Multiple Codebook Information Hiding Based On Minimum Distortion Criterion

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
Computer Information Science Department
Polytechnic University
Six, Metrotech Center, Brooklyn, NY 11201
e-mail: ramkumar@vip.poly.edu

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

Abstract — **We present an alternate implementation of multiple codebook hiding method introduced in [1]. 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 this sense, multiple codebook hiding resembles optimal binning technique as it offers information hider a choice in the selection of the codeword and the extraction requires a search over all codebooks. With this method, the performance of practical oblivious hiding methods that employ scalar quantization procedures is improved through better codeword selection by introducing a dependency to the host signal in embedding process at an added complexity. In the proposed scheme, embedder searches among a set of unitary transformations of the host signal that minimizes the squared error distance between the embedded and detected message signals. 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 derive the upper bound on the probability of error in detecting the wrong message signal for both single and multiple codebook cases. Both analytical and simulation results show that hiding rate is improved with the use of multiple codebooks.**

I. Introduction

Information hiding addresses the determination of achievable limits and the design of methods for conveying a message signal through a host signal in a reliable and transparent manner. One conservative assumption of information hiding is that the detector has no access to the host signal, namely oblivious information hiding (also known as public or blind information hiding). Analysis of oblivious information hiding problem is developed by projecting the earlier theoretical studies in communication with side information and combining them with game-theoretic formulations. From power constrained channel communications perspective, this problem is analogous to the one where encoder wants to set a reliable communication with decoder at the highest transmission rate in the presence of some side information, in the form of channel's state, that's only available to the encoder. However, this approach does not include the intelligent operations that maliciously intends to distort the information hidden signal (stego signal). Therefore, information hiding is modeled as a mutual information game between encoder-decoder (embedder-detector) and attacker whose exact formulation depends on the amount of knowledge one party has about the strategy of the other and vice versa.

Shannon [2], introduced the first discrete memoryless channel with side information, in form of varying channel states from a finite set, casually known to the encoder. He proved that this channel is equivalent (in terms of capacity) to a channel with no side information, employing the same output alphabet and an expanded input alphabet. Gelfand *et al.* in [3] derived the capacity of a similar channel by removing the casuality condition at the encoder such that at any transmission time encoder has the whole channel state information for all times. Costa [4], was the first to present an information-theoretic analysis of a problem that applies to oblivious information hiding. He proved that the transmission capacity of a additive white Gaussian noise (AWGN) channel with the side information at the encoder is the same whether this information is known to the decoder or not for Gaussian distributed state information and codeword assumptions and calculated that the corresponding capacity is equal to that of an AWGN channel with the same signal to noise ratio. Later research gained a considerable momentum by reinterpreting these results within the context of oblivious information hiding. In Refs. [5] [6], the oblivious information hiding problem is formulated as a mutual information game between embedder-detector and the attacker. It is shown that the solution for hiding capacity varies with the setting of the game and Costa's framework corresponds to the upper bound on the coding capacity of all versions of the game since attacker has a fixed strategy (AWGN) and this is known to both encoder and decoder.

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 [4] outlined an encoder decoder structure that achieves the capacity through the use of optimal codebook. Despite the impractical solution, Costa's framework proved useful in categorizing the existing practical hiding methods according to their performances based on the type of codebook they utilize. Refs. [7] [8] [9] [10] proposed schemes that are similar in principle and have better robustness vs. rate trade offs than the conventional information hiding methods (*i.e.* spread transform methods).

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. Each codebook is assumed to be generated through a unitary transformation of the host signal. Among a set of unitary transformations of the host signal embedder picks the one that is expected to yield minimum squared error distance between the embedded message signal and the extracted signal in the detector at a fixed embedding distortion and a presumed noise level. The set of transformation bases are known to both embedder and detector, however detector has to extract the message signal from a number of transformations of the received stego signal (or a distorted version of it) in a blind manner. We show that for AWGN channel, Gaussian distributed host signal and squared error distortion measure, the increase in probability of error due to use of multiple codebooks is compensated by a reduction (in probability of error) due to the embedder's ability to adapt the codeword to the host signal. We derived the upper bound on the probability of error in detecting a wrong message signal in terms of the number of codebooks and codeword size. The embedding-detection scheme described in Ref. [9] 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 analysis we used the notation given in Table 1. For the general case all signals are assumed to be random vectors of size N. In the next section we present the motivation behind the multiple codebook hiding. We describe the information hiding model and the utilized scheme in Section III and derive performance analysis methodology for the one codebook and multiple codebook hiding cases in Section IV. Performance results are presented in Section V.

$\mathbf{S}$	Host signal vector
$\mathbf{X}$	Codeword
$\hat{\mathbf{S}}$	Information hidden signal vector
$\mathbf{Z}$	Channel noise vector
$\mathbf{Y}$	Distorted $\hat{S}$
$\mathbf{W}_m$	Signal vector corresponding to message m to be conveyed
$\hat{\mathbf{W}}_m$	Extracted signal vector when $\mathbf{W}_m$ is embedded
$d_{m,j}$	The squared error distance between $\hat{\mathbf{W}}_m$ and $\mathbf{W}_j$
$\hat{\mathbf{W}}_m^i$	Extracted signal vector from T_i transformation of $\mathbf{Y}$ when $\mathbf{W}_m$ is embedded
$\tilde{\mathbf{W}}_m^i$	Extracted signal vector from T_i transformation of $\hat{\mathbf{S}}$ when $\mathbf{W}_m$ is embedded
$d_{m,j}^i$	The squared error distance between $\hat{\mathbf{W}}_m^i$ and $\mathbf{W}_j$
$\tilde{d}_{m,m}^i$	The normalized correlation between $\tilde{\mathbf{W}}_m^i$ and $\mathbf{W}_m$

Table 1: The notation used in the analysis

II. Motivation

Practical information hiding methods can be grouped into three based on the design of the codebook within the framework introduced by Costa. In type-I methods stego signal is generated by adding a mapping of the message signal or a non-uniform scaled version of it to the host signal. These methods suffer dramatically from low hiding rates due to the non-optimal design which assumes S as a noise and tries to cancel it. Therefore, they are only preferable when the attack is too severe. Type-II methods are characterized by the use of quantizer structures in embedding and detection. The performance of this type of methods is optimal, immune to host signal interference, only if there's no attack and it degrades drastically as the attack gets severe. The shortcoming in codebook design of type-I and type-II methods is due to the channel independent nature unlike the optimal codebook

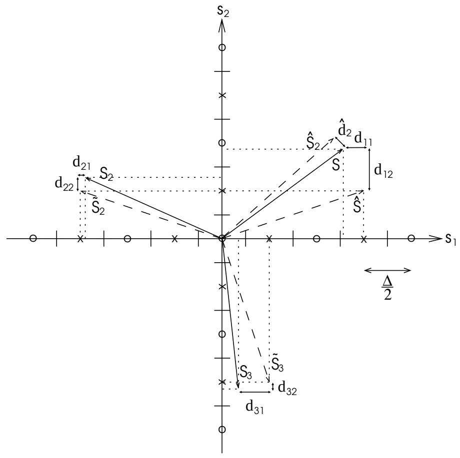

Figure 1: Representation for 3 codebook embedding of two binary symbols into the host signal vector $S = (s_1, s_2)$ using scalar quantizers.

design described in [4]. Type-III methods enable information hider in designing the operation characteristics of the methods so that the performance, in terms of hiding rate, is maximized based on the noise level. Therefore, type-III is the generalization of type-I and type-II.

The solution of a dual problem, source coding with side information at the decoder, in distributed source coding is adapted to oblivious information hiding in [7]. The codebook generation follows that of the Costa's optimal codebook through the use of optimal quantizers. The schemes proposed in [8] [9] [10] employed scalar quantization procedures for codebook generation. These methods fall into type-III categorization and compensate the drawback of type-II methods by enhancing the functionality of embedder by added processing such as thresholding and distortion compensation. In these methods type-II embedded signal undergoes the particular processing which is adjusted to maximize the hiding rate at the expected noise level. Due to simple and host signal independent structures, they're suitable for practical implementation.

Multiple codebook hiding improves hiding rate vs. robustness performance of type-III methods by better codeword selection through deciding on a transformation of the host signal for embedding. The essence of multiple codebook hiding is as follows. In Fig. 1 two binary symbols corresponding to $\times$ are embedded into the host signal vector $S = (s_1, s_2)$ and into two transformed (rotated) versions S_2 and S_3 trough the use of scalar quantizers. Embedding is the translation of the host signal to the nearest centroid of corresponding symbol (type-II embedding). Among the three embedded signals $\hat{S}$, $\tilde{S}_2$, and $\tilde{S}_3$, with the corresponding embedding distortions $\sqrt{d_{11}^2 + d_{12}^2}$, $\sqrt{d_{21}^2 + d_{22}^2}$ and $\sqrt{d_{31}^2 + d_{32}^2}$, the one that yields the smallest embedding distortion in the signal domain is picked which in this case is the stego signal $\hat{S}_2$ with an embedding distortion of $\hat{d}_2$. Correspondingly, it's possible to embed a message symbol into the host signal at lower distortion levels with an added complexity of transformation. When this phenomenon is incorporated with type-III methods the advantage due to multiple codebook hiding arises.

For a given embedding distortion, type-III methods improve the performance of type-II methods at an expected noise level by increasing the distance between the centroids, so that higher noise levels can be sustained. The consequent increase in the embedding distortion is balanced through processing distortion as a result of which permitted embedding distortion amount is maintained. Since it's possible to embed the message symbols at lower embedding distortions, as shown in Fig. 1, type-III embedder can make use of this either by further increasing the the separation between the centroids at a fixed processing distortion or by reducing the the processing distortion at a fixed separation of centroids. In both cases, the detection performance improves as a result.

With multiple codebook hiding, the source of the improvement in the performance is due to the selection of a transformation of the host signal which enables embedding with lower distortions than the permitted amount. On the other hand, since the transformation used for embedding is not known at the detector (but the set of available transformations), message signal extraction is done from all transformations of the received signal, therefore, probability of detection error increases. In the following sections, we show that the swing in the balance between the opposing factors of the decrease in probability of detection error due to the increase in the separation of centroids at the embedder and of the increase in the probability of detection error due to extraction from among a number of transformations at the detector is in favor of the former.

III. Information Hiding Model

In a generic information hiding scenario a message indexed by m, from an alphabet $\mathcal{M}$, known to both embedder and detector with $1 \leq m \leq M$, is mapped out to a sequence $\mathbf{W} \in \Re^N$. Sequence $\mathbf{W}$ must be embedded into the host signal, $\mathbf{S}$, without any perceptual distortion. Embedder, $\mathcal{E}$, modifies signal $\mathbf{S}$ with respect to $\mathbf{W}$ within the distortion constraint and generates the stego signal $\hat{\mathbf{S}}$. The difference signal, $\mathbf{X}$, between $\hat{\mathbf{S}}$ and $\mathbf{S}$ is the embedding distortion (codeword) corresponding to message m, $\mathbf{X} = \hat{\mathbf{S}} - \mathbf{S}$. Detector, $\mathcal{D}$, extracts signal $\hat{\mathbf{W}}$ from $\hat{\mathbf{S}}$ or from an "attacked" version $\mathbf{Y}$ of $\hat{\mathbf{S}}$. Embedding and detection 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

$$I = T_i^T T_i, \; i = 1, \ldots, L, \tag{1}$$

where I is the identity matrix and ^{T} is the matrix transpose operation. One selection criterion for 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 $\mathbf{S_i} = T_i \mathbf{S}$, $i = 1, \ldots, L$, let k, $1 \leq k \leq L$ represent the index of transform basis which will be used for embedding. Uninformed of particular T_k used for embedding, detector generates L transforms of signal $\mathbf{Y}$ and extracts message in a blind manner.

The overall multiple codebook information hiding is outlined below in an additive channel model

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

where $m \in \mathcal{M}$ is the index of the hidden message, $\mathbf{X} = \mathbf{X}_k$ is the distortion introduced by the type-III embedder for the chosen transformation basis T_k, and $\mathbf{Z}$ is the intrusion of the attacker. $\mathcal{W}$ is a one to one mapping from m to $\mathbf{W}$ which transforms message m into a better representation for embedding. The signal $\mathbf{W}$ is embedded into $\mathbf{S}_k$, one of the L transformations $\mathbf{S}_i = T_i \mathbf{S}$, $i = 1, \ldots, L$ whose selection criterion will be elaborated in the following section. Then, the stego signal $\hat{\mathbf{S}}_k$ is inverse transformed to signal domain, $\mathbf{S} + \mathbf{X}_k$. At the detector the signal $\mathbf{Y}$, which is the distorted stego signal with the additive noise $\mathbf{Z}$, is transformed with all transformation bases and the signal $\hat{\mathbf{W}}$ is extracted and mapped to $\hat{m}$, without knowing $\mathbf{S}_k$ is the embedded signal. 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 $\mathbf{S}_k$ determines signal vector $\hat{\mathbf{S}}_k$.

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 a perceptual sense, it may be deployed in accordance with the findings of compression methods due to the ease in analytical tractability. In the rest of the analysis, embedding and attack distortions are quantified as $P_E = \frac{1}{N}||\mathbf{X}||^2$ and $P_Z = \frac{1}{N}||(\mathbf{Y} - \hat{\mathbf{S}})||^2$, respectively. Imposing restrictions on the distortions introduced by the information hider and attacker, such that these distortions have much less power than the host signal $\mathbf{S}$ ($\frac{1}{N}||\mathbf{S}||^2 \gg P_E$ and P_Z), will keep $\mathbf{S}$ more or less intact and simplify the problem [5].

As discussed in Ref. [5] and Ref. [6] for an i.i.d. Gaussian distributed host signal and memoryless embedding scheme, white Gaussian noise is the worst (optimal) additive attack. These assumptions on the host signal and attack are in accordance with the ones of [4] and provide means for characterizing hiding rate vs. robustness features of a given method. In the rest of the analysis, all intrusions of the attacker to watermarked signal is represented by AWGN.

The results of multiple codebook hiding are based on the type-III method described in [9], however, the methodology applies applies to all type-III methods. The method considers binary signal embedding through scalar quantization with thresholding as the processing. The embedding quantizers are characterized by two parameters. Namely, quantization interval Δ and the threshold β where $0 < \beta \leq \Delta$. Embedding is defined as the translation of the host signal in the direction of the nearest centroid, associated with the symbol to be embedded, with the amount of shifting limited to $\pm\frac{\beta}{2}$. The extraction of the message signal is by mapping the stego signal (or the distorted version) over a periodic triangular function whose peak locations coincide with the centroids of embedding quantizers corresponding to binary symbols.

Fig. 2 displays the embedding process of two signals, $w_0 = -\frac{\Delta}{4}$ and $w_1 = \frac{\Delta}{4}$, to the signal s. Embedding w_0 into s generates the $\hat{s}_0$. Whereas, embedding w_1 into s generates $\hat{s}_\beta$ rather than $\hat{s}_1$ due to the thresholding by $\frac{\beta}{2}$. Similarly, the extracted signals corresponding to $\hat{s}_0$ and $\hat{s}_\beta$ are $\hat{w}_0 = -\frac{\Delta}{4}$ and $\hat{w}_1 < \frac{\Delta}{8}$, respectively. The error in the latter is due to the thresholding processing which which adds non-invertibility to the $(\mathcal{E}, \mathcal{D})$ pair.

IV. Single and Multiple Codebook Embedding-Detection

In the considered single codebook hiding scenario, a binary signal sequence $\mathbf{W}_m$ corresponding to the message m, where

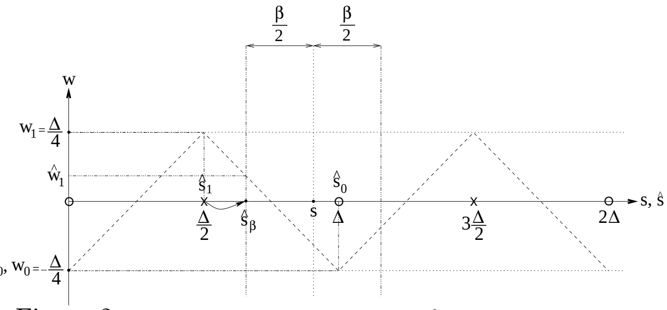

Figure 2: Embedding and detection of w_0 and w_1 into s.

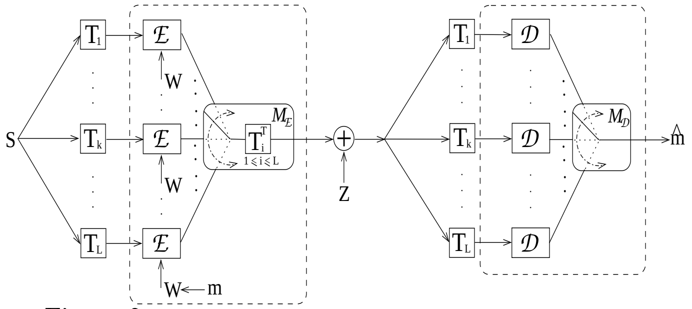

Figure 3: Multiple codebook embedding and detecting.

the whole message set and the corresponding binary sequences are known to both embedder and detector, is embedded into the host signal $\mathbf{S}$ by a type-III hiding method. At the detector, the signal $\hat{\mathbf{W}}_m$ is extracted from the received distorted signal $\mathbf{Y}$, $\hat{\mathbf{W}}_m = \mathcal{D}(T_i\mathbf{Y})$ and matched to the corresponding message $\hat{m}$.

In multiple codebook hiding, $\mathbf{W_m}$ is embedded into L transformations of the host signal, $\mathbf{S}_i = T_i\mathbf{S}$ $i = 1, \ldots, L$, by the same type-III method. Since the embedding and detection functions of type-III methods are not inverses of each other, the signal $\mathbf{W}_m$ embedded into $\mathbf{S}_i$ and the extracted signal $\tilde{\mathbf{W}}^i_m$ are not the same. Therefore, embedder has to decide on the transformation basis T_k, $1 \leq k \leq L$, that'll maximize the detection statistics. We proposed two ways for choosing the transformation basis T_k for embedding. In the first one, the sequence $\mathbf{W}_m$ is embedded into $\mathbf{S}_i$, $i = 1, \ldots, L$, based on the *minimum distance* criterion. The index that yields the smallest mean squared error distance between $\mathbf{W}_m$ and $\tilde{\mathbf{W}}^i_m$, $k = \arg\min_i\{d_i\}$, $i = 1, \ldots, L$ where $d_i = \| \mathbf{W}_m - \tilde{\mathbf{W}}^i_m \|$, 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. The embedder picks the transform basis T_k that yields the highest correlation between $\mathbf{W}_m$ and $\tilde{\mathbf{W}}^i_m$ at the embedder, $k = \arg\max_i\{\rho_i\}$, $i = 1, \ldots, L$ where $\rho_i = \frac{\mathbf{W}_m^T\tilde{\mathbf{W}}^i_m}{|\mathbf{W}_m||\tilde{\mathbf{W}}^i_m|}$ as discussed in [1]. In this paper we use the former approach.

Within multiple codebook hiding, the stego signal is generated by inverse transforming the embedded signal, $\mathbf{S}_k + \mathbf{X}_k = \mathcal{E}(T_k\mathbf{S}, \mathbf{W}_m)$, back to signal domain, $T_k^T(\mathbf{S}_k + \mathbf{X}_k)$. The codeword corresponding to transform basis T_k for the message m to be conveyed is $\mathbf{X} = T_k^T\mathbf{X}_k$. Detector extracts $\hat{\mathbf{W}}^i_m = \mathcal{D}(T_i\mathbf{Y})$ for all k, $1 \leq k \leq M$ and computes the distance metrics between the message signals $\mathbf{W}_j$, $1 \leq j \leq M$, and the extracted message signals. Detected message $\hat{m}$ is the one that yields the smallest distance.

Fig. 3 displays an L codebook embedding and detection scheme. In the block diagram, $\mathbf{W}$ is the watermark signal to be embedded. Embedding block $M_{\mathcal{E}}$ decides on the transform basis T_i, $1 \leq i \leq L$ to be used for embedding by computing the distances. Then, it transmits the codeword corresponding to the signal $\mathbf{W}$ and the host signal $\mathbf{S}$. The detector $M_{\mathcal{D}}$ extracts the signal $\hat{\mathbf{W}}$ and matches it to one of the signals $\mathbf{W}_j$, $1 \leq j \leq M$ with index $\hat{m}$. A detection error occurs whenever m and $\hat{m}$ are not the same.

IV.I **Probability of Error for Single Codebook Embedding**

Let $\mathbf{W}_m^T = [W_{m_1}, \ldots, W_{m_N}]$ be a length N i.i.d. zero mean binary random vector, representing the message m and $\hat{\mathbf{W}}_m^T = [\hat{W}_{m_1}, \ldots, \hat{W}_{m_N}]$ be the extracted real valued signal at the detector. $\hat{\mathbf{W}}_m$ is also an i.i.d. zero mean random vector since the embedding and detection processes are memoryless, host signal $\mathbf{S}$ is i.i.d. and channel noise is white and zero mean. When only one codebook is used, the embedder employs an $M \times N$ sized codebook composed of M length of N codewords. The probability of error is the result of $\hat{\mathbf{W}}_m$ having the smallest distance with any of $[\mathbf{W}_1; \ldots; \mathbf{W}_M]$ other than $\mathbf{W}_m$.

The distance $d_{m,j}$ is defined between the received watermark signal $\hat{\mathbf{W}}_m$ and the signal $\mathbf{W}_j$. Then, an event E_j that detector will pick $\hat{m}$ as the detected message m is denoted as

$$\mathrm{E}_j = \{\mathrm{p}(d_{m,j} \leq d_{m,m})\}, j = 1, \ldots, M \text{ and } j \neq m. \quad (3)$$

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

$$\mathrm{E} = \bigcup_{j=1, j\neq m}^{M} \mathrm{E}_j. \quad (4)$$

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

$$\mathrm{P}_e^{one} = \Pr\{\mathrm{E}\} \leq \sum_{j=1, j\neq m}^{M} \Pr\{\mathrm{E}_j\} \quad (5)$$

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

$$\mathrm{P}_e^{one} \leq \sum_{j=1, j\neq m}^{M} \mathrm{p}(d_{m,j} \leq d_{m,m}). \quad (6)$$

In Eq. (6), $d_{m,j}$ and $d_{m,m}$ are random variables corresponding to computed distance metrics for two separate cases. First case occurs, when the extracted signal vector has zero covariance matrix with the message signal vector. The statistics of the metric $d_{m,j}$ between $\hat{\mathbf{W}}_m$ and $\mathbf{W}_j$, $\forall\ m, j \mid m \neq j$, can be computed for large N using Central Limit Theorem as $d_{m,j} \sim \mathcal{N}(\frac{\Delta^2}{12}, \frac{\Delta^4}{N180})$. The second case refers to computation of the distance metric between two somewhat dependent vectors. Statistics of $d_{m,m}$, $\forall m$, can be computed for the given Δ, β and expected noise variance while knowing the internal processing of the embedder detector. Therefore, the random variable (r.v.) $d_{m,j}$ is generalized as

$$d_{m,j} \sim \begin{cases} \mathcal{N}(\frac{\Delta^2}{12}, \frac{\Delta^4}{N180}),\ 1 \leq m, j \leq M \text{ if } m \neq j \\ \mathcal{N}(m_{d_{dep}}, \sigma^2_{d_{dep}}),\ 1 \leq m, j \leq M \text{ if } m = j. \end{cases} \quad (7)$$

In the rest of the analysis, we will drop the first subscript of $d_{m,j}$ and assume m is the index of the transmitted message for all the cases for the sake of generality. Eq. (6) can be rewritten using Eq. (7) as

$$\mathrm{P}_\mathrm{e}^\mathrm{one} \quad \leq \quad \sum_{j=1, j\neq m}^{M} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{d_j}(d_j \leq d_m) f_{d_m}(d_m)\, dd_j\, dd_m,$$

$$\leq \sum_{j=1,j\neq m}^{M} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{d_m} f_{d_j}(d_j) dd_j \right) f_{d_m}(d_m) dd_m. \quad (8)$$

Inner integral in Eq. (8) can be expressed in terms of Gaussian Q function where $Q(x) = \frac{1}{\sqrt{(2\pi)}} \int_x^\infty e^{-\frac{t^2}{2}} dt$. Additionally, since statistics of d_j is independent of the index j for $j \neq m$, the sum operator in Eq. (8) can be dropped and P_e^{one} simplifies to

$$\mathrm{P}_e^{one} \leq (M-1) \int_{-\infty}^{\infty} \left(1 - Q\left(\frac{d_m - m_{d_{ind}}}{\sigma_{d_{ind}}}\right) \right) f_{d_m}(d_m) dd_m. \quad (9)$$

IV.II **Probability of Error for Multiple Codebook Embedding**

Assuming $\mathbf{W}_m$ is the embedded watermark signal by using the basis T_k, the detector will extract the signals $[\hat{\mathbf{W}}_m^i; \ldots; \hat{\mathbf{W}}_m^L]$ from the L transformations of the received signal, $\mathbf{Y}_i = T_i \mathbf{Y}$, $1 \leq i \leq L$. Let $d_{m,j}^i$ represent the distance between the signal $\mathbf{W}_m$ embedded into S_k and the signal $\hat{\mathbf{W}}_m^i$ detected form Y_i. Among all index pairs (i, j), the j index of the pair that minimizes $d_{m,j}^i$, for $1 \leq j \leq M$ and $1 \leq i \leq L$, is the detected message $\hat{m}$, $\hat{m} = \arg_j \min_{i,j}[d_{m,j}^i]$, $1 \leq j \leq M$ and $1 \leq i \leq L$.

Probability of error for multiple codebook embedding, P_e^{mul}, is due to any of the distance values $d_{m,j}^i$, $1 \leq j \leq M$, $j \neq m$ and $1 \leq i \leq L$, being smaller than $d_{min} = \min[d_{m,m}^1, \ldots, d_{m,m}^L]$. Compared to the one codebook case, probability of error is expected to increase with respect to the number of codebooks since there are L times more distance values that may be smaller than d_{min}, On the other hand, as d_{min} is expected to have lower mean than $d_{m,m}$ the probability of error for each comparison of the distances will be reduced. Defining $d_{m,j}^i$ as the distance between the received watermark signal $\hat{\mathbf{W}}_m^i$, extracted using T_i, and the signal $\mathbf{W}_j$, then, an event E_j^i that the detector will pick $\hat{m}$ instead of m is denoted as (similar to Eq. (3))

$$\mathrm{E}_j^i = \{\mathrm{p}(d_{m,j}^i \leq d_{min})\}, i = 1\ldots, L,\ j = 1, \ldots, M \text{ and } j \neq m. \quad (10)$$

The event E^{mul} that the detector makes an error is,

$$\mathrm{E}^{mul} = \bigcup_{i=1}^{L} \bigcup_{j=1,j\neq m}^{M} \mathrm{E}_j^i. \quad (11)$$

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

$$\mathrm{P}_e^{mul} = \Pr\{\mathrm{E}^{mul}\} \leq \sum_{i=1}^{L} \sum_{j=1,j\neq m}^{M} \Pr\{\mathrm{E}_j^i\} \quad (12)$$

The union bound on the probability of error for multiple codebook embedding, P_e^{mul}, can be found as

$$\mathrm{P}_e^{mul} \leq \sum_{i=1}^{L} \sum_{j=1,j\neq m}^{M} \Pr(d_{m,j}^i \leq d_{min}). \quad (13)$$

The advantage of multiple codebook embedding stems from the difference in the distributions of the random variables $d_{m,m}$ and d_{min} (in Eq. (6) and Eq. (13) respectively).

The distribution of the random variable $d_{m,j}^i$ can be found, $1 \leq j \leq M$ and $1 \leq i \leq L$, for the general case by considering a message m is embedded through the use of transform basis $\mathbf{T}_k$. When the embedding and extraction are done at the same transform domain, T_k is the transform basis used for both embedding and detection, then $d_{m,j}^k$ reduces to $d_{m,j}$ and becomes equivalent in statistics. For the case embedding and detection transform basis are different, extracted signal will be irrelevant to all message signals and can be considered to be independent with all message signals. Hence, distribution function of $d_{m,j}^i$ is expressed as,

$$d_{m,j}^i \sim \begin{cases} \mathcal{N}(\frac{\Delta^2}{12}, \frac{\Delta^4}{N180}),\ 1 \leq j \leq M \text{ if } i \neq k, \\ \mathcal{N}(\frac{\Delta^2}{12}, \frac{\Delta^4}{N180}),\ 1 \leq j \leq M \text{ if } i = k \text{ and } j \neq m, \\ \mathcal{N}(m_{d_{dep}}, \sigma_{d_{dep}}^2),\ 1 \leq j \leq M \text{ if } i = k \text{ and } j = m. \end{cases} \quad (14)$$

The probability density function of the r.v. d_{min} is given as

$$d_{min} = \min[d_{m,m}^1, \ldots, d_{m,m}^L], \quad (15)$$

where $d_{m,m}^i$ are independent and Gaussian distributed random variables, $d_{m,j}^i \sim \mathcal{N}(m_{d_{dep}}, \sigma_{d_{dep}}^2)$.

The probability of error for multiple codebooks given in Eq. (13) can be rewritten using the above results by dropping the first subscript referring to the transmitted message m as

$$\begin{aligned} \mathrm{P}_e^{mul} &\leq \sum_{i=1}^{L} \sum_{j=1,j\neq m}^{M} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{d_j^i}(d_j^i \leq d_{min}) f_{d_{min}}(d_{min}) \\ &\quad dd_j^i\, dd_{min}, \quad (16) \\ &\leq \sum_{i=1}^{L} \sum_{j=1,j\neq m}^{M} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{d_{min}} f_{d_j^i}(d_j^i) dd_j^i \right) f_{d_{min}}(d_{min}) \\ &\quad dd_{min}, \quad (17) \end{aligned}$$

where $d_j^i \sim \mathcal{N}(\frac{\Delta^2}{12}, \frac{\Delta^4}{N180})$. Since the inner integral in Eq. (17) is the Gaussian Q function and does not depend on the index j, Eq. (17) can be simplified to

$$\mathrm{P}_e^{mul} \leq L(M-1) \int_{-\infty}^{\infty} Q\left(\frac{d_{min} - m_{d_{ind}}}{\sigma_{d_{ind}}}\right) f_{d_{min}}(d_{min}) dd_{min}. \quad (18)$$

V. Results

Figs. 4, 5 and 6 display the *union bound* on the probability of error vs. robustness computed by numerically solving Eq. (18), through deriving $m_{d_{dep}}$ and $\sigma_{d_{dep}}^2$, for various codebook numbers and sizes of $M \times N$. The corresponding robustness measure $R = \frac{P_E}{\sigma_Z^2}$ 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 [11]. Givens rotations provide orthogonal transformations in $\Re^{N \times N}$ that rotate each vector with a fixed angle. In other words, signal space is spanned by the sets of basis vectors which are rotated with respect to each other by the chosen angle.

Hadamard transform matrix of size $N \times N$ and its negated version are combined into $2N \times N$ binary valued matrix to generate the 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 $E[\mathbf{W}_i^T \mathbf{W}_j] = 0$, $i \neq j$ and $i \neq j + N$.

We fixed the embedding distortion P_E and optimized the embedding parameter Δ for each $R = \frac{P_E}{\sigma_Z^2}$ value. The Δ value used for embedding is also revealed to the detector but not the β. We assumed the host signal $\mathbf{S}$ and channel noise $\mathbf{Z}$ are i.i.d. zero mean Gaussian vectors with the variances σ_S^2 and σ_Z^2, respectively, satisfying $\sigma_S^2 \gg P_E$ and σ_Z^2.

The simulations are performed 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 $\mathbf{W}_m$ vector of length N to $\mathbf{S}$. The watermarked signal $\hat{\mathbf{S}}$ is passed through an AWGN channel with the noise variance selected in a way that $R = \frac{P_E}{\sigma_Z^2}$ is satisfied for some discrete values of R. The detector extracts the signal $\hat{\mathbf{W}}_m$ and computes its distance to each of the watermark signals in order to detect the hidden message $\hat{m}$. If the extracted message $\hat{m}$ at the detector is the same with m, it is called a success, and otherwise, an error. The resulting probability of success values are used to compute the hiding rate of the system within an *M-ary symmetric channel* assumption.

We performed embedding with up to 25 codebooks and various codebook sizes of 64×32, 128×64, 256×128. The results are evaluated within $0.1 \leq R \leq 0.8$ range of embedding power to noise power ratios. Fig. 7 display the hiding rates for $L = 4$ and varying N values. The increase in the watermark signal size N improves the detection statistics as the variance of the r.v. representing the distance between two signals decreases. Fig. 8 display the hiding rates for $N = 128$ and $L = 1, 3, 5, 9, 14, 25$. It is observed from these performance simulations that the multiple codebook embedding has superior hiding rates than the corresponding single codebook case (for the same N).

References

[1] H. T. Sencar, M. Ramkumar and A. N. Akansu, "Multiple Codebook Information Hiding," CISS-2002 Conference.

[2] C. E. Shannon, "Channels with side information at the transmitter," *IBM Journal of Research and Development*, Vol. 2, pp. 289-293.

[3] 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.

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

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

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

[7] J. Chou, S. S. Pradhan, L. E. Ghaoui and K. Ramchandran, "A Robust Optimization Solution to the Data Hiding Problem using Distributed Source Coding Principles," *Proc SPIE: Image and Video Communications and Processing*, Vol. 3974, 2000.

[8] B. Chen and G. W. Wornell, "Preprocessed and postprocessed quantization index modulation methods for digital watermarking," *Proc SPIE: Security and Watermarking of Multimedia Contents II*, Vol. 3971, pp. 48-54, 2000.

[9] 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.

[10] 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.

[11] D. S. Watkins, "Fundamentals of Matrix Computations," John Wiley & Sons, New York, 1991.

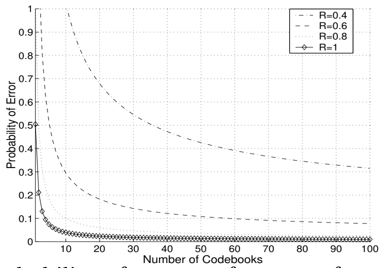

Figure 4: Probability of error performance for multiple codebook embedding and detection for M=64 and N=32.

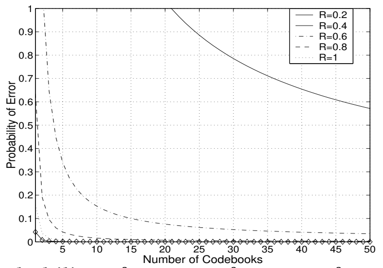

Figure 5: Probability of error performance for multiple codebook embedding and detection for M=128 and N=64.

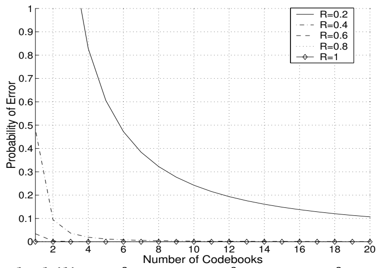

Figure 6: Probability of error performance for multiple codebook embedding and detection for M=256 and N=128.

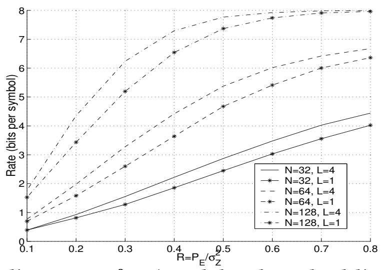

Figure 7: Hiding rates for 4-codebook embedding and detecting with varying watermark signal sizes $N = 32$, $N = 64$ and $N = 128$.

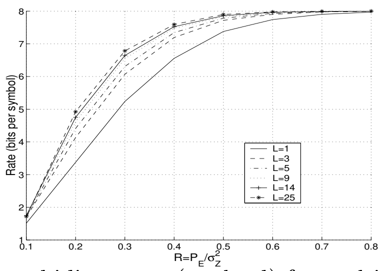

Figure 8: Data hiding rates (payload) for multiple codebook embedding and detection for various number of codebooks $L = 1, 3, 5, 9, 14, 25$ and $N = 128$.