Multi-Hop Routing in Covert Wireless Networks

Azadeh Sheikholeslami*, Majid Ghaderi†, Don Towsley‡, Boulat A. Bash§, Saikat Guha¶, Dennis Goeckel*
*Dept. of Elec. and Comp. Engineering (ECE), Univ. of Massachusetts, Amherst
† Department of Computer Science , University of Calgary
‡College of Info. and Comp. Sciences (CICS), Univ. of Massachusetts, Amherst
§ Raytheon BBN Technologies, Cambridge, MA
¶ The Univ. of Arizona, College of Optical Sciences, Tucson, AZ

***Abstract*—In covert communication, Alice tries to communicate with Bob without being detected by a warden Willie. When the distance between Alice and Bob becomes large compared to the distance between Alice and Willie(s), the performance of covert communication will be degraded. In this case, multi-hop message transmission via intermediate relays can help to improve performance. Hence, in this work multi-hop covert communication over a moderate size network and in the presence of multiple collaborating Willies is considered. The relays can transmit covertly using either a single key for all relays, or different independent keys at the relays. For each case, we develop efficient algorithms to find optimal paths with maximum throughput and minimum end-to-end delay between Alice and Bob. As expected, employing multiple hops significantly improves the ability to communicate covertly versus the case of a single-hop transmission. Furthermore, at the expense of more shared key bits, analytical results and numerical simulations demonstrate that multi-hop covert communication with different independent keys at the relays has better performance than multi-hop covert communication with a single key.**

I. Introduction

Due to the broadcast nature of wireless networks, any node near a transmitter can overhear the message. Thus, providing security for wireless communications is of central importance and has attracted particular attention. Various security schemes have been developed to protect the content of a message from an unintended recipient [1]–[6]; however, there are security scenarios where the *existence* of a transmission (or the transmitter) is to be kept hidden from adversaries. In such adversarial scenarios, traditional security approaches are no longer effective, and the communicating parties should seek low probability of detection approaches, which have been studied recently and termed "covert communication" [7]–[17]. Consider a wireless communication scenario when Alice (the transmitter) wants to send a message to Bob (the intended receiver) such that an attentive adversary Willie is not aware of the transmission. In [7], it is shown that using a pre-shared key between Alice and Bob, it is possible to transmit $\mathcal{O}(\sqrt{n})$ information bits covertly over n channel uses such that Willie is not aware of the existence of communication. Moreover, it is not possible to transmit $\omega(\sqrt{n})$ bits over n channel uses covertly: if the transmitter transmits $\omega(\sqrt{n})$ bits either Willie can detect the communication, or Bob will not be able to decode the message with a (arbitrarily) low probability of error.

In [10], [12] the constant in front of $\sqrt{n}$ for the number of bits transmitted covertly over memoryless channels is characterized. It is shown that the number of bits that can go through the channel without being detected by Willie has a direct relationship with the distance between the probability distribution functions of the received signals at Bob when no communication occurs and when Alice is transmitting. Also, it has an inverse relationship with the distance between the probability distribution functions of the received signals at Willie when no communication occurs and when Alice is transmitting [10], [12].

In an environment with AWGN channels when Alice and Bob are located far from each other, in order to make the probability of error at Bob sufficiently small, Alice should use a high transmit

This work was sponsored by the National Science Foundation (NSF) under grants ECCS-1309573 and CNS-1564067, and DARPA under contract number HR0011-16-C-0111.

power. However, this increases the probability of being detected by Willie, especially if Willie is close to Alice and thus receives a strong signal, and/or if multiple collaborating Willies are present and try to detect any transmission. In order to solve this problem, in [9] Alice and Bob use artificial noise from friendly chatterers to increase the noise level of the wireless environment to help them hide their communication even when Willie is close to Alice and when multiple collaborating Willies are present. However, this approach requires some friendly system nodes in the network who are not concerned to transmit openly and reveal their existence and/or their locations. Hence, when all system nodes prefer to hide their existence and/or their locations, the scheme of [9] cannot be used. In this case, in order to facilitate covert communication, we propose utilizing the friendly system nodes to establish a multi-hop path from Alice to Bob. On this path, the distance between intermediate relays is short and thus each relay can transmit with a small transmit power in order to decrease its probability of being detected. Also, the multi-hop path from Alice to Bob can take detours to avoid Willies. That is, the routing algorithm can be designed so that it chooses relays that are less susceptible to being detected by Willies.

In this paper, we consider multi-hop covert communication between Alice and Bob in the presence of multiple Willies. In order to consider the most powerful adversary scenario, we assume all Willies are collaborating to detect any transmission of Alice and the intermediate relays. A message generated by Alice travels hop-by-hop until it is delivered to Bob. For covert communication, Alice and the intermediate relays use a key to encode the message. We consider two scenarios. In the first scenario, we consider the case that a single key is used by Alice and all relays at all hops to encode the message. While this approach is simple and does not require separate keys at each hop, it can increase the probability of being detected because the exact same codeword is transmitted over multiple links and is observed by Willies. In the second scenario, we consider employing independent keys at the relays. Each relay encodes the received message with an independent key and then forwards it to the next relay. In this case, independent codewords are transmitted over different links, and thus the signals being observed by the Willies at different hops are independent.

We consider two performance metrics, namely, throughput and end-to-end delay over the path from Alice to Bob. We develop algorithms to find the path with the maximum throughput and the path with minimum end-to-end delay between Alice and Bob for the case of a single key and the case of independent keys at the relays. We compare the performances of all algorithms numerically as we vary the network parameters.

The rest of this paper is organized as follows. In Section II a short summary of covert communications, the system model, the covertness criteria, and the multi-hop strategies used in this work are explained. In Sections III and IV, multi-hop covert communication with a single key and with independent keys at the relays are considered, respectively, and for each case algorithms to establish an optimal path from Alice to Bob are proposed. The proposed algorithms are studied and compared numerically in Section V. Concluding remarks are discussed in Section VI.

II. Prerequisites

A. Covert Communication or Communication with Low Probability of Detection

Consider a transmitter Alice, a receiver Bob, and a warden Willie. Alice wants to transmit a message to Bob such that Willie is not aware of the communication. Willie uses his observations of the channel to detect whether Alice transmits or not. Suppose H_1 is the hypothesis that Alice transmits a signal, and H_0 is the hypothesis that no communication occurs. Willie's probability of detection error consists of two components: the probability of missed detection (Willie declares no communication when Alice transmits) denoted by $\mathbb{P}_{MD}^{W} = \mathbb{P}(H_0|H_1 \text{ is correct})$, and the probability of false alarm (Willie declares communication when no communication takes place) denoted by $\mathbb{P}_{FA}^{W} = \mathbb{P}(H_1|H_0 \text{ is correct})$. Hence, considering equal prior probabilities, the total detection error of Willie is:

$$\mathbb{P}_{e}^{W} = \frac{\mathbb{P}_{FA}^{W} + \mathbb{P}_{MD}^{W}}{2}. \tag{1}$$

In covert communication, the goal is to prevent Willie from using his observations of the channel to make the probability of detection error $\mathbb{P}_{e}^{W}$ arbitrarily small. In order to reach this goal, Alice

and Bob pre-share a secret key, based on which Alice selects a codebook from an ensemble of codebooks. Assume that the channel between Alice and Willie experiences some sort of uncertainty (e.g. it is an AWGN channel). The codebooks that Alice chooses from are low power codebooks such that Willie, without knowing the key, cannot decide with arbitrarily low probability of detection error that whether his observation is a signal transmitted by Alice or a result of the uncertainty of the channel. In [7], it is shown that the power of the signal transmitted over n channel uses should be of order of $\frac{1}{\sqrt{n}}$, which allows transmission of $\mathcal{O}(\sqrt{n})$ covert bits over n channel uses. Note that unlike conventional communication, in covert communication throughput changes with the number of channel uses n, and is on the order of $\frac{1}{\sqrt{n}}$. In [10], it is shown that for covert communication the number of key bits shared between Alice and Bob should be on the order of $\sqrt{n}$, and using this key, Bob can decode the message with arbitrarily low probability of error.

B. System Model

We consider a wireless network that consists of multiple (friendly) system nodes which are distributed arbitrarily. The set of (friendly) system nodes is denoted by $\mathcal{T} = \{T_1, \ldots, T_N\}$, where N is the number of such nodes in the network. In addition to the system nodes, multiple collaborating Willies, i.e. the wardens that want to detect any communication in the network, are present. The set of Willies is denoted by $\mathcal{W} = \{W_1, \ldots, W_M\}$, where M is the number of Willies. Willies collaborate and use all of their observations (across Willies, across transmissions and across time) to attempt to determine whether any of the system nodes transmitted or not.

The channel between nodes is an additive white Gaussian noise (AWGN) channel with path-loss exponent α, where $\alpha = 2$ corresponds to free space, and $\alpha > 2$ corresponds to a terrestrial environment. Any transmitter X in the network attempts to transmit a message by employing a Gaussian codebook [7], and its transmitted signal is given by $[f_1, f_2, \ldots, f_n]$, where $f_j \sim \mathcal{N}(0, 1)$ and n is the length of each codeword. The signal that a receiver Y receives is,

$$Z_j^{(Y)} = \frac{\sqrt{P_X} f_j}{d_{X,Y}^{\alpha/2}} + N_j^{(Y)},\ j = 1, \ldots, n, \qquad (2)$$

where P_X is the transmit power of node X, $d_{X,Y}$ is the distance between transmitter X and receiver Y, and $N_j^{(Y)} \sim \mathcal{N}(0, \sigma_Y^2)$ is AWGN at the receiver. The signal that Willie W_k receives is,

$$Z_j^{(W_k)} = \frac{\sqrt{P_X} f_j}{d_{X,W_k}^{\alpha/2}} + N_j^{(W_k)},\ j = 1, \ldots, n, \qquad (3)$$

where d_{X,W_k} is the distance between the transmitter X and Willie W_k, and $N_j^{(W_k)} \sim \mathcal{N}(0, \sigma_{W_k}^2)$ is AWGN at Willie W_k. Throughout this paper, it is assumed that the distances between the system nodes and the distances between the system nodes and the Willies are known to the system nodes and the Willies. In the case that the knowledge of the locations of the Willies is not complete, we can use lower bounds on the distances between the transmitters and the Willies to obtain bounds on the allowable transmit powers such that Willies cannot detect the communication (similar to the analysis presented in [15]).

C. Covertness Criteria and Covert Throughput

A transmission in the presence of Willies is considered covert when for any $\epsilon > 0$, the sum of probabilities of detection errors of their joint decision is lower bounded as,

$$\mathbb{P}_{FA}^W + \mathbb{P}_{MD}^W \geq 1 - \epsilon, \qquad (4)$$

for sufficiently large n (recall that n is the length of the codewords) [7]. The joint probability distribution function of Willies' observations when a transmission occurs is given by $\mathbb{Q}_1$, and the joint probability distribution function of Willies' observations when no transmission occurs is given by $\mathbb{Q}_0$. Suppose Willies perform the optimal test. Thus, using Pinsker's inequality [18], [19],

$$\mathbb{P}_{FA}^W + \mathbb{P}_{MD}^W \geq 1 - \sqrt{\frac{1}{2}\mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0)}, \qquad (5)$$

where $\mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0)$ is the relative entropy between $\mathbb{Q}_1$ and $\mathbb{Q}_0$. Hence, combining (4) and (5), an alternative covertness criteria is to bound the relative entropy:

$$\mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0) \leq \delta, \qquad (6)$$

where $\delta = 2\epsilon^2$. That is, if $\mathbb{Q}_1$ and $\mathbb{Q}_0$ are such that $\mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0) \leq \delta$, it is guaranteed that the communication is covert, i.e. $\mathbb{P}_{FA}^W + \mathbb{P}_{MD}^W \geq 1 - \epsilon$. In this paper, we consider (6) as our covertness criteria.

A transmission from a transmitter X to a receiver Y is considered reliable if as the block-length n goes to infinity, the average error probability of receiving a message at receiver Y approaches zero. We define *covert throughput* as the rate of reliable communication between a transmitter X and a receiver Y such that the communication is hidden from warden Willies. In this paper, we use the terms "throughput" and "covert throughput" interchangeably. For an AWGN channel, it has been shown that if the transmitter uses zero mean Gaussian input symbols with average power P_X, any covert throughput less than

$$C = \frac{1}{2}\log\left(1 + \frac{P_X}{\sigma_Y^2 d_{X,Y}^{\alpha}}\right), \tag{7}$$

can be achieved reliably [19]. Note that C depends on the Willies' distances to the transmitter through P_X. As mentioned before, in order to guarantee covertness we should have $P_X = \mathcal{O}(\frac{1}{\sqrt{n}})$. Since P_X becomes very small as n tends to ∞, the approximation

$$C \approx \frac{P_X}{2\sigma_Y^2 d_{X,Y}^{\alpha}},$$

is tight for large n, and hence we will employ

$$C = \frac{P_X}{2\sigma_Y^2 d_{X,Y}^{\alpha}}, \tag{8}$$

for our network design.

D. Multi-Hop Strategies

As mentioned in Section I, in order to improve the performance of communication between Alice and Bob, we consider multi-hop transmission. Alice, the source node, transmits a message to Bob, the destination node, in a multi-hop fashion. Let a path from Alice to Bob be denoted by $\Pi = (\ell_1, \ldots, \ell_H)$, where H is the number of hops of the path from Alice to Bob, and $\ell_i = (S(\ell_i), D(\ell_i))$ is the link between node $S(\ell_i)$ and node $D(\ell_i)$ along the path, where $S(\ell_1)$ is Alice and $D(\ell_H)$ is Bob.

We consider two approaches: multi-hop communication *with a single key (SK)* and multi-hop communication *with independent keys at the relays (IK)*. With a single key, all relays use the same key to encode the message, i.e. a message received from the previous relay is sent to the next relay using the same key. Hence, the exact same codeword is transmitted over every hop on the path from Alice to Bob. On the other hand, with independent keys at the relays, the message is re-encoded at each hop with a different key sequence such that the codeword sent over each link is independent of the codewords sent over other links of the path. For each approach, we first optimize transmission along a given path Π between Alice and Bob. In particular, we find the optimal powers that should be allocated to each relay along the path such that the end-to-end covertness constraint is satisfied, and the desired performance metric (covert throughput or end-to-end delay) is optimized. Then, for each case, we exploit these results to develop a routing algorithm that not only allocates the optimal powers to the relays, but also finds the optimal path Π^* from the set $\mathbf{\Pi}$ of all possible paths between Alice and Bob.

E. Key Distribution

Considering the fact that we need $\mathcal{O}(\sqrt{n})$ number of key-bits to encode each message, it seems quite challenging to exchange such a long key sequence in an adversarial environment. In particular, for multi-hop covert communication with independent keys at the relays (IK) many such long key sequences are needed. Fortunately, in practice (very) short key sequences shared between the relays are sufficient to generate the (very) long key sequences required for covert communication. As described in [4, Section II], the relays can use the short key sequences as the initial keys for a stream-cipher generating scheme to generate the long key sequences. For instance, a stream-cipher generating scheme called Trivium [20] with an 80-bit initial key can generate a 2^{64}-bit key sequence [4].

III. Covert Communication with a Single Key (SK)

In this section, we consider multi-hop covert communication with a single key. Consider an H-hop path $\Pi = (\ell_1, \ldots, \ell_H)$ between Alice and Bob. Every relay $S(\ell_i)$ forwards the message to the next relay $D(\ell_i)$ using the same key until it is delivered to the destination, Bob.

Consider an arbitrary Willie W_k observing the message transmission over Π. Since W_k can observe the transmission of every relay along the path, it can use its observations across hops to decide whether a transmission occurs or not. This is equivalent to the

case that H cooperating Willies $W_{k_1}, \ldots, W_{k_H}$ are present at the location of Willie W_k, such that W_{k_i} monitors the transmission of only the i^{th} hop. Then, $W_{k_1}, \ldots, W_{k_H}$ use their observations to constitute the total observation of Willie W_k over all hops. Hence, Willie W_k's observations of the i^{th} hop ($k = 1, \ldots, M$ and $i = 1, \ldots, H$) under hypothesis H_1 are described as:

$$Z_j^{(i,k)} = \frac{\sqrt{P_i} f_{i,j}}{d_{i,k}^{\alpha/2}} + N_j^{(W_k)},\ j = 1, 2, \ldots, n, \quad (9)$$

where P_i is the transmit power of relay $S(\ell_i)$, $f_{i,j}$ is the symbol that is sent over the i^{th} hop during the j^{th} symbol period, and $d_{i,k}$ is the distance from relay $S(\ell_i)$ to Willie W_k. Since, in this case, the same key is used to encode the message at every relay, the same symbol is sent over different hops ($f_{i,j} = f_j$) and thus, under hypothesis H_1, we have,

$$Z_j^{(i,k)} = \frac{\sqrt{P_i} f_j}{d_{i,k}^{\alpha/2}} + N_j^{(W_k)},\ j = 1, 2, \ldots, n. \quad (10)$$

Under hypothesis H_0, the Willies observations are given by,

$$Z_j^{(i,k)} = N_j^{(W_k)},\ j = 1, 2, \ldots, n. \quad (11)$$

Willies, with their collective observations over all Willies (k), hops (i) and symbol periods (j), attempt to detect message transmission.

A. Covertness Analysis of Covert Communication with a Single Key

Suppose $\mathbb{Q}_0$ is the joint probability distribution of Willies' observations over M Willies, over H hops, and over n channel uses under hypothesis H_0, and $\mathbb{Q}_1$ is the joint probability distribution of Willies' observations over M Willies, over H hops and over n channel uses under hypothesis H_1. Hence, $\mathbb{Q}_0$ is a zero-mean multivariate Gaussian probability distribution function with covariance matrix

$$\Sigma_0 = S \otimes I_{n \times n}, \quad (12)$$

where S is an $HM \times HM$ diagonal matrix

$$S = \text{diag}\ (\underbrace{\sigma_{W_1}^2, \ldots, \sigma_{W_1}^2}_{H \text{ times}}, \ldots, \sigma_{W_M}^2). \quad (13)$$

Note that each $\sigma_{W_k}^2,\ k = 1, \ldots, M$ is repeated H times in S because each Willie W_k observes the transmission of the same message over all H hops. On the other hand, $\mathbb{Q}_1$ is a zero-mean multivariate Gaussian probability distribution function with covariance matrix

$$\Sigma_1 = (S + UU^T) \otimes I_{n \times n}, \quad (14)$$

where U is a column vector with HM elements,

$$U = \left[\frac{\sqrt{P_1}}{d_{1,1}^{\alpha/2}}, \ldots, \frac{\sqrt{P_H}}{d_{H,1}^{\alpha/2}}, \ldots, \frac{\sqrt{P_1}}{d_{1,M}^{\alpha/2}}, \ldots, \frac{\sqrt{P_H}}{d_{H,M}^{\alpha/2}} \right]^T. \quad (15)$$

Suppose the Willies apply the optimal hypothesis test. Since $\mathbb{Q}_1$ and $\mathbb{Q}_0$ are multivariate Gaussian distributions, the relative entropy between them is given by (Appendix A),

$$\mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0) = \frac{1}{2} \Big(\text{Tr}\big(\Sigma_0^{-1} \Sigma_1\big) - \dim(\Sigma_0) - \ln \frac{|\Sigma_1|}{|\Sigma_0|} + (\mu_0 - \mu_1)^T \Sigma_0^{-1} (\mu_0 - \mu_1) \Big), \quad (16)$$

where μ_0 is the mean of $\mathbb{Q}_0$, μ_1 is the mean of $\mathbb{Q}_1$, $|\Sigma_0|$ is the determinant of Σ_0, and $\dim(\Sigma_0)$ is dimension of Σ_0. Replacing μ_0, μ_1, Σ_0 and Σ_1 in (16) and performing some algebraic manipulations (Appendix B), the relative entropy in (5) can be written as,

$$\mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0) = \frac{n}{2} \left(\sum_{\substack{\ell_i \in \Pi \\ W_k \in \mathcal{W}}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^{\alpha}} - \ln \left(1 + \sum_{\substack{\ell_i \in \Pi \\ W_k \in \mathcal{W}}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^{\alpha}} \right) \right). \quad (17)$$

Using the inequality $\ln(1 + x) \geq x - \frac{x^2}{2}$ for $x \geq 0$,

$$\mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0) \leq \frac{n}{4} \left(\sum_{\ell_i \in \Pi} \sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^{\alpha}} \right)^2. \quad (18)$$

Combining (6) and (18), if the following condition is satisfied,

$$\frac{n}{4} \left(\sum_{\ell_i \in \Pi} \sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^{\alpha}} \right)^2 \leq \delta, \quad (19)$$

then covertness is guaranteed. Equivalently, (19) can be written as

$$\sum_{\ell_i \in \Pi} \sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^{\alpha}} \leq \gamma_1, \quad (20)$$

where $\gamma_1 = 2\sqrt{\frac{\delta}{n}}$.

B. Maximum Throughput Covert Routing with a Single Key

In this section, first we find the optimal power allocation to relays of a given path Π between Alice and Bob to maximize the throughput of covert communication over Π. Then, we design a routing algorithm that computes the optimal path with maximum throughput from the set $\mathbf{\Pi}$ of all possible paths between Alice and Bob. While the set of all possible paths has exponential number of paths in it, we will present a routing algorithm that can find the optimal path in polynomial time.

1) Maximum Throughput of a Given Path: We consider maximizing the throughput of covert communication between Alice and Bob over a given path Π. In other words, we maximize the minimum throughput over all links in Π such that the constraint in (20) is satisfied:

$$\max \left(\min_i C_i\right),\ i=1,\dots,H \quad \text{s.t.} \sum_{\ell_i\in\Pi}\sum_{W_k\in\mathcal{W}} \frac{P_i}{\sigma^2_{W_k} d^\alpha_{i,k}} \le \gamma_1, \tag{21}$$

where C_i is the throughput achieved over link ℓ_i between relays $S(\ell_i)$ and $D(\ell_i)$, and, per Section II, we will employ

$$C_i = \frac{P_i}{2\sigma_i^2 d_i^\alpha}, \tag{22}$$

where σ_i^2 is the variance of AWGN at $D(\ell_i)$, and d_i is the length of the link ℓ_i. In the following, we show that $\min_i C_i$ subject to the covertness constraint $\sum_{\ell_i\in\Pi}\sum_{W_k\in\mathcal{W}} \frac{P_i}{\sigma^2_{W_k} d^\alpha_{i,k}} \le \gamma_1$ is maximized when all links $\ell_i = (S(\ell_i), D(\ell_i)) \in \Pi$ have the same covert throughput, i.e. $C_1 = \cdots = C_H$. First, let us restate (21),

$$\max \left(\min_i C_i\right),\ i=1,\dots,H \text{ s.t. } \sum_{\ell_i\in\Pi} a_i C_i \le \gamma_1, \tag{23}$$

where,

$$a_i = \sum_{W_k\in\mathcal{W}} \frac{2\sigma_i^2 d_i^\alpha}{\sigma^2_{W_k} d^\alpha_{i,k}}. \tag{24}$$

Suppose $C_{(1)} \le C_{(2)} \le \cdots \le C_{(H)}$ are ordered C_i's such that $C_{(1)} = \min_i C_i$. Since $C_{(1)} \le C_i,\ \forall\ i = 1,\cdots,H$,

$$C_{(1)} \sum_i a_i \le \sum_i C_i a_i \le \gamma_1,$$

and thus,

$$C_{(1)} \le \frac{\gamma_1}{\sum_i a_i}.$$

Now it remains to show that this upper-bound is achievable. This is achieved if

$$C_1 = \cdots = C_H = \frac{\gamma_1}{\sum_i a_i}. \tag{25}$$

Hence, the maximum covert throughput of a given path in the presence of multiple Willies is given by,

$$\begin{aligned} C_{\text{SK}} &= \frac{\gamma_1}{\sum_{\ell_i\in\Pi}\sum_{W_k\in\mathcal{W}} \frac{2\sigma_i^2 d_i^\alpha}{\sigma^2_{W_k} d^\alpha_{i,k}}} \\ &= \frac{\sqrt{\delta}}{\sum_{\ell_i\in\Pi}\sum_{W_k\in\mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma^2_{W_k} d^\alpha_{i,k}}} \frac{1}{\sqrt{n}}. \end{aligned} \tag{26}$$

Using (8), the optimal power that a relay $S(\ell_i) \in \Pi$ should transmit with to obtain the maximum covert throughput in (26) is,

$$\begin{aligned} P_i &= 2\sigma_i^2 d_i^\alpha C_{\text{SK}} \\ &= \frac{2\sqrt{\delta}\sigma_i^2 d_i^\alpha}{\sum_{\ell_j\in\Pi}\sum_{W_k\in\mathcal{W}} \frac{\sigma_j^2 d_j^\alpha}{\sigma^2_{W_k} d^\alpha_{j,k}}} \frac{1}{\sqrt{n}}. \end{aligned}$$

2) MT-SK Routing Algorithm: In this section, we find the optimal path with maximum throughput between Alice and Bob. From (26), the path that maximizes the covert throughput is the path that minimizes $\sum_{\ell_i\in\Pi}\sum_{W_k\in\mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma^2_{W_k} d^\alpha_{i,k}}$. Let us define the "cost" of a maximum covert throughput path Π with a single key as,

$$\omega_{\text{MT-SK}}(\Pi) = \sum_{\ell_i\in\Pi}\sum_{W_k\in\mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma^2_{W_k} d^\alpha_{i,k}}, \tag{27}$$

and the cost of communication over a link $\ell_i = (S(\ell_i), D(\ell_i))$ as,

$$\omega_{\text{MT-SK}}(\ell_i) = \sum_{W_k\in\mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma^2_{W_k} d^\alpha_{i,k}}. \tag{28}$$

Since the cost of each link does not depend on other links, we can obtain the minimum cost (maximum throughput) path by assigning the cost $\omega_{\text{MT-SK}}(\ell_i)$ to every potential link ℓ_i of the network, and solving a shortest-path problem. There are several classical shortest path algorithms that can be used for this purpose. In this paper, we use Dijkstra's algorithm.

C. Minimum Delay Covert Routing with a Single Key

In this section, first we find the optimal power allocation to minimize the end-to-end delay of a given path when transmitting a message from Alice to Bob in the presence of multiple Willies. Then we design a routing algorithm to choose the path with minimum end-to-end delay between Alice and Bob.

1) Minimum Delay of a Given Path: Suppose we have a multi-hop path from Alice to Bob. Here our goal is to minimize the end-to-end delay of transmitting a message covertly over a given path Π from Alice to Bob, such that the constraint in (20) is satisfied. We define the average delay of the i^{th} link, denoted by Δ_i, as the inverse of the link covert throughput, $\Delta_i = \frac{1}{C_i}$, and thus the end-to-end delay can be written as,

$$\Delta_{\text{SK}}(\Pi) = \sum_{\ell_i \in \Pi} \Delta_i = \sum_{\ell_i \in \Pi} \frac{1}{C_i}.$$

Hence, we want to solve the following problem,

$$\min \Delta_{\text{SK}}(\Pi), \quad \text{s.t.} \quad \sum_{\ell_i \in \Pi} \sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \leq \gamma_1. \tag{29}$$

From (8),

$$\Delta_i = \frac{1}{C_i} = \frac{2\sigma_i^2 d_i^{\alpha}}{P_i}. \tag{30}$$

Let us define,

$$b_i = \sum_{W_k \in \mathcal{W}} \frac{2\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{i,k}}. \tag{31}$$

Substituting b_i in (29), our optimization problem is,

$$\min \Delta_{\text{SK}}(\Pi), \ \text{s.t.} \quad \sum_{\ell_i \in \Pi} \frac{b_i}{\Delta_i} \leq \gamma_1. \tag{32}$$

In (32) the optimization objective is linear and the constraint is a convex set, and thus (32) is a convex optimization problem. Hence, a point that minimizes $\Delta_{\text{SK}}(\Pi)$ in (32) is a global minimum. Since left side of the constraint in (32) is a decreasing function of Δ_i and our goal is to minimize $\sum_{\ell_i \in \Pi} \Delta_i$, the constraint is active and becomes

$$\sum_{\ell_i \in \Pi} \frac{b_i}{\Delta_i} = \gamma_1. \tag{33}$$

In order to solve this optimization problem, we use the Lagrange multipliers technique. Thus, we should solve the following Lagrangian equations and the constraint (33) simultaneously,

$$\frac{\partial}{\partial \Delta_i} \left\{ \sum_{j=1}^{H} \Delta_j + \lambda \left(\sum_{j=1}^{H} \frac{b_j}{\Delta_j} - \gamma_1 \right) \right\} = 0, \ i = 1, \ldots, H.$$

Taking the derivatives of the Lagrangian functions the following equations are obtained,

$$1 - \lambda \frac{b_i}{\Delta_i^2} = 0, \ i = 1, \ldots, H. \tag{34}$$

Substituting Δ_i from (34) into (33), we obtain,

$$\lambda = \frac{1}{\gamma_1^2} \left(\sum_i \sqrt{b_i} \right)^2 \tag{35}$$

Hence, after substituting λ from (35) into (34), Δ_i is given by,

$$\Delta_i = \frac{1}{\gamma_1} \sqrt{b_i} \sum_{j=1}^{H} \sqrt{b_j}. \tag{36}$$

Therefore, we have,

$$\sum_{i=1}^{H} \Delta_i = \frac{1}{\gamma_1} \left(\sum_{j=1}^{H} \sqrt{b_j} \right)^2. \tag{37}$$

From (31), (32) and (37), the minimum end-to-end delay over a given path Π can be written as,

$$\sum_{i=1}^{H} \Delta_i = \frac{2}{\gamma_1} \left(\sum_{\ell_i \in \Pi} \sqrt{\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{i,k}}} \right)^2. \tag{38}$$

In order to obtain the minimum delay, relay $S(\ell_i)$ along the path Π should transmit to relay $D(\ell_i)$ with power,

$$P_i = \frac{2\sigma_i^2 d_i^{\alpha}}{\Delta_i}, \tag{39}$$

where from (31) and (36),

$$\Delta_i = \left(\frac{1}{\sqrt{\delta}} \sqrt{\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{i,k}}} \sum_{\ell_j \in \Pi} \sqrt{\sum_{W_k \in \mathcal{W}} \frac{\sigma_j^2 d_j^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{j,k}}} \right) \sqrt{n}. \tag{40}$$

2) MD-SK Routing Algorithm: In this section, our goal is to find the path Π with minimum end-to-end delay from Alice to Bob. From (38), in order to find the path with minimum delay, we should find a path for which

$$\Delta_{\text{SK}}(\Pi) = \sum_{\ell_i \in \Pi} \Delta_i = \frac{2}{\gamma_1} \left(\sum_{\ell_i \in \Pi} \sqrt{\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma_{W_k}^2 d_{i,k}^\alpha}} \right)^2,$$

is minimum. Let us define the cost of covert communication to minimize the end-to-end delay with a single key (MD-SK) of a path Π as,

$$\omega_{\text{MD-SK}}(\Pi) = \sum_{\ell_i \in \Pi} \sqrt{\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma_{W_k}^2 d_{i,k}^\alpha}} \tag{41}$$

This can be attained by assigning the link cost:

$$\omega_{\text{MD-SK}}(\ell_i) = \sqrt{\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma_{W_k}^2 d_{i,k}^\alpha}} \tag{42}$$

to every potential link ℓ_i in the network. Clearly, a path Π that minimizes $\omega_{\text{MD-SK}}(\Pi) = \sum_{\ell_i \in \Pi} \omega_{\text{MD-SK}}(\ell_i)$ also minimizes the end-to-end delay $\Delta_{\text{SK}}(\Pi)$. Hence, the problem is reduced to a shortest path problem with link costs $\omega_{\text{MD-SK}}(\ell_i)$ given by (42).

IV. Covert Communication with Independent Keys at the Relays (IK)

In this section, we consider multi-hop covert communication between Alice and Bob in the presence of multiple collaborating Willies with independent keys at the relays. In this approach, each relay along the path between Alice and Bob re-encodes the message with a different key, and then forwards it to the next relay until it is delivered to the destination, Bob. Since the message is encoded with different and independent keys at each hop, unlike the previous approach the signal sent over each hop is independent of the signal sent over other hops.

A. Covertness Analysis of Covert Communication with Independent Keys

Suppose the message is sent over a path Π from Alice to Bob, and $\mathbb{Q}_0$ is the joint probability distribution of Willies' observations over all hops and n channel uses under hypothesis H_0, and $\mathbb{Q}_1$ is the joint probability distribution of Willie's observations over all hops and n channel uses under hypothesis H_1. Since the message is encoded with different independent key sequences at each hop, the codewords sent over different hops are independent. Also, the noise is AWGN and thus is independent across different hops. Hence,

$$\mathbb{Q}_0 = \prod_{\ell_i \in \Pi} \mathbb{Q}_0^i, \text{ and, } \mathbb{Q}_1 = \prod_{\ell_i \in \Pi} \mathbb{Q}_1^i \tag{43}$$

where $\mathbb{Q}_0^i$ and $\mathbb{Q}_1^i$ are the joint probability distributions of Willies' observations over the i^{th} hop under hypotheses H_0 and H_1, respectively. Suppose Willies apply the optimal hypothesis test to make a decision on the end-to-end communication. From the independence of the observations across hops, the end-to-end relative entropy between $\mathbb{Q}_1$ and $\mathbb{Q}_0$ is,

$$\mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0) = \sum_{\ell_i \in \Pi} \mathbb{D}(\mathbb{Q}_1^i \| \mathbb{Q}_0^i). \tag{44}$$

At each hop, Willies combine their observations across Willies and across time. Using the same approach as in Section III-A, the relative entropy between $\mathbb{Q}_1^i$ and $\mathbb{Q}_0^i$ is,

$$\mathbb{D}(\mathbb{Q}_1^i \| \mathbb{Q}_0^i) = \frac{n}{2} \left(\sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^\alpha} - \ln \left(1 + \sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^\alpha} \right) \right) \tag{45}$$

Using the fact that $\ln(1+x) \geq x - \frac{x^2}{2}$ for $x \geq 0$,

$$\mathbb{D}(\mathbb{Q}_1^i \| \mathbb{Q}_0^i) \leq \frac{n}{4} \left(\sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^\alpha} \right)^2. \tag{46}$$

Hence, from (44) and (46) the end-to-end relative entropy between $\mathbb{Q}_1$ and $\mathbb{Q}_0$ can be written as,

$$\begin{aligned} \mathbb{D}(\mathbb{Q}_1 \| \mathbb{Q}_0) &= \sum_{\ell_i \in \Pi} \mathbb{D}(\mathbb{Q}_1^i \| \mathbb{Q}_0^i) \\ &\leq \frac{n}{4} \sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^\alpha} \right)^2. \end{aligned} \tag{47}$$

Combining (6) and (47), in order to guarantee the end-to-end covertness it suffices to have,

$$\frac{n}{4} \sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma_{W_k}^2 d_{i,k}^\alpha} \right)^2 \leq \delta. \tag{48}$$

Setting $\gamma_2 = \frac{4\delta}{n}$ the covertness constraint (47) can be written as,

$$\sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2 \leq \gamma_2. \tag{49}$$

B. Comparison of Multi-Hop Covert Communication with a Single Key and with Independent Keys

Consider the covertness constraint of covert communication with a single key (19) in Section III-A and let,

$$B_{\text{SK}} = \frac{n}{4} \left(\sum_{\ell_i \in \Pi} \sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2. \tag{50}$$

Also, consider the covertness constraint of covert communication with independent keys at the relays (51) in Section IV-A and let,

$$B_{\text{IK}} = \frac{n}{4} \sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2. \tag{51}$$

Clearly, for the same path Π and same powers P_i, $B_{\text{IK}} < B_{\text{SK}}$. Hence, with the same covertness constraint δ, communication with independent keys approach compared to communication with a single key approach allows higher powers while maintaining covertness, which results in higher throughput and lower delay. Hence, we expect covert communication with independent keys to have better performance than covert communication with a single key. We will show this in more detail with simulations for various parameters of the network in Section V. Note that the better performance of the scheme with independent keys comes at the expense of more key bits.

C. Maximum Throughput Covert Communication with Independent Keys

In this section, we characterize the optimum power allocation to the relays along a given path Π from Alice to Bob to maximize the covert throughput when using independent keys at relays. Also, we design a routing algorithm that computes the maximum throughput path.

1) Maximum Throughput of a Given Path: Here we find the optimal power allocation on a given path Π between Alice and Bob to maximize the covert throughput. In order to find the maximum covert throughput, we should maximize the minimum throughput over all hops in Π such that the constraint in (49) is satisfied,

$$\begin{aligned} &\max \left(\min_i C_i \right), \ i = 1, \ldots, H \\ &\text{s.t.} \sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{P_i}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2 \leq \gamma_2. \end{aligned} \tag{52}$$

This optimization is similar to the optimization in Section III-B, and can be solved in the same way. The detailed solution is presented in Appendix C. Hence, the maximum covert throughput of a given path with independent keys at the relays is,

$$\begin{aligned} C_{\text{IK}} &= \frac{\sqrt{\gamma_2}}{\sqrt{\sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{2\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2}} \\ &= \frac{\sqrt{\delta}}{\sqrt{\sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2}} \frac{1}{\sqrt{n}}. \end{aligned} \tag{53}$$

Each relay along the path Π should transmit its message to the next relay with optimal power,

$$\begin{aligned} P_i &= 2\sigma_i^2 d_i^{\alpha} C_{\text{IK}} \\ &= \frac{2\sigma_i^2 d_i^{\alpha} \sqrt{\delta}}{\sqrt{\sum_{\ell_j \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{\sigma_j^2 d_j^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{j,k}} \right)^2}} \frac{1}{\sqrt{n}}. \end{aligned} \tag{54}$$

2) MT-IK Routing Algorithm: From (53) in order to find the maximum throughput path Π between Alice and Bob, we should find the path Π for which $\sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{2\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2$ is minimum. Hence, define the cost of covert communication to maximize the throughput with independent keys at the relays (MT-IK) of a path Π as,

$$\omega_{\text{MT-IK}}(\Pi) = \sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2. \tag{55}$$

Assign the following link cost $\omega(\ell_i)$ to every potential link in the network:

$$\omega_{\text{MT-IK}}(\ell_i) = \left(\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k} d^{\alpha}_{i,k}} \right)^2 \tag{56}$$

and find the shortest path Π with link costs $\omega_{\text{MT-IK}}(\ell_i)$ using any shortest path routing algorithm.

D. Minimum Delay Covert Routing with Independent Keys

In this section, first we find the suitable power allocation to minimize the end-to-end delay of covert communication over a given path. Next, we propose a routing algorithm to find the minimum end-to-end delay path from the set of all paths $\mathbf{\Pi}$ between Alice and Bob.

1) Minimum Delay of a Given Path: Here the goal is to minimize the end-to-end delay of a given path Π such that the constraint in (49) is satisfied,

$$\min\ \Delta_{\text{IK}}(\Pi), \quad \text{s.t.} \sum_{\ell_i\in\Pi}\left(\sum_{W_k\in\mathcal{W}}\frac{P_i}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\right)^2 \le \gamma_2. \tag{57}$$

Define,

$$h_i = \left(\sum_{W_k\in\mathcal{W}}\frac{2\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\right)^2. \tag{58}$$

Substituting $\Delta_{\text{IK}}(\Pi)=\sum_{i=1}^{H}\Delta_i$ and h_i in (57), our optimization problem is,

$$\min\sum_{i=1}^{H}\Delta_i, \quad \text{s.t.} \quad \sum_{\ell_i\in\Pi}\frac{h_i}{\Delta_i^2}\le\gamma_2. \tag{59}$$

The objective function in (59) is linear and the constraint is a convex set. Hence, (59) is a convex optimization problem and any point that minimizes the objective function is a global minimum as well. Using the same reasoning as in Section III-C1, the constraint in (59) is active and thus the inequality constraint in (59) can be substituted by the following equality constraint,

$$\sum_{\ell_i\in\Pi}\frac{h_i}{\Delta_i^2}=\gamma_2. \tag{60}$$

In order to solve this optimization problem, we use the Lagrange multipliers technique. Thus, we should solve the following equations and the constraint (60) simultaneously,

$$\frac{\partial}{\partial\Delta_i}\left\{\sum_{j=1}^{H}\Delta_j+\lambda\left(\sum_{j=1}^{H}\frac{h_j}{\Delta_j^2}-\gamma_2\right)\right\}=0, \quad i=1,\dots,H.$$

Setting the derivatives to zero, we have,

$$1-2\lambda\frac{h_i}{\Delta_i^3}=0,\ i=1,\dots,H,$$

and thus,

$$\Delta_i=(2\lambda h_i)^{1/3},\ i=1,\dots,H. \tag{61}$$

Substituting Δ_i from (61) into (60),

$$\lambda=\frac{1}{(2\gamma_2)^{3/2}}\left(\sum_i {h_i}^{1/3}\right)^{3/2}. \tag{62}$$

Hence, by substituting λ from (62) into (61) we have,

$$\Delta_i=\frac{1}{\sqrt{\gamma_2}}{h_i}^{1/3}\left(\sum_{j=1}^{H}{h_j}^{1/3}\right)^{1/2}. \tag{63}$$

Thus, the minimum end-to-end delay of sending a message covertly from Alice to Bob over a given path Π is,

$$\begin{aligned}\Delta_{\text{IK}}(\Pi)&=\sum_{i=1}^{H}\Delta_i\\&=\frac{1}{\sqrt{\gamma_2}}\sum_{i=1}^{H}{h_i}^{1/3}\left(\sum_{i=1}^{H}{h_i}^{1/3}\right)^{1/2}\\&=\frac{1}{\sqrt{\gamma_2}}\left(\sum_{i=1}^{H}{h_i}^{1/3}\right)^{3/2}\\&=\frac{2}{\sqrt{\gamma_2}}\left(\sum_{\ell_i\in\Pi}\left(\sum_{W_k\in\mathcal{W}}\frac{\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\right)^{1/3}\right)^{3/2}.\end{aligned} \tag{64}$$

In order to attain the minimum end-to-end delay, relay $S(\ell_i)$ should transmit with power,

$$P_i=\frac{2\sigma_i^2 d_i^{\alpha}}{\Delta_i} \tag{65}$$

where,

$$\Delta_i=\frac{1}{\sqrt{\gamma_2}}\left(\sum_{W_k\in\mathcal{W}}\frac{2\sigma_i^2 d_i^{\alpha}}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\right)^{1/3}\left(\sum_{\ell_j\in\Pi}\left(\sum_{W_k\in\mathcal{W}}\frac{2\sigma_j^2 d_j^{\alpha}}{\sigma^2_{W_k}d^{\alpha}_{j,k}}\right)^{1/3}\right)^{1/2}. \tag{66}$$

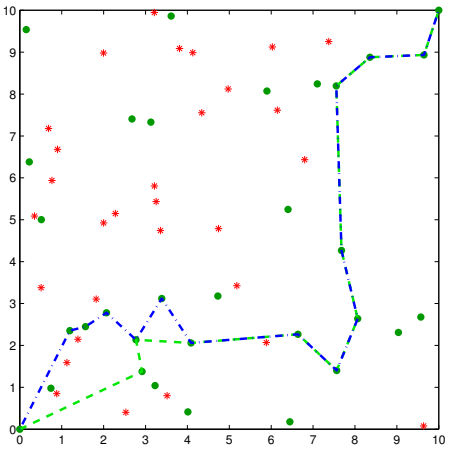

Fig. 1. A snapshot of the network when 30 system nodes (green circles) and 30 Willies (red stars) are present in the network. All nodes are distributed uniformly at random over the network. The covertness factor is set to $\delta = 0.05$, and the path-loss exponent to $\alpha = 3$. The path that achieves the maximum throughput with the MT-SK algorithm is shown by blue dash-dot lines, and the path that achieves the maximum throughput with the MT-IK algorithm is shown by green dashed lines. The maximum covert throughput using the MT-SK algorithm is $0.0024/\sqrt{n}$ and the maximum covert throughput using the MT-IK algorithm is $0.0056/\sqrt{n}$.

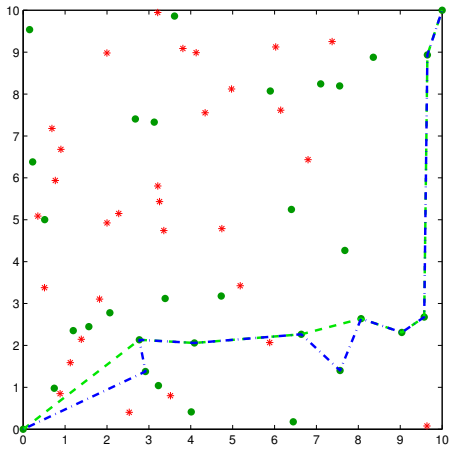

Fig. 2. A snapshot of the network when 30 system nodes (green circles) and 30 Willies (red stars) are present in the network. All nodes are distributed uniformly at random over the network. The covertness factor is set to $\delta = 0.05$, and the path-loss exponent to $\alpha = 3$. The path achieving the minimum delay with the MD-SK algorithm is shown by blue dash-dotted lines, and the path achieving the minimum delay with the MD-IK algorithm is shown by green dashed lines. The minimum end-to-end delay using the MD-SK algorithm is $2448.8\sqrt{n}$ and the minimum end-to-end delay using the MD-IK algorithm is $1048.1\sqrt{n}$.

2) MD-IK Routing Algorithm: In order to compute the path with maximum throughput Π between Alice and Bob, we should find the path for which the end-to-end delay,

$$\Delta_{\text{IK}}(\Pi) = \frac{2}{\sqrt{\gamma_2}} \left(\sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma_{W_k}^2 d_{i,k}^\alpha} \right)^{1/3} \right)^{3/2}, \tag{67}$$

is minimized. Define the cost of covert communication with minimum end-to-end delay using independent keys at the relays (MD-IK) over a path Π as,

$$\omega_{\text{MD-IK}}(\Pi) = \sum_{\ell_i \in \Pi} \left(\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma_{W_k}^2 d_{i,k}^\alpha} \right)^{1/3}. \tag{68}$$

Assign the link cost $\omega_{\text{MD-IK}}(\ell_i)$ to every link ℓ_i in the network,

$$\omega_{\text{MD-IK}}(\ell_i) = \left(\sum_{W_k \in \mathcal{W}} \frac{\sigma_i^2 d_i^\alpha}{\sigma_{W_k}^2 d_{i,k}^\alpha} \right)^{1/3}, \tag{69}$$

and apply any shortest-path algorithm to find the path with minimum cost from Alice to Bob, which is the desired path Π^*.

V. Numerical Results

In this section, we evaluate and compare the performance of the routing algorithms proposed in this paper numerically. A wireless network on a $d \times d$ square on the 2-D plane with corners $(0,0), (0,d), (d,0), (d,d)$ is considered. In all simulations, Alice (source) is located at point $(0,0)$ and Bob (destination) is located at point (d,d). Multiple friendly system nodes and multiple Willies are distributed uniformly at random over the network. We consider fully connected networks. For all routing algorithms, we assign the link costs described in previous sections to every link in the network, and then apply the Dijkstra's algorithm to find the shortest (minimum cost) path from Alice to Bob in each case. Since Dijkstra's algorithm is a polynomial time algorithm, the computational complexity of the proposed algorithms is also polynomial in the size of the network, and hence the proposed routing algorithms are efficient.

Figs. 1 and 2 show one snapshot of the network when 30 system nodes and 30 Willies are present. In both figures we set the path-loss exponent to $\alpha = 3$ and the covertness factor to $\delta = 0.05$. In Fig. 1, the maximum throughput paths obtained by the MT-

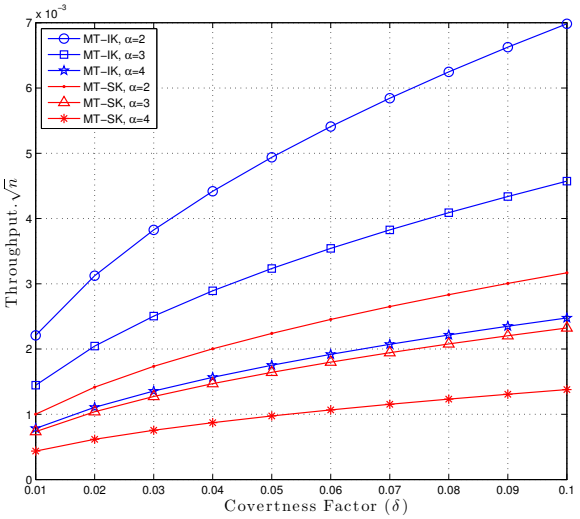

Fig. 3. Maximum throughput versus covertness factor δ, where 30 system nodes and 30 Willies are present in the network.

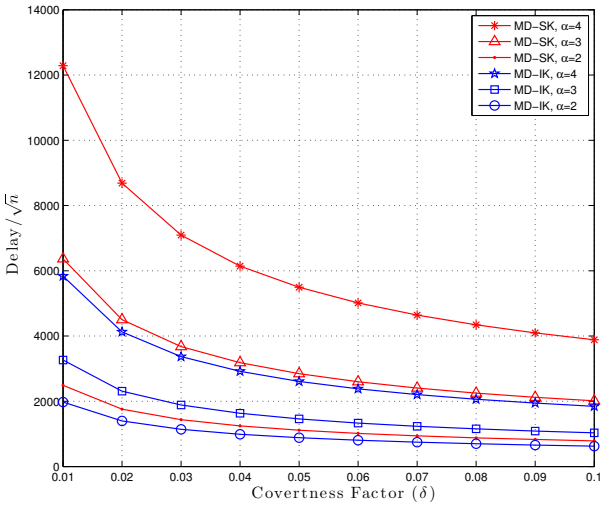

Fig. 4. End-to-end delay versus covertness factor δ, where 30 system nodes and 30 Willies are present in the network.

SK and MT-IK algorithms are shown. In this case, the covert throughput of one-hop communication from Alice to Bob with covertness factor of $\delta = 0.05$ is $8.4476 \times 10^{-5}/\sqrt{n}$. When using multi-hop communication, the maximum covert throughput of the MT-SK algorithm is $0.0024/\sqrt{n}$ and the maximum covert throughput of the MT-IK algorithm is $0.0056/\sqrt{n}$. Thus, both MT-SK and MT-IK algorithms improve the performance of one hop covert communication significantly, since they both choose paths that avoid Willies, and allocate the covertness factor to the links on each path to increase the covert throughput. As expected from Section IV-B, MT-IK offers a higher covert throughput compared to MT-SK.

In Fig. 2, the minimum delay paths selected by the MD-SK and MD-IK algorithms are shown. The minimum delay of one-hop covert communication from Alice to Bob, which is defined as the inverse of the covert throughput of the link from Alice to Bob, is $11838\sqrt{n}$. The minimum end-to-end delay of MD-SK is $2448.8\sqrt{n}$, and the minimum end-to-end delay of MD-IK is $1048.1\sqrt{n}$. Thus, both the MD-SK and MD-IK algorithms improve the performance of one-hop covert communication significantly. Again, both paths avoid Willies by taking detours. Because of the different allocation of the covertness factors to the links along each path, the optimal paths and the optimal end-to-end delays are different, and, as expected, the MD-IK algorithm offers a smaller end-to-end delay than the MD-SK algorithm.

In the remainder of this section, we consider the effect of different parameters of the network on the performance of the MT-SK, MT-IK, MD-SK, and MD-IK algorithms. We average our results over 100 randomly generated realizations of the network with different seeds, and with uniform distribution of system nodes and Willies. Our performance metric is the average throughput over different realizations of the network for the MT-SK and MT-IK algorithms, and the average end-to-end delay over different realizations of the network for the MD-SK and MD-IK algorithms. In order to have precise comparisons, we use the same placements of the system nodes and Willies in different cases.

Effect of the covertness factor δ. Fig. 3 shows the maximum throughput of MT-SK and MT-IK versus the covertness factor δ, when 30 system nodes and 30 Willies are present in the network, and δ changes from 0.01 to 0.1. As can be seen, the performance of MT-IK for different δs and for different path-loss exponents α is better than MT-SK. Also, as expected, as δ increases, higher covert throughputs can be achieved.

In Fig. 4, the minimum end-to-end delay of MD-SK and MD-IK versus the covertness factor δ is shown, where 30 system nodes and 30 Willies are present in the network, and δ changes from 0.01 to 0.1. It is apparent that using the MD-IK algorithm,

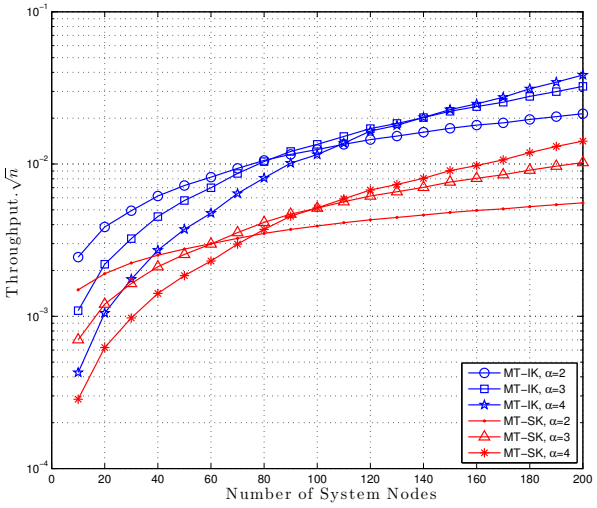

Fig. 5. Throughput versus the number of system nodes, when 30 Willies are present in the network and the covertness factor $\delta = 0.05$.

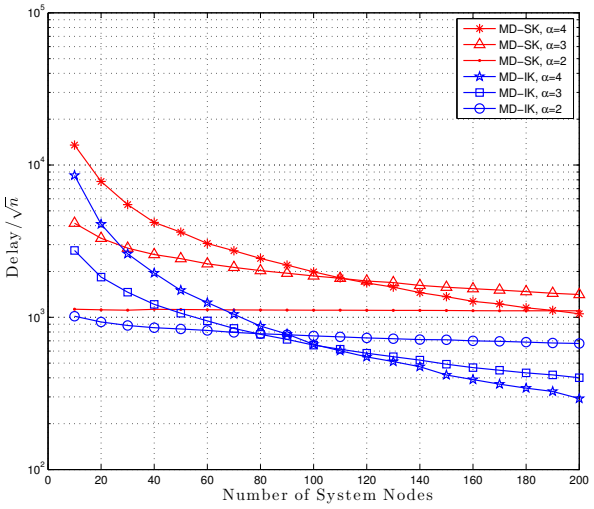

Fig. 6. End-to-end delay versus number of system nodes, when 30 Willies are present in the network and the covertness factor $\delta = 0.05$.

a smaller end-to-end delay can be achieved. As δ increases, each relay can transmit with higher power (higher throughput), and thus the end-to-end delay decreases for both MD-SK and MD-IK.

Effect of the number of system nodes. In Fig. 5, the maximum covert throughput of MT-SK and MT-IK versus the number of system nodes, when $\delta = 0.05$ and 30 Willies are present in the network, are shown. As the number of system nodes increases, the maximum throughputs that MT-SK and MT-IK can achieve increase, since the path can take more detours to avoid Willies. The maximum covert throughput of MT-IK is always larger than the maximum covert throughput of MT-SK. Further, as the number of system nodes increases, the maximum throughput of MT-SK increases slowly. The reason is that for MT-SK, when the number of hops increases, the same signal is sent over a higher number of hops and thus it will be more likely that the Willies can detect the communication.

The performance curves of the schemes as a function of path-loss exponent intersect at some points. That is, for a small number of system nodes, the covert throughput when α is small is higher, but as the number of system nodes increases, the covert throughput when α is larger becomes higher. The reason is that when the number of system nodes is small, Alice and Bob have very few choices of nodes to construct a path, and thus the optimal path might not be able to avoid Willies effectively. Hence, the throughput when α is small is higher because smaller α leads to smaller signal attenuation and thus higher throughput. But when the number of system nodes becomes larger, Alice and Bob have many choices to construct a path that can avoid Willies. And, when the path-loss exponent is large, the higher attenuation of the environment makes the transmission of each relay local (because each relay has a smaller broadcast range), which helps the system nodes to avoid Willies more effectively. Thus, when the number of system nodes is higher, we have a higher covert throughput for larger path-loss exponents.

Fig. 6 shows the end-to-end delay of the MT-SK and MT-IK algorithms versus the number of system nodes, when $\delta = 0.05$ and 30 Willies are present in the network. It can be seen that the performance of MT-IK is better than the performance of MT-SK. As the number of system nodes increases, the end-to-end delay decreases because the routing algorithm has a larger set from which to choose the relays so as to minimize the end-to-end delay.

For all algorithms, when the number of system nodes is small, for smaller α we have better performances. The reason is that a smaller path-loss exponent means less attenuation and thus higher throughput. However, as the number of system nodes increases, the optimal path can take advantage of more system nodes to take detours and avoid Willies. In this case, the performance of the algo-

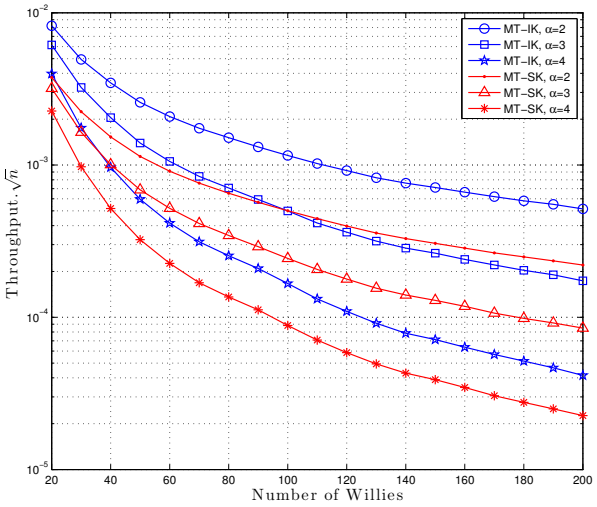

Fig. 7. Throughput versus the number of Willies, when 30 system nodes are present in the network and the covertness factor $\delta = 0.05$.

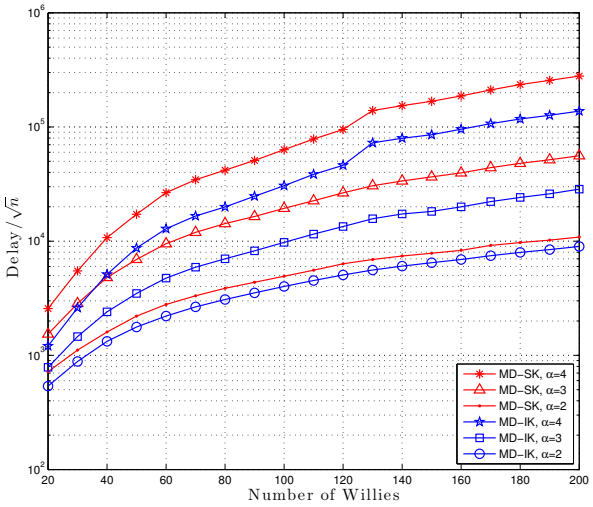

Fig. 8. End-to-end delay versus the number of Willies, when 30 system nodes are present in the network and the covertness factor $\delta = 0.05$.

rithms when the path-loss exponent is large is better compared to when the path-loss exponent is small. The reason is that when the path-loss exponent is large the effect of each Willie is local and taking detours can improve the performance to a greater extent.

Effect of Number of Willies. The effect of the number of Willies on the maximum covert throughput achieved by MT-SK and MT-IK is shown in Fig. 7. In this figure, $\delta = 0.05$ and 30 system nodes are present in the network. With both algorithms, as the number of Willies increases, the maximum throughput of covert communication decreases. In all situations considered in this figure, the performance of the MT-IK algorithm is better than the performance of MT-SK algorithm, as expected.

The end-to-end delay versus the number of Willies is shown in Fig. 8, when 30 system nodes are present in the network and $\delta = 0.05$. As expected, the end-to-end delay of transmission from source to destination increases as the number of Willies increases, because with more Willies the throughput of communication at each link becomes smaller, and the optimum path should take more detours to avoid Willies, resulting in a larger number of hops. It can be seen that MD-IK always has a smaller end-to-end delay than MD-SK.

VI. Concluding Remarks

In this paper, multi-hop covert communication over an arbitrary network in an AWGN environment and in the presence of multiple collaborating Willies has been considered. We developed maximum throughput and minimum end-to-end delay routing algorithms for *a single key for all relays* approach, and for *independent keys at the relays* approach. We have shown that using these multi-hop algorithms improves the performance of traditional one-hop covert communication from Alice to Bob substantially. Each proposed routing algorithm is straightforward to implement, and finds the optimal path in polynomial time in the size of the network.

We have shown mathematically and via simulations that for different network parameters the performance (throughput and delay) of routing algorithms with independent keys is better compared to that of routing algorithms with a single key for all relays. Note that the better performance of routing with independent keys is gained at the expense of a higher number of key bits used for covert communication. As mentioned in Section II-E, the long key sequences can be generated from short keys pre-shared between the system nodes. In this paper, we pictured a scenario where the system nodes are co-located and share keys, and then they are distributed over the network area and use their pre-shared keys for covert communication.

An exciting direction for future work is to consider covert wireless key distribution.

APPENDIX A

The relative entropy between $\mathbb{Q}_1$ and $\mathbb{Q}_0$ can be written as,

$$\mathbb{D}(\mathbb{Q}_1\|\mathbb{Q}_0) = \mathbb{E}_{\mathbb{Q}_1}\{\log \mathbb{Q}_1 - \log \mathbb{Q}_0\}. \tag{70}$$

Then,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}_1}\{\log \mathbb{Q}_1\} \\
&= \frac{\dim(\Sigma_1)}{2}\log(2\pi) + \frac{1}{2}\log|\Sigma_1| \\
&\quad + \frac{1}{2}\mathbb{E}_{\mathbb{Q}_1}\{(x-\mu_1)\Sigma_1^{-1}(x-\mu_1)\} \\
&= \frac{\dim(\Sigma_1)}{2}\log(2\pi) + \frac{1}{2}\log|\Sigma_1| + \frac{1}{2}\operatorname{Tr}\{\Sigma_1^{-1}\Sigma_1\} \\
&\quad + \frac{1}{2}(\mu_1-\mu_1)\Sigma_1^{-1}(\mu_1-\mu_1) \qquad (71) \\
&= \frac{\dim(\Sigma_1)}{2}\log(2\pi) + \frac{1}{2}\log|\Sigma_1| + \frac{1}{2}\dim(\Sigma_1), \qquad (72)
\end{aligned}$$

where (71) follows from [21, Section 8.2.2], and,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}_1}\{\log \mathbb{Q}_0\} \\
&= \mathbb{E}_{\mathbb{Q}_1}\Big\{ -\frac{\dim(\Sigma_0)}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_0| \\
&\quad - \frac{1}{2}(x-\mu_0)\Sigma_0^{-1}(x-\mu_0)\Big\} \\
&= -\frac{\dim(\Sigma_0)}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_0| \\
&\quad - \frac{1}{2}\mathbb{E}_{\mathbb{Q}_1}\{(x-\mu_0)\Sigma_0^{-1}(x-\mu_0)\} \\
&= -\frac{\dim(\Sigma_0)}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_0| - \frac{1}{2}\operatorname{Tr}\{\Sigma_0^{-1}\Sigma_1\} \\
&\quad + \frac{1}{2}(\mu_1-\mu_0)\Sigma_0^{-1}(\mu_1-\mu_0), \qquad (73)
\end{aligned}$$

where (73) follows from [21, Section 8.2.2]. Combining (70), (72) and (73),

$$\begin{aligned}
\mathbb{D}(\mathbb{Q}_1\|\mathbb{Q}_0) = \frac{1}{2}\Bigg(&\operatorname{Tr}\{\Sigma_0^{-1}\Sigma_1\} + (\mu_1-\mu_0)\Sigma_0^{-1}(\mu_1-\mu_0) \\
&+ \log\frac{|\Sigma_0|}{|\Sigma_1|} - \dim(\Sigma_1)\Bigg).
\end{aligned}$$

APPENDIX B

Here we calculate the relative entropy between $\mathbb{Q}_1$ and $\mathbb{Q}_0$ of Section III-A. The relative entropy of two multivariate Gaussian random variable $\mathbb{Q}_1 = \mathcal{N}(\mu_1, \Sigma_1)$ and $\mathbb{Q}_0 = \mathcal{N}(\mu_0, \Sigma_0)$ is,

$$\begin{aligned}
\mathbb{D}(\mathbb{Q}_1\|\mathbb{Q}_0) = \frac{1}{2}\Big(&\operatorname{Tr}\big(\Sigma_0^{-1}\Sigma_1\big) + (\mu_0-\mu_1)^T\Sigma_0^{-1}(\mu_0-\mu_1) \\
&- \dim(\Sigma_0) - \ln\left(\frac{|\Sigma_1|}{|\Sigma_0|}\right)\Big). \qquad (74)
\end{aligned}$$

We use the same approach as in [9, Appendix] to calculate each term of (74). The first term can be written as,

$$\begin{aligned}
\operatorname{Tr}\big(\Sigma_0^{-1}\Sigma_1\big) &= n\operatorname{Tr}\big(S^{-1}(S+UU^T)\big) \\
&= n\operatorname{Tr}\big(I_{HM\times HM} + S^{-1}UU^T\big) \\
&= nHM + n\sum_{i=1}^{H}\sum_{k=1}^{M}\frac{P_i}{\sigma_{W_k}^2 d_{i,k}^{\alpha}}. \qquad (75)
\end{aligned}$$

The second term vanishes because $\mu_0 = \mu_1 = 0$. The third term is,

$$\dim(\Sigma_0) = \dim(S\otimes I_{n\times n}) = nHM.$$

The forth term can be calculated as,

$$\begin{aligned}
|\Sigma_0| &= |S\otimes I_{n\times n}| \\
&= |S|^n|I_{n\times n}|^{HM} \\
&= |S|^n. \qquad (76)
\end{aligned}$$

and,

$$\begin{aligned}
|\Sigma_1| &= |(S+UU^T)\otimes I_{n\times n}| \\
&= |S+UU^T|^n|I_{n\times n}|^{HM} \\
&= |S|^n|I+S^{-1}UU^T|^n \\
&= |S|^n(I+U^TS^{-1}U)^n \\
&= |\Sigma_0|\left(1+\sum_{i=1}^{H}\sum_{k=1}^{M}\frac{P_i}{\sigma_{W_k}^2 d_{i,k}^{\alpha}}\right)^n. \qquad (77)
\end{aligned}$$

Thus,

$$
\begin{aligned}
&\mathbb{D}(\mathbb{Q}_1\|\mathbb{Q}_0)\\
&=\frac{1}{2}\Big(\operatorname{Tr}\big(\Sigma_0^{-1}\Sigma_1\big)+(\mu_0-\mu_1)^T\Sigma_0^{-1}(\mu_0-\mu_1)\\
&\quad-\ln\frac{|\Sigma_1|}{|\Sigma_0|}-\dim(\Sigma_0)\Big)\\
&=\frac{n}{2}\Big(HM+\sum_{i=1}^{H}\sum_{k=1}^{M}\frac{P_i}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\\
&\quad-\ln\Big(1+\sum_{i=1}^{H}\sum_{k=1}^{M}\frac{P_i}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\Big)-HM\Big)\\
&=\frac{n}{2}\Big(\sum_{i=1}^{H}\sum_{k=1}^{M}\frac{P_i}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\\
&\quad-\ln\Big(1+\sum_{i=1}^{H}\sum_{k=1}^{M}\frac{P_i}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\Big)\Big)
\end{aligned}
$$

APPENDIX C

In this appendix we present the solution of the following optimization problem:

$$
\begin{aligned}
&\max\Big(\min_i C_i\Big),\ i=1,\dots,H\\
&\text{s.t.}\sum_{\ell_i\in\Pi}\left(\sum_{W_k\in\mathcal{W}}\frac{P_i}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\right)^2\le\gamma_2. \qquad (78)
\end{aligned}
$$

We claim that $\max\left(\min_i C_i\right)$ in (78) is obtained when all links $\ell_i\in\Pi$ have the same throughput. Let us define,

$$
g_i=\left(\sum_{W_k\in\mathcal{W}}\frac{2\sigma_i^2d_i^{\alpha}}{\sigma^2_{W_k}d^{\alpha}_{i,k}}\right)^2. \qquad (79)
$$

Hence, we should maximize $\min_i C_i$ such that $\sum_i C_i^2 g_i \le \gamma_2$. Suppose $C_{(1)}=\min_i C_i,\ i=1,\dots,H$. We have,

$$
\gamma_2\ge\sum_i C_i^2g_i\ge C_{(1)}^2\sum_i g_i,
$$

and thus,

$$
C_{(1)}\le\sqrt{\frac{\gamma_2}{\sum_i g_i}}.
$$

Setting

$$
C_1=\cdots=C_H=\sqrt{\frac{\gamma_2}{\sum_i g_i}}, \qquad (80)
$$

proves the claim.

REFERENCES

[1] C. Paar and J. Pelzl, *Understanding cryptography: a textbook for students and practitioners*. Springer, 2010.

[2] A. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[3] C. Cachin and U. Maurer, "Unconditional security against memory-bounded adversaries," *Advances in Cryptology*, pp. 292–306, 1997.

[4] A. Sheikholeslami, D. Goeckel, and H. Pishro-Nik, "Everlasting secrecy by exploiting non-idealities of the eavesdropper's receiver," *IEEE Journal of Selected Areas in Communications*, vol. 31, no. 9, pp. 1828–1839, 2013.

[5] A. Sheikholeslami, M. Ghaderi, H. Pishro-Nik, and D. Goeckel, "Energy-efficient secrecy in wireless networks based on random jamming," *IEEE Transactions on Communications*, vol. 65, no. 6, pp. 2522–2533, 2017.

[6] A. Sheikholeslami, D. Goeckel, and H. Pishro-Nik, "Jamming based on an ephemeral key to obtain everlasting security in wireless environments," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6072–6081, 2015.

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

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

[9] R. Soltani, B. Bash, D. Goeckel, S. Guha, and D. Towsley, "Covert single-hop communication in a wireless network with distributed artificial noise generation," in *Proc. IEEE Allerton Conference*, 2014, pp. 1078–1085.

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

[11] 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.

[12] 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.

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

[14] A. Sheikholeslami, B. A. Bash, D. Towsley, D. Goeckel, and S. Guha, "Covert communication over classical-quantum channels," in *Proc. IEEE ISIT*, 2016, pp. 2064–2068.

[15] 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.

[16] B. He, S. Yan, X. Zhou, and V. K. Lau, "On covert communication with noise uncertainty," *IEEE Communications Letters*, vol. 21, no. 4, pp. 941–944, 2017.

[17] J. Hu, S. Yan, X. Zhou, F. Shu, J. Li, and J. Wang, "Covert communication achieved by a greedy relay in wireless networks," *arXiv preprint arXiv:1708.00905*, 2017.

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

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

[20] M. Robshaw and O. Billet, "New stream cipher designs: The eSTREAM finalists," vol. 4986. Springer, 2008.

[21] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, p. 15, 2008.