Limits of Reliable Communication with Low Probability of Detection on AWGN Channels

Boulat A. Bash, Dennis Goeckel, Don Towsley

Abstract

We present a square root limit on the information rate with low probability of detection (LPD) over additive white Gaussian noise (AWGN) channels. Specifically, if the transmitter has AWGN channels to a receiver and a warden, both with non-zero noise power, we prove that $o(\sqrt{n})$ bits can be sent from the transmitter to the receiver in n channel uses while lower-bounding the warden's probability of detection error by $1-\epsilon$ for any $\epsilon > 0$. Moreover, in most practical scenarios, a lower bound on the noise power on the warden's channel to the transmitter is known and $\mathcal{O}(\sqrt{n})$ bits can be covertly sent in n channel uses. Conversely, attempting 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 $n \to \infty$.

I. Introduction

Securing information transmitted over wireless links is of paramount concern for consumer, industrial, and military applications. The taxonomy of network security classifies secure communication into two distinct categories: *low probability of intercept* (LPI) communication and *low probability of detection* (LPD) communication [1]. In recent years, the wireless networking community has made tremendous strides in the former area, securing data transmitted in wireless networks from interception by an untrusted eavesdropper using various encryption and key exchange protocols. However, the latter area, LPD communication, which concerns the prevention of transmissions from being *detected* in the first place, has been relatively underexplored.

Consider a node that tries to send data on a wireless channel to another node so that the *presence* of this transmission is not detected by an eavesdropping third party. There are many real-life scenarios where this is preferable to standard cryptographic security. 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. Thus, the study of covert communications over LPD channels is extremely important.

We examine the fundamental limits of covert communication over wireless channels subject to additive white Gaussian noise (AWGN). In our scenario, Alice communicates with Bob over an AWGN channel, while passive eavesdropper Warden Willie attempts to detect her transmission. The channel between Willie and Alice is also subject to AWGN and Willie is passive in that he does not actively jam Alice's channel. 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 he detects covert communication, Willie can potentially shut the channel down or otherwise punish Alice. If the noise on the channel between Willie and Alice has non-zero power, Alice

B. A. Bash and D. Towsley are with the Computer Science Department, University of Massachusetts, Amherst, Massachusetts.
D. Goeckel is with the Electrical and Computer Engineering Department, University of Massachusetts, Amherst, Massachusetts.

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.

can communicate with Bob while tolerating a certain probability of detection, which she can drive down by transmitting with low enough power. Thus, Alice potentially transmits non-zero mutual information across the covert channel to Bob in n uses of the channel.

Our problem is related to steganography, which considers hiding information by altering the properties of fixed-size, finite-alphabet covertext objects (such as images or software binary code) with imperfect steganography systems allowing some fixed probability of detection of hidden information. The square root law of steganography in the passive warden environment states that $\mathcal{O}(\sqrt{n})$ symbols in the covertext of size n may safely be modified to hide an $\mathcal{O}(\sqrt{n}\log n)$-bit steganographic message [2, Ch. 13], where the $\log n$ factor stems directly from the fact that transmission to Bob is noiseless [2, Ch. 8]. In our scenario, Alice uses the noise on her channel to Willie instead of the statistical properties of the covertext to hide information. However, having to code against the noise on her channel to Bob allows only $\mathcal{O}(\sqrt{n})$ bits to be covertly sent in n uses of the LPD channel.[1] The mathematics of statistical hypothesis testing yield a square root law in both problems, but as answers to different questions due to the fundamental differences in the communication channels. This relationship is discussed further at the end of Section III.

We state our main result that limits the capacity of the covert channel between Alice and Bob using asymptotic notation where $f(n) = \mathcal{O}(g(n))$ denotes an asymptotically tight upper bound on $f(n)$ (i.e. there exist constants $m, n_0 > 0$ such that $0 \leq f(n) \leq mg(n)$ for all $n \geq n_0$), $f(n) = o(g(n))$ denotes an upper bound on $f(n)$ that is not asymptotically tight (i.e. for any constant $m > 0$, there exists constant $n_0 > 0$ such that $0 \leq f(n) < mg(n)$ for all $n \geq n_0$), and $f(n) = \omega(g(n))$ denotes a lower bound on $f(n)$ that is not asymptotically tight (i.e. for any constant $m > 0$, there exists constant $n_0 > 0$ such that $0 \leq mg(n) < f(n)$ for all $n \geq n_0$) [3, Ch. 3.1]:

Theorem (Square root law)**.** *Suppose the channel between Alice and each of Bob and Willie experiences 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 reliably send $o(\sqrt{n})$ information bits to Bob in n channel uses while lower-bounding Willie's probability of detection error by $1 - \epsilon$. 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 the same error bound. 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).*

After introducing our system framework and hypothesis testing background in Section II, we prove the achievability of the square root law in Section III. We then prove the converse in Section IV. We discuss the mapping to the continuous-time channel and the relationship to previous work in Section V, and conclude in Section VI.

II. Prerequisites

A. System Framework

Alice and Bob construct a covert communications system, with all the details known to Willie except for a secret key that is shared before communication. This follows "best practices" in security system design as the security of our system depends only on the key [4]. Note that, if

[1]The capacity of a *noiseless* LPD channel between Alice and Bob would be infinite due to it being continuously-valued, and a noiseless channel between Alice and Willie would preclude the existence of the LPD channel between Alice and Bob.

information-theoretic *secrecy* (LPI communication) was desired, a sufficiently long key trivially provides such through the employment of a one-time pad [5], but this is not sufficient for LPD communication.

We use the discrete-time AWGN channel model with real-valued symbols (and defer discussion of the mapping to a continuous-time channel to Section V). Our formal system framework is depicted in Figure 1. 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. $z_i^{(w)} \sim \mathcal{N}(0, \sigma_w^2)$. Willie uses statistical hypothesis tests on $\mathbf{y}_w$ to determine whether Alice has communicated, which we discuss next.

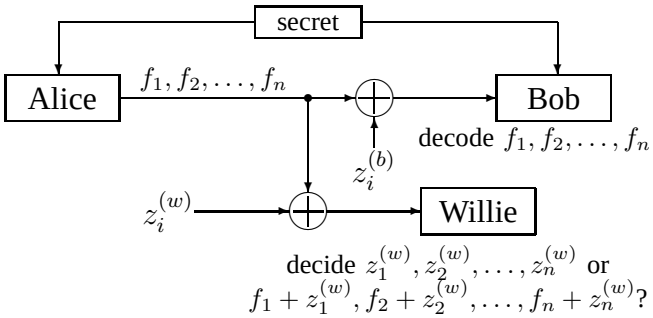

Fig. 1. System framework: Alice and Bob share a secret before the transmission. Alice encodes information into a vector of real symbols $\mathbf{f} = \{f_i\}_{i=1}^n$ and transmits it on an AWGN channel to Bob, while Willie attempts to classify his vector of observations of the channel from Alice $\mathbf{y}_w$ 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.

B. Hypothesis Testing

Willie expects vector $\mathbf{y}_w$ of n channel readings 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. In this case each sample is i.i.d. $y_i^{(w)} \sim \mathcal{N}(0, \sigma_w^2)$. The alternate hypothesis H_1 is that Alice is transmitting, which corresponds to samples $y_i^{(w)}$ coming from a different distribution. Willie can tolerate some false positives, or cases when his statistical test incorrectly accuses Alice. This rejection of H_0 when it is true is known as the type I error (or false alarm), and, following the standard nomenclature, we denote its probability by α [6]. Willie's test may also miss Alice's covert transmissions. Acceptance of H_0 when it is false is known as the type II error (or missed detection), and we denote its probability by β. The lower bound on the sum $\alpha + \beta$ characterizes the necessary trade-off between the false alarms and the missed detections in the design of a hypothesis test.

III. Achievability of Square Root Law

Willie's objective is 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 Alice does not transmit (i.e. when H_0 is true) as $\mathbf{P}_0$, and the probability distribution of the observations when Alice transmits (i.e. 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 $\alpha + \beta$ [6, Ch. 13]. The following holds for such a test:

Fact 1 (Theorem 13.1.1 in [6])**.** *For the optimal test,*

$$\alpha + \beta \;=\; 1 - \mathbb{V}_T(\mathbf{P}_0, \mathbf{P}_1)$$

where $\mathbb{V}_T(\mathbf{P}_0, \mathbf{P}_1)$ is the total variation distance between $\mathbf{P}_0$ and $\mathbf{P}_1$ defined as follows:

Definition 1 (Total variation distance [6])**.** *The* total variation distance *between two continuous probability measures* $\mathbf{P}_0$ *and* $\mathbf{P}_1$ *is*

$$\mathbb{V}_T(\mathbf{P}_0, \mathbf{P}_1) \;=\; \frac{1}{2}\|p_0(x) - p_1(x)\|_1 \tag{1}$$

where $p_0(x)$ *and* $p_1(x)$ *are densities of* $\mathbf{P}_0$ *and* $\mathbf{P}_1$*, respectively, and* $\|a - b\|_1$ *is the* $\mathcal{L}_1$ *norm.*

Since total variation lower-bounds the error of all hypothesis tests Willie can use, a clever choice of $\mathbf{f}$ allows Alice to limit Willie's detector performance. 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 Pinsker's Inequality:

Fact 2 (Pinsker's Inequality (Lemma 11.6.1 in [7]))**.**

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

where relative entropy $\mathcal{D}(\mathbf{P}_0 \| \mathbf{P}_1)$ is defined as follows:

Definition 2. *The* relative entropy *(also known as* Kullback-Leibler divergence*) between two probability measures* $\mathbf{P}_0$ *and* $\mathbf{P}_1$ *is:*

$$\mathcal{D}(\mathbf{P}_0 \| \mathbf{P}_1) \;=\; \int_{\mathcal{X}} p_0(x) \ln \frac{p_0(x)}{p_1(x)} dx \tag{2}$$

where $\mathcal{X}$ *is 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:

Fact 3 (Relative Entropy Product)**.** *From the chain rule for relative entropy [7, Eq. (2.67)]:*

$$\mathcal{D}(\mathbf{P}_0^n \| \mathbf{P}_1^n) \;=\; n\mathcal{D}(\mathbf{P}_0 \| \mathbf{P}_1)$$

Relative entropy is related to hypothesis testing via the Chernoff-Stein Lemma [7, Ch. 11.8] as an exponent in the expression for α given β and visa-versa, and can be used to analyze the hypothesis test performance, as is commonly done by the steganography community [2], [8]. However, lower-bounding $\alpha + \beta$ has a natural signal processing interpretation via the receiver operating characteristic (ROC) curve [9, Ch. 2.2.2], which plots probability of detection $1 - \beta$ versus α. Since $1 - \beta \geq \alpha$ and $\alpha + \beta \geq 1 - \epsilon$, small ϵ implies that the ROC curve lies very close to the line of no-discrimination (the diagonal line where $1 - \beta = \alpha$) over the entire domain of α because $\alpha + \epsilon \geq 1 - \beta \geq \alpha$. We thus state the achievability theorem under an average power constraint as follows:

Theorem 1.1 (Achievability)**.** *Suppose Willie's channel is subject to AWGN with average 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 reliably 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 over 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 length 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 $\{W_k\}_{k=1}^{2^{nR}}$, 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 used only to send a single message and is the secret that is not revealed to Willie, though he knows how it is constructed, including the value of P_f. The length of this secret is discussed in the remark following the proof of Theorem 1.2.

The channel between Alice and Willie is corrupted by AWGN with power σ_w^2. Willie applies statistical hypothesis testing on a vector of n channel readings $\mathbf{y}_w$ to decide whether Alice transmitted. Next we show how Alice can limit the performance of Willie's methods.

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. His null hypothesis H_0 is that Alice did not transmit and 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 $\alpha + \beta = 1 - \mathbb{V}_T(\mathbf{P}_0, \mathbf{P}_1)$, 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 the sum of the error probabilities by upper-bounding the total variation distance: $\mathbb{V}_T(\mathbf{P}_0, \mathbf{P}_1) \leq \epsilon$.

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 $\mathbf{P}_w = \mathcal{N}(0, \sigma_w^2)$. Recall that Willie does not know the codebook. Therefore, 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 vector $\mathbf{y}_w$, where $y_i^{(w)} \sim \mathcal{N}(0, P_f + \sigma_w^2) = \mathbf{P}_s$ is i.i.d. , and thus, $\mathbf{P}_1 = \mathbf{P}_s^n$. By Facts 2 and 3:

$$\mathbb{V}_T(\mathbf{P}_w^n, \mathbf{P}_s^n) \leq \sqrt{\frac{1}{2}\mathcal{D}(\mathbf{P}_w^n \| \mathbf{P}_s^n)} = \sqrt{\frac{n}{2}\mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s)}$$

The relative entropy follows as:

$$\mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s) \;=\; \frac{1}{2}\left[\ln\left(1 + \frac{P_f}{\sigma_w^2}\right) - \frac{P_f}{P_f + \sigma_w^2}\right]$$

While the expression for $\mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s)$ has a closed form, its Taylor series expansion with respect to P_f around $P_f = 0$ is more useful. While the zeroth and first order terms are zero, the second order term is:

$$\frac{P_f^2}{2!} \times \left.\frac{\partial^2 \mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s)}{\partial P_f^2}\right|_{P_f=0} \;=\; \frac{P_f^2}{4\sigma_w^4}$$

Relative entropy being locally quadratic is well-known [10, Ch. 2.6]; in fact $\left.\frac{\partial^2 \mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s)}{\partial P_f^2}\right|_{P_f=0} = \frac{1}{2\sigma_w^4}$ is the Fisher information that an observation of noise carries about its power. Now, the third order term is:

$$\frac{P_f^3}{3!} \times \left.\frac{\partial^3 \mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s)}{\partial P_f^3}\right|_{P_f=0} \;=\; -\frac{P_f^3}{3\sigma_w^6} \leq 0$$

If $P_f < \sigma_w^2$, then the Taylor series converges and we can apply Taylor's Theorem to upper-bound relative entropy with the second order term. The upper bound we seek is:

$$\mathbb{V}_T(\mathbf{P}_w^n, \mathbf{P}_s^n) \;\leq\; \frac{P_f}{2\sigma_w^2}\sqrt{\frac{n}{2}} \tag{3}$$

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 $f(n) = \hat{\sigma}_w^2$ (a conservative lower bound is the thermal noise power of the best receiver currently available). 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 n large enough, $P_f < \sigma_w^2$ satisfies the Taylor series convergence criterion, and Alice obtains the upper bound $\mathbb{V}_T(\mathbf{P}_w^n, \mathbf{P}_s^n) \leq \epsilon$, limiting the performance of Willie's detector.

Since Alice's symbol power P_f is a decreasing function of the codeword length n, the standard channel coding results for constant power (and constant rate) do not directly apply. Thus, 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 $\mathbf{c}(W_i)$, $i \neq k$. The decoding error probability, averaged over all codebooks, is then:

$$\begin{aligned}
\mathbf{P}_e &= \mathbb{E}_{\mathbf{c}(W_k)}\left[\mathbf{P}\left(\cup_{i=0,i\neq k}^{2^{nR}} E_i(\mathbf{c}(W_k))\right)\right] \\
&\leq \mathbb{E}_{\mathbf{c}(W_k)}\left[\sum_{i=0,i\neq k}^{2^{nR}} \mathbf{P}\left(E_i(\mathbf{c}(W_k))\right)\right] \qquad (4)\\
&= \sum_{i=0,i\neq k}^{2^{nR}} \mathbb{E}_{\mathbf{c}(W_k)}\left[\mathbf{P}\left(E_i(\mathbf{c}(W_k))\right)\right] \qquad (5)
\end{aligned}$$

where $\mathbb{E}_X[\cdot]$ denotes the expectation operator over random variable X and (4) follows from the union bound. Let $\mathbf{d} = \mathbf{c}(W_k) - \mathbf{c}(W_i)$. Then $\|\mathbf{d}\|_2$ is the distance between two codewords, where $\|\cdot\|_2$ is the $\mathcal{L}_2$ norm. Since codewords are independent and Gaussian, $d_j \sim \mathcal{N}(0, 2P_f)$ for $j = 1, 2, \ldots, n$ and $\|\mathbf{d}\|_2^2 = 2P_f U$, where $U \sim \chi_n^2$, with χ_n^2 denoting the chi-squared distribution with n degrees of freedom. Therefore, by [11, Eq. (3.44)]:

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

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$. Since $Q(x) \leq \frac{1}{2}e^{-x^2/2}$ [12, Eq. (5)] and $P_f = \frac{cf(n)}{\sqrt{n}}$:

$$\begin{aligned}
\mathbb{E}_U\left[Q\left(\sqrt{\frac{P_f U}{2\sigma_b^2}}\right)\right] &\leq \mathbb{E}_U\left[\exp\left(-\frac{cf(n)U}{4\sqrt{n}\sigma_b^2}\right)\right] \\
&= \int_0^\infty \frac{e^{-\frac{cf(n)u}{4\sqrt{n}\sigma_b^2}-\frac{u}{2}} 2^{-\frac{n}{2}} u^{\frac{n}{2}-1}}{\Gamma(n/2)} du \\
&= 2^{-n/2}\left(\frac{1}{2} + \frac{cf(n)}{4\sqrt{n}\sigma_b^2}\right)^{-n/2} \qquad (6)
\end{aligned}$$

where (6) is from the substitution $v = u\left(\frac{1}{2} + \frac{cf(n)}{4\sqrt{n}\sigma_b^2}\right)$ and the definition of the Gamma function $\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}dx$. Since $\frac{1}{2} + \frac{cf(n)}{4\sqrt{n}\sigma_b^2} = 2^{\log_2\left(\frac{1}{2}+\frac{cf(n)}{4\sqrt{n}\sigma_b^2}\right)}$:

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

Hence, the summand in (5) does not depend on i, and (5) becomes:

$$\mathbf{P}_e \;\leq\; 2^{nR-\frac{n}{2}\log_2\left(1+\frac{cf(n)}{2\sqrt{n}\sigma_b^2}\right)} \tag{7}$$

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 averaged over all codebooks 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 $nR \leq \frac{\sqrt{n}\rho cf(n)}{4\sigma_b^2\ln 2}$, approaching equality as n gets large, Bob receives $o(\sqrt{n})$ bits in n channel uses, and $\mathcal{O}(\sqrt{n})$ bits in n channel uses if $f(n) = \hat{\sigma}_w^2$. ■

Unlike Shannon's coding theorem for AWGN channels [7, Theorem 9.1.1], we cannot select a codebook that performs better than average, as that would violate the i.i.d. condition needed to limit Willie's detection ability. If such a codebook is desirable, the construction of Theorem 1.2 can be employed using the modification given by the remark following its proof. This construction also satisfies both the peak and the average power constraints, as demonstrated below.

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 reliably 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 length 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, this single-use codebook is not revealed to Willie, though he knows how it is constructed, including the value of a. While here the entire codebook is secretly shared between Alice and Bob, in the remark following the proof we discuss how to reduce the amount of shared secret information.

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, $\mathbf{P}_s = \frac{1}{2}\left(\mathcal{N}(-a, \sigma_w^2) + \mathcal{N}(a, \sigma_w^2)\right)$, and, with $\mathbf{P}_w = \mathcal{N}(0, \sigma_w^2)$, we have:

$$\mathcal{D}(\mathbf{P}_w\|\mathbf{P}_s) = \int_{-\infty}^{\infty} \frac{e^{-\frac{x^2}{2\sigma_w^2}}}{\sqrt{2\pi}\sigma_w} \ln \frac{e^{-\frac{x^2}{2\sigma_w^2}}}{\frac{1}{2}\left(e^{-\frac{(x+a)^2}{2\sigma_w^2}} + e^{-\frac{(x-a)^2}{2\sigma_w^2}}\right)} dx$$

There is no closed-form expression for $\mathcal{D}(\mathbf{P}_w\|\mathbf{P}_s)$, but it can be expanded using the Taylor series with respect to a around $a = 0$. While the zeroth through third order terms are zero, the fourth order term is:

$$\frac{a^4}{4!} \times \left.\frac{\partial^4 \mathcal{D}(\mathbf{P}_w\|\mathbf{P}_s)}{\partial a^4}\right|_{a=0} \;=\; \frac{a^4}{4\sigma_w^4}$$

While the fifth order term is zero, for the sixth order term we obtain:

$$\frac{a^6}{6!} \times \left. \frac{\partial^6 \mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s)}{\partial a^6} \right|_{a=0} = -\frac{a^6}{3\sigma_w^6} < 0$$

If $a < \sigma_w$, then the Taylor series converges and we can apply Taylor's Theorem and upper-bound relative entropy with the fourth order term. The upper bound we seek is:

$$\mathbb{V}_T(\mathbf{P}_w^n, \mathbf{P}_s^n) \leq \frac{a^2}{2\sigma_w^2}\sqrt{\frac{n}{2}} \tag{8}$$

Since the power of Alice's covert symbol is $a^2 = P_f$, (8) is identical to (3) 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 $\mathbb{V}_T(\mathbf{P}_w^n, \mathbf{P}_s^n) \leq \epsilon$, limiting the performance of Willie's detector.

As in Theorem 1.1, we cannot directly apply the standard constant-power channel coding results to our system where the symbol power is a decreasing function of the codeword length. 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 $\mathbf{P}_e \leq 2^{nR - n(1-\mathcal{H}(p_e))}$ [13], where $\mathcal{H}(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy function. We expand the analysis in [14, Section I.2.1] to characterize the rate R. The Taylor series of $e^{-t^2/2}$ alternates, and the Taylor series expansion of $p_e = Q\left(\frac{a}{\sigma_b}\right) = \frac{1}{2} - \int_0^{\frac{a}{\sigma_b}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$ with respect to a around $a = 0$ (which converges since a is small for large n) yields an upper bound: $p_e \leq \frac{1}{2} - \frac{1}{\sqrt{2\pi}}\left(\frac{a}{\sigma_b} - \frac{a^3}{6\sigma_b^3}\right) \triangleq p_e^{(UB)}$. Since $\mathcal{H}(p)$ is a monotonically increasing function on the interval $\left[0, \frac{1}{2}\right]$, $\mathcal{H}(p_e) \leq \mathcal{H}(p_e^{(UB)})$. The odd terms of the Taylor series expansion of $\mathcal{H}(p_e^{(UB)})$ with respect to a around $a = 0$ are zero, and, thus, $\mathcal{H}(p_e^{(UB)}) = 1 - \frac{a^2}{\sigma_b^2 \pi \ln 2} + \mathcal{O}(a^4)$. Since $a^2 = \frac{cf(n)}{\sqrt{n}}$, $\mathbf{P}_e \leq 2^{nR - \frac{\sqrt{n} cf(n)}{\sigma_b^2 \pi \ln 2} + \mathcal{O}(1)}$. Since $f(n) = \omega(1/\sqrt{n})$, 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 averaged over all codebooks decays exponentially to zero as n increases and Bob obtains $nR = o(\sqrt{n})$ bits in n channel uses, and $\mathcal{O}(\sqrt{n})$ bits in n channel uses if $f(n) = \hat{\sigma}_w^2$. ■

Remarks

Employing the best codebook: Following the standard argument [7, p. 204], there must be at least one codebook that performs at least as well as the average. Consider this "best" binary codebook, but now assume that it is public (i.e. known to Willie). Theorem 1.2 shows that Alice can use it to transmit $\mathcal{O}(\sqrt{n})$ bits to Bob in n channel uses with exponentially-decaying probability of error. However, since the codebook is public, Willie can use it to detect Alice's transmissions by performing the same decoding as Bob.

Now, suppose that, prior to communication, Alice and Bob generate and share binary vector $\mathbf{k}$ where $p_{\mathbf{K}}(\mathbf{k}) = \prod_{i=1}^{n} p_K(k_i)$ with $p_K(0) = p_K(1) = \frac{1}{2}$. Alice XORs $\mathbf{k}$ and the binary representation of the codeword $\mathbf{c}(W_k)$, resulting in an equiprobable transmission of $-a$ and a when Alice transmits a covert symbol during the i^{th} symbol period. Provided $\mathbf{k}$ is never re-used and is kept secret from Willie, the i.i.d. assumption for the vector $\mathbf{y}_w$ in Theorem 1.2 holds

without the need to exchange an entire secret codebook between Alice and Bob. Bob decodes by XORing $\mathbf{k}$ with the output of the hard-decision device prior to applying the ML decoder.

While the square root law implies that the shared secret here is quadratic in the length $M = \mathcal{O}(\sqrt{n})$ of a message, we can construct a coding scheme that, on average, requires an $\mathcal{O}(\sqrt{n}\log n)$-bit secret in two stages. First, Alice and Bob randomly select the symbol periods that they will use for their transmission by flipping a biased coin n times and selecting the i^{th} symbol period with probability τ. Denote the number of selected symbol periods by η and note that $\mathbb{E}[\eta] = \tau n$. Alice and Bob then use the best public binary codebook with codewords of length η on these selected η symbol periods. They also generate and share a random binary vector $\mathbf{k}$ where $p_{\mathbf{K}}(\mathbf{k}) = \prod_{i=1}^{\eta} p_K(k_i)$ with $p_K(0) = p_K(1) = \frac{1}{2}$. Alice XORs $\mathbf{k}$ and the binary representation of the codeword $\mathbf{c}(W_k)$. The symbol location selection is independent of both the symbol and the channel noise. When Alice is transmitting a codeword, the distribution of each of Willie's observations is $\mathbf{P}_s = (1-\tau)\mathcal{N}(0,\sigma_w^2) + \frac{\tau}{2}\left(\mathcal{N}(-a,\sigma_w^2) + \mathcal{N}(a,\sigma_w^2)\right)$ and, thus,

$$\mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s) \;=\; \int_{-\infty}^{\infty} \frac{e^{-\frac{x^2}{2\sigma_w^2}}}{\sqrt{2\pi}\sigma_w} \ln \frac{e^{-\frac{x^2}{2\sigma_w^2}}/\sqrt{2\pi\sigma_w^2}}{\frac{(1-\tau)e^{-\frac{x^2}{2\sigma_w^2}}}{\sqrt{2\pi}\sigma_w} + \frac{\tau}{2}\left(\frac{e^{-\frac{(x+a)^2}{2\sigma_w^2}}}{\sqrt{2\pi\sigma_w^2}} + \frac{e^{-\frac{(x-a)^2}{2\sigma_w^2}}}{\sqrt{2\pi\sigma_w^2}}\right)} dx$$

There is no closed-form expression for $\mathcal{D}(\mathbf{P}_w \| \mathbf{P}_s)$, but a Taylor series expansion with respect to a around $a = 0$ yields the following bound:

$$\mathbb{V}_T(\mathbf{P}_w^n, \mathbf{P}_s^n) \;\leq\; \frac{\tau a^2}{2\sigma_w^2}\sqrt{\frac{n}{2}} \tag{9}$$

The only difference in (9) from (8) is τ in the numerator. Thus, if Alice sets the product $\tau a^2 \leq \frac{cf(n)}{\sqrt{n}}$, 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. Now let's fix a and, thus, set $\tau = \mathcal{O}(1/\sqrt{n})$. Since, on average, τn symbol periods are selected, it takes (again, on average) $\mathcal{O}(\sqrt{n})$ positive integers to enumerate the selected symbols. There are n total symbols, and, thus, it takes at most $\log(n)$ bits to represent each selected symbol location and $\mathcal{O}(\sqrt{n}\log n)$ bits to represent all the locations of selected symbols. Also, the average length of key $\mathbf{k}$ is $\mathcal{O}(\sqrt{n})$ bits. Thus, on average, Alice and Bob need to exchange $\mathcal{O}(\sqrt{n}\log n)$ bits under this coding scheme. The possibility of LPD communication with a key linear to the message length and a detailed consideration of the key length in general is an open problem that we defer to the future work.

Relationship with Square Root Law in Steganography: The LPD communication problem is related to the problem of steganography. A comprehensive review of steganography is available in a book by Fridrich [2]. In finite-alphabet imperfect steganographic systems at most $\mathcal{O}(\sqrt{n})$ symbols in the original covertext of length n may safely be modified to hide a steganographic message of length $\mathcal{O}(\sqrt{n}\log n)$ bits [2, Ch. 13] [15]. This result was extended to Markov covertext [16] and was shown to either require a key linear to the length of the message [17] or encryption of the message prior to embedding [18].

The square root law in steganography has the same form as our square root law because both laws follow from the relative entropy being locally quadratic [10, Ch. 2.6]:

$$\mathcal{D}(\mathbf{P}_0 \| \mathbf{P}_1) \;=\; \frac{\delta^2}{2}\mathcal{J}(\theta) + \mathcal{O}(\delta^3)$$

where $\mathcal{J}(\theta) = \int_{\mathcal{X}} \left(\frac{\partial}{\partial\theta}\ln f(x;\theta)\right)^2 f(x;\theta)dx$ is the Fisher information associated with parameter θ, and $\mathbf{P}_0$ and $\mathbf{P}_1$ are probability measures with density functions from the same family over

the support $\mathcal{X}$, but with parameters differing by δ: $p_0(x) = f(x;\theta)$ and $p_1(x) = f(x;\theta+\delta)$. Fisher information is thus used as a metric for steganographic security [19], [20].

In a typical steganography scenario with a passive warden, coding techniques similar to Hamming codes allow embedding of $\log(n)$ bits per changed symbol [2, Ch. 8], which make hiding $\mathcal{O}(\sqrt{n}\log n)$ bits in n symbols possible. However, due to the noise on the channel between Alice and Bob, and the resultant need for error correction, the LPD channel only allows $\mathcal{O}(\sqrt{n})$ bits to be transmitted in n channel uses, as we prove in the following section.

IV. Converse

Here, as in the proof of 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. Alice encodes each message W_k arbitrarily into n symbols at the rate $R = M/n$ symbols/bit. For an upper bound on the reduction in entropy, the messages are chosen equiprobably.

Willie observes all n of Alice's channel uses. To strengthen the converse, he is oblivious to her signal properties. Nevertheless, even with Willie's knowledge limited, Alice cannot transmit a message with $\omega(\sqrt{n})$ bits of information in n channel uses without either being detected by Willie or having Bob suffer a non-zero decoding error.

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: Suppose Alice employs an arbitrary codebook $\{\mathbf{c}(W_k), k = 1, 2, \ldots, 2^{nR}\}$. 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. First, we show how Willie can bound the errors α and β of this test as a function of Alice's signal parameters. Then we show that if Alice's codebook prevents Willie's test from detecting her, Bob cannot decode her transmissions without error.

To perform the test, Willie collects a row 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 on his channel. Thus, $y_i^{(w)} \sim \mathcal{N}(0, \sigma_w^2)$, and the mean and the variance of S when H_0 is true are:

$$\mathbb{E}[S] = \sigma_w^2 \tag{10}$$

$$\mathrm{Var}[S] = \frac{2\sigma_w^4}{n} \tag{11}$$

Suppose Alice transmits codeword $\mathbf{c}(W_k) = \{f_i^{(k)}\}_{i=1}^n$. Then Willie's vector of observations $\mathbf{y}_{w,k} = \{y_i^{(w,k)}\}_{i=1}^n$ contains readings of mean-shifted noise $y_i^{(w,k)} \sim \mathcal{N}(f_i^{(k)}, \sigma_w^2)$. The mean of each squared observation is $\mathbb{E}[y_i^2] = \sigma_w^2 + \left(f_i^{(k)}\right)^2$ and the variance is $\mathrm{Var}[y_i^2] = \mathbb{E}[y_i^4] -$

$\left(\mathbb{E}\left[y_i^2\right]\right)^2 = 4\left(f_i^{(k)}\right)^2\sigma_w^2 + 2\sigma_w^4$. Denote the average symbol power of codeword $\mathbf{c}(W_k)$ by $P_k = \frac{\mathbf{c}(W_k)\mathbf{c}^T(W_k)}{n}$. Then the mean and variance of S when Alice transmits codeword $\mathbf{c}(W_k)$ are:

$$\mathbb{E}\left[S\right] = \sigma_w^2 + P_k \tag{12}$$

$$\mathrm{Var}\left[S\right] = \frac{4P_k\sigma_w^2 + 2\sigma_w^4}{n} \tag{13}$$

The variance of Willie's test statistic (13) is computed by adding the variances conditioned on $\mathbf{c}(W_k)$ of the squared individual observations $\mathrm{Var}\left[y_i^2\right]$ (and dividing by n^2) since the noise on the individual observations is independent.

The probability distribution for the vector of Willie's observations depends on which hypothesis is true. 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. While $\mathbf{P}_1^{(k)}$ is conditioned on Alice's codeword, we show that the average symbol power $P_k = \frac{\mathbf{c}(W_k)\mathbf{c}^T(W_k)}{n}$ of the codeword $\mathbf{c}(W_k)$ determines its detectability by this detector, and that our result applies to all codewords with power of the same order.

If H_0 is true, then S should be close to (10). 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 (10) and (11) with Chebyshev's Inequality [7, Eq. (3.32)]:

$$\begin{aligned}\alpha &= \mathbf{P}_0\left(S \geq \sigma_w^2 + t\right)\\ &\leq \mathbf{P}_0\left(|S - \sigma_w^2| \geq t\right)\\ &\leq \frac{2\sigma_w^4}{nt^2}\end{aligned}$$

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.

Suppose Alice transmits codeword $\mathbf{c}(W_k)$. Then the probability of a miss $\beta^{(k)}$ given t is the probability that $S < \sigma_w^2 + t$, which we bound using (12) and (13) with Chebyshev's Inequality:

$$\begin{aligned}\beta^{(k)} &= \mathbf{P}_1^{(k)}\left(S < \sigma_w^2 + t\right)\\ &\leq \mathbf{P}_1^{(k)}\left(\left|S - \sigma_w^2 - P_k\right| \geq P_k - t\right)\\ &\leq \frac{4P_k\sigma_w^2 + 2\sigma_w^4}{(\sqrt{n}P_k - d)^2}\end{aligned} \tag{14}$$

If $P_k = \omega(1/\sqrt{n})$, $\lim_{n\to\infty}\beta^{(k)} = 0$. Thus, with enough observations, Willie can detect with arbitrarily low error probability Alice's codewords with average symbol power $P_k = \frac{\mathbf{c}(W_k)\mathbf{c}^T(W_k)}{n} = \omega(1/\sqrt{n})$. Note that Willie's detector is oblivious to any details of Alice's codebook construction.

By (14), if Alice desires to lower-bound the sum of the probabilities of error of Willie's statistical test by $\alpha + \beta \geq \zeta > 0$, her codebook must contain low-power codewords; in particular, a fraction $\gamma > 0$ of the codewords must have symbol power $P_{\mathcal{U}} = \mathcal{O}(1/\sqrt{n})$. Let's denote this set of codewords as $\mathcal{U}$ and examine the probability of Bob's decoding error $\mathbf{P}_e$. The probability that a message from set $\mathcal{U}$ is sent is $\mathbf{P}\left(\mathcal{U}\right) = \gamma$, as all messages are equiprobable. We bound $\mathbf{P}_e = \mathbf{P}_e\left(\mathcal{U}\right)\mathbf{P}\left(\mathcal{U}\right) + \mathbf{P}_e\left(\overline{\mathcal{U}}\right)\mathbf{P}\left(\overline{\mathcal{U}}\right) \geq \gamma\mathbf{P}_e\left(\mathcal{U}\right)$, where $\overline{\mathcal{U}}$ is the complement of $\mathcal{U}$ and $\mathbf{P}_e\left(\mathcal{U}\right)$ is the probability of decoding error when a message from $\mathcal{U}$ is sent:

$$\mathbf{P}_e\left(\mathcal{U}\right) = \frac{1}{|\mathcal{U}|}\sum_{W\in\mathcal{U}}\mathbf{P}_e\left(\mathbf{c}(W)\text{ sent}\right) \tag{15}$$

where $\mathbf{P}_e\left(\mathbf{c}(W) \text{ sent}\right)$ is the probability of error when codeword $\mathbf{c}(W)$ is transmitted, $|\cdot|$ denotes the set cardinality operator, and (15) holds because all messages are equiprobable.

When Bob uses the optimal decoder, $\mathbf{P}_e\left(\mathbf{c}(W) \text{ sent}\right)$ is the probability that Bob decodes the received signal as $\hat{W} \neq W$. This is the probability of a union of events E_j, where E_j is the event that sent message W is decoded as some other message $W_j \neq W$:

$$\begin{aligned} \mathbf{P}_e\left(\mathbf{c}(W) \text{ sent}\right) &= \mathbf{P}\left(\cup_{j=1, W_j \neq W}^{2^{nR}} E_j\right) \\ &\geq \mathbf{P}\left(\cup_{W_j \in \mathcal{U}\setminus\{W\}} E_j\right) \triangleq \mathbf{P}_e^{(\mathcal{U})} \end{aligned} \tag{16}$$

where the inequality in (16) is due to the observation that the sets in the second union are contained in the first. From the decoder perspective, this is due to the decrease in the decoding error probability if Bob knew that the message came from $\mathcal{U}$ (reducing the set of messages on which the decoder can err).

Our analysis of $\mathbf{P}_e^{(\mathcal{U})}$ uses Cover's simplification of Fano's inequality similar to the proof of the converse to the coding theorem for Gaussian channels in [7, Ch. 9.2]. Since we are interested in $\mathbf{P}_e^{(\mathcal{U})}$, we do not absorb it into ϵ_n as done in (9.37) of [7]. Rather, we explicitly use:

$$H(W|\hat{W}) \leq 1 + (\log_2 |\mathcal{U}|)\mathbf{P}_e^{(\mathcal{U})} \tag{17}$$

where $H(W|\hat{W})$ denotes the entropy of message W conditioned on Bob's decoding $\hat{W}$ of W.

Noting that the size of the set $\mathcal{U}$ from which the messages are drawn is $\gamma 2^{nR}$ and that, since each message is equiprobable, the entropy of a message W from $\mathcal{U}$ is $H(W) = \log_2 |\mathcal{U}| = \log_2 \gamma + nR$, we utilize (17) and carry out steps (9.38)–(9.53) in [7] to obtain:

$$\mathbf{P}_e^{(\mathcal{U})} \geq 1 - \frac{P_{\mathcal{U}}/2\sigma_b^2 + 1/n}{\frac{\log_2 \gamma}{n} + R} \tag{18}$$

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. Since $\gamma > 0$, $\mathbf{P}_e$ is bounded away from zero if Alice tries to beat Willie's simple hypothesis test. ■

Goodput of Alice's Communication

Define the goodput $G(n)$ of Alice's communication as the average number of bits that Bob can receive from Alice over n channel uses with non-zero probability of a message being undetected as $n \to \infty$. Since only $\mathcal{U}$ contains such messages, by (18), the probability of her message being successfully decoded by Bob is $\mathbf{P}_s^{(\mathcal{U})} = 1 - \mathbf{P}_e^{(\mathcal{U})} = \mathcal{O}\left(\frac{1}{\sqrt{n}R}\right)$ and the goodput is $G(n) = \gamma \mathbf{P}_s^{(\mathcal{U})} Rn = \mathcal{O}(\sqrt{n})$. Thus, Alice cannot break the square root law using an arbitrarily high transmission rate and retransmissions while keeping the power below Willie's detection threshold.

V. Discussion

A. Mapping to Continuous-time Channel

We employ a discrete-time model throughout the paper. However, whereas this is a common assumption made without loss of generality in standard communication theory, it is important to consider whether some aspect of the LPD problem has been missed by starting in discrete-time.

Consider the standard communication system model, where Alice's (baseband) continuous-time waveform would be given in terms of her discrete-time transmitted sequence by:

$$x(t) = \sum_{i=1}^{n} f_i \; p(t - iT_s)$$

where T_s is the symbol period and $p(\cdot)$ is the pulse shaping waveform. Consider a (baseband) system bandwidth constraint of W Hz. Now, if Alice chooses $p(\cdot)$ ideally as $\mathrm{sinc}(2Wt)$, where $\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$, then the natural choice of $T_s = 1/2W$ results in no intersymbol interference (ISI). From the Nyquist sampling criterion, both Willie (and Bob) can extract all of the information from the signaling band by sampling at a rate of $2W$ samples/second, which then leads directly to the discrete-time model of Section II and suits our demonstration of the fundamental limits to Alice's covert channel capabilities. However, when $p(\cdot)$ is chosen in a more practical fashion, for example, as a raised cosine pulse with some excess bandwidth, then sampling at a rate higher than $2W$ has utility for signal detection even if the Nyquist ISI criterion is satisfied. In particular, techniques involving cyclostationary detection are now applicable, and we consider such a scenario a promising area for future work.

B. Relationship to Previous Work

The relationship of our work to steganography has already been discussed in the remark at the end of Section III. Here we relate our problem to other work in communication.

The LPD communication problem is related to that of establishing a cognitive radio (CR) network [21]. An aspect of the CR problem is limiting the interference from the secondary users' radios to the primary users of the network. The LPD problem with a passive warden can be cast within this framework by having primary users only listen [22]. However, the properties of the secondary signal that allow smooth operation of the primary network are very different from those of an undetectable signal. While there is a lot of work on the former topic, we are not aware of work by the CR community on the latter issue.

Analytical evaluation of LPD communication has been sparse. Hero studies LPI/LPD channels [1] in a multiple-input multiple-output (MIMO) setting. However, he focuses on the constraints (s.t. power, fourth moment, etc.) that the LPD communication over a MIMO channel should enforce given the kind of information the adversary possesses and on the signaling methods that maximize the throughput of the channel given those constraints. 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. It is notable that the LPI portion of his work has drawn significant attention, while the LPD portion has been largely overlooked.

VI. Conclusion

Practitioners have always known that LPD communication requires one to use low power in order to blend in with the noise on the eavesdropping warden's channel. However, the specific requirements for achieving LPD communication and resulting achievable performance have not been analyzed prior to this work. We quantified the conditions for existence and maintenance of an LPD channel by proving 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 $\mathcal{O}(\sqrt{n})$.

There are a number of avenues for future research. The key efficiency and, specifically, LPD communication with a key linear in message length is an open theoretical research problem.

Practical network settings and the implications of the square root law on the covert transmission of packets under additional constraints such as delay should be analyzed. The impact of dynamism in the network should also be examined, as well as more realistic scenarios that include channel artifacts such as fading and interference from other nodes. One may be able to improve LPD communication by employing nodes that perform friendly jamming. 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] A. O. Hero, "Secure space-time communication," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3235–3249, Dec. 2003.

[2] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*, 1st ed. New York, NY, USA: Cambridge University Press, 2009.

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

[4] 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.

[5] C. E. Shannon, "Communication theory of security," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

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

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

[8] C. Cachin, "An information-theoretic model for steganography," *Information and Computation*, vol. 192, no. 1, pp. 41–56, 2004.

[9] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. New York: John Wiley & Sons, Inc., 2001.

[10] S. Kullback, *Information Theory and Statistics*. New York, NY: Wiley, 1959.

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

[12] 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.

[13] 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.

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

[15] A. D. Ker, "A capacity result for batch steganography," *IEEE Signal Processing Letters*, vol. 14, no. 8, pp. 525–528, Aug. 2007.

[16] 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.

[17] A. D. Ker, "The square root law requires a linear key," in *Proceedings of the 11th ACM workshop on Multimedia and security*, ser. MM&Sec '09. New York, NY, USA: ACM, 2009, pp. 85–92.

[18] ——, "The square root law does not require a linear key," in *Proceedings of the 12th ACM workshop on Multimedia and security*, ser. MM&Sec '10. New York, NY, USA: ACM, 2010, pp. 213–224.

[19] T. Filler and J. Fridrich, "Fisher information determines capacity of ϵ-secure steganography," in *Information Hiding*, ser. Lecture Notes in Computer Science, S. Katzenbeisser and A.-R. Sadeghi, Eds. Springer Berlin / Heidelberg, 2009, vol. 5806, pp. 31–47.

[20] A. D. Ker, "Estimating steganographic fisher information in real images," in *Information Hiding*, ser. Lecture Notes in Computer Science, S. Katzenbeisser and A.-R. Sadeghi, Eds. Springer Berlin / Heidelberg, 2009, vol. 5806, pp. 73–88.

[21] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, First Qtr 2009.

[22] W. Ren, A. Swami, and Q. Zhao, "Coexistence, connectivity and delay in heterogeneous networks," in *Proceedings of the 27th Army Science Conference*, 2011.