Authors:

Reza Samadi (M'2011) is graduated from Sharif University of technology in 2006. Now he is pursuing his PhD in Ferdowsi University of Mashhad. His research interest covers information security, physical layer security, and applications of information theory.
(Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran. Post code: 91779-48944, e-mail: rsamadi@ieee.org)

Seyed Alireza Seyedin was born in Mashhad. He received the B.S. degree in Electronics Engineering from Isfehan University of Technology, Isfehan, Iran in 1986, and the M.E. degree in Control and Guidance Engineering from Roorkee University, Roorkee, India in 1992, and the Ph.D. degree from the University of New South Wales, Sydney, Australia in 1996. He has been an Associate Professor with the Department of Electrical Engineering, the Ferdowsi University of Mashhad, Mashhad, Iran. His research interest includes image processing, computer vision, signal processing, and pattern recognition. In these fields, specially, he is interested in image analysis, motion detection and estimation in image sequences, autonomous vehicles, and diverse applications of the radon transform.
(Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran. Post code: 91779-48944, e-mail: Seyedin@um.ac.ir)

Improving the Security and Robustness of the Scalar Costa Scheme

***Abstract*—**Unintentional attacks on watermarking schemes lead to degrade the watermarking channel, while intentional attacks try to access the watermarking channel. Therefore, watermarking schemes should be robust and secure against unintentional and intentional attacks respectively. Usual security attack on watermarking schemes is the Known Message Attack (KMA). Most popular watermarking scheme with structured codebook is the Scalar Costa Scheme (SCS). The main goal of this paper is to increase security and robustness of SCS in the KMA scenario. To do this, SCS model is extended to more general case. In this case, the usual assumption of an infinite Document to Watermark Ratio (DWR) is not applied. Moreover watermark is assumed to be an arbitrary function of the quantization noise without transgressing orthogonality as in the Costa's construction. Also, this case is restricted to the structured codebooks. The fundamental trade-off is proved between security and robustness of Generalized SCS (GSCS) in the KMA scenario. Based on this trade-off and practical security attack on SCS, a new extension of SCS is proposed which is called Surjective-SCS (SSCS). In the absence of robustness attack, the SSCS has more security than SCS in the same DWR. However, the SSCS achieves more security and robustness than SCS only in low Watermark to Noise Ratio (WNR) regime or low rate communications.

***Index Terms*—**Achievable rate, Flat-host assumption, Known Message Attack, Scalar Costa Scheme, security, Trade-off, Watermarking.

I. INTRODUCTION

A. Motivation

The rapid development of digital information technology has made multimedia distribution over public networks easy and popular. However, it is along with some serious problems such as unlimited duplication, illegally redistribution etc. To prevent these illegal operations, the watermarking is used to hide a verification message into multimedia. Watermarking schemes should be robust and secure against unintentional and intentional attacks respectively. The popular theoretic measure to evaluate the robustness is the achievable rate over AWGN attack channel. Communication model of digital watermarking is well fitted to problem of communication with side information at the encoder [1]; that in the case of independent and identically distributed (i.i.d), Gaussian host (side information) is reduced to the well-known Costa Scheme [2]. Deterministic and structured implementation of the Costa Scheme for digital watermarking is the quantization-based schemes which are suboptimal implementation. Lattice-based data hiding, which is known as Distortion Compensation-Dither Modulation (DC-DM) [3], is the simplest dither quantizer; and Scalar DC-DM, which is known as Scalar Costa Scheme (SCS) [4], is the most popular one which is used in practice for its simplicity and good performance.

In analyzing performance of quantization-based data hiding, the fundamental assumption is that the host pdf can be approximated to be flat in each quantization cell [3-18] (hereafter, flat-host assumption). This implies an infinite document to watermark ratio (DWR). When this approximation is not valid, Quantization-based data hiding schemes offer larger achievable rate versus Spread Spectrum (SS) schemes [19] for any Watermark to Noise Ratio (WNR). However, these schemes have lower security [5-6], so there is a demand for more secure quantization-based data hiding schemes. Our main motivation is to increase security of scalar

quantization-based data hiding scheme, hold or improve its achievable rate.

B. Literature Review

The security of a symmetric watermarking scheme (like SCS) may be enhanced by secure coding [7-8], host-dependent keys [9-10], host-dependent codebook (randomized codebook) [11-13] or secure embedding [14-18]. Although, secure coding is applicable to any watermarking schemes, it doesn't prevent the unauthorized removal attack [7]. Host-dependent keying suffers from key synchronization at the decoder side [10]. Also, Randomizing codebook increases the entropy (security) of the codebook by means of non-structured codebooks, so it makes the estimation more difficult, but it increases computational complexity too [12].

Secure embedding keeps the structure of the lattice codebook and increases the security by changing the law of embedding. One method is to make a secret rotation to the embedding lattice [14]. However it is easy to model secret rotation as secure coding followed lattice embedding. Other method is to use a Look-Up Table (LUT) along with SCS [15]. This method achieves less probability of error but more embedding distortion. Another method to increase security is the use of Spread Transform Dither Modulation (STDM) in conjunction with LUT [16]. However, Spread Transform and LUT techniques are applicable to any watermarking scheme. More recent works to increase the security of scs [17-18] are suggested to define the decoding region based on the distribution matching of the host signals to predefined watermarked host signals. However this method can only be perfectly secure against Watermark Only Attack (WOA) where, the adversary uses a pool of hosts watermarked with the same key. Nevertheless, in watermarking the adversary usually has access to a pool of independent messages and corresponding hosts watermarked with the same key which is called Known Message Attack (KMA).

C. Contribution

The important problem of secure embedding of SCS in KMA scenario is not addressed before in the literature which is the main scope of this paper. To do this, SCS model is extended to more general case. Then, the possibility of jointly increasing security and robustness of SCS is explored through this case. In this case, the flat-host assumption is not applied to make our analysis general. Moreover watermark is assumed to be an arbitrary function of quantization noise. However this case is restricted to the structured codebooks. The rationale behind considering this general case is the schemes in [18][20]. The differences between these schemes and this case are the ignoring of flat-host assumption and considering arbitrary law for watermark as a function of quantization noise.

The fundamental trade-off is proved between security and robustness of Generalized SCS (GSCS) in KMA scenario. As a special case, it is shown that the security level of SCS is much lower than before was derived. Then, a new extension of SCS is proposed which is called Surjective-SCS (SSCS). The rationale behind the developing of new scheme is resistance to previous estimation attack on SCS; and the form of trade-off between transparency-security and robustness of GSCS in KMA scenario. The SSCS achieves more security and robustness than SCS while it keeps transparency and computational cost.

D. Paper Organization

Formal definitions of digital watermarking criteria's are reviewed in section II. Section III provides the accurate security analysis of GSCS in KMA scenario, and fundamental trade-off between security and achievable rate. As a special case, the security analysis of SCS in KMA scenario is examined. New scheme is proposed in section IV with the proof of its superiority than SCS in term of security and achievable rate. Finally section V concludes the paper and gives some remarks and future research lines.

II. PROBLEM FORMULATION

The theoretical model of digital watermarking followed in this paper is shown in Fig. 1. This model is the same one which is used in [3]. Coefficients x_i are i.i.d sequences of scalar features which are extracted from original digital content by Discrete Cosine Transform, Discrete Wavelet Transform, Fast Fourier Transform or other spatial/temporal transformations. The embedder hides an equiprobable watermark message $m_i \in \{0,...,p-1\}$ in x_i using secret key dependent and deterministic embedder $f(.)$ with secret key k_i yielding a watermark w_i. Then watermark w_i is added with host x_i yielding watermarked host y_i. Watermarked host y_i undergoes channel attack and is added by AWGN noise n_i. Detector receives noisy watermarked host z_i and having k_i should estimate embedded message m_i.

Embedding distortion (watermark power) is computed as $D_w = E\{|w|^2\}$. Transparency is measured by host variance to watermark power ratio $\lambda = \sigma_x^2 / D_w$. Attack channel is parameterized by watermark power to noise variance ratio $\zeta = D_w / \sigma_n^2$. Based on λ and ζ, two parameters, DWR and WNR can be defined respectively as $DWR = 10\log_{10}(\lambda)$ and $WNR = 10\log_{10}(\zeta)$.

As a robustness measure, the achievable rate can be computed by maximizing mutual information over embedding function $f(.)$, while λ and ζ are fixed.

$$R(\lambda,\zeta) = \max_{f(.)} I(Z;M \mid K) \tag{1}$$

In security evaluation, the purpose of attacker is to disclose the secret parameters and then implement the tampering attack. According to Kerckhoffs' principle all details of the watermarking technique except the so-called secret key parameter of the embedding and decoding processes are publicly known. For evaluating security we use security level definition

from [21]. The security level of a secrecy system is said to be the effort that attacker requires for estimating the secret key. Also in this paper, we only concentrate on KMA scenario where the attacker has access to the pool of independent messages and corresponding watermarked host when all watermarked with the same secret key. By using residual entropy as information theoretic measure of security level [22], the γ-security level [5] is defined as the number of observation N_0 that attacker needs to holds inequality (2). The security level of SCS in KMA scenario by using flat-host assumption, is derived theoretically in [5] only for $\alpha_{SCS} \geq 0.5$ where α_{SCS} is distortion compensation parameter. In [5] authors show that the trade-off between security and achievable rate of SCS is controlled by α_{SCS}.

$$h(K \mid Y^{N_0}, M^{N_0}) \leq \gamma \tag{2}$$

In this paper, we propose a security and robustness analysis of Generalized SCS after removing restrictive assumptions like flat-host assumption, limited values of distortion compensation parameter and specific embedding law. Then, we propose a new scheme to achieve more security and robustness than SCS while it keeps transparency. An easy way to increase security is to decrease α_{SCS}. Although this choice decreases achievable rate, but as we prove in next section, security doesn't increase so much for small α_{SCS} and it is still large gap to security level of SS scheme. In recent work [18], author chooses an extension of SCS by changing the embedding law which achieves perfect secrecy in WOA scenario, while obtains more achievable rate than SCS. This result motivates us to analyze the relation between security and achievable rate of Generalized SCS in KMA scenario, and then design more secure and robust embedding law.

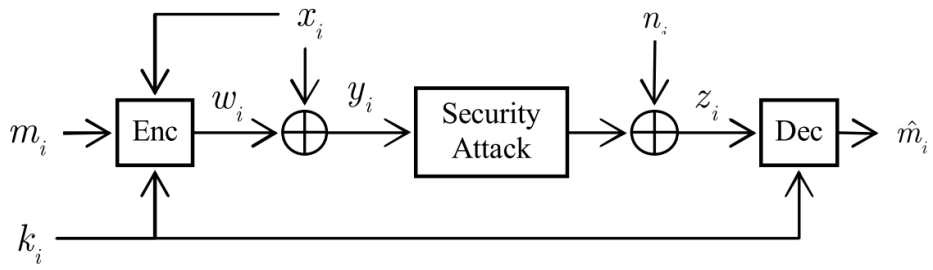


Fig. 1. Theoretical model for additive side-informed watermarking, including security attack and robustness attack/noisy channel

III. Security Analysis of General SCS

SCS can be extended into GSCS without transgressing orthogonality as in Costa's construction [2] and its implementation SCS [4]. Obviously, watermark can be an arbitrary function of quantization error and not just like SCS. Notice that in GSCS, watermark is still nearly orthogonal to x and host rejection is still possible. In GSCS, watermark can be written as

$$w = G(e_{m,k}(x)) = T(e_{m,k}(x)) - e_{m,k}(x) \tag{3}$$

where

$$\begin{aligned} e_{m,k}(x) &= x - \Delta\frac{m}{p} - k - Q_{\Delta}(x - \Delta\frac{m}{p} - k) \\ &\triangleq x - Q_{m,k}(x) \end{aligned} \tag{4}$$

And Q_{Δ} is uniform scalar quantizer over period $[-\Delta/2, \Delta/2]$, Δ is quantization step size and m is to-be-transmitted message symbol. Also $Q_{m,k}$ is shifted uniform scalar quantizer with its centroids distributed along the points which defined by a shifted lattice over scaled integer numbers.

$$\Lambda_{m,k} \triangleq \Delta\mathbb{Z} + \Delta\frac{m}{p} + k \tag{5}$$

The secret key k is dither and security of the embedder relies only on the randomization of the codebook via a dithering process. Previous works concerning on performance analysis of

SCS assume that secret key k is statistically distributed uniformly over period $[-\Delta/2, \Delta/2]$, therefore the Schuchman condition is satisfied [23] and error signal $e_{m,k}(x)$ is nearly orthogonal to x, m. In the remainder, we use this assumption to derive theoretical results.

Two variant of GSCS (else than SCS) are used before in literature. In [20], authors define $G(.)$ as a transform function from uniform distribution to Gaussian one, then they derive simple theoretical expression for probability of error and show that their scheme achieve lower error rate than SCS for a large WNR range. They call it Gaussian DC-DM (GDC-DM). In recent work [18], author model pdf of host and watermarked host using flat host assumption in the manner that scheme to be perfectly secure in WOA scenario, then obtain function $T(.)$ using optimal distribution matching and call it Soft-SCS. In both scheme GDC-DM and Soft-SCS, function $T(.)$ is nonlinear which increases computational complexity.

Power of watermark in GSCS can be computed using crypto lemma [24] by Forney as bellow:

$$D_w = \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} (T(e)-e)^2 de \tag{6}$$

For the sake of simplicity in deriving pdf of Watermarked host, we assume that $T(.)$ is odd function as in [4], [20] and [18]. Also without loss of generality, we assume that $|T(e)| \leq |e|$ holds, because if it doesn't hold, then we can translate the codebook (both host and watermarked host) by $\Delta/2$, so assumption will be hold for new embedder $T_1(e)$ as follows [18]:

$$T_1(e) = \Delta/2 - T(\Delta/2 - e) \tag{7}$$

For computing the pdf of watermarked host, we use same approach as used in [19] by taking into account that SCS can be thought of as a random variable transformation. Form (3) we

have equality

$$Q_{m,k}(y) = Q_{m,k}(x) + Q_{m,k}(T(e_{m,k}(x))) = Q_{m,k}(x) \tag{8}$$

Therefore inverse GSCS would be as

$$x = T^{-1}(y - Q_{m,k}(y)) + Q_{m,k}(y) \tag{9}$$

Now consider the following pdf transformation [25] resulting from (3).

$$p_Y(y \mid M = m, K = k) = \frac{p_X(x)}{T'(x - Q_{m,k}(x))} \tag{10}$$

Substituting (8) and (9) into (10) we have following pdf of watermarked host which is used frequently during this paper.

$$\begin{aligned} & p_Y(y \mid M = m, K = k) \\ & \quad = \frac{p_X(T^{-1}(y - Q_{m,k}(y)) + Q_{m,k}(y))}{T'(T^{-1}(y - Q_{m,k}(y)))} \end{aligned} \tag{11}$$

Now we are ready for analyzing the security of GSCS as the sense in (2). First we analyze security for $N_0 = 1$ observation which gives a simple closed form. Theoretical security analysis for $N_0 = 1$ observation helps us to compare it with previous theoretical result.

A. Residual Entropy of $N_0 = 1$ observation

Residual entropy of secret key conditioned on watermarked host and to-be-transmitted message can be written as follows:

$$\begin{aligned} h(K \mid Y, M) &= h(K) - I(K; Y, M) \\ &= h(K) - I(K; Y \mid M) \\ &= h(K) - h(Y \mid M) + h(Y \mid M, K) \end{aligned} \tag{12}$$

The first term is equal to $\log_2(\Delta)$. In order to compute the second term $h(Y \mid M)$, we need to know the pdf of watermarked host conditioned on message which can be computed via (11)

as follows. The proof is the same as one used in [23] plus some simplifications.

$$p_Y(y \mid M = m) = \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} p_X(y + r - T(r))dr \tag{13}$$

Obviously this conditioned pdf is independent of message $p_Y(y \mid M = m) = p_Y(y)$, i.e. in GSCS, watermarked host don't leak any information about to-be-transmitted message to attacker. Also, this form corresponds to previous exact result in [26], so it testifies (11).

The second term $h(Y \mid M, K)$ can be derived via (11) as follows. The proof is in Appendix-A.

$$h(Y \mid M, K) = \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \log(T'(e))de + h(X) \tag{14}$$

So residual entropy of secret key conditioned on watermarked host and to-be-transmitted message can be written as

$$\begin{aligned} h(K \mid Y, M) &= \log(\Delta) + h(X) - h(Y) \\ &+ \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \log(T'(e))de \end{aligned} \tag{15}$$

The term $h(X) - h(Y)$ for Gaussian host, can be bounded. Because x and w are independent (Schuchman condition) and $y = x + w$, by using Power Entropy Inequality (PEI) lemma [27] and some simplifications, we have:

$$h(X) - h(Y) \le -0.5 \log_2(1 + \frac{2^{2h(W)}}{2\pi e \sigma_x^2}) \tag{16}$$

Also by using maximum entropy lemma [27] we have:

$$h(X) - h(Y) \ge -0.5 \log_2(1 + \frac{D_w}{\sigma_x^2}) \tag{17}$$

Thus the term $h(X)-h(Y)$ is negative and dual bounded, also lower and upper bound can be converged if distribution of watermark tends to Gaussian shape (using maximum entropy lemma [27]). The third term in (15) is also negative, because $|T(e)| \leq |e|$. From (15) and (17) we have:

$$\lim_{T(e)\to e, DWR\to+\infty} I(K;Y,M) = \lim_{DWR\to+\infty} h(Y)-h(X) = 0 \tag{18}$$

So, perfect secrecy for GSCS in KMA scenario is possible if both $T(e)\to e$ and $DWR\to+\infty$, but this choice leads to zero embedding distortion (from (6)) and consequently zero achievable rate, since achievable rate is an increasing function of embedding distortion. This shows trade-off between security and achievable rate in GSCS in KMA scenario (at least for one observation).

B. Residual Entropy of $N_0 \geq 1$ observations

Residual entropy of secret key conditioned on watermarked hosts and messages can be evaluated as:

$$\begin{aligned} h(K \mid Y^{N_0}, M^{N_0}) &= h(K) - I(K;Y^{N_0}, M^{N_0}) \\ &= h(K) - I(K;Y^{N_0} \mid M^{N_0}) \\ &= h(K) - h(Y^{N_0} \mid M^{N_0}) \\ &\quad + h(Y^{N_0} \mid M^{N_0}, K) \end{aligned} \tag{19}$$

Notice that Y^{N_0} conditioned on M^{N_0} and K, are composed of i.i.d sequence X_i, using (14) we can write:

$$h(Y^{N_0} \mid M^{N_0}, K) = N_0 \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \log(T'(e))de + N_0 h(X) \tag{20}$$

For second term, we should first compute the pdf of watermarked hosts conditioned on

messages. It can be written as follows:

$$\begin{aligned} p(y^{N_0} \mid m^{N_0}) &= \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} p(y^{N_0} \mid m^{N_0}, k) p_K(k) dk \\ &= \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \prod_{i=1}^{N_0} \frac{p_X(T^{-1}(y_i - Q_{m,k}(y_i)) + Q_{m,k}(y_i))}{T'(T^{-1}(y_i - Q_{m,k}(y_i)))} dk \end{aligned} \tag{21}$$

Unfortunately, deriving theoretical closed form for this pdf is intricate, so we resort to numerical integration for deriving this pdf. The first solution is to compute (21) numerically, then replacing in (19) and integrating over $\{y^{N_0}, m^{N_0}\}$ using Monte-Carlo integration. This solution is not accurate for large N_0 due to big error in large dimension in Monte-Carlo integration. The more exact solution is to average $h(K \mid y^{N_0}, m^{N_0})$ over large number N of outcomes $\{y^{N_0}, m^{N_0}\}_{i=1}^{N}$ instead of integration over them as

$$\begin{aligned} h(K \mid Y^{N_0}, M^{N_0}) &= E_{Y^{N_0}, M^{N_0}} \{h(K \mid y^{N_0}, m^{N_0})\} \\ &\cong \frac{1}{N} \sum_{i=1}^{N} h(K \mid \{y^{N_0}, m^{N_0}\}_i) \end{aligned} \tag{22}$$

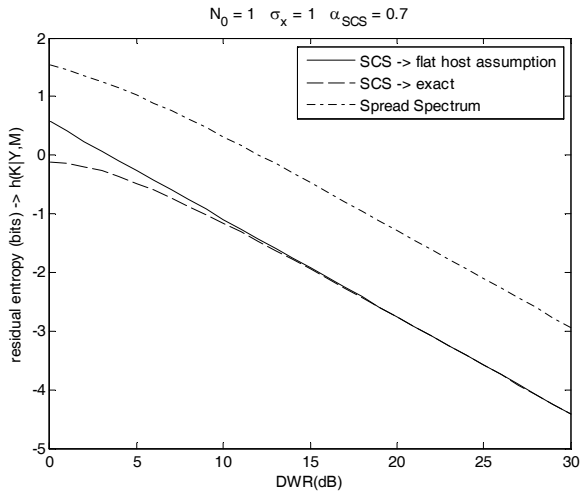
which is application of *weak law of large numbers* in approximation used above. As N increases, variance of error tends to zero if outcomes $\{y^{N_0}, m^{N_0}\}_i$ are independent. This is true because $\{y^{N_0}\}_i$ in iteration i are function of $\{x^{N_0}\}_i$, $\{m^{N_0}\}_i$ and *same secret key* k_i, where all of them are mutually independent. In practice $N \approx 500$ is sufficient to get accurate result.

C. *Investigation on SCS as a special case*

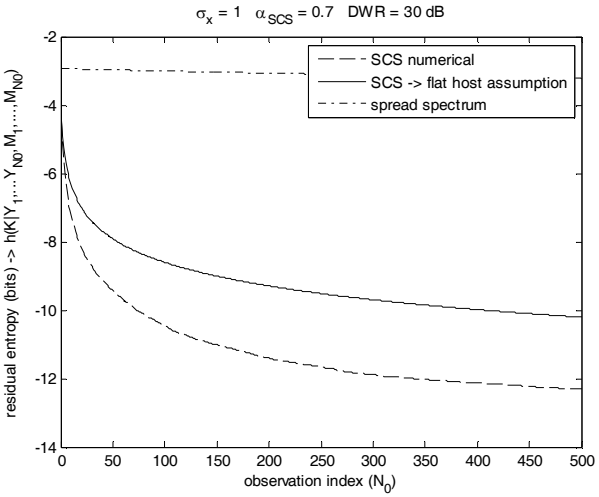
GSCS can be easily simplified to SCS by substituting $T_{SCS}(e) = (1 - \alpha_{SCS})e$ in GSCS. A good approximation for residual entropy conditioned on watermarked host and message for SCS Gaussian host in $N_0 = 1$ observation is as follows:

$$h(K \mid Y, M) \cong \log_2(\Delta_{SCS}) + \log_2(1 - \alpha_{SCS}) - \frac{0.684}{\lambda} \tag{23}$$

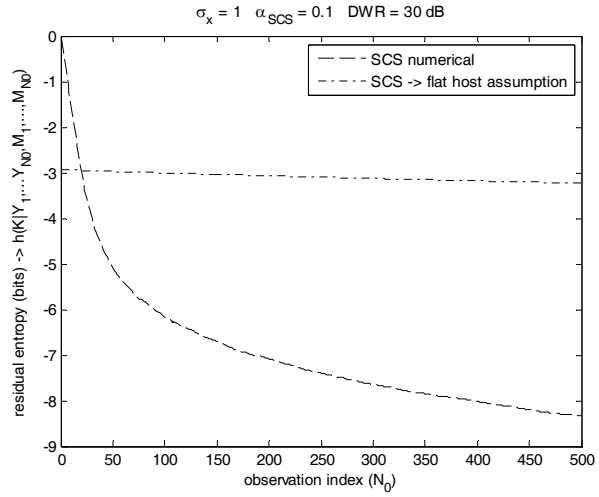
The proof comes after using (6), (13) and (15) and some simplifications. Notice that previous security analysis of SCS [5] ignores host statistics and didn't drive the third term in (23). It is because of using flat-host assumption which implies infinite host variance, so substituting $\lambda = +\infty$ in (23) gets previous result in [5]. Comparison between exact security analysis of SCS proposed here and security analysis based on flat-host assumption [5] is plotted in Fig. 2(a). For $N_0 \geq 1$ we resort to numerical method as in (22). Comparison between previous theoretic result based on flat-host assumption [5] and our result is sketched in Fig. 2(b) and Fig. 2(c), also theoretic result for SS scheme [28] is sketched for comparison. Accurate result for $N_0 = 1$ and numerical result for $N_0 \geq 1$ shows large difference between the security level of SCS derived here and the security level of SCS by flat-host assumption [5], e.g. as depicted in Fig. 2(b), previous result in [5] says that we need $N_0 \simeq 500$ observations to make residual entropy lower than -10.1, but we showed here that attacker only needs $N_0 \simeq 70$ observations to do this. Other key point is that even for low α_{SCS}, there is still large gap between security level of SCS and SS scheme.



(a)



(b)



(c)

Fig. 2. Security analysis of SCS for Gaussian host and $N_0 \geq 1$ observation

IV. SURJECTIVE SCS

Security analysis made in section III, clearly shows that there is a fundamental trade-off between security and achievable rate of GSCS in KMA scenario. So the perfect secrecy as defined in [21] and [29] or like one developed in [18] for improving security of SCS in WOA scenario, is not possible. As a result, we can only improve *security level* as defined in (2) in KMA scenario.

Security level of GSCS is completely dependent on the shape of function $T(.)$. In previous

estimation attack on SCS [5], author benefits from the weakness that distribution of watermarked host doesn't cover the quantization cell for fixed message and secret key. It is clear by substituting $T_{SCS}(e) = (1-\alpha_{SCS})e$ in (11). To make connection with GSCS, the weakness in SCS comes arise from this fact, that function $T(.)$ in SCS is not surjective over its domain, i.e. for some η there is no e such that $T(e) = \eta$. To overcome this weakness, we propose new scheme which called Surjective-SCS (SSCS) as follows. The value for α and β is computed through numerical optimization.

$$T_{SSCS}(e) = \begin{cases} \dfrac{e + \beta\dfrac{\Delta}{2}}{1-\beta}, & -\dfrac{\Delta}{2} \le e \le -x_\beta \\ (1-\alpha)e, & |\, e \,| \le x_\beta = \dfrac{\Delta}{2}\dfrac{\beta}{\alpha+\beta-\alpha\beta} \\ \dfrac{e - \beta\dfrac{\Delta}{2}}{1-\beta}, & x_\beta \le e \le \dfrac{\Delta}{2} \end{cases} \tag{24}$$

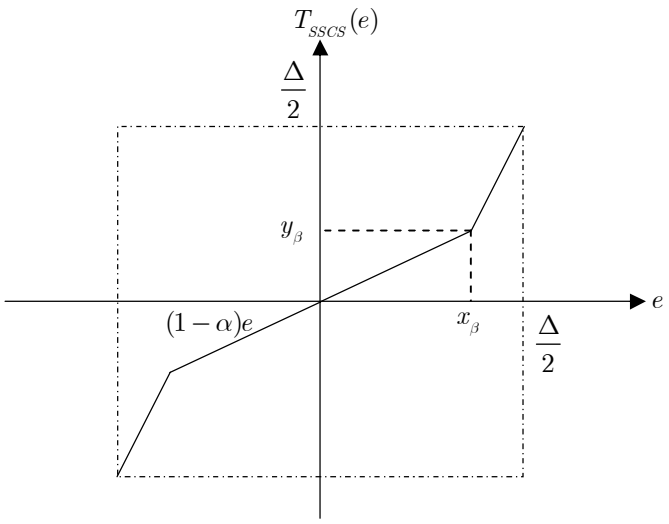


Fig. 3. Surjective-SCS

Function $T_{SSCS}(.)$ is illustrated in Fig. 3. Notice that SSCS with $\beta = 1$ is equal to SCS, also SSCS with $\beta \neq 1$ by definition is completely secure against estimation attack developed in [5]. But as stated in [5], attacker may use other non-convex estimation attack. To analyze security of SSCS against every estimation attack, we use theoretic security analysis of GSCS made in

previous section and show that, SSCS is always more secure than SCS in the same DWR and WNR.

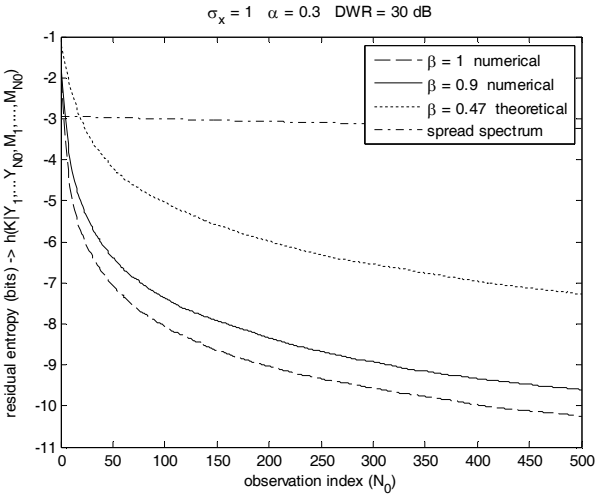
For $N_0 = 1$ observation, (15) simplifies to

$$h(K \mid Y, M) = \log(\Delta) - \frac{0.684}{\lambda} + \frac{\beta \log(1-\alpha) - \alpha(1-\beta)\log(1-\beta)}{\alpha + \beta - \alpha\beta} \tag{25}$$

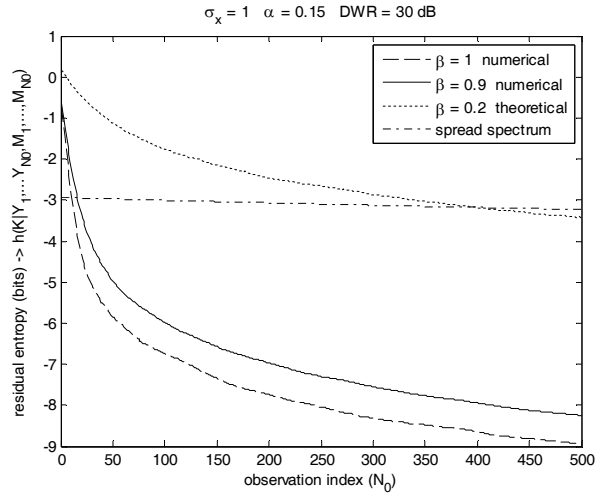
The proof is same as (23). Now we compare residual entropy of SCS and SSCS in (23) and (25). It is simple to find the region $(\alpha_{SCS}, \alpha, \beta)$, which residual entropy of SCS is smaller than SSCS in same DWR. This region is derived from below which come after using (6) for SCS and SSCS and substituting in (23) and (25).

$$\log(\frac{1-\alpha_{SCS}}{\alpha_{SCS}}) \le -\log(\frac{\alpha\beta}{\alpha + \beta - \alpha\beta}) + \frac{\beta \log(1-\alpha) - \alpha(1-\beta)\log(1-\beta)}{\alpha + \beta - \alpha\beta} \tag{26}$$

For $N_0 \ge 1$, we use (22) to compare security level of SSCS and SCS. Comparison for some α, β, Δ is shown in Fig. 4. Simulation results state that, security level of SSCS is greater than SCS. As discussed in previous section, by reducing α, we can't achieve more security than SS, but here we can see that, for low α by reducing β, we can achieve more security even than SS. However reducing β, may lead to lower robustness in same DWR, so in following, we add AWGN attack channel and compare achievable rate of SSCS with SCS in same DWR and WNR.



(a)



(b)

Fig. 4. Residual entropy of SSCS ($\beta = 1$ is equal to SCS)

With $\beta \neq 1$, we introduce more symbol interference (self-noise) which happens in SCS only when $\alpha \leq 0.5$, but we show that this more symbol interference increases achievable rate as in [30]. We compute achievable rate to find the inherent performance limits of SSCS. To investigate this, we compute achievable rate from (1) and use the same approach as used in [19] by ignoring flat-host assumption. The results in this section are derived for Gaussian hosts and binary signaling. Comparison of achievable rate between SSCS and SCS are sketched in Fig. 5. Optimum encoder parameters which maximize achievable rate are in Fig. 6. For WNR

larger than -2 dB, $\beta = 1$ (SCS) is optimum. SSCS obtains more achievable rate than SCS only for WNR smaller than -2 dB, so we compare results only for negative WNRs. It is worthy to note that, we compare maximum achievable rate of SSCS and SCS by fixing $k = -\Delta / 4$ in both SCS and SSCS. As discussed in [19], for lattices like in (5), using different dither may incur a loss of performance and comparison of maximum achievable rate is meaningless.

Finally, we compare security level of SSCS and SCS after adding Gaussian attack, to make a connection with achievable rate. Residual entropy of SSCS and SCS are compared in Fig. 7. As discussed in the first of this section, we can conclude that security level of SSCS is greater than SCS for WNR smaller than -2 dB.

We can conclude this section with four statements, 1) both security and achievable rate of SSCS is more than SCS for WNR smaller than -2 dB; 2) security level of SSCS in comparison with SCS increases as WNR decreases, e.g. in WNR=-8 dB as illustrated in Fig. 4(a) and Fig. 6, security level of SSCS is much higher than SCS while keeps other criteria, but in WNR=-15 dB as illustrated in Fig. 4(b) and Fig. 6, security level of SSCS is very much higher than SCS or SS while keeps other criteria; 3) by reducing β, we can fill the gap between the security level of SCS and SS while have more achievable rate than both of them; and 4) However $\beta = 1$ is optimum for WNR larger than -2 dB from achievable rate point of view, yet we can use another β and increase the security level of SCS as wish if we accept loss of achievable rate.

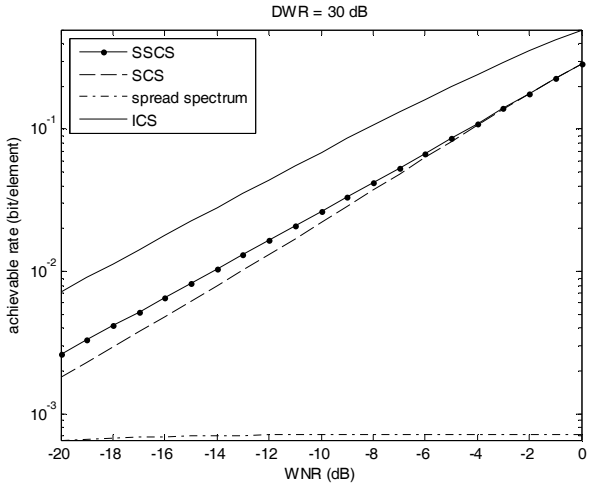


Fig. 5. Achievable rate of SSCS and SCS

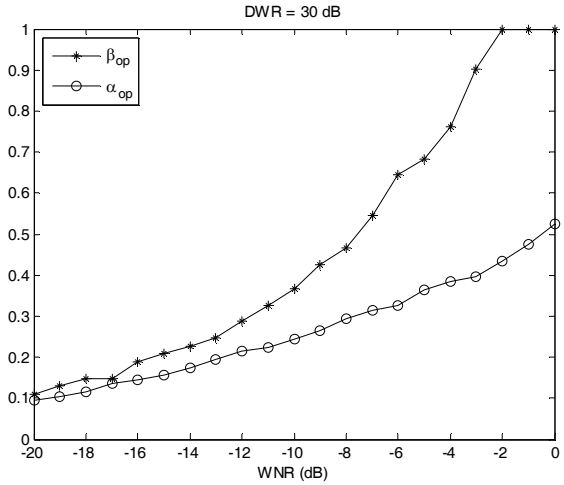
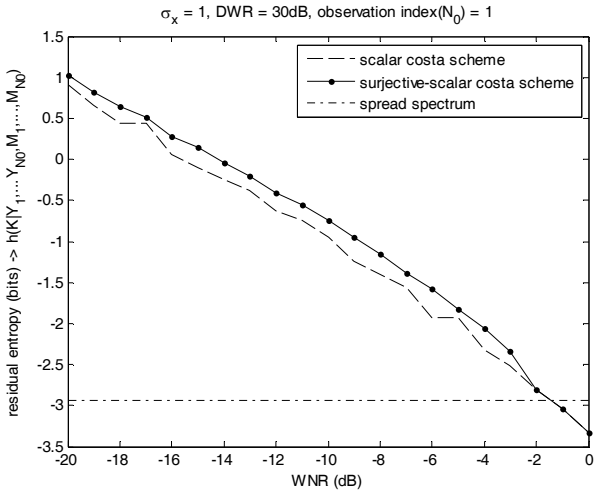
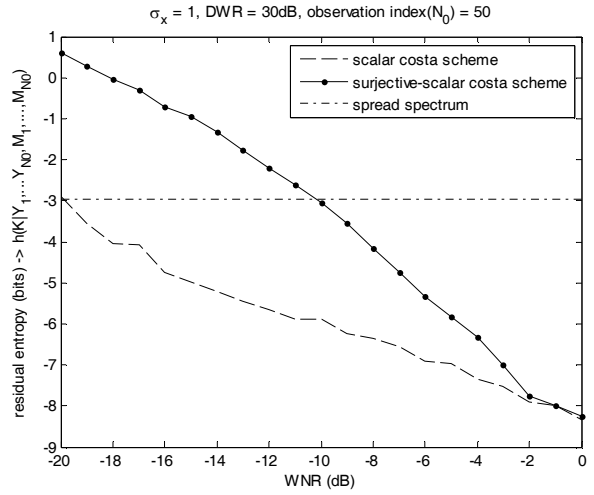


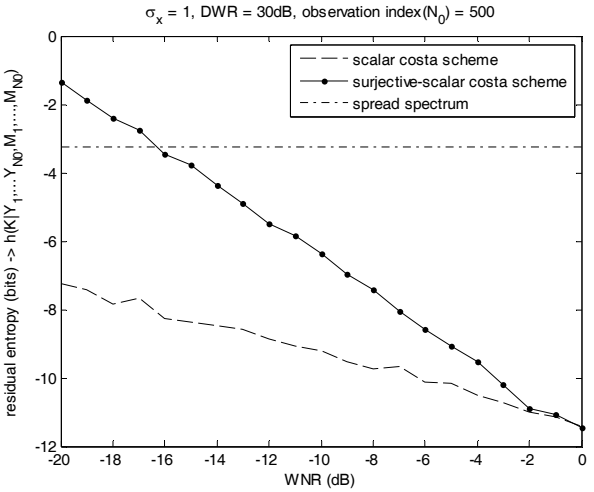
Fig. 6. Optimum encoder parameter in SSCS



(a)



(b)



(c)

Fig. 7. Residual entropy comparison between SSCS and SCS

V. CONCLUSION AND FUTURE WORK

The object of this paper is to increase security and robustness of SCS KMA scenario and to fill the gap to SS scheme, while it keeps other watermarking criteria, which has not done before in the literature. To do this, we proposed security analysis by ignoring flat-host assumption, which is applicable to any extension of SCS. Also we showed that, although analyzing the performance of SCS based on flat-host assumption, underestimates SCS achievable rate, but it overestimates SCS security level. Then we proposed new scheme which

is called SSCS.

SSCS increases security and achievable rate of SCS in low WNR as much as Soft-SCS or Spread Transform-SCS (STSCS) in same DWR without using any spreading techniques. Also SSCS keeps simplicity and computational complexity which is desirable from implementation point of view. Other aspect key is the Controllable trade-off between security and achievable rate of SSCS in KMA scenario, like scheme developed in [31] for improving security of Circular Watermarking (CW) [29] in WOA scenario. The application of SSCS in multimedia is against the situation where attacker in passive mode, has large number of observations and wants to disclose the secret key and implement powerful attacks (security attack), and simultaneously in active mode, adds strong noise to decrease embedding bit rate (robustness attack).

Future works include: 1) design suboptimum and low cost decoder for SSCS to work in practical applications especially in high noise level situations and then replacing SCS by SSCS in applications of SCS which they needs high security level; 2) use (11) to get theoretic performance of SCS or any extension of SCS in existing applications; 3) extension to multidimensional case by extending lattice embedding in (3) and introducing additional self-noise like (24) to get more performance in low WNR. Notice that, the function $T(.)$ in multidimensional case can be more complex because of dependency between the dimensions.

Our main contributions are the following:

- Theoretic proof for the trade-off between security and achievable rate of GSCS,
- An experimental setup for evaluating the security levels for GSCS,
- Joint security and achievable rate enhancement of SCS in KMA scenario.

APPENDIX

A. *Entropy of watermarked host conditioned on message and secret key for GSCS*

In this appendix to overcome the lack of space needed for long formula, we use notation $p_Y(y \mid m,k)$ instead of $p_Y(y \mid M=m, K=k)$ and $h(Y \mid m,K)$ instead of $h(Y \mid M=m,K)$. First we have:

$$h(Y \mid M,K) = \frac{1}{p}\sum_{m=0}^{p-1} h(Y \mid m,K) \tag{27}$$

The term $h(Y \mid m,K)$ can be written as follows:

$$\begin{aligned} h(Y \mid m,K) &= E_{Y,K|M=m}\{-\log_2(p_Y(y \mid m,k))\} \\ &= -\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} p_K(k)\int_{-\infty}^{+\infty} p_Y(y \mid m,k)\log_2(p_Y(y \mid m,k))dydk \end{aligned} \tag{28}$$

For uniform secret key k, we have:

$$h(Y \mid m,K) = \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h(Y \mid m,k)dk \tag{29}$$

The term $h(Y \mid m,k)$ can be simplified as follows. The proof is the same as (13) or the one used in [23] plus some simplifications.

$$\begin{aligned} h(Y \mid m,k) &= h(X) \\ &\quad + \int_{-\infty}^{+\infty} p_X(e+\Delta\frac{m}{p}+k)\log_2(T'(e-Q_\Delta(e)))de \end{aligned} \tag{30}$$

After substituting (30) in (29) and then in (27), we obtain the intended result in (14).

ACKNOWLEDGMENT

The authors are indebted to Prof. Gonzalez, Dr. Furon, Dr. Cayre, and Dr. Bas for their valuable helps.

REFERENCES

[1] I. Cox, M. Miller and A. Mckellips, "Watermarking as communications with side information," *Proceedings of the IEEE,* vol. 87, no. 7, pp. 1127-1141, 1999.

[2] M. Costa, "Writing on dirty paper (Corresp.)," *Information Theory, IEEE Transactions on*, vol.29, no.3, pp. 439- 441, May 1983.

[3] B. Chen, and G.W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *Information Theory, IEEE Transactions on*, vol.47, no.4, pp.1423-1443, May 2001.

[4] J.J. Eggers, R. Bauml, R. Tzschoppe, and B. Girod, "Scalar Costa scheme for information embedding," *Signal Processing, IEEE Transactions on* , vol.51, no.4, pp. 1003- 1019, Apr 2003.

[5] L. Perez-Freire, F. Perez-Gonzalez, T. Furon, and P. Comesana, "Security of Lattice-Based Data Hiding Against the Known Message Attack," *Information Forensics and Security, IEEE Transactions on*, vol.1, no.4, pp.421-439, Dec. 2006.

[6] L. Perez-Freire, and F. Perez-Gonzalez, "Security of Lattice-Based Data Hiding Against the Watermarked-Only Attack," *Information Forensics and Security, IEEE Transactions on* , vol.3, no.4, pp.593-610, Dec. 2008.

[7] I. Cox, G. Doerr and T. Furon, "Watermarking is not cryptography," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* vol. 4283 LNCS, pp. 1-15, 2006.

[8] N. Liu, P. Amin, and K.P. Subbalakshmi, "Security and Robustness Enhancement for Image Data Hiding," *Multimedia, IEEE Transactions on* , vol.9, no.3, pp.466-474, April 2007.

[9] M. Holliman, N. Memon, and M. Yeung, "On the need for image dependent keys in

watermarking," *Proceedings of IEEE Content Security and Data Hiding in Digital Media,* 1999.

[10] I. Nasir, F. Khelifi, J. Jiang and S. Ipson, "Robust image watermarking via geometrically invariant feature points and image normalisation," *IET Image Processing,* vol. 6, no. 4, pp. 354-363, 2012.

[11] Fei. Chuhong, D. Kundur, and R.H. Kwong, "Analysis and design of secure watermark-based authentication systems," *Information Forensics and Security, IEEE Transactions on*, vol.1, no.1, pp. 43- 55, March 2006.

[12] P. Comesana, F. Balado, and F. Perez-Gonzalez, "A novel interpretation of content authentication," In Edward J. Delp III and Ping W. Wong, editors, *Security, Steganography, and Watermarking of Multimedia Contents IX*, volume 6505, San Jose, California, USA, January 2007.

[13] F. Perez-Gonzalez, C. Mosquera, M. Barni and A. Abrardo, "Rational Dither Modulation: A high-rate data-hiding method invariant to gain attacks," *IEEE Transactions on Signal Processing,* vol. 53, no. 10 II, pp. 3960-3975, 2005.

[14] P. Moulin, and A.K. Goteti, "Block QIM watermarking games," *Information Forensics and Security, IEEE Transactions on* , vol.1, no.3, pp.293-310, Sept. 2006.

[15] M. Wu, "Joint security and robustness enhancement for quantization based data embedding," *Circuits and Systems for Video Technology, IEEE Transactions on* , vol.13, no.8, pp. 831- 841, Aug. 2003.

[16] A. Piva, T. Bianchi, and A. De Rosa, "Secure Client-Side ST-DM Watermark Embedding," *Information Forensics and Security, IEEE Transactions on* , vol.5, no.1, pp.13-26, March 2010.

[17] P. Bas, "Informed secure watermarking using optimal transport," *Acoustics, Speech and*

Signal Processing (ICASSP), 2011 IEEE International Conference on , pp.1848-1851, 22-27 May 2011.

[18] P. Bas, "Soft-SCS: improving the security and robustness of the Scalar-Costa-Scheme by optimal distribution matching," *In Information Hiding, IH'11, Lecture Notes in Computer Science*, Prague, Czech Republic, May 18-20, 2011.

[19] L. Perez-Freire, F. Perez-Gonzalez, and S. Voloshynovskiy, "An accurate analysis of scalar quantization-based data hiding," *Information Forensics and Security, IEEE Transactions on*, vol.1, no.1, pp. 80- 86, March 2006.

[20] F. Perez-Gonzalez, F. Balado, and J.R.H. Martin, "Performance analysis of existing and new methods for data hiding with known-host information in additive channels," *Signal Processing, IEEE Transactions on*, vol.51, no.4, pp. 960- 980, Apr 2003.

[21] F. Cayre, C. Fontaine, and T. Furon, "Watermarking security: theory and practice," *Signal Processing, IEEE Transactions on*, vol.53, no.10, pp. 3976- 3987, Oct. 2005.

[22] C.E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.,* vol. 28, no. 4, pp. 656–715, 1949.

[23] L. Schuchman, "Dither Signals and Their Effect on Quantization Noise," *Communication Technology, IEEE Transactions on*, vol.12, no.4, pp.162-165, December 1964.

[24] R. Zamir, and M. Feder, "On lattice quantization noise," *Information Theory, IEEE Transactions on*, vol.42, no.4, pp.1152-1159, Jul 1996.

[25] A. Papoulis, and U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2002.

[26] Y. Wang, and P. Moulin, "Steganalysis of block-structured stegotext," in *Proc. Security, Stenography Watermarking of Multimedia Contents*, vol. 5306, pp. 477–488, San Jose, CA, 2004.

[27] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 2006.

[28] L. Perez-Freire, and F. Perez-Gonzalez, "Spread-Spectrum Watermarking Security," *Information Forensics and Security, IEEE Transactions on*, vol.4, no.1, pp.2-24, March 2009.

[29] F. Cayre, and P. Bas, "Kerckhoffs-Based Embedding Security Classes for WOA Data Hiding," *Information Forensics and Security, IEEE Transactions on* , vol.3, no.1, pp.1-15, March 2008.

[30] U. Erez, S. Shamai, R. Zamir, "Capacity and lattice strategies for canceling known interference," *Information Theory, IEEE Transactions on* , vol.51, no.11, pp. 3820-3833, Nov. 2005.

[31] J. Cao, and J. Huang, "Controllable Secure Watermarking Technique for Tradeoff Between Robustness and Security," *Information Forensics and Security, IEEE Transactions on*, vol.7, no.2, pp.821-826, April 2012.