Information-Theoretic Bounds for Steganography in Multimedia

Hassan Y. El-Arsh, Amr Abdelaziz, *Member, IEEE,* Ahmed Elliethy, *Member, IEEE,*
Hussein A. Aly, *Senior Member, IEEE,* and T. Aaron Gulliver, *Senior Member, IEEE*

Abstract—Steganography in multimedia aims to embed secret data into an innocent looking multimedia cover object. This embedding introduces some distortion to the cover object and produces a corresponding stego object. The embedding distortion is measured by a cost function that determines the detection probability of the existence of the embedded secret data. A cost function related to the maximum embedding rate is typically employed to evaluate a steganographic system. In addition, the distribution of multimedia sources follows the Gibbs distribution which is a complex statistical model that restricts analysis. Thus, previous multimedia steganographic approaches either assume a relaxed distribution or presume a proposition on the maximum embedding rate and then try to prove it is correct. Conversely, this paper introduces an analytic approach to determining the maximum embedding rate in multimedia cover objects through a constrained optimization problem concerning the relationship between the maximum embedding rate and the probability of detection by any steganographic detector. The KL-divergence between the distributions for the cover and stego objects is used as the cost function as it upper bounds the performance of the optimal steganographic detector. An equivalence between the Gibbs and correlated-multivariate-quantized-Gaussian distributions is established to solve this optimization problem. The solution provides an analytic form for the maximum embedding rate in terms of the WrightOmega function. Moreover, it is proven that the maximum embedding rate is in agreement with the commonly used Square Root Law (SRL) for steganography, but the solution presented here is more accurate. Finally, the theoretical results obtained are verified experimentally.

I. INTRODUCTION

Steganography is used to embed data within innocent looking cover objects such as images, audio, video, and text [1], to produce a stego object that looks similar to the original object but with hidden data embedded. This implies some distortion of the cover object which can be exploited by a warden to detect if there is hidden data or not using steganalysis. The goal of steganography is to hide the maximum amount of data subject to a minimum probability of detection by a warden. Thus, steganography has two competing goals: embedding rate and undetectability.

Undetectability is concerned with determining how to alter a cover object to embed data without detectable distortion to the object. The distortion is measured by a cost function

H. Y. El-Arsh, Amr Abdelaziz, A. Elliethy, and H. A. Aly are with Dept. of Computer Engineering, Military Technical College, Cairo, Egypt. (e-mail: hassan.yakout@ymail.com, amrashry@mtc.edu.eg, a.s.elliethy@mtc.edu.eg, haly@ieee.org). T. Aaron Gulliver is with the Dept. of Electrical and Computer Engineering, University of Victoria, Victoria, BC Canada. (e-mail:agullive@ece.uvic.ca)

which can be modeled as the difference between the cover and stego objects [2], or the sensitivity of the warden to changes to the cover object [3]. The embedding rate is defined as the ratio of the number of hidden data bits to the number of cover data bits for a given undetectability [4, Ch. 4]. Maximizing the embedding rate for a given undetectability is fundamental to any steganographic algorithm. Multimedia objects are excellent choices for steganography due to their rich structural features and their ubiquity in business and daily life [5], [6].

A relation between undetectability and embedding rate can be used to evaluate the performance of a multimedia steganographic approach, but this requires a suitable cost function. Multimedia has been shown to follow the Gibbs distribution [7]–[9], which is a complex statistical model that makes analysis difficult. For example, in [4, Ch.13] the Kullback–Leibler (KL) divergence (also called relative entropy), was used as a measure of cover distortion, however this requires a precise model of the cover object.

Previous steganography cost functions are either not accurate for all multimedia types or are accurate for only a certain class of embedding or detection techniques [10]. Further, the complexity of the Gibbs distribution has resulted in the use of relaxed distributions such as the Gaussian distribution [10]–[17], or a presumed preposition on the maximum embedding rate for which the correctness must be demonstrated [11], [12]. Thus, the corresponding results may not reflect the true performance because of these approximations.

To partially address the above problems, a constrained optimization problem was proposed in [18] that models the undetectability using KL divergence and the embedding rate with the mutual information between the cover object and stego object distributions. This problem was solved to obtain an analytic relation between the embedding rate and undetectability. However, a relaxed statistical model (Multivariate Quantized Gaussian Distribution (MQGD)), was used for the cover and stego objects. Thus, an achievability proof that verifies the reliable extraction of the embedded message at the receiver when the exact cover realization is not present and the probability of decoding error is nonzero was not obtained.

In this paper, a constrained optimization problem that provides an upper bound on the maximum embedding rate that a multimedia steganographic method can achieve reliably (i.e., with low probability of error), with an achievability proof for a given undetectability. The cover object is modeled using an equivalence between the Gibbs distribution and the Correlated Multivariate Quantized Gaussian Distribution (CMQGD). The corresponding optimization problem is solved to obtain an analytic form for the maximum embedding rate in terms of

the WrightOmega function. It is also shown that the maximum embedding rate agrees with the commonly used Square Root Law (SRL) for steganography [4, Ch.13]. Finally, this theoretical maximum embedding rate is compared with other rates in the literature [10], [16], [17]. The results obtained demonstrate that the upper bound presented here is small compared to the embedding rates for the steganographic methods. One reason is that this bound is for an optimal detector, which is not the case in the literature.

The rest of this paper is organized as follows. Section II presents the related work. The main definitions and assumptions for the proposed constrained optimization problem are given in Section III. Section IV provides an equivalence relation between the Gibbs distribution and CMQGD. The solution of the proposed constrained optimization problem is given in Section V and the achievability proof is presented in Section VI. Section VII considers the relationship between the SRL and results given here. An interpretation is given in Section VIII and discussed in Section IX. Finally, some concluding remarks are given in Section X.

II. RELATED WORK

Previous steganography approaches employed the KL divergence as a cost function and derived a theoretical bound (under different assumptions), on the maximum number of bits that can be embedded in a class of multimedia cover objects for a specified probability of detection by a warden. In [11], continuous Gaussian distributed cover objects were considered (AWGN channels), and the achievability and converse were derived. The technique in [11] was generalized in [12] for Multivariate Continuous Gaussian Distributed (MCGD) cover objects (MIMO Gaussian channels). Although these approaches employ an accurate model for the cover objects with the KL divergence cost function, a preposition on the maximum embedding rate is assumed.

A systematic technique for constructing the distortion function based on the relationship between the steganographic Fisher information and KL divergence was given in [15]. The embedded data was modeled as an additive image distortion, and individual pixel costs were determined adaptively to minimize the KL divergence for embedding using least-significant bit matching. A zero mean independent MQGD was used as the cover object model considering the noise from imaging sensors. It was shown that the security is better than with the Highly Undetectable Steganograpy (HUGO) algorithm [19]. These results were improved in [13] using a Multivariate Generalized Gaussian (MVGG) model combined with a new variance estimator. This improved the steganographic performance by allowing larger amplitude embeddings in complex (highly textured) image regions due to the larger tail of the MVGG model. Content-adaptive pentary steganography utilizing a MVGG model provides comparable performance to pentary coded HiLL [20] and S-UNIWARD [21] against the maxSRMd2 [22] and SRM [23] feature-based steganalysis methods. Although these techniques provide an accurate analytic relation between the embedding rate and cost function, the results are based on a relaxed cover model.

In [16], a Gaussian Markov Random Field (GMRF) model was employed which is a non-additive model with a four-element cross neighborhood. This model has low-dimensional clique structures that can capture the interdependencies among spatially contiguous pixels. The cost function minimizes the KL divergence between the cover object and stego object. The cover image is split into two disjoint sub-images which are conditionally independent. Then, an alternating iterative optimization technique is employed to achieve an efficient embedding which minimizes the KL divergence. Results were presented which show that the proposed approach outperforms the MiPOD [14] and HiLL techniques in terms of the secure payload against SRM and maxSRMd2. This approach was extended in [17] to higher dimension clique structures (eight element GMRF). Although analytic relationships between the embedding rate and cost function were given in [16], [17], they are based on a relaxed model for the cover object.

The approach in [10] considers gray-scale images as cover objects and assumes a MQGD for the cover object, hidden message, and stego object. Theoretical bounds for the Bayes, Minimax, and Neyman-Pearson likelihood ratio tests were obtained. This approach maximizes the steganalysis error related to a specified payload. Although an analytic relation between the embedding rate and cost function was given, the scope of the cost function is limited to only three types of detectors and a relaxed model (MQGD), is used.

The Square Root Law (SRL) [4, Ch.13] is regarded as the first analytic model of the relationship between the embedding rate and undetectability in a global framework for multimedia. The SRL states that the embedding rate for any steganographic technique is proportional to the square root of the number of cover elements. For example, assume cover object A contains x elements and cover object B contains $4x$ elements. If we can embed y bits of a secret message within A with a probability of detection $\mathcal{P}$ using a given technique then for the same probability of detection we can only embed $2y$ bits within B. Further, the embedding rate per cover element in A is $\frac{y}{x}$, but for B the corresponding embedding rate is $\frac{y}{2x}$. The SRL is valid for any steganographic approach regardless of the embedding and steganalysis methods. However, it is only a proportional result for a given cost function and so the scaling constant is unknown.

III. SYSTEM MODEL AND PROBLEM STATEMENT

Elements of a cover object are assumed to be the communication channel between two entities: sender $\mathbf{A}$ and receiver $\mathbf{B}$, and an eavesdropper $\mathbf{E}$ monitors this channel. Fig. 1 gives the general communication model for steganography. In this paper, we consider an innocent cover object $\mathbf{c} = [c_1, c_2, \ldots, c_n]$ and a stego object $\mathbf{s} = [s_1, s_2, \ldots s_n]$ with probability distributions P_c and P_s, respectively, where n is the number of elements in the cover and stego objects and $c_i, s_j \in \mathbb{V}$ where $\mathbb{V}$ is the set of allowed cover values. For example, if the cover object is a gray-scale image, then $\mathbb{V} = \{0, 1, \ldots, 255\}$ is the set of possible pixel values. Note that $\mathbb{V}$ is not limited to pixels, but can also be, for example, discrete cosine transform (DCT) coefficients [24] or video motion vectors [2], [25]. Similarly,

TABLE I: List of Symbols

|	
$\mathbf{A}$	Sender
$\mathbf{B}$	Receiver
$\mathbf{E}$	Eavesdropper
$\mathbb{V}$	Alphabet of cover and stego elements
$\mathbb{X}$	Alphabet of the original message
$\mathbb{M}$	Alphabet of the coded message
$\mathbf{c}$	Cover object
$\mathbf{s}$	Stego object
$\mathbf{x}$	Original message
$\mathbf{m}$	Coded message
P_D	Probability of steganalysis detection at $\mathbf{E}$
P_e	Probability of steganalysis error at $\mathbf{E}$
P_E	Average probability of steganalysis error at $\mathbf{E}$
P_B	Probability of decoding error at $\mathbf{B}$
P_c	Joint probability distribution of the cover object
p_c	Continuous version of P_c
P_s	Joint probability distribution of the stego object
p_s	Continuous version of P_s
P_m	Joint probability distribution of the coded message
p_m	Continuous version of P_m
c_i	ith cover element
s_i	ith stego element
P_{c_i}	Probability distribution of the ith cover object
p_{c_i}	Continuous version of P_{c_i}
P_{s_i}	Probability distribution of the ith stego object
p_{s_i}	Continuous version of P_{s_i}
$\Sigma_c, \Sigma_s, \Sigma_m$	Covariance matrices for P_c, P_s and P_m, respectively, and also valid for p_c, p_s and p_m, respectively
$\mathrm{tr}(X)$	Trace of matrix X
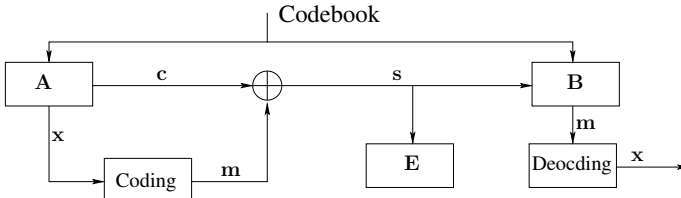
Fig. 1: The steganography communication model.

we consider a message $\mathbf{x} \in \mathbb{X}^k$, where k is the number of message elements and $\mathbb{X}$ is the set of possible message values. The stego object $\mathbf{s}$ is obtained by embedding a coded message $\mathbf{m} \in \mathbb{M}^n$ via an addition process, i.e. $\mathbf{s} = \mathbf{c} + \mathbf{m}$. The coded message $\mathbf{m}$ is obtained from $\mathbf{x}$ using a codebook that maps $\mathbf{x} \mapsto \mathbf{m}$ and has the same probability distribution as P_m.

A. Problem Statement

Consider the scenario in which $\mathbf{A}$ has access to a multimedia cover object $\mathbf{c}$ drawn from a distribution P_c which is known by both $\mathbf{B}$ and $\mathbf{E}$. To conceal the secret information within $\mathbf{c}$, $\mathbf{A}$ transmits a perturbed version of $\mathbf{c}$ called the stego object $\mathbf{s}$ and the perturbations are known to only $\mathbf{A}$. The goal is for $\mathbf{A}$ to minimize the probability of detection, P_D, by $\mathbf{E}$.

The fundamental problem is then to quantify for a given P_c how much information can be exchanged reliably between $\mathbf{A}$ and $\mathbf{B}$ while maintaining a given P_D by $\mathbf{E}$. Since $\mathbf{A}$ does not know the detection strategy at $\mathbf{E}$, the transmission strategy should be designed considering an optimal detector at $\mathbf{E}$. In addition, no prior knowledge of the exact realization of the cover object is assumed at $\mathbf{B}$ or $\mathbf{E}$. In other words, the cover object realization is selected at random by $\mathbf{A}$.

We assume that $\mathbf{E}$ employs a binary hypothesis test for $\mathcal{H}_0$ and $\mathcal{H}_1$ where $\mathcal{H}_0$ means $\mathbf{A}$ has not embedded a secret message and $\mathcal{H}_1$ means $\mathbf{A}$ has embedded a secret message. Thus, there are two types of errors.

- Type **I** error: Deciding $\mathcal{H}_1$ when $\mathcal{H}_0$ is true, called false positive (false alarm). We denote the probability of this type of error by α.
- Type **II** error: Deciding $\mathcal{H}_0$ when $\mathcal{H}_1$ is true, called false negative (miss detection). We denote the probability of this type of error by β.

Assuming equal priors for the optimal hypothesis test at $\mathbf{E}$, the total probability of error is

$$P_e = 1 - P_D = \alpha + \beta = 2P_E, \tag{1}$$

where P_E is the average probability of error for $\mathbf{E}$ for equal priors. According to [26]

$$P_e = 1 - \mathcal{V}(P_s, P_c), \tag{2}$$

where $\mathcal{V}(P_s, P_c)$ is the total variation distance between P_s and P_c defined as [26]

$$\mathcal{V}(P_s, P_c) = \frac{1}{2} \|P_s - P_c\|_1, \tag{3}$$

and $\|.\|_1$ is the L_1 norm. As an analytic form of the L_1 norm is typically complex, the KL divergence can be employed [27, Ch. 11] to obtain

$$\mathcal{V}(P_s, P_c) \leq \sqrt{\frac{1}{2}\mathcal{D}(P_s \parallel P_c)}, \tag{4}$$

where

$$\mathcal{D}(P_s \parallel P_c) = \sum_{j=1}^{\dot{n}} P_{s_j} \ln \frac{P_{s_j}}{P_{c_j}}, \tag{5}$$

and P_{s_j} and P_{c_j} are the probability mass function bins for the cover and stego objects, respectively, and $\dot{n}$ is the number of bins. To guarantee a low detection probability at $\mathbf{E}$, $\mathbf{A}$ must bound $\mathcal{V}(P_s, P_c)$ by some ϵ based on the desired probability of detection. From (4), $\mathbf{A}$ should design a steganographic technique such that

$$2 \times \mathcal{V}(P_s, P_c)^2 \leq \mathcal{D}(P_s \parallel P_c) \leq 2\epsilon^2. \tag{6}$$

On the other hand, $\mathbf{B}$ must decode the message $\mathbf{x}$ from $\mathbf{m}$ sent by $\mathbf{A}$. Thus, the codebook must be selected to maximize the mutual information between P_s and P_m

$$I(P_m; P_s) = H(P_m) - H(P_m \mid P_s), \tag{7}$$

where $H(P_m)$ is the entropy of P_m and $H(P_m \mid P_s)$ is the conditional entropy of P_m given P_s. Thus, the proposed constrained optimization problem can be written as

$$\underset{P_m}{\operatorname{argmax}} \quad I(P_m; P_s) \quad \text{s.t.} \quad \mathcal{D}(P_s \parallel P_c) \leq 2\epsilon^2. \tag{8}$$

Note that $I(P_m; P_s)$ in (7) can be written as [27, Ch. 2]

$$I(P_m; P_s) = H(P_s) - H(P_s \mid P_m) = H(P_s) - H(P_c), \tag{9}$$

and because $H(P_c)$ is known, maximizing $I(P_m; P_s)$ is equivalent to maximizing $H(P_s)$. Thus, the proposed optimization problem in (8) can be rewritten as

$$\underset{P_m}{\operatorname{argmax}} \quad H(P_s) \quad \text{s.t.} \quad \mathcal{D}(P_s \parallel P_c) \leq 2\epsilon^2. \tag{10}$$

IV. Stochastic Multimedia Model

In this section, the multimedia distribution model, P_c, is given. We first define Markov Random Field (MRF) and Gibbs distribution.

MRF is a multidimensional random process which generalizes the single dimensional Markov random process [9]. Let $\ddot{X}$ be a coordinate system in R^N and $\rho(i)$ a function representing the neighbourhood of element $i \in \ddot{X}$ such that $i \notin \rho(i)$ and $i \in \rho(j)$ iff $j \in \rho(i)$. For example, the neighbourhood can be defined as the immediate left, right, top and bottom neighbours of i. Let $\ddot{Y}$ be the neighbourhood system representing the set of neighbourhoods of all elements in $\ddot{X}$.

A random field ξ over $\ddot{X}$ is a multidimensional random process where each element $i \in \ddot{X}$ is assigned a random variable ξ_i and $\mathfrak{f}_i$ is the associated realization for all $i \in \ddot{X}$. ξ is called a MRF if it satisfies the following conditions

- Positivity: $P(\xi = \mathfrak{f}) > 0$, $\forall \mathfrak{f} \in \ddot{S}$, and
- Markovianity: $P(\xi_i = \mathfrak{f}_i | \xi_{j_.} = \mathfrak{f}_j, \forall j \neq i) = P(\xi_i = \mathfrak{f}_i | \xi_j = \mathfrak{f}_j, \forall j \in \rho(i)), \forall i \in \ddot{X}, \forall \mathfrak{f} \in \ddot{S}$,

where P is the probability measure and $\ddot{S}$ is the state space for ξ.

The Hammersley-Clifford Theorem [9] states that ξ is a MRF over $\ddot{X}$ with respect to $\ddot{Y}$ if and only if its probability distribution follows the Gibbs distribution with respect to $\ddot{X}$ and $\ddot{Y}$. To explain the Gibbs distribution, a clique must be defined. A clique ω is a correlated group of neighbouring elements where $\omega \subset \ddot{X}$ with respect to $\ddot{Y}$, i.e. ω consists of a single element or multiple elements which are neighbours. Further, $\omega \in \ddot{\Omega}$ where $\ddot{\Omega}$ denotes the set of all cliques.

According to [7]–[9], multimedia objects can be modeled using the Gibbs distribution. This implies the above described properties (positivity and Markovianity), so the elements of a multimedia object can be organized in cliques. The Gibbs distribution employs an energy function $U_\mathfrak{f}$ and a temperature constant $\hat{T}$ to obtain the probability

$$P(\xi = \mathfrak{f}) = \frac{1}{Z} e^{-U_\mathfrak{f}/\hat{T}}, \tag{11}$$

where

$$Z = \sum_{k=1}^{M} e^{-U_k/\hat{T}}, \tag{12}$$

is the normalization denominator called the partition function and M is the number of system states (number of elements in $\ddot{S}$). The energy function is modeled as

$$U_\mathfrak{f} = \sum_{\omega \in \ddot{\Omega}} V_\mathfrak{f}(\omega), \tag{13}$$

where $V_\mathfrak{f}(\omega)$ is the potential function for $\mathfrak{f}$ calculated only within the clique $\omega \in \ddot{\Omega}$.

The Gibbs distribution requires the partitioning function which is intractable to obtain for most models. Therefore, a CMQGD is used here to model the Gibbs distribution. This is implemented by modeling (11) with a CMQGD by modeling $V_\mathfrak{f}(\omega)$ in (13) as

$$V_\mathfrak{f}(\omega) = \frac{\hat{T}}{2} (\mathfrak{f}_\omega - \mu_\omega) \Sigma_\omega^{-1} (\mathfrak{f}_\omega - \mu_\omega)^T, \tag{14}$$

where $\mathfrak{f}_\omega$ is the realization of the random field ξ within clique ω, i.e. $\mathfrak{f}_i \in \mathfrak{f}_\omega \quad \forall i \in \omega$, $\mathfrak{f}_\omega \subset \mathfrak{f}$, and μ_ω and Σ_ω are the mean vector of $\xi_\omega \subset \xi$ and the corresponding covariance matrix within the clique ω, respectively. This equivalence is valid for a wide class of Gibbs distributions in which the potential function can be expressed in a quadratic form of the exponent of an equivalent Gaussian distribution. This is based on the following.

- $V_\mathfrak{f}(\omega)$ can be mapped to any function using only the elements within the clique ω [9].
- The product of $(\mathfrak{f}_\omega - \mu_\omega)$ and $(\mathfrak{f}_\omega - \mu_\omega)^T$ (the distance from the mean for clique ω), controls the probability of a given realization (state) $\mathfrak{f}_\omega$, and hence whether $P(\xi = \mathfrak{f})) \in \omega$ for a given Σ_ω. In other words, the local coherence condition of multimedia elements, $P(\xi = \mathfrak{f}) \propto -U_\mathfrak{f} \propto -V_\mathfrak{f}(\omega)$, is satisfied. This results in low energy states having a higher probability than high energy states, which is one of the properties of the Gibbs distribution [7].
- Z, which is the normalization denominator, is equal to

$$\sqrt{2\pi} |\Sigma_\omega| = \sum_{\omega \in \ddot{\Omega}} e^{-\sum \frac{\hat{T}}{2} (\mathfrak{f}_\omega - \mu_\omega) \Sigma_\omega^{-1} (\mathfrak{f}_\omega - \mu_\omega)^T}.$$

It should be noted that other assumptions for the potential function $V_\omega(\mathfrak{f})$ can be used to model the Gibbs distribution.

V. Main Results

In this section, the solution of the optimization problem (10) in Section III is presented using the assumptions given below.

- Each clique within the cover is considered as a single cover element c_i with unknown distribution P_{c_i}.
- The equivalence relation between Gibbs and Gaussian distributions in Section IV is used to model the cover object, i.e. the joint distribution of P_c, as a CMQGD with mean $\vec{\mu_c}$ and covariance matrix Σ_c. No assumptions are made on the marginal distribution P_{c_i} of c_i.
- The cover statistical parameters $(\vec{\mu_c}, \Sigma_c)$ are known to **A**, **B**, and **E**.
- Although no assumptions are made on P_s and P_m, minimizing the KL divergence for a Gaussian distributed cover object requires that the corresponding stego object also be Gaussian distributed [28]. Thus, the embedding operation will produce a CMQGD P_s with parameters $(\vec{\mu_s}, \Sigma_s)$. Although the zero mean Gaussian distribution case is considered in [28], the proof can be extended to any Gaussian distribution as shown in Appendix A.
- The codebook realization from P_m, i.e. the codewords used between **A** and **B**, is shared only between **A** and **B** and is not known to **E**.

A. Solution of (10)

To solve (10), we must determine $H(P_c)$, $H(P_s)$ and the KL divergence $\mathcal{D}(P_s \parallel P_c)$. First, the KL divergence between two quantized distributions is evaluated using (16) using the following lemma.

Lemma 1. *[27, Ch. 11] The KL divergence between two uniformly quantized distributions $F(\dot{X})$ and $G(\dot{X})$ is bounded as*

$$\mathcal{D}(F(\dot{X}) \parallel G(\dot{X})) \leq \mathcal{D}(f(\dot{x}) \parallel g(\dot{x})) \quad (15)$$

where f and g are continuous versions of F and G, respectively, and $\dot{X}$ is the uniformly quantized version of the continuous random variable $\dot{x}$.

From [12], the KL divergence for the MCGD is defined as:

$$\mathcal{D}(p_s \parallel p_c) = \frac{1}{2}\Big(\text{tr}(\Sigma_c^{-1}\Sigma_s) + (\vec{\mu_c} - \vec{\mu_s})^T \Sigma_c^{-1}(\vec{\mu_c} - \vec{\mu_s}) \\ - \ln\frac{|\Sigma_s|}{|\Sigma_c|} - n \Big), \quad (16)$$

From Lemma 1 and (16), the KL divergence for a CMQGD can be bounded as

$$\mathcal{D}(P_s \parallel P_c) \leq \mathcal{D}(p_s \parallel p_c),$$

$$\mathcal{D}(P_s \parallel P_c) \leq \frac{1}{2}\Big(\text{tr}(\Sigma_c^{-1}\Sigma_s) + (\vec{\mu_c} - \vec{\mu_s})^T \Sigma_c^{-1}(\vec{\mu_c} - \vec{\mu_s}) \\ - \ln\frac{|\Sigma_s|}{|\Sigma_c|} - n \Big), \quad (17)$$

where n is the number of cover elements. Thus, we can use the upper bound on $\mathcal{D}(P_s \parallel P_c)$ (i.e. substitute $\mathcal{D}(P_s \parallel P_c)$ with $\mathcal{D}(p_s \parallel p_c)$)) to evaluate the optimization problem in (10). From (17), let $\ddot{A} = \Sigma_c^{-1}\Sigma_s$, then we have

$$|\ddot{A}| = |\Sigma_c^{-1}||\Sigma_s| = \frac{|\Sigma_s|}{|\Sigma_c|}.$$

As Σ_c^{-1} and Σ_s are symmetric and positive semi-definite

$$|\ddot{A}| = \prod \lambda_i, \quad \text{tr}(\ddot{A}) = \sum \lambda_i,$$

and $\text{tr}(\ddot{A})$ is monotonically increasing with $|\Sigma_s|$ where $\sum \lambda_i$ and $\prod \lambda_i$ are the sum and product of the eigenvalues of $\ddot{A}$, respectively. Then from the arithmetic-geometric inequality we have

$$\frac{1}{\ddot{n}}\text{tr}(\ddot{A}) \geq |\ddot{A}|^{\frac{1}{\ddot{n}}},$$

where $\ddot{n}$ is the number of eigenvalues. As $x > \ln(x)$ for any positive real number x

$$\frac{1}{\ddot{n}}\text{tr}(\ddot{A}) > \frac{1}{\ddot{n}}\ln(|\ddot{A}|),$$
$$\text{tr}(\ddot{A}) > \ln(|\ddot{A}|),$$
$$\text{tr}(\Sigma_c^{-1}\Sigma_s) > \ln\left(\frac{|\Sigma_s|}{|\Sigma_c|}\right), \quad (18)$$

Then from (15), (17) and (18), $\mathcal{D}(p_s \parallel p_c)$ is monotonically increasing with $|\Sigma_s|$. Moreover, from (52) and (54), $H(P_s)$ and $h(p_s)$ are also monotonically increasing with $|\Sigma_s|$, where $h(p_s)$ is the differential entropy of the continuous CMQGD p_s and Σ_s is the only variable in (54) as $2\pi e$ and b can be considered as constants.

Then, substituting $\mathcal{D}(P_s \parallel P_c)$ with $\mathcal{D}(p_s \parallel p_c)$ in (10), the Lagrangian of (10) is

$$\mathcal{L}(p_s, \lambda) = \mathcal{L}(\vec{\mu_s}, \Sigma_s, \lambda) = h(p_s) - \lambda(\mathcal{D}(p_s \parallel p_c) - 2\epsilon^2), \quad (19)$$

where λ is the Lagrange multiplier.

The solution of the optimization problem in (19) is given in the following theorem.

Theorem 1. *With the given assumptions, the solution of (19) is obtained as follows.*

1) *The distribution of P_m must be CMQGD for both the cover and stego objects, with mean and covariance matrix*
   a) $\vec{\mu_m} = 0$, *and*
   b) $\Sigma_m = \Big(-W(-\frac{4\epsilon^2}{n} - 1 - i\pi) - 1 \Big)\Sigma_c$,

   for steganalysis with $P_D \leq \sqrt{\frac{\mathcal{D}(P_s \parallel P_c)}{2}} \leq \epsilon$) where $W(.)$ is the WrightOmega function [29].

2) *The maximum achievable embedding rate is*

$$I(P_s; P_m) = \frac{n}{2}\ln(-W(-\frac{4\epsilon^2}{n} - 1 - i\pi)),$$

   for sufficiently large n.

Proof. To find the optimal parameters for (19), set $\frac{\partial}{\partial\vec{\mu_s}}\mathcal{L} = 0$ and $\frac{\partial}{\partial\Sigma_s}\mathcal{L} = 0$. With the definitions of $h(p_s)$ and $\mathcal{D}(P_s \parallel P_c)$ in (52) in Appendix B and (16), respectively, we have

$$\frac{\partial}{\partial\vec{\mu_s}}h(p_s) = 0, \quad (20)$$

then from (52) and [30, (57)]:

$$\frac{\partial}{\partial\Sigma_s}h(p_s) = \frac{1}{2}\frac{\partial}{\partial\Sigma_s}\Big(\ln(2\pi\epsilon) + \ln\Sigma_s \Big),$$
$$= \frac{1}{2}(\Sigma_s^{-1})^T, \quad (21)$$

then from [30, (37,81)]:

$$\frac{\partial}{\partial\vec{\mu_s}}\mathcal{D}(p_s \parallel p_c) = \frac{-1}{2}\Big(\Sigma_c^{-1}(\vec{\mu_c} - \vec{\mu_s}) + (\Sigma_c^{-1})^T(\vec{\mu_c} - \vec{\mu_s}) \Big), \quad (22)$$

then from [30, (57,100,101)]:

$$\frac{\partial}{\partial\Sigma_s}\mathcal{D}(p_s \parallel p_c) = \frac{1}{2}\Big(\Sigma_c^{-1} - \Sigma_s^{-1} \Big). \quad (23)$$

As Σ_c and Σ_s are symmetric matrices, (22) can be rewritten as

$$\frac{\partial}{\partial\vec{\mu_s}}\mathcal{D}(p_s \parallel p_c) = -\Sigma_c^{-1}(\vec{\mu_c} - \vec{\mu_s}). \quad (24)$$

From (20) and (24), as $\frac{\partial}{\partial\vec{\mu_s}}\mathcal{L} = 0$, we have

$$\vec{\mu_s} = \vec{\mu_c}. \quad (25)$$

Further, from (21) and (23), as $\frac{\partial}{\partial\Sigma_s}\mathcal{L} = 0$, we have

$$(\lambda + 1)\Sigma_s^{-1} = \lambda\Sigma_c^{-1},$$
$$\Sigma_s = \frac{\lambda + 1}{\lambda}\Sigma_c,$$
$$\Sigma_s = a\Sigma_c. \quad (26)$$

Let $a = \frac{\lambda+1}{\lambda}$ be the embedding factor. From (26), it is clear that Σ_s has the same eigenvectors as Σ_c and the eigenvalues of Σ_s are a scaled version of the eigenvalues of Σ_c by the embedding factor a.

Let the characteristic functions of the cover and stego objects be

$$\psi_c(\vec{t}) = e^{i\vec{\mu_c}^T \vec{t} + \frac{1}{2}\vec{t}^T \Sigma_c \vec{t}},$$
$$\psi_s(\vec{t}) = e^{i\vec{\mu_s}^T \vec{t} + \frac{1}{2}\vec{t}^T \Sigma_s \vec{t}}.$$

As $\mathbf{s} = \mathbf{c} + \mathbf{m}$, this implies $P_s = P_c \oplus P_m$, where $\oplus$ denotes convolution. Thus, the characteristic function for the message codebook distribution can be expressed as

$$\psi_m(\vec{t}) = \frac{\psi_s(\vec{t})}{\psi_c(\vec{t})} = e^{i(\vec{\mu_s}^T - \vec{\mu_c}^T)\vec{t} + \frac{1}{2}\vec{t}^T (\Sigma_s - \Sigma_c)\vec{t}}.$$

Then from (25), we have

$$\psi_m(\vec{t}) = e^{\frac{1}{2}\vec{t}^T (\Sigma_s - \Sigma_c)\vec{t}} = e^{\frac{1}{2}\vec{t}^T (\Sigma_m)\vec{t}}. \tag{27}$$

and

$$\mu_m = 0, \tag{28}$$

so the message codebook can be modeled as a CMQGD with zero mean and covariance matrix

$$\Sigma_m = \Sigma_s - \Sigma_c$$
$$= (a - 1)\Sigma_c. \tag{29}$$

The embedding factor a is obtained as follows. From (26), we have

$$\text{tr}(\Sigma_c^{-1}\Sigma_s) = na, \tag{30}$$

and

$$\ln \frac{|\Sigma_c|}{|\Sigma_s|} = -n\ln(a). \tag{31}$$

Substituting (25), (30) and (31) in (16) gives

$$an - n\ln(a) = 2\mathcal{D}(p_s \parallel p_c) + n,$$
$$a - \ln(a) = \frac{2}{n}\mathcal{D}(p_s \parallel p_c) + 1. \tag{32}$$

The solution of (32) can be obtained from [29] as

$$a = -W(-\frac{2}{n}\mathcal{D}(p_s \parallel p_c) - 1 - i\pi). \tag{33}$$

Then, substituting the upper bound $2\epsilon^2$ for $\mathcal{D}(p_s \parallel p_c)$ from (19) in (33), we have

$$a = -W(-\frac{4\epsilon^2}{n} - 1 - i\pi), \tag{34}$$

Equations (28), (29) and (34) complete the proof of part 1).

For part 2), from (9), and (53) and (54) in Appendix B, we have

$$I(p_s; p_m) = h(p_s) - h(p_c).$$

Then from (52) in Appendix B

$$I(p_s; p_m) = \frac{1}{2}\ln(2\pi e|\Sigma_s|) - \frac{1}{2}\ln(2\pi e|\Sigma_c|).$$

and from (29) and (31)

$$I(p_s; p_m) = \frac{n}{2}\ln(a) + \frac{1}{2}\ln(2\pi e|\Sigma_c|) - \frac{1}{2}\ln(2\pi e|\Sigma_c|)$$
$$= \frac{n}{2}\ln(a). \tag{35}$$

Thus

$$I(p_s; p_m) = \frac{n}{2}\ln\left(-W(-\frac{4\epsilon^2}{n} - 1 - i\pi)\right). \tag{36}$$

which completes the proof of part 2). ∎

Note that although $\mathcal{D}(p_s \parallel p_c) \neq \mathcal{D}(p_c \parallel p_s)$, Theorem 1 is also valid for $\mathcal{D}(p_c \parallel p_s)$. The proof of this is given in the following lemma.

Lemma 2. $\vec{\mu_m}$ and Σ_m in Theorem 1 are also valid for $\mathcal{D}(p_c \parallel p_s)$.

Proof. Equation (16) can be modified as

$$\mathcal{D}(p_c \parallel p_s) = \frac{1}{2}\bigg(\text{tr}(\Sigma_s^{-1}\Sigma_c) + (\vec{\mu_s} - \vec{\mu_c})^T \Sigma_s^{-1}(\vec{\mu_s} - \vec{\mu_c})$$
$$+ \ln\frac{|\Sigma_s|}{|\Sigma_c|} - n\bigg). \tag{37}$$

As a consequence, (23) and (24) become

$$\frac{\partial}{\partial\Sigma_s}\mathcal{D}(p_c \parallel p_s) = \frac{1}{2}\big(-(\Sigma_s^{-1})^T\Sigma_c(\Sigma_s^{-1})^T + (\Sigma_s^{-1})^T\big), \tag{38}$$

$$\frac{\partial}{\partial\vec{\mu_s}}\mathcal{D}(p_c \parallel p_s) = \Sigma_s^{-1}(\vec{\mu_s} - \vec{\mu_c}). \tag{39}$$

Then following the same steps as in the proof of Theorem 1 using (20), (21), (38), and (39), the results in (28), (29), (33), and (34) in Theorem 1 follow. ∎

VI. ACHIEVABILITY

In this section, we prove that the maximum embedding rate $I(P_s; P_m)$ in (36) is achievable with a low probability of decoding error P_B at **B**. An achievability constraint is not included in the optimization problem (10) and (19) as P_s and P_m are obtained after this problem is solved.

The standard random coding argument [27, Ch. 7] is used as the basis for the proof. Although the minimum distance decoder (maximum likelihood estimator), is the optimal detector [27, Ch. 7], it is difficult to use this detector to calculate P_B as we do not have the marginal distribution for each cover element, only the joint distribution for the entire cover. Thus, a jointly typical decoder is considered at **B**.

The approach employed in the achievability proof is the same as in the proof of the Channel Capacity Theorem [27, Theorem 9.1.1], to obtain P_B. Although this theorem constrains the average transmission power, in our case we do not need this constraint as Σ_m is very small due to the low probability of detection required at **E** in (10). Two types of errors can be defined if the ith codeword is transmitted.

- The transmitted and received codewords are not jointly typical. We denote this error by $\hat{E}_i$.
- The received codeword is jointly typical with an incorrect codeword. This error is denoted by $\sum_{j=1,j\neq i}^{\mathcal{K}} E_j$ where $\mathcal{K}$ is the number of codewords.

Then (9.26) in [27] is modified as

$$P_B \leq P(\hat{E}_i) + \sum_{j=1,j\neq i}^{\mathcal{K}} P(E_j), \tag{40}$$

where $P(x)$ is the probability of event x. Continuing with the approach in [27] gives

$$P_B \leq 2\delta, \tag{41}$$

where $\delta > 0$ is an arbitrary small number. Thus, (41) proves that when n is sufficiently large and for an embedding rate $R \leq I(P_s; P_m) - 2\epsilon$, P_B is sufficiently small. Therefore, a codebook $(\mathcal{K}, n, \epsilon)$ exists that achieves an embedding rate $R \leq I(P_s; P_m) - 2\epsilon$ at **B** with low P_B.

VII. RELATIONSHIP TO THE SRL

In this section, we establish the relationship between (34) and (36) in terms of the WrightOmega function and the results for the SRL in [11], [12] and [4, Ch.13]. According to [31, Theorem 1], a special case $\mathcal{W}_{-1}(-e^{-u-1})$ of the Lambert W function $\mathcal{W}(.)$ can be bounded as

$$-1 - \sqrt{2u} - u < \mathcal{W}_{-1}(-e^{-u-1}) < -1 - \sqrt{2u} - \frac{2}{3}u. \tag{42}$$

From [29] and (56) in Appendix C, setting $u = \frac{4\epsilon^2}{n}$ in (34) gives

$$W(-u - 1 - i\pi) = \mathcal{W}_{-1}(e^{-u-1-i\pi}),$$
$$= \mathcal{W}_{-1}(-e^{-u-1}),$$

and then using (42)

$$W(-u - 1 - i\pi) > -1 - \sqrt{2u} - u,$$
$$-W(-u - 1 - i\pi) < 1 + \sqrt{2u} + u,$$
$$a < 1 + \sqrt{\frac{8\epsilon^2}{n}} + \frac{4\epsilon^2}{n}.$$
$$\ln(a) < \ln(1 + \sqrt{\frac{8\epsilon^2}{n}} + \frac{4\epsilon^2}{n})$$

As $\ln(1 + x) \leq x$, then

$$\ln(a) < \sqrt{\frac{8\epsilon^2}{n}} + \frac{4\epsilon^2}{n}$$
$$n \ln(\sqrt{a}) < \frac{n}{2}(\sqrt{\frac{8\epsilon^2}{n}} + \frac{4\epsilon^2}{n}),$$
$$n \ln(\sqrt{a}) < \sqrt{2\epsilon^2}\sqrt{n} + 2\epsilon^2. \tag{43}$$

From (35) and (43), we have

$$I(p_s; p_m) < \sqrt{2}\,\epsilon\,\sqrt{n} + 2\epsilon^2, \tag{44}$$

which establishes the relationship between (36) and the SRL. This shows that the results given here are in agreement with those established for the SRL in [4, Ch.13] in the context of image steganography and in [11], [12] in the context of covert communications. Based on the justification in Section IV for modeling the Gibbs distribution with a CMQGD and (44), (36) can be considered as an analytic form of the SRL. ∎
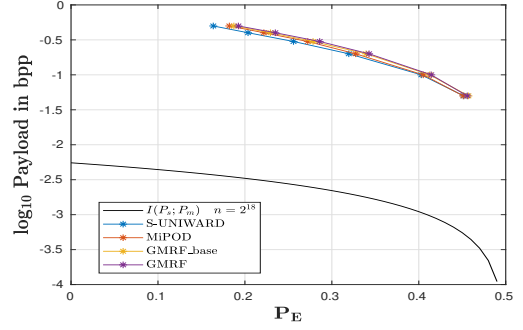

Fig. 2: Payload versus P_E from [16] for the steganographic methods S-UNIWARD, MIPOD, GMRF_BASE and GMRF with steganalysis using SRM.

VIII. BOUND INTERPRETATION

In this section, the experimental results for the steganographic methods in [10], [16], and [17] using the BOSSbase 1.01 dataset [32] are compared with the results given here. Specifically, the payload (bits per pixel) of each steganographic method is determined versus P_E. Then, these results are compared to the theoretical upper bounds for the payload from (36). As the size of each image of the BOSSbase 1.01 dataset is 512×512 pixels, we set $n = 512 \times 512 = 2^{18}$ in (36) to determine the maximum achievable embedding rate for independent cover elements, i.e. when each clique represents only a single pixel, which is an extreme case for a cover object to provide the maximum achievable embedding rate. The relationship between the achievable embedding rate and the clique size is investigated in Section IX.

Figures 2 to 5 present the results from [16] and the upper bound in (35) with a set to the lower bound in (59) to obtain an upper bound for the payload versus P_E. Figures S-1 to S-4 and Figures S-5 to S-11 for the methods in [17] and [10], respectively, present the average payload (in bits per pixel) for the steganographic methods S-UNIWARD, MIPOD, GMRF BASE, GMRF, MiPOD and HILL versus P_E for steganalysis using SRM and maxSRMd2. These results show that the upper bound is small compared to the results for the steganographic methods in the literature. The reason is that the steganographic detectors employed are not optimal for the steganographic methods. In other words, the steganalysis methods used in [10], [16], [17] are suboptimal compared to the KL divergence $\mathcal{D}(p_s \parallel p_c)$ given previously which can be regarded as an upper bound for any steganographic detector. Consequently, using a suboptimal detector may be misleading as it will result in a lower probability of detection error, P_E, for a given embedding rate compared to the upper bound. Further, using a suboptimal embedding technique may result in lower embedding rates than the upper bound for the same detection probability.

IX. DISCUSSION

From (44), $I(P_s; P_m) \propto \sqrt{n}$ which implies that the upper bound for the embedding rate occurs when each clique contains a single element, i.e. when n is large. This occurs in cases such as a highly textured image or a motion field of a sandstorm video as in Fig. 6. Conversely, the lowest
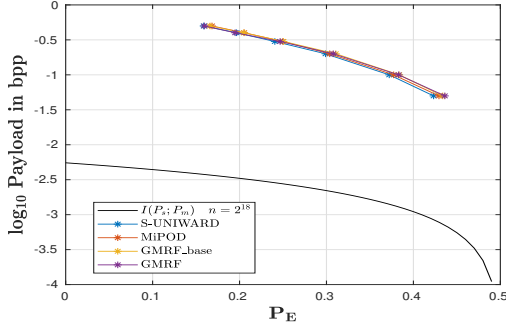
Fig. 3: Payload versus P_E from [16] for the steganographic methods S-UNIWARD, MIPOD, GMRF_BASE and GMRF with steganalysis using maxSRMd2.
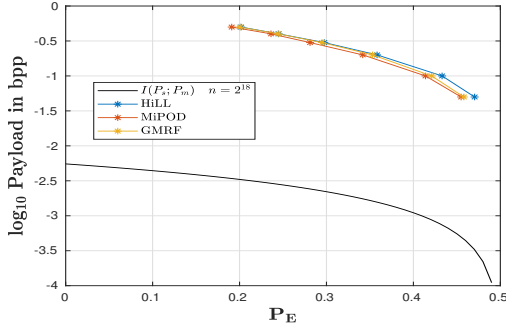

Fig. 4: Payload versus P_E from [16] for the steganographic methods MiPOD and GMRF (both enhanced using low-pass filtered cost) and HILL with steganalysis using SRM.

embedding rate occurs when all elements of the cover are correlated, i.e. the entire cover can be considered as a single clique so that $n = 1$. This can occur in cases such as an image of a clear sky or a motion field of a stationary video scene with global camera motion as in Fig. 7, both of which imply a small number of cliques. In other words, the largest value occurs when all elements of the cover are independent, whereas the smallest value occurs when the cover contains only a single clique.

It should be noted that the proof of the information-theoretic upper bound for $I(p_s; p_m)$ in (36) is not constructive due to the random coding argument. In other words, it does not provide a means of designing a steganographic method or codebook to achieve the upper bound. Only an analytic existence proof
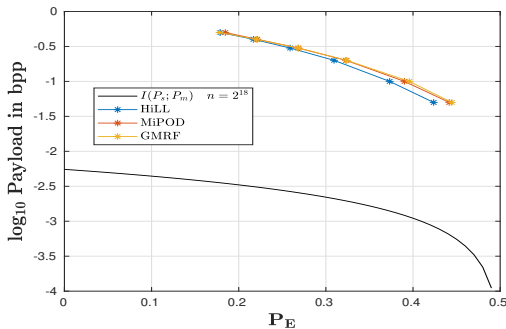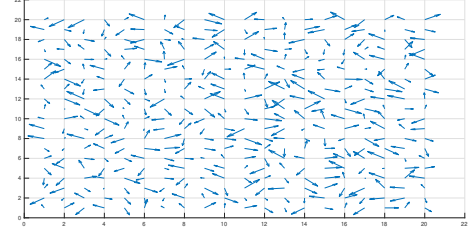

Fig. 5: Payload versus P_E from [16] for the steganographic methods MiPOD and GMRF (both enhanced using low-pass filtered cost) and HILL with steganalysis using maxSRMd2.

(a) A highly textured image (6184.pgm) from [33] with added zero mean Gaussian noise (variance 0.05).
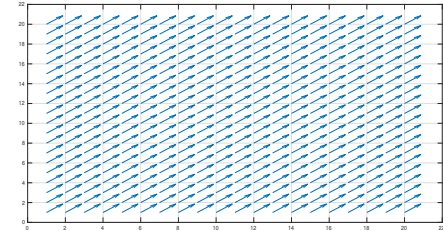

(b) A non-correlated motion field.

Fig. 6: Examples of covers allowing high embedding rates.

(a) A smooth image.

(b) A highly correlated motion field.

Fig. 7: Examples of covers allowing low embedding rates.

was provided here so further work is needed to construct steganographic methods close to the bound. However, the following can be inferred from Theorem 1 for steganography and steganalysis methods.

- The achievable embedding rate is monotonically increasing with the variance of the cover elements (or the entropy of the cover), as shown in Figures 6 and 7.
- The achievable embedding rate with respect to a given level of risk, i.e. best steganalysis performance, is monotonically increasing with the number of cover elements.
- The message elements should be mapped to a codebook with a distribution similar to that of the cover with zero mean.

X. Conclusion

An information-theoretic upper bound was presented for data embedding with multimedia covers for a given level of security. The contributions of this paper are as follows. First, a justification was given for using a CMQGD to approximate the intractable Gibbs distribution. Second, an analytic form for the SRL was derived. Third, an analytic form was given not only for the case when the cover is known by both the sender **A** and receiver **B**, but also for the practical case when **B** only knows the cover distribution. This was obtained using the achievability proof in Section VI. Fourth, the bound was used with several types of steganographic detectors and multimedia. Fifth, model parameters were given for the optimal message codebook in an analytic form. Note that the results presented for cover objects with a continuous distribution are exact closed-form solutions.

Appendix A
Extension of Theorem 2 in [28] for Non-zero Mean Gaussian Distributions

The proof of Theorem 2 in [28] for zero mean Gaussian distributions can be extended to the case of non-zero mean Gaussian distributions, i.e. $p_0(y) \sim \mathcal{N}(\mu, \sigma_w)$. Following the approach in [28] except that $p_0(y) = \frac{1}{\sigma_w\sqrt{2\pi}}e^{\frac{-1}{2\sigma_w^2}(y-\mu)^2}$, we have

$$\int_{-\infty}^{\infty} p_1(y)\ dy = \frac{e^{-\rho_0-1}}{\sigma_w\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma_w^2}(y-\mu)^2}e^{-\rho_1 y^2}\ dy,$$

$$= \frac{e^{-\rho_0-1-\frac{\mu}{2\sigma_w^2}}}{\sigma_w\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2(\rho_1+\frac{1}{2\sigma_w^2})}e^{\frac{\mu}{\sigma_w^2}y}\ dy.$$

Then using [34, (3.323.2)] gives

$$\int_{-\infty}^{\infty} e^{-a^2 x^2 \pm bx}dx = \frac{\sqrt{\pi}}{a}e^{\frac{b^2}{4a^2}},$$

so

$$\int_{-\infty}^{\infty} p_1(y)\ dy = \frac{1}{\sqrt{2\rho_1\sigma_w^2+1}}e^{\rho_0-1-\frac{\mu}{2\sigma_w^2}+\frac{\mu^2}{4\rho\sigma_w^4+2\sigma_w^2}}. \quad (45)$$

$$(46)$$

Equation (45) is an extension of (23) in [28]. Similarly, for (26) in [28] we have

$$\int_{-\infty}^{\infty} y^2 p_1(y)\ dy = \frac{e^{-\rho_0-1}}{\sigma_w\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{\frac{-1}{2\sigma_w^2}(y-\mu)^2}e^{-\rho_1 y^2}\ dy,$$

$$= \frac{e^{-\rho_0-1-\frac{\mu}{2\sigma_w^2}}}{\sigma_w\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{-y^2(\rho_1+\frac{1}{2\sigma_w^2})}$$

$$e^{\frac{\mu}{\sigma_w^2}y}\ dy.$$

Then from [34, (3.462.8)]

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2}e^{2bx}dx = \frac{1}{2a}\sqrt{\frac{\pi}{a}}\left(1+2\frac{b^2}{a}\right)e^{\frac{b^2}{a}},$$

which gives

$$\int_{-\infty}^{\infty} y^2 p_1(y)\ dy = \frac{e^{-\rho_0-1-\frac{\mu}{2\sigma_w^2}}\sigma_w^2}{(1+2\rho_1\sigma_w^2)^{\frac{3}{2}}}\left(1+\frac{\mu}{4\sigma_w^4(\rho_1+\frac{1}{2\sigma_w^2})}\right)$$

$$e^{\frac{\mu^2}{4\sigma_w^4(\rho_1+\frac{1}{2\sigma_w^2})}}. \quad (47)$$

Substituting (17-b) from [28] in (45) we have

$$e^{\rho_0-1} = \frac{1}{\sqrt{2\rho_1\sigma_w^2+1}}e^{-\frac{\mu}{2\sigma_w^2}+\frac{\mu^2}{4\rho\sigma_w^4+2\sigma_w^2}}, \quad (48)$$

which is a generalization of (25) at [28]. In addition, substituting (17-c) from [28] in (45) and (47) results in

$$P_y = \frac{2\sigma_w^2(1+2\rho_1\sigma_w^2)+\mu}{2(2\rho_1\sigma_w^2+1)^2},$$

and solving for ρ_1 gives

$$\rho_1 = \frac{\sigma_w^2-2P_y+\sqrt{\sigma_w^4+2\mu P_y}}{4\sigma_w^2 P_y}, \quad (49)$$

which generalizes (28) in [28]. Now, substituting (48) and (49) in (22) in [28], we have

$$p_1(y) = \frac{1}{\sqrt{2\pi P_y}}e^{\frac{-1}{2P_y^2}(y-\mu)^2}. \quad (50)$$

which generalizes (29) in [28]. Using this expression in the remainder of the proof of Theorem 2 in [28] proves the generalization for non-zero mean Gaussian distributions, and this can be extended to multivariate Gaussian distributions such as CMQGD. ■

Appendix B
Entropy of a Multivariate Quantized Gaussian Distribution

Definition 1. *The entropy of P_c and P_s.*

Theorem 8.3.1 in [27] defines the entropy of a quantized distribution as

$$H(P) \approx h(p) + b, \quad (51)$$

where p is a continuous distribution, P is a quantized version of p, b is the number of quantization bits, $H(P)$ is the entropy of the quantized distribution and $h(p)$ is the differential entropy of the continuous distribution. The entropy of CMQGD p_g is defined in [12] as

$$h(p_g) = \frac{1}{2}\ln(2\pi e|\Sigma_g|), \quad (52)$$

so the entropy of P_c and P_s can be approximated w.r.t. the corresponding continuous versions of P_c and P_s as

$$H(P_c) \approx \frac{1}{2}\ln(2\pi e|\Sigma_c|) + b, \quad (53)$$

$$H(P_s) \approx \frac{1}{2}\ln(2\pi e|\Sigma_s|) + b. \quad (54)$$

Appendix C
Additional Properties of the Embedding Factor a

Corollary 1. *a cannot be a complex number.*

Note that $a \in \mathbb{R}$ although $W(.) \in \mathbb{C}$. From [29], the relationship between the Lambert W Function $\mathcal{W}(.)$ and the WrightOmega function $W(.)$ can be expressed as

$$W(x) = \mathcal{W}_{\left\lceil\frac{\text{Im}(x)-\pi}{2\pi}\right\rceil}(e^x). \quad (55)$$

If $x = -\frac{4\epsilon^2}{n} - 1 - i\pi$ as in (34), then $\text{Im}(x) = \pi$ and therefore (55) can be rewritten as

$$W(x) = \mathcal{W}_{-1}(e^x), \tag{56}$$

which is a special case of the Lambert W Function where $\mathcal{W}_{-1}(.) \in \mathbb{R}$ [35]. Thus, $a \in \mathbb{R}$, and this also applies to a in (33).

Corollary 2. *a cannot be less than 1.*

As Σ_s, Σ_c and Σ_m are positive semi-definite matrices, from (29) a must be greater than 1 with equality when no embedding occurs from (26) and (29). Thus

$$a \geq 1. \tag{57}$$

Corollary 3. *The relation between P_E and a.*

From (1)-(4) and (6), we have

$$2\left(1 - 2P_E\right)^2 \leq \mathcal{D}(P_s \parallel P_c) \leq 2\epsilon^2, \tag{58}$$

Then, as the WrightOmega function in the form $-W(-\gamma - 1 - i\pi)$ is monotonically increasing with γ [29], where γ is a positive constant, from (33), (34) and (58) we have that

$$-W\left(-\frac{4\left(1 - 2P_E\right)^2}{n} - 1 - i\pi\right) \leq a \leq -W\left(-\frac{4\epsilon^2}{n} - 1 - i\pi\right). \tag{59}$$

REFERENCES

[1] P. Joseph and S. Vishnukumar, "A study on steganographic techniques," in *Global Conference on Communication Technologies*, Thuckalay, India, Apr. 2015, pp. 206–210.

[2] H. A. Aly, "Data hiding in motion vectors of compressed video based on their associated prediction error," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pp. 14–18, Mar. 2011.

[3] A. Westfeld, "F5—a steganographic algorithm," in *Information Hiding*. Berlin, Heidelberg: Springer Berlin Heidelberg, Oct. 2001, pp. 289–302.

[4] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge, GB: Cambridge University Press, 2009.

[5] J. M. Jenifer, S. R. Ratna, J. S. Loret, and D. M. Gethsy, "A survey on different video steganography techniques," in *International Conference on Trends in Electronics and Informatics*. Tirunelveli, India: IEEE, May 2018, pp. 627–632.

[6] Y. Tew and K. Wong, "An overview of information hiding in H.264/AVC compressed video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 2, pp. 305–319, Feb. 2014.

[7] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[8] E. Dubois and J. Konrad, "Motion estimation and motion-compensated filtering of video signals," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, Minneapolis, MN, USA, Apr. 1993, pp. 95–98.

[9] J. Konrad, "Motion detection and estimation," in *The Essential Guide to Video Processing*, A. Bovik, Ed. Boston, MA, USA: Academic Press, 2009, pp. 31–67.

[10] M. Sharifzadeh, M. Aloraini, and D. Schonfeld, "Adaptive batch size image merging steganography and quantized Gaussian image steganography," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 867–879, Jul. 2020.

[11] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1921–1930, Sep. 2013.

[12] A. Abdelaziz and C. E. Koksal, "Fundamental limits of covert communication over MIMO AWGN channel," in *IEEE Conference on Communications and Network Security*, Las Vegas, NV, USA, Oct. 2017, pp. 1–9.

[13] V. Sedighi, J. Fridrich, and R. Cogranne, "Content-adaptive pentary steganography using the multivariate generalized Gaussian cover model," in *Media Watermarking, Security, and Forensics 2015*, A. M. Alattar, N. D. Memon, and C. D. Heitzenrater, Eds., vol. 9409, International Society for Optics and Photonics. San Francisco, CA, USA: SPIE, Mar 2015, pp. 144–156.

[14] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 221–234, Feb. 2016.

[15] J. Fridrich and J. Kodovský, "Multivariate Gaussian model for designing additive distortion for steganography," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, CA, May 2013, pp. 2949–2953.

[16] W. Su, J. Ni, X. Hu, and J. Fridrich, "Image steganography with symmetric embedding using Gaussian Markov random field model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 3, pp. 1001–1015, Mar. 2020.

[17] Y. Tong, J. Ni, and W. Su, "Image steganography using an eight-element neighborhood Gaussian Markov random field model," in *Digital Forensics and Watermarking*, H. Wang, X. Zhao, Y. Shi, H. J. Kim, and A. Piva, Eds. Chengdu, China: Springer International Publishing, 2020, pp. 247–255.

[18] H. Y. El-Arsh, A. Abdelaziz, A. Elliethy, and H. A. Aly, "Fundamental limits of steganographic capacity for multivariate-quantized-Gaussian-distributed multimedia," in *IEEE International Conference on Image Processing (ICIP)*, Abu Dhabi, UAE, Oct. 2020, pp. 1246–1250.

[19] T. Pevný, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Information Hiding*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 161–177.

[20] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP Journal on Information Security*, vol. 2014, no. 1, Jan. 2014.

[21] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *IEEE International Conference on Image Processing*, Paris, France, Oct. 2014, pp. 4206–4210.

[22] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich, "Selection-channel-aware rich model for steganalysis of digital images," in *2014 IEEE International Workshop on Information Forensics and Security*, Atlanta, GA, USA, Dec. 2014, pp. 48–53.

[23] J. Kodovsky and J. Fridrich, "Steganalysis of JPEG images using rich models," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 8303, pp. 7–19, Feb. 2012.

[24] J. Jia, Z. Xiang, L. Wang, and Y. Xu, "An adaptive JPEG double compression steganographic scheme based on irregular DCT coefficients distribution," *IEEE Access*, vol. 7, pp. 119 506–119 518, Jul. 2019.

[25] D. Xu, "Commutative encryption and data hiding in HEVC video compression," *IEEE Access*, vol. 7, pp. 66 028–66 041, May 2019.

[26] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. New York, NY, USA: Springer Science & Business Media, 2006.

[27] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: John Wiley & Sons, Ltd, 2006.

[28] S. Yan, Y. Cong, S. V. Hanly, and X. Zhou, "Gaussian signalling for covert communications," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3542–3553, Jul. 2019.

[29] R. M. Corless and D. J. Jeffrey, "The Wright ω function," in *Proceedings of the Joint International Conferences on Artificial Intelligence, Automated Reasoning, and Symbolic Computation*, ser. AISC '02/Calculemus '02. Berlin, Heidelberg: Springer-Verlag, 2002, pp. 76–89.

[30] K. B. Petersen and M. S. Pedersen, "The Matrix Cookbook," Nov. 2012, version 20121115. [Online]. Available: http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html

[31] I. Chatzigeorgiou, "Bounds on the Lambert function and their application to the outage analysis of user cooperation," *IEEE Communications Letters*, vol. 17, no. 8, pp. 1505–1508, Aug. 2013.

[32] P. Bas, T. Filler, and T. Pevný, "Break our steganographic system: the ins and outs of organizing BOSS," in *Information Hiding*. Berlin, Heidelberg: Springer, 2011, pp. 59–70.

[33] "2nd BOWS contest (Break Our Watermarking System)," The 10000 original images used during Ep.3, http://bows2.ec-lille.fr/. [Online]. Available: http://bows2.ec-lille.fr/BOWS2OrigEp3.tgz

[34] "3–4 - definite integrals of elementary functions," in *Table of Integrals, Series, and Products*, 7th ed., A. Jeffrey, D. Zwillinger, I. Gradshteyn, and I. Ryzhik, Eds. Boston, MA, USA: Academic Press, 2007, pp. 247–617.

[35] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, Dec. 1996.