Covert Communications in Multi-Channel Slotted ALOHA Systems

Azadeh Sheikholeslami, Majid Ghaderi and Dennis Goeckel

Abstract—The fundamental limits of covert communication, where a message is sent from transmitter Alice to intended recipient Bob without detection by an attentive adversary warden Willie, has been considered extensively in recent years at the physical layer. The covert throughput depends critically on the warden's understanding of the characteristics of the radio environment and the type of receiver that he employs, and, as expected, the throughput increases when the warden has some uncertainty about the environment or some non-idealities in his receiver. In this paper, we consider the covert throughput when the adversary is only able to observe the medium access control (MAC) layer in a wireless communication system. In particular, given that the system has a rate of λ packets per slot transmitted over n channels by allowable system users, we study the allowable rate λ_a by covert users while maintaining covertness from an attentive warden observing the channel status in a slotted ALOHA system. We characterize performance for wardens with different abilities to discern the number of packets on a given channel, ranging from simple receivers that detect only whether there was a packet present to complicated receivers that can determine the number of packets involved in any collision, and also consider intended recipients Bob with varying abilities to perform multi-packet reception. In contrast to prior work in covert communications, the application considered motivates the consideration of results for finite (often small) observation vector lengths n at the adversary. Numerical results are provided both to illustrate the tightness of our achievability regions for the packet transmission rate of the covert transmitters and to demonstrate the covert throughput of the system as a function of λ and n.

Index Terms—Covert communications, Wireless system security, Multi-channel ALOHA.

---◆---

1 INTRODUCTION

Security and privacy are major concerns of modern communication networks. Much of the work on secrecy with cryptographic and information-theoretic approaches has considered hiding the content of a message transmitted by Alice intended for legitimate receiver Bob from an eavesdropper Eve. However, there are applications where even the presence of a message can convey meaningful information to the adversary. For example, radio transmissions can be used as a surrogate for the presence of military activity, or any communication between dissidents that is detected by an authoritarian government might be reason to shut down all communications. And, in particular, the Snowden disclosures indicate that the "meta-data" revealing whom is talking to whom can be of significant interest to an observer. This motivates the study of undetectable communications, which has been termed "covert communications" in recent literature. In the covert communications scenario, Alice tries to communicate to legitimate receiver Bob without detection of the presence of that message by an attentive and capable adversary denoted warden Willie.

Undetectable communications has been of great interest historically. At the physical layer, spread spectrum tech-

niques have been traditionally employed when such low probability of detection (LPD) communication was of interest. In computer networks, covert channels have attracted significant interest [2], [3], which can be classified into covert timing channels [4]–[6] and covert storage channels [7], [8]. Various practical aspects and security issues related to covert channels have been considered in the literature. For example, covert channels in cloud computing environments, where multiple virtual machines share the same physical server, have been extensively studied as a means for sharing small, but sensitive data (e.g., a secret key) [9]–[11].

However, a fundamental investigation of the achievable throughput of such a system was not established until [12], and then independently and formally in [13], [14]. In [13], [14], additive white Gaussian noise (AWGN) channels from Alice to each of Bob and Eve were considered, and it was established that, in n channel uses, $O(\sqrt{n})$ bits[1] could be transmitted reliably from Alice to Bob while bounding warden Willie's detection error probability arbitrarily close to one-half; conversely, $\omega(\sqrt{n})$ bits[2] cannot be transmitted reliably while being kept covert from Willie. The work of [13], [14] motivated significant further work on the characterization of covert point-to-point communication system performance as $n \to \infty$. In particular, performance limits with scaling constants for covert communications for point-to-point links, including discrete-memoryless channels (DMCs) and AWGN channels, were rapidly characterized [15]–[17]. Further work has begun to consider multiple

- A. Sheikholeslami is with the Department of Computer Science, Boston College, Chestnut Hill, MA. Email: sheikhoa@bc.edu

- M. Ghaderi is with the Department of Computer Science, University of Calgary, Calgary, Alberta. Email: mghaderi@ucalgary.ca

- D. Goeckel is with the Department of Elecectrical and Computer Engineering, University of Massachusetts, Amherst, MA. Email: goeckel@ece.umass.edu

1. Let $f(n)$ and $g(n)$ be two functions, and $\exists k > 0$ and $\exists N > 0$, such that $\forall n > N, |f(n)| \leq |k.g(n)|$, then $f(n) = O(g(n))$.

2. Let $f(n)$ and $g(n)$ be two functions, and $\forall k > 0, \exists N > 0$, such that $\forall n > N, 0 \leq k.g(n) \leq f(n)$, then $f(n) = \omega(g(n))$.

access channels [18] and coding schemes that can achieve the limits of covert communications [19]–[22].

Recent work has demonstrated that the achievable covert throughput depends critically on the capabilities of Willie's receiver and his knowledge of the operating environment. In [13], [14], the authors assume an ideal situation for Willie: Willie is able to employ an optimal receiver on the received physical layer signal, and the characteristics of the channel from Alice to Willie are known perfectly at Willie. In particular, [13], [14] assume that Willie knows both the timing of any potential message from Alice and also the variance of the AWGN affecting his received signal. In contrast, [23] considers what happens when the timing of the message is not known, and [24]–[26] consider what happens when the level of the background noise is not known. These works have demonstrated that there is a significant increase in covert throughput in the presence of such uncertainty, although [27] shows that the background noise level can generally be estimated so as to limit the covert throughput. To keep the adversary from estimating the background noise, the authors in [28] employ an uninformed jammer to improve the covert throughput.

Motivated by this extensive work at the physical layer, in this paper we move one layer up in the protocol stack to consider the covert throughput when the adversary warden Willie is viewing the behavior of the medium access control (MAC) to attempt to detect the presence of covert users. In particular, consider a slotted ALOHA system where each of N allowable system users randomly transmit a packet in each slot independently with probability p_{t_1} and thus with aggregate rate $\lambda = Np_{t_1}$. If a user transmits a packet in a given time slot, the user chooses one of the n channels at random. The warden Willie, with knowledge of N and p_{t_1}, has a statistical characterization of the number of packets on the n channels when only allowable users are present. If covert users are present, suppose that each of M covert users sends a packet in each slot independently with probability p_{t_2} and thus with aggregate rate $\lambda_a = Mp_{t_2}$; if a covert user sends a packet, they choose one of the n channels at random to do such. Willie attempts to detect the presence of such covert users by observing the number of packets on each of the n channels and determining whether these observations indicate a deviation from the statistical model of the behavior when only allowable users are present. We consider here values for λ_a as a function of λ that maintain covertness from a warden Willie as a function of his receiver complexity. In particular, we will consider a range of receivers at Willie, from a simple receiver able only to determine whether each channel contains a packet or not, to a complicated receiver that is able to determine the exact number of packets on each channel. Given the allowable λ_a determined by the covertness constraint, we then consider the covert throughput as a function of the complexity of the intended recipient Bob's receiver.

We will assume that p_{t_1} and p_{t_2} are small and that N and M are large. Given these assumptions, the number of packets from either the legitimate or covert users in a given slot will be modeled as a Poisson random variable with means λ and λ_a, respectively. This mathematical formulation makes the results applicable in other contexts, ranging from the covert throughput on optical communication channels [29] to packet insertion on Poisson channels [30]. However, in contrast to prior work in covert communications, the length n of the observation vector at Willie here is not the number of symbols in a codeword, which has (reasonably) been taken to infinity in previous works on the foundations of covert communication, but rather the number of channels in the system. Hence, rather than being concerned with only scaling results, we are instead interested in achievable covert throughput results for finite (and possibly very small) n. To our knowledge, these are the first non-asymptotic results in the recent study of the foundations of covert communications. Mathematically, this means that we are unable to rely on concentration inequalities [13]–[17] or laws of large numbers [31]–[33] to aid us in the derivation of our results.

The probability of error at the adversary Willie being lower bounded is a general condition of covertness, as defined precisely in Section 2 below. The probability of error of Willie's optimal receiver can be related to the total variation between the distribution of Willie's observations when Alice is not transmitting and the distribution of Willie's observations when Alice is transmitting. The total variation can be further upper bounded by the Kullback-Leibler (KL) distance, and, since the KL distance is generally more amenable to analysis than the total variation, particularly for the vector observation case generally of interest, a vast majority of prior work has focused on upper bounding the KL distance to insure covertness. However, the Poisson probability mass functions that arise in the model here are more easily addressed through total variation for $n = 1$, and we employed such in our preliminary work [1]. When turning to the multi-channel case, one might expect that we would need to employ the KL distance, since the vector KL distance in the case of independent components can be readily expressed as the sum of the KL distances in each of the components. However, the sum of the total variation in each dimension provides an upper bound to the total variation of the vector total variation [34], and hence a total variation approach can also be applied in the vector case. We will see that approaches based on both total variation and KL distance will require further bounding to derive clean analytical results for the achievable regions, reinforcing that it is not clear a priori which will lead to a larger achievable region. In fact, we will observe that the metric that leads to the larger achievable region for the covert traffic intensity λ_a will depend on the system traffic λ and the number of channels n; hence, achievable regions for λ_a under both a total variation constraint and a KL distance constraint will be considered in Section 3. Section 4 presents numerical results to demonstrate the tightness of our achievable covert rate regions under both constraints for different Willie's detection capability and different number of channels.

Although knowing the achievable regions for the covert rate λ_a is useful in designing a covert ALOHA system with n channels and load λ, the quantity that shows the actual amount of covert data passing through the system is the covert throughput. The covert throughput depends on the covert rate λ_a, and is limited by the collisions at Bob. Hence, there is a trade-off between an overloaded system with large λ (more covertness), and a system with small load λ (less collisions). In Section 5, we study the covert

throughput of transmission from Alice to Bob, and will find the system nodes' transmission rate λ that maximizes the covert throughput. Also, we will study the effect of number of channels n being utilized by Alice and the system nodes, and will find the number of channels that lead to the maximum covert throughput. Finally, Section 6 presents the conclusions and future work.

2 SYSTEM MODEL AND METRICS

2.1 System Model

Consider a multi-channel time-slotted random access system with N legitimate users $u_1, u_2, \ldots, u_N$, called system users, contending for n independent channels. Likewise, M covert users $c_1, c_2, \ldots c_M$ may (or may not) access the medium while trying to avoid detection by a network monitor termed Willie. In each time slot, each user independently flips a coin with probability of heads p_{t_1} (p_{t_2} for covert users) and if the result is heads, it transmits a packet in that time slot on a channel chosen uniformly at random. Hence, there is the possibility that multiple users will transmit in a given channel in a time slot, causing a collision. Throughout this work, we will assume that the number of users is large enough and p_{t_1} and p_{t_2} small enough such that:

- The number of transmissions in each time slot for each type of user can be approximated by a Poisson random variable. That is, the number of packet transmissions in a given time slot from system (allowed) users is Poisson with rate $\lambda = N p_{t_1}$, and the number of packet transmissions in a given time slot over the network from covert users (when they are present) is Poisson with rate $\lambda_a = M p_{t_2}$.
- Given λ, λ_a, n, and whether the covert users are present (or not), the number of packets in a given channel is independent of that in other channels. Strictly, for a finite number of system users and covert users, there would be dependence between the numbers of packets on distinct channels, as a user using one channel would preclude that user from using another channel. However, this dependence disappears as N and M become large for fixed λ and λ_a.

Traditionally, packets involved in a collision would simply be discarded and re-transmission would be required. However, with advances in multi-user detection, e.g. successive interference cancellation, multiple receiving antennas, capture effect, and etc. [35]–[39], it is often possible for an advanced receiver to recover multiple packets from a collision. The ability of a receiver to *receive* multiple simultaneous transmissions is referred to as Multi-Packet Reception in the literature [40]. To detect covert communications, the network monitor is only interested in detecting the number of packets transmitted in a time slot. In other words, the network monitor does not need to decode and recover the content of the packets, which requires a more sophisticated receiver. We will denote the ability to *detect* multiple simultaneous transmissions as Multi-Packet Detection (MPD). A receiver is called a K-MPD detector if it can detect up to K packet transmissions in a time slot on a channel. Let s denote

the state of a channel in a given time slot. Specifically, a K-MPD detector can detect one of the following $K+2$ channel states in a given time slot:

- Idle state ($s = 0$): No packets were on the channel during the time slot.
- Packet states ($s = 1, 2, \ldots, K$): Exactly s packets were detected on the channel during the time slot.
- Collision state ($s = K + 1$): In this state, more than K packets are transmitted over the channel.

That is, a K-MPD detector can always determine the number of packets involved in a collision as long as that number is less than or equal to K; when the number of packets involved in the collision is more than K, the receiver only knows that there were more than K packets on the channel but not the exact number. We will consider detectors ranging from a 0-MPD detector to a K-MPD detector, $1 \le K < \infty$, to an ∞-MPD detector. The 0-MPD detector, which is the weakest detector we will consider at Willie, can only determine whether there was any packets on the channel or not, whereas an ∞-MPD detector, which is the most powerful detector we consider at Willie, can determine the exact number of packets involved in any collision.

Let H_0 be the hypothesis that the covert users are not present, and H_1 the hypothesis that the covert users are present. We will denote $\mathbb{P}_0$ and $\mathbb{P}_1$ as the distribution of observed channel states at Willie given H_0 and H_1, respectively. Given the above assumptions, for state s ($0 \le s \le K + 1$), we have

$$\mathbb{P}_0(s) = \begin{cases} \frac{(\frac{\lambda}{n})^s e^{-\lambda/n}}{s!}, & s \le K, \\ \sum_{l=K+1}^{\infty} \frac{(\frac{\lambda}{n})^l e^{-\lambda/n}}{l!}, & s = K + 1, \end{cases} \quad (1)$$

and,

$$\mathbb{P}_1(s) = \begin{cases} \frac{(\frac{\lambda+\lambda_a}{n})^s e^{-(\lambda+\lambda_a)/n}}{s!}, & s \le K, \\ \sum_{l=K+1}^{\infty} \frac{(\frac{\lambda+\lambda_a}{n})^l e^{-(\lambda+\lambda_a)/n}}{l!}, & s = K + 1. \end{cases} \quad (2)$$

Willie makes such observations on each of n channels. As noted earlier, we will assume that N and M are large enough that the components of this vector observation containing the number of packets on each of the n channels given H_0 or H_1 can be modeled as mutually independent, which means that Willie will make his decision based on a vector observation for which the product distributions $\mathbb{P}_0^n$ and $\mathbb{P}_1^n$ characterize the observations under H_0 and H_1, respectively.

2.2 Metrics

Covertness Condition. The covert users' goal is to prevent Willie from determining whether the covert users are active or not. The condition they employ for such is that the probability of error of Willie's optimal decision is close to that which he would have obtained if he made the decision while ignoring his observations.

In particular, if the probability that the covert users is active (inactive) is $P(H_1)$ ($P(H_0)$), then Willie's probability of error is:

$$\mathbb{P}_E = \mathbb{P}_{FA}^W P(H_0) + \mathbb{P}_{MD}^W P(H_1),$$

where Willie's probability of false alarm is denoted by $\mathbb{P}_{FA}^W$, and his probability of missed detection is denoted by $\mathbb{P}_{MD}^W$. If Willie ignores his observations, he will always choose the most likely hypothesis and thus have an error probability of $\min(P(H_0), P(H_1))$. Noting that

$$\mathbb{P}_{FA}^W P(H_0) + \mathbb{P}_{MD}^W P(H_1) \geq (P_{FA}^W + P_{MD}^W)\min(P(H_0), P(H_1)),$$

a transmission will be defined to be *covert* when $\mathbb{P}_{FA}^W + \mathbb{P}_{MD}^W > 1 - \epsilon$ for any $\epsilon > 0$ [41], hence guaranteeing that Willie's performance is close to that obtained when he ignores his observations.

We define *covert rate* as the maximum rate[3] of packet transmission by the covert users, λ_a, such that their transmission is covert from Willie, as a function of the rate λ of packet transmission by the system users. We will also consider *covert throughput*, which is the reliable throughput obtained by the covert users as a function of λ, although we hasten to note that λ is a system parameter and hence not available for optimization by the covert users.

Total Variation. The total variation distance between two probability mass functions gives the maximum difference over all events in the probability assigned to the event by the corresponding probability measures, and is defined for $\mathbb{P}_0^n$ and $\mathbb{P}_1^n$ as:

$$d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n) = \frac{1}{2}\sum_{\mathbf{s}}|\mathbb{P}_0^n(\mathbf{s}) - \mathbb{P}_1^n(\mathbf{s})|, \tag{3}$$

where the sum is over all n-dimensional vectors of channel states (denoted by $\mathbf{s}$) in the support of $\mathbb{P}_0^n \cup \mathbb{P}_1^n$. An optimal receiver at Willie can be characterized in terms of the total variation as [42]:

$$\mathbb{P}_{FA}^W + \mathbb{P}_{MD}^W = 1 - d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n).$$

Hence, if

$$d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n) < \epsilon, \tag{4}$$

covertness is maintained. Because the sum over the n-dimensional support in (3) is challenging to use, total variation is often difficult to employ except in the $n = 1$ case, an example of which is done in the single channel version of this work in [1]. However, since we are interested in achievable rates, an upper bound for the total variation distance is sufficient. The following result from [34] characterizes an upper bound on the total variation distance between product distributions:

Theorem 1. *(Total Variation Bound) Let $\mathbb{P}_0^n$ and $\mathbb{P}_1^n$ be the products of probability distributions $\mathbb{P}_0$ and $\mathbb{P}_1$, respectively. The total variation distance between $\mathbb{P}_0^n$ and $\mathbb{P}_1^n$ is bounded as:*

$$d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n) \leq n d_{TV}(\mathbb{P}_0, \mathbb{P}_1). \tag{5}$$

Hence, in order to obtain covertness, it is sufficient to maintain

$$d_{TV}(\mathbb{P}_0, \mathbb{P}_1) < \frac{\epsilon}{n}. \tag{6}$$

Kullback-Leibler (K-L) Distance. The Kullback-Leibler distance (K-L distance), which provides another measure of the difference between two probability mass functions, is

3. Rate is defined as the long term average of the number of packets per time slot.

broadly used in information theory under the term relative entropy [43], and is defined as:

$$d_{KL}(\mathbb{P}_0, \mathbb{P}_1) = \sum_x \mathbb{P}_0(x)(\ln\mathbb{P}_0(x) - \ln\mathbb{P}_1(x)).$$

Pinsker's inequality gives the K-L distance operational meaning in detection theory [43]:

$$d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n) \leq \sqrt{\frac{d_{KL}(\mathbb{P}_0^n, \mathbb{P}_1^n)}{2}}, \tag{7}$$

where $d_{KL}(\mathbb{P}_0^n, \mathbb{P}_1^n)$ is the KL distance between $\mathbb{P}_0^n$ and $\mathbb{P}_1^n$. In particular, since the K-L distance can be used to provide an upper bound on the total variation, and hence a lower bound on Willie's probability of error, it can be used to establish achievability results for covert communications as we seek to do here. In particular, using (7), the condition

$$\sqrt{\frac{d_{KL}(\mathbb{P}_0^n, \mathbb{P}_1^n)}{2}} < \epsilon, \tag{8}$$

or equivalently,

$$d_{KL}(\mathbb{P}_0^n, \mathbb{P}_1^n) < 2\epsilon^2, \tag{9}$$

is sufficient to maintain covertness. The KL distance between product of probability distributions $\mathbb{P}_0^n$ and $\mathbb{P}_1^n$ is [43]:

$$d_{KL}(\mathbb{P}_0^n, \mathbb{P}_1^n) = n d_{KL}(\mathbb{P}_0, \mathbb{P}_1). \tag{10}$$

This property makes the KL distance a very useful tool in analyzing covertness when facing with multiple independent and identically distributed random variables. In particular, (9) becomes

$$d_{KL}(\mathbb{P}_0, \mathbb{P}_1) < \frac{2\epsilon^2}{n}. \tag{11}$$

By applying Pinsker's inequality to the distribution of a single channel, one can observe that the condition in (6) leads to a tighter lower bound on the error probability at Willie than the condition in (11) for $n = 1$, and hence (6) should lead to a larger region of achievable rates for Alice than than in (11) for the $n = 1$ case. However, for larger n, it is not clear which criterion is tighter, and thus it is not apparent a priori which of the metrics will lead to higher achievable rates for $n > 1$. Furthermore, we will employ significant bounding when employing each of these metrics to find achievable rates for Alice. Hence, we explore achievable rates under both metrics in the next section.

3 COVERT RATE ANALYSIS

In this section, we present the main results of the paper. We will consider two covertness metrics: total variation distance and KL distance.

3.1 Covert Rate with Total Variation Distance Metric

In this subsection, we characterize the total variation distance of the state of a channel conditioned on the presence or absence of covert transmissions when Willie has different detection capabilities. We consider four different scenarios, and will find an achievable region for the transmission rate of the covert users in each case:

1) *0-MPD detector*: Willie can only distinguish between an idle channel and a busy channel. This is the least capable Willie, and thus the allowable covert rates will be the largest (or at least no smaller) compared to those of other scenarios.

2) *1-MPD detector*: Willie can distinguish between an idle channel and one packet transmission, but when more than one packet is transmitted over the channel, he only knows that more than one packet was on the channel.

3) *K-MPD detector*: Willie can detect up to K simultaneous transmissions. If more than K packets are transmitted over the channel, he only detects a collision.

4) *∞-MPD detector*: Willie can determine the exact number of packets on the channel. This is the most capable Willie, and thus the allowable covert rates are the smallest (or at least no larger) compared to the other scenarios.

Note that, in any security scheme, the goal is to protect against a certain class of adversaries. In this work, we provide security (covertness) against a class of Willies that is only able to observe the medium access control (MAC) layer in a wireless communication system. In the following, in each case we characterize the probability distribution of the channel state conditioned on the presence or absence of covert users. Recall that the channel state denotes the number of packet transmissions in a time slot. The conditional state probability distributions are then used to compute the total variation distance, and consequently the covert transmission rate.

Theorem 2 (Willie with 0-MPD capability). *In the presence of a 0-MPD Willie observing n channels with total system nodes' rate λ, if $\lambda \geq n \ln\left(\frac{n}{\epsilon}\right)$, the transmission of covert users is covert for any $\lambda_a > 0$. Otherwise, if $\lambda \leq n \ln\left(\frac{n}{\epsilon}\right)$, the transmissions of covert users is covert if*

$$\lambda_a \leq n \ln \frac{1}{1 - \frac{\epsilon}{n} e^{\lambda/n}}.$$

Proof. When Willie has 0-MPD capability, he can only determine if the channel is busy or not. If the channel is busy, Willie is not able to determine how many packets are being transmitted in a time slot.

In this case, the probability distributions of the network states observed by Willie, $\mathbb{P}_0$ and $\mathbb{P}_1$, are Bernoulli distributions. Let $\mathcal{S} = \{s_0, s_1\}$ denote the set of states of each Bernoulli process, where s_0 and s_1 indicate that the channel is, respectively, idle and busy. We have:

$$\mathbb{P}_0\{s = s_0\} = 1 - \mathbb{P}_0\{s = s_1\} = e^{-\lambda/n},$$
$$\mathbb{P}_1\{s = s_0\} = 1 - \mathbb{P}_1\{s = s_1\} = e^{-(\lambda+\lambda_a)/n}. \quad (12)$$

Hence, the total variation distance between $\mathbb{P}_0^n$ and $\mathbb{P}_1^n$ is,

$$d_{TV}(\mathbb{P}_0, \mathbb{P}_1) = \frac{1}{2} \sum_{s \in \mathcal{S}} |\mathbb{P}_1(s) - \mathbb{P}_0(s)|$$
$$= \frac{1}{2} \left(\left| e^{-\lambda/n} - e^{-(\lambda+\lambda_a)/n} \right| + \left| 1 - e^{-\lambda/n} - (1 - e^{-(\lambda+\lambda_a)/n}) \right| \right)$$
$$= e^{-\lambda/n}(1 - e^{-\lambda_a/n}), \quad (13)$$

where the last equality holds because for $\lambda_a \geq 0$, we have $1 - e^{-\lambda_a/n} \geq 0$. Using Theorem 1, the covertness condition is satisfied if

$$d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n) \leq n e^{-\lambda/n}(1 - e^{-\lambda_a/n}) \leq \epsilon. \quad (14)$$

The term $1 - e^{-\lambda_a/n}$ in the above equation is always less than one. Hence, if $\frac{\epsilon}{n} e^{\lambda/n} \geq 1$, any covert rate $\lambda_a \geq 0$ can be obtained. Otherwise, in order to maintain covertness,

$$\lambda_a \leq n \ln \frac{1}{1 - \frac{\epsilon}{n} e^{\lambda/n}}, \quad (15)$$

should be satisfied. $\square$

Theorem 3 (Willie with 1-MPD capability). *In the presence of a 1-MPD Willie observing n channels with total system nodes' rate λ, if λ satisfies*

$$\frac{\epsilon e^{\lambda/n}}{1 + \max\{\frac{\lambda}{n}, 1 - \frac{\lambda}{n}\}} \geq n,$$

the transmission of covert users is covert for any $\lambda_a \geq 0$. Otherwise, the covertness is maintained if

$$\lambda_a \leq -n \ln \left(1 - \frac{\epsilon e^{\lambda/n}}{n + \max\{n - \lambda, \lambda\}} \right).$$

Proof. See Appendix A. $\square$

In the next theorem, we characterize the covert rate when Willie has K-MPD capability. In this case, in order to be able to find a closed form expression for the covert rate, we only consider $\lambda \geq nK$.

Theorem 4 (Willie with K-MPD capability, $\lambda \geq nK$). *In the presence of a K-MPD Willie observing n channels with total system nodes' rate $\lambda \geq nK$, if*

$$\frac{\epsilon e^{\lambda/n}}{\sum_{k=0}^{K} \frac{(\lambda/n)^k}{k!}} \geq n,$$

transmission is covert for any $\lambda_a \geq 0$. Otherwise, covertness is maintained if

$$\lambda_a \leq -n \ln \left(1 - \frac{\epsilon e^{\lambda/n}}{n \sum_{k=0}^{K} \frac{(\lambda/n)^k}{k!}} \right).$$

Proof. See Appendix B. $\square$

Theorem 5 (Willie with ∞-MPD capability). *In the presence of an ∞-MPD Willie observing n channels with total system nodes' rate λ, if,*

$$\lambda_a \leq \max \left\{ \epsilon, \frac{e\epsilon^2}{2n} + \frac{\epsilon\sqrt{2e\lambda}}{n} \right\},$$

then transmissions of cover users are covert.

Proof. See Appendix C. $\square$

Overloaded Willie. With the exception of the case of an ∞-MPD receiver at Willie, there is an intensity λ of system traffic above which Willie cannot detect the covert communication, regardless of λ_a. Intuitively, this is because the channel hits the largest state of Willie's detector so often just from the system traffic that the presence of the covert users rarely changes his observation. Next, we establish this intuition more formally. Consider $n = 1$. Observe that

Willie's optimal strategy will result in him always choosing H_1 if the system state is $s = K+1$ if he is trying to minimize $P_{FA} + P_{MD}$; if this were not true, his strategy would choose H_0 regardless of the system state and thus $P_{FA} + P_{MD} = 1$. Given that he always chooses H_1 when $s = K + 1$,

$$P_{FA} + P_{MD} \geq P_{FA} \geq P(s = K + 1|H_0) = \sum_{l=K+1}^{\infty} \frac{\lambda^l e^{-\lambda}}{l!}. \tag{16}$$

Thus, if

$$\sum_{l=K+1}^{\infty} \frac{\lambda^l e^{-\lambda}}{l!} \geq 1 - \epsilon, \tag{17}$$

λ_a can be arbitrary, which recovers the condition on λ for λ_a to be arbitrary for $n = 1$ in Theorems 2 and 4.

For $n > 1$, let M be the event that Willie's receiver observes its maximum on each channel; that is, the event that $\{s_1 = K + 1\} \cup \{s_2 = K + 1\} \cup \ldots \cup \{s_n = K + 1\}$. Then, following the reasoning above:

$$P(M) = 1 - P(\bigcup_{i=1}^{n} \{s_i \leq K\}) \tag{18}$$

$$\geq 1 - \sum_{i=1}^{K} P(s_i \leq K) \tag{19}$$

$$= (1 - n) \sum_{l=0}^{K} \frac{(\lambda/n)^l e^{-\lambda/n}}{l!}, \tag{20}$$

which is greater than $1 - \epsilon$, allowing arbitrary covert traffic λ_a, for the conditions given in Theorems 2 and 4 for $n > 1$.

3.2 Covert Rate with KL Distance Metric

The additive property of the KL distance for independent distributions makes it a useful metric in analyzing covertness in the literature [43]. In particular, it can be a useful metric when considering multiple independent channels. Hence, in this section we characterize the KL distance conditioned on the presence or absence of covert transmissions when Willie has different detection capabilities. Since KL distance is a complicated function of the probability distributions of the channel states observed by Willie, in this section we only consider Willie with 0-MPD and ∞-MPD detection capabilities.

Theorem 6 (Willie with 0-MPD capability). *In the presence of a 0-MPD Willie observing n channels with total system nodes' rate λ, if*

$$\lambda \geq n \ln \left(\frac{n}{2\epsilon^2} + 1 \right),$$

any $\lambda_a \geq 0$ can be obtained. Otherwise, if

$$\lambda_a \leq n \ln \left(\frac{1}{1 - \frac{2\epsilon^2}{n}(e^{\lambda/n} - 1)} \right),$$

then the transmissions of covert users are covert.

Proof. See Appendix D. □

Theorem 7 (Willie with ∞-MPD capability). *In the presence of an ∞-MPD Willie observing n channels with total system nodes' rate λ, if*

$$\lambda_a \leq 2\epsilon\sqrt{\lambda},$$

then the transmissions of covert users are covert.

Proof. See Appendix E. □

Willie with ∞-MPD capability is the most capable Willie, and thus the covert rates in this case are the smallest compared to the covert rates in other scenarios. In fact, in contrast to the other scenarios, we see a square root law similar to that in [41] and subsequent work [29] at the physical layer.

4 NUMERICAL EXAMPLES

In this section, we numerically evaluate the bounds obtained in previous sections. In all figures we set $\epsilon = 0.1$ as the covertness parameter[4].

4.1 Covert rate of single-channel ALOHA

First, we study how the covert rate λ_a depends on the transmission rate of the system nodes in a single-channel ALOHA network configuration[5]. Whereas we have provided closed-form analytical expressions for the covert rate achievable by the covert users, it is possible to evaluate the exact covert rate numerically. This exact covert rate along with the covert rate with TV distance, and the covert rate with KL distance, versus the system nodes' transmission rate are shown in Fig. 1, with different levels of Willie's MPD capability. These figures show the accuracy of our analytical bounds. Also, they demonstrate that as expected from results of Section 3, when Willie has 0-MPD, 1-MPD, or 5-MPD capability, the covert rate λ_a can be arbitrarily large for sufficiently large λ. Whereas at first this seems surprising (or even erroneous), the reason is clear: if λ is much larger than K, then Willie's detector observes the event of more than K packets occurring with high probability in each time slot, even when the covert users are not present. When the covert users are present, they are unlikely to cause a change in the observation at Willie, hence providing covertness. *A more interesting observation is that the square root law of [41] is not observed at small λ for the 0-MPR and 1-MPR detector, hence indicating the degree to which only a noisy view of the collision status of the channel can hide the presence of covert users.*

4.2 Covert rate of multi-channel ALOHA

In this section, we study examples of multi-channel ALOHA networks. In Figs. 2-4, the exact and lower bound (evaluated based on TV distance and/or KL distance in Section 3) for the covert rate λ_a versus system nodes rate λ for different Willie capabilities are shown. In each figure, a single channel network, a two channel network, and a ten channel network are considered.

4. As in Section 2.2, $\epsilon = 1 - (\mathbb{P}_{FA}^W + \mathbb{P}_{MD}^W)$, and thus a small ϵ is desirable.
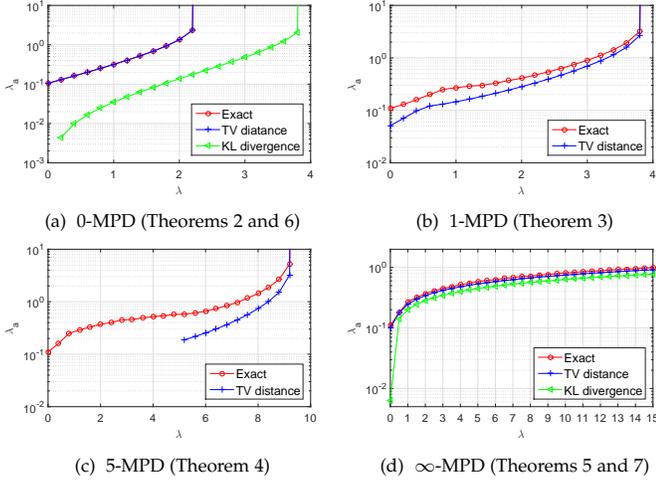5. This section is presented in [1].

(a) 0-MPD (Theorems 2 and 6)

(b) 1-MPD (Theorem 3)

(c) 5-MPD (Theorem 4)

(d) ∞-MPD (Theorems 5 and 7)

Fig. 1: Single channel ($n = 1$) covert rate λ_a versus system nodes rate λ.

(a) $n = 1$

(b) $n = 2$

(c) $n = 5$

(d) $n = 10$

Fig. 2: Covert rate λ_a versus the rate of system nodes λ for different number of channels when Willie employs a 0-MPD detector. As n increases, the achievable covert rate λ_a for a given rate λ decreases.

In Figs. 2 and 3, it can be seen that as λ gets larger, the covert users can hide a higher number of covert packets in the traffic of legitimate users and thus λ_a becomes higher. Also, for $n = 1$ and $n = 2$, when λ is larger than a threshold, Willie is overwhelmed and thus the communication is covert for any λ_a. This threshold for $n = 5$ and $n = 10$ is higher and is not depicted in Figs. 2 and 3.

Another observation is that as the number of channels n increases, for a given λ, the covert rate λ_a becomes smaller. Hence, increasing the number of channels does not help to improve the covert rate. However, in the next section, we will show that a larger number of covert channels can help to improve the covert throughput.

In Fig. 4, for Willie with ∞-MPD capability, as is shown in Section 3, the exact λ_a and the KL bound are independent of the number of channels n. However, the total variation (TV) bound depends on n, and since the total variation bound becomes looser as n increases, the difference between the exact λ_a and the total variation λ_a increases too. Hence, in general for the ∞-MPD case, the KL distance is a more accurate metric and should be used in covertness analysis.

In order to calculate the covert throughput, in addition to the covertness constraint we should consider the success probability of packet transmissions in the presence of collisions. We will consider this in the next section.

5 COVERT THROUGHPUT ANALYSIS

Suppose covert transmitters and a receiver (Bob) are trying to convey as many covert packets as possible. The number of covert packets that Bob can receive successfully is determined by the covertness constraint, the number of collisions at Bob's receiver, and Bob's multiple packet reception (MPR) capability. The achievable transmission rate of covert transmitters in different network settings is obtained in Section 3, and numerical examples are studied in Section 4. In this section, we analyze the covert throughput of the network by considering Bob's capability in successful reception of packets. Given the throughput, other performance metrics
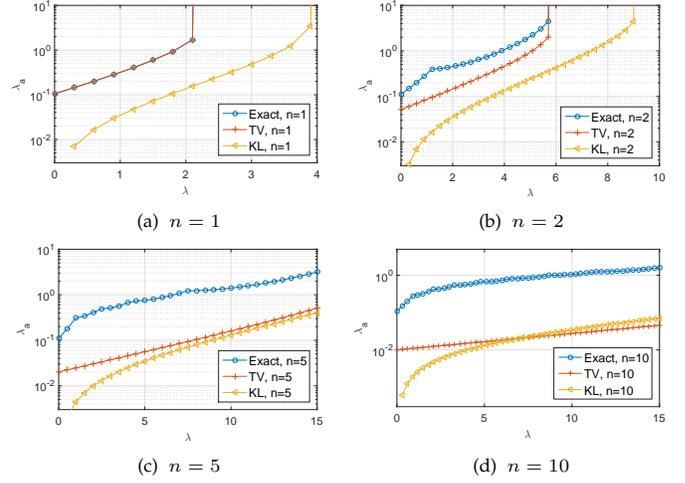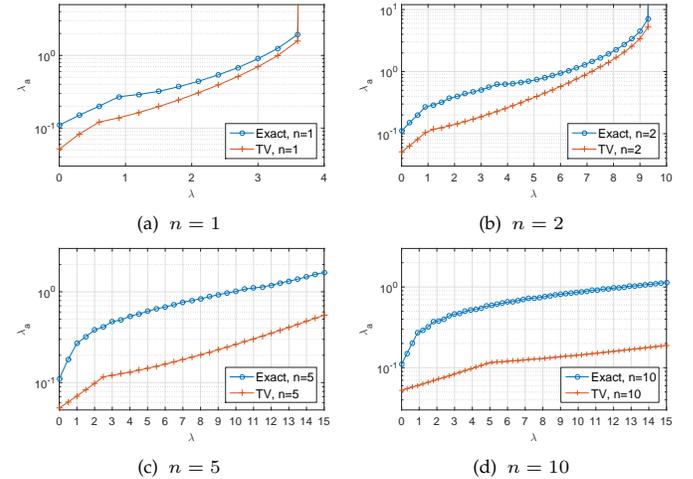

(a) $n = 1$

(b) $n = 2$

(c) $n = 5$

(d) $n = 10$

Fig. 3: Covert rate λ_a versus the rate of system nodes λ for different number of channels when Willie employs a 1-MPD detector.

of the covert users are readily derived. For instance, the average delay,

$$\text{delay} \propto \frac{1}{\text{Throughput}},$$

and the loss probability,

$$P_{\text{Loss}} = 1 - \frac{\text{rate}}{\text{Throughput}},$$

are both monotonic functions of the throughput. In the following, we will consider a receiver Bob with different multiple packet reception capabilities: Bob with single packet reception capability (1-MPR), two packet reception capability (2-MPR), and infinity packet reception capability (∞-MPR).
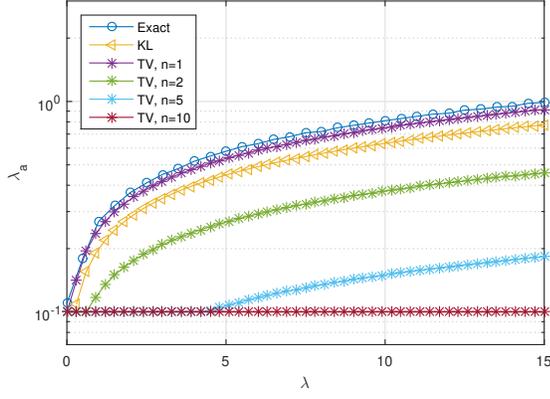
Fig. 4: Covert rate λ_a versus the rate of system nodes λ for different number of channels ($n = 1, 2, 5$ and 10) when Willie employs an ∞-MPD detector. Note that the KL bound does not depend on the number of channels (Theorem 7).

5.1 Bob with single packet reception capability (1-MPR)

Suppose Bob's receiver has 1-MPR capability, i.e. it can receive and decode a packet successfully if only one packet is transmitted in a time slot on a given channel. The success probability is the probability that on a given channel, in a given time slot, only one packet from the covert users and no packet from the system users, is transmitted. Hence, the covert throughput of a channel is given by

$$\tau^{\text{1-MPR}} = \frac{\lambda_a e^{-(\lambda+\lambda_a)/n}}{n}. \tag{21}$$

Consequently, the covert throughput of a network consisting of n channels is

$$T^{\text{1-MPR}} = \lambda_a e^{-(\lambda+\lambda_a)/n}. \tag{22}$$

First, assuming that there is no restriction on λ_a, we find the values of λ_a for which $T^{\text{1-MPR}}$ is maximized:

$$\frac{\partial T^{\text{1-MPR}}}{\partial \lambda_a} = e^{-(\lambda+\lambda_a)/n} - \frac{\lambda_a e^{-(\lambda+\lambda_a)/n}}{n} = 0, \tag{23}$$

which results in $\lambda_a = n$.

Next, we will study the covert throughput when Willie has different detection capabilities.

5.1.1 Willie with 0-MPD receiver

From Theorem 2, if $\epsilon e^{\lambda/n} \geq n$, λ_a is not limited by the covertness constraint, and thus the maximum covert throughput is obtained when $\lambda_a = n$. Thus, when $\epsilon e^{\lambda/n} \geq n$,

$$T^{\text{1-MPR}}_{\text{0-MPD}} = \frac{n e^{-\lambda/n}}{e}. \tag{24}$$

However, if $\epsilon e^{\lambda/n} \leq n$, the transmission is covert if $\lambda_a \leq \epsilon e^{\lambda/n}$. Since (22) is increasing in λ_a for any $\lambda_a \leq n$, it is also increasing for $\lambda_a \leq \epsilon e^{\lambda/n}$. Hence, the maximum covert throughput is obtained when $\lambda_a = \epsilon e^{\lambda/n}$, i.e.

$$T^{\text{1-MPR}}_{\text{0-MPD}} = \frac{\epsilon}{n} \exp\left(\frac{-\epsilon e^{\lambda/n}}{n}\right). \tag{25}$$
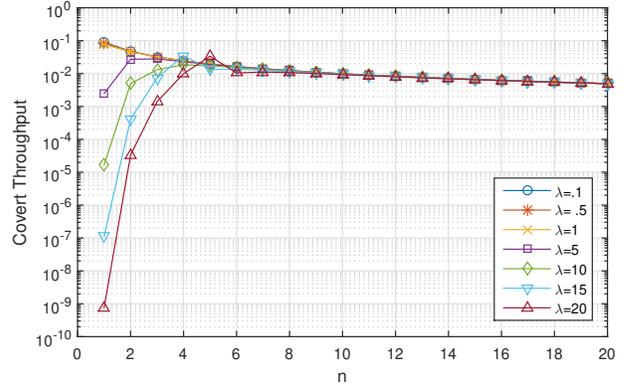

Fig. 5: Covert throughput versus number of channels for different values of system nodes' transmission rate λ. Willie employs a 0-MPD detector and the intended receiver (Bob) employs a 1-MPR receiver.

In Fig. 5, covert throughput versus the number of channels for different values of the system nodes' transmission rate λ is shown. For small values of λ, increasing the number of channels does not help to improve the covert throughput. However, as λ increases, the number of collisions in the network increases and thus having more channels can help to improve the covert throughput.

5.1.2 Willie with 1-MPD receiver

From Theorem 3, if ϵ and λ are such that

$$\frac{\epsilon e^{\lambda/n}}{n + \max\{\lambda, n - \lambda\}} \geq 1, \tag{26}$$

the transmission is covert for any $\lambda_a \geq 0$. Hence, from (23) the maximum covert throughput is obtained when $\lambda_a = n$, i.e.,

$$T^{\text{1-MPR}}_{\text{1-MPD}} = \frac{n e^{-\lambda/n}}{e}. \tag{27}$$

If ϵ and λ are such that

$$\frac{\epsilon e^{\lambda/n}}{n + \max\{\lambda, n - \lambda\}} < 1, \tag{28}$$

the covert rate that maximizes the throughput is:

$$\lambda_a = \min\left\{ n \ln\left(\frac{n + \max\{n - \lambda, \lambda\}}{n + \max\{n - \lambda, \lambda\} - \epsilon e^{\lambda/n}} \right), n \right\}. \tag{29}$$

and the covert throughput is obtained by substituting λ_a from (29) into (22).

In Fig. 6, the curves of covert throughput versus number of channels for different system nodes' transmission rates are depicted. Similar to the case of Willie with 0-MPD capability, for small values of λ, increasing the number of channels does not help to improve the covert throughput, but as λ increases having more channels can help to improve the covert throughput.

5.1.3 Willie with ∞-MPD receiver

In Section 4, it is shown that for the case of ∞-MPD, KL distance gives a tighter lower bound of covert rate
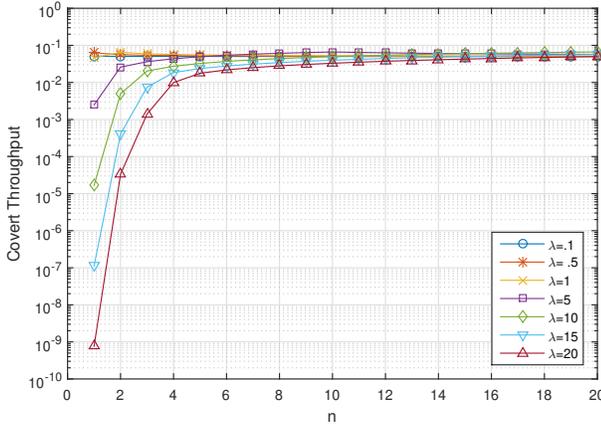
Fig. 6: Covert throughput versus number of channels for different values of system nodes' transmission rate λ. Willie employs a 1-MPD detector and Bob employs a 1-MPR receiver.
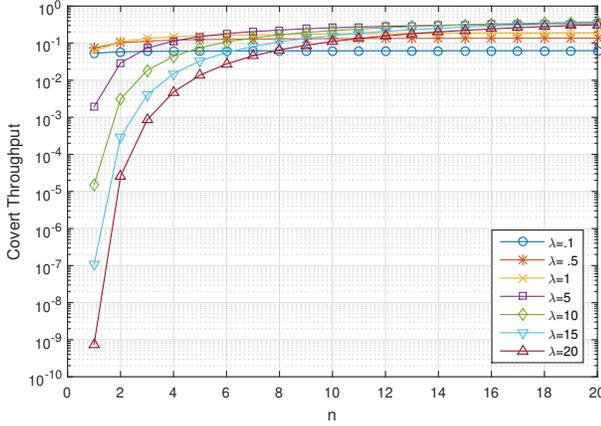

Fig. 7: Covert throughput versus the number of channels for different values of system nodes' transmission rate λ. In this case, Willie employs an ∞-MPD detector and Bob employs a 1-MPR receiver.

λ_a compared to TV distance. Hence, in this case we use Theorem 7 for λ_a, i.e.,

$$\lambda_a \leq 2\epsilon\sqrt{\lambda}. \tag{30}$$

Let ϵ be sufficiently small such that $\lambda \leq \frac{1}{4\epsilon^2}$, and,

$$\lambda_a \leq 2\epsilon\sqrt{\lambda} \leq n, \ \forall n \geq 1, \tag{31}$$

and thus (22) is increasing in λ_a. Hence, the maximum throughput is obtained when $\lambda_a = 2\epsilon\sqrt{\lambda}$,

$$T^{\text{1-MPR}}_{\infty\text{-MPR}} = 2\epsilon\sqrt{\lambda}e^{-(\lambda+2\epsilon\sqrt{\lambda})/n}. \tag{32}$$

In Fig. 7, covert throughputs versus the number of network channels for different system nodes' transmission rates are shown. In this setting, $\epsilon = 0.1$ and thus (32) can be used for $\lambda \leq 25$. Similar to Willie with 0-MPD and 1-MPD detector, for small values of λ increasing the number of channels does not help, but as λ, and consequently the number of collisions, increases more channels will help to obtain a higher covert throughput.

5.2 Bob with multiple packet reception capability (2-MPR)

Now, suppose Bob has 2-MPR capability, i.e. if two packets are transmitted over one channel simultaneously, Bob can receive and decode both packets successfully. Let L is the random variable associated with the number of covert packets received at Bob. Hence, the expected number of packets successfully received at Bob is given by,

$$
\begin{aligned}
E\{L\} &= \sum_{\ell=1}^{2} LP(\text{success}, L = \ell) \\
&= \frac{\lambda_a e^{-\frac{\lambda_a}{n}}}{n} \left(e^{-\frac{\lambda}{n}} + \frac{\lambda e^{-\frac{\lambda}{n}}}{n} \right) + 2 \times \left(\frac{\lambda_a}{n} \right)^2 \frac{e^{-\frac{\lambda_a}{n}}}{2} \left(e^{-\frac{\lambda}{n}} \right) \\
&= \frac{\lambda_a e^{-\frac{\lambda+\lambda_a}{n}}}{n} \left(1 + \frac{\lambda + \lambda_a}{n} \right). \tag{33}
\end{aligned}
$$

Hence, the covert throughput of a channel is given by

$$\tau^{\text{2-MPR}} = \frac{\lambda_a e^{-(\lambda+\lambda_a)/n}}{n} \left(1 + \frac{\lambda + \lambda_a}{n} \right), \tag{34}$$

and the covert throughput of the network consisting of n channels is,

$$T^{\text{2-MPR}} = \lambda_a e^{-(\lambda+\lambda_a)/n} \left(1 + \frac{\lambda + \lambda_a}{n} \right). \tag{35}$$

Now, assuming that there is no restriction on λ_a, $T^{\text{2-MPR}}$ is maximized when:

$$
\begin{aligned}
\frac{\partial T^{\text{2-MPR}}}{\partial \lambda_a} &= e^{-(\lambda+\lambda_a)/n} \left[1 + \frac{\lambda + \lambda_a}{n} \left(1 - \frac{\lambda_a}{n} \right) \right] \tag{36} \\
&= 0.
\end{aligned}
$$

We need to find a λ_a that maximizes $T^{\text{2-MPR}}$ for any value of λ. For $\lambda_a \leq n$,

$$\frac{\partial T^{\text{2-MPR}}}{\partial \lambda_a} \geq 0,$$

and thus, $T^{\text{2-MPR}}$ is increasing in λ_a for $\lambda_a \leq n$. Now, let $\lambda_a = n(1 + \delta)$ when $\delta > 0$ is arbitrary. Substituting $\lambda_a = n(1 + \delta)$ in (36) yields,

$$
\begin{aligned}
\frac{\partial T^{\text{2-MPR}}}{\partial \lambda_a} &= e^{-(1+\delta+\lambda_a/n)} \left[1 + \frac{\lambda + n(1+\delta)}{n} \left(1 - \frac{n(1+\delta)}{n} \right) \right] = \\
&= e^{-(1+\delta+\lambda_a/n)} \left[1 - \delta - \delta^2 - \delta(\frac{\lambda}{n}) \right].
\end{aligned}
$$

For $\lambda > (1 - \delta - \delta^2)\frac{n}{\delta}$,

$$\frac{\partial T^{\text{2-MPR}}}{\partial \lambda_a} < 0,$$

and thus, $T^{\text{2-MPR}}$ is decreasing in λ_a for $\lambda_a > n$. Hence, when there is no restriction on λ_a, $\lambda_a = n$ maximizes $T^{\text{2-MPR}}$.

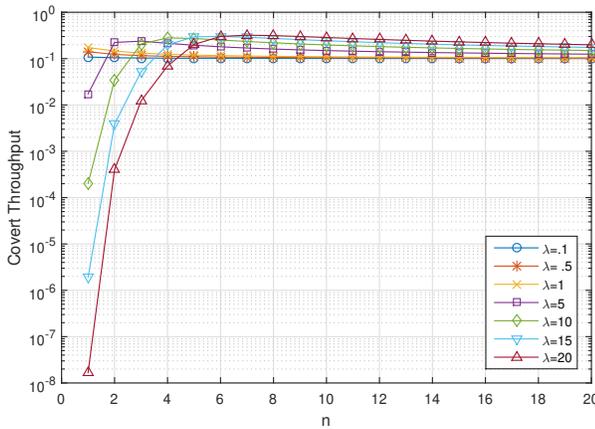Next, consider Willie with different detection capabilities:

Fig. 8: Covert throughput versus number of channels for different values of system nodes' transmission rate λ. Willie employs a 0-MPD detector and the intended receiver (Bob) employs a 2-MPR receiver.

Fig. 9: Covert throughput versus number of channels for different values of system nodes' transmission rate λ. Willie employs a 1-MPD detector and Bob employs a 2-MPR receiver.

5.2.1 Willie with 0-MPD receiver

From Theorem 2, if $\epsilon e^{\lambda/n} \geq n$, λ_a is not restricted by covertness considerations, and thus we can set $\lambda_a = n$. Substituting this into (36),

$$T_{0\text{-MPD}}^{2\text{-MPR}} = \frac{ne^{-\lambda/n}}{e}\left(2 + \frac{\lambda}{n}\right). \tag{37}$$

If $\epsilon e^{\lambda/n} \leq n$, covertness is maintained if $\lambda_a \leq \epsilon e^{\lambda/n}$. Eq. (35) is increasing in λ_a for $\lambda_a \leq n$. Hence, it is increasing for $\lambda_a \leq \epsilon e^{\lambda/n} \leq n$, and thus the maximum covert throughput is obtained when $\lambda_a = \epsilon e^{\lambda/n}$,

$$T_{0\text{-MPD}}^{2\text{-MPR}} = \epsilon e^{-\frac{1}{n}\epsilon e^{\lambda/n}}\left(1 + \frac{\lambda + \epsilon e^{\lambda/n}}{n}\right). \tag{38}$$

In Fig. 8, the covert throughput versus the number of channels for different values of system nodes transmission rate λ is shown. Similar to previous cases, for larger system nodes transmission rates, using more channels can help to improve the covert throughput. Further, comparing Fig. 5 and Fig. 8, when Bob employs a 2-MPR receiver, larger λ leads to achieving a higher covert throughput.

5.2.2 Willie with 1-MPD receiver

From Theorem 3, if ϵ and λ are such that

$$\frac{\epsilon e^{\lambda/n}}{n + \max\{\lambda, n - \lambda\}} \geq 1, \tag{39}$$

the communication is covert for any $\lambda_a > 0$. Hence, using (36), the covert rate $\lambda_a = n$ is achievable, and

$$T_{1\text{-MPD}}^{2\text{-MPR}} = \frac{ne^{-\lambda/n}}{e}\left(2 + \frac{\lambda}{n}\right). \tag{40}$$

If ϵ and λ are such that

$$\frac{\epsilon e^{\lambda/n}}{n + \max\{\lambda, n - \lambda\}} < 1, \tag{41}$$

the covert rate,

$$\lambda_a = \min\left\{n \ln\left(\frac{n + \max\{n - \lambda, \lambda\}}{n + \max\{n - \lambda, \lambda\} - \epsilon e^{\lambda/n}}\right), n\right\},$$
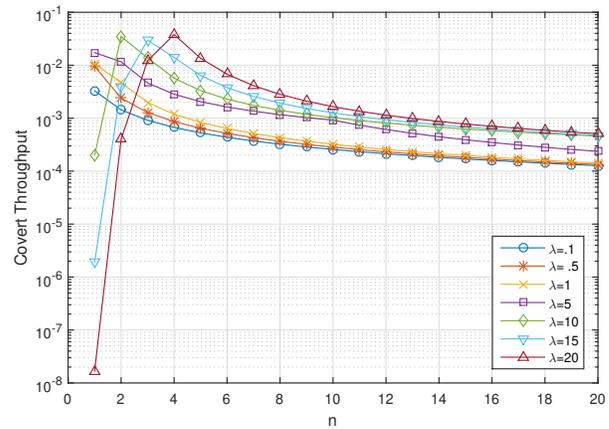
is achievable, and the covert throughput is obtained by substituting λ_a from the above equation into (35).

In Fig. 9, the curves of covert throughput versus number of channels for different system nodes' transmission rates are depicted. Similar to the case of Willie with 0-MPD capability, for small values of λ, increasing the number of channels does not help to improve the covert throughput, but as λ increases having more channels can help to improve the covert throughput.

5.2.3 Willie with ∞-MPD receiver

Using the same reasoning as in Section 5.1.3, when ϵ is sufficiently small such that $\lambda \leq \frac{1}{4\epsilon^2}$, the covert rate

$$\lambda_a = 2\epsilon\sqrt{\lambda}, \ \forall n \geq 1, \tag{42}$$

can be achieved. Hence, substituting (42) into (35),

$$T_{\infty\text{-MPD}}^{2\text{-MPR}} = 2\epsilon\sqrt{\lambda}e^{-(\lambda + 2\epsilon\sqrt{\lambda})/n}\left(1 + \frac{\lambda + 2\epsilon\sqrt{\lambda}}{n}\right). \tag{43}$$

In Fig. 10, covert throughput versus number of network channels for different system nodes' transmission rates are shown. The covertness factor $\epsilon = 0.1$, and thus (43) holds for $\lambda \leq 25$.

5.3 Bob and Willie with unbounded reception capabilities

Let us consider the extreme case that Bob is able to apply an ∞-MPR receiver, and Willie is able to apply an ∞-MPD detector. In this case, the expected number of covert packets that is transmitted over one channel in a time slot is $\frac{\lambda_a}{n}$, and thus the channel throughput is,

$$\tau^{\infty\text{-MPR}} = \frac{\lambda_a}{n}. \tag{44}$$

Thus, the throughput of the n channel network is,

$$T^{\infty\text{-MPR}} = \lambda_a. \tag{45}$$

The network throughput is monotonically increasing in λ_a, and thus a larger λ_a leads to a higher covert throughput.
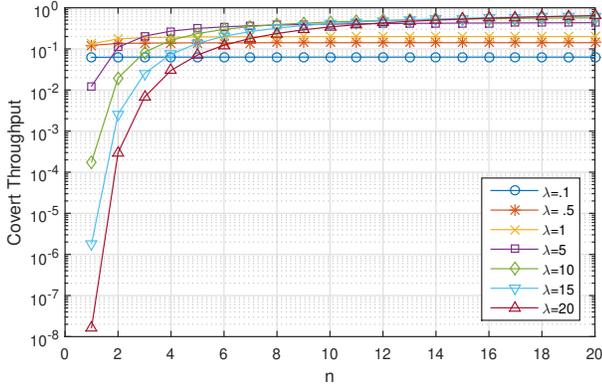
Fig. 10: Covert throughput versus the number of channels for different values of system nodes' transmission rate λ. In this case, Willie employs a ∞-MPD detector and Bob employs a 2-MPR receiver.

From Theorem 7, when Willie has ∞-MPD detection capability, in order to maintain covertness λ_a is bounded as,

$$\lambda_a \leq 2\epsilon\sqrt{\lambda}.$$

Hence,

$$T^{\infty\text{-MPR}}_{\infty\text{-MPD}} = 2\epsilon\sqrt{\lambda}. \tag{46}$$

This shows that when both Bob and Willie have unbounded reception capabilities, there is a square root relationship between the covert throughput and the transmission rate of the system nodes, and the covert throughput does not depend on the number of channels.

6 CONCLUSIONS AND FUTURE WORK

The fundamental limits of covert communications have been considered extensively in recent years at the physical layer for the scenario of covert transmitter Alice, receiver Bob, and capable and attentive warden Willie who attempts to detect Alice. Here, we consider for the first time the medium access control (MAC), where a number of covert users are attempting to access the channel without detection by warden Willie, who is observing the channel collision process. We consider a variety of receivers at Willie, ranging from one that can only determine whether the channel was idle or busy (0-MPD), to one that always knows the number of packets involved in a collision (∞-MPD). In the latter case, the results follow much of what has been found at the physical layer, where the rate of the covert users is restricted roughly to the square root of the rate of the system users. However, for a K-MPD detector, $K < \infty$, the throughput grows much faster than the square root of λ, thus indicating the degree to which Willie's blindness to the channel state allows for covert transmission.

Our results also reveal that, while the achievable covert rate generally increases as the system traffic rate increases, the covert throughput has a more subtle relationship with the system traffic rate. Specifically, depending on the number of channels, there exists a traffic rate that maximizes the covert throughput. This finding can be used to design end-to-end covert communication strategies in *network* scenarios

where Alice and Bob can choose a set of intermediary relays across the network to help them convey their messages covertly. By strategically choosing those relays that are located optimally, with respect to the system traffic rate in their neighborhood and considering Willies' receiver capabilities, a covert routing algorithm can maximize end-to-end covert throughput for Alice and Bob. In the current paper, we were able to derive the covert rate when Willie has a K-MPD receiver only for $\lambda \geq nK$. The general covert rate and covert throughput analysis (for any λ) can be considered in future research.

APPENDIX A
PROOF OF THEOREM 3

When Willie has 1-MPD capability, he can detect an idle channel ($s = s_0$), a single packet transmission ($s = s_1$), or more than one packet transmission ($s = s_2$). Hence,

$$
\begin{aligned}
\mathbb{P}_0\{s = s_0\} &= e^{-\lambda/n}, \\
\mathbb{P}_0\{s = s_1\} &= \frac{\lambda}{n}e^{-\lambda/n}, \\
\mathbb{P}_0\{s = s_2\} &= 1 - \left(\mathbb{P}_0\{s = s_0\} + \mathbb{P}_0\{s = s_1\}\right), \tag{47}
\end{aligned}
$$

and,

$$
\begin{aligned}
\mathbb{P}_1\{s = s_0\} &= e^{-(\lambda+\lambda_a)/n}, \\
\mathbb{P}_1\{s = s_1\} &= \frac{\lambda + \lambda_a}{n}e^{-(\lambda+\lambda_a)/n}, \\
\mathbb{P}_1\{s = s_2\} &= 1 - \left(\mathbb{P}_1\{s = s_0\} + \mathbb{P}_1\{s = s_1\}\right). \tag{48}
\end{aligned}
$$

Let $\mathcal{S} = \{s_0, s_1, s_2\}$ denote the set of states of each process. The total variation distance between $\mathbb{P}_0$ and $\mathbb{P}_1$ is,

$$
\begin{aligned}
d_{TV}(\mathbb{P}_0, \mathbb{P}_1) &= \frac{1}{2}\sum_{s\in\mathcal{S}}|\mathbb{P}_1(s) - \mathbb{P}_0(s)| \\
&= \frac{1}{2}\left(\left|e^{-\lambda/n} - e^{-(\lambda+\lambda_a)/n}\right| + \left|\frac{\lambda}{n}e^{-\lambda/n} - \frac{\lambda+\lambda_a}{n}e^{-(\lambda+\lambda_a)/n}\right| \right.\\
&\quad \left. + \left|1 - (1+\frac{\lambda}{n})e^{-\lambda/n} - 1 + (1+\frac{\lambda+\lambda_a}{n})e^{-(\lambda+\lambda_a)/n}\right|\right) \\
&= \frac{1}{2}\left(\left|e^{-\lambda/n} - e^{-(\lambda+\lambda_a)/n}\right| + \left|\frac{\lambda}{n}e^{-\lambda/n} - \frac{\lambda+\lambda_a}{n}e^{-(\lambda+\lambda_a)/n}\right| \right.\\
&\quad \left. + \left|(1+\frac{\lambda}{n})e^{-\lambda/n} - (1+\frac{\lambda+\lambda_a}{n})e^{-(\lambda+\lambda_a)/n}\right|\right) \\
&\leq \left|e^{-\lambda/n} - e^{-(\lambda+\lambda_a)/n}\right| + \left|\frac{\lambda}{n}e^{-\lambda/n} - \frac{\lambda+\lambda_a}{n}e^{-(\lambda+\lambda_a)/n}\right| \\
&= e^{-\lambda/n}(1 - e^{-\lambda_a/n}) + e^{-\lambda/n}\left|\frac{\lambda}{n}(1 - e^{-\lambda_a/n}) - \frac{\lambda_a}{n}e^{-\lambda_a/n}\right|,
\end{aligned}
$$

where the inequality is the triangle inequality. We know $|x| \leq \max\{x, -x\}$. Hence,

$$
\begin{aligned}
d_{TV}(\mathbb{P}_0, \mathbb{P}_1) &\leq e^{-\lambda/n}(1 - e^{-\lambda_a/n}) \\
&+ e^{-\lambda/n}\max\left\{\frac{\lambda}{n}(1 - e^{-\lambda_a/n}) - \frac{\lambda_a}{n}e^{-\lambda_a/n},\right. \\
&\quad \left. -\frac{\lambda}{n}(1 - e^{-\lambda_a/n}) + \frac{\lambda_a}{n}e^{-\lambda_a/n}\right\} \\
&\leq e^{-\lambda/n}(1 - e^{-\lambda_a/n}) + e^{-\lambda/n}\max\left\{\frac{\lambda}{n}(1 - e^{-\lambda_a/n}),\right. \\
&\quad \left. -\frac{\lambda}{n}(1 - e^{-\lambda_a/n}) + 1 - e^{-\lambda_a/n}\right\} \\
&= e^{-\lambda/n}(1 - e^{-\lambda_a/n}) + e^{-\lambda/n}(1 - e^{-\lambda_a/n})\max\left\{\frac{\lambda}{n}, 1 - \frac{\lambda}{n}\right\},
\end{aligned}
$$

where the second inequality follows from the fact that for $\lambda_a \geq 0$, $\frac{\lambda_a}{n} e^{-\lambda_a/n} \geq 0$, and $\frac{\lambda_a}{n} e^{-\lambda_a/n} \leq 1 - e^{-\lambda_a/n}$. The equality follows because $1 - e^{-\lambda_a/n} \geq 0$. Therefore, the covertness condition is satisfied if

$$d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n) \leq n e^{-\lambda/n}(1 - e^{-\lambda_a/n})(1 + \max\{\frac{\lambda}{n}, 1 - \frac{\lambda}{n}\}) \leq \epsilon. \tag{49}$$

The term $1 - e^{-\lambda_a/n} \leq 1$ for any $\lambda_a \geq 0$. Hence, if

$$\frac{\epsilon e^{\lambda/n}}{1 + \max\{\frac{\lambda}{n}, 1 - \frac{\lambda}{n}\}} \geq n, \tag{50}$$

for any $\lambda_a \geq 0$ transmission is covert. Otherwise, from (49) transmission is covert if,

$$\lambda_a \leq -n \ln\left(1 - \frac{\epsilon e^{\lambda/n}}{n + \max\{n - \lambda, \lambda\}}\right). \tag{51}$$

APPENDIX B
PROOF OF THEOREM 4

In this case, suppose Willie is able to detect up to K simultaneous packet transmissions. If more than K packets are transmitted, then Willie will only detect a collision event. Thus, the network state observed by Willie is one of $K + 2$ states denoted by $s_0, s_1, \cdots, s_K, s_{K+1}$, where s_i, $i = 0, 1, \ldots, K$, indicates that Willie detected i concurrent transmissions. Let $\mathcal{S} = \{s_0, s_1, \cdots, s_{K+1}\}$ denote the set of states of each channel process. Hence,

$$d_{TV}(\mathbb{P}_0, \mathbb{P}_1) = \frac{1}{2} \sum_{s \in \mathcal{S}} |\mathbb{P}_1(s) - \mathbb{P}_0(s)|$$

$$= \frac{1}{2}\left(\sum_{k=0}^{K}\left|\frac{(\frac{\lambda}{n})^k e^{-\lambda/n}}{k!} - \frac{(\frac{\lambda+\lambda_a}{n})^k e^{-(\lambda+\lambda_a)/n}}{k!}\right|\right.$$

$$\left. + \left|1 - \sum_{k=0}^{K}\frac{(\frac{\lambda}{n})^k e^{-\lambda/n}}{k!} - 1 + \sum_{k=0}^{K}\frac{(\frac{\lambda+\lambda_a}{n})^k e^{-(\lambda+\lambda_a)/n}}{k!}\right|\right)$$

$$\leq \sum_{k=0}^{K} \frac{e^{-\lambda/n}}{k!}\left|(\frac{\lambda}{n})^k - (\frac{\lambda+\lambda_a}{n})^k e^{-\lambda_a/n}\right|, \tag{52}$$

where the inequality is the triangle inequality. Because of the absolute value in (52), it is hard to find an upper-bound for $d_{TV}(\mathbb{P}_0, \mathbb{P}_1)$ for arbitrary λ. However, when $\frac{\lambda}{n} \geq K$, the term in the absolute value is greater than zero for any $\lambda_a \geq 0$, which makes the analysis of (52) easier. In this case,

$$d_{TV}(\mathbb{P}_0, \mathbb{P}_1) \leq \sum_{k=0}^{K} \frac{e^{-\lambda/n}}{k!}\left|\left(\frac{\lambda}{n}\right)^k - \left(\frac{\lambda+\lambda_a}{n}\right)^k e^{-\lambda_a/n}\right|$$

$$= \sum_{k=0}^{K} \frac{e^{-\lambda/n}}{k!}\left(\left(\frac{\lambda}{n}\right)^k - \left(\frac{\lambda+\lambda_a}{n}\right)^k e^{-\lambda_a/n}\right)$$

$$\leq \sum_{k=0}^{K} \frac{e^{-\lambda/n}}{k!}\left(\left(\frac{\lambda}{n}\right)^k - \left(\frac{\lambda}{n}\right)^k e^{-\lambda_a/n}\right)$$

$$= \sum_{k=0}^{K} \frac{e^{-\lambda/n}}{k!}\left(\frac{\lambda}{n}\right)^k (1 - e^{-\lambda_a/n}). \tag{53}$$

Thus, if the following relation holds,

$$d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n) \leq n \sum_{k=0}^{K} \frac{e^{-\lambda/n}}{k!}\left(\frac{\lambda}{n}\right)^k (1 - e^{\lambda_a/n}) \leq \epsilon, \tag{54}$$

then covertness is maintained. Using the same reasoning as in Theorems 2 and 3, the statement is proved.

APPENDIX C
PROOF OF THEOREM 5

An ∞-MPD Willie can determine the number of packets being transmitted simultaneously over the channel. In this case, the probability distributions of the channel states observed by Willie are given by two Poisson probability distributions: $\mathbb{P}_0 = Poisson(\frac{\lambda}{n})$, and $\mathbb{P}_1 = Poisson(\frac{\lambda+\lambda_a}{n})$. Using the upper-bound for total variation distance between two Poisson distributions in [44], it then follows that

$$d_{TV}(\mathbb{P}_0, \mathbb{P}_1) \leq \min\left\{\frac{\lambda_a}{n}, \sqrt{\frac{2}{e}}\left(\sqrt{\frac{\lambda+\lambda_a}{n}} - \sqrt{\frac{\lambda}{n}}\right)\right\}. \tag{55}$$

Consequently, the covertness is maintained if,

$$d_{TV}(\mathbb{P}_0^n, \mathbb{P}_1^n) \leq \min\left\{\lambda_a, \sqrt{\frac{2n}{e}}\left(\sqrt{\lambda+\lambda_a} - \sqrt{\lambda}\right)\right\} \leq \epsilon.$$

After some algebraic manipulations, the statement of the theorem follows.

APPENDIX D
PROOF OF THEOREM 6

Similar to Theorem 2, Willie can only determine if the channel is busy or not. If the channel is busy, Willie is not able to determine how many packets are being transmitted in a time slot. Using the probability distributions from (12),

$$d_{KL}(\mathbb{P}_0, \mathbb{P}_1) = \sum_{s \in \mathcal{S}} \mathbb{P}_1(s) \ln\left(\frac{\mathbb{P}_1(s)}{\mathbb{P}_0(s)}\right)$$

$$= e^{-(\lambda+\lambda_a)/n} \ln \frac{e^{-(\lambda+\lambda_a)/n}}{e^{-\lambda/n}}$$

$$\quad + \left(1 - e^{-(\lambda+\lambda_a)/n}\right)\ln\frac{1 - e^{-(\lambda+\lambda_a)/n}}{1 - e^{-\lambda/n}}$$

$$= -\frac{\lambda_a}{n}e^{-(\lambda+\lambda_a)/n} + \left(1 - e^{-(\lambda+\lambda_a)/n}\right)\ln\frac{1 - e^{-(\lambda+\lambda_a)/n}}{1 - e^{-\lambda/n}}. \tag{56}$$

The logarithmic term can be written as,

$$\ln\left(\frac{1 - e^{-(\lambda+\lambda_a)/n}}{1 - e^{-\lambda/n}}\right) = \ln\left(1 + e^{-\lambda/n}\frac{1 - e^{-\lambda_a/n}}{1 - e^{-\lambda/n}}\right). \tag{57}$$

Using the $\ln(1+x) \leq x$ for $x \geq 0$ results in the following relation,

$$d_{KL}(\mathbb{P}_0, \mathbb{P}_1)$$

$$\leq -\frac{\lambda_a}{n}e^{-(\lambda+\lambda_a)/n} + \left(1 - e^{-(\lambda+\lambda_a)/n}\right)e^{-\lambda/n}\frac{1 - e^{-\lambda_a/n}}{1 - e^{-\lambda/n}}$$

$$\leq e^{-\lambda/n}\frac{(1 - e^{-(\lambda+\lambda_a)/n})(1 - e^{-\lambda_a/n})}{1 - e^{-\lambda/n}}$$

$$\leq \frac{1 - e^{-\lambda_a/n}}{e^{\lambda/n} - 1}. \tag{58}$$

Therefore, to upper bound $d_{KL}(\mathbb{P}_0^n, \mathbb{P}_1^n)$ to satisfy the covertness constraint of (11) with some $\epsilon \geq 0$, the following relation should be satisfied,

$$\frac{1 - e^{-\lambda_a/n}}{e^{\lambda/n} - 1} \leq \frac{2\epsilon^2}{n}, \tag{59}$$

or equivalently,

$$1 - e^{-\lambda_a/n} \leq \frac{2\epsilon^2}{n}(e^{\lambda/n} - 1). \tag{60}$$

This condition holds for any $\lambda_a \geq 0$ if the right side of (60) is not less than one, i.e. λ, n, and ϵ are such that,

$$\lambda \geq n \ln \left(\frac{n}{2\epsilon^2} + 1 \right). \tag{61}$$

Otherwise, if (61) does not hold, the covert rates

$$\lambda_a \leq n \ln \left(\frac{1}{1 - \frac{2\epsilon^2}{n}(e^{\lambda/n} - 1)} \right), \tag{62}$$

are achievable.

APPENDIX E
PROOF OF THEOREM 7

Suppose Willie has ∞-MPD capability, i.e. he can determine how many packets are being transmitted simultaneously over the channel. In this case, the probability distributions of the channel states observed by Willie are given by two Poisson probability distributions: $\mathbb{P}_0 = Poisson(\frac{\lambda}{n})$, and $\mathbb{P}_1 = Poisson(\frac{\lambda+\lambda_a}{n})$. It then follows that

$$\begin{aligned} d_{KL}(\mathbb{P}_0, \mathbb{P}_1) &= \frac{\lambda_a}{n} - \frac{\lambda}{n} \ln \left(1 + \frac{\lambda_a}{\lambda} \right) \\ &\leq \frac{\lambda_a}{n} - \frac{\lambda}{n} \left(\frac{\lambda_a}{\lambda} - \frac{1}{2} \left(\frac{\lambda_a}{\lambda} \right)^2 \right) \\ &= \frac{\lambda_a^2}{2\lambda n}, \end{aligned} \tag{63}$$

where the inequality is true since $\frac{\lambda_a}{\lambda} \geq 0$. Therefore, to maintain covertness, i.e.

$$d_{KL}(\mathbb{P}_0^n, \mathbb{P}_1^n) = n d_{KL}(\mathbb{P}_0, \mathbb{P}_1) \leq \frac{\lambda_a^2}{2\lambda} \leq 2\epsilon^2, \tag{64}$$

λ_a is bounded as,

$$\lambda_a \leq 2\epsilon\sqrt{\lambda}. \tag{65}$$

REFERENCES

[1] A. Sheikholeslami, M. Ghaderi, and D. Goeckel, "Covert communications in packet collision channels," in *Proc. IEEE WCNC*, April 2019.

[2] S. Zander, G. Armitage, and P. Branch, "A survey of covert channels and countermeasures in computer network protocols," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 3, pp. 44–57, 2007.

[3] B. Carrara and C. Adams, "Out-of-band covert channels—a survey," *ACM Computing Surveys (CSUR)*, vol. 49, no. 2, pp. 1–36, 2016.

[4] S. Cabuk, C. E. Brodley, and C. Shields, "Ip covert timing channels: design and detection," in *Proc. ACM conference on Computer and communications security*, 2004, pp. 178–187.

[5] S. Gianvecchio and H. Wang, "Detecting covert timing channels: an entropy-based approach," in *Proc. ACM conference on computer and communications security*, 2007, pp. 307–316.

[6] ——, "An entropy-based approach to detecting covert timing channels," *IEEE Transactions on Dependable and Secure Computing*, vol. 8, no. 6, pp. 785–797, 2010.

[7] M. Wolf, "Covert channels in lan protocols," in *Local Area Network Security Workshop*. Springer, 1989, pp. 89–101.

[8] R. A. Kemmerer and P. A. Porras, "Covert flow trees: A visual approach to analyzing covert storage channels," *IEEE Transactions on Software Engineering*, vol. 17, no. 11, p. 1166, 1991.

[9] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage, "Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds," in *Proc. of ACM conference on Computer and communications security*, 2009, pp. 199–212.

[10] A. Lemay, J. M. Fernandez, and S. Knight, "A modbus command and control channel," in *2016 Annual IEEE Systems Conference (SysCon)*, 2016, pp. 1–6.

[11] C. Alcaraz, G. Bernieri, F. Pascucci, J. Lopez, and R. Setola, "Covert channels-based stealth attacks in industry 4.0," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3980–3988, 2019.

[12] V. Korzhik, G. Morales-Luna, and M. H. Lee, "On the existence of perfect stegosystems," in *Proc. 4th IWDW*, Siena, Italy, Sep. 2005, pp. 30–38.

[13] B. Bash, D. Goeckel, and D. Towsley, "Square root law for communication with low probability of detection on AWGN channels," in *Proc. IEEE ISIT*, July 2012, pp. 448–452.

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

[15] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: Hiding messages in noise," in *Proc. IEEE ISIT*, 2013, pp. 2945–2949.

[16] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2334–2354, 2016.

[17] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3493–3503, 2016.

[18] K. S. K. Arumugam and M. R. Bloch, "Covert communication over a k-user multiple-access channel," *IEEE Transactions on Information Theory*, vol. 65, no. 11, pp. 7020–7044, 2019.

[19] Q. Zhang, M. Bakshi, and S. Jaggi, "Computationally efficient deniable communication," in *Proc. IEEE ISIT*, July 2016, pp. 2234–2238.

[20] G. Frèche, M. R. Bloch, and M. Barret, "Polar codes for covert communications over asynchronous discrete memoryless channels," in *51st Annual CISS*, March 2017, pp. 1–1.

[21] M. R. Bloch and S. Guha, "Optimal covert communications using pulse-position modulation," in *IEEE ISIT*, June 2017, pp. 2825–2829.

[22] I. A. Kadampot, M. Tahmasbi, and M. Bloch, "Multilevel-coded pulse-position modulation for covert communications," in *Proc. IEEE ISIT*, June 2018.

[23] B. A. Bash, D. Goeckel, and D. Towsley, "Covert communication gains from adversary's ignorance of transmission time," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8394–8405, 2016.

[24] S. Lee and R. Baxley, "Achieving positive rate with undetectable communication over AWGN and Rayleigh channels," in *Proc. IEEE ICC*, June 2014, pp. 780–785.

[25] S. Lee, R. J. Baxley, M. A. Weitnauer, and B. Walkenhorst, "Achieving undetectable communication," *IEEE Journal on Selected Topics in Signal Processing*, vol. 9, no. 7, pp. 1195–1205, Oct 2015.

[26] Z. Liu, J. Liu, Y. Zeng, J. Ma, and Q. Huang, "On covert communication with interference uncertainty," in *Proc. IEEE ICC*, 2018, pp. 1–6.

[27] D. Goeckel, B. Bash, S. Guha, and D. Towsley, "Covert communications when the warden does not know the background noise power," *IEEE Communications Letters*, vol. 20, no. 2, pp. 236–239, 2016.

[28] T. V. Sobers, B. A. Bash, S. Guha, D. Towsley, and D. Goeckel, "Covert communication in the presence of an uninformed jammer," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6193–6206, 2017.

[29] B. A. Bash, A. H. Gheorghe, M. Patel, J. L. Habif, D. Goeckel, D. Towsley, and S. Guha, "Quantum-secure covert communication on bosonic channels," *Nature communications*, vol. 6, 2015.

[30] R. Soltani, D. Goeckel, D. Towsley, and A. Houmansadr, "Fundamental limits of covert packet insertion," *CoRR*, vol. abs/1903.11640, 2019. [Online]. Available: http://arxiv.org/abs/1903.11640

[31] J. Hu, S. Yan, X. Zhou, F. Shu, and J. Li, "Covert wireless communications with channel inversion power control in Rayleigh fading," *CoRR*, vol. abs/1803.07812, 2018. [Online]. Available: http://arxiv.org/abs/1803.07812

[32] K. Shahzad, X. Zhou, and S. Yan, "Covert communication in fading channels under channel uncertainty," in *IEEE 85th VTC Spring*, June 2017, pp. 1–5.

[33] S. Yan, B. He, X. Zhou, Y. Cong, and A. L. Swindlehurst, "Delay-intolerant covert communications with either fixed or random transmit power," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 1, pp. 129–140, Jan 2019.

[34] W. Hoeffding and J. Wolfowitz, "Distinguishability of sets of distributions," *The Annals of Mathematical Statistics*, vol. 29, no. 3, pp. 700–718, 1958.

[35] J. G. Andrews, "Interference cancellation for cellular systems: a contemporary overview," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 19–29, 2005.

[36] E. Casini, R. De Gaudenzi, and O. D. R. Herrero, "Contention resolution diversity slotted ALOHA (CRDSA): An enhanced random access schemefor satellite access packet networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1408–1419, 2007.

[37] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted ALOHA," *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 477–487, 2010.

[38] E. Paolini, G. Liva, and M. Chiani, "Coded slotted aloha: A graph-based method for uncoordinated multiple access," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6815–6832, 2015.

[39] F. Babich and M. Comisso, "Impact of segmentation and capture on slotted ALOHA systems exploiting interference cancellation," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2878–2892, 2019.

[40] L. Tong, Q. Zhao, and G. Mergen, "Multipacket reception in random access wireless networks: From signal processing to optimal medium access control," *IEEE Communications Magazine*, vol. 39, no. 11, pp. 108–112, 2001.

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

[42] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*. Springer Science & Business Media, 2006.

[43] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[44] P. Ruzankin, "On the rate of Poisson process approximation to a Bernoulli process," *Journal of Applied Probability*, vol. 41, no. 1, pp. 271–276, 2004.