Floating Signal Constellations for Multimedia Steganography

Mahalingam Ramkumar, Ali N. Akansu and Xiaodong Cai
Department of Electrical and Computer Engineering
New Jersey Center for Multimedia Research
New Jersey Institute of Technology
University Heights, Newark, NJ 07102.
ali@megahertz.njit.edu

Abstract— **Conventional spread spectrum communication methods essentially map a discrete symbol from an alphabet to a length-N real valued sequence, which represents a point in the N-dimensional signal constellation. The real valued sequence is in turn transmitted over a communications channel. However, communication schemes for the purpose of multimedia steganography (or data hiding) have to transmit the real valued sequence corresponding to a point in the signal constellation *superimposed* on the original or cover content (without affecting the fidelity of the original content noticeably). The detector therefore, has to *estimate the origin* of the signal constellation, and thereafter proceed to decode the transmitted symbol. In this paper, we explore some efficient solutions for signaling methods employing floating signal constellations.**

I. Introduction

Data hiding [1] or *steganography* is the art of hiding a *message signal* in a *host signal*, without any *perceptual distortion* of the host signal. It is a form of communication utilizing *subliminal* channels. Like any form of communication, steganographic communication relies on two primary resources - bandwidth and power. The bandwidth available for steganographic communication is the *bandwidth of the host signal.* The available power is the *permitted distortion of the host signal.* The limited power available, and comparatively large bandwidth of multimedia host signals like images / video / audio, immediately suggests some form of *spread spectrum* communication.

Consider a digital baseband spread spectrum communication scheme, denoted by $(\mathcal{S}, \mathcal{S}^{-1})$, where an arbitrary symbol indexed by $m : 1 \leq m \leq M$ from an alphabet $\mathcal{M}$ of equi-probable symbols, is mapped to a sequence $\mathbf{s} \in \Re^N$. The sequence $\mathbf{s}$ is in turn transmitted over a channel characterized by additive noise $\mathbf{n} \sim \mathcal{N}[0, \sigma_n^2 \mathbf{I}]$. The corrupted vector at the receiver is $\tilde{\mathbf{s}} = \mathbf{s} + \mathbf{n}$. The mappings

$$\mathcal{S} : m \rightarrow \mathbf{s} \in \Re^N \quad \mathcal{S}^{-1} : \tilde{\mathbf{s}} \rightarrow \tilde{m} \qquad (1)$$

may be performed by using a $M \times N$ codebook $\mathbf{C}$ at the transmitter and the receiver. The transmitted vector $\mathbf{s}$ is one of the possible M codewords. The codebook should be chosen such that the the codewords, $\mathbf{s}_i$, $1 \leq i \leq M$, and $\sum_{j=1}^{N} s_i^2(j) = NP, \forall\, i$ are *maximally separated.* The receiver determines the element of the codebook closest to the corrupted vector $\tilde{\mathbf{s}} = \mathbf{s} + \mathbf{n}$ to obtain an estimate of the transmitted symbol.

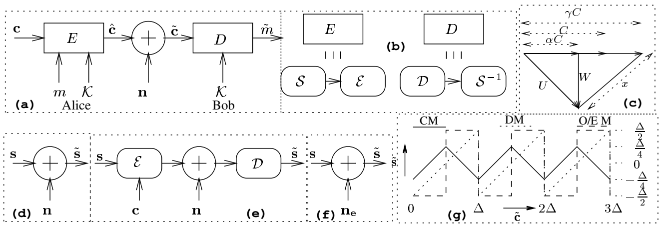

Fig. 1. (a) Steganographic communication between Alice and Bob. $\mathcal{K}$ is a *secret key* used for the communication. (b) Decomposition of embedder E and detector D into $(\mathcal{S}, \mathcal{E})$ and $(\mathcal{D}, \mathcal{S}^{-1})$ respectively. (c) Geometric interpretation of the relationship between α and γ. (d) A simple additive noise channel. (e) Oblivious data hiding (f) Equivalent additive noise channel. (g) Three periodic functions for the mapping $\tilde{\mathbf{s}} = \mathcal{D}(\hat{\mathbf{c}})$ - CP (Continuous Periodic), DM (Dither Modulation), and O/E M (Odd / Even Modulation).

A. Steganographic Communication

Now consider a communication scenario between Alice and Bob depicted in Figure 1 (a). Alice wishes to communicate a symbol index m over a *pre-existing channel meant for the transmission of some host signal* $\mathbf{c} \in \Re^N$. To achieve this Alice is allowed to distort the host signal "imperceptibly". Alice therefore "intercepts" the signal $\mathbf{c}$ and modifies it slightly to obtain the *stego* signal (signal with embedded information) $\hat{\mathbf{c}}$, such that $\mathrm{d}(\mathbf{c}, \hat{\mathbf{c}}) \leq \mathrm{NP}$, where $\mathrm{d}(.,.)$ is some meaningful metric. Alice then transmits the *stego* signal $\hat{\mathbf{c}}$ over the channel. The steganographic communication scheme requires an embedder E at the transmitter and a detector D at the receiver. The stego signal $\hat{\mathbf{c}}$ is obtained as $\hat{\mathbf{c}} = E(\mathbf{c}, m, \mathcal{K})$, where $\mathcal{K}$ is the *key* needed by the detector to decipher the secret communication. We expect $\hat{\mathbf{c}}$ to undergo some modification (due to additive noise $\mathbf{n}$ in the channel) before it reaches the receiver (detector D), where the hidden symbol is extracted. Let $\tilde{\mathbf{c}} = \hat{\mathbf{c}} + \mathbf{n}$ be the input to the receiver. *Non oblivious* data hiding methods need the original content $\mathbf{c}$ for extracting the hidden symbol. On the other hand, *oblivious* detection methods can extract the hidden symbol without any knowledge of the original content $\mathbf{c}$.

It can be easily seen that non-oblivious data hiding can be implemented as a conventional spread spectrum communication scheme outlined in the previous section where $\hat{\mathbf{c}} = \mathbf{c} + \mathbf{s}$ and $\tilde{\mathbf{s}} = \tilde{\mathbf{c}} - \mathbf{c}$, and as earlier, a codebook $\mathbf{C}$ is used for the mappings $\mathcal{S} : m \rightarrow \mathbf{s}$ and

$\mathcal{S}^{-1} : \tilde{\mathbf{s}} \rightarrow \tilde{m}$. However, if the receiver does not have access to the original content $\mathbf{c}$, the problem of designing a good steganographic communication scheme is an interesting one. In this paper, we explore various options for the design of sophisticated oblivious steganographic communication schemes.

II. Oblivious Data Hiding

In [2] Costa considered a power constrained communication scenario modeled as $\tilde{\mathbf{c}} = \mathbf{c} + \mathbf{w} + \mathbf{n}$, where $\mathbf{c}, \mathbf{w}, \mathbf{n} \in \Re^N$, and $\mathbf{c} \sim \mathcal{N}[0, \sigma_c^2 \mathbf{I}], \mathbf{w} \sim \mathcal{N}[0, P\mathbf{I}]$ and $\mathbf{n} \sim \mathcal{N}[0, \sigma_n^2 \mathbf{I}]$ are i.i.d. Furthermore, $\mathbf{c}$, $\mathbf{w}$, and $\mathbf{n}$ are independent. In the above model $\mathbf{w}$ is power constrained (variance P), and $\mathbf{n}$ is the noise in the channel. $\tilde{\mathbf{c}}$ is the signal received at the receiver, which does not have access to $\mathbf{c}$. This is exactly the same as the steganographic communication model where $\mathbf{c}$ is the content, $\mathbf{n}$ is the additive noise in the channel, and $\mathbf{w} = \hat{\mathbf{c}} - \mathbf{c}$, is the permitted distortion of the host signal.

Once again, if $\mathbf{c}$ was known to the receiver, the situation is one of transmitting a signal $\mathbf{w}$ of variance P over a channel with noise variance σ_n^2. For Gaussian $\mathbf{w}, \mathbf{n}$ one could achieve a capacity [3] of $NR^* = \frac{N}{2} \log_2 \left(1 + \frac{P}{\sigma_n^2}\right)$ bits. Costa however argued that one could achieve capacity NR^* *even if* $\mathbf{c}$ *is not available at the receiver*! The argument is based on the capacity of a discrete memoryless channel $\tilde{C} = C + W + N_o$, where the random state C is known only to the encoder. Let $U = f(C, W)$ be an auxiliary random variable. Gel'fand *et. al* [4] have shown that the capacity in this case is given by

$$R = \max_{p(u,w|c)} \left\{ I(U; \tilde{C}) - I(U; C) \right\}. \qquad (2)$$

Costa considered auxiliary variable U of form $U = W + \alpha C$, to obtain

$$R(\alpha) = \frac{1}{2} \log_2 \left(\frac{P(P + \sigma_c^2 + \sigma_n^2)}{P\sigma_c^2(1 - \alpha^2) + \sigma_n^2(P + \alpha^2 \sigma_c^2)} \right) \text{ bits.} \qquad (3)$$

where $R(\alpha)$ is maximum for $\alpha^* = \frac{P}{P + \sigma_n^2}$, or $R(\alpha^*) = \frac{1}{2} \ln \left(1 + \frac{P}{\sigma_n^2}\right) = R^*$ bits, which is the same as the case when $\mathbf{c}$ is known to the detector. Thus oblivious data hiding can do as well as non oblivious data hiding.

The encoder and decoder share a codebook of $2^{NI(U;\tilde{C})}$ codewords [1] distributed amongst 2^{NR^*} bins. Some reasonable choices [1] indicate that $I(U; \tilde{C}) \approx (20 \text{ to } 60)R^*$. This implies that while conventional spread-spectrum communication schemes employ a codebook of size $2^{NR^*} \times N$, the size of codebook needed for ideal the signaling scheme for blind steganography is about $2^{20NR^*} \times N$ to $2^{60NR^*} \times N$. In using the ideal method, to transmit a symbol m, where $1 \leq m \leq 2^{NR^*}$, Alice chooses a codeword $\mathbf{u}_0$ from the m^{th} bin. For sufficiently large N, it is guaranteed that a codeword $\mathbf{u}_0$ can be found in bin m which satisfies

$$(\mathbf{u}_0 - \alpha^* \mathbf{c})^T \mathbf{c} = \mathbf{w}^{\mathrm{T}} \mathbf{c} \leq \delta \quad \textstyle\sum_{i=1}^N w^2(i) \leq NP, \qquad (4)$$

where $\delta \rightarrow 0$ for $N \rightarrow \infty$. The signal at the detector is $\tilde{\mathbf{c}} = \mathbf{c} + \mathbf{w} + \mathbf{n}$. Let $\mathbf{u}_i, i = 1 \cdots 2^{NI(U;\tilde{C})}$ be the i^{th} codeword in $\mathbf{U}$, and $d_i = (\mathbf{u}_i - \alpha \tilde{\mathbf{c}})^T \tilde{\mathbf{c}}$. Further, let $q = \arg\min_{\mathrm{i}} \mathrm{d_i}$. The receiver's estimate of m, viz. $\tilde{m}$ is the bin in which the codeword $\mathbf{u}_q$ resides. It can be easily shown that for the codeword $\mathbf{u}_0$, $d_q = d_0 = (1 - \alpha)P - \alpha\sigma_n^2 + (1 - \alpha)\delta + t$, where t are terms that go to zero for large N. As $\alpha = \frac{P}{P + \sigma_n^2}$, $(1 - \alpha)P - \alpha\sigma_n^2 = 0$ and $d_0 = (1 - \alpha)\delta + t$. Note that the detector *needs to know the noise variance.*

In [5], Chou *et. al.* obtain the orthogonality of $\mathbf{c}$ and $\mathbf{w}$ by choosing the codebook $\mathbf{U}$ as rate-distortion optimized quantized version of a scaled version of $\mathbf{c}$, or $\gamma\mathbf{c}$. As the quantization error x is orthogonal to $\mathbf{U}$, $\mathbf{w}$ will be orthogonal to $\mathbf{c}$, and will have a power P for (see Figure 1 (c)) for $x^2 = P + (\gamma - \alpha)^2 \sigma_c^2 = \gamma^2 \sigma_c^2 - (\alpha^2 \sigma_c^2 + P)$, or $\gamma = \frac{P + \alpha^2 \sigma_c^2}{\alpha \sigma_c^2}$.

For most applications of data hiding however, it may be impractical to use such large codebook sizes. Additionally, for most applications, it is unreasonable to expect the detector to know the actual noise variance. Communication schemes are typically designed assuming "worst-case" scenarios. A reasonable direction to take is to determine how well a steganographic communication scheme can perform if we *restrict the complexity of the signaling method used for data hiding to be the same as conventional spread-spectrum communication schemes.* Furthermore, the detector should not need to know the actual noise variance. To achieve this, we limit the size of the codebook used to $2^{NR} \times N$. We now have a signaling scheme employing a codebook $\mathbf{C}$ described by (1). $\mathcal{S}$ converts the symbol indexed by m to $\mathbf{s} \in \Re^N$. The embedder E and detector D are split into two parts - $(\mathcal{S}, \mathcal{E})$ and $(\mathcal{D}, \mathcal{S}^{-1})$ as shown in Figure 1 (b). Furthermore, $\mathcal{E}, \mathcal{D}$ are restricted to be scalar operations. The overall hiding and detection scheme can therefore be described by

$$\begin{array}{lll} \mathbf{s} = \mathcal{S}(m) & \hat{c}(i) = \mathcal{E}(c(i), s(i)) & \text{embedding} \\ \tilde{s}(i) = \mathcal{D}(\tilde{c}(i)) & \tilde{m} = \mathcal{S}^{-1}(\tilde{\mathbf{s}}) & \text{detection,} \end{array} \qquad (5)$$

for $1 \leq i \leq N$. In the rest of the paper, we shall assume that we have a good "conventional" spread spectrum scheme $(\mathcal{S}, \mathcal{S}^{-1})$ (or equivalently, a "good" codebook $\mathbf{C}$) available. All we need to do is to come up with good alternatives for the pair $\mathcal{E}, \mathcal{D}$. The overall signaling technique can now be seen as one employing a *floating signal constellation.* An N-dimensional constellation is defined by $(\mathcal{S}, \mathcal{S}^{-1})$ (with respect to a known origin). The job of $\mathcal{E}$ is to "shift the origin" of the constellation in the space of $\mathbf{c} \in \Re^N$ to obtain $\hat{\mathbf{c}}$ such that $\mathrm{d}(\mathbf{c}, \hat{\mathbf{c}})$ is minimum. Simultaneously, $\mathcal{E}$ has to guarantee that the detector $\mathcal{D}$ will be able to "estimate" the origin of

[1] The codewords are the finite states of the auxiliary random variable U.

the constellation. The real vector $\mathbf{s}$ transmitted by Alice (or generated by $\mathcal{S}$) is received by Bob (or input to $\mathcal{S}^{-1}$) as $\tilde{\mathbf{s}}$. Therefore a good measure for the efficiency of $(\mathcal{E}, \mathcal{D})$ would be how "close" $\tilde{\mathbf{s}}$ is expected to be to $\mathbf{s}$, for a given noise $f_{\mathrm{N}_\circ}(n)$ in the channel. Good choices of $(\mathcal{E}, \mathcal{D})$ should obviously minimize $E[d(s(k), \tilde{s}(k))] \forall k$ subject to the constraint $E[d(\hat{c}(k), c(k))] \leq NP$, where E[.] denotes the expectation operation.

Consider the simple additive noise channel of Figure 1 (d). Let $\mathbf{n} \sim [f_{\mathrm{N}_\circ}(n), \sigma_n^2 \mathbf{I}]$ be additive noise in the channel. In such a scenario, for even $f_{\mathrm{N}_\circ}(n)$, it can be easily shown [6] that the expected value of normalized inner product of $\mathbf{s}$ and $\tilde{\mathbf{s}}$, viz. ρ, and the noise variance σ_n^2 are related as

$$\rho^2 = \frac{P}{P+\sigma_n^2} \quad \text{or} \quad \sigma_n^2 = \frac{P(1-\rho^2)}{\rho^2}. \tag{6}$$

Now consider oblivious data hiding scenario in Figure 1 (e). The signaling scheme $\mathcal{S}, \mathcal{S}^{-1}$ can consider the situation in Figure 1 (e) as an "equivalent" channel shown in Figure 1 (f), where similar to (6), $\sigma_{n_e}^2 = \frac{P(1-\rho^2)}{\rho^2}$ (note that ρ in this case is the normalized correlation of $\mathbf{s}$ and $\tilde{\mathbf{s}}$ in Figures 1 (e) and (f) while ρ in (6) corresponds to $\mathbf{s}$ and $\tilde{\mathbf{s}}$ in Figure 1 (d)). We would expect $\sigma_{n_e}^2 > \sigma_n^2$. We may consider $\sigma_{n_e}^2$ as the variance of the *equivalent* additive noise. The difference $\sigma_{n_e}^2 - \sigma_n^2$ may then be considered as the *penalty* paid for choosing a suboptimal steganographic communication scheme. With this model, henceforth, our aim is the choice of $(\mathcal{E}, \mathcal{D})$ such that $\sigma_{n_e}^2 - \sigma_n^2$ is as small as possible.

The choice of normalized correlation as the measure of the merit of a system, and the correspondence between correlation and equivalent noise also helps in comparison of achievable capacities of the suboptimal steganographic communication methods. While *ideal oblivious* steganographic methods can achieve a capacity of $NR^* = \frac{N}{2}\log_2(1+\frac{P}{\sigma_n^2})$, (which is the same as the capacity of non-oblivious steganographic communications), the non-ideal (and significantly less complex) methods can only achieve a capacity of $NR = \frac{N}{2}\log_2(1+\frac{P}{\sigma_{n_e}^2})$ bits. Obviously, as $\sigma_{n_e}^2 > \sigma_n^2$, $R < R^*$.

III. Choices for $\mathcal{E}, \mathcal{D}$

A. Type I

A simple (and very inefficient) alternative is to consider the host signal $\mathbf{c}$ as noise. Typically, the energy of the host signal, viz. $N\sigma_c^2 = \sum_{i=1}^{N} c^2(i)$ is much higher than the channel noise energy $N\sigma_n^2$ and the permitted distortion NP. For such systems, the equivalent noise is $\sigma_{n_e}^2 = \sigma_c^2 + \sigma_n^2$, which is obviously much greater than σ_n^2. We shall refer to these methods, where $\hat{\mathbf{c}} = \mathcal{E}(\mathbf{c}, \mathbf{s}) = \mathbf{c} + \mathbf{s}$ and $\tilde{\mathbf{s}} = \mathcal{D}(\tilde{\mathbf{c}}) = \tilde{\mathbf{c}}$ as Type I methods. While Type I methods are ideal for non-oblivious data hiding ($\tilde{\mathbf{s}} = \mathcal{D}(\tilde{\mathbf{c}}) = \tilde{\mathbf{c}} - \mathbf{c}$), they are far from optimal for oblivious data hiding. An alternate interpretation [7] of the inadequacy of Type I methods, based on the model of (3) is that for Type I methods $\alpha = 0$, or $U = W$. Substituting $\alpha = 0$ in (3) we can see that $R(0) = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma_n^2+\sigma_c^2}\right)$ bits per sample.

B. Type II

The main shortcoming of oblivious Type I methods is that they assume that the "noise" due to the original content is due to the entropy of $\mathbf{c}$, or $h(\mathbf{c})$. However, the actual entropy of $\mathbf{c}$ at the detector is $h(\mathbf{c}|\tilde{\mathbf{c}})$. Typically as $\sigma_c^2 >> \sigma_n^2$ and $\sigma_c^2 >> P$, $\tilde{\mathbf{c}}$ is still "close" to $\mathbf{c}$. Thus $h(\mathbf{c}|\tilde{\mathbf{c}})$ is substantially smaller that $h(\mathbf{c})$.

Many methods that utilize this fact have been proposed in literature. Such methods, which we shall refer to as Type II, can be characterized by $\mathcal{E}, \mathcal{D}$ which are *exact inverses.* Mathematically, $\hat{\mathbf{c}} = \mathcal{E}(\mathbf{c}, \mathbf{s})$ and $\mathbf{s} = \mathcal{D}(\hat{\mathbf{c}})$ (Note that for oblivious Type I data hiding $\mathcal{D}(\hat{\mathbf{c}}) = \hat{\mathbf{c}} = \mathbf{c} + \mathbf{s} \neq \mathbf{s}$). Another interesting characteristic of Type II methods is that it is *impossible to recover the original* $\mathbf{c}$ from $\hat{\mathbf{c}}$ (for Type I methods $\mathbf{c}$ can be recovered as $\hat{\mathbf{c}} - \mathbf{s}$).

In fact even the earliest data hiding methods which modified least significant bits (LSBs) of image pixels to hide information [8] fall into this category (Type II). Most methods proposed in literature use some form of quantizer to implement $\mathcal{E}$ and $\mathcal{D}$. The method that modifies only the LSB can be considered as using a quantizer with step size of 2. Wang *et. al.* [9] and Wu *et. al.* [10] modify wavelet / DCT coefficients so that they *quantize* to an even or odd value depending on the bit to be embedded. The first formal treatment of such data hiding techniques was proposed by Chen *et. al.* [11]. The form of quantizers used for implementation of $\mathcal{E}$ and $\mathcal{D}$ may also be seen as *periodic functions.* The methods which quantize coefficients to odd or even indices can be considered as using square wave periodic function (O/E M) in Figure 1 (g). On the other hand the dither modulation (DM) scheme proposed by Chen *et. al.* employs a sawtooth periodic function. The general notion that hard decisions are a result of *discontinuities* in the periodic function prompted us to use the *continuous* triangular periodic function (CP) in [6], [12]. As $\mathcal{E}, \mathcal{D}$ are exact inverses, the host signal noise is *completely eliminated* as long as there is *no additive noise* in the channel. However, the performance of Type II systems deteriorate very rapidly when the SNR (ratio of signal power P, the permitted distortion of the host signal and additive channel noise power σ_n^2) is low, as indicated by the squares in Figure 2 (left). This can be explained as follows. By using Type II methods, one attempts to "predict" a substantial part of $\mathbf{c}$ from the received signal $\tilde{\mathbf{c}}$ to reduce its entropy. Δ, the quantizer step-size or the period of the periodic function used, can be seen as the *tightness* of the prediction. However, the choice of Δ is completely dictated by P (more specifically, $P = \frac{\Delta^2}{12}$). For high SNRs one can afford to make tight predictions. However, as the SNR reduces, the predictions become extremely unreliable.

Another interpretation [7] of Type II methods is that the auxiliary random variable U is obtained as $U = W + C$ (or $\alpha = 1$). Substituting $\alpha = 1$ in (3), we can see that the capacity of of Type II methods is given by $R(1) = \frac{1}{2}\log_2\left(\frac{P}{P+\sigma_c^2} + \frac{P}{\sigma_n^2}\right)$. Note that the capacity goes to zero when $\frac{P}{\sigma_n^2} = \frac{\sigma_c^2}{P+\sigma_c^2} \approx 1$.

C. Spread Transform Dither Modulation

In [13], Chen *et. al* proposed a spread transform dither modulation (ST-DM) method in which the vector $\mathbf{c} \in \Re^N$ is split into $\frac{N}{L}$ blocks $[\mathbf{c}_1\, \mathbf{c}_2\, \ldots, \mathbf{c}_{N/L}]$, each of length L. A spreading vector $\mathbf{u} \in \Re^L, \mathbf{u}^T\mathbf{u} = 1$ is used to obtain a vector $\mathbf{x} \in \Re^{N/L}$, where $x(i) = \mathbf{c}_i^T\mathbf{u}$, $1 \le i \le \frac{N}{L}$. The invertible Type II embedding scheme operates on $\mathbf{x}$ instead of $\mathbf{c}$, thus reducing the bandwidth from N to $\frac{N}{L}$. The signature sequence in this case is $\mathbf{d} \in \Re^{\frac{N}{L}}$. The embedder $\mathcal{E}$ obtains

$$\hat{x}(i) = \mathcal{E}(x(i), d(i)) \quad \hat{\mathbf{c}}_i = \mathbf{c}_i + (\hat{x}(i) - x(i))\mathbf{u}. \tag{7}$$

It can be easily seen that $\mathcal{D}(\hat{\mathbf{c}}_i^T\mathbf{u}) = d(i)$. The authors reported that ST-DM performed significantly better than the DM method proposed earlier in [11].

D. Optimal ST - Type II

The reason for the improved performance of ST-DM is due to an advantageous trade-off of bandwidth for SNR. The ST-DM method results in an L fold reduction in bandwidth and an L fold increase in SNR. The total distortion power available, viz. NP, is now spread over $\frac{N}{L}$ coefficients of $\mathbf{x}$. Thus for embedding in $\frac{N}{L}$ coefficients we can afford to increase Δ by a factor of $\sqrt{L}$. This enables the Type II system to continue operating in the high SNR region, where it performs reasonably well.

The capacity of any communication scheme in general, can be represented as a function of bandwidth (N) and SNR, or $C = Nf(\text{SNR})$. By sacrificing bandwidth and increasing SNR, the capacity of the ST-Type II method takes the form $C_{STII} = \frac{N}{L}f(L \times \text{SNR})$. For the CP Type II system, the expected value of the normalized correlation for any SNR (or given Δ and σ_n^2) is given by [1], [6]

$$\rho = \frac{2\sum_{i=0}^{\infty}\int_{\frac{i\Delta}{2}}^{\frac{(i+1)\Delta}{2}}(-1)^i(\frac{(2i+1)\Delta}{4} - n)f_{\mathrm{N}_\circ}(n)\mathrm{d}n}{\sqrt{2\sum_{i=0}^{\infty}\int_{\frac{i\Delta}{2}}^{\frac{(i+1)\Delta}{2}}(\frac{(2i+1)\Delta}{4} - n)^2 f_{\mathrm{N}_\circ}(n)\mathrm{d}n.}} \tag{8}$$

Instead of $\Delta^2 = 12P$ for the original Type-II, for ST-Type II we have $\Delta^2 = 12LP$. This value of ρ can be used to evaluate the variance of the *equivalent* additive noise, $\sigma_{n_e}^2 = \frac{\frac{\Delta^2}{12}(1-\rho^2)}{\rho^2}$, which in turn can be used to estimate the capacity as $C_{STII} = \frac{N}{2L}\log_2(1 + \frac{\Delta^2}{12\sigma_{n_e}^2})$ bits. The optimal choice of L for different SNRs is displayed in Table II. The dotted line in Figure 2 (left) is the plot of achievable capacity for the optimal ST-Type II. In general, lower the SNR, higher will be the optimal value of L. The capacity for $L = 1$ (Type II) and optimal choice of L (ST-II) are plotted in Figure 2(left - squares (Type II), dotted line (optimal ST-II)).

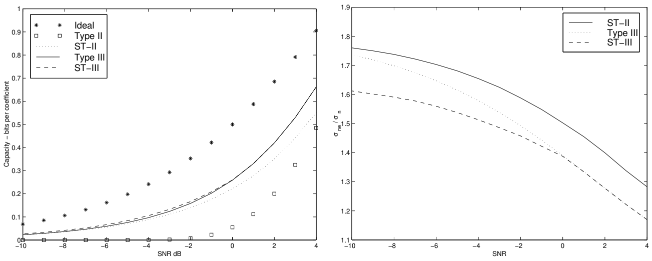

Fig. 2. Left: Capacity of oblivious data hiding schemes for different choices of $\mathcal{E}, \mathcal{D}$. Right: Ratio $\frac{\sigma_{n_e}}{\sigma_n}$ for various SNR's for optimal ST-II, optimal Type III and optimal ST-III. For an ideal system, $\frac{\sigma_{n_e}}{\sigma_n} = 1$.

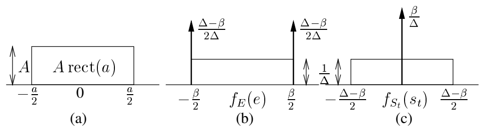

Fig. 3. (a) Rectangular function. Probability distribution of (b) embedding distortion $\mathbf{e}$ and (c) thresholding noise.

E. Type III

In [6] we further proposed a modification of Type II (called Type III) in which we removed the invertibility condition for $\mathcal{E}, \mathcal{D}$. In other words, for Type III systems, $\mathcal{D}(\hat{\mathbf{c}}) \neq \mathbf{s}$, where $\hat{\mathbf{c}} = \mathcal{E}(\mathbf{c}, \mathbf{s})$. We also showed that Type III systems are in fact generalizations of both Type I and Type II systems. Unlike Type II systems which are characterized by a period Δ, the proposed Type III system in [6] was characterized by a period Δ and a threshold β. In the proposed Type III system, the distortion introduced (which in Type II is uniformly distributed between $\pm\frac{\Delta}{2}$) is hard limited to $\pm\frac{\beta}{2}$ where $\beta \le \Delta$. Obviously, for $\Delta = \beta$ Type III systems become Type II. However, what is not so readily obvious is that as $\Delta \to \infty$ (and β is finite), Type III systems become Type I!

The distortion $\mathbf{e} \sim f_E(e)$ introduced by the embedding function $\mathcal{E}$ is displayed in Figure 3 (a). Limiting the distortion also introduces an additional noise $\mathbf{s}_t = \mathbf{s} - \mathcal{D}(\hat{\mathbf{c}})$ for the purpose of detecting the signature. The thresholding noise $\mathbf{s}_t \sim f_{S_T}(s_t)$ (Figure 3 (c)) is independent of the channel noise $\mathbf{n}$. The probability distribution $f_Z(z)$, of the total noise $\mathbf{z} = \mathbf{n} + \mathbf{s}_t$, is obtained by convolution of $f_{\mathrm{N}_\circ}(n)$ and $f_{S_t}(s_t)$ [6]. The optimal choice of Δ and β for a given permitted distortion $P = \frac{\beta^2}{12\Delta}(3\Delta - 2\beta)$, is obtained by maximizing the expected value of the normalized correlation given by (8), where $f_{N_\circ}$ is replaced with f_Z (see [6] for details). If we define $k = \frac{\Delta}{\sqrt{12P}}$, for $k = 1$, Type III systems become Type II. As $k \to \infty$ Type III systems become Type I. The optimal choice of

TABLE I

OPTIMAL VALUES OF $k = \frac{\Delta}{\Delta_0}$ FOR DIFFERENT SNRS

SNR	k	SNR	k	SNR	k
5	1.22	4	1.31	3	1.41
2	1.53	1	1.69	0	1.86
-1	2.07	-2	2.29	-3	2.56
-4	2.86	-5	3.20	-6	3.62

TABLE II

OPTIMAL L / k FOR DIFFERENT SNRS FOR ST-II AND ST-III.

SNR (dB)	4	3	2	1	0
ST-II (L)	2	2	3	3	4
ST-III (L)	1	1	1	1	2
ST-III (k)	1.31	1.41	1.53	1.69	1.41
SNR (dB)	-1	-2	-3	-4	-5
ST-II (L)	5	7	8	11	13
ST-III (L)	2	2	3	4	5
ST-III (k)	1.54	1.68	1.57	1.53	1.54

k for different SNRs is tabulated in Table I. The achievable capacities for Type III are shown in Figure 2 (left - continuous line). The performance of Type III is significantly better than Type II (which is not surprising, since Type II is a special case of Type III). Note that Type III also performs better than ST-Type II, even though the difference is only marginal for SNRs less than -5 dB.

Even though the fall in capacity with SNR for Type III methods is less severe than Type II methods, Type III methods can still gain marginally by trading of bandwidth for SNR (particularly for SNRs lower than 0 dB). This is illustrated by the dashed line for the Spread Transform Type III (ST-III) in Figure 2 (left). The optimal choice of parameters for ST-III involves choice of k and L. The optimal choices k and L for ST-III are shown in Table II. The difference between the performance of ST-II, Type III and ST-III is perhaps more evident in Figure 2 (right) which is a plot of $\frac{\sigma_{n_e}}{\sigma_n}$ vs. SNR, for ST-II, Type III and the ST-III methods. Note that for ideal methods the ratio should approach unity.

IV. CONCLUSIONS

In this paper we have addressed various options for signaling schemes for multimedia steganography. We revisited an information theoretic approach proposed and solved in [2] to design ideal oblivious data hiding schemes. However, such schemes were found to be impractical. We therefore, limited ourselves to practical schemes for which steganographic communications have the same complexity as conventional spread spectrum communications.

This restriction is imposed by splitting the overall signaling scheme into two *independent* parts. One part defines the signal constellation, and the other, a scalar embedding/ detection operation, *estimates the origin of the floating signal constellation.* In particular, this paper addressed various options for the scalar embedding and detector operators $\mathcal{E}, \mathcal{D}$. The data hiding capacities of many such practical schemes along with the capacity of (non-practical) ideal scheme (represented by stars) are plotted in Figure 2 (left). From Figure 2 (right), it is seen that such practical schemes can come within 3 dB of the impractical ideal scheme for which $\sigma_{n_e} = \sigma_n$.

Two solutions to the problem of rapidly deteriorating performance of Type II methods with reduction in SNR have been proposed in literature. Spread transform methods sacrifice bandwidth to improve SNR. On the other hand, Type III methods make use of non-invertible $\mathcal{E}, \mathcal{D}$. We revisited the spread transform method proposed in [13], and then expanded the work to address the issue of "optimally sacrificing bandwidth for increasing SNR" to improve the overall performance. Comparison of the performance of such optimal spread transform and Type III methods (Section III D) showed that Type III performed better than spread transform. It was also shown that spread transform methods used in *conjunction* with Type III methods improve the performance of Type III methods marginally.

REFERENCES

[1] M. Ramkumar, *Data Hiding in Multimedia : Theory and Applications*, Doctoral Dissertation, Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, January 2000. Available for download from http://www.njcmr.org/~mxr0096.

[2] M.H.M. Costa, "Writing on Dirty Paper", *IEEE Trans. on Information Theory*, **IT-29**, pp 439-441, May 1983.

[3] T. M. Cover, J. A. Thomas, *Elements of Information Theory*, Second Edition, John-Wiley and Sons Inc, 1991.

[4] S.I. Gel'fand, M.S. Pinsker, "Coding for Channel with Random Parameters," *Problems of Control and Information Theory*, vol **9**(1), pp 19-31, 1980.

[5] J. Chou, S. S. Pradhan, K. Ramchandran "On the duality between data hiding and distributed source coding", *Proc. 33rd Annual Asilomar conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1999.

[6] M.Ramkumar, A.N. Akansu, "Self-Noise Suppression Schemes for Blind Image Steganography", *Proc. of SPIE: Multimedia Systems and Applications II (Photonics East '99)*, vol **3845**, Boston, MA, Sep. 1999.

[7] P. Moulin, J. A. O'Sullivan, "Information-Theoretic Analysis of Information Hiding," preprint, Sep. 1999. Available from http://www.ifp.uiuc.edu/ moulin/paper.html.

[8] R.G. van Schyndel, A.Z. Tirkel, C.F. Osborne, "A Digital Watermark", *IEEE International Conference on Image Processing*, vol **2** , pp 86-90, 1994.

[9] H.-J.M. Wang, P.-C. Su, C.-C.J. Kuo, "Wavelet Based Digital Image Watermarking", *Optics Express*, vol **3**, No 12, pp 491 - 496, Dec 1998.

[10] M.Wu, B. Liu, "Watermarking for Image Authentication", *Proceedings of IEEE International Conference on Image Processing*, October 4-7, 1998, Chicago, Illinois, USA, vol **2**, pp 437 - 441.

[11] B. Chen, G.W. Wornell, "Digital Watermarking and Information Embedding Using Dither Modulation", *IEEE Second Workshop on Multimedia Signal Processing*, Redondo Beach, California, pp 273-278, Dec 1998.

[12] M.Ramkumar, A.N. Akansu, A.A.Alatan, "A Robust Data Hiding Schemes for Images Using DFT", *IEEE International Conference on Image Processing*, **II**, pp 211-215, October 1999.

[13] B. Chen, G. W. Wornell, " Provably robust digital watermarking," *Proc. of SPIE: Multimedia Systems and Applications II*, vol **3845**, Boston, MA, pp. 43-54, Sept. 1999.