volume 3845, pages 43–54, Boston, MA, USA, September 1999.

[2] M. H. M. Costa. Writing on dirty paper. *IEEE Transactions on Information Theory*, 29(3):439–441, May 1983.

[3] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York, 1991.

[4] 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(7):1127–1141, July 1999.

[5] J. J. Eggers and B. Girod. Quantization effects on digital watermarks. *Signal Processing*, 81(2):239–263, February 2001.

[6] 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*, pages 4/1–4/6, London, UK, April 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. Kutter. Watermarking resisting to translation, rotation and scaling. In *Proc. of SPIE: Multimedia systems and application*, volume 3528, pages 423–431, Boston, USA, November 1998.

[9] P. Moulin and M. K. Mihçak. The data-hiding capacity of image sources. preprint, June 2001.

[10] P. Moulin and J. A. O'Sullivan. Information-theoretic analysis of information hiding. Preprint, September 1999.

[11] I. B. Ozer, M. Ramkumar, and A. N. Akansu. A new method for detection of watermarks in geometrically distorted images. In *Proceedings of the IEEE Intl. Conference on Speech and Signal Processing 2000 (ICASSP 2000)*, Istanbul, Turkey, June 2000.

[12] F. A. P. Petitcolas and M. G. Kuhn. StirMark image watermark benchmark software. Technical report, available at http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/, October 1998.

[13] J. K. Su, J. J. Eggers, and Bernd Girod. Analysis of digital watermarks subjected to optimum linear filtering and additive noise. *Signal Processing, Special Issue on Information-Theoretic Issues in Digital Watermarking*, 81(6), June 2001.