

Multi-Antenna Jamming in Covert Communication

Ori Shmuel, Asaf Cohen, Omer Gurewitz

Ben-Gurion University of the Negev, {shmuelor, coasaf, gurewitz}@bgu.ac.il

Abstract—Covert communication conceals transmission of messages from Alice to Bob out of a watchful adversary, Willie, who tries to determine if a transmission took place or not. While covert communication in a basic, vanilla setting where all variables are known to Willie, results in the well-known square-root law, when a jammer is present and assists Alice by creating uncertainty in Willie’s decoder, a strictly positive transmission rate is possible.

In this work, we analyze the case where the jammer is equipped with multiple antennas. Specifically, we analyze the effect of multiple antennas at the jammer on Alice’s transmission power and consequently on the transmission rate. We consider both cases, one in which the channel knowledge is known and one in which it is unknown by the jammer. We formulate several optimization problems for the transmission strategies of the jammer, to maximize his assistance to Alice, in terms of maximizing a ratio between Willie’s and Bob’s noise variances.

When the channel information is known to the jammer, we show that the optimal strategy of the jammer is to perform beamforming towards a single direction with all his available power. This direction though, is not trivial, since it reflects an optimal tradeoff point between minimizing the interference at Bob and maximizing the interference at Willie. When the channel knowledge is unknown, we show that the optimal strategy of the jammer is either to transmit isotropically to all directions or to the null-space of Bob, where this choice depends on certain channel conditions. This is in contrast to current schemes in the literature. Furthermore, we extend the optimization problems to the case where Bob is also equipped with multiple antennas, and provide insightful results, shown to be asymptotically optimal, accompanied by simulations.

I. INTRODUCTION

In covert communication (also known as Low Probability of Detection - LPD) Alice tries to reliably communicate a message to Bob, such that a watchful adversary Willie remains unaware of the presence of the communication. To make this possible, Alice may use the fact that the channels between all participants are subject to some kind of noise, and therefore she can try to hide her communication within the margin of uncertainty at Willie’s decoder. In fact, for AWGN channels, it was shown in [1] that Alice can covertly transmit $O(\sqrt{n})$ bits in n channel uses (a.k.a the square root law). Extensions for binary symmetric, discrete memoryless, multiple access and fading channels were done in [2]–[6], respectively.

This law essentially means that the transmission rate goes to zero with n ; however, subsequent works showed that $O(n)$ bits in n channel uses can be achieved, namely, a strictly positive rate, if Willie suffers from some kind of uncertainty in his received noise power [7]–[9]. The uncertainty may be a result of inaccurate knowledge of his noise or a result of an active node which confuses Willie (e.g., a jammer that varies his noise power randomly). The ability to achieve a strictly

positive rate is of great importance since not only Alice can transmit a meaningful amount, but existing coding schemes can be used instead of designing special codes which are suitable only for covert communication.

The limits of covert communication in a multiple-antenna setting were first established in [10]. Therein, it was shown that in case Alice is equipped with multiple antennas, her best strategy is to perform beamforming towards Bob, which results with a multiplicative constant gain to the square root law, by the number of independent paths between her and Bob. However, the case where such a communication channel includes a jammer which is equipped with multiple antennas is still open and remains unclear under various settings. For example, the knowledge the jammer has on the Channel State Information (in particular the CSI of Willie), and his preference on which user to assist, may affect his strategy, and the resulting rates, significantly.

In this work, we analyze the effect of multiple antennas at the jammer on covert communication, while assuming the jammer chooses to assist Alice and Bob. For simplicity, we assume that Alice and Willie have a single antenna. Note that a strictly positive power (hence, rate) can be achieved when there is a jammer with a single antenna ([9]) and that the limits for MIMO settings without a jammer were already examined in [10]. Therefore, adding more antennas to Alice and Willie won’t contribute much additional insight.

The jammer’s assistance comes in the form of transmitting Artificial Noise (AN) while using all his multiple antennas. Thus, his transmission strategy is reflected by the covariance matrix for the random vector he chooses. In a way, this covariance matrix defines the power allocation and the directions for that allocation. Accordingly, we analyze the behavior of Alice’s transmission power as a function of the jammer’s strategy which, to the best of our knowledge, was not yet examined under the settings of covert communication. Recently, and independently with this work, covert communication in the presence of a multi-antenna jammer was studied in [11]. Therein, the authors *fixed a certain jamming strategy*, of transmitting AN diffusely into the null-space of Bob, and optimized the covert rate by optimizing Alice’s strategy (also having multiple antennas). In this work, we show that such a strategy is not always optimal, for both cases of known and unknown CSI of Willie, and show that under certain scenarios, transmitting AN outside the null-space of Bob, although it adds noise to his reception, is still better, since it is able to significantly jam Willie’s receiver.

The notion of AN transmission was also considered for wiretap channels, where the AN was used to improve the

arXiv:2001.03146v2 [cs.IT] 26 Jul 2020

secrecy rate of the system. These works can be roughly divided into two cases, which differ by the source of the AN transmission. In the first case, it is assumed that the transmitter (Alice) is equipped with multiple antennas and she performs simultaneously the secret transmission and the AN. E.g., [12]–[17]. In the second case, there exist an additional node(s) in the system, e.g., a relay(s) [18]–[21] or a dedicated jammer [22], [23], which output AN to degrade the eavesdropper’s channel. At the heart of the analysis of these mentioned works, one can find several optimization problems depending on the models’ assumptions. A major factor in all is the CSI, which is known or unknown to the nodes in the system. Regardless of this knowledge, it was shown in [13] that the transmission of AN can guarantee a minimal secrecy rate. This result is consistent with the fact that a positive rate is also achievable in covert communication problems when there exists uncertainty at Willie’s detector. Similarly to the wiretap channel problems, in this work, we are faced also with optimization problems for the optimal AN transmission strategy of the jammer. However, these optimization problems are originated from different settings, and therefore require different treatment due to the covertness requirement.

Main Contribution: In this work, we explore the effect of multiple antennas of an uninformed jammer on covert communication. We analyze the covert rate as a function of the jammer’s transmission strategy. We choose to use Bob’s received SNR as our figure of merit since in most cases the actual rate of any coding scheme is an increasing function of the SNR. This is, of course, true for the AWGN channel as well. Specifically, we consider the optimization problems that arise when maximizing the received SNR at Bob by optimizing the jammer’s strategy, defined by his covariance matrix. As part of the solution for these optimization problems, we concentrate on a covertness-achieving transmission scheme (construction) for Alice. Specifically, we provide the transmission power as a function of the jammer’s transmission strategy for the cases where CSI on Willie is globally known or unknown. We show that this power ensures that the system is covert; that is, Willie has nothing better to do besides guessing if communication occurred or not. We show that multiple antennas at the jammer provide additional gain for Alice’s transmission power with respect to the case of a single antenna jammer. Furthermore, when the CSI of Willie is unknown, we show that the transmission power of Alice is an increasing function of the number of antennas the jammer has.

Provided that the system is covert, i.e., assuming Alice uses the covertness achieving construction, we solve the optimization problem which turns out to be not trivial, as the jammer’s strategy affects Bob’s SNR in opposite directions. Nevertheless, under this construction, we solve both optimization problems for the cases of known and unknown CSI of Willie. The solutions essentially describe the directions and power allocations the jammer’s transmission should use. When the CSI is known, the optimal transmission strategy for the jammer is to transmit in a single direction, with all of his available power. In general, this direction is not trivial, since

it takes into account both the channel coefficients to Bob as well as those to Willie. Thus, the transmission direction has a tradeoff between being orthogonal to Bob and in the direction of Willie. When the CSI of Willie is unknown, yet the CSI of Bob is known to the jammer, we show that transmitting the AN diffusely into the null-space of Bob is not always optimal.

Furthermore, we consider the case where Bob is also equipped with multiple antennas and uses a linear receiver. This provides even further depth and complexity to the optimizations in question, as Bob takes an active part in the decoding process. We extend the optimization problems for this case and present several transmission strategies that the jammer should take to assist Alice as much as possible, as well as the detection strategies Bob should employ when he has multiple antennas. The suggested strategies are compared to the optimal solution using numerical results. For the low SNR regime, some of the suggested strategies are shown to be asymptotically optimal.

II. SYSTEM MODEL

We consider a system in which Alice ("a" in the channel coefficients notation) wishes to communicate covertly with Bob ("b") while Willie ("w") remains uncertain of this communication. Also, we assume that there is a third participant, the jammer ("j"), who assists Alice and Bob. This paper focuses on covert communication with a positive covert rate. Thus, we assume that the uninformed jammer assists Alice in creating uncertainty at Willie’s decoder continuously, regardless of whether or not Alice transmits (Similarly to [9]). The model is depicted in Figure 1.

Alice and Willie are equipped with a single antenna, while Bob and the jammer are equipped with M and N antennas, respectively. The channel between all participants is subject to block fading and AWGN. In this setting, Willie tries to detect whether transmission by Alice was made or not, by performing a statistical hypothesis test on his received signal. In this test, the null hypothesis \mathcal{H}_0 is that Alice does not transmit, while the hypothesis \mathcal{H}_1 is that Alice transmits. Throughout, lower case letters represent random variables, bold lower case represent random vectors, and bold upper case represent random matrices. Thus, under each of the hypotheses, the received signals at Bob and Willie in the i th channel use are

$$\begin{aligned} \mathcal{H}_1 : \quad & \mathbf{y}_b[i] = x[i]\mathbf{h}_{ab} + \mathbf{H}_{jb}\mathbf{u}[i] + \mathbf{n}_b[i] \\ & y_w[i] = x[i]h_{aw} + \mathbf{u}[i]^T\mathbf{h}_{jw} + n_w[i] \\ \mathcal{H}_0 : \quad & \mathbf{y}_b[i] = \mathbf{H}_{jb}\sqrt{P_j}\mathbf{u}[i] + \mathbf{n}_b[i] \\ & y_w[i] = \sqrt{P_j}\mathbf{u}[i]^T\mathbf{h}_{jw} + n_w[i], \end{aligned} \quad (1)$$

where $x[i]$ is the complex symbol transmitted by Alice in the i th channel use, with average power P_a (i.e., $\mathbb{E}[|x[i]|^2] = P_a$) and $\mathbf{u}[i] = (w_1[i], w_2[i], \dots, w_N[i])^T$ is the vector transmitted by the jammer at the i th channel use, with a covariance matrix $\mathbf{\Sigma} = \mathbb{E}[\mathbf{u}[i]\mathbf{u}[i]^\dagger]$ such that $\mathbb{E}[\mathbf{u}[i]^\dagger\mathbf{u}[i]] = 1$. $\mathbf{u}[i]$ is multiplied by $\sqrt{P_j}$ where P_j is the total transmission power of the jammer.

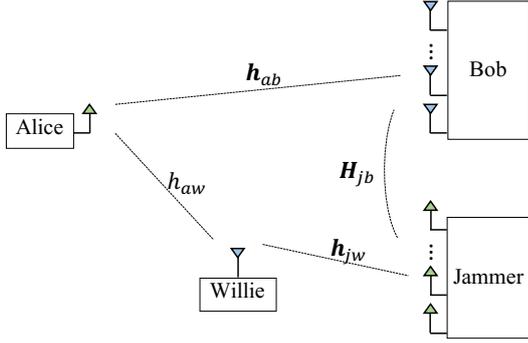


Fig. 1. Covert communication model where an independent jammer who assist the communication by transmitting AN.

The channel coefficients between Alice, Willie and Bob are $h_{aw} \in \mathbb{C}$ and $h_{ab} \in \mathbb{C}^M$, respectively. $h_{jw} \in \mathbb{C}^N$ and $\mathbf{H}_{jb} \in \mathbb{C}^{M \times N}$ are the channel coefficients from the jammer to Willie and Bob, respectively. These channel coefficients are assumed to have a zero-mean complex Gaussian distribution with unit variance and are considered to be fixed for the period of n channel uses (a slot). In addition, both Bob and Willie endure complex additive Gaussian noise denoted by $\mathbf{n}_b \sim \mathcal{CN}(0, \sigma_b^2 \mathbf{I}_{M \times M})$ and $n_w \sim \mathcal{CN}(0, \sigma_w^2)$.

The CSI of Willie affects the ability of the jammer to contribute to the covert communication between Alice and Bob. We thus separate the scenarios where the CSI of Willie is globally known or unknown. When the CSI is known, we assume all nodes know h_{aw} and h_{jw} . When the CSI is unknown, only the statistics of h_{aw} and h_{jw} are known to all nodes. In any case, we assume that Willie has full CSI of all channel coefficients and that Alice, Bob and the jammer know \mathbf{H}_{jb} and h_{ab} .

The jammer's assistance comes in the form of AN with a total power P_j using his multiple antennas. Since he has N antennas, he is free to choose how to allocate P_j and to which direction to transmit. Hence, the transmitted vector by the jammer is $\mathbf{u}[i] \sim \mathcal{CN}(0, \mathbf{\Sigma})$ where the total power P_j is allocated according to $\mathbf{\Sigma}$. Again, such AN assistance does not impair the assumption that the jammer is uninformed. Note that the case in which the jammer has some kind of knowledge on when Alice is transmitting results in a very different problem, for which other type of assistance may be optimal.

To create uncertainty at Willie's decoder, the jammer must vary his total power P_j , independently in each n channel uses [8], [9], [24]. That is, the value of P_j must be unknown to Willie and changed in a way which he cannot estimate efficiently. Therefore, following similar assumptions as in [9], [24], we assume that P_j is a uniform r.v. on $[0, P_{max}]$ with the probability density function (pdf) given by,

$$f_{P_j}(x) = \begin{cases} \frac{1}{P_{max}} & \text{if } 0 \leq x \leq P_{max} \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

and it is redrawn every n channel uses independently. The

realization of P_j in each slot is denoted by p_j . Throughout, we assume that P_{max} is a fixed parameter of the system and it is known to all. Note also that if Willie had not known the channel coefficients, the jammer could have used a noise distribution with constant variance, since the uncertainty would arise from the random channel coefficients [9].

As stated above, the jammer allocates its power P_j in each slot according to the covariance matrix $\mathbf{\Sigma}$. His goal is to choose $\mathbf{\Sigma}$ which assists Alice and Bob as much as possible. Without loss of generality, the jammer transmits with $d \leq N$ directions with the fraction of power allocated to each direction represented by the vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)^T$ such that $\xi_l > 0$ and $\sum_{l=1}^d \xi_l = 1$. Thus, the covariance matrix $\mathbf{\Sigma}$ is represented as $\mathbf{\Sigma} = \mathbf{V}\mathbf{X}\mathbf{V}^\dagger$ where \mathbf{X} is a diagonal matrix with $\boldsymbol{\xi}$ as its elements and \mathbf{V} is a matrix with the corresponding directions as its columns. Note that under this definition, the matrix \mathbf{X} is of dimension $d \times d$ and \mathbf{V} is of dimension $N \times d$.

Similar to previous works on covert communication ([1], [8], [9], [24]), we assume that Alice and Bob share a codebook which is used only once and is not revealed to Willie; however, Willie knows its statistics.

Remark 1: Since the jammer assists Alice and Bob, this model and its analysis are also suitable, under some assumptions, to the situation where Bob is equipped with a full-duplex transceiver and he is the source of the AN. Moreover, in this case, any assumptions on the channel knowledge between them become trivial.

A. Covert Criteria

Upon receiving the vector \mathbf{y}_w , Willie performs hypothesis testing to determine if transmission by Alice took place or not. That is, he tries to distinguish between the two hypotheses \mathcal{H}_0 and \mathcal{H}_1 . The optimal test for that manner, in terms of minimizing the probability of error, is to apply the Neyman-Pearson criterion, resulting in the likelihood ratio test:

$$\frac{\mathbb{P}_1}{\mathbb{P}_0} > \eta, \quad (3)$$

$$\mathbb{P}_0 < \eta$$

where \mathbb{P}_0 and \mathbb{P}_1 are the probability distributions of Willie's observations under the hypotheses \mathcal{H}_0 and \mathcal{H}_1 , respectively, and η is a threshold which Willie can choose to trade-off between his two error probabilities, P_{MD} and P_{FA} . P_{MD} is the probability of miss detection in case a transmission occurred and P_{FA} is the probability of false alarm in case a transmission did not occur. In this work, we follow the standard criteria for covertness in the covert communication literature [1], [8], [9], [25]. Namely, we have the following.

Definition 1 (Covert Communication criteria): Alice's transmission is considered covert if

$$P_{MD} + P_{FA} \geq 1 - \epsilon, \quad (4)$$

where $\epsilon > 0$ is the covertness requirement.

Throughout, we assume that ϵ is a fixed parameter of the system and it is known to all. Note that this criterion is

acceptable for the following reason. Willie can easily choose a strategy with $P_{FA} = 0$ and $P_{MD} = 1$, by simply declaring \mathcal{H}_0 at all times, regardless of his channel measurements. Analogously, $P_{FA} = 1$ and $P_{MD} = 0$ are achieved by always declaring \mathcal{H}_1 . Requiring $P_{MD} + P_{FA} \geq 1 - \epsilon$ is therefore equivalent to forcing Willie to at most time-share between these two trivial strategies.

B. Performance Metric and Reliability

In this work, we examine the *received SNR at Bob such that Alice's transmission is covert for a given covertness requirement ϵ* . In what follows, we describe Bob's received SNR for fixed Σ, p_j and a fixed transmission power from Alice. First, for the case where Bob has a single antenna (i.e., $M = 1$) and then for the case of multiple antennas which Bob may use beneficially. Such a distinction will help us focus only on the jammer's strategy first. Clearly, the SNR is affected by Alice's transmission power, regardless of the way it is chosen. In this work, we set this power such that the covertness constraint is met. Specifically, in Corollary 2 given later on, we provide a sufficient condition for Alice's power, and prove that the resulting value is a function of solely Σ , hence we write $P_a = P_a(\Sigma)$. Thus, in case Alice transmitted with some power $P_a(\Sigma)$, the received SNR of Bob is given as follows:

$$SNR_b^1 = \frac{P_a(\Sigma)|h_{ab}|^2}{p_j \mathbf{h}_{jb}^\dagger \Sigma \mathbf{h}_{jb} + \sigma_b^2} = \frac{P_a(\mathbf{V}, \mathbf{X})|h_{ab}|^2}{p_j \mathbf{h}_{jb}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jb} + \sigma_b^2}, \quad (5)$$

where \mathbf{H}_{jb} , in this case, is a single column matrix denoted as \mathbf{h}_{jb} .

When Bob has multiple antennas, he can take an active part in the communication strategy. For example, by steering his antennas away from the jammer. We assume that Bob uses a linear receiver. In this case, Bob performs a linear operation on the received signal; specifically, Bob projects the received vector onto a subspace which on one hand diminishes the effect of the AN from the jammer and on the other intensifies Alice's transmission. That is, we have

$$\mathbf{c}^T \mathbf{y}_b[i] = x[i] \mathbf{c}^T \mathbf{h}_{ab} + \mathbf{c}^T \mathbf{H}_{jb} \sqrt{p_j} \mathbf{u}[i] + \mathbf{c}^T \mathbf{n}_b[i],$$

where \mathbf{c} denotes the linear filter. The output of this receiver has SNR

$$\frac{P_a(\Sigma) |\mathbf{c}^\dagger \mathbf{h}_{ab}|^2}{\mathbf{c}^\dagger (p_j \mathbf{H}_{jb} \Sigma \mathbf{H}_{jb}^\dagger + \sigma_b^2 \mathbf{I}) \mathbf{c}} = \frac{P_a(\mathbf{V}, \mathbf{X}) |\mathbf{c}^\dagger \mathbf{h}_{ab}|^2}{\mathbf{c}^\dagger (p_j \mathbf{H}_{jb} \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{H}_{jb}^\dagger + \sigma_b^2 \mathbf{I}) \mathbf{c}}. \quad (6)$$

The SNR is sufficient to derive the covert rate for the block fading and AWGN channels as it is an injective (and increasing) function of the SNR [26]. However, considering the CSI knowledge assumptions taken in this work, the SNR expressions in (5) and (6) are random as they are a function of the varying AN noise variance P_j . Thus, to show that there exist a positive covert rate R (without guaranteeing the maximum rate achievable) such that Bob can decode the transmission successfully with a probability of error that goes to zero, we note the following. First, p_j is unknown to Bob

and Alice; however, since the statistics of P_j is known to all, Alice and Bob can derive a non-trivial bound as they know P_{max} . Second, when the CSI of Willie is known, the jammer's strategy, i.e., Σ , can be computed directly from the specific realization of h_{aw} and \mathbf{h}_{jw} and the bound on p_j . When the CSI of Willie is unknown, we chose to maximize a lower bound on the SNR, hence show that the jammer's strategy is characterized only by the rank used for Σ . Specifically, under the assumptions made in this work, it is either $N - 1$ or N , depending on the value of p_j . Accordingly, from the above, Alice and Bob can derive a lower bound on the SNR and agree on the covert rate such that the communication is reliable.

We note that the SNR expressions above and the suggested analysis and results in this work can be used also for the scenario in which the channel coefficients \mathbf{h}_{ab} and \mathbf{H}_{jb} are unknown to Alice and Bob and only their statistics are known. In this case, there is a non-zero probability that the received SNR will be lower than any predefined value. For such fading channels, one of the acceptable performance metrics is the ϕ -outage capacity. This is the maximal transmission rate R such that the outage probability $P_r^{out}(R)$ is less than ϕ [26, Ch. 5]. Thus, for a given outage probability ϕ , we say that Alice and Bob can communicate reliably and covertly with a covert rate R and outage ϕ , if the covert criteria (4) is satisfied and $R < C_\phi$, where C_ϕ is the channel ϕ -outage capacity from Alice to Bob. Now, to find such R , one needs to devise a different lower bound on the SNR by using the known distribution of the channel's fading coefficients along with the bound on p_j . For example, one can use a global lower bound for the SNR, which is a function of the channel's fading coefficients, given in Section III-B, for that manner.

III. TRANSMISSION STRATEGIES AND ALICE'S TRANSMISSION POWER FOR $M = 1$

In this section, we analyze the jammer's AN transmission strategies along with the construction of a covertness achieving transmission scheme by Alice. Note that under our performance metric, Alice's transmission scheme is eventually measured by the power, $P_a(\Sigma)$, for a given ϵ and as a function of the jammer's strategy Σ .

We start by formulating an optimization problem for the AN transmission strategy, to maximize Bob's received SNR, provided that the system is covert. We distinguish between the cases of known and unknown CSI of Willie, which influences the transmission strategy. As mentioned in Section II, the Jammer's strategy is reflected by the covariance matrix $\Sigma = \mathbf{V} \mathbf{X} \mathbf{V}^\dagger$, which can be described by the matrices \mathbf{V} and \mathbf{X} alone since it assumed that the jammer uses all of his power, p_j , for transmission.

Accordingly, we define the problem as follows

$$\begin{aligned} & \max_{\mathbf{V}, \mathbf{X}} \frac{P_a(\mathbf{V}, \mathbf{X}) |h_{ab}|^2}{p_j \mathbf{h}_{jb}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jb} + \sigma_b^2}, \\ & \text{s.t.} \quad 0 \leq \xi_l \leq 1, \quad \sum_{l=1}^N \xi_l = 1, \\ & \quad \quad P_{MD} + P_{FA} \geq 1 - \epsilon. \end{aligned} \quad (7)$$

Proposition 1: Let $P_a^{c,}(\mathbf{V}, \mathbf{X})$ be the optimal transmission power of Alice to assure covertness, i.e., satisfying (4). Thus, the optimization problem (7) is equivalent to*

$$\begin{aligned} \max_{\mathbf{V}, \mathbf{X}} & \frac{P_a^{c,*}(\mathbf{V}, \mathbf{X}) |h_{ab}|^2}{p_j \mathbf{h}_{jb}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jb} + \sigma_b^2}, \\ \text{s.t.} & \quad 0 \leq \xi_l \leq 1, \quad \sum_{l=1}^N \xi_l = 1. \end{aligned} \quad (8)$$

Unfortunately, the optimization problem in (8) is non-convex and does not seem to have a closed form solution. This is mainly since P_{MD} and P_{FA} are affected by both the jammer's strategy (\mathbf{V}, \mathbf{X}) as well as Alice's choice of power. Thus, the utility (SNR) and the constraint (covertness) interplay non-trivially.

Consequently, we suggest a sub-optimal solution by providing a transmission scheme by Alice which assures a sufficient condition for covertness, for every \mathbf{V} and \mathbf{X} . This will allow us to provide an explicit expression for Alice's transmission power, ensuring that the system is covert as a function of \mathbf{V} and \mathbf{X} . That is, we first make sure (4) is satisfied and Willie is unable to decide if transmission occurred. We then find the optimal strategy by optimizing on \mathbf{V} and \mathbf{X} . Thus, in the remainder of this section, *an optimal strategy for the jammer refers to a one solving (8), yet with a lower bound $P_a^c(\mathbf{V}, \mathbf{X}) \leq P_a^{c,*}(\mathbf{V}, \mathbf{X})$, which is given in the following lemma.* Specifically, Lemma 1 below expresses Alice's transmission power, $P_a(\mathbf{V}, \mathbf{X})$, as a function of any specific strategy, i.e., a fixed $\mathbf{\Sigma}$ and p_j by the jammer, such that the communication is covert.

Lemma 1: Assume a block fading AWGN channel and a jammer with N antennas, who transmits AN with a fixed covariance matrix $\mathbf{\Sigma}$ and a total transmission power of p_j . For a given covertness requirement, $\epsilon > 0$, as long as Alice transmits with power

$$P_a \leq \frac{\epsilon P_{max}}{4|h_{aw}|^2} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}, \quad (9)$$

the system is covert, i.e., (4) applies and Willie is unable to decide if transmission occurred, and Alice's rate is strictly positive.

Proof of Lemma 1: Construction: We employ a Gaussian random codebook consisting of M messages of size n . That is, the codebook is generated by independently drawing symbols from a zero-mean complex Gaussian distribution with variance P_a and it is assumed to be used only once. When Alice wishes to transmit, she picks a codeword and transmits its n symbols as the sequence $\{x[i]\}_{i=1}^n$. Such a construction was also used in other covert communication works for the AWGN channel such as, [1], [9], [24]. In what follows, under this construction, given ϵ and a fixed $\mathbf{\Sigma}$, we show how to set P_a such that the covertness requirement (4) is satisfied. Note that this construction will be used also in the proof of Lemma 2.

Optimal hypothesis testing: The optimal test for Willie to distinguish between \mathcal{H}_0 and \mathcal{H}_1 , given in (3), can be written as

$$\frac{\mathbb{P}_1}{\mathbb{P}_0} = \frac{\prod_{i=1}^n f_{y_w[i]}^1}{\prod_{i=1}^n f_{y_w[i]}^0} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{>}} \eta, \quad (10)$$

where $f_{y_w[i]}^0$ and $f_{y_w[i]}^1$ are the probability distributions of Willie's observation in a single channel use under the hypotheses \mathcal{H}_0 and \mathcal{H}_1 , respectively. Note that we may write the joint distributions \mathbb{P}_0 and \mathbb{P}_1 as a multiplication of the marginal distributions since both the channel uses and the code are *i.i.d.* In particular, under \mathcal{H}_0 and given p_j , $f_{y_w[i]}^0$ is $\mathcal{CN}(0, \sigma_w^0)$, and under \mathcal{H}_1 and given p_j , $f_{y_w[i]}^1$ is $\mathcal{CN}(0, \sigma_w^1)$ where,

$$\begin{aligned} \sigma_w^0 &= \sigma_w^2 + p_j \mathbf{h}_{jw}^\dagger \mathbf{\Sigma} \mathbf{h}_{jw}, \\ \sigma_w^1 &= \sigma_w^2 + p_j \mathbf{h}_{jw}^\dagger \mathbf{\Sigma} \mathbf{h}_{jw} + P_a |h_{aw}|^2. \end{aligned} \quad (11)$$

The terms in the last line above reflect the self-noise power of Willie, the received AN power and the transmission power of Alice, respectively. It would be beneficial to express σ_w^i , $i \in \{0, 1\}$ in (11) using the representation $\mathbf{\Sigma} = \mathbf{V} \mathbf{X} \mathbf{V}^\dagger$ as follows

$$\begin{aligned} \sigma_w^0 &= \sigma_w^2 + p_j \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}, \\ \sigma_w^1 &= \sigma_w^2 + p_j \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw} + P_a |h_{aw}|^2. \end{aligned} \quad (12)$$

In [9], by using Fisher-Neyman factorization and likelihood ratio ordering techniques, it was shown that the above optimal ratio test is eventually an energy test on the average received power. This applies to our model as well, since from Willie's point of view, the models are the same. Specifically, the average received power, P_w^{rav} , is compared with a threshold τ ,

$$P_w^{rav} \triangleq \frac{1}{n} \sum_{i=1}^n |y_w[i]|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{>}} \tau. \quad (13)$$

One can realize that, given p_j , the average received power P_w^{rav} is a Gamma r.v. with parameters $k = n$ and $\theta = \frac{\sigma_w^i}{n}$ for $i = 0, 1$, i.e. $P_w^{rav} \sim \Gamma(n, \frac{\sigma_w^i}{n})$. Furthermore, as $n \rightarrow \infty$, by the weak law of large numbers, P_w^{rav} converges in probability to σ_w^i .

Covertness achieving P_a : Willie compares P_w^{rav} to a threshold τ ; however, this threshold depends on the distribution of P_j and thus may be optimized by Willie. The following analysis shows that for any threshold τ that Willie sets for himself, which is not known to anyone but Willie, there exists a construction by Alice such that (4) holds. Specifically, we bound each of the probabilities P_{MD} and P_{FA} for a given value of p_j resulting with sufficient conditions for covertness. Following similar conditioning arguments as in [9], for a fixed p_j , we have

$$P_{FA}(p_j) = \mathbb{P}_r(P_w^{rav} \geq \tau | \mathcal{H}_0, p_j). \quad (14)$$

Recall that given p_j and \mathcal{H}_0 , $P_w^{rav} \sim \text{Gamma}(n, \frac{\sigma_w^0}{n})$ and as $n \rightarrow \infty$, P_w^{rav} converges in probability to σ_w^0 . Thus, for any given covertness requirement $\epsilon > 0$, one can find $\delta^0(\epsilon, n) \rightarrow 0$ as $n \rightarrow \infty$ such that

$$\Pr(\sigma_w^0 - \delta^0 \leq P_w^{rav} \leq \sigma_w^0 + \delta^0 | \mathcal{H}_0, p_j) > 1 - \frac{\epsilon}{2}.$$

Note that the dependence of δ^0 on ϵ and n is omitted to ease notation. Since

$$\begin{aligned} \Pr(P_w^{rav} \geq \sigma_w^0 - \delta | \mathcal{H}_0, p_j) \\ &\geq \Pr(P_w^{rav} \geq \sigma_w^0 - \delta^0 | \mathcal{H}_0, p_j) \\ &> \Pr(\sigma_w^0 - \delta^0 \leq P_w^{rav} \leq \sigma_w^0 + \delta^0 | \mathcal{H}_0, p_j), \end{aligned}$$

for any $\delta(\epsilon) \geq \delta^0(\epsilon)$, then for any $\tau < \sigma_w^0 - \delta(\epsilon)$ we have

$$\Pr(P_w^{rav} \geq \tau | \mathcal{H}_0, p_j) > 1 - \frac{\epsilon}{2}. \quad (15)$$

Similarly for P_{MD} given p_j ,

$$P_{MD}(p_j) = \Pr(P_w^{rav} \leq \tau | \mathcal{H}_1, p_j). \quad (16)$$

Now, $P_w^{rav} \sim \text{Gamma}(n, \frac{\sigma_w^1}{n})$ and as $n \rightarrow \infty$, P_w^{rav} converges in probability to σ_w^1 . Thus, one can find $\delta^1(\epsilon, n) \rightarrow 0$ as $n \rightarrow \infty$ such that

$$\Pr(\sigma_w^1 - \delta^1 \leq P_w^{rav} \leq \sigma_w^1 + \delta^1 | \mathcal{H}_1, p_j) > 1 - \frac{\epsilon}{2}.$$

Again, since

$$\begin{aligned} \Pr(P_w^{rav} \leq \sigma_w^1 + \delta | \mathcal{H}_1, p_j) \\ &\geq \Pr(P_w^{rav} \leq \sigma_w^1 + \delta^1 | \mathcal{H}_1, p_j) \\ &> \Pr(\sigma_w^1 - \delta^1 \leq P_w^{rav} \leq \sigma_w^1 + \delta^1 | \mathcal{H}_1, p_j), \end{aligned}$$

for any $\delta(\epsilon) \geq \delta^1(\epsilon)$, then for any $\tau > \sigma_w^1 + \delta(\epsilon)$ we have

$$\Pr(P_w^{rav} \leq \tau | \mathcal{H}_1, p_j) > 1 - \frac{\epsilon}{2}. \quad (17)$$

Set $\delta(\epsilon, n) = \max\{\delta^0(\epsilon, n), \delta^1(\epsilon, n)\}$ where we note that $\delta(\epsilon, n) \rightarrow 0$ as $n \rightarrow \infty$ as well. Since σ_w^0, σ_w^1 and δ implicitly depend on p_j , define the set of intervals $\mathcal{P} = \{p_j : \sigma_w^0 - \delta < \tau < \sigma_w^1 + \delta\}$. Thus, for all $p_j \notin \mathcal{P}$, either (15) or 17 are satisfied and we have,

$$P_{MD}(p_j) + P_{FA}(p_j) \geq 1 - \frac{\epsilon}{2}. \quad (18)$$

Since P_j is a uniform r.v. we have

$$\begin{aligned} \Pr(\mathcal{P}) &= \Pr(\sigma_w^0 - \delta < \tau < \sigma_w^1 + \delta) \\ &= \Pr\left(\frac{\tau - \sigma_w^2 - P_a |h_{aw}|^2 - \delta}{\mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}} \leq P_j \leq \frac{\tau - \sigma_w^2 + \delta}{\mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}}\right) \\ &\leq \frac{P_a |h_{aw}|^2 + 2\delta}{P_{max} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}}, \end{aligned} \quad (19)$$

where the inequality in the last line is due to the possibility that the boundaries in the probability computation are outside the support of P_j (i.e. outside $[0, P_{max}]$). Note that the uniform assumption on P_j is critical for the result being independent of τ . Therefore, if we set $P_a = \frac{\epsilon P_{max}}{4|h_{aw}|^2} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}$ and

$\delta(\epsilon) = \frac{\epsilon P_{max}}{8} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}$, for sufficiently large n , we are left with

$$\Pr(\mathcal{P}) \leq \frac{\epsilon}{2}.$$

Considering all the above in order we have,

$$\begin{aligned} P_{MD} + P_{FA} &= \mathbb{E}_{P_j}[P_{MD}(P_j) + P_{FA}(P_j)] \\ &\geq \mathbb{E}_{P_j}[P_{MD}(P_j) + P_{FA}(P_j) | \mathcal{P}^c] \Pr(\mathcal{P}^c) \\ &> 1 - \epsilon. \end{aligned} \quad (20)$$

The above shows that as long as Alice transmits with power $P_a = \frac{\epsilon P_{max}}{4|h_{aw}|^2} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}$ the system is covert. The rate of Alice can be obtained by using P_a above in Bob's SNR, which can be lower bounded by a constant, providing a positive rate. ■

Lemma 1 provides the transmission power of Alice as a function of Willie's CSI and the jammer's strategy, in a form of a multiplicative gain $\mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw} / |h_{aw}|^2$. This multiplicative gain is with respect to the case of a single antenna jammer as given in [9], which equals $P_a = \frac{P_{max}\epsilon}{4}$ if we adjust their model assumptions to ours. Note that although h_{aw} and \mathbf{h}_{jw} are known, in the beginning of each slot, they are redrawn and thus P_a should be updated accordingly. This means that the shared codebook needs to be updated as well. However, in the next section we provide analysis for this random gain and the corresponding received SNR. Using this analysis, one can first realize that the probability of this gain to be equal to zero, is zero. This eventually ensures a positive covert rate. Furthermore, one can compute the probability of this gain to be above some value. That is, define an outage criteria such that Alice would not transmit if the rate is too low. Using this criteria, Alice and Bob can agree on a covert rate.

Our main results for $M = 1$ are presented in the following subsections with respect to the cases of known or unknown CSI of Willie.

A. Known CSI of Willie

Theorem 1: Assume a block fading AWGN channel and a jammer with N antennas, who transmits AN with covariance matrix Σ and a total transmission power of p_j . For a given covertness requirement $\epsilon > 0$ with globally known CSI of Willie, the optimal strategy for the jammer, i.e., the solution for the maximization problem in (8) while Alice is transmitting with the power given in Lemma 1, is the following power allocation,

$$\Sigma = \mathbf{v}^* \mathbf{v}^{*\dagger}, \quad (21)$$

where

$$\mathbf{v}^* = \frac{\mathbf{q}}{\|\mathbf{q}\|} \quad (22)$$

and \mathbf{q} is the eigenvector which corresponds to the highest eigenvalue of the matrix

$$\left(\mathbf{h}_{jb} \mathbf{h}_{jb}^\dagger + \sigma \mathbf{I}\right)^{-1} \left(\mathbf{h}_{jw} \mathbf{h}_{jw}^\dagger\right), \quad (23)$$

where $\sigma = \frac{\sigma_b^2}{p_j}$.

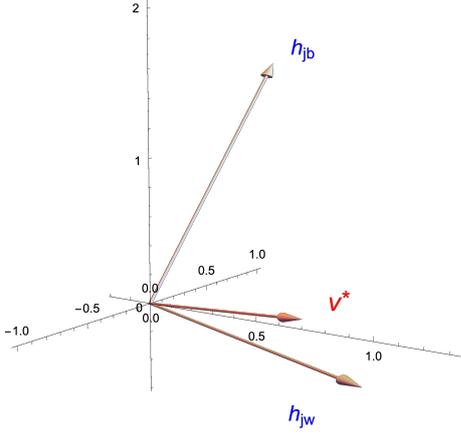


Fig. 2. An example for the optimal direction of the AN transmission by the jammer for a specific realization of the channel fading coefficients for the case of $M = 1$ and $N = 3$.

Hence the optimal strategy is to transmit AN in a single direction. The optimal direction, \mathbf{v}^* , depends on both channel vectors \mathbf{h}_{jw} and \mathbf{h}_{jb} . While it is hard to gain insight on the actual direction from (23) one can gain intuition from Equation (28) in the proof of Theorem 1, where, it is clearly seen that \mathbf{v}^* , on one hand, should be close to the direction of Willie, i.e., maximize the projection on \mathbf{h}_{jw} , while on the other hand it should be orthogonal to Bob as much as possible, i.e., minimize the projection on \mathbf{h}_{jb} . Figure 2 depicts a visualization for a specific channel realization for $N = 3$ antennas. One can easily observe that the optimal direction is indeed pointed towards Willie while being close to orthogonal to the direction of Bob.

Proof of Theorem 1: Setting Alice's transmission power in the maximization problem (8) according to (9) guarantees that the transmission is covert. Then, we have

$$\max_{\mathbf{V}, \mathbf{X}} \frac{\frac{\epsilon P_{max}}{4|h_{aw}|^2} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw} |h_{ab}|^2}{p_j \mathbf{h}_{jb}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jb} + \sigma_b^2}}. \quad (24)$$

To solve the above, we first find the optimal \mathbf{X} for any given \mathbf{V} . Accordingly, for a fixed \mathbf{V} , similar to [27], we further simplify (24) as follows

$$= c_1 \frac{P_{max}}{p_j} \max_{\mathbf{X}} \frac{\mathbf{w}^\dagger \mathbf{X} \mathbf{w}}{\mathbf{b}^\dagger \mathbf{X} \mathbf{b} + \sigma}, \quad (25)$$

where $c_1 = \epsilon |h_{ab}|^2 / 4 |h_{aw}|^2$ depends on Alice's channels, $\sigma = \frac{\sigma_b^2}{p_j}$, $\mathbf{w} = \mathbf{V}^\dagger \mathbf{h}_{jw}$ and $\mathbf{b} = \mathbf{V}^\dagger \mathbf{h}_{jb}$. Note that

$$\frac{\mathbf{w}^\dagger \mathbf{X} \mathbf{w}}{\mathbf{b}^\dagger \mathbf{X} \mathbf{b} + \sigma} = \frac{\sum_{l=1}^N \xi_l w_l^2}{\sum_{l=1}^N \xi_l b_l^2 + \sigma}.$$

We wish to show that the optimal power allocation ξ^* (which defines \mathbf{X}), for a fixed \mathbf{w} and \mathbf{b} , is a unit vector. To this end, we examine two indices i and j in ξ^* which have power allocations (ξ_i, ξ_j) such that $\xi_i + \xi_j = \rho_{ij}$, where $\rho_{ij} > 0$ is some constant. We will show first that either $\xi_i = \rho_{ij}$ or $\xi_j = \rho_{ij}$ must occur, hence, eventually, the optimal power

allocation is a unit vector (since for any two indices i and j , it is always better to move the allocation from one to the other, until one is exhausted).

The optimization problem on ξ can be written as follows,

$$\begin{aligned} \max_{\xi} f(\xi) &\triangleq \max_{\xi} \frac{\sum_{l=1}^N \xi_l w_l^2}{\sum_{l=1}^N \xi_l b_l^2 + \sigma} \\ &= \max_{\xi} \frac{\sum_{l \neq i, j} \xi_l w_l^2 + w_i^2 \xi_i + w_j^2 \xi_j}{\sum_{l \neq i, j} \xi_l b_l^2 + b_i^2 \xi_i + b_j^2 \xi_j + \sigma} \\ &= \max_{\xi} \frac{\sum_{l \neq i, j} \xi_l w_l^2 + w_i^2 \xi_i + w_j^2 (\rho_{ij} - \xi_i)}{\sum_{l \neq i, j} \xi_l b_l^2 + b_i^2 \xi_i + b_j^2 (\rho_{ij} - \xi_i) + \sigma} \\ &= \max_{\xi} \frac{\sum_{l \neq i, j} \xi_l w_l^2 + \xi_i (w_i^2 - w_j^2) + w_j^2 \rho_{ij}}{\sum_{l \neq i, j} \xi_l b_l^2 + \xi_i (b_i^2 - b_j^2) + b_j^2 \rho_{ij} + \sigma}. \end{aligned}$$

Taking the derivative of the above according to ξ_i shows that the function $f(\xi)$ is either monotonically increasing or monotonically decreasing with ξ_i depending on the sign of $(w_i^2 - w_j^2) \left(\sum_{l \neq i, j} \xi_l b_l^2 + \sigma + b_j^2 \rho_{ij} \right) - (b_i^2 - b_j^2) \left(\sum_{l \neq i, j} \xi_l w_l^2 + w_j^2 \rho_{ij} \right)$.

Thus, for every two indices i, j , if $f(\xi)$ is monotonically decreasing in ξ_i , set $\xi_j = \rho_{ij}$. On the other hand, if $f(\xi)$ is monotonically increasing in ξ_i , set $\xi_j = 0$. This essentially means that the optimal strategy of the jammer is to allocate all his power towards a single direction.

In order to find this direction, which is the corresponding eigenvector \mathbf{v} , we may write the unit rank matrix Σ as $\Sigma = \mathbf{v} \mathbf{v}^\dagger$. Note that \mathbf{v} is constrained to have a unit norm, i.e., $\mathbf{v}^\dagger \mathbf{v} = 1$. Returning to the maximization problem in (24), we have,

$$c_1 \frac{P_{max}}{p_j} \max_{\mathbf{V}, \mathbf{X}} \frac{\mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}}{\mathbf{h}_{jb}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jb} + \sigma} \quad (26)$$

$$= c_1 \frac{P_{max}}{p_j} \max_{\mathbf{v}} \frac{\mathbf{h}_{jw}^\dagger \mathbf{v} \mathbf{v}^\dagger \mathbf{h}_{jw}}{\mathbf{h}_{jb}^\dagger \mathbf{v} \mathbf{v}^\dagger \mathbf{h}_{jb} + \sigma \mathbf{v}^\dagger \mathbf{v}} \quad (27)$$

$$= c_1 \frac{P_{max}}{p_j} \max_{\mathbf{v}} \frac{\mathbf{v}^\dagger \mathbf{h}_{jw} \mathbf{h}_{jw}^\dagger \mathbf{v}}{\mathbf{v}^\dagger \mathbf{h}_{jb} \mathbf{h}_{jb}^\dagger \mathbf{v} + \sigma \mathbf{v}^\dagger \mathbf{v}} \quad (28)$$

$$= c_1 \frac{P_{max}}{p_j} \max_{\mathbf{v}} \frac{\mathbf{v}^\dagger \mathbf{h}_{jw} \mathbf{h}_{jw}^\dagger \mathbf{v}}{\mathbf{v}^\dagger (\mathbf{h}_{jb} \mathbf{h}_{jb}^\dagger + \sigma \mathbf{I}) \mathbf{v}} \quad (29)$$

$$= c_1 \frac{P_{max}}{p_j} \max_{\mathbf{v}} \frac{\mathbf{v}^\dagger \mathbf{W} \mathbf{v}}{\mathbf{v}^\dagger \mathbf{B} \mathbf{v}}, \quad (30)$$

where, $\mathbf{W} = \mathbf{h}_{jw} \mathbf{h}_{jw}^\dagger$, and, $\mathbf{B} = \mathbf{h}_{jb} \mathbf{h}_{jb}^\dagger + \sigma \mathbf{I}$. The above maximization problem is also known as the Rayleigh quotient [28], which is maximized by the eigenvector that corresponds to the highest eigenvalue of the matrix $\mathbf{B}^{-1} \mathbf{W}$. Denote this vector by \mathbf{q} , the optimal \mathbf{v} is,

$$\mathbf{v}^* = \frac{\mathbf{q}}{\|\mathbf{q}\|}.$$

■

B. Unknown CSI of Willie

We revisit Equation (9). When \mathbf{h}_{jw} and h_{aw} are unknown to Alice, she cannot determine her transmission power in order to stay covert. Moreover, without \mathbf{h}_{jw} , it is not clear how to perform the optimization on $\mathbf{h}_{jw}^\dagger \boldsymbol{\Sigma} \mathbf{h}_{jw}$, in order to maximize Bob's SNR. We thus take the following course of action. We first lower bound Bob's SNR by a separable function, that will facilitate the optimization on $\boldsymbol{\Sigma}$. We then provide a sufficient condition for Alice's power, independent of \mathbf{h}_{jw} , yet using the optimal $\boldsymbol{\Sigma}$ which resulted from the above optimization. The result is a strictly positive transmission power for Alice, a strictly positive SNR of Bob, ϵ -covert transmission, with the optimal jamming strategy given the lower bound.

Accordingly, the following proposition provides a global lower bound on Bob's SNR, for any realization of the channel's coefficients, if Alice uses the same code construction as given in the proof of Lemma 1.

Proposition 2: Using the code construction given in the proof of Lemma 1, for any realization of Willie's CSI, Bob's received SNR is lower bounded by

$$\frac{c_1 P_{max}}{p_j} \frac{\|\mathbf{h}_{jw}\|^2 \min_{i=1,\dots,d} \xi_i}{\|\mathbf{h}_{jb}\|^2 \max_{i=1,\dots,d} \xi_i + \sigma}. \quad (31)$$

Proof: Setting Alice's transmission power in the maximization problem (8) according to (9) while using the notations of Theorem 1, we have

$$\frac{c_1 P_{max}}{p_j} \frac{\mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}}{\mathbf{h}_{jb}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jb} + \sigma}. \quad (32)$$

For any vector \mathbf{h} and matrices \mathbf{V} , \mathbf{X} , we have [28, Theorem 2.4.3]:

$$\|\mathbf{h}\|^2 \min_{i=1,\dots,d} \xi_i \leq \mathbf{h}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h} \leq \|\mathbf{h}\|^2 \max_{i=1,\dots,d} \xi_i, \quad (33)$$

and the bound follows immediately. \blacksquare

For covert communication, a natural choice would be to maximize the lower bound, for example, to ensure that the SNR is above a certain level with very high probability. This may be a very important promise since it enables Alice to communicate while remaining undetected. We thus assume for the rest of this section that the jammer strategy is subject to maximizing the lower bound on Bob's SNR given in (31). The following proposition provides the jammer strategy for this case.

Proposition 3: To maximize the lower bound in (31), the jammer should direct the AN towards d directions, with all of his available power equally divided between them. That is $\boldsymbol{\Sigma} = \mathbf{V} \mathbf{X} \mathbf{V}^\dagger$, with \mathbf{V} being a unitary matrix of dimension $N \times d$ with rank d and $\mathbf{X} = \frac{1}{d} \mathbf{I}_{d \times d}$.

Proof: The proof follows since $\sum_i^d \xi_i = 1$ and we wish to maximize the smallest ξ_i and minimize the highest ξ_i , thus letting $\xi_i = \frac{1}{d}$ for all i will maximize the lower bound in (31). \blacksquare

Note that the above proposition is consistent with Theorem 1, in which the jammer transmits in a single direction, i.e.,

$d = 1$ for the case of known CSI. On the other hand, when the CSI of Willie is unknown (\mathbf{h}_{jw} and h_{aw} are not known and random), to maximize the SNR, the jammer strategy should follow Proposition 3. In fact, transmitting AN uniformly in all directions is consistent with AN transmission strategies for the case of unknown CSI towards the adversary in the context of wiretap channel, e.g., [12], [13]. However, while the reason for transmitting uniformly across the basis of the null-space of Bob was intuitive in [12], [13], herein we prove that this choice actually maximises the lower bound on Bob's SNR.

Considering the above, and similarly to the previous subsection, we first express Alice's transmission power, $P_a^c(\mathbf{V}, \mathbf{X})$, which assures covertness for the case of unknown CSI of Willie.

Lemma 2: Assume a block fading AWGN channel and a jammer with N antennas, who transmits AN with his total transmission power, p_j , equally divided into d arbitrary directions. I.e., $\text{rank}(\mathbf{X}) = d$. For a given covertness requirement $\epsilon > 0$ when the CSI of Willie is unknown, as long as Alice transmits with power

$$P_a \leq \frac{\epsilon P_{max}}{12e \ln \frac{6}{\epsilon}} \left(\frac{\epsilon}{6}\right)^{\frac{1}{d}}, \quad (34)$$

the system is covert. I.e., (4) applies and Alice's rate is strictly positive.

Proof of Lemma 2: The construction is similar to the one given in Lemma 1. However, now, to obtain the covertness achieving power P_a , one needs to consider the channel's statistics for h_{aw} and \mathbf{h}_{jw} as well as P_j when computing the probability $\Pr(\mathcal{P})$ in equation (19). To do so, we upper bound $\Pr(\mathcal{P})$ and find P_a such that (4) applies. Specifically, we can write for all $\alpha > 0$

$$\begin{aligned} \Pr(\mathcal{P}) &= \Pr\left(\mathcal{P} \mid |h_{aw}|^2 > \alpha\right) \Pr(|h_{aw}|^2 > \alpha) + \\ &\quad \Pr\left(\mathcal{P} \mid |h_{aw}|^2 \leq \alpha\right) \Pr(|h_{aw}|^2 \leq \alpha) \quad (35) \\ &\leq \Pr(|h_{aw}|^2 > \alpha) + \Pr\left(\mathcal{P} \mid |h_{aw}|^2 \leq \alpha\right). \end{aligned}$$

Denoting $\Psi = \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}$ we can further upper bound the above for all $\beta > 0$

$$\begin{aligned} \Pr(\mathcal{P}) &\leq \Pr(|h_{aw}|^2 > \alpha) + \Pr\left(\mathcal{P} \mid |h_{aw}|^2 \leq \alpha\right) \\ &= \Pr(|h_{aw}|^2 > \alpha) \\ &\quad + \Pr\left(\mathcal{P} \mid |h_{aw}|^2 \leq \alpha, \Psi < \beta\right) \Pr(\Psi < \beta) \\ &\quad + \Pr\left(\mathcal{P} \mid |h_{aw}|^2 \leq \alpha, \Psi \geq \beta\right) \Pr(\Psi \geq \beta) \quad (36) \\ &\leq \Pr(|h_{aw}|^2 > \alpha) + \Pr(\Psi < \beta) \\ &\quad + \Pr\left(\mathcal{P} \mid |h_{aw}|^2 \leq \alpha, \Psi \geq \beta\right). \end{aligned}$$

Recall the definition of $\mathcal{P} = \{p_j : \sigma_w^0 - \delta < \tau < \sigma_w^1 + \delta\}$. If we substitute α and β for $|h_{aw}|^2$ and Ψ , respectively, we can upper bound the third term in the equation above and thus increase the probability. That is, set α and β in (19). This

replacement can only increase the length of the considered interval. Since the distribution is uniform, we have

$$\begin{aligned}
\Pr(\mathcal{P}) &\leq \Pr(|h_{aw}|^2 > \alpha) + \Pr(\Psi < \beta) \\
&\quad + \Pr\left(\mathcal{P} \mid |h_{aw}|^2 \leq \alpha, \Psi \geq \beta\right) \\
&\leq \Pr(|h_{aw}|^2 > \alpha) + \Pr(\Psi < \beta) \\
&\quad + \Pr\left(\frac{\tau - \sigma_w^2 - P_a\alpha - \delta}{\beta} \leq P_j \leq \frac{\tau - \sigma_w^2 + \delta}{\beta}\right) \\
&\leq \Pr(|h_{aw}|^2 > \alpha) + \Pr(\Psi < \beta) + \frac{2\delta + P_a\alpha}{\beta P_{max}}.
\end{aligned} \tag{37}$$

Since $|h_{aw}|^2$ is a $\chi^2(2)$ r.v., by setting $\alpha = 2 \ln \frac{\epsilon}{e}$ we have,

$$\Pr(|h_{aw}|^2 > \alpha) = \frac{\epsilon}{6}. \tag{38}$$

To obtain the distribution of Ψ , recall that the elements of the diagonal matrix \mathbf{X} are all equal to $1/d$, hence, $\Psi = \frac{1}{d} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{V}^\dagger \mathbf{h}_{jw} = \frac{1}{d} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{V}^\dagger \mathbf{h}_{jw}$. Since \mathbf{V} is a unitary matrix, $\mathbf{h}_{jw}^\dagger \mathbf{V}$ is distributed as a complex Gaussian vector of length d with independent variables. Accordingly, $d \cdot \Psi$ is a $\chi^2(2d)$ r.v.. This is also consistent with [29, Theorem 2, Ch.1], using a different approach. Thus, we compute

$$\begin{aligned}
\Pr(\Psi < \beta) &= \Pr(d\Psi < d\beta) \\
&\stackrel{(a)}{\leq} \left(\frac{d\beta}{2d} e^{(1-\frac{d\beta}{2d})}\right)^d \\
&\stackrel{(b)}{\leq} \left(\frac{\beta}{2} e\right)^d \\
&\stackrel{(c)}{\leq} \frac{\epsilon}{6}.
\end{aligned} \tag{39}$$

Where (a) follows from the Chernoff bound which requires that $0 < \beta < 2$ and (b) follows for such β . (c) is by setting $\beta = \frac{2}{e} \sqrt{\frac{\epsilon}{6}}$. Note that this choice of β satisfies the Chernoff's bound requirement for all $\epsilon \in (0, 1)$ and $d \in \mathbb{Z}^+ \setminus \{0\}$.

Thus, setting $P_a = \frac{\beta}{\alpha} \frac{\epsilon}{12} P_{max}$ and $\delta = \frac{\epsilon}{24} \beta P_{max}$ we have,

$$\Pr(\mathcal{P}) \leq \frac{\epsilon}{2}.$$

Continuing similar to the proof of Lemma 1, the covertness criteria $P_{MD} + P_{FA} \geq 1 - \epsilon$ holds. ■

In the following corollary, we present the covertness achieving transmission power for Alice in the case where the jammer has a single antenna (i.e., $N = 1$). We note that this result is different from the one obtained in [9], as the model assumptions are different. In [9], it was assumed that Willie does not know his own CSI, thus the jammer may use AN with constant variance $P_j = P_{max}$. Of course, this eventually affects Alice's power. In this work, we assume a stringent model where Willie knows his own CSI, therefore the communicating parties cannot assume that the channel randomization is enough to confuse Willie. Thus, the jammer must employ AN with varying power.

Corollary 1: Assume a block fading AWGN channel and a jammer with a single antenna and a total transmission power of p_j , where the CSI of Willie is unknown. For a given

covert requirement $\epsilon > 0$, as long as Alice transmits with power

$$P_a \leq \frac{\ln 1 - \frac{\epsilon}{6}}{\ln \frac{\epsilon}{6}} \frac{\epsilon}{12} P_{max}, \tag{40}$$

the system is covert. I.e., (4) applies and Alice's rate is strictly positive.

Proof: Following the same steps as the proof of Lemma 2, when the jammer is equipped with a single antenna, the received power of the AN at Willie is $p_j |h_{jw}|^2$. Thus, in the proof of Lemma 2, $\Psi = |h_{jw}|^2$, which is distributed as a $\chi^2(2)$ r.v.. Therefore, we may set $\beta = 2 \ln \frac{6}{6-\epsilon}$, which results in $\Pr(\Psi < \beta) = \frac{\epsilon}{6}$. ■

A more strict upper bound for Alice's transmission power can be obtained following to the inequality $\frac{\ln 1-x}{\ln x} > x\sqrt{x}$, for $0 < x < 1$, which results with $P_a \leq \left(\frac{\epsilon}{6}\right)^{2.5} \frac{P_{max}}{2}$.

Corollary 2: Given the jammer's strategy Σ , for the case of known and unknown CSI of Willie, for any fixed ϵ and P_{max} , a covertness achieving, strictly positive, power of Alice is attainable.

The proof follows immediately from the proofs of Lemmas 1 and 2 for the case of known and unknown CSI of Willie, respectively.

As Lemma 2 suggests, the covertness-achieving power of Alice for the case of unknown CSI, assuming equal power allocation, only depends on the number of directions the jammer transmits into. Again, such an assumption is motivated by the desire to promise minimal value for Alice's transmission power. Accordingly, the optimal strategy under such assumptions is given in the following theorem.

Theorem 2: Assume a block fading AWGN channel, and a jammer with N antennas, who transmits AN with covariance matrix Σ such that his total transmission power, p_j , is equally divided into d arbitrary directions. For a given covertness requirement $\epsilon > 0$ with unknown CSI of Willie, the optimal strategy for the jammer, is the following power allocation (for $N \geq 2$ antennas),

$$\Sigma = \mathbf{V}^* \mathbf{X}^* \mathbf{V}^{*\dagger}, \tag{41}$$

where in case $\sigma = \frac{\sigma_b^2}{P_j} < \frac{\|\mathbf{h}_{jb}\|^2}{N\left(\frac{\epsilon}{6} - \frac{1}{N(N-1)} - 1\right)}$,

$$\mathbf{V}^* = \mathbf{Q} \quad \text{and} \quad \mathbf{X}^* = \frac{1}{N-1} I_{(N-1) \times (N-1)}, \tag{42}$$

otherwise,

$$\mathbf{V}^* = \left(\mathbf{Q}, \frac{\mathbf{h}_{jb}}{\|\mathbf{h}_{jb}\|} \right) \quad \text{and} \quad \mathbf{X}^* = \frac{1}{N} I_{N \times N}, \tag{43}$$

where \mathbf{Q} is a matrix whose columns span the null-space of \mathbf{h}_{jb} .

The above result implies that when Willie's CSI is unknown, the jammer should transmit his AN either to the null-space of Bob or isotropically to all directions. The choice depends on the magnitude of the channel \mathbf{h}_{jb} . That is, for fixed N, σ_b^2 , and p_j , if $\|\mathbf{h}_{jb}\|$ is high, the optimal strategy would be to transmit only to the null-space. On the other hand, when $\|\mathbf{h}_{jb}\|$ is low, transmitting to all directions (including Bob's)

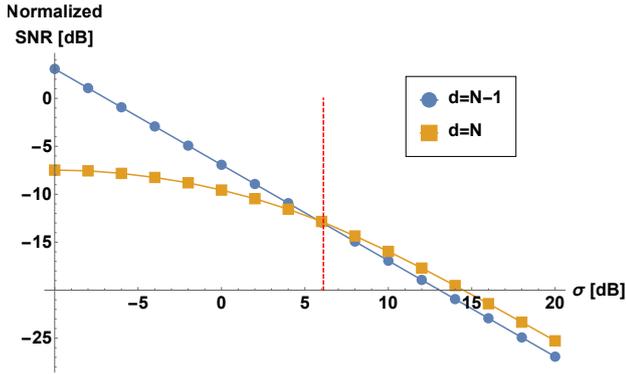


Fig. 3. Simulation results for Bob's received SNR as a function of σ for $N = 4$ antennas at the jammer, where the CSI of Willie is unknown. The SNR values are normalized so the leading constant coefficients is equal to 1.

would be better in the sense of maximizing the covert rate. An interpretation for the above is possible by considering the parameter σ which reflects the ratio between the "bad" noise power that Bob endures due to his own antenna noise and the "good" AN power which helps Alice and consequently helps Bob's SNR. When σ is low Bob endures more "bad" noise and thus the jammer should avoid adding more noise power in the direction of Bob. Note also that Theorem 2 refers only to the case of a multiple antenna jammer (i.e. $N \geq 2$). $N = 1$ is not considered since no optimization can be performed. Moreover, not transmitting AN is not possible as it results with zero covert rate (a square root law).

Figure 3 depicts the turning point of the average SNR value when comparing the strategy of transmitting the AN isotropically to all directions against transmitting only in the null-space. The simulation was performed for $\epsilon = 0.05$ and $N = 4$ antennas at the jammer. As described in Theorem 2, the figure depicts that around the point of $\sigma = \mathbb{E} \left[\frac{\|\mathbf{h}_{jb}\|^2}{N \left(\frac{\epsilon}{6}^{-1/N(N-1)} - 1 \right)} \right] = \frac{8}{4 \left(\frac{\epsilon}{6}^{-1/12} - 1 \right)} \approx 6.1 [dB]^1$ the jammer should switch strategy in order to maximize the SNR.

Proof of Theorem 2: Setting Alice's transmission power in the maximization problem (8) according to (34) guarantees that the transmission is covert. Then, considering that \mathbf{X} is a diagonal matrix with $\xi_l = \frac{1}{d}$ for all l , we have

$$\max_{\mathbf{V}, d} \frac{\frac{\epsilon P_{max}}{12e \ln \frac{\epsilon}{6}} \left(\frac{\epsilon}{6} \right)^{\frac{1}{d}} |h_{ab}|^2}{p_j \mathbf{h}_{jb}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jb} + \sigma_b^2},$$

$$s.t. \quad \xi_l = \frac{1}{d}, \quad l = 1, \dots, d.$$

The above can be simplified to

$$c_2 \frac{P_{max}}{p_j} \max_{d, \mathbf{b}} \frac{(\epsilon')^{\frac{1}{d}}}{\sum_{l=1}^d \frac{1}{d} b_l^2 + \sigma}, \quad (44)$$

where $c_2 = \frac{\epsilon |h_{ab}|^2}{2e \ln \epsilon'}$, $\epsilon' = \frac{\epsilon}{6}$, $\sigma = \frac{\sigma_b^2}{p_j}$ and $\mathbf{b} = \mathbf{V}^\dagger \mathbf{h}_{jb}$. Let \mathbf{Q} be a matrix whose columns span the null-space of \mathbf{h}_{jb} , and

¹ $\|\mathbf{h}_{jb}\|^2$ is distributed as a χ_{2N}^2 r.v.

let \mathbf{V} be a matrix with any d columns taken from the columns of \mathbf{Q} , i.e., \mathbf{V} is of dimension $N \times d$. Since the numerator is an increasing function of d and in the denominator we can have $b_l^2 = 0$ for $l = 1, \dots, N-1$ due to the choice of \mathbf{V} , any $d < N-1$, will result in a lower SNR. Thus, one needs to consider only the cases of $d = N-1$ and $d = N$, for which the target function in (44) is equal to

$$\begin{cases} \frac{(\epsilon')^{\frac{1}{N-1}}}{\sigma} & \text{if } d = N-1 \\ \frac{(\epsilon')^{\frac{1}{N}}}{\frac{1}{N} \|\mathbf{h}_{jb}\|^2 + \sigma} & \text{if } d = N. \end{cases} \quad (45)$$

When $\sigma > \frac{\|\mathbf{h}_{jb}\|^2}{N(\epsilon'^{-1/N(N-1)} - 1)}$ the SNR is larger for the case where $d = N$. Thus, in this case, \mathbf{V} must span the whole space. So, by setting $\mathbf{V}^* = \left(\mathbf{Q}, \frac{\mathbf{h}_{jb}}{\|\mathbf{h}_{jb}\|} \right)$ we have $b_N = \|\mathbf{h}_{jb}\|$ and the resulting value for the target function in (45). Otherwise, it is better to transmit in $d = N-1$ directions and only in the null-space of Bob, i.e., $\mathbf{V}^* = \mathbf{Q}$ which completes the proof. ■

IV. AN TRANSMISSION AND DETECTION STRATEGIES FOR $M > 1$ ANTENNAS AT BOB

In this section, we analyze the AN transmission strategy of the jammer and the detection strategy of Bob in case Bob has $M > 1$ antennas. When Bob is equipped with multiple antennas, the AN transmission and the detection strategies of the jammer and Bob are coupled and depend on the joint information both have.

Accordingly, we formulate a global optimization problem on both the transmission of the jammer and the detection of Bob where, similar to Section III, we restrict ourselves to constructions which satisfy the covertness requirement (4) for the cases of known and unknown CSI of Willie, respectively.

Recall that Bob employs a linear receiver. Thus, following the SNR expression in (6), and assuming that the system is covert (i.e. Alice is transmitting with a covertness achieving power $P_a^c(\Sigma)$), we may write the global optimization problem as follows,

$$\max_{\mathbf{V}, \mathbf{X}, \mathbf{c}} \frac{P_a^c(\mathbf{V}, \mathbf{X}) |\mathbf{c}^\dagger \mathbf{h}_{ab}|^2}{\mathbf{c}^\dagger \left(p_j \mathbf{H}_{jb} \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{H}_{jb}^\dagger + \sigma_b^2 \mathbf{I} \right) \mathbf{c}} \quad (46)$$

$$s.t. \quad 0 \leq \xi_l \leq 1 \text{ and, } \sum_{l=1}^N \xi_l = 1.$$

Note that \mathbf{c} should be optimized together with Σ (i.e., with \mathbf{V} and \mathbf{X}) which leads to yet another non-convex optimization problem. In the following subsections, we suggest strategies for Bob and the jammer which decentralized the global optimization above to obtain practical solutions for the cases of known and unknown CSI of Willie, respectively. The numerical results of the suggested strategies are shown to be very close to the optimal solution of (46), which can be acquired numerically.

A. Known CSI of Willie

The following lemma asserts that regardless of Bob's filter, under the covertness construction given in Lemma 1, the jammer should direct his AN to a single direction.

Lemma 3: For any linear filter \mathbf{c} , the optimal strategy for the jammer, i.e., the solution to (46) when the CSI of Willie is known, is to transmit all his available power towards a single direction.

Proof: Substituting Alice's transmission power given in (9) in (46), the maximization problem in (46) can be reduced as follows,

$$\begin{aligned} & \max_{\mathbf{V}, \mathbf{X}, \mathbf{c}} |\mathbf{c}^\dagger \mathbf{h}_{ab}|^2 \frac{\frac{\epsilon P_{max}}{4|h_{aw}|^2} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}}{\mathbf{c}^\dagger \left(p_j \mathbf{H}_{jb} \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{H}_{jb}^\dagger + \sigma_b^2 \mathbf{I} \right) \mathbf{c}} \\ &= \max_{\mathbf{V}, \mathbf{X}, \mathbf{c}} |\mathbf{c}^\dagger \mathbf{h}_{ab}|^2 \frac{\frac{\epsilon P_{max}}{4|h_{aw}|^2} \mathbf{h}_{jw}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{h}_{jw}}{\mathbf{c}^\dagger p_j \mathbf{H}_{jb} \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{H}_{jb}^\dagger \mathbf{c} + \sigma_b^2 \|\mathbf{c}\|^2} \\ &= \max_{\mathbf{b}', \mathbf{X}, \mathbf{c}} c_3 \frac{P_{max}}{p_j} |\mathbf{c}^\dagger \mathbf{h}_{ab}|^2 \frac{\mathbf{w}^\dagger \mathbf{X} \mathbf{w}}{\mathbf{b}'^\dagger \mathbf{X} \mathbf{b}' + \sigma \|\mathbf{c}\|^2}, \end{aligned} \quad (47)$$

where $c_3 = \epsilon/4|h_{aw}|^2$, $\sigma = \frac{\sigma_b^2}{P_j}$, $\mathbf{w} = \mathbf{V}^\dagger \mathbf{h}_{jw}$ and \mathbf{b}' is of the form $\mathbf{b}' = \mathbf{V}^\dagger \mathbf{H}_{jb}^\dagger \mathbf{c}$ where \mathbf{V} is any unitary matrix. Following the same analysis performed in the proof of Theorem 1, it can be shown that for any fixed \mathbf{w} , \mathbf{b}' and \mathbf{c} the optimal power allocation ξ^* (i.e., the diagonal of \mathbf{X}) is a unit vector. ■

Lemma 3 agrees with the conclusion of Theorem 1 for a single direction strategy (when $M = 1$) and shows that since Bob is equipped with several antennas yet a fixed filter, the jammer should transmit in a single direction. Accordingly, using Lemma 3, the maximization problem in (46) can be reduced as follows,

$$c_3 \frac{P_{max}}{p_j} \max_{\mathbf{v}, \mathbf{c}} \frac{\left(\mathbf{h}_{jw}^\dagger \mathbf{v} \mathbf{v}^\dagger \mathbf{h}_{jw} \right) |\mathbf{c}^\dagger \mathbf{h}_{ab}|^2}{\mathbf{c}^\dagger \left(\mathbf{H}_{jb} \mathbf{v} \mathbf{v}^\dagger \mathbf{H}_{jb}^\dagger + \sigma \mathbf{I} \right) \mathbf{c}}. \quad (48)$$

Inspired by existing works on MIMO communication, we suggest and analyze sub-optimal solutions for the above optimization problem. Specifically, we consider the null steering (also known as "masked beamforming") and Maximal Ratio Combiner (MRC) techniques ([30], [31]) which Bob and the jammer may perform independently to simplify the optimization problem. That is, instead of solving the global optimization for both strategies simultaneously, we fix one and perform optimization on the other. In a way, these sub-optimal schemes decentralized the global optimization problem by forcing the jammer and Bob to rely only on themselves and their ability to enhance the covert rate. Later on, in Corollary 7, we show that some of the suggested strategies are asymptotically optimal as σ grows.

In what follows we present analysis for the sub-optimal strategies described above. We solve the optimization problem in (48) first for a fixed \mathbf{c} and then for a fixed \mathbf{v} . Then, we suggest strategies for the jammer and Bob given a fixed \mathbf{c} or a fixed \mathbf{v} . Specifically, we first consider a fixed linear filter,

\mathbf{c} , as an MRC filter, to match the channel between Alice and Bob, ignoring the AN. Then fix \mathbf{v} by the jammer, to be in the direction of Willie. Lastly, we consider the strategy of canceling the AN at Bob's receiver by either the jammer or Bob where, depending on the number of antennas each has, they may agree in advance upon the responsibility for the cancelation task.

Fixed \mathbf{c} :

For any fixed \mathbf{c} we have the following theorem.

Theorem 3: For a fixed linear filter \mathbf{c} , the strategy for the jammer, i.e., the solution to (48) when the CSI of Willie is known, is the following power allocation,

$$\mathbf{\Sigma} = \mathbf{v}^* \mathbf{v}^{*\dagger}, \quad (49)$$

where \mathbf{v}^* is the eigenvector which corresponds to the highest eigenvalue of the matrix $\mathbf{B}^{-1} \mathbf{W}$, where $\mathbf{W} = \mathbf{h}_{jw} \mathbf{h}_{jw}^\dagger$ and, $\mathbf{B} = \tilde{\mathbf{h}} \tilde{\mathbf{h}}^\dagger + \sigma \|\mathbf{c}\|^2 \mathbf{I}$ for, $\tilde{\mathbf{h}} = \mathbf{H}_{jb}^\dagger \mathbf{c}$.

Proof: Assuming \mathbf{c} is fixed, the optimization problem in (48) is reduced to,

$$\begin{aligned} & \max_{\mathbf{v}} \frac{\mathbf{h}_{jw}^\dagger \mathbf{v} \mathbf{v}^\dagger \mathbf{h}_{jw}}{\mathbf{c}^\dagger \left(\mathbf{H}_{jb} \mathbf{v} \mathbf{v}^\dagger \mathbf{H}_{jb}^\dagger + \sigma \mathbf{I} \right) \mathbf{c}} \\ &= \max_{\mathbf{v}} \frac{\mathbf{v}^\dagger \mathbf{h}_{jw} \mathbf{h}_{jw}^\dagger \mathbf{v}}{\tilde{\mathbf{h}}^\dagger \mathbf{v} \mathbf{v}^\dagger \tilde{\mathbf{h}} + \sigma \|\mathbf{c}\|^2} \\ &= \max_{\mathbf{v}} \frac{\mathbf{v}^\dagger \mathbf{h}_{jw} \mathbf{h}_{jw}^\dagger \mathbf{v}}{\mathbf{v}^\dagger \left(\tilde{\mathbf{h}} \tilde{\mathbf{h}}^\dagger + \sigma \|\mathbf{c}\|^2 \mathbf{I} \right) \mathbf{v}}, \end{aligned}$$

which is the Rayleigh quotient maximization problem [28]. Note that we neglect the constants in (48) since we are interested only in \mathbf{v} . Thus, similarly to the proof of Theorem 1, setting $\mathbf{W} = \mathbf{h}_{jw} \mathbf{h}_{jw}^\dagger$ and, $\mathbf{B} = \tilde{\mathbf{h}} \tilde{\mathbf{h}}^\dagger + \sigma \|\mathbf{c}\|^2 \mathbf{I}$, where $\tilde{\mathbf{h}} = \mathbf{H}_{jb}^\dagger \mathbf{c}$, the optimal \mathbf{v} is the eigenvector which corresponds to the highest eigenvalue of the matrix $\mathbf{B}^{-1} \mathbf{W}$. ■

Fixed \mathbf{v} :

For any fixed \mathbf{v} we have the following theorem.

Theorem 4: For a fixed \mathbf{v} , the strategy for Bob, i.e., the solution to (48) when the CSI of Willie is known, is the linear filter \mathbf{c}^ which is the eigenvector that corresponds to the highest eigenvalue of the matrix $\mathbf{B}^{-1} \mathbf{W}$, where $\mathbf{W} = \mathbf{h}_{ab} \mathbf{h}_{ab}^\dagger$ and $\mathbf{B} = \mathbf{H}_{jb} \mathbf{v} \mathbf{v}^\dagger \mathbf{H}_{jb}^\dagger + \sigma \mathbf{I}$.*

Proof: Assuming \mathbf{v} is fixed the optimization problem in (48) is reduced to,

$$\begin{aligned} & \max_{\mathbf{c}} \frac{|\mathbf{c}^\dagger \mathbf{h}_{ab}|^2}{\mathbf{c}^\dagger \left(\mathbf{H}_{jb} \mathbf{v} \mathbf{v}^\dagger \mathbf{H}_{jb}^\dagger + \sigma \mathbf{I} \right) \mathbf{c}} \\ &= \max_{\mathbf{c}} \frac{\mathbf{c}^\dagger \mathbf{h}_{ab} \mathbf{h}_{ab}^\dagger \mathbf{c}}{\mathbf{c}^\dagger \left(\mathbf{H}_{jb} \mathbf{v} \mathbf{v}^\dagger \mathbf{H}_{jb}^\dagger + \sigma \mathbf{I} \right) \mathbf{c}}, \end{aligned}$$

which is again the Rayleigh quotient maximization problem [28]. Thus, setting $\mathbf{W} = \mathbf{h}_{ab} \mathbf{h}_{ab}^\dagger$ and $\mathbf{B} = \mathbf{H}_{jb} \mathbf{v} \mathbf{v}^\dagger \mathbf{H}_{jb}^\dagger + \sigma \mathbf{I}$, the optimal \mathbf{c} is the eigenvector which corresponds to the highest eigenvalue of the matrix $\mathbf{B}^{-1} \mathbf{W}$. ■

As a consequence of Theorems 3 and 4 we now present the strategies in the following corollaries. The corollaries provide

the specific linear filter \mathbf{c} and the direction vector \mathbf{v} of Bob and the jammer to be set in (48).

Corollary 3: When Bob employs a MRC for the channel between him and Alice while ignoring the AN, i.e., $\mathbf{c} = \mathbf{h}_{ab}$, the solution to (48) is $\mathbf{v}^* = \frac{\mathbf{v}}{\|\mathbf{v}\|}$, where \mathbf{v} is the eigenvector which corresponds to the highest eigenvalue of the matrix $\mathbf{B}^{-1}\mathbf{W}$, where $\mathbf{W} = \mathbf{h}_{jw}\mathbf{h}_{jw}^\dagger$, $\mathbf{B} = \tilde{\mathbf{h}}\tilde{\mathbf{h}}^\dagger + \sigma\|\mathbf{c}\|^2\mathbf{I}$ and, $\tilde{\mathbf{h}} = \mathbf{H}_{jb}^\dagger\mathbf{c}$.

Corollary 4: When the jammer directs his AN towards Willie, i.e., $\mathbf{v} = \frac{\mathbf{h}_{jw}}{\|\mathbf{h}_{jw}\|}$, the solution to (48) is the linear filter \mathbf{c}^* which is the eigenvector that corresponds to the highest eigenvalue of the matrix $\mathbf{B}^{-1}\mathbf{W}$, where $\mathbf{W} = \mathbf{h}_{ab}\mathbf{h}_{ab}^\dagger$ and $\mathbf{B} = \mathbf{H}_{jb}\mathbf{v}\mathbf{v}^\dagger\mathbf{H}_{jb}^\dagger + \sigma\mathbf{I}$.

As mentioned above, the cancelation of the AN at Bob's receiver depends on the number of antennas that Bob and the jammer have. If Bob has more antennas than the jammer, i.e., $M > N$, then he can find a null-space for the columns of \mathbf{H}_{jb} such that the linear filter \mathbf{c} will satisfy $\mathbf{c}^T\mathbf{H}_{jb} = 0$. In other words, Bob projects the received vector $\mathbf{y}_b[i]$ at the i th channel use onto a subspace spanned by the null-space of \mathbf{H}_{jb} . That is, we have

$$\begin{aligned}\mathbf{c}^T\mathbf{y}_b[i] &= x[i]\mathbf{c}^T\mathbf{h}_{ab} + \mathbf{c}^T\mathbf{H}_{jb}\mathbf{v}[i] + \mathbf{c}^T\mathbf{n}_b[i] \\ &= x[i]\mathbf{c}^T\mathbf{h}_{ab} + \mathbf{c}^T\mathbf{n}_b[i].\end{aligned}$$

The above implies that the vector $\tilde{\mathbf{h}}$ in Theorem 3 is equal to zero. Note that Bob may choose \mathbf{c} to improve the SNR while still restricting \mathbf{c} to being in null-space of \mathbf{H}_{jb} . Consequently, we have the following corollary.

Corollary 5: When Bob cancels the AN and chooses \mathbf{c} such that,

$$\mathbf{c} = \arg \max_{\mathbf{c} \in \text{nullspace}(\mathbf{H}_{jb})} \frac{|\mathbf{c}^\dagger\mathbf{h}_{ab}|^2}{\sigma_b^2\|\mathbf{c}\|^2}. \quad (50)$$

The optimal strategy for the jammer is the following power allocation,

$$\Sigma = \mathbf{v}^*\mathbf{v}^{*\dagger}, \quad (51)$$

where,

$$\mathbf{v}^* = \frac{\mathbf{h}_{jw}}{\|\mathbf{h}_{jw}\|}. \quad (52)$$

Proof: Since Bob cancels the received AN from the jammer, following the notations of Theorem 3, we have $\mathbf{B}^{-1} = \mathbf{B} = \sigma\|\mathbf{c}\|^2\mathbf{I}$ and $\mathbf{W} = \mathbf{h}_{jw}\mathbf{h}_{jw}^\dagger$. Since the matrix $\mathbf{B}\mathbf{W}$ is of rank 1, it has only a single non-zero eigenvalue which \mathbf{h}_{jw} is its corresponding eigenvector. ■

Corollary 5 essentially shows that when Bob can cancel the AN, the jammer can transmit his AN towards Willie without degrading the covert rate between Alice and Bob.

On the other hand, if the jammer has more antennas than Bob, i.e., $N > M$, then, the jammer may construct his covariance matrix such that the received AN power at Bob will fall first in the null-space of the columns of \mathbf{H}_{jb}^T and then as close as possible towards the direction of Willie. This strategy is different than the optimal one obtained in Theorem 1 since here the direction is always orthogonal to Bob. Thus, setting \mathbf{v} such that it satisfies $\mathbf{H}_{jb}\mathbf{v} = 0$, the jammer is able to

cancel the AN at Bob. Similarly, \mathbf{v} can be optimized further while restricted to be in the null-space. Accordingly we have the following corollary.

Corollary 6: When the jammer cancels the AN and chooses \mathbf{v} such that,

$$\mathbf{v} = \arg \max_{\mathbf{v} \in \text{nullspace}(\mathbf{H}_{jb}^T)} \left(\mathbf{h}_{jw}^\dagger\mathbf{v}\mathbf{v}^\dagger\mathbf{h}_{jw} \right). \quad (53)$$

The optimal strategy for Bob is to perform a MRC for the channel between Alice and himself. That is,

$$\mathbf{c}^* = \mathbf{h}_{ab}. \quad (54)$$

Proof: Since the jammer cancels the AN in advance, following the notations of Theorem 4, we have $\mathbf{B}^{-1} = \mathbf{B} = \sigma\mathbf{I}$ and $\mathbf{W} = \mathbf{h}_{ab}\mathbf{h}_{ab}^\dagger$. Since the matrix $\mathbf{B}\mathbf{W}$ is of rank 1, it has only a single non-zero eigenvalue which \mathbf{h}_{ab} is its corresponding eigenvector (this solution is also known as the matched filter [31]). ■

Corollary 6 essentially states that once the jammer is able to cancel his AN in advance at Bob's receiver, Bob only performs maximal ratio combining to maximize his received SNR.

Lastly, we have the following corollary which states that, as σ grows (the low SNR regime), corollaries 3 and 4 are asymptotically optimal.

Corollary 7: For $\sigma \rightarrow \infty$, the strategies suggested in corollaries 3 and 4 are asymptotically optimal in the sense of approaching the global optimum of (48).

Proof: At the limit of $\sigma \rightarrow \infty$, the effect of the AN noise term in the denominator of the SNR expression in (48) becomes negligible thus (48) is reduced to

$$\max_{\mathbf{v}, \mathbf{c}} \frac{\left(\mathbf{h}_{jw}^\dagger\mathbf{v}\mathbf{v}^\dagger\mathbf{h}_{jw} \right) |\mathbf{c}^\dagger\mathbf{h}_{ab}|^2}{\sigma\|\mathbf{c}\|^2}, \quad (55)$$

which is maximized when Bob employ an MRC and the AN noise transmission is in the direction of Willie. ■

Remark 2 (The case $M = N$): In case the jammer and Bob have the same number of antennas, the sub-optimal schemes suggested herein are not applicable due to the fact that \mathbf{H}_{jb} and \mathbf{H}_{jb}^\dagger are square matrices and thus a null-space does not exist for them. Therefore, other types of schemes must be performed. We note that in any case, one can use either Corollary 3 or Corollary 4 as a possible scheme.

Simulation results which compare the normalized average SNR value, as a function of σ , are depicted in Figure 4 for the cases of $M < N$, $M = N$ and $M > N$. The figures compare the optimal solution for the global optimization problem (48) (solved using Mathematica 12) with the suggested sub-optimal schemes given in Corollaries 3,4,5 and 6.

We note that the strategy of canceling the AN given in Corollaries 5 and 6 are depicted in Figures 4(a) and 4(c) for the cases of $M < N$ and $M > N$, respectively. When $M = N$ neither Bob nor the jammer can cancel the AN as the matrices \mathbf{H}_{jb} and \mathbf{H}_{jb}^T are square. Thus, only corollaries 3 and 4 are depicted in Figure 4(b).

One can observe several interesting insights from the simulation results. The first is that when $M < N$, choosing the

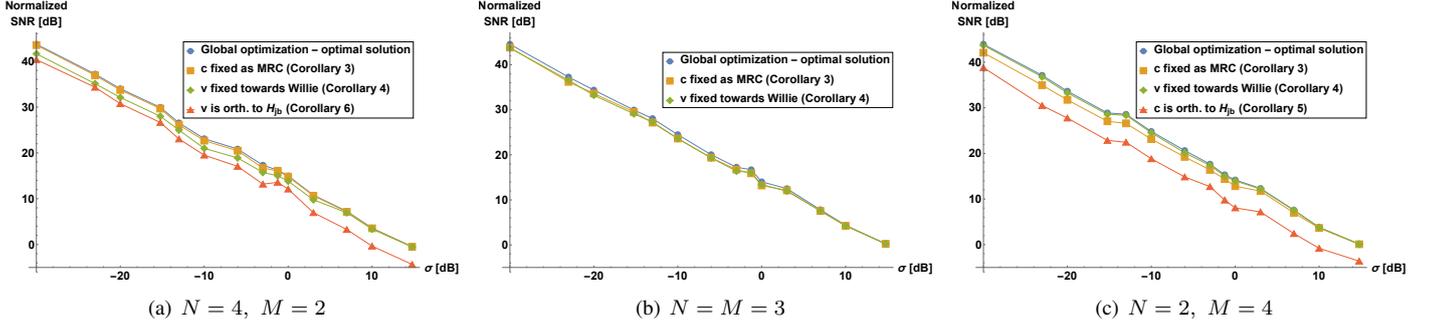


Fig. 4. Simulation results which compare the optimal solution for the global optimization in Equation (48) with the suggested sub-optimal schemes in Corollaries 3,4,5 and 6 as a function of σ . The three graphs are for a different constellations of the number of antennas. The SNR values, under ϵ -covertness, are normalized so the leading constant coefficient is equal to 1.

strategy that fixes \mathbf{c} to be the MRC and optimizing only on \mathbf{v} (Corollary 3) performs very well compared to the optimal solution of 48. On the other hand, when $M > N$, it is better to choose the strategy which fixes \mathbf{v} to be in the direction of Willie and optimize only on \mathbf{c} (Corollary 4). This can be explained by the number of DoF which are left for the optimization. That is, for example, when $M < N$ and we fix \mathbf{c} we are left with N variables in the optimization problem which can provide a more accurate search for the optimum. A piece of additional evidence for that is the fact that the two curves unite when $M = N$.

The second observation is that as σ grows (low SNR scenario), consistent to Corollary 7, the sub-optimal schemes, with the exception of the AN cancelation scheme, coincide with the optimal SNR making them asymptotically optimal schemes.

Finally, one can observe that the strategy of canceling the AN is always worse than MRC or directing the noise towards Willie. The reason lies in the fact that once the AN is canceled, the inner products in the numerator are restricted to be only in the specific null-space of Bob (as can be seen in equations (50) and (53)). This evidently, restricts much the SNR value as the figures depict. Moreover, even when $\sigma \rightarrow \infty$ this restriction remains and thus there is no convergence to the optimal SNR.

B. Unknown CSI of Willie

When the CSI of Willie is unknown, similar to the $M = 1$ case, Alice transmits with the power given in Lemma 2 in order to guarantee a covert transmission and the jammer transmits with all of his available power equally divided to d directions. Accordingly, we have the following global optimization problem for Bob and the jammer,

$$\max_{\mathbf{V}, d, \mathbf{c}} \frac{\frac{\epsilon P_{max}}{12e \ln \frac{6}{\epsilon}} \left(\frac{\epsilon}{\delta}\right)^{\frac{1}{d}} |\mathbf{c}^\dagger \mathbf{h}_{ab}|^2}{\mathbf{c}^\dagger \left(p_j \mathbf{H}_{jb} \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{H}_{jb}^\dagger + \sigma_b^2 \mathbf{I} \right) \mathbf{c}} \quad (56)$$

$$s.t. \quad \xi_l = \frac{1}{d}, \quad l = 1, \dots, d.$$

Again, to avoid the complexity in (56), we restrict the attention to several sub-optimal solutions. Specifically, in case

Bob has $M > N$ antennas, he can cancel the AN by projecting it onto the null-space of \mathbf{H}_{jb} and choosing \mathbf{c} which is closest to \mathbf{h}_{ab} . The jammer's strategy for this case is given in the following theorem.

Theorem 5: Assume $M > N$ and set:

$$\mathbf{c}^* = \arg \max_{\mathbf{c} \in \text{nullspace}(\mathbf{H}_{jb})} \frac{|\mathbf{c}^\dagger \mathbf{h}_{ab}|^2}{\sigma_b^2 \|\mathbf{c}\|^2}. \quad (57)$$

Then the optimal strategy for the jammer when the jammer has no CSI of Willie and he transmits with all of his available power, p_j , equally divided to all directions, is the following power allocation,

$$\Sigma = \frac{1}{N} \mathbf{I}_{N \times N}. \quad (58)$$

Proof: Once the interference term in the denominator of (56) is gone, the jammer should maximize the number of direction d which is N regardless of the specific directions. ■

The above theorem implies that once Bob can cancel the noise, the jammer should try to impair Willie's detection as much as possible. This is different from the result in Theorem 2 since now Bob takes an active part in the reception which enables the jammer to concentrate on Willie and to transmit in every direction (as Willie's channel is unknown).

In case the jammer has more antennas than Bob ($M < N$), a naive solution would be to cancel the AN by setting the matrix \mathbf{V} to span the null-space of \mathbf{H}_{jb}^T . In this case the rank of \mathbf{V} is $N - M$, i.e., $d = N - M$, which affects the numerator in (56) directly and especially when N is on the same order as M . However, according to Corollary 6, we know that in case the jammer can completely remove the AN interference, Bob's optimal linear filter is the MRC \mathbf{c}^* as given in the corollary. In fact, since Bob uses a linear filter cancelling his AN should require only one DoF. We thus suggest a scheme for which the jammer relies upon this choice of Bob for the linear filter \mathbf{c}^* and can null the equivalent one dimensional channel and benefiting from the remaining DoF he has. This strategy is given in the following theorem.

Theorem 6: For the fixed linear filter $\mathbf{c}^ = \mathbf{h}_{ab}$, the optimal strategy for the jammer when the CSI of Willie is unknown*

and the jammer transmits AN with covariance matrix Σ such that his total transmission power, p_j , is equally divided into d arbitrary directions, is the following power allocation,

$$\Sigma = \mathbf{V}^* \mathbf{X}^* \mathbf{V}^{*\dagger}, \quad (59)$$

where in case $\sigma = \frac{\sigma_b^2}{p_j} < \frac{\|\tilde{\mathbf{h}}\|^2}{N \left(\frac{\epsilon}{\delta} - \frac{1}{N(N-1)} - 1 \right) \|\mathbf{c}^*\|^2}$,

$$\mathbf{V}^* = \tilde{\mathbf{Q}} \quad \text{and} \quad \mathbf{X}^* = \frac{1}{N-1} I_{(N-1) \times (N-1)}. \quad (60)$$

Otherwise,

$$\mathbf{V}^* = \left(\tilde{\mathbf{Q}}, \frac{\tilde{\mathbf{h}}}{\|\tilde{\mathbf{h}}\|} \right) \quad \text{and} \quad \mathbf{X}^* = \frac{1}{N} I_{N \times N}, \quad (61)$$

where $\tilde{\mathbf{h}} = \mathbf{H}_{jb}^\dagger \mathbf{c}^*$ and $\tilde{\mathbf{Q}}$ is a matrix whose columns span the null-space of $\tilde{\mathbf{h}}$.

Proof: Considering that $\mathbf{c}^* = \mathbf{h}_{ab}$ is fixed and using notations from the proof of Theorem 2, the maximization problem (56) can be reduced to,

$$\begin{aligned} & \max_{\mathbf{V}, d} \frac{(\epsilon')^{\frac{1}{d}}}{\mathbf{c}^{*\dagger} \mathbf{H}_{jb} \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \mathbf{H}_{jb}^\dagger \mathbf{c}^* + \sigma \|\mathbf{c}^*\|^2} \\ & = \max_{\mathbf{V}, d} \frac{(\epsilon')^{\frac{1}{d}}}{\tilde{\mathbf{h}}^\dagger \mathbf{V} \mathbf{X} \mathbf{V}^\dagger \tilde{\mathbf{h}} + \sigma \|\mathbf{c}^*\|^2}, \end{aligned}$$

where $\tilde{\mathbf{h}} = \mathbf{H}_{jb}^\dagger \mathbf{c}^*$. The above can be simplified further to

$$\max_{d, \mathbf{b}} \frac{(\epsilon')^{\frac{1}{d}}}{\sum_{l=1}^d \frac{1}{d} b_l^2 + \sigma \|\mathbf{c}^*\|^2}$$

where $\mathbf{b} = \mathbf{V}^\dagger \tilde{\mathbf{h}}$. We result with a similar to the expression in the proof of Theorem 2 and thus the rest follows similarly. ■

V. CONCLUSION

In this work, we considered the problem of covert communication in the presence of a jammer who is equipped with multiple antennas. The jammer assists Alice and Bob by transmitting AN that creates uncertainty at Willie's detector and enables Alice and Bob to communicate with a positive rate. The transmission strategies which affect this rate are examined and analyzed in the form of optimization problems for the cases of full CSI and partial CSI at the jammer. The use of multiple antennas provides a multiplicative gain to the covertness achieving transmission power with respect to the case of a single antenna jammer. Specifically, under certain assumptions, this gain is an increasing function of the number of antennas at the jammer.

When Bob is equipped with a single antenna the optimal strategy for a jammer with full CSI is beamforming the AN with all available power to a specific direction, which is close to the direction of Willie while being orthogonal to Bob as much as possible. On the other hand, in the case of partial CSI, when the jammer does not know the channel toward Willie, the optimal AN transmission strategy is to transmit isotropically to all directions or to the null-space of Bob, where the choice

depends on certain channel conditions. Finally, when Bob is equipped with multiple antennas and is assumed to employ a linear receiver, several sub-optimal schemes are suggested along with numerical results. These results show that the strategy of letting the jammer transmit his AN towards Willie while letting Bob employ MRC is an asymptotically optimal solution for the global optimization problem which includes both Willie's and Bob's strategies.

ACKNOWLEDGMENT

The authors would like to thank Alejandro Cohen for discussions that helped prompt this work.

REFERENCES

- [1] 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.
- [2] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: Hiding messages in noise," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 2945–2949.
- [3] 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.
- [4] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2334–2354, 2016.
- [5] K. S. K. Arumugam and M. R. Bloch, "Keyless covert communication over multiple-access channels," in *Information Theory (ISIT), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 2229–2233.
- [6] M. Tahmasbi, A. Savard, and M. R. Bloch, "Covert capacity of non-coherent rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 66, no. 4, pp. 1979–2005, 2019.
- [7] P. H. Che, M. Bakshi, C. Chan, and S. Jaggi, "Reliable deniable communication with channel uncertainty," in *Information Theory Workshop (ITW), 2014 IEEE*. IEEE, 2014, pp. 30–34.
- [8] S. Lee, R. J. Baxley, M. A. Weitnauer, and B. Walkenhorst, "Achieving undetectable communication," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 7, pp. 1195–1205, 2015.
- [9] 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.
- [10] A. Abdelaziz and C. E. Koksal, "Fundamental limits of covert communication over mimo awgn channel," in *2017 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2017, pp. 1–9.
- [11] M. Forouzes, P. Azmi, N. Mokari, and D. Goeckel, "Covert communication using null space and 3d beamforming: Uncertainty of willie's location information," *IEEE Transactions on Vehicular Technology*, 2020.
- [12] A. Khisti and G. Wornell, "Secure transmission with multiple antennas: The MISOME wiretap channel," *arXiv preprint arXiv:0708.4219*, 2007.
- [13] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE transactions on wireless communications*, vol. 7, no. 6, pp. 2180–2189, 2008.
- [14] A. Mukherjee and A. L. Swindlehurst, "Fixed-rate power allocation strategies for enhanced secrecy in MIMO wiretap channels," in *2009 IEEE 10th Workshop on Signal Processing Advances in Wireless Communications*. IEEE, 2009, pp. 344–348.
- [15] A. L. Swindlehurst, "Fixed SINR solutions for the MIMO wiretap channel," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2009, pp. 2437–2440.
- [16] A. Khisti and G. Wornell, "Secure transmission with multiple antennas II: The MIMOME wiretap channel," *arXiv preprint arXiv:1006.5879*, 2010.
- [17] S.-H. Tsai and H. V. Poor, "Power allocation for artificial-noise secure MIMO precoding systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 13, pp. 3479–3493, 2014.

- [18] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE transactions on signal processing*, vol. 58, no. 3, pp. 1875–1888, 2010.
- [19] J. Huang and A. L. Swindlehurst, "Cooperative jamming for secure communications in MIMO relay networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4871–4884, 2011.
- [20] H.-M. Wang, Q. Yin, and X.-G. Xia, "Distributed beamforming for physical-layer security of two-way relay networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3532–3545, 2012.
- [21] S. Luo, J. Li, and A. P. Petropulu, "Uncoordinated cooperative jamming for secret communications," *IEEE transactions on information forensics and security*, vol. 8, no. 7, pp. 1081–1090, 2013.
- [22] S. A. A. Fakoorian and A. L. Swindlehurst, "Solutions for the MIMO Gaussian wiretap channel with a cooperative jammer," *IEEE Transactions on signal Processing*, vol. 59, no. 10, pp. 5013–5022, 2011.
- [23] X. Chen, J. Chen, H. Zhang, Y. Zhang, and C. Yuen, "On secrecy performance of multiantenna-jammer-aided secure communications with imperfect csi," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8014–8024, 2015.
- [24] K. Shahzad, X. Zhou, S. Yan, J. Hu, F. Shu, and J. Li, "Achieving covert wireless communications using a full-duplex receiver," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8517–8530, 2018.
- [25] 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.
- [26] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
- [27] S. Shafiee and S. Ulukus, "Achievable rates in Gaussian MISO channels with secrecy constraints," in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*. IEEE, 2007, pp. 2466–2470.
- [28] A. M. Mathai and S. B. Provost, *Quadratic forms in random variables: theory and applications*. M. Dekker New York, 1992.
- [29] R. J. Pavur, "Quadratic forms involving the complex gaussian," Ph.D. dissertation, Texas Tech University, 1980.
- [30] L. C. Godara, "Application of antenna arrays to mobile communications. ii. beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.
- [31] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.