Multiple Codebook Information Hiding

H. T. Sencar
Electrical and Computer
Engineering Department
New Jersey Institute of Technology
University Heights, Newark, NJ
07102-1982
e-mail: taha.sencar@njit.edu

M. Ramkumar
AVWAY.COM Inc.
IDT Corporation
520 Broad Street, Newark, NJ
07102
e-mail: ramkumar@corp.idt.net

A. N. Akansu
Electrical and Computer
Engineering Department
New Jersey Institute of Technology
& NJCMR
University Heights, Newark, NJ
07102-1982
e-mail: ali@megahertz.njit.edu

Abstract — **We present a new embedding methodology for data hiding applications by employing multiple codebooks for embedding the watermark signal. The use of multiple codebooks offers a freedom in the choice of the codeword that is more "friendly" with the host signal for a particular message to be conveyed. In the proposed scheme, embedder searches among a set of unitary transformations of the cover signal that maximizes the correlation between the embedded and detected watermark signals at a fixed embedding distortion. The transform bases set is known to the embedder and detector while particular transform basis used for embedding is not revealed to the detector. We also derive the closed form expressions for the upper bound on the probability of error in detecting the wrong watermark signal for both single and multiple codebook cases. Simulation results show that hiding rate is improved with the use of multiple codebooks.**

I. Introduction

Extensive availability of multimedia data and advanced processing tools brought data hiding applications into the realm of content providers and researchers to improve the tractability of their digital media assets. These concerns mainly evolved around issues like copyright protection and content authentication. Common to all of these motives is the likely presence of intruders willing to modify contents that has undergone such a processing with the intention of nullifying aforementioned efforts. Information hiding (Data hiding) solutions promise to address some of these concerns.

Gelfand *et al.* in [1] derived the capacity of a discrete memoryless channel with side information at the encoder such that at any transmission time encoder has the whole channel state information for all times. Their works have become a cornerstone for oblivious data hiding applications. Later research gained a considerable momentum by reinterpreting side information as the multimedia content that carries hidden information. Costa [2], was the first to present an information-theoretic analysis of the problem that applies to oblivious data hiding. He also proved that the capacity of the Gaussian channel is the same whether the side information is known to the decoder or not by using the results of [1] for *iid* Gaussian distributed cover-signal and side information.

In Ref. [3], Moulin *et al.* gave complementary analysis for hiding capacity considering all aspects of data hiding within a game theoretic perspective. That study also highlights some design criteria for practical systems. Cohen *et al.* [4] present a detailed discussion and results of hiding capacity assuming a Gaussian distributed cover-signal and squared error distortion.

Although fundamental limits of hiding rate are known for additive white Gaussian noise (AWGN) attack and mean squared error distortion measure, practical algorithms that achieve these limits are not well established yet. Costa in [2] outlined an encoder decoder structure that achieves the capacity by utilizing a codebook, although it was far beyond being practical. Ramkumar *et al.* [5], Eggers *et al.* [6] and Chen *et al.* [7] proposed schemes that are similar in principle and have better robustness vs. rate trade offs than the conventional data hiding methods.

Assuming a fixed distortion measure that is in compliance with the perceptual properties of the cover-signal, information hider has two degrees of freedom to improve hiding rate vs. robustness characteristics. These are of designing the codebooks and the embedder-detector set that utilizes them. In this paper, we investigate multiple codebook embedding technique. The use of multiple codebooks provides with the choice of the codebook that has favorable distortion properties. In general, embedders are nonlinear functions. Consequently, codewords from different codebooks corresponding to the same message may have different embedding distortions. In a typical application, embedding distortion is limited by some fixed value derived from the distortion measure, and the hider uses all available resources in order to increase the robustness. In other words, at the same level of embedding distortion, a message may be represented by codewords with different strengths and correspondingly different detection statistics. Embedder picks the codeword that is expected to yield highest correlation at the detector at a fixed embedding distortion. Each codebook is assumed to be generated through a unitary transformation of the cover-signal. We show that for AWGN channel, Gaussian distributed cover-signal and squared error distortion measure, the increase in probability of error due to use of multiple codebooks is compensated by an exponential reduction (in probability of error) due to the embedder's ability to adapt the codeword to the cover-signal. We derived closed form expressions for the upper bound on the probability of error in terms of the number of codebooks and codeword size. The embedder described in Ref. [5] is incorporated with the proposed methodology. However, the concept is applicable to a wide range of embedders.

In the text, we denote vectors with boldfaced characters, random variables with capital letters and their realizations with the corresponding lower case letters. In the next section

we present details of the data hiding model used. We describe the data hiding scheme in Section III and derive performance analysis methodology for the one codebook and multiple codebook embedding in the Section IV. Performance results are presented in Section V.

II. Data Hiding Model

In a generic data hiding application a message indexed by m, from an alphabet $\mathcal{M}$, $1 \leq m \leq M$, is mapped out to a sequence $W \in \Re^N$. Sequence W (watermark signal) must be embedded into the cover-signal, S, without any perceptual distortion. Embedder, $\mathcal{E}$, modifies signal S with respect to W within some distortion constraint and generates the stego-signal (watermarked signal) $\hat{S}$. The difference signal, X, between $\hat{S}$ and S is the embedding distortion corresponding to message m, $X = \hat{S} - S$. Detector, $\mathcal{D}$, extracts signal $\hat{W}$ from $\hat{S}$ or from an "attacked" version Y of $\hat{S}$. Signals S, W, X, $\hat{S}$, Y, and $\hat{W}$ are generalized to be vectors of length N. Embedder and detector may be scalar or vector operations that operates on these vectors based on the choice of designer.

By multiple codebook embedding we assume the presence of L number of $N \times N$ unitary transform bases at the embedder and detector

$$I = \mathbf{T}_i^T \mathbf{T}_i, \ i = 1, \ldots, L, \tag{1}$$

where I is the identity matrix and ^{T} is the matrix transpose operation. One selection criterion for $\mathbf{T}_i$, $i = 1, \ldots, L$, is that all transformations of a vector are maximally separated from each other in $\Re^N$ with respect to a pre-designated distance measure. Among L possible transformations of $S_i = \mathbf{T}_i S$, $i = 1, \ldots, L$, let k, $1 \leq k \leq L$ represent the index of transform basis which will be used for embedding. Uninformed of particular $\mathbf{T}_k$ used for embedding, detector generates L transforms of signal Y and extracts message in a blind manner.

One consequence of using multiple codebooks is that embedding is not strictly a scalar operation because for a message m to be conveyed choice of S_k determines signal vector $\hat{S}_k$. The overall information hiding system in an additive channel model is outlined below

$$\begin{aligned}
\mathcal{W} &: \quad m \longrightarrow W, \\
\hat{S_k} &= \mathcal{E}(\mathbf{T}_k S, W) = S_k + \hat{X}_k, \\
Y &= T_k^T(\mathcal{E}(\mathbf{T}_k S, W)) + Z = S + X_k + Z, \\
\hat{W}_i &= \mathcal{D}(\mathbf{T}_i^T Y), \ i = 1, \ldots, L \\
\mathcal{W}^{-1} &: \quad \hat{W}_i \longrightarrow \hat{m},
\end{aligned} \tag{2}$$

where $m \in \mathcal{M}$ is the index of the hidden message, $X = X_k$ is the distortion introduced by the embedder for the chosen transformation basis T_k, and Z is the intrusion of the attacker. $\mathcal{W}$ is a one to one mapping from m to W which transforms message m into a better representation for embedding. The embedder, $\mathcal{E}$, and the detector, $\mathcal{D}$, may be linear or nonlinear and not necessarily invertible functions ($\mathcal{D}(\mathcal{E}(S, W)) \neq W$).

Not evident in the model is the distortion constraints imposed on information hider and attacker. We assume mean squared error distance as the measure of distortions introduced by information hider and attacker. Although power of a difference signal is not a true distance in perceptual sense, it may be deployed in accordance with the findings of multimedia compression methods due to the ease in analytical tractability. Imposing restrictions on the distortions introduced by the information hider and attacker, such that these distortions have much less power than the cover signal S ($\frac{1}{N}\sum_{j=1}^{j=N} S_j^2 \gg \frac{1}{N}\sum_{j=1}^{j=N} X_j^2$ and $\frac{1}{N}\sum_{j=1}^{j=N} S_j^2 \gg \frac{1}{N}\sum_{j=1}^{j=N} Z_j^2$), will keep the original content more or less intact and simplify the problem.

Analysis of data hiding rate is developed by projecting the earlier theoretical studies in channel communication with side information. Gelfand *et. al* [1] considered a discrete memoryless channel with an input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$, both of which depend on a given side information from a finite set $\mathcal{S}$ where $\mathcal{X}, \mathcal{Y}, \mathcal{S}, \in \Re^N$. Channel capacity is expressed in terms of random variables $X \in \mathcal{X}$, $Y \in \mathcal{Y}$, $S \in \mathcal{S}$ and an additional auxiliary random variable $U \in \mathcal{U}$, $\mathcal{U}$ being a finite alphabet in $\Re^N$, given the conditional joint probability density $p(u, x|s)$ as

$$R = \max_{p(u,x|s)} (I(U, Y) - I(U, S)). \tag{3}$$

Costa in Ref. [2] had a design of $U = X + \alpha S$ by assuming codewords satisfying the power constraint $\frac{1}{N}\sum_{j=1}^{j=n} X_j^2 \leq P$ and independent random variables X, S, Z with probability distribution functions $X \sim \mathcal{N}(0, P)$, $S \sim \mathcal{N}(0, \sigma_S^2)$, $Z \sim \mathcal{N}(0, \sigma_Z^2)$, respectively. He showed that by setting $\alpha = \frac{P}{P+\sigma_Z^2}$ the optimal codebook that achieves the capacity $C = \frac{1}{2}\log_2(1 + P/\sigma_Z^2)$ is designed. This is the capacity of AWGN channel where S is known to both encoder and decoder.

In data hiding applications, channel is unpredictably nonlinear. Practically, it is impossible to model the channel considering the vast variety of attack scenarios and their combinations, [8]. As discussed in [3] and [4] assuming independent Gaussian distributed cover signal and codeword white Gaussian noise is the optimal attack. These assumptions provide means for characterizing hiding rate vs. robustness features of a method. In the rest of the analysis all intrusions of the attacker to watermarked signal is represented by AWGN. Corresponding hiding rates can be computed assuming an *Mary symmetric* channel with transition probabilities $\mathrm{p}(j|m) = 1 - \mathrm{P_e}$ for $j = m$ and $\mathrm{p}(j|m) = \frac{\mathrm{P_e}}{M-1}$ for $j \neq m$.

III. Embedder

Data hiding methods may be categorized into three main types based on how $(\mathcal{E}, \mathcal{D})$ are designed. **Type-I** methods are very common and simple to implement. Basically, stego-signal is generated by adding the watermark signal or a non-uniform scaled version of it to the cover signal. These methods suffer from dramatically low hiding rates because of the non-optimal design which assumes S as a noise and tries to cancel it. Type-I methods are preferable only when the attack is too severe. However, they are ideal only for applications for which the cover-signal is present at the detector.

Type-II methods are characterized by the use of quantizer structures in the embedding and detection, [9], [10], [11]. Unlike Type-I methods, watermark signal has restricted values and detection is a many to one mapping such that stego-signal values apart from each other at certain amounts have the same detected value with respect to a periodic pattern. The distortion, X, is a function of S and W. Also, the embedder, $\mathcal{E}$, and detector, $\mathcal{D}$, are inverses of each other. Disadvantage is that the system performs well only if the attack is not severe. These can also be employed with oblivious data hiding systems at considerable hiding rates.

Type-I and Type-II methods are equivalent to designs of $U = X$, $\alpha = 0$ and $U = X + \alpha S$, $\alpha = 1$, respectively. These two choices of U correspond to two extremes in hiding rate

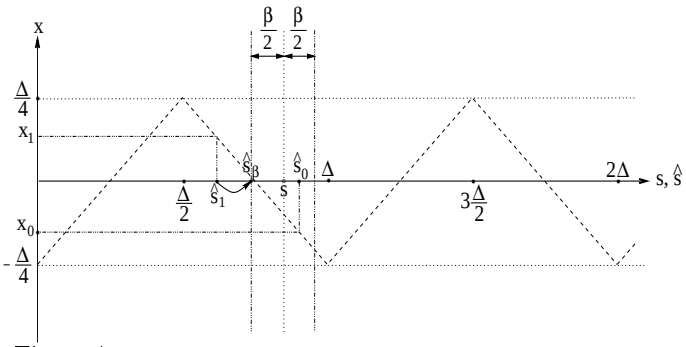

Figure 1: Representation for embedding the watermark signal into the cover-signal.

vs. robustness curves. Within the context of data hiding this fact may restated as Type-I and Type-II methods having preferable performances at "severe attack" and "no attack" cases. An optimal design is the one that designer has control over the operating characteristics of the method. In Refs. [5] [6] [7] modifications to the Type-II methods are proposed by removing the invertibility condition on the set $(\mathcal{E}, \mathcal{D})$. In **Type-III** methods added non-invertibility is designed in a particular way that hiding rate is maximized for a presumed attack level. Type-III methods utilize the design $U = X + \alpha S$, $0 < \alpha < 1$ which gives information hider a freedom to adapt the codeword to the channel. Type-III data hiding is optimal for oblivious data hiding applications.

The embedder being utilized by the data hiding technique is a quantizer characterized by a pair of parameters, period Δ and threshold β where $0 < \beta \leq \Delta$. The form of quantizer used for implementation is a periodic continuous triangular function. Watermark signal values are limited by the peak values of the periodic function. Embedding is a translation of the input coefficient values by introducing distortions thresholded to $\pm\frac{\beta}{2}$ such that the mapping of embedded coefficient over the periodic function has a minimum Euclidean distance to the watermark signal. The period Δ and threshold β of the quantizer are dictated by the preassigned embedding distortion, above which perceptual features of the cover signal will be considered changed. Among the Δ, β pairs that meet the distortion constraint, the one that maximizes hiding rate for a presumed distortion amount is picked. Detection of the watermark signal is similar to embedding. The stego-signal is mapped over the periodic function with the same Δ as used for embedding and with the fixed threshold $\beta = \Delta$, which adds non-invertibility to the $\mathcal{E}, \mathcal{D}$ set.

Figure 1 represents embedding of two different watermark signal coefficients, x_0 and x_1, to signal coefficient s. Embedding x_0 into s generates the stego-signal $\hat{s}_0$. Whereas, embedding x_1 into s generates $\hat{s}_\beta$ rather than $\hat{s}_1$ due to thresholding by $\pm\frac{\beta}{2}$.

As $\Delta \to \infty$, for some finite β, hiding rate vs. robustness characteristics of this scheme gets similar to Type-I methods. When $\Delta = \beta$, this scheme becomes a Type-II method. For all other fractions of $\frac{\beta}{\Delta}$, scheme performance is optimal for different attack powers. The distortion X introduced by this embedder is a non-linear function of S and W, $X = \mathcal{E}(S, W) - S$. However, it is statistics can be computed given the probability distributions of S and W. Similarly, the statistics of the distortion X_t on X as a consequence of the uninvertibility condition introduced by the Type-III method, due to thresholding X by $\beta < \Delta$, can be computed.

IV. Multiple Codebooks

A codebook is the collection of sequences, codewords, each of which is generated through an intelligent combination of the cover-signal S and watermark-signal W corresponding to one of the M possible messages (i.e. information of $\log_2 M$ bits to be conveyed.) Every codeword is required to comply with a perceptual distortion constraint. Ultimately, embedder generates the codebook for a fixed cover signal and a presumed attack level. Then, it picks the codeword that is pointed by the message index m (i.e. decimal number that $\log_2 M$ bit sequence corresponds to) which is consequently delivered to the channel by adding it to S. Detector's function is to decode the message m which might be distorted intentionally or unintentionally.

Use of multiple codebook helps embedder in choosing the codeword among an increased number of possible sequences all of which satisfy the power constraint. In other words, employing multiple codebooks corresponds to a simplified way of generating U sequences for practical applications where number of sequences in each bin (as described within random coding argument) is increased by the number of codebooks.

In a multiple codebook embedding each codebook is generated using a particular unitary transformation of the cover-signal for the same message set. The embedder that makes use of L codebooks embeds sequence W_m corresponding to message m with L transformations of the cover signal, $S_i = \mathbf{T}_i S$ $i = 1, \ldots, L$, consecutively. Then, it decides on the transform basis $\mathbf{T}_k$, $1 \leq k \leq L$ that maximizes the detection statistics. The codeword corresponding to the transform basis $\mathbf{T}_k$, $X_k = \mathcal{E}(\mathbf{T}_k S, W_m) - S_k$, is transmitted after backward transformation $\hat{S} = \mathbf{T}_k^T(X_k + S_k)$. Detector extracts $\hat{W} = \mathcal{D}(\mathbf{T}_i Y)$ for all k values, $1 \leq k \leq L$, from the received signal Y.

In a practical method, watermark signal detection is followed by matching the extracted signal to one of the known watermark signals in order to decide on the sent message. Given that the detector has no knowledge of the internal processing of the channel, use of normalized correlation is a practical approach but non-optimal. Normalized correlation is a similarity measure between two vectors which can be geometrically interpreted as the cosine of the angle between the vectors. Thus, detector computes the normalized correlation between the extracted vector $\hat{W}$ and all W sequences corresponding to M messages, and the message that yields the highest normalized correlation is defined as the sent message.

The embedder can decide on the transform basis $\mathbf{T}_k$, $1 \leq k \leq L$, for embedding in two ways. In the first one, sequence W_m is embedded into S_i , i=1,...,L, based on the *minimum distortion* criterion. The index that yields the smallest distortion, $k = \arg\min_i\{d_i\}$, $i = 1, \ldots, L$ where $d_i = \frac{1}{N}\sum_{j=1}^{j=N}(X_{i_j})^2$, is chosen as the index of the transform basis, T_k. Alternately, the embedder can use normalized correlation as the decision metric to choose the transform basis, *maximum correlation* criterion. In general, the amount of distortion that can be introduced to S is limited. In this case, for each S_i, embedding parameters are chosen such that the resulting embedding distortion, P_E, is the same. Since the embedding and detection functions are not the inverses of each other, the embedded watermark signal W_m and the detected watermark signal $\tilde{W}_m$ are not the same. The embedder picks the transform basis $\mathbf{T}_k$ that yields the highest correlation between W_m and $\tilde{W}_m$ at the embedder, $k = \arg\max_i\{\rho_i\}$, $i = 1, \ldots, L$ where $\rho_i = \frac{W_m^T \tilde{W}_m}{|W_m||\tilde{W}_m|}$. In this paper we use the

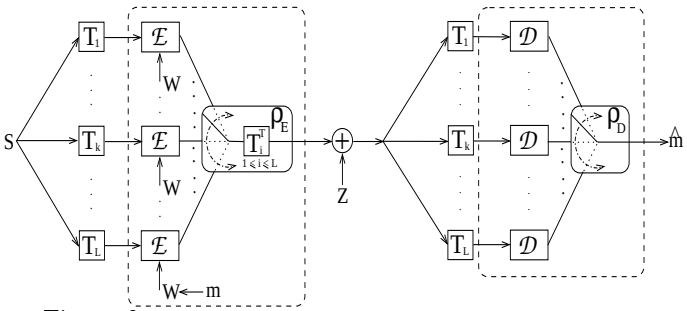

Figure 2: Multiple codebook embedding and detecting.

latter due to the ease in its analytical tractability.

Figure 2 displays L codebook embedding and detecting. In the block diagram W is the watermark signal to be embedded. Embedding block ρ_E decides on the transform basis $\mathbf{T}_i$, $1 \leq i \leq L$ for embedding by computing the normalized correlation. Then, it transmits the codeword corresponding to watermark signal W and cover-signal S. Detector ρ_D extracts the signal $\hat{W}_m$ and matches it to one of the watermark signals W_m, $1 \leq l \leq M$ with index $\hat{m}$. An error occurs whenever m and $\hat{m}$ are not the same.

IV.I Probability of Error for Single Codebook embedding

Let $\mathbf{W_m} = [W_{m,1}, \ldots, W_{m,N}]$ be a length N *iid* zero mean binary distributed random vector, representing the message m information hider is willing to convey and $\hat{\mathbf{W}}_\mathbf{m} = [\hat{W}_{m,1}, \ldots, \hat{W}_{m,N}]$ be the extracted real valued signal at the detector. $\hat{\mathbf{W}}_\mathbf{m}$ is also an *iid* zero mean random vector since embedding and detection are memoryless, cover-signal S is *iid* and channel noise is white zero mean. Furthermore, Type-II and Type-III embedder-detector sets will force $\hat{\mathbf{W}}_\mathbf{m}$ to be *iid* due to quantizers involved in both embedding and detection. When only one codebook is used embedder employs an $M \times N$ sized codebook composed of M length N codewords. Probability of error is the result of $\hat{\mathbf{W}}_\mathbf{m}$ having the highest correlation with any of $[\mathbf{W_1}, \ldots, \mathbf{W_M}]$ other than $\mathbf{W_m}$.

The normalized correlation $\rho_{i,j}$ is defined between the received watermark signal $\hat{\mathbf{W}}_\mathbf{i}$ and the signal $\mathbf{W_j}$ than an event E_j that detector will pick $\hat{m}$ as the detected message is denoted as

$$\mathrm{E_j} = \{\mathrm{p}(\rho_{m,j} \geq \rho_{m,m})\}, j = 1, \ldots, M \text{ and } j \neq m. \quad (4)$$

Therefore, the event E that detector makes an error is expressed as

$$\mathrm{E} = \bigcup_{j=1, j\neq m}^{M} \mathrm{E_j}. \quad (5)$$

Hence, the probability of detecting a wrong message is found as

$$\mathrm{P_e^{one}} = Pr\{E\} \leq \sum_{j=1, j\neq m}^{M} \Pr\{\mathrm{E_j}\}. \quad (6)$$

The upper bound (*union bound*) on probability of error for one codebook, $\mathrm{P_e^{one}}$, can be expressed as

$$\mathrm{P_e^{one}} \leq \sum_{j=1, j\neq m}^{M} \mathrm{p}(\rho_{\mathbf{m,j}} \geq \rho_{\mathbf{m,m}}). \quad (7)$$

Similarly, if $\mathbf{W_i}$ and $\mathbf{W_j}$ are two *iid* random vectors with zero covariance matrix and $\mathbf{W_i}$ is the embedded watermark signal than the extracted signal $\hat{\mathbf{W}}_\mathbf{i}$ will also have a zero covariance matrix with $\mathbf{W_j}$. Random variable $\rho_{\mathbf{i,j}}$ can be generalized to,

$$\rho_{\mathbf{i,j}} \sim \begin{cases} \mathcal{N}(0, \frac{1}{N}), \ 1 \leq i, j \leq M \text{ if } i \neq j \\ \mathcal{N}(m_{\rho_{dep}}, \sigma^2_{\rho_{dep}}), \ 1 \leq i, j \leq M \text{ if } i = j, \end{cases} \quad (8)$$

where $m_{\rho_{dep}}$ and $\sigma^2_{\rho_{dep}}$ can be computed given the statistics of X, X_t and Z.

In the rest of the analysis we will drop the first sub-script of $\rho_{i,j}$ and assume m is the index of the sent message for all cases. Eq. (7) can be rewritten using Eq. (8),

$$\begin{aligned} \mathrm{P_e^{one}} &\leq \sum_{j=1, j\neq m}^{M} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\rho_\mathbf{j}}(\rho_j \geq \rho_m) f_{\rho_\mathbf{m}}(\rho_m) d\rho_j d\rho_m, \\ &\leq \sum_{j=1, j\neq m}^{M} \int_{-\infty}^{\infty} \left(\int_{\rho_m}^{\infty} f_{\rho_\mathbf{j}}(\rho_j) d\rho_j \right) f_{\rho_\mathbf{m}}(\rho_m) d\rho_m. \quad (9) \end{aligned}$$

Inner integral in Eq. (9) can be expressed in terms of Gaussian Q function ($Q(x) = \frac{1}{\sqrt{(2\pi)}} \int_x^\infty e^{-\frac{t^2}{2}} dt$.) Additionally, since statistics of ρ_j is independent of the index j for $j \neq m$, sum operator in Eq. (9) can be dropped and it simplifies to

$$\mathrm{P_e^{one}} \leq (M-1) \int_{-\infty}^{\infty} Q_{\rho_j}(\rho_m \sqrt{N}) f_{\rho_\mathbf{m}}(\rho_m) d\rho_m. \quad (10)$$

IV.II Probability of Error for Multiple Codebook embedding

When L codebooks are used for embedding, one of the watermark signals $[\mathbf{W_1}, \ldots, \mathbf{W_M}]$ is embedded in one of the L transformations of the cover signal $S_i = \mathbf{T}_i S$, $1 \leq i \leq L$. We will use the superscript ^{i} to denote the transform basis $\mathbf{T}_i$ used for embedding and detection.

The embedder will decide on the transform basis $\mathbf{T}_i$, $1 \leq i \leq L$, based on the correlation between the embedded watermark signal W_m^i and detected watermark signal $\tilde{W}_m^i$ at the embedder. Since embedder-detector set is not invertible in the watermark signal W_m^i even in the absence of channel noise detected W_m^i will be different from $\tilde{W}_m^i$ due to thresholding noise X_t, $\mathcal{D}(\mathcal{E}(S_i, W_m^i)) \neq \hat{W}_m^i$. Among the watermark signals $[\mathbf{W_m^i}, \ldots, \mathbf{W_m^L}]$ the one that yields the highest correlation, $\max[\tilde{\rho}_\mathbf{m}^\mathbf{i}, \ldots, \tilde{\rho}_\mathbf{m}^\mathbf{L}]$, will be embedded where $\tilde{\rho}_\mathbf{m}^\mathbf{i}$ is the normalized correlation between $\mathbf{W_m^i}$ and $\tilde{\mathbf{W}}_\mathbf{m}^\mathbf{i}$. Accordingly, the index for transform basis is set by the index of the highest correlation, $\arg\max_i[\tilde{\rho}_\mathbf{m}^\mathbf{i}]$, $i = 1, \ldots, L$.

Assuming $\mathbf{W_m^k}$ is the sent signal embedded using basis $\mathbf{T}_k$, the detector will extract the watermark signals $[\hat{\mathbf{W}}_\mathbf{m}^\mathbf{i}, \ldots, \hat{\mathbf{W}}_\mathbf{m}^\mathbf{L}]$ from the L back transformations of the received signal, $Y_i = \mathbf{T}_i^T Y$, $1 \leq i \leq L$. Let $\rho_{m,j}^i$ represent the normalized correlation between the signal $\mathbf{W_m^k}$ embedded into S_k and the signal $\hat{\mathbf{W}}_\mathbf{m}^\mathbf{i}$ detected form Y_i. Among all indices i, j that maximize $\rho_{m,j}^i$ for $1 \leq j \leq M$ and $1 \leq i \leq L$, j is the detected message $\hat{m}$, $\hat{m} = \arg_j \max_{i,j}[\rho_{\mathbf{m,j}}^\mathbf{i}]$, $1 \leq j \leq M$ and $1 \leq i \leq L$.

Probability of error for multiple codebook embedding, $\mathrm{P_e^{mul}}$, is due to any of the normalized correlation values $\rho_{m,j}^i$, $1 \leq j \leq M$, $j \neq m$ and $1 \leq i \leq L$ being greater than $\rho_{max} = \max[\rho_{m,m}^1, \ldots, \rho_{m,m}^L]$. Compared to the one codebook case, probability of error is expected to increase with the number of codebooks because there are L times more normalized

correlation values that can exceed ρ_{max}. Defining $\rho^i_{m,j}$ as the normalized correlation between the received watermark signal $\hat{\mathbf{W}}^{\mathbf{i}}_{\mathbf{m}}$ generated using $\mathbf{T}_i$ and the signal $\mathbf{W_j}$ than an event E^i_j that detector will prefer $\hat{m}$ to m as the detected message is denoted as (similar to Eq. (4))

$$\mathrm{E}^{\mathrm{i}}_{\mathrm{j}} = \{\mathrm{p}(\rho^i_{m,j} \geq \rho_{max})\}, i = 1 \ldots, L,\ j = 1, \ldots, M \text{ and } j \neq m. \tag{11}$$

The event $\mathrm{E}^{\mathrm{mul}}$ that detector makes an error is

$$\mathrm{E}^{\mathrm{mul}} = \bigcup_{i=1}^{L} \bigcup_{j=1, j\neq m}^{M} \mathrm{E}^{\mathrm{i}}_{\mathrm{j}}. \tag{12}$$

Hence, the probability of detecting a wrong message is obtained as

$$\mathrm{P}^{\mathrm{mul}}_{\mathrm{e}} = Pr\{E^{mul}\} \leq \sum_{i=1}^{L} \sum_{j=1, j\neq m}^{M} \Pr\{\mathrm{E}^{\mathrm{i}}_{\mathrm{j}}\} \tag{13}$$

The union bound on the probability of error for multiple codebook embedding, $\mathrm{P}^{\mathrm{mul}}_{\mathrm{e}}$, can be found as

$$\mathrm{P}_{\mathrm{e}}{}^{mul} \leq \sum_{i=1}^{L} \sum_{j=1, j\neq m}^{M} \Pr(\rho^{\mathbf{i}}_{\mathbf{m},\mathbf{j}} \geq \rho_{\mathbf{max}}). \tag{14}$$

The advantage of multiple codebook embedding stems from the difference in the distributions of the random variables $\rho_{m,m}$ and ρ_{max} (in Eq. (7) and Eq. (14), respectively.)

The distributions of $\rho^{\mathbf{i}}_{\mathbf{m},\mathbf{j}}$ for $1 \leq j \leq M$ and $1 \leq i \leq L$, assuming message m is embedded using transform basis $\mathbf{T}_k$, is shown as

$$\rho^{\mathbf{i}}_{\mathbf{m},\mathbf{j}} \sim \begin{cases} \mathcal{N}(0, \frac{1}{N}),\ 1 \leq j \leq M \text{ if } i \neq k, \\ \mathcal{N}(0, \frac{1}{N}),\ 1 \leq j \leq M \text{ if } i = k \text{ and } j \neq m \\ \mathcal{N}(m_{\rho_{dep}}, \sigma^2_{\rho_{dep}}),\ 1 \leq j \leq M \text{ if } i = k \text{ and } j = m. \end{cases}$$

The probability density function of the r.v. ρ_{max} is determined using,

$$\rho_{max} = \max[\rho^{\mathbf{1}}_{\mathbf{m},\mathbf{j}}, \ldots, \rho^{\mathbf{L}}_{\mathbf{m},\mathbf{j}}], \tag{15}$$

where $\rho^{\mathbf{i}}_{\mathbf{m},\mathbf{j}}$ are *iid* Gaussian distributed random variables, $\rho^{\mathbf{i}}_{\mathbf{m},\mathbf{j}} \sim \mathcal{N}(m_{\rho_{dep}}, \sigma^2_{\rho_{dep}})$.

The probability of error for multiple codebooks given in Eq. (14) can be rewritten using the above results by dropping the first sub-script referring to sent message m,

$$\begin{aligned} \mathrm{P}^{\mathrm{mul}}_{\mathrm{e}} \quad &\leq \quad \sum_{i=1}^{L} \sum_{j=1, j\neq m}^{M} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\rho^{\mathbf{i}}_{\mathbf{j}}}(\rho^i_j \geq \rho_{max}) f_{\rho_{\mathbf{max}}}(\rho_{max}) \\ &\quad d\rho^i_j d\rho_{max}, \\ &\leq \quad \sum_{i=1}^{L} \sum_{j=1, j\neq m}^{M} \int_{-\infty}^{\infty} \left(\int_{\rho_{max}}^{\infty} f_{\rho^{\mathbf{i}}_{\mathbf{j}}}(\rho^i_j) d\rho^i_j \right) f_{\rho_{\mathbf{max}}}(\rho_{max}) \\ &\quad d\rho_{max}, \end{aligned} \tag{16}$$

where $\rho^{\mathbf{i}}_{\mathbf{j}} \sim \mathcal{N}(0, \frac{1}{N})$. Since the inner integral in Eq. (16) is the Gaussian Q function and does not depend on the index j, Eq. (16) can be simplified to

$$\mathrm{P}^{\mathrm{mul}}_{\mathrm{e}} \leq L(M-1) \int_{-\infty}^{\infty} Q_{\rho^i_j}(\rho_{max}\sqrt{N}) f_{\rho_{\mathbf{max}}}(\rho_{max}) d\rho_{max}. \tag{17}$$

V. Results

Figures 3, 4 and 5 display the *union bound* on the probability of error vs. robustness computed by numerically solving Eq. (10) and (17) for various codebook numbers and sizes of $M \times N$. The corresponding robustness measure $R = \frac{P_E}{\sigma^2_Z}$ is the ratio of the embedding distortion power to the channel noise distortion power. However, an exact comparison of single and multiple codebook embedding schemes is not possible for the actual probability of errors, results indicate that the upper bound on probability of error decreases exponentially to zero for multiple codebook embedding scheme.

We implemented multiple codebook embedding by designing a set of transform bases $T_1, \ldots, T_L$ using Givens rotations. Givens rotations provide orthogonal transformations in $\Re^{N\times N}$ that rotates each vector with a fixed angle. A particular transform basis T_k is obtained by the consecutive multiplication of $\frac{N(N-1)}{2}$ number of orthogonal matrices all with determinant 1 so that resulting T_k is unitary. Each orthogonal matrix is derived from the identity matrix by introducing $\cos\theta_k$ terms at (i,i) and (j,j) locations with $\sin\theta_k$ and $-\sin\theta_k$ terms at (i,j) and (j,i) locations in order to rotate (i,j) coordinate plane with the designated angle θ_k. Rotation angles $\theta_k,\ k = 1, \ldots, L$ can be chosen by uniformly sampling 2π, $\theta_k = (k-1)\frac{2\pi}{L}$.

Hadamard transform matrix of size $N \times N$ and its negated version are combined into $2N \times N$ binary valued matrix to generate the orthogonal watermark signals. Every row of the combined matrix is indexed from 1 to M and assigned to one of the watermark signal vectors $\mathbf{W_m}$, $1 \leq m \leq M = 2N$, such that $E[\mathbf{W_i W_j}] = 0, i \neq j$. Watermark signals are BPSK modulated and scaled to $\frac{\Delta}{4}$ and $-\frac{\Delta}{4}$, for maximum separation, before embedding. Setting watermark signal size to N and number of messages to M, the size of the codebooks that will be utilized by embedder is fixed to $M \times N$ where $M = 2N$.

We fixed the embedding distortion P_E and optimized the embedding parameter Δ for each $R = P_E/\sigma^2_Z$ value. The latter is also revealed to the detector. With proper selection of β value, P_E is adjusted to the designated distortion amount. We assumed the cover signal S and channel noise Z are *iid* zero mean Gaussian vectors with variances σ^2_S and σ^2_Z, respectively, satisfying $\sigma^2_S \gg P_E$ and $\sigma^2_S \gg P_Z = \sigma^2_Z$.

The simulations are done by embedding and detecting randomly chosen messages with the use of different number of codebooks L. The embedder chooses the message m, $1 \leq m \leq M$ and embeds the corresponding W_m vector of length N to the cover-signal S. Signal S is passed through AWGN channel with noise variance selected in way that $R = \frac{P_E}{P_Z}$ is satisfied for a range of R values. The detector extracts the signal $\hat{W}_m$ and uses normalized correlation to match it to message $\hat{m}$. If the extracted message $\hat{m}$ at the detector is same with m, it's called a success and otherwise an error. Resultant probability of success values are used to compute the hiding rate of the system within an *Mary symmetric channel* assumption.

We performed embedding with up to 14 codebooks and codebook sizes of 32×64, 64×128, 128×256. Results are evaluated within $0.1 \leq R \leq 0.8$ range of embedding power to noise power ratios. Figures 6 and 7 display the probability of success and corresponding hiding rates for L=4 and varying N values. The increase in the watermark signal size N improves the detection statistics because normalized correlation gives more reliable results with the larger signal sizes. Figures 8 and 9 display the probability of success and corresponding hiding rates for $N = 128$ and $L = 1, 3, 5, 14$.

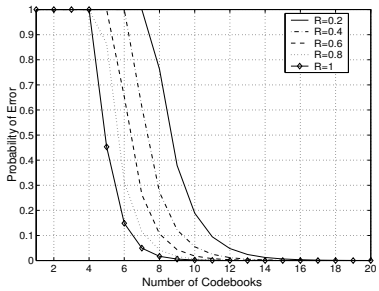

Figure 3: Multiple codebook embedding and detection, N=64 and M=128.

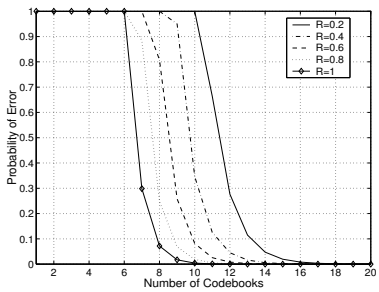

Figure 4: Multiple codebook embedding and detection, N=1280 and M=2560.

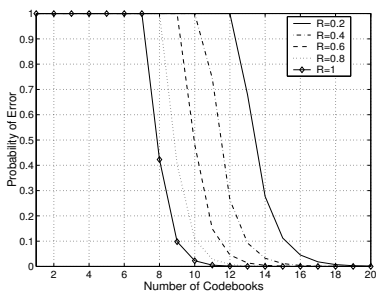

Figure 5: Multiple codebook embedding and detection, N=8192m and M=16384.

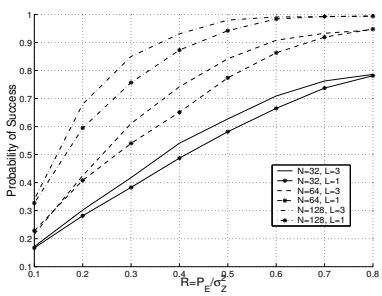

Figure 6: Probability of success performance for 4-codebook embedding and detection for various watermark signal sizes of $N = 32, 64, 128$.

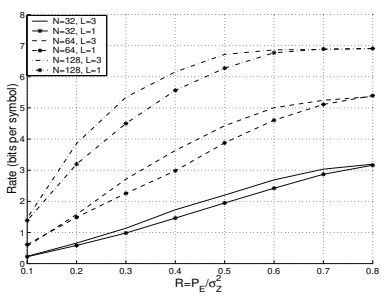

Figure 7: Data hiding rates for 4-codebook embedding and detection for various watermark signal sizes of $N = 32, 64, 128$.

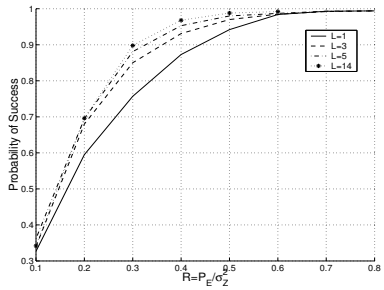

Figure 8: Probability of success performance for multiple codebook embedding and detection, $L = 1, 3, 5, 14$ and $N = 128$.

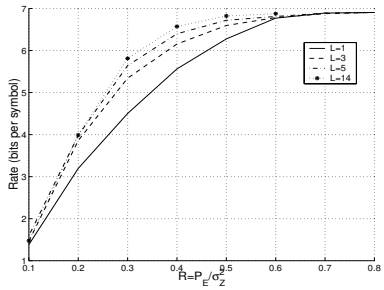

Figure 9: Data hiding rates for multiple codebook embedding and detection for $L = 1, 3, 5, 14$ and $N = 128$.

It is observed that for the data hiding scheme described in Section III, as the ratio of $\frac{P_E}{\sigma_Z^2}$ changes in between 0.1 and 0.8 multiple codebook embedding has higher hiding rates. However, the concept is trivially applicable to all Type-III data hiding schemes.

References

[1] S.I. Gel'fand and M.S. Pinsker, "Coding for Channel with Random Parameters," *Problems of Control and Information Theory*, Vol. 9, No. 1, pp. 19-31, 1980.

[2] M. Costa, "Writing on Dirty Paper," *IEEE Trans. Info. Thy.*, Vol. 29, No. 3, pp. 439-441, 1983.

[3] P. Moulin and J.A. O'Sullivan, "Information-Theoretic Analysis of Information Hiding," http://www.ifp.uiuc.edu/~moulin/paper.html

[4] A. S. Cohen and A. Lapidoth, "The Gaussian Watermarking Game," http://www.mit.edu/people/acohen/Pubs.

[5] M. Ramkumar and A. N. Akansu, "Self-Noise Suppression Schemes for Blind Image Steganography," *Proc. of SPIE: Multimedia Systems and Applications II (Photonics East'99)*, Vol. 3845, pp. 55-68, 1999.

[6] J. J. Eggers, J. K. Su, and B. Girod, "A Blind Watermarking Scheme Based on Structured Codebooks," *IEE Colloq. Secure Images and Image Authentication*, Vol. 4, pp. 1-6, 2000.

[7] B. Chen and G. W. Wornell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," *IEEE Tran. Info. Thy.*, Vol. 47, No. 4, pp. 1423-1443, 2001.

[8] M. Kutter and F. A. P. Petitcolas, "A fair benchmark for image watermarking systems,"*Proc. of SPIE: Security and Watermarking of Multimedia Contents*, Vol. 3657, pp. 226-239, 1999.

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

[10] M. Wu and B. Liu, "Watermarking for image authentication," *Proc. IEEE International Conference On Image Processing*, Vol. 2, pp. 437-441, 1998.

[11] B. Chen and G.W. Wornell, "Provably Robust Digital Watermarking," *Proc. IEEE Second Workshop on Multimedia Signal Processing*, pp. 43-54, 1998.