Square Root Law for Communication with Low Probability of Detection on AWGN Channels

Boulat A. Bash*, Dennis Goeckel†, Don Towsley*

*Department of Computer Science, University of Massachusetts, Amherst, Massachusetts 01003–9264
†Electrical and Computer Engineering Department, University of Massachusetts, Amherst, Massachusetts 01003–9292

***Abstract*—We present a square root limit on *low probability of detection* (LPD) communication over additive white Gaussian noise (AWGN) channels. Specifically, if a warden has an AWGN channel to the transmitter with non-zero noise power, we prove that $o(\sqrt{n})$ bits can be sent from the transmitter to the receiver in n AWGN channel uses with probability of detection by the warden less than ϵ for any $\epsilon > 0$, and, if a lower bound on the noise power on the warden's channel is known, then $\mathcal{O}(\sqrt{n})$ bits can be covertly sent in n channel uses. Conversely, trying to transmit more than $\mathcal{O}(\sqrt{n})$ bits either results in detection by the warden with probability one or a non-zero probability of decoding error as $n \to \infty$. Further, we show that LPD communication on the AWGN channel allows one to send a non-zero symbol on *every* channel use, in contrast to what might be expected from the square root law found recently in image-based steganography.**

I. Introduction

Securing information transmitted over wireless links is of paramount concern for consumer, industrial, and military applications. Traditional encryption and key exchange protocols secure data from interception by an untrusted third party. However, there are many real-life situations where it is imperative to prevent the transmission from being *detected* in the first place, as encrypted data arouses suspicion, and even the most theoretically robust encryption can often be defeated by a determined adversary using non-computational methods such as side-channel analysis. In spite of its importance, *low probability of detection* (LPD) communication has been relatively underexplored. In this work we examine the fundamental limitations of LPD communication over wireless links.

In our scenario, Alice communicates with Bob over a channel subject to additive white Gaussian noise (AWGN), while Willie attempts to detect her transmission (without actively jamming Alice's channel). The channel between Willie and Alice is also subject to AWGN. Alice sends low-power covert signals to Bob that Willie attempts to classify as either noise on his channel from Alice or Alice's signals to Bob. If the noise on the channel between Willie and Alice has non-zero power, Alice can communicate with Bob while tolerating a certain probability of detection, which she can drive down by transmitting with low enough power.

Our problem is related to the problem of imperfect steganography, but the two problems are not the same. Steganography considers hiding information by altering the properties of fixed-size, finite-alphabet covertext objects (like images or software binary code) with imperfect steganography systems allowing a fixed probability of detection of hidden information. Covertext can be considered a type of lossless finite-alphabet channel. However, the square root law recently found in this environment [1], which states that $\mathcal{O}(\sqrt{n})$ symbols in the original covertext of size n may safely be modified to hide a message, is limited in its scope. The continuous-valued channel allows us to spread hidden information across *every* symbol used in the transmission, thus showing that a direct application of the steganographic result quickly leads to contradiction and demonstrating the distinction between the two problems. In fact, our square root law can be viewed as generalizing the square root law for imperfect steganography.

We state our main result that limits mutual information on the covert channel between Alice and Bob using asymptotic notation where $f(n) = \mathcal{O}(g(n))$ denotes an asymptotically tight upper bound on $f(n)$, and $f(n) = o(g(n))$ and $f(n) = \omega(g(n))$ denote upper and lower bounds, respectively, that are not asymptotically tight [2, Ch. 3.1]:

Theorem (Square root law)**.** *Suppose the channel between Alice and each of Bob and Willie experiences independent additive white Gaussian noise (AWGN) with power $\sigma_b^2 > 0$ and $\sigma_w^2 > 0$, respectively, where σ_b^2 and σ_w^2 are constants. Then, for any $\epsilon > 0$ and unknown σ_w^2, Alice can send $o(\sqrt{n})$ information bits to Bob in n channel uses while maintaining a probability of detection of Alice's transmission by Willie of less than ϵ. Moreover, if Alice can lower-bound $\sigma_w^2 \geq \hat{\sigma}_w^2$, she can send $\mathcal{O}(\sqrt{n})$ bits in n channel uses while maintaining a probability of detection of less than ϵ. Conversely, if Alice attempts to transmit $\omega(\sqrt{n})$ bits in n channel uses, then, as $n \to \infty$, either Willie detects her with arbitrary low probability of error or Bob cannot decode her message reliably (i.e. with arbitrary low probability of decoding error).*

This research was sponsored by the National Science Foundation under grants CNS-0905349 and CNS-1018464, and by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

After introducing our discrete-time channel model and hypothesis testing background in Section II, we sketch the proofs of the achievability and the converse of the square root law in Sections III and IV, respectively. Detailed proofs and remarks are available in [3]. We discuss the mapping to the continuous-time channel and the implications of channel fading on our results, as well as the relationship to previous work in Section V, and conclude in Section VI.

II. Prerequisites

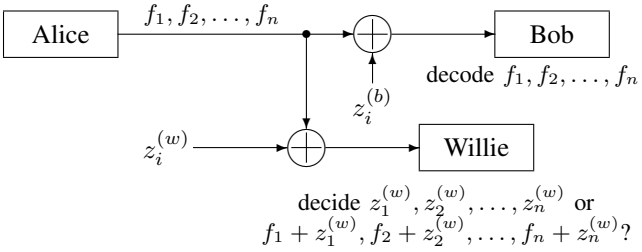

Fig. 1. System framework: Alice encodes information into a vector of real symbols $\mathbf{f} = \{f_i\}_{i=1}^n$ and transmits it on an AWGN channel to Bob, while Willie attempts to classify his vector of observations of the channel from Alice $\mathbf{y}_w$ as either an AWGN vector $\mathbf{z}_w = \{z_i^{(w)}\}_{i=1}^n$ or a vector $\{f_i + z_i^{(w)}\}_{i=1}^n$ of transmissions corrupted by AWGN.

A. Channel Model

We use the discrete-time AWGN channel model with real-valued symbols depicted in Figure 1 (and defer discussion of the mapping to a continuous-time channel as well as a fading channel to Section V). Alice transmits a vector of n real-valued symbols $\mathbf{f} = \{f_i\}_{i=1}^n$. Bob receives vector $\mathbf{y}_b = \{y_i^{(b)}\}_{i=1}^n$ where $y_i^{(b)} = f_i + z_i^{(b)}$ with an independent and identically distributed (i.i.d.) $z_i^{(b)} \sim \mathcal{N}(0, \sigma_b^2)$. Willie observes vector $\mathbf{y}_w = \{y_i^{(w)}\}_{i=1}^n$ where $y_i^{(w)} = f_i + z_i^{(w)}$, with i.i.d. $z_i^{(w)} \sim \mathcal{N}(0, \sigma_w^2)$.

B. Hypothesis Testing

Willie expects vector $\mathbf{y}_w$ to be consistent with his channel noise model. He performs a statistical hypothesis test on this vector, with the null hypothesis H_0 being that Alice is not covertly communicating. This corresponds to each sample $y_i^{(w)} \sim \mathcal{N}(0, \sigma_w^2)$ i.i.d. The alternate hypothesis H_1 is that Alice is transmitting, which corresponds to samples $y_i^{(w)}$ from a different distribution. Willie tolerates some cases when his test incorrectly accuses Alice. Following the standard nomenclature, we denote the probability of such rejection of H_0 when it is true by α [4]. Willie's test may also accept H_0 when it is false and miss Alice's covert transmission. We denote the probability of a miss by β. The sum $\alpha + \beta$ determines the performance of a hypothesis test [4].

III. Achievability of Square Root Law

In our scenario, Alice and Bob construct a covert communications system, with all the details known to Willie except for a secret key shared before communication. This follows "best practices" in security system design, as its security depends only on the key [5]. Since this work concerns the limits of covert communication, key size is not a constraint and we defer the study of key efficiency to future work.

Willie tries to determine whether Alice transmitted covert data given the vector of observations $\mathbf{y}_w$ of his channel from Alice. Denote the probability distribution of Willie's channel observations when H_0 is true as $\mathbf{P}_0$, and when H_1 is true as $\mathbf{P}_1$. To strengthen the achievability result, we assume that Alice's channel input distribution, as well as the distribution of AWGN on the channel between Alice and Willie are known to Willie. Then $\mathbf{P}_0$ and $\mathbf{P}_1$ are known to Willie, and he can construct an optimal statistical hypothesis test that minimizes the sum of error probabilities $\alpha + \beta$ [4, Ch. 13]. Then:

Fact 1 (Theorem 13.1.1 in [4])**.** *For the optimal test:*

$$\alpha + \beta \quad = \quad 1 - TV(\mathbf{P}_0, \mathbf{P}_1)$$

where $TV(\mathbf{P}_0, \mathbf{P}_1) = \frac{1}{2}\int_{-\infty}^{\infty} |p_0(x) - p_1(x)|dx$ is the *total variation distance* between $\mathbf{P}_0$ and $\mathbf{P}_1$ and $p_0(x)$ and $p_1(x)$ are densities of $\mathbf{P}_0$ and $\mathbf{P}_1$, respectively. Unfortunately, the total variation metric is unwieldy for the products of probability measures, which are used in the analysis of the vectors of observations. We thus use [6, Lemma 11.6.1]:

$$\frac{1}{2}\left(\int_{-\infty}^{\infty} |p_0(x) - p_1(x)|dx\right)^2 \leq D(\mathbf{P}_0 \| \mathbf{P}_1) \tag{1}$$

where relative entropy $D(\mathbf{P}_0\|\mathbf{P}_1) = \int_{\mathcal{X}} p_0(x) \ln \frac{p_0(x)}{p_1(x)} dx$ with $\mathcal{X}$ being the support of $p_1(x)$. If $\mathbf{P}^n$ is the distribution of a sequence $\{X_i\}_{i=1}^n$ where each $X_i \sim \mathbf{P}$ is i.i.d., then $D(\mathbf{P}_0^n\|\mathbf{P}_1^n) = nD(\mathbf{P}_0\|\mathbf{P}_1)$. Now we are ready to prove the achievability theorem under an average power constraint.

Theorem 1.1 (Achievability)**.** *Let Willie's channel be subject to AWGN with power $\sigma_w^2 > 0$. Then Alice can maintain Willie's sum of the probabilities of detection errors $\alpha + \beta \geq 1 - \epsilon$ for any $\epsilon > 0$ while covertly transmitting $o(\sqrt{n})$ bits to Bob in n uses of an AWGN channel if σ_w^2 is unknown and $\mathcal{O}(\sqrt{n})$ bits in n channel uses if she can lower-bound $\sigma_w^2 \geq \hat{\sigma}_w^2$.*

Proof: **Construction:** Alice's channel encoder takes input in blocks of size M bits and encodes them into codewords of length n at the rate of $R = M/n$ bits/symbol. We employ random coding arguments and independently generate 2^{nR} codewords $\{\mathbf{c}(W_k), k = 1, 2, \ldots, 2^{nR}\}$ from $\mathbb{R}^n$ for messages W_k, each according to $p_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n p_X(x_i)$, where $X \sim \mathcal{N}(0, P_f)$ and P_f is defined later. The codebook is the secret key shared between Alice and Bob, and is not revealed to Willie.[1] However, Willie knows how it is constructed, including the value of P_f, and he uses statistical hypothesis testing on n channel readings $\mathbf{y}_w$ to decide whether Alice transmitted.

Analysis: Consider the case when Alice transmits codeword $\mathbf{c}(W_k)$. Suppose that Willie employs a detector that implements an optimal hypothesis test on his n channel readings.

[1] Willie's lack of knowledge of the codebook is critical to our result, as the sparsity of the codewords implies that, if the codeword is correctly decoded by Willie, then the transmission is detected.

His null hypothesis H_0 is that he observed noise on his channel. His alternate hypothesis H_1 is that Alice transmitted and he observed Alice's codeword corrupted by noise. By Fact 1, the sum of the probabilities of Willie's detector's errors is expressed by $\alpha + \beta = 1 - TV(\mathbf{P}_0, \mathbf{P}_1)$, where the total variation distance is between the distribution $\mathbf{P}_0$ of n noise readings that Willie expects to observe under his null hypothesis and the distribution $\mathbf{P}_1$ of the covert codeword transmitted by Alice corrupted by noise. Alice can lower-bound $\alpha + \beta$ by upper-bounding $TV(\mathbf{P}_0, \mathbf{P}_1) \leq \epsilon$.

The realizations of noise $z_i^{(w)}$ in vector $\mathbf{z}_w$ are zero-mean i.i.d. Gaussian random variables with variance σ_w^2, and, thus, $\mathbf{P}_0 = \mathbf{P}_w^n$ where $\mathbf{P}_w = \mathcal{N}(0, \sigma_w^2)$. Since he does not know the codebook, Willie's probability distribution of the transmitted symbols is of zero-mean i.i.d. Gaussian random variables with variance P_f. Since noise is independent of the transmitted symbols, when Alice transmits, Willie observes $\mathbf{y}_w$, where $y_i^{(w)} \sim \mathcal{N}(0, P_f + \sigma_w^2) = \mathbf{P}_s$ is i.i.d., and $\mathbf{P}_1 = \mathbf{P}_s^n$. Then, using (1), the properties of relative entropy, and the Taylor series expansion of $D(\mathbf{P}_w \| \mathbf{P}_s)$ with respect to P_f around $P_f = 0$, we obtain the upper bound:

$$TV(\mathbf{P}_w^n, \mathbf{P}_s^n) \leq \frac{P_f}{2\sigma_w^2}\sqrt{\frac{n}{2}} \tag{2}$$

Suppose Alice sets her average covert symbol power $P_f \leq \frac{cf(n)}{\sqrt{n}}$, where $c = 2\epsilon\sqrt{2}$. In most practical scenarios Alice can lower-bound $\sigma_w^2 \geq \hat{\sigma}_w^2$ and set $f(n) = \hat{\sigma}_w^2$. If σ_w^2 is unknown, select $f(n)$ such that $f(n) = o(1)$ and $f(n) = \omega(1/\sqrt{n})$ (the latter condition is used to bound Bob's decoding error probability). In either case, for large n, $P_f < \sigma_w^2$ satisfies the Taylor series convergence criterion for (2) to be valid, and Alice upper-bounds $TV(\mathbf{P}_w^n, \mathbf{P}_s^n) \leq \epsilon$, limiting the performance of Willie's detector.

As standard results for constant symbol power are not directly applicable to our system where P_f is a function of codeword size n, we examine the probability $\mathbf{P}_e$ of Bob's decoding error averaged over all possible codebooks. Let Bob employ a maximum-likelihood (ML) decoder (i.e. minimum distance) to process the received vector $\mathbf{y}_b$ when $\mathbf{c}(W_k)$ was sent. The decoder suffers an error event $E_i(\mathbf{c}(W_k))$ when $\mathbf{y}_b$ is closer to another codeword $\mathbf{c}(W_i)$, $i \neq k$:

$$\begin{aligned}\mathbf{P}_e &= \mathbf{E}_{\mathbf{c}(W_k)}\left[\mathbf{P}\left(\cup_{i=0, i\neq k}^{2^{nR}} E_i(\mathbf{c}(W_k))\right)\right] \\ &\leq \sum_{i=0, i\neq k}^{2^{nR}} \mathbf{E}_{\mathbf{c}(W_k)}\left[\mathbf{P}\left(E_i(\mathbf{c}(W_k))\right)\right] \end{aligned} \tag{3}$$

where (3) follows from the union bound and the linearity of expectation. The distance between two codewords is $d = \|\mathbf{c}(W_k) - \mathbf{c}(W_i)\|_2$, where $\|\cdot\|_2$ is the $\mathcal{L}^2$ norm. Since codewords are independent and Gaussian, $\mathbf{c}(W_k) - \mathbf{c}(W_i) \sim \mathcal{N}(0, 2P_f)$ and $d^2 = 2P_f U$, where $U \sim \chi_n^2$, with χ_n^2 denoting the chi-squared distribution with n degrees of freedom. Therefore, by [7, (3.44)]:

$$\mathbf{E}_{\mathbf{c}(W_k)}\left[\mathbf{P}\left(E_i(\mathbf{c}(W_k))\right)\right] = \mathbf{E}_U\left[Q\left(\sqrt{\frac{P_f U}{2\sigma_b^2}}\right)\right]$$

where $Q(x) = \int_x^\infty \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \leq \frac{1}{2}e^{-x^2/2}$ [8, (5)]. Taking the expectation yields:

$$\mathbf{E}_{\mathbf{c}(W_k)}\left[\mathbf{P}\left(E_i(\mathbf{c}(W_k))\right)\right] \leq 2^{-\frac{n}{2}\log_2\left(1+\frac{cf(n)}{2\sqrt{n}\sigma_b^2}\right)}$$

The summand in (3) does not depend on i, and (3) becomes:

$$\mathbf{P}_e \leq 2^{nR - \frac{n}{2}\log_2\left(1+\frac{cf(n)}{2\sqrt{n}\sigma_b^2}\right)}$$

Since $f(n) = \omega(1/\sqrt{n})$, if rate $R = \frac{\rho}{2}\log_2\left(1 + \frac{cf(n)}{2\sqrt{n}\sigma_b^2}\right)$ for a constant $\rho < 1$, as n increases, the probability of Bob's decoding error decays exponentially to zero and Bob obtains $nR = \sqrt{n}\frac{\rho}{2}\log_2\left(1 + \frac{cf(n)}{2\sqrt{n}\sigma_b^2}\right)^{\sqrt{n}}$ covert bits in n channel uses. Since $nR \leq \frac{\sqrt{n}\rho c f(n)}{4\sigma_b^2 \ln 2}$, approaching equality as n gets very large, Bob receives $nR = o(\sqrt{n})$ bits in n channel uses, and $nR = \mathcal{O}(\sqrt{n})$ bits in n channel uses if $f(n) = \hat{\sigma}_w^2$. ■

Implications of a peak power constraint

Since most practical systems are peak-power constrained, we show that the square root law holds for the binary input Gaussian output channel using a proof similar to that of Theorem 1.1.

Theorem 1.2 (Achievability under a peak power constraint)**.** *Suppose Alice's transmitter is subject to the peak power constraint b and Willie's channel is subject to AWGN with power $\sigma_w^2 > 0$. Then Alice can maintain Willie's sum of the probabilities of detection errors $\alpha + \beta \geq 1 - \epsilon$ for any $\epsilon > 0$ while covertly transmitting $o(\sqrt{n})$ bits to Bob over n uses of an AWGN channel if σ_w^2 is unknown and $\mathcal{O}(\sqrt{n})$ bits in n channel uses if she can lower-bound $\sigma_w^2 \geq \hat{\sigma}_w^2$.*

Proof: **Construction:** Alice encodes the input in blocks of size M bits into codewords of length n at the rate $R = M/n$ bits/symbol with the symbols drawn from alphabet $\{-a, a\}$, where a satisfies the peak power constraint $a^2 < b$ and is defined later. We independently generate 2^{nR} codewords $\{\mathbf{c}(W_k), k = 1, 2, \ldots, 2^{nR}\}$ for messages W_k from $\{-a, a\}^n$ according to $p_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n p_X(x_i)$, where $p_X(-a) = p_X(a) = \frac{1}{2}$. As in the proof of Theorem 1.1, the codebook is a secret key shared between Alice and Bob, but Willie knows how it is constructed, including the value of a.

Analysis: When Alice transmits a covert symbol during the i^{th} symbol period, she transmits $-a$ or a equiprobably by construction and Willie observes the covert symbol corrupted by AWGN. Therefore, $\mathbf{P}_s = \frac{1}{2}\left(\mathcal{N}(-a, \sigma_w^2) + \mathcal{N}(a, \sigma_w^2)\right)$. Again, using (1), the properties of relative entropy, and the Taylor series expansion of $D(\mathbf{P}_w \| \mathbf{P}_s)$ with with respect to a around $a = 0$, we obtain the upper bound:

$$TV(\mathbf{P}_w^n, \mathbf{P}_s^n) \leq \frac{a^2}{2\sigma_w^2}\sqrt{\frac{n}{2}} \tag{4}$$

Since the power of Alice's covert symbol is $a^2 = P_f$, (4) is identical to (2) and Alice sets $a^2 \leq \frac{cf(n)}{\sqrt{n}}$, where c and $f(n)$ are defined as in Theorem 1.1. Then, for n large enough, $a < \sigma_w$ satisfies the Taylor series convergence criterion, and

Alice obtains the upper bound $TV(\mathbf{P}_w^n, \mathbf{P}_s^n) \leq \epsilon$, limiting the performance of Willie's detector.

Like in Theorem 1.1, we cannot apply standard constant-power channel coding results. Thus, we upper-bound Bob's decoding error probability by analyzing a suboptimal decoding scheme. Suppose Bob uses a hard-decision device on each received covert symbol $y_i^{(b)} = f_i + z_i^{(b)}$ via the rule $\hat{f}_i = \left\{a \text{ if } y_i^{(b)} \geq 0; -a \text{ otherwise}\right\}$, and applies an ML decoder on its output. The effective channel for the encoder/decoder pair is a binary symmetric channel with cross-over probability $p_e = Q(a/\sigma_b)$ and the probability of the decoding error averaged over all possible codebooks is $\mathbf{P}_e \leq 2^{nR - n(1-\mathcal{H}(p_e))}$ [9], where $\mathcal{H}(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy function. Extending the analysis in [10, Section I.2.1], we first use the Taylor series expansion of p_e with respect to a around $a = 0$ to upper-bound $p_e \leq p_e^{(UB)}$. Since $\mathcal{H}(p_e) \leq \mathcal{H}(p_e^{(UB)})$ on the interval $[0, \frac{1}{2}]$, we perform Taylor series expansion of $\mathcal{H}(p_e^{(UB)})$ with respect to a around $a = 0$ to obtain $\mathbf{P}_e \leq 2^{nR - \frac{\sqrt{n}cf(n)}{\sigma_b^2 \pi \ln 2} + \mathcal{O}(1)}$. As $f(n) = \omega(1/\sqrt{n})$, if rate $R = \frac{\rho c f(n)}{\sqrt{n}\sigma_b^2 \pi \ln 2}$ bits/symbol for a constant $\rho < 1$, the probability of Bob's decoding error decays exponentially to zero as n increases and Bob obtains $nR = o(\sqrt{n})$ bits in n channel uses, and $nR = \mathcal{O}(\sqrt{n})$ bits in n channel uses if $f(n) = \hat{\sigma}_w^2$. ∎

Relationship with Square Root Law in Steganography

It has recently been shown that in finite-alphabet imperfect steganographic systems at most $\mathcal{O}(\sqrt{n})$ symbols in the original covertext of size n may safely be modified to hide a steganographic message [1]. From the steganographic perspective, our covertext is the noise on Willie's channel to Alice. But our result does not obey their converse, as we can modify *all* symbols in our covertext, highlighting the different nature of the problem scenarios. To demonstrate the richness of our scenario and the generality of our square root law, we construct a codebook where roughly τn out of n of symbols are used to carry the message and when Alice is transmitting a codeword, the distribution of each of Willie's observations is $\mathbf{P}_s = (1-\tau)\mathcal{N}(0, \sigma_w^2) + \tau\mathcal{N}(0, P_f + \sigma_w^2)$. Again, using (1), the properties of relative entropy, and the Taylor series expansion of $D(\mathbf{P}_w \| \mathbf{P}_s)$ with respect to P_f around $P_f = 0$ yields the following bound:

$$TV(\mathbf{P}_w^n, \mathbf{P}_s^n) \quad \leq \quad \frac{\tau P_f}{2\sigma_w^2}\sqrt{\frac{n}{2}} \tag{5}$$

The only difference in (5) from (2) is τ in the numerator. Thus, if Alice sets the product $\tau P_f = \frac{cf(n)}{\sqrt{n}}$, with c and $f(n)$ as previously defined, she limits the performance of Willie's detector. This product is the average symbol power used by Alice. It is easy to verify that in the peak power constrained scenario Alice should set product $\tau a^2 = \frac{cf(n)}{\sqrt{n}}$ and that the number of bits that Alice can covertly transmit to Bob obeys the square root bounds.

IV. Converse

Here, as in the achievability, the channel between Alice and Bob is subject to AWGN of power σ_b^2. Alice's objective is to covertly transmit a message W_k that is $M = \omega(\sqrt{n})$ bits long to Bob in n channel uses with arbitrarily small probability of decoding error as n gets large. For an upper bound on the reduction in entropy, the messages are chosen equiprobably. Alice encodes each message W_k arbitrarily into n symbols at the rate $R = M/n$ symbols/bit. In the converse that follows Willie observes all n of Alice's channel uses, but, to strengthen the result, he is oblivious to her signal properties.

Theorem 2. *If over n channel uses, Alice attempts to transmit a covert message to Bob that is $\omega(\sqrt{n})$ bits long, then, as $n \to \infty$, either Willie can detect her with arbitrarily low sum of error probabilities $\alpha + \beta$, or Bob cannot decode with arbitrarily low probability of error.*

Proof: To detect Alice's covert transmissions, Willie performs the following hypothesis test:

$$\begin{aligned} H_0: &\quad y_i^{(w)} = z_i^{(w)},\ i = 1, \ldots, n \\ H_1: &\quad y_i^{(w)} = f_i + z_i^{(w)},\ i = 1, \ldots, n \end{aligned}$$

Rejection of H_0 means that Alice is covertly communicating with Bob. To perform the test, Willie collects a vector of n independent readings $\mathbf{y}_w$ from his channel to Alice and generates the test statistic $S = \frac{\mathbf{y}_w \mathbf{y}_w^T}{n}$ where $\mathbf{x}^T$ denotes transpose of vector $\mathbf{x}$. Under the null hypothesis H_0 Alice does not transmit and Willie reads AWGN. Thus, $y_i^{(w)} \sim \mathcal{N}(0, \sigma_w^2)$, and the mean and the variance of S are:

$$\mathbf{E}[S] \quad = \quad \sigma_w^2 \tag{6}$$

$$\mathbf{var}[S] \quad = \quad \frac{2\sigma_w^4}{n} \tag{7}$$

Suppose Alice transmits codeword $\mathbf{c}(W_k)$ with the average power per symbol $P_k = \frac{\mathbf{c}(W_k)\mathbf{c}^T(W_k)}{n}$. Then the mean and variance of S when Alice transmits message W_k are:

$$\mathbf{E}[S] \quad = \quad \sigma_w^2 + P_k \tag{8}$$

$$\mathbf{var}[S] \quad = \quad \frac{4P_k\sigma_w^2 + 2\sigma_w^4}{n} \tag{9}$$

The variance of Willie's test statistic (9) is computed by adding the variances conditioned on $\mathbf{c}(W_k)$ of the squared individual observations $\mathbf{var}\left[y_i^2\right]$ (and dividing by n^2) since the noise on the individual observations is independent.

Denote $\mathbf{P}_0$ as the distribution when H_0 holds, and $\mathbf{P}_1^{(k)}$ when H_1 holds with Alice transmitting message W_k. If H_0 is true, then S should be close to (6). Willie picks some threshold t and compares the value of S to $\sigma_w^2 + t$. He accepts H_0 if $S < \sigma_w^2 + t$ and rejects it otherwise. Suppose that he desires false positive probability α^*, which is the probability that $S \geq \sigma_w^2 + t$ when H_0 is true. We bound it using (6) and (7) with Chebyshev's Inequality [6, (3.32)]:

$$\alpha \quad \leq \quad \mathbf{P}_0\left(|S - \sigma_w^2| \geq t\right) \leq \frac{2\sigma_w^4}{nt^2}$$

Thus, to obtain α^*, Willie sets $t = \frac{d}{\sqrt{n}}$, where $d = \frac{\sqrt{2}\sigma_w^2}{\sqrt{\alpha^*}}$ is a constant. As n increases, t decreases, which is consistent with Willie gaining greater confidence with more observations.

Now let's bound the probability of a miss β given t, which is the probability that $S < \sigma_w^2 + t$ when Alice transmits W_k. We use Chebyshev's Inequality with (8) and (9):

$$\beta \leq \mathbf{P}_1^{(k)} \left(\left| S - \sigma_w^2 - P_k \right| \geq P_k - t \right) \leq \frac{4P_k\sigma_w^2 + 2\sigma_w^4}{(\sqrt{n}P_k - d)^2} \quad (10)$$

If $P_k = \omega(1/\sqrt{n})$, $\lim_{n\to\infty} \beta = 0$. Thus, with enough observations, Willie can detect with arbitrarily low error probability Alice's codewords with average symbol power $P_k = \omega(1/\sqrt{n})$. Note that Willie's detector is oblivious to any details of Alice's codebook construction.

By (10), if Alice wants to lower-bound the sum of the probabilities of error of Willie's statistical test by $\alpha + \beta \geq \zeta > 0$, she must use low-power codewords; in particular, a fraction $\gamma > 0$ of the codewords must have $P_{\mathcal{U}} = \mathcal{O}(1/\sqrt{n})$. Denoting this set of codewords by $\mathcal{U}$, we can lower-bound the probability of Bob's decoding error $\mathbf{P}_e \geq \gamma \mathbf{P}_e^{(\mathcal{U})}$, where $\mathbf{P}_e^{(\mathcal{U})}$ is the probability of decoding error when a message from $\mathcal{U}$ is sent. Focusing on $\mathcal{U}$, we adapt the proof of the converse to the coding theorem for Gaussian channels in [6, Ch. 9.2] to obtain:

$$\mathbf{P}_e^{(\mathcal{U})} \quad \geq \quad 1 - \frac{P_{\mathcal{U}}/2\sigma_b^2 + 1/n}{\frac{\log_2 \gamma}{n} + R} \quad (11)$$

Since Alice transmits $\omega(\sqrt{n})$ bits in n channel uses, her rate is $R = \omega(1/\sqrt{n})$ bits/symbol. However, $P_{\mathcal{U}} = O(1/\sqrt{n})$, and, as $n \to \infty$, $\mathbf{P}_e^{(\mathcal{U})}$ is bounded away from zero. Thus, $\mathbf{P}_e$ is bounded away from zero if Alice tries to beat Willie's simple hypothesis test. ■

V. Discussion

A. Mapping to Continuous-time Channel

Consider the mapping of the discrete-time model employed throughout this paper to the physical (continuous-time) channel. For a system that has a (baseband) bandwidth constraint of W Hz, if Alice employs the optimal bandlimited pulse shape $\mathrm{sinc}(2Wt)$, all of the information is extracted from the channel by Willie (and Bob) by sampling at rate $2W$ samples/second. This results in the discrete-time model of Section II, and the results presented here apply directly. When the pulse shape is instead chosen to have some excess bandwidth, then sampling at a rate higher than $2W$ has utility for Willie when attempting to detect Alice's transmission. Although we find it unlikely that the asymptotics considered here will be altered, techniques involving cyclostationary detection are applicable and could potentially impact practical system implementations.

B. Fading and Shadowing

Fading and shadowing will impact both the capacity of the channel from Alice to Bob and the ability for Willie to detect Alice's transmission. There are a number of different models that could be employed to incorporate these effects. However, while these models will have an impact as we move toward practical systems, they have little impact on the asymptotic results presented here.

C. Relationship to Previous Work

Analytical evaluation of LPD communication has been sparse. Hero studies LPD channels [11] in a multiple-input multiple-output (MIMO) setting. While he recognizes that an LPD communication system is constrained by average power, he does not analyze the constraint asymptotically and, thus, does not obtain the square root law. Unlike LPD communication, much analytical work has been done on steganography. As noted in the remark in Section III, the square root law was found in finite-alphabet imperfect steganography [1]. However, although their goal is the same as ours (hiding information with low probability of detection by Willie), their model based on hiding information in finite-alphabet images is very different from ours. Our scenario is arguably richer, and its additional degree of freedom in the choice of transmission power allows Alice to alter all n symbols used in transmission while maintaining a fixed detection probability, which stands in contrast to the finite-alphabet steganography result.

VI. Conclusion

We proved that the LPD communication is subject to a square root law in that the number of bits that can be covertly transmitted in n channel uses is $\mathcal{O}(\sqrt{n})$. An interesting result in our work is the fact that one can use all of the n symbols with positive power to transmit the covert messages. A promising direction of future research is the extension of this work to a practical networked setting. Eventually, we would like to answer this fundamental question: is it possible to establish and maintain a "shadow" wireless network in the presence of both active and passive wardens?

References

[1] T. Filler, A. D. Ker, and J. Fridrich, "The square root law of steganographic capacity for markov covers," *Media Forensics and Security*, vol. 7254, no. 1, 2009.

[2] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. Cambridge, Massachusetts: MIT Press, 2001.

[3] B. A. Bash, D. Goeckel, and D. Towlsey, "Square root law for communication with low probability of detection on awgn channels," University of Massachusetts, Tech. Rep. UM-CS-2012-003, in submission to IEEE Transactions on Wireless Communications.

[4] E. Lehmann and J. Romano, *Testing Statistical Hypotheses*, 3rd ed. New York: Springer, 2005.

[5] A. J. Menezes, S. A. Vanstone, and P. C. V. Oorschot, *Handbook of Applied Cryptography*, 1st ed. Boca Raton, FL, USA: CRC Press, Inc., 1996.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Hoboken, NJ, 2002.

[7] U. Madhow, *Fundamentals of Digital Communication*. Cambridge, UK: Cambridge University Press, 2008.

[8] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 840–845, Jul. 2003.

[9] A. Barg and G. D. Forney, Jr., "Random codes: minimum distances and error exponents," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2568–2573, Sep. 2002.

[10] E. E. Majani, "A model for the study of very noisy channels, and applications," Ph.D. dissertation, California Institute of Technology, 1988.

[11] A. O. Hero, "Secure space-time communication," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3235–3249, Dec. 2003.