Covert Communication with Finite Blocklength in AWGN Channels

Shihao Yan†, Biao He‡, Yirui Cong†, and Xiangyun Zhou†

†Research School of Engineering, The Australian National University, Canberra, ACT, Australia
‡Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong

Emails: {shihao.yan, yirui.cong, xiangyun.zhou}@anu.edu.au, eebiaohe@ust.hk

Abstract—Covert communication is to achieve a reliable transmission from a transmitter to a receiver while guaranteeing an arbitrarily small probability of this transmission being detected by a warden. In this work, we study the covert communication in AWGN channels with finite blocklength, in which the number of channel uses is finite. Specifically, we analytically prove that the entire block (all available channel uses) should be utilized to maximize the effective throughput of the transmission subject to a predetermined covert requirement. This is a nontrivial result because more channel uses results in more observations at the warden for detecting the transmission. We also determine the maximum allowable transmit power per channel use, which is shown to decrease as the blocklength increases. Despite the decrease in the maximum allowable transmit power per channel use, the maximum allowable total power over the entire block is proved to increase with the blocklength, which leads to the fact that the effective throughput increases with the blocklength.

I. INTRODUCTION

As people become more dependent on wireless devices to share private information, crucial concerns on the security and privacy of wireless communications are emerging since a large amount of confidential information (e.g., email/bank account information and password, credit card details) is transferred over wireless networks. In addition to the secrecy and integrity of the transmitted information, in some scenarios a user may wish to transmit messages over wireless networks without being detected. This is due to the fact that for example the exposure of this transmission may disclose the user’s location information, which probably violates the privacy of the user. Therefore, covert communication is attracting an increasing amount of research interests recently (e.g., [1–3]). In covert communication, a transmitter (Alice) intends to communicate with a legitimate receiver (Bob) without being detected by a warden (Willie), who is observing this communication.

In fact, covert communication was addressed by spread spectrum techniques in the early 20th century and a review on spread spectrum techniques can be found in [4]. However, the performance limit of covert communication has not been fully examined in the literature and recently attracts much research attention. Considering additive white Gaussian noise (AWGN) channels, a square root law has been derived in [5], which states that Alice can transmit no more than \( O\left(\sqrt{n}\right) / n \) average number of bits that can be covertly and reliably transmitted per channel use asymptotically approaches zero). However, in some scenarios a positive rate has been proved to be achievable (e.g., [7, 9–13]). For example, it is proved that a positive rate can be obtained when Willie has uncertainty about the receiver noise variance in AWGN channels [11, 13], when Willie does not exactly know the receiver noise model in BSC channels [7], or when Willie lacks knowledge of his channel characteristics in AWGN and block fading channels [12–14]. In addition to the noise or channel uncertainty, as proved in [10] a positive rate can also be achieved when Willie has uncertainty on the time instant of the communication.

In the square root law we have \( O(\sqrt{n})/n \to 0 \) as \( n \to \infty \), which states that the rate is asymptotically zero (i.e., the average number of bits that can be covertly and reliably transmitted per channel use asymptotically approaches zero). However, in some scenarios a positive rate has been proved to be achievable (e.g., [7, 9–13]). For example, it is proved that a positive rate can be obtained when Willie has uncertainty about the receiver noise variance in AWGN channels [11, 13], when Willie does not exactly know the receiver noise model in BSC channels [7], or when Willie lacks knowledge of his channel characteristics in AWGN and block fading channels [12–14]. In addition to the noise or channel uncertainty, as proved in [10] a positive rate can also be achieved when Willie has uncertainty on the time instant of the communication.

In the literature as seen in the aforementioned works, only [11] mentioned the impact of finite samples (i.e., finite \( n \)) on the detection performance at Willie. It is numerically shown that with noise uncertainty at Willie there may exist an optimal number of samples that maximizes the communication rate subject to \( \xi \geq 1 - \epsilon \), where \( \xi \) is the sum of false positive and miss detection rates at Willie and \( 0 < \epsilon \leq 1 \) is an arbitrarily small number. Besides the detection performance at Willie, finite \( n \) also has significant impact on the maximal achievable rate \( R \) of the channel from Alice to Bob (i.e., the maximal achievable rate decreases as \( n \) decreases for a fixed decoding error probability \( \delta \) [15], which has not been considered in the literature of covert communication (including [11]). Therefore, the impact of finite \( n \) on covert communication has not been well examined. This leaves a significant gap in our understanding of the performance limit of practical covert communication, since in practice the length of a codeword is always finite. For example, to achieve transmission efficiency (e.g., short delay) we may require the codeword to be short (e.g., in the order of 100 channel uses) for vehicle-to-vehicle communication or real-time video processing [16].
A. Our Contributions

Considering AWGN channels, we study the impact of finite $n$ on both the maximal achievable rate at Bob and detection performance at Willie in covert communication. To this end, noting that the decoding error probability $\delta$ is not negligible when $n$ is finite, we first propose to adopt the effective throughput $\eta$ (i.e., $\eta = nR(1 - \delta)$) subject to $\xi \geq 1 - \epsilon$, as the performance metric to evaluate covert communication. As can be seen from the definition of $\eta$, it explicitly captures the tradeoff among $n$, $R$, and $\delta$ for a given covert requirement.

We consider a maximum blocklength of $N$ channel uses, in which the covert information needs to be transmitted. Hence, the actual number of channel uses $n$ is constrained by $n \leq N$. Although a larger $n$ offers more observations to Willie for detecting the transmission, we analytically prove that the optimal value of $n$ that maximizes $\eta$ subject to the given covert requirement is $N$ (i.e., the entire block with all available channel uses). We also determine the maximum allowable transmit power per channel use (denoted by $P^*$) that achieves the maximum $\eta$. Our examination shows that $P^*$ decreases as $N$ increases, which is due to the fact that increasing $N$ forces Alice to allocate less power for each channel use to meet the covert requirement. Nevertheless, we show that the maximum allowable total transmit power (i.e., $NP^*$) increases as $N$ increases, which leads to the fact that the effective throughput of the communication from Alice to Bob increases as $N$ increases. The results in this paper, for the first time, provide important insights on the design of covert communication.

Notations: Scalar variables are denoted by italic symbols. Vectors and matrices are denoted by lower-case and upper-case boldface symbols, respectively. Given a vector $\mathbf{x}$, $x[i]$ denotes the $i$-th element of $\mathbf{x}$. The expectation is denoted by $E\{\cdot\}$ and $\mathcal{CN}(0, \sigma^2)$ denotes the circularly-symmetric complex normal distribution with zero mean and variance $\sigma^2$.

II. SYSTEM MODEL

A. Channel Model

The system model of interest for covert communication is illustrated in Fig. 1, where each of Alice, Bob, and Willie is equipped with a single antenna. We assume the channel from Alice to Bob and the channel from Alice to Willie are only subject to AWGN. In the covert communication, Alice transmits $n$ complex-valued symbols $x[i]$ ($i = 1, 2, \ldots, n$) in each codeword to Bob, while Willie is passively collecting $n$ observations on Alice’s transmission to detect her presence (i.e., whether Alice is transmitting). In this work, we consider that the length of a codeword is constrained by a maximum blocklength denoted by $N$. Thus, we have $n \leq N$ as a constraint on $n$. We denote the AWGN at Bob and Willie as $r_b[i]$ and $r_w[i]$, respectively, where $r_b[i] \sim \mathcal{CN}(0, \sigma_b^2)$, $r_w[i] \sim \mathcal{CN}(0, \sigma_w^2)$. $\sigma_b^2$ and $\sigma_w^2$ are the noise variances at Bob and Willie, respectively. In addition, we assume that $x[i]$, $r_b[i]$, and $r_w[i]$ are mutually independent. We denote the transmit power of Alice as $P$ (i.e., $E\{|x[i]|^2\} = P$). Furthermore, we assume that Alice adopts Gaussian signaling, i.e., $x[i] \sim \mathcal{CN}(0, P)$.

B. Channel Coding Rate for Finite Blocklength

The received signal at Bob for each signal symbol is given by

$$y_b[i] = x[i] + r_b[i].$$

As pointed out by [15], the decoding error probability at Bob is not negligible when $n$ is finite. As such, for a given decoding error probability $\delta$ the channel coding rate of the channel from Alice to Bob can be approximated by [15, 17]

$$R \approx \log_2(1 + \gamma_b) - \frac{1}{n(\gamma_b + 1)^2} \left(\frac{Q^{-1}(\delta)}{\ln(2)} + \frac{\log_2(n)}{2n}\right),$$

where $\gamma_b = P/\sigma_b^2$ is the signal-to-noise ratio (SNR) at Bob, and $Q^{-1}(\cdot)$ is the inverse Q-function. Equivalently, for a given channel coding rate $R$, the decoding error probability at Bob is given by

$$\delta = Q\left(\sqrt{n(1 + \gamma_b)} \left(\ln(1 + \gamma_b) + \frac{1}{2} \ln(n) - R \ln 2\right)\right) \left(\gamma_b + 2\right)\sqrt{\gamma_b(\gamma_b + 2)}.$$

C. Binary Hypothesis Testing at Willie

In order to detect Alice’s presence, Willie is to distinguish the following two hypotheses

$$\begin{align*}
\mathcal{H}_0 : & \quad y_w[i] = r_w[i] \\
\mathcal{H}_1 : & \quad y_w[i] = x[i] + r_w[i],
\end{align*}$$

where $\mathcal{H}_0$ denotes the null hypothesis where Alice is not transmitting, $\mathcal{H}_1$ denotes the alternative hypothesis where Alice is transmitting, and $y_w[i]$ is the received signal at Willie. Following the assumptions detailed in Section II-A, we have the likelihood functions of $y_w[i]$ under $\mathcal{H}_0$ and $\mathcal{H}_1$ as $f(y_w[i] | \mathcal{H}_0) = \mathcal{CN}(0, \sigma_w^2)$ and $f(y_w[i] | \mathcal{H}_1) = \mathcal{CN}(0, P + \sigma_w^2)$, respectively. In the cover communication, the ultimate goal of Willie is to minimize the total error rate, which is given by

$$\xi = P_F + P_M,$$
where \( P_F \equiv \Pr(D_1 | H_0) \) is the false positive rate, \( P_M \equiv \Pr(D_0 | H_1) \) is the miss detection rate, \( D_1 \) and \( D_0 \) are the binary decisions that infer whether Alice is present or not, respectively. We assume that Willie knows both \( P \) and \( \sigma_w \) exactly, and thus the optimal test that minimizes \( \xi \) is the likelihood ratio test with \( \lambda = 1 \) as the threshold, which is given by

\[
P_F \equiv \prod_{i=1}^{n} f(y_{w_i}[i]|H_1) \geq \frac{D_1}{P_M} \quad \text{and} \quad P_M \equiv \prod_{i=1}^{n} f(y_{w_i}[i]|H_0) \leq \frac{D_0}{P_F}.
\]

After performing some algebraic manipulations, (6) can be reformulated as

\[
T \equiv \frac{1}{n} \sum_{i=1}^{n} |y_{w_i}[i]|^2 \geq \frac{\gamma(n)}{\Gamma(n)},
\]

where \( T \) is the average power of each received symbol at Willie and \( \Gamma \) is the threshold for \( T \), which is given by

\[
\Gamma \equiv \left( \frac{P + \sigma_w^2}{P} \right)^{\frac{n}{2}} \ln \left( \frac{P + \sigma_w^2}{\sigma_w^2} \right).
\]

As per (6) and (7), we note that the radiometer is indeed the likelihood ratio test with \( \xi \) given by \([5, 18, 19]\), where \( \xi \) is the false positive rate and miss detection rate are given by \([9, 11]\).

\[
\xi = 1 \frac{1}{2} D(P_0 \| P_1),
\]

\[
\max_{n,P} \eta, \quad \text{s.t.} \quad D(P_0 \| P_1) \leq 2\epsilon^2, \quad n \leq N.
\]

\[1\]We note that \( \lambda = 1 \) is due to the unknown or equal \( a \ priori \) probabilities, i.e., \( P_0 \) and \( P_1 \) are unknown or equal, where \( P_0 \) is the \( a \ priori \) probability that \( H_0 \) is true, \( P_1 \) is the \( a \ priori \) probability that \( H_1 \) is true, and \( P_0 + P_1 = 1 \). If both \( P_0 \) and \( P_1 \) are known, the total error rate is reformulated as \( \xi \equiv P_0 P_F + P_1 P_M \) and the optimal test that minimizes this reformulated \( \xi \) is the likelihood ratio test with \( \lambda = P_1/P_0 \). We also note that the assumption of equal \( a \ priori \) probabilities is commonly adopted in the literature of covert communication (e.g., \([5, 10, 11]\)).
Theorem 1: The optimal values of n and P that maximize the effective throughput $\eta$ subject to $D(P_0||P_1) \leq 2\epsilon^2$ and $n \leq N$ are derived as

$$n^* = N,$$

$$P^* = \sigma_w^2 + P^* \left[ \ln \left( \frac{P^*}{\sigma_w^2} + 1 - 2\epsilon^2 N \right) \right],$$

where $P^*$ is the solution to the fixed-point equation (19).

Proof: The detailed proof is provided in Appendix.

In Fig. 2 we plot the maximum allowable total transmit power $NP^*$ versus $\epsilon$ for different values of $N$, where $\sigma_w^2 = \sigma_{\xi}^2 = 1$.

**IV. NUMERICAL RESULTS**

In this section, we provide numerical results on the effective throughput subject to $\xi \geq 1 - \epsilon$ to verify our analysis on the covert communication with $D(P_0||P_1) \leq 2\epsilon^2$ as the constraint.

In Fig. 2 we plot the maximum allowable total transmit power $NP^*$ over the entire block versus $\epsilon$. In this figure and the following figures, the curves for $\xi \geq 1 - \epsilon$ are achieved by numerically evaluating the false positive and detection rates as per (9) and (10), respectively. In this figure, we observe that the $NP^*$ with $\xi \geq 1 - \epsilon$ as the constraint is higher than that with $D(P_0||P_1) \leq 2\epsilon^2$ as the constraint. This is due to the fact that the equality in (12) cannot be achieved in the considered system model, and hence $D(P_0||P_1) \leq 2\epsilon^2$ is a more strict constraint than $\xi \geq 1 - \epsilon$. We also observe that $NP^*$ increases (hence the effective throughput increases) as $N$ increases, which can be explained by our Theorem 1. Finally, we observe that $NP^*$ decreases (hence the effective throughput decreases) as $\epsilon$ decreases, which demonstrates the tradeoff between the covert requirement and the achievable effective throughput (e.g., a more strict covert requirement leads to a smaller effective throughput).

In Fig. 3, we plot $NP^*$, $\eta$, $D(P_0||P_1)$ versus $N$, where $\sigma_w^2 = \sigma_{\xi}^2 = 1$, $\delta = 0.01$, and $\epsilon = 0.1$. We define $\delta^* = \delta(P^*, N, R^*)$ and denote the maximum effective throughput as $\eta^*$, which is achieved by substituting $P^*$, $n^*$, $R^*$, and $\delta^*$ into (14).
power to each channel use in order to maintain the same level of
covertness, but also reduces the decoding error probability
in the communication from Alice to Bob, which turns out to
improve the effective throughput of the covert communication.

In Fig. 4, we plot the effective throughput $\eta$ subject to
$\xi \geq 1 - \epsilon$ versus the decoding error probability $\delta$. We first observe that as $N$ increases,
the optimal value of $\delta$ decreases as $N$ increases. As shown in Fig. 3 (c), the maximum allowable
transmit power $P^*$ decreases as $N$ increases. As per (2), we
observe that the optimal channel coding rate $R^*$ decreases
as $N$ increases. And we also plot the maximum effective throughput
per channel use (i.e., $\eta^*/N$) versus $N$ in Fig. 5. In this figure,
we first observe that as $N$ increases $\eta^*/N$ increases, which
is consistent with our observation found in Fig. 3 (d). We
also observe that as $\epsilon$ increases slightly (e.g., from 0.02 to
0.08) $\eta^*/N$ significantly increases. This demonstrates that the
achievable effective throughput is very sensitive to the covert requirement.

V. CONCLUSION

This work investigated the covert communication with finite
blocklength (i.e., a finite number of channel uses $n \leq N$)
over AWGN channels. We proved that the effective throughput
of covert communication is maximized when all available
channel uses are utilized, i.e., $n^* = N$. To guarantee the
same level of covertness, the maximum allowable transmit
power per channel use decreases as $N$ increases, while the
maximum allowable total transmit power over all channel uses
increases as $N$ increases. In contrast, we found that both the
effective throughput and the effective throughput per channel
use increase as $N$ increases. This is due to the fact that
increasing $N$ not only reduces the transmit power allocated
to each channel use, but also decreases the decoding error probability of the communication from Alice to Bob.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Coun-
cil’s Discovery Projects (DP150103905).

APPENDIX

We present our proof of Theorem 1 in the following 6 steps.

Step 1: We note that $\eta$ and $D(\mathbb{E}_0||\mathbb{E}_1)$ are both mono-
tonically increasing functions of $P$ and $\gamma$. As such, we can
conclude that the equality in the constraint (16) is always met
in order to maximize $\eta$. Thus, we have $D(\mathbb{E}_0||\mathbb{E}_1) = 2\epsilon^2$ and
following (13) we have

$$n = \frac{2\epsilon^2}{f(\gamma_w)},$$

where

$$f(\gamma_w) \triangleq \frac{D(\mathbb{E}_0||\mathbb{E}_1)}{\gamma_w} = \ln (\gamma_w + 1) - \frac{\gamma_w}{\gamma_w + 1},$$

and $\gamma_w = P/\sigma_b^2$ is the SNR at Willie.

Step 2: We note $f(0) = 0$ and we derive the first derivative of
$f(\gamma_w)$ with respect to $\gamma_w$ as

$$\frac{\partial f(\gamma_w)}{\partial \gamma_w} = \frac{\gamma_w}{(\gamma_w + 1)^2} \geq 0,$$

which leads to the fact that $f(\gamma_w)$ is a monotonically increasing
function of $\gamma_w$. With the constraint $D(\mathbb{E}_0||\mathbb{E}_1) = 2\epsilon^2$, $n$
is a monotonically decreasing function of $f(\gamma_w)$ as per (21),
which results in that $n$ is a monotonically decreasing function of
$\gamma_w$ (thus of $P$).

Step 3: Instead of directly proving $n^* = N$ for maximizing
the effective throughput, we next prove that $n^* = N$ maximi-
izes $n\gamma_w$ (i.e., maximizes $nP$) under the constraint (21) in
the remaining steps. This is due to the fact that \( nP \) is the total transmit power for the \( n \) channel uses and the effective throughput increases as the total transmit power increases [15].

**Step 4:** We next prove that either \( n = 1 \) or \( n = N \) maximizes \( n\gamma_w \). To this end, in the following we first show that \( n\gamma_w \) initially decreases and then increases with \( n \). Following (21) and (22), we have

\[
 n\gamma_w = \frac{2e^2}{g(\gamma_w)},
\]

where \( g(\gamma_w) \) is given by

\[
g(\gamma_w) = \frac{\ln(1+\gamma_w)}{\gamma_w} - \frac{1}{1+\gamma_w}. \quad (25)
\]

We then derive the first derivative of \( g(\gamma_w) \) with respect to \( \gamma_w \) as

\[
\frac{\partial g(\gamma_w)}{\partial \gamma_w} = \frac{h(\gamma_w)}{\gamma_w^2}\left(1 + \frac{1}{\gamma_w}\right)^2, \quad (26)
\]

where

\[
h(\gamma_w) = 2\gamma_w^2 + \gamma_w - (1 + \gamma_w)^2\ln(1 + \gamma_w). \quad (27)
\]

We note that there are only two solutions to \( h(\gamma_w) = 0 \) for \( \gamma_w \geq 0 \), including \( \gamma_w = 0 \) and \( \gamma_w = \gamma_w^* \). We also note that as \( \gamma_w \to \infty \) we have \( h(\gamma_w) \to -\infty \). Then, we can conclude that \( h(\gamma_w) \geq 0 \) for \( \gamma_w < \gamma_w^* \) and \( h(\gamma_w) \leq 0 \) for \( \gamma_w \geq \gamma_w^* \). As such, noting \( \gamma_w^* \) is a monotonically decreasing function of \( \gamma_w \) under the constraint (21), which is proved following (23). Therefore, we conclude that \( n\gamma_w \) first decreases and then increases as \( \gamma_w \) increases (i.e., \( n\gamma_w \) has one minimum value but no maximum value). We recall that \( n \) is a monotonically decreasing function of \( \gamma_w \) as noted above. As such, \( n\gamma_w \) is the maximum value obtained when \( n = 1 \) or \( n = N \).

**Step 5:** We next prove that \( n = N \) (not \( n = 1 \)) maximizes \( n\gamma_w \). Substituting \( \gamma_w^* \) into (21), we have \( n^* = 2e^2/\gamma_w^* \). For \( 0 < \epsilon < 0.4835 \), we have \( n^* < 1 \) due to \( f(\gamma_w^*) > 0.4675 \). When \( n^* < 1 \), \( n\gamma_w \) increases with \( n \) due to \( n \geq 1 \). As such, \( n\gamma_w \) is the optimal value of \( n \) that maximizes \( n\gamma_w \) as \( n \) increases. We note that \( n\gamma_w \) is the maximum value obtained when \( n = 1 \) or \( n = N \).

We obtain this maximum value by solving \( f(\gamma_w^*) = 2e^2 \). The maximum value of \( \gamma_w \) that guarantees \( f(\gamma_w^*) = 2e^2 \) (i.e., \( n = 1 \)) is obtained when \( \epsilon = 0.5 \) since \( f(\gamma_w^*) \) is a monotonically increasing function of \( \gamma_w \) as noted above. We obtain this maximum value by solving \( f(\gamma_w^*) = 0.4835 \). We obtain this minimum value by solving \( f(\gamma_w) = (0.4835)^2 \) as \( n\gamma_w^* > 2.32 \) when \( n = 2 \). As such, we have \( n\gamma_w < 2.3145 \) when \( n = 1 \) and \( n\gamma_w > 2.32 \) when \( n = 2 \). As such, \( n\gamma_w \) is the maximum value obtained when \( n = 1 \) or \( n = N \). We recall that \( n\gamma_w \) is the maximum value obtained when \( n = 1 \) or \( n = N \).

We have

\[
\gamma_w^* \approx 2.1626 \text{ by numerically solving } h(\gamma_w) = 0.
\]

**Step 6:** So far, we have proved \( n^* = N \). Then, substituting \( n^* = N \) into (21), we obtain the fixed-point equation in (19).

### References

[1] P. H. Che, S. Kadhe, M. Bakshe, C. Chan, S. Jaggi, and A. Sprintson, “Reliable, deniable and hidable communication: A quick survey,” in Proc. IEEE Inf. Theory Workshop, Nov. 2014, pp. 227–231.

[2] B. A. Bash, D. Goeckel, D. Towsley, and S. Guha, “Hiding information in noise: Fundamental limits of covert wireless communication,” IEEE Commun. Mag., vol. 53, no. 12, pp. 26–31, Dec. 2015.

[3] B. He, S. Yan, X. Zhou, and V. Lau, “On covert communication with noise uncertainty,” IEEE Commun. Lett., to appear, Dec. 2016.

[4] M. K. Simon, Jim K. Omura, Robert A. Scholtz, and Barry K. Levitt, Spread Spectrum Communications Handbook, McGraw-Hill, 1994.

[5] B. A. Bash, D. Goeckel, and D. Towsley, “Limits of reliable communication with low probability of detection on AWGN channels,” IEEE J. Sel. Areas Commun., vol. 31, no. 9, pp. 1921–1930, Sep. 2013.

[6] L. Wang, G. Wornell, and L. Zheng, “Fundamental limits of communication with low probability of detection,” IEEE Trans. Inf. Theory, vol. 62, no. 6, pp. 3493–3503, Jun. 2016.

[7] P. H. Che, M. Bakshe, and S. Jaggi, “Reliable deniable communication: Hiding messages in noise, in Proc. IEEE Int’l. Symp. Info. Theory, Jul. 2013, pp. 2945–2949.

[8] M. R. Bloch, “Covert communication over noisy channels: A resolvability perspective,” IEEE Trans. Inf. Theory, vol. 62, no. 5, pp. 2334–2354, May 2016.

[9] Z. Lee and R. Baxley, “Achieving positive rate with undetectable communication over AWGN and Rayleigh channels,” in Proc. IEEE ICC, 2014, Jun. 2014, pp. 780–785.

[10] B. A. Bash, D. Goeckel, and D. Towsley, LPD communication when the warden does not know when, in Proc. IEEE Int. Symp. Inf. Theory, Jul. 2014, pp. 606–610.

[11] Z. Lee, R. J. Baxley, M. A. Weitnauer, and B. Walkerhorst, “Achieving undetectable communication,” IEEE J. Sel. Topics Signal Process., vol. 9, no. 7, pp. 1195–1205, Oct. 2015.

[12] T. Sobers, B. Bash, D. Goeckel, S. Guha, and D. Towsley, “Covert communication with the help of an unformed jammer achieves positive rate,” in Proc. Asilomar Conf. on Signals, Systems, and Comput., Nov. 2015, pp. 625–629.

[13] D. Goeckel, B. Bash, S. Guha, and D. Towsley, Covert communications when the warden does not know the background noise power, IEEE Commun. Lett., vol. 20, no. 2, pp. 236–239, Feb. 2016.

[14] K. Shahzad, X. Zhou, and S. Yan, “Covert communication in fading channels under channel uncertainty,” in Proc. IEEE VTC Spring, to appear, Feb. 2017.

[15] Y. Polyanyskyy, H. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” IEEE Trans. Inf. Theory, vol. 56, no. 5, pp. 2307–2359, May 2010.

[16] B. Makki, T. Svensson, and M. Zorzi, “Finite block-length analysis of spectrum sharing networks using rate adaptation,” IEEE Trans. Commun., vol. 63, no. 8, pp. 2823–2835, Aug. 2015.

[17] G. Ozcan and M. C. Gursoy, “Throughput of cognitive radio systems with finite blocklength codes,” IEEE J. Sel. Areas Commun., vol. 31, no. 11, pp. 2541–2554, Nov. 2013.

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

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

[20] X. Zhou, M. R. McKay, B. Maham, and A. Hjortnæs, “Rethinking the secrecy outage formulation: A secure transmission design perspective,” IEEE Commun. Lett., vol. 15, no. 3, pp. 302–304, Mar. 2011.

[21] S. Yan, N. Yang, G. Geraci, R. Malaney, and J. Yuan, “Optimization of code rates in SISO/M wiretap channels,” IEEE Trans. Wireless Commun., vol. 14, no. 11, pp. 6377–6388, Nov. 2015.