Achieving Covert Wireless Communications Using a Full-Duplex Receiver

Khurram Shahzad, Student Member, IEEE, Xiangyun Zhou, Senior Member, IEEE, Shihao Yan, Member, IEEE, Jinsong Hu, Student Member, IEEE, Feng Shu, Member, IEEE, and Jun Li, Senior Member, IEEE

Abstract—Covert communications hide the transmission of a message from a watchful adversary while ensuring a certain decoding performance at the receiver. In this paper, a wireless communication system under fading channels is considered where covertness is achieved by using a full-duplex receiver. More precisely, the receiver of covert information generates artificial noise with a varying power causing uncertainty at the adversary, Willie, regarding the statistics of the received signals. Given that Willie’s optimal detector is a threshold test on the received power, we derive a closed-form expression for the optimal detection performance of Willie averaged over the fading channel realizations. Furthermore, we provide guidelines for the optimal choice of artificial noise power range, and the optimal transmission probability of covert information to manage the detection errors at Willie. Our analysis shows that the transmission of artificial noise, although causing self-interference, provides the opportunity of achieving covertness but simultaneously need to be managed carefully. We also demonstrate that the prior transmission probability of 0.5 is not always the best choice for achieving the maximum possible covertness when the covert transmission probability and artificial noise power can be jointly optimized.

Index Terms—Physical layer security, covert wireless communications, low probability of detection, artificial noise, full-duplex.

I. INTRODUCTION

A. Background

The wireless air interface is open and accessible to both legitimate and illegitimate users. This creates reasonable concerns over the security and privacy of information transmitted over the air. The recent remarkable increase in the amount of information conveyed using the wireless medium has spurred an interest in both research and academic communities regarding the development of new mechanisms, enhancing the privacy and integrity of wirelessly transmitted data. In recent years, physical layer security [2], [3] has emerged as an alternative to traditional cryptographic ways of securing wireless information, where the mechanisms of key exchange and distribution impose varied challenges, especially in dynamic network environments. Physical layer security techniques exploit the uncertainties and lack of predictability of the wireless channel, minimizing the information obtained by an unauthorized eavesdropper. Under the varied circumstances where users communicate over the wireless medium, situations exist where not only the privacy and integrity of the information are important, but the users may also wish to avert any invigilation, hiding the very existence of their communication. Such situations, although, commonplace in military applications, are now also arising in non-military applications, relating to civil unrest and even monitoring of people’s daily activities. Thus, catering for such security concerns is a need of the moment and motivates the recent interest in covert communications [4], [5].

Covert communications intend to obscure the existence of any wireless transmission from a watchful adversary, referred to as Willie in recent literature of covert communications, while guaranteeing a certain decoding performance at the intended receiver. The low probability of detection (LPD) communications have drawn significant research attention and are materializing as a promising prospect for shielding the future wireless communication networks from unapproved probing and access. In this regard, the fundamental limits of
covert communications were established in [6], presenting a square root limit on the amount of information transmitted reliably and with low probability of detection over additive white Gaussian noise (AWGN) channels. This work has been further extended to binary symmetric channels (BSCs) [7], discrete memoryless channels (DMCs) [8] and multiple access channels (MACs) [9]. Although under the square root law, the average number of covert bits per channel use asymptotically reaches zero, a positive covert rate has been shown to be achievable in a number of cases. These include the situations of Willie’s uncertainty in the knowledge of noise power [10]–[13], Willie’s ignorance of transmission time [14], and presence of a continuously transmitting jammer in the environment [15]. The case when additional friendly nodes generating artificial noise are present in the environment, causing confusion at Willie regarding the received signal statistics, is presented in [16], while analysis of covert transmissions under finite blocklengths, imperfect channel state information and in one-way relay networks is presented in [17]–[20], respectively.

B. Our Approach and Contribution

In this work, we make use of a full-duplex (FD) receiver to achieve covert communication. Specifically, the FD receiver generates artificial noise (AN) with a randomized transmit power, causing a deliberate confusion and affecting the decisions at Willie regarding the presence of any covert transmissions. Although the use of AN and jamming signals for enhancing physical layer security has been widely advocated in the literature [21]–[25], and references therein, to the best of our knowledge, it has not been studied before in the context of covert communications. The use of a FD receiver generating AN provides a cover for the covert transmission, and offers a multitude of benefits as compared to the use of a separate, independent jammer. Being equipped with an FD receiver, we can exercise a better control over the power used for transmitting AN, hence a better management of system resources to achieve the said purpose of security is achievable. Furthermore, while Willie will face a strong interference, the self-interference at the FD receiver can be greatly suppressed by the well-developed self-interference cancellation techniques [26], [27], providing a significant advantage to the covert communication pair.

In our considered scenario, covert transmissions can occur in multiple blocks of time, and Willie is performing the detection on a block-to-block basis. In this case, the a priori transmission probability becomes an interesting and important parameter affecting Willie’s detection performance as well as the overall throughput of covert communications. A general assumption in the literature regarding the a priori probability of covert transmission is that there is a 50% chance that transmission occurs in a block of interest. This assumption is understood as a good choice for covertness, since it renders Willie’s knowledge of Alice’s transmission uninformative and is equivalent to assuming that Willie has no prior knowledge on whether Alice transmits or not [6], [14]. We show in this work that an a priori probability of 0.5 is not always the best choice in our considered scenario; rather a joint adjustment of this probability with other system parameters can offer a better covert performance.

The main contributions of this work are summarized as follows:

- We show that the use of an FD receiver is an effective way of achieving covert communication over fading wireless channels, where the FD receiver is designed to transmit AN with varying power to cause confusion at Willie.
- Under the assumption of a radiometer (power-detector) at Willie, we analytically derive the optimal detection threshold of Willie’s radiometer and obtain its optimal detection performance in terms of the minimum detection error probability.
- For a given covert rate requirement, we provide the design guidelines on optimal choices for the range of AN transmit power at the FD receiver and the optimal a priori probability of covert transmission in order to maximize the expected detection error probability at Willie.
- Our analysis reveals that an a priori transmission probability of 0.5 is not always the best choice. Increasing this transmission probability beyond 0.5 gives more room to increase the AN transmit power for maintaining the same rate requirement. Thus whether to allow such a change in the transmission probability can be the difference between achieving strong covertness and achieving almost no covertness at all.

C. Related Works

Our work is closely related to [15], where a jammer is assumed to be present in the environment. Although the jammer does not closely coordinate with the covert transmitter, it is allowed to transmit continuously and the received power at Willie due to the jammer changes randomly from slot to slot. In this case, the covert communication pair has no control over the jammer’s transmit power level. In contrast, although we also consider randomizing the AN power in each slot, our focus is on optimizing the AN transmit power range, since this choice affects the information decoding at the intended receiver through self-interference. This important optimization is made possible because the AN is transmitted by the FD receiver, and hence, controllable by the covert communication pair. Moreover, instead of satisfying a given covertness constraint, we present our analysis on the choice of AN transmit power range to achieve the maximum possible covertness while meeting a given rate requirement.

Furthermore, performance of communication systems with randomly distributed interferers has been studied extensively in the literature [28]–[30]. More recently, a study on covert communications in the presence of a Poisson distributed field of interferers has been presented in [31], where leveraging the total received interference, the effect of density and transmit powers of the interferers on the covert throughput is analyzed. Our work differs from [31] in that we consider AN generated by the FD receiver, hence allowing design and optimization of AN power with other design parameters. Thus, while the authors in [31] study the covert performance for a given
interference scenario, we take a design approach and provide guidelines on the optimal choice of parameters for achieving
covertness.

D. Organization

The rest of this paper is organized as follows: Section II
details our communication scenario, proposed scheme and the
assumptions used in this paper. Section III explains Willie’s
approach for detection of any covert transmissions, deriving
the conditions for possibility of any covert communications
and the optimal settings at Willie. Using the knowledge
of Willie’s approach, Section IV discusses the parameters that
affect the achievable performance of the proposed communica-
tion scheme while Section V addresses the optimal design of
all the system parameters that we have in our control to achieve
the best possible performance in covertness. Section VI pro-
vides numerical results validating our analysis and provides
further insights on the impact of AN and a priori probabilities,
and finally, Section VII draws some concluding remarks.

II. SYSTEM MODEL

A. Communication Scenario

A covert wireless communication system is considered,
as shown in Fig. 1, where a transmitter (Alice) possesses
sensitive information that needs to be sent to an information
receiver (Bob). Bob operates in full-duplex mode, and Alice
seeks to transmit covertly to Bob with the aid of artificial
noise (AN) generated by Bob. Under these circumstances,
an adversary (Willie) silently listens to the communication
environment and tries to detect any covert transmission from
Alice to Bob. We use the subscripts \(a\), \(b\) and \(w\) to represent
the terms associated with Alice, Bob and Willie, respectively.
It is assumed that Willie has complete knowledge of the carrier
frequency of any transmissions, associated antenna gains and
the distances between all the nodes.

A communication slot is defined as a block of time over
which the transmission of a message from Alice to Bob
is completed. Each slot contains \(n\) symbol periods and we
assume that \(n\) is large enough, i.e., \(n \to \infty\). The slot
boundaries are perfectly synchronized among all the users, and
we consider fading wireless channels where the channel coeffi-
cients remain constant in one slot, changing independently
from one slot to another, i.e., quasi-static Rayleigh fading
channels are considered. The channel between any two users \(i\)
and \(j\) is represented by \(h_{ij}\), where the channel gain is assumed
to encompass the combined antenna gain of transmit/receive
antennas and the distance between the two users as well. The
mean of \(|h_{ij}|^2\) over different communication slots is
denoted by \(1/\lambda_{ij}\), where subscript \(ij\) can be \(ab\), \(aw\), \(bw\) or \(bb\).
Hence, the Alice-Bob, Alice-Willie and Bob-Willie channels
are denoted by \(h_{ab}\), \(h_{aw}\), and \(h_{bw}\), respectively, while the self-
interference channel of Bob is denoted by \(h_{bb}\). We note here
that \(h_{bb}\) is the loop interference channel at Bob and is modelled
via the Rayleigh fading distribution under the assumption that
any line-of-sight component is efficiently reduced by antenna
isolation and the major effect comes from scattering [32].
Regarding the channel knowledge, it is assumed that Bob
knows \(h_{ab}\), while Willie possesses complete knowledge of \(h_{aw}
\) and \(h_{bw}\). Hence, the availability of knowledge regarding \(h_{aw}\)
and \(h_{bw}\) at Willie represents the worst case scenario from the
perspective of covert communication design.

The complex additive Gaussian noise at Bob and Willie’s
receiver is denoted by \(n_b \sim CN(0, \sigma^2_b)\) and \(n_w \sim CN(0, \sigma^2_w)\),
respectively. Each of Alice and Willie is equipped with a single
antenna, while apart from a receiving antenna, Bob also has
an additional antenna for the transmission of AN. Due to its
full-duplex nature, Bob suffers from self-interference, causing
a degradation in the signal-to-noise ratio (SNR) of the message
signal received from Alice [33], [34]. Since the generated
AN signal is known to Bob, the self-interfering signal, acting
as noise for Bob’s receiver, can be rebuilt and eliminated
up to a certain extent by using efficient techniques of self-
interference cancellation [26], [35]. However, owing to compu-
tational limitations and practical non-idealities, we assume that
perfect cancellation of self-interference is not achievable [36].
The self-interference cancellation coefficient is denoted by \(\phi\),
where \(0 < \phi \leq 1\) corresponds to different cancellation levels
of interfering AN signals. The residual interfering link is also
modelled as Rayleigh fading channel, following a common
assumption in the literature [32], [37].

The transmit power of Alice and Bob is denoted by \(P_a\)
and \(P_b\), respectively. When Alice transmits, the signal received at
Bob for each symbol period is given by

\[
y_b(i) = \sqrt{P_a}h_{ab}x_a(i) + \sqrt{\phi P_b}h_{bb}x_b(i) + n_b(i),
\]

where \(i = 1, \ldots, n\) represents the symbol index. Here, \(x_a\) and
\(x_b\) represent the signals transmitted by Alice and Bob, respect-
ively, satisfying \(E[x_a(i)x_a^*(i)] = 1\) and \(E[x_b(i)x_b^*(i)] = 1\).
We also consider an average power constraint on Bob’s
transmit power, denoted by \(P_{aw}\). We follow the common
assumption that a secret of sufficient length is shared between
Alice and Bob [6], [15], which while unknown to Willie,
enables Bob to know Alice’s strategy. Employing random
coding arguments, Alice generates codewords of length \(n\),
by independently drawing symbols from a zero-mean complex
Gaussian distribution with unit variance. Here, each codebook
is known to Alice and Bob and is used only once. When Alice
transmits in a slot, she selects the codeword corresponding to
her message and transmits the resulting sequence.
of transmission by Alice in each slot. Due to the shared secret between Alice and Bob, this constitutes a form of a time-hopping system. Here, Willie looks to make a decision regarding Alice’s transmission in each slot as he is interested in knowing for each individual slot that whether Alice transmitted or not. This means that Willie is not only interested in “whether” a transmission happens but also “when” it happens, i.e., in which slot. Note that if Willie is only interested in whether transmission happens but not in when it happens, he needs to make only a single decision after observing all slots. Such a scenario has been considered in [14], where the slot selection is kept secret from Willie and he looks to make a single decision regarding Alice’s transmission over all possible slots.

The knowledge of “when” a transmission happens not only improves upon Willie’s effectiveness in detecting covert transmission, but also gives him the capability of taking an action at the required time rather than waiting for the end of observation interval before intervening (although the corresponding action by Willie is beyond the scope of this work). Consider a scenario where Alice and Bob agree upon a certain “pattern” in choosing the slots over which covert messages are sent. Once Willie is able to detect the pattern based on his per slot decisions, it becomes easier for him to efficiently predict the slots over which future transmissions will happen.\footnote{Although the proposed scheme will help Willie in being able to predict any such pattern, this prediction is beyond the scope of this work and is thus not considered here.}

\subsection*{D. Priors and Performance Metrics}

Willie faces a decision as to whether or not Alice sent any covert information to Bob. As a result, Willie faces a binary hypothesis testing problem. The null hypothesis, $H_0$, states that Alice did not transmit while the alternative hypothesis, $H_1$, states that Alice did transmit, sending covert information to Bob. We define the probability of false alarm (or Type I error) as the probability that Willie makes a decision in favor of $H_1$, while $H_0$ is true, denoted by $P_{FA}$. Similarly, the probability of missed detection (or Type II error) is defined as the probability of Willie making a decision in favor of $H_0$, while $H_1$ is true, and is denoted by $P_{MD}$. We denote by $\pi_0$ and $\pi_1$ the a priori probabilities of hypothesis $H_0$ and $H_1$, respectively. The detection error probability at Willie is given by

$$P_E = \pi_0P_{FA} + \pi_1P_{MD},$$

which serves as a measure of covertness. In the recent literature, the assumption of both hypotheses being presented with an equal a priori probability has been widely adopted [10], [14]. The knowledge of a priori probabilities helps Willie improve his detection performance [6, Fact 4], as his assumption of $\pi_0 = \pi_1 = \frac{1}{2}$ implies that his observations are of little use to him and his decisions are akin to a random guess about the transmission state of Alice.

In this work, we instead consider general, i.e., not necessarily equal priors, and assume that Willie happens to know them. Since $P_E \leq \min(\pi_0, \pi_1)$, achieving covert communication guarantees that $P_E$ is in close proximity of $\min(\pi_0, \pi_1)$.\footnote{Although the proposed scheme will help Willie in being able to predict any such pattern, this prediction is beyond the scope of this work and is thus not considered here.}
III. DETECTION SCHEME AT WILLIE

The signals received at Willie under the two possible hypotheses for each symbol period are given by

\[ y_w(i) = \begin{cases} \sqrt{P_a} h_{aw} x_a(i) + \sqrt{P_b} h_{bw} x_b(i) + n_w(i), & \text{Alice tx.} \\ \sqrt{P_b} h_{bw} x_b(i) + n_w(i), & \text{else.} \end{cases} \]

(4)

From the independent and identically distributed (i.i.d.) nature of Willie’s received vector, \( y_w \), each element of \( y_w \), i.e., \( y_w(i) \) has a distribution given by

\[ \begin{array}{ll} CN(0, |h_{aw}|^2 P_a + |h_{bw}|^2 P_b + \sigma_w^2), & \text{Alice tx.} \\ CN(0, |h_{bw}|^2 P_b + \sigma_w^2), & \text{else.} \end{array} \]

(5)

We note while that the distribution of \( P_b \), is known to Willie, its value in a given slot is not known. Based on his observation vector \( y_w = [y_w(1), \ldots, y_w(n)] \), Willie has to make a decision regarding Alice’s actions in each communication slot. We assume that Willie uses a radiometer as his detector [10], [15] due to its low complexity and ease of implementation. When Willie has the statistical knowledge of his observations, this assumption is justified and the optimality of radiometer can be proved along the same lines as the proof of Lemma 3 in [15] using Fisher-Neyman Factorization Theorem [39] and Likelihood Ratio Ordering concepts [40].

While adopting a radiometer, the total received power at Willie, \( \sum_{i=1}^{n} |y_w(i)|^2 \), is a sufficient statistic for Willie’s test. Since any one-to-one transformation of a sufficient statistic is also sufficient, the term \( 1/n \sum_{i=1}^{n} |y_w(i)|^2 \) is also a sufficient statistic. Thus Willie conducts a threshold test on the average power received in a slot, given by

\[ P_w \overset{D_1}{\underset{D_0}{\geq}} \gamma, \]

(6)

where \( P_w = 1/n \sum_{i=1}^{n} |y_w(i)|^2 \) is the average power received at Willie in a slot, \( D_0 \) and \( D_1 \) are defined as the events that Willie makes a decision in the favor of \( H_0 \) and \( H_1 \), respectively, and \( \gamma \) is Willie’s detector threshold, which can be optimized to minimize the detection error probability. The average power at Willie in a slot under hypothesis \( H_0 \) is given by

\[ P_w(H_0) = \lim_{n \to \infty} \left( |h_{bw}|^2 P_b + \sigma_w^2 \right) \frac{\chi^2_{2n}}{n} \]

\[ = |h_{bw}|^2 P_b + \sigma_w^2, \]

(7)

where \( \chi^2_{2n} \) represents a chi-squared random variable with \( 2n \) degrees of freedom and from the Strong Law of Large Numbers, we know that \( \chi^2_{2n}/n \to 1 \) almost surely. Similarly, the average power at Willie in a slot under hypothesis \( H_1 \) is

\[ P_w(H_1) = |h_{bw}|^2 P_b + |h_{aw}|^2 P_a + \sigma_w^2. \]

(8)

We first analyze the condition under which Willie has non-zero probability of making detection errors and based on that, we find the optimal setting for Willie’s detector threshold. It should be noted here that the analysis of Willie’s detection error probability presented in the following proposition is for given channel realizations as Willie possesses the full knowledge of his channel from Alice and Bob.

**Proposition 1:** Willie has a non-zero probability of making detection errors when:

\[ \frac{|h_{aw}|^2}{|h_{bw}|^2} \leq \frac{P_{\text{max}} - P_{\text{min}}}{P_a}. \]

(9)

When (9) holds, the optimal choice for Willie’s detector’s threshold is

\[ \gamma^* = \begin{cases} |h_{bw}|^2 P_{\text{min}} + |h_{aw}|^2 P_a + \sigma_w^2, & \text{if } \pi_1 \geq \pi_0 \\ |h_{bw}|^2 P_{\text{max}} + \sigma_w^2, & \text{otherwise}, \end{cases} \]

(10)

and the corresponding minimum detection error probability at Willie is given by

\[ P^*_E = \begin{cases} \pi_0 \left[ 1 - \frac{|h_{aw}|^2 P_a}{|h_{bw}|^2 (P_{\text{max}} - P_{\text{min}})} \right], & \text{if } \pi_1 \geq \pi_0 \\ \pi_1 \left[ 1 - \frac{|h_{bw}|^2 P_a}{|h_{aw}|^2 (P_{\text{max}} - P_{\text{min}})} \right], & \text{otherwise.} \end{cases} \]

(11)

Proof: The detailed proof is provided in Appendix. \( \blacksquare \)

Remark 1: From Proposition 1, when (9) does not hold, Willie will have zero probability of making a detection error by setting the threshold \( \gamma \) in the interval \( |h_{bw}|^2 P_{\text{max}} + \sigma_w^2 < \gamma \leq |h_{bw}|^2 P_{\text{min}} + |h_{aw}|^2 P_a + \sigma_w^2 \). We also note here that although Willie’s receiver noise variance, \( \sigma_w^2 \), is required for the calculation of the optimal threshold for Willie’s detector, its value does not affect the minimum detection error probability at Willie. This can be attributed to the fact that as \( n \to \infty \), there is no uncertainty at Willie regarding the noise statistics and hence it does not contribute to an increase or decrease in the detection error probability at Willie.

IV. PERFORMANCE OF COVERT COMMUNICATION

In this section, we present those system metrics which affect the performance of our proposed covert transmission scheme. We note that the square root law presented by Bash et al. [6] holds given Willie has perfect statistical knowledge of the test statistics. It has been shown in prior works [10], [12], [15] that uncertainties present (or intentionally introduced) in the test statistics under both the null and alternative hypotheses at Willie result in a positive rate. In this work, the randomness in Bob’s transmit power introduces the required uncertainty at Willie, and hence we are able to achieve a positive covert rate. Here, we first calculate the outage probability for the transmission from Alice to Bob, and then present a measure that helps in quantifying the performance of our presented covert scheme.

A. Transmission Outage Probability From Alice to Bob

The signal-to-interference-plus-noise ratio (SINR) at Bob, in case Alice transmits, is given by

\[ \text{SINR}_b = \frac{|h_{ab}|^2 P_a}{\phi |h_{bb}|^2 P_b + \sigma_b^2}. \]

(12)

We assume a pre-determined rate from Alice to Bob, and denote it by \( R_{ab} \). Due to the random nature of \( h_{ab}, h_{bb} \) and \( P_b \), a transmission outage from Alice to Bob occurs when \( C_{ab} < R_{ab} \), where \( C_{ab} \) is the channel capacity from Alice to Bob.
Lemma 1: The transmission outage probability from Alice to Bob is given by
\[ \delta_{ab} = 1 - \frac{\lambda_{ab} \exp(-\lambda_{ab} \mu \sigma_b^2)}{(\lambda_{ab} \phi \mu P_{\text{max}} P_{\text{min}}) \ln \left( \frac{\lambda_{bb} + \lambda_{ab} \phi \mu P_{\text{max}}}{\lambda_{bb} + \lambda_{ab} \phi \mu P_{\text{min}}} \right)}, \] (13)
where \( \mu \triangleq \frac{(2R_{ab} - 1)}{P_a}. \)

Proof: From the definition of transmission outage probability, we have
\[ \delta_{ab} = \mathbb{P} [C_{ab} < R_{ab}] \]
\[ = \mathbb{P} \left[ \left| h_{ab} \right|^2 P_a \phi |h_{ba}|^2 P_b + \sigma_b^2 < 2R_{ab} - 1 \right] \]
\[ = \int_{P_{\text{min}}}^{P_{\text{max}}} \int_{0}^{\infty} \int_{0}^{\sigma_b^2} f_{\left| h_{ab} \right|^2}(x) f_{\left| h_{ba} \right|^2}(y) f_{P_b}(z) \times dx \, dy \, dz \]
\[ = \int_{P_{\text{min}}}^{P_{\text{max}}} \int_{0}^{\infty} \left[ 1 - \exp \left( -\lambda_{ab} \phi \left| h_{ab} \right|^2 P_b - \sigma_b^2 \right) \right] \times f_{\left| h_{ba} \right|^2}(y) f_{P_b}(z) \, dy \, dz \]
\[ = \int_{P_{\text{min}}}^{P_{\text{max}}} \left[ 1 - \frac{\lambda_{ab} \exp \left( -\lambda_{ab} \phi \sigma_b^2 \right)}{\lambda_{bb} + \lambda_{ab} \phi \mu P_b} \right] f_{P_b}(z) \, dz \]
\[ = 1 - \frac{1}{P_{\text{max}} - P_{\text{min}}} \int_{P_{\text{min}}}^{P_{\text{max}}} \left[ \frac{\lambda_{ab} \exp \left( -\lambda_{ab} \phi \sigma_b^2 \right)}{\lambda_{bb} + \lambda_{ab} \phi \mu} \right] \, dz, \]
(14)
and using the solution from [41] for the general form of integral \( \int_{A+C} f \, dx = \frac{\lambda \nu}{2} \nu \) for the second term gives the desired result.

B. Expected Detection Error Probability at Willie

Since Alice and Bob are unaware of their instantaneous channel to Willie, we consider the expected value of detection error probability at Willie, \( \mathbb{E}_E \), over all possible realizations of \( h_{aw} \) and \( h_{bw} \) as the measure of coverness from the viewpoint of Alice and Bob, and this expected detection error probability at Willie is denoted by \( \mathbb{E}_E \).

**Lemma 2:** Under the optimal detection threshold setting, the expected detection error probability at Willie is given by
\[ \mathbb{E}_E = \left\{ \begin{array}{ll}
\pi_0 \left[ 1 + t \ln t - t^2 \right], & \text{if } \pi_1 \geq \pi_0 \\
\pi_1 \left[ 1 + t \ln t - t^2 \right], & \text{otherwise},
\end{array} \right. \]
(15)
where \( t \triangleq (\lambda_{bw} P_a - \lambda_{aw} (P_{\text{max}} - P_{\text{min}})). \)

Proof: For the case of \( \pi_1 \geq \pi_0 \), and under the condition of Willie making detection errors, given by \( |h_{bw}|^2 P_{\text{max}} + \sigma_{aw}^2 \geq |h_{bw}|^2 P_{\text{min}} + |h_{aw}|^2 P_a + \sigma_{aw}^2 \), we have
\[ \mathbb{E}_E = \pi_0 \left\{ \int_{0}^{\infty} \int_{0}^{\infty} \frac{|h_{bw}|^2 (P_{\text{max}} - P_{\text{min}})}{P_a} \mathbb{P}_E^* \cdot f_{|h_{bw}|^2}(x) f_{|h_{bw}|^2}(y) \, dx \, dy \right\}, \]
(16)
which, using the law of total expectation, can also be written as
\[ \mathbb{E}_E = \pi_0 \left\{ \mathbb{P} \left[ |h_{bw}|^2 \leq \frac{|h_{bw}|^2 (P_{\text{max}} - P_{\text{min}})}{P_a} \right] \right\}, \]
(17)
where
\[ \mathbb{P} \left[ |h_{bw}|^2 \leq \frac{|h_{bw}|^2 (P_{\text{max}} - P_{\text{min}})}{P_a} \right] \]
\[ = \int_{0}^{\infty} \int_{0}^{\infty} f_{|h_{bw}|^2}(x) f_{|h_{bw}|^2}(y) \, dx \, dy \]
\[ = \int_{0}^{\infty} \left[ 1 - \exp \left( -\lambda_{aw} \left( \frac{P_{\text{max}} - P_{\text{min}}}{P_a} \right) y \right) \right] \times \lambda_{bw} \exp \left( -\lambda_{bw} y \right) \, dy \]
\[ = \frac{\lambda_{bw} P_b}{\lambda_{bw} P_a + \lambda_{aw} (P_{\text{max}} - P_{\text{min}})}, \] (18)
and
\[ \mathbb{E} \left[ \mathbb{P}_E \left[ |h_{aw}|^2 \leq \frac{|h_{aw}|^2 (P_{\text{max}} - P_{\text{min}})}{P_a} \right] \right] \]
\[ = 1 - \frac{P_a}{P_{\text{max}} - P_{\text{min}}} \int_{0}^{\infty} \int_{0}^{\infty} \frac{|h_{bw}|^2 (P_{\text{max}} - P_{\text{min}})}{P_a} \, x \, dy \]
\[ = 1 - \frac{\lambda_{bw} P_a}{\lambda_{bw} (P_{\text{max}} - P_{\text{min}})} \ln \left( \frac{1 + \lambda_{aw} \left( \frac{P_{\text{max}} - P_{\text{min}}}{P_a} \right)}{\lambda_{bw} P_a} \right) \]
\[ - \frac{\lambda_{bw} P_a}{\lambda_{bw} (P_{\text{max}} - P_{\text{min}})}, \] (19)
and putting in these expressions into (17) gives the desired result.

The case for \( \pi_0 > \pi_1 \) follows along the same lines, hence concluding the proof.

**Remark 2:** We make a few observations regarding the effect of \( P_{\text{max}} \) and \( P_a \) on Willie’s detection performance. Firstly, as \( P_{\text{max}} \to \infty \), the probability of Willie making detection errors approaches \( \pi_0 \) or \( \pi_1 \), in respective cases, which represents the maximum of \( \mathbb{E}_E \). Secondly, if Alice’s transmit power \( P_a \to \infty \), then \( t \to 1 \) and \( \mathbb{E}_E \to 0 \). Thus for a given set of \( \{P_{\text{min}}, P_{\text{max}}\} \), Alice can be “loud” enough to be heard by Willie.

V. COVERT COMMUNICATION DESIGN

In majority of the recent literature in covert communications, the detection error probability is used to measure the level of coverness under the assumption of equal priors. However, in this work, we propose a different framework and instead of putting a constraint on the error probability to achieve a said coverness, we look to maximize it under the given system model. Hence, from Alice and Bob’s perspective,
the objective is to achieve the best possible covertness in transmission, while being subject to an average power constraint and satisfying a given effective covert rate requirement which we denote by $\tau$. In this section, we consider optimal choices for the parameters in our control to achieve the said purpose.

Although Alice’s transmit power, $P_a$, is assumed to be fixed in this work, to make the problem feasible, we assume that the value of $P_a$ at least satisfies the rate requirement from Alice to Bob when no AN is transmitted by Bob. The rest of the design parameters that affect the performance of covert communication in our system model are the distribution parameters of Bob’s AN power, $\{P_{min}, P_{max}\}$, and the a priori probabilities of Alice’s transmission, $\{\pi_0, \pi_1\}$, with $\pi_0 = 1 - \pi_1$.

We state our main problem as following:

\[ \textbf{P1} \text{ maximize } \frac{E}{\pi_1 P_{min} + P_{max}} \]
subject to $\pi_1 R_{ab}(1 - \delta_{ab}) \geq \tau$, 
$P_{min} + P_{max} \leq 2P_{avg}$, 

while 

$P_a \geq \frac{\lambda_{ab} \sigma_b^2 (2R_{ab} - 1)}{\ln(R_{ab}/\tau)}$  \hspace{1cm} (21)

is assumed for feasibility. Here, the expression for $\frac{E}{\pi_1}$ is given in (15) under Lemma 2, $\delta_{ab}$ is the transmission outage probability and is a function of $\{P_{min}, P_{max}\}$, $\tau$ is the minimum required covert rate, and $P_{avg}$ is the average transmit power for Bob’s AN. We solve $\textbf{P1}$ in a step-by-step manner, as this approach not only provides the globally optimal solution, but also provides further insights in the role of different parameters in achieving the said purpose of covertness.

\section{A. Optimal Minimum AN Power}

For a given average transmit power at Bob, we look to minimize the value of transmission outage probability, in order to satisfy the covert rate requirement, corresponding to the first constraint in (20), while maximizing $\frac{E}{\pi_1}$. Under these conditions, in this subsection, we consider finding the optimal minimum AN power at Bob, $P_{min}$, for any given maximum AN power, $P_{max}$, and prior probabilities of Alice’s transmission, $\{\pi_0, \pi_1\}$.

Proposition 2: The optimal choice of $P_{min}$ to maximize the expected detection error probability at Willie, $\frac{E}{\pi_1}$, while satisfying the effective covert rate requirement from Alice to Bob is given by $P_{min} = 0$.

Proof: We first consider the maximization of $\frac{E}{\pi_1}$ where the optimal choice of $P_{min}$ should maximize $\kappa(t) = 1 + t \ln t - t^2$ under both cases of $\pi_1 \geq \pi_0$ and $\pi_1 < \pi_0$, as per (15). To first determine the monotonicity of $\frac{E}{\pi_1}$ w.r.t. $P_{min}$, we consider the derivatives of $\kappa(t)$ w.r.t. $t$, given by $\frac{\partial \kappa(t)}{\partial t} = 1 + \ln t - 2t$ and $\frac{\partial^2 \kappa(t)}{\partial t^2} = \frac{1}{t} - 2$. Since $P_{max} \geq P_{min}$, we have

\[ \frac{\partial \kappa}{\partial P_{min}} = \frac{\lambda_{aw} \lambda_{bw} P_a}{[\lambda_{bw} P_a + \lambda_{aw} (P_{max} - P_{min})]^2} \geq 0. \]

For $t \in [0, 1)$, the first derivative of $\kappa(t)$ w.r.t. $t$ increases for $0 \leq t < 1/2$ and decreases for $1/2 \leq t < 1$, with the maximum value of $-\ln 2$, occurring at $t = 1/2$. Using this and the fact that $\frac{\partial \kappa}{\partial P_{min}} \geq 0$, it can be concluded that $\kappa(t)$, and resultantly, $\frac{E}{\pi_1}$ is a decreasing function of $P_{min}$. Hence, the optimal choice in this regard is the minimum possible value of $P_{min}$, which is zero.

We next consider the covert rate constraint, where the outage probability $\delta_{ab}$ is represented as

\[ \delta_{ab} = 1 - \lambda_{bb} \exp(-\lambda_{ab} \mu \sigma_b^2) v(x), \]

and here 

\[ v(x) = \frac{1}{y - x} \ln \left( \frac{\lambda_{bb} + y}{\lambda_{bb} + x} \right). \]

Considering the first derivative of $v(x)$ w.r.t. $x$, we have

\[ \frac{\partial v(x)}{\partial x} = \frac{1}{(y - x)^2} \left[ \ln \left( \frac{\lambda_{bb} + y}{\lambda_{bb} + x} \right) - \frac{y - x}{\lambda_{bb} + x} \right]. \]

Here, $\frac{\partial v(x)}{\partial x}$ depends on

\[ l(x) \triangleq \ln \left( \frac{\lambda_{bb} + y}{\lambda_{bb} + x} \right) - \frac{y - x}{\lambda_{bb} + x} \]

\[ = \ln \left( 1 + \frac{y - x}{\lambda_{bb} + x} \right) - \frac{y - x}{\lambda_{bb} + x} \leq 0, \]

where the second line in (26) is due to the logarithmic inequality, $\ln(1 + a) \leq a, \forall a \geq -1$. Thus $v(x)$ is always a decreasing function of $x$, and resultantly, $\delta_{ab}$ is always an increasing function of $P_{min}$. From the covert rate constraint, we can write

\[ \delta_{ab} \leq 1 - \frac{\tau}{\pi_1 R_{ab}}, \]

and hence, to satisfy this constraint, $P_{min}$ is upper bounded by a value which can be found by solving (27) at equality. This concludes the proof.

As a result of Proposition 1, we can simplify the transmission outage probability at Bob and the expected detection error probability at Willie as

\[ \delta_{ab} = 1 - \lambda_{bb} \exp(-\lambda_{ab} \mu \sigma_b^2) \ln \frac{\lambda_{bb} + \lambda_{ab} \mu \sigma_b^2}{\lambda_{bb}} P_{max} \]

and

\[ \frac{E}{\pi_1} = \begin{cases} \pi_0 \left[ 1 + s \ln s - s^2 \right], & \text{if } \pi_1 \geq \pi_0 \\ \pi_1 \left[ 1 + s \ln s - s^2 \right], & \text{otherwise,} \end{cases} \]

respectively, where

\[ s = \frac{\lambda_{bw} P_a}{(\lambda_{bw} P_a + \lambda_{aw} P_{max})}. \]

\section{B. Optimal Priors for Alice’s Transmission}

Once the optimal value of $P_{min}$ has been found, the task from Alice and Bob’s perspective is to find the optimal a priori probabilities of Alice’s transmission and Bob’s maximum possible transmit power, $P_{max}$. In this subsection, we consider
finding the optimal choice of Alice’s \emph{a priori} transmission probabilities for a given \( P_{\text{max}} \). We state this problem as:

\textbf{P1.1} maximize \( \mathbb{P}_E \)  
subject to \( \pi_1 R_{ab}(1 - \delta_{ab}) \geq \tau, \)  
(31)

where the expression for \( \mathbb{P}_E \) is now given by (29), and the feasibility condition of (21) is still held. The solution to problem P1.1 is presented in the following:

\textbf{Proposition 3: The optimal choice of a priori probabilities for Alice’s transmission, as a function of maximum AN power, } \( P_{\text{max}} \), is given by

\[ \pi_1^* (P_{\text{max}}) = \max \left( \frac{1}{2}, \frac{\tau}{R_{ab}(1 - \delta_{ab}(P_{\text{max}}))} \right), \]  
(32)

and \( \pi_0^* = 1 - \pi_1^* \).

\textbf{Proof:} We consider the two cases for \( \mathbb{P}_E \) individually. Note here that \( \delta_{ab} \) is now a function of \( P_{\text{max}} \) only.

1) \( \pi_1 < \pi_0 \): In this case, \( \pi_1 < 1/2 \), and using the constraint in P1.1, we have \( R_{ab}(1 - \delta_{ab}(P_{\text{max}})) \leq \pi_1 < 1/2 \), which can only happen when \( R_{ab}(1 - \delta_{ab}(P_{\text{max}})) \leq 1/2 \). Also in this case, \( \frac{\partial \mathbb{P}_E}{\partial \pi_1} = 1 + s \ln s - s^2 \geq 0 \) for \( s \in [0,1] \).

2) \( \pi_1 \geq \pi_0 \): Here, \( \pi_1 \geq 1/2 \), and due to the constraint in P1.1, \( \pi_1 \geq \max \left( \frac{1}{2}, \frac{\tau}{R_{ab}(1 - \delta_{ab}(P_{\text{max}}))} \right) \). Also, in this case, \( \frac{\partial \mathbb{P}_E}{\partial \pi_1} = -(1 + s \ln s - s^2) \leq 0 \) for \( s \in [0,1] \).

Combining these two cases gives the desired result. \( \square \)

From Proposition 3, it is evident that the optimal value of \( \pi_1 \) is dependent upon the choice of \( P_{\text{max}} \). Thus to satisfy a given covert rate requirement, any choice of \( P_{\text{max}} \) at Bob, directly affecting the transmission outage probability through self-interference, will determine whether \( \pi_1^* \) is equal to 0.5 or not. Since the purpose of our covert scheme is to maximize the detection error at Willie while satisfying the rate requirement, it presents an interesting interplay of our choice of these parameters.

\section{C. Optimal Maximum AN Power}

Once the optimal priors for Alice’s transmission i.e., \( \{\pi_0^*, \pi_1^*\} \) have been found in terms of \( P_{\text{max}} \), the expected detection error probability at Willie becomes

\[ \mathbb{P}_E(\pi_1^*) = \begin{cases} \frac{1}{2} \kappa(s), & \text{if } \frac{\tau}{R_{ab}(1 - \delta_{ab})} \leq \frac{1}{2} \\ (1 - \frac{\tau}{R_{ab}(1 - \delta_{ab})}) \kappa(s), & \text{else} \end{cases} \]  
(33)

where again, \( \kappa(s) = (1 + s \ln s - s^2) \), and \( s \) is as defined earlier in (30). We now consider finding the optimal value for Bob’s maximum transmit power, \( P_{\text{max}} \), under the average power constraint. This problem is stated as

\textbf{P1.2} maximize \( \mathbb{P}_E(\pi_1^*) \)  
subject to \( \pi_1^* R_{ab}(1 - \delta_{ab}) \geq \tau, \)  
\( P_{\text{max}} \leq 2P_{\text{avg}}. \)  
(34)

We note here that in the statement of P1.2 above, \( \mathbb{P}_E \) from (29) has now been replaced by \( \mathbb{P}_E(\pi_1^*) \) in (33) and the feasibility condition of (21) is still held. Following the step-by-step approach, and due to the monotonicity of \( \mathbb{P}_E \) w.r.t \( P_{\text{min}} \) and \( \pi_1 \), P1.1 is now reduced to P1.2. The solution to this problem is presented in the following proposition.

\textbf{Proposition 4: The optimal value for Bob’s maximum transmit power under an average power constraint, } \( P_{\text{avg}} \), is given by

\[ P_{\text{max}}^* = \begin{cases} 2P_{\text{avg}}, & \text{if } 2P_{\text{avg}} \leq P_{\text{max}}^\dagger \\ P_{\text{max}}^\dagger, & \text{otherwise} \end{cases} \]  
(35)

where \( P_{\text{max}}^\dagger \) is the solution of \( \delta_{ab}(P_{\text{max}}) = 1 - \frac{2\tau}{R_{ab}} \) for \( P_{\text{max}} \) and \( P_{\text{max}}^* \) is the solution to

\[ \text{maximize} \quad \left( 1 - \frac{\tau}{R_{ab}(1 - \delta_{ab}(P_{\text{max}}))} \right) (1 + s \ln s - s^2) \]

subject to \( P_{\text{max}}^1 \leq P_{\text{max}} \leq 2P_{\text{avg}}. \)

\textbf{Proof:} We first show the monotonicity of \( \delta_{ab} \) w.r.t \( P_{\text{max}} \).

Here, \( \delta_{ab} \) can be written as

\[ \delta_{ab} = 1 - \lambda_{ab} \exp(-\lambda_{ab} \mu \sigma_b^2) u(x), \]

where \( u(x) = \frac{1}{2} \ln \left( \frac{\lambda_b + x}{\lambda_a} \right) \) and \( x \geq 0 \).

We note that

\[ \frac{\partial u(x)}{\partial x} = \frac{1}{2} \left( \frac{1}{\lambda_b + x} - \ln \left( \frac{\lambda_b + x}{\lambda_b} \right) \right), \]

(38)

which depends on \( m(x) = \frac{x}{\lambda_b + x} - \ln \left( \frac{\lambda_b + x}{\lambda_b} \right) \). Here, \( m(0) = 0 \) and \( \frac{\partial m(x)}{\partial x} = -\frac{1}{\lambda_b + x} \leq 0 \), thus \( m(x) \) decreases monotonically with \( x \), giving \( m(x) \leq m(0) \) for \( x \geq 0 \), and consequently, \( \frac{\partial \mathbb{P}_E}{\partial s} \leq 0 \). As a result, \( \delta_{ab} \) is a monotonically increasing function of \( P_{\text{max}} \).

Next, we consider the optimal choice of \( P_{\text{max}} \) under the two cases of \( \mathbb{P}_E \), keeping in view the change in \( \delta_{ab} \) w.r.t \( P_{\text{max}} \).

1) \( \frac{\tau}{R_{ab}(1 - \delta_{ab})} \leq \frac{1}{2} \), we have

\[ \delta_{ab} \leq 1 - \frac{2\tau}{R_{ab}}. \]  
(39)

Due to a monotonic increase in \( \delta_{ab} \) w.r.t \( P_{\text{max}} \), the optimal value of \( P_{\text{max}} \) has to satisfy \( P_{\text{max}} \leq P_{\text{max}}^1 \), where \( P_{\text{max}}^1 \) is the solution of (39) at equality. Combining with the average power constraint, we have \( P_{\text{max}} \leq \min \left( 2P_{\text{avg}}, P_{\text{max}}^1 \right) \).

Now we consider the monotonicity of \( \mathbb{P}_E(\pi_1^*) = \frac{1}{2}(1 + s \ln s - s^2) \) w.r.t \( P_{\text{max}} \).

Here, \( \frac{\partial \mathbb{P}_E(\pi_1^*)}{\partial s} = \frac{1}{2} \left( 1 + \ln s - 2s \right) \) and \( \frac{\partial \mathbb{P}_E(\pi_1^*)^2}{\partial s} = \frac{1}{2} - 1 \), where \( s = \frac{\lambda_{aw} P_{aw} + \lambda_{aw} \mu \sigma_b^2 P_a}{\lambda_{aw} P_{aw} + \lambda_{aw} \mu \sigma_b^2 P_a} \).

Also,

\[ \frac{\partial s}{\partial P_{\text{max}}} = -\frac{\lambda_{aw} \lambda_{aw} P_a}{(\lambda_{aw} P_a + \lambda_{aw} \mu \sigma_b^2 P_a)^2} \leq 0. \]  
(40)

We note here that \( s \in [0,1), \) \( \frac{\partial \mathbb{P}_E(\pi_1^*)}{\partial s} \) increases for \( 0 \leq s < 1/2 \) and decreases for \( 1/2 < s < 1 \), with a maximum value of \( -\frac{1}{2} \ln 2 \). Since \( \frac{\partial \mathbb{P}_E(\pi_1^*)}{\partial s} \leq 0 \), \( \mathbb{P}_E(\pi_1^*) \) is an increasing function of \( P_{\text{max}} \), and hence the best possible choice in this case is

\[ P_{\text{max}} = \min \left( 2P_{\text{avg}}, P_{\text{max}}^1 \right). \]
2) For $\frac{\tau}{R_{ab}(1-\delta_{ab})} > \frac{1}{2}$, we have

$$\delta_{ab} > 1 - \frac{2\tau}{R_{ab}}, \quad (41)$$

and in this case,

$$P_E^*(\pi_1^*) = \left(1 - \frac{\tau}{R_{ab}(1-\delta_{ab})}\right) (1 + s \ln s - s^2). \quad (42)$$

Since $\delta_{ab}$ increases monotonically in $P_{\text{max}}$, hence to satisfy (41), $P_{\text{max}} > P_{\text{avg}}$, and resultantly, the optimal choice lies between $P_{\text{max}}^1$ and $2P_{\text{avg}}$, where $P_{\text{max}}^1$ is as defined earlier. Let $\tilde{P}_E^*(\pi_1^*) = p(x)q(x)$, where $p(P_{\text{max}}) = \left(1 - R_{\text{avg}}(1-\delta_{ab}(P_{\text{max}}))\right)$ and $q(P_{\text{max}}) = (1 + s \ln s - s^2)$. We note here that $\tilde{P}_E^*(\pi_1^*)$ is not a monotonic function of $P_{\text{max}}$, since as $P_{\text{max}}$ increases, $p(P_{\text{max}})$ decreases while $q(P_{\text{max}})$ increases. Thus there may exist an optimal value of $P_{\text{max}}$ that maximizes $\tilde{P}_E^*(\pi_1^*)$, which motivates the optimization

$$P_{\text{max}}^* = \text{maximize} \quad \tilde{P}_E^*(\pi_1^*). \quad (43)$$

We note that the optimization problem in (43) is of one dimension and can be solved by methods of efficient numerical search.

Combining the two cases, the optimal value for $P_{\text{max}}$ is found, thus completing the proof.

Remark 3: The approach we have taken in solving $P_1$ guarantees that the obtained solution is globally optimal. Specifically, we first solved for the optimal $P_{\text{min}}$ for any value of $\pi_1$ and $P_{\text{max}}$. We next solved for the optimal $\pi_1$ as a function of any given $P_{\text{max}}$. Finally the optimal $P_{\text{max}}$ is obtained. The globally optimal solution to $P_1$ can thus be summarized as $P_{\text{max}}^* = 0$, $\pi_1^* = \text{max} \left(1/2, \tau/[R_{ab}(1-\delta_{ab}(P_{\text{max}}^*))]\right)$ and $I_{\text{max}}$ as given in (35) – (36).

As discussed in Remark 2, increasing the value of $P_{\text{max}}$ helps improve our covert performance, but on the other hand, $P_{\text{max}}$ also affects the covertly conveyed information through self-interference. Proposition 4 tells us that while satisfying the average power constraint, the optimal choice of $P_{\text{max}}$ satisfies the covert rate requirement with a bare equality, while maximizes the expected detection error probability at Willie. There might exist values of $\tau$ for which the choice of $P_{\text{max}}$ satisfying the rate constraint under $\pi = \frac{1}{2}$ are not the optimal choice of maximizing $\tilde{P}_E^*$. Under such a scenario, the freedom of adjusting $\pi_1$ helps us in meeting the rate requirement while keeping $\tilde{P}_E^*$ as high as possible.

VI. NUMERICAL RESULTS AND DISCUSSION

In this section, we present the numerical results and study the performance of our proposed scheme in achieving covert-ness while satisfying a given covert rate requirement. Unless otherwise stated, we set the transmit power at Alice $P_a = 10$ dB, a pre-determined rate for Alice to Bob transmission $R_{ab} = 1$, Bob and Willie’s receiver noise power $\sigma_b^2 = \sigma_w^2 = -10$ dB and Bob’s self-interference cancellation coefficient $^2$

$^2$Self-interference passive suppression of roughly 34 – 44 dB for FD systems has been reported in the literature [34], while a combination of passive suppression and active cancellation resulting in a total self-interference suppression of 90 dB has also been demonstrated [26].

Fig. 3. Optimal maximum transmit power of Bob’s AN, $P_{\text{max}}^*$, versus the covert rate requirement from Alice to Bob, $\tau$, for varying values of $P_a$.

Fig. 4. Optimal maximum transmit power of Bob’s AN, $P_{\text{max}}^*$, versus the covert rate requirement from Alice to Bob, $\tau$, for varying values of Bob’s self-interference cancellation coefficient, $\phi$.

$\phi = 0.01$. The average power constraint on Bob’s AN power is 40 dB, while for simplicity, the means of all fading channels are considered as $1/\lambda_{ab} = 1/\lambda_{aw} = 1/\lambda_{bw} = 1/\lambda_{bb} = 1$.

We first show the effect of $P_a$ and $\phi$ on the optimal maximum transmit power for Bob’s AN for varying covert transmission rate requirements, as demonstrated in Fig. 3 and Fig. 4, respectively. In Fig. 3 with a fixed value of $\phi$, a higher value of $P_a$ allows a higher value of $P_{\text{max}}^*$ to maximize the detection error probability at Willie, whilst satisfying the given rate requirement. In Fig. 4, with a fixed value of $P_a$ in the feasible range, a higher value of $\phi$ (i.e., poorer self-interference cancellation) requires a lower value of $P_{\text{max}}^*$ (i.e., less self-interference) to satisfy the same rate requirement. We note here that in such circumstances, a reduced $P_{\text{max}}^*$ for a given $P_a$ will adversely affect the achievable covertness.
We next consider the effect of $P_a$ and $\phi$ on the optimal transmission probability of Alice’s covert transmissions for varying covert transmission rate requirements, as demonstrated in Fig. 5 and Fig. 6, respectively. From Fig. 5, we see that for a given $P_a$, a choice of $\pi_1 = 1/2$ is optimal up to a certain value of $\tau$, but a further increase in $\tau$ results in an increase in optimal $\pi_1$. For a given $P_a$, a rate requirement can be met by decreasing the value of $P_{\text{max}}^*$, but it will in return decrease the achievable covertness. Keeping in view the results shown in Fig. 3 and Fig. 4, the optimal solution dictates that instead of making a drastic change in $P_{\text{max}}^*$, a better choice is to decrease $P_{\text{max}}^*$ a little while $\pi_1$ can be increased to meet the rate requirement.

As the value of $P_a$ is increased, the same effect appears for a little higher value of $\tau$. Fig. 6 shows the effect of increasing $\phi$ on the optimal $\pi_1$ for a given value of $P_a$, which is inverse of what is observed for increasing $P_a$. Since an increase in $\phi$ will have a detrimental effect on the transfer of covert information, thus to keep the covertness high and to satisfy the rate requirement, an increase in $\pi_1$ is desired for an even lower value of $\tau$. To further demonstrate the effect of $\phi$, we show the expected detection error probability at Willie for different values of $\phi$ in Fig. 7. For a fixed $P_a$ and a given $\phi$, as $\tau$ increases, $P_{\text{max}}^*$ at Bob has to decrease in order to reduce Bob’s self-interference. A lower value of $P_{\text{max}}^*$ will result in a lower $P_{E}^*$, since it decreases the confusion in received signal statistics at Willie. This effect is shown more explicitly in Fig. 8, where we show the effect of $\phi$ on the performance of proposed covert scheme through the expected minimum detection error probability at Willie $P_{E}^*$. It should be noted here that a value of $\phi = 0$ corresponds to a perfect cancellation of the self-interference while $\phi = 1$ refers to no cancellation or suppression at all, representing the worst case scenario for the FD receiver. For a higher value of $\phi$, Bob has...
to reduce $P_{\text{max}}^*$ to satisfy a certain rate requirement, which in effect, reduces the achievable covertness.

Last but not least, we investigate the advantage of our proposed scheme of jointly optimizing $\tau_1$ and $P_{\text{max}}$ over a benchmark scheme of only optimizing $P_{\text{max}}$ while keeping $\tau_1 = 0.5$. Fig. 9 shows the overall performance of our proposed scheme in terms of the expected detection error probability at Willie versus the covert rate requirement from Alice to Bob. For $\tau \in [0, 0.44]$, the proposed joint optimization scheme performs the same as the benchmark scheme and there is no discernable difference in $P_E^*$ for the two schemes. However, for $\tau \geq 0.44$, the optimal $\tau_1$ starts to deviate from 0.5, as shown in Fig. 5 and Fig. 6. Here, the $P_E^*$ achieved by the joint optimization scheme reduces gradually as the rate requirement increases, but the $P_E^*$ for the benchmark scheme drops sharply, and at $\tau = 0.5$, the benchmark scheme offers $P_E^* \approx 0.005$, which means almost no covertness at all. Thus for $\tau \geq 0.44$, the proposed joint optimization scheme provides a significant gain in the achievable covertness.

VII. CONCLUSION

In this paper, we have considered the potential of achieving covert communication using a full-duplex receiver that generates artificial noise to cause detection errors at a watchful adversary Willie. Considering a radiometer as the detector of choice at Willie, we have analyzed the conditions under which Willie makes detection errors, and characterized Willie’s optimal detection performance conditioned over the fading channel realizations. From the perspective of covert communication pair, we have provided design guidelines for the optimal choice of transmit power of full-duplex receiver’s artificial noise. Owing to the self-interference of the full-duplex receiver, these power levels need to be controlled carefully, otherwise they affect the transfer of any covert information. We have also shown that contrary to a commonly adopted assumption, the \textit{a priori} transmission probabilities of 0.5 are not always the optimal choice to achieve the best possible covertness.

A limitation of the current work lies in its assumption of $n \rightarrow \infty$. Future work will focus on scenarios where $n$ is finite. This will provide more precise results relating to future communication systems, especially for delay-intolerant systems.

APPENDIX

PROOF OF PROPOSITION 1

Using the definition of incorrect decisions at Willie, we have

\[
\mathbb{P}_F = \mathbb{P}[D_1|H_0] = \mathbb{P}[P_w > \gamma|H_0]
\]

\[
= \mathbb{P}[h_{bw}^2 P_b + \sigma_w^2 \geq \gamma|H_0] = \mathbb{P}
\]

\[
= \begin{cases}
1, & \text{if } \gamma - \frac{\sigma_w^2}{|h_{bw}|^2} \leq P_{\text{min}} \\
0, & \text{if } P_{\text{min}} < \gamma - \frac{\sigma_w^2}{|h_{bw}|^2} \leq P_{\text{max}} \\
\end{cases}
\]

and

\[
\mathbb{P}_{MD} = \mathbb{P}[D_0|H_1] = \mathbb{P}[P_w < \gamma|H_1]
\]

\[
= \mathbb{P}[h_{bw}^2 P_b + |h_{aw}|^2 P_a + \sigma_w^2 \geq \gamma]
\]

\[
= \mathbb{P}
\]

\[
= \begin{cases}
0, & \text{if } \nu \leq P_{\text{min}} \\
1, & \text{if } P_{\text{min}} < \nu \leq P_{\text{max}} \\
\end{cases}
\]

where $\nu \triangleq 2\frac{\gamma - \sigma_w^2}{|h_{bw}|^2} P_a - \frac{|h_{bw}|^2 P_{\text{min}} - \sigma_w^2}{|h_{bw}|^2}$.

We also note that

- $|h_{bw}|^2 P_{\text{min}} + \sigma_w^2 \leq |h_{bw}|^2 P_{\text{max}} + |h_{aw}|^2 P_a + \sigma_w^2$, but the relationship between $|h_{bw}|^2 P_{\text{min}} + |h_{aw}|^2 P_a + \sigma_w^2$ and $|h_{bw}|^2 P_{\text{max}} + \sigma_w^2$ is unclear.

- For a choice of $\gamma < |h_{bw}|^2 P_{\text{min}} + \sigma_w^2$, $\mathbb{P}_{FA} = 1$, $\mathbb{P}_{MD} = 0$ and hence $\mathbb{P}_E = \pi_0$.

- For a choice of $\gamma > |h_{bw}|^2 P_{\text{max}} + |h_{aw}|^2 P_a + \sigma_w^2$, $\mathbb{P}_{FA} = 0$, $\mathbb{P}_{MD} = 1$ and hence $\mathbb{P}_E = \pi_1$. 
which has a decreasing partial derivative w.r.t. $\gamma$, given by

$$
\frac{\partial P_E}{\partial \gamma} = \frac{\pi_0 - \pi_1}{|h_{bw}|^2 (P_{max} - P_{min})} = \begin{cases} 
0 & \text{if } \pi_1 \geq \pi_0 \\
< 0 & \text{otherwise.}
\end{cases}
$$

(52)

Based on the knowledge of $\pi_0$ and $\pi_1$, Willie can choose the optimal value of $\gamma$.

The corresponding $P_E$ for the choice of optimal threshold, $\gamma^*$, can be found by using the appropriate expressions of $P_E$ from Case-II, hence concluding the proof.

REFERENCES


Khuurram Shahzad (S’16) received the B.E. and M.S. degrees in electrical engineering from the National University of Sciences and Technology (NUST), Pakistan, in 2004 and 2008, respectively, and the M.S. degree in wireless communications from the University of Maryland, College Park, MD, USA, in 2013. He is currently pursuing the Ph.D. degree with the Research School of Engineering, Australian National University, Canberra, ACT, Australia. In 2010, he was awarded the Fulbright Scholarship for graduate studies in electrical engineering at the University of Maryland for his M.S. degree. He was involved in different teaching capacities at NUST from 2008 to 2010 and from 2013 to 2016. His research interests lie in the areas of wireless communications, physical layer security, and covert communications.

Xiangyun Zhou (SM’17) received the Ph.D. degree from Australian National University (ANU), Canberra, ACT, Australia, in 2010. He is currently a Senior Lecturer at ANU. His research interests are in the fields of communication theory and wireless networks. He was a recipient of the Best Paper Award at ICC 2011 and the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2016. He was the Chair of the ACT Chapter of the IEEE Communications Society and Signal Processing Society from 2013 to 2014. He was named the Best Young Researcher in the Asia-Pacific Region in 2017 by the IEEE ComSoc Asia-Pacific Board. He has served as symposium/track and workshop co-chair for major IEEE conferences. He has been serving as an Editor for various IEEE journals, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE WIRELESS COMMUNICATIONS LETTERS, and the IEEE COMMUNICATIONS LETTERS. He also served as a Guest Editor for the IEEE Communications Magazine’s feature topic on wireless physical layer security in 2015.

Shihao Yan (M’15) received the B.S. degree in communication engineering and the M.S. degree in communication and information systems from Shandong University, Jinan, China, in 2009 and 2012, respectively, and the Ph.D. degree in electrical engineering from The University of New South Wales, Sydney, NSW, Australia, in 2015. From 2015 to 2017, he was a Post-Doctoral Research Fellow with the Research School of Engineering, The Australian National University, Canberra, ACT, Australia. He is currently a Research Fellow with the School of Engineering, Macquarie University, Sydney, NSW, Australia. His current research interests are in the areas of wireless communications and statistical signal processing, including physical layer security, covert communications, and location spoofing detection.
Jinsong Hu received the B.S. degree from the Nanjing University of Science and Technology, Nanjing, China, in 2013, where he is currently pursuing the Ph.D. degree with the School of Electronic and Optical Engineering. From 2017 to 2018, he was a Visiting Ph.D. Student with the Research School of Engineering, Australian National University, Canberra, ACT, Australia. His research interests include array signal processing, covert communications, and physical layer security. He served as a TPC member for the IEEE ICC 2019.

Feng Shu (M’16) was born in 1973. He received the B.S. degree from the Fuyang Teaching College, Fuyang, China, in 1994, the M.S. degree from the Xi’ian University, Xi’an, China, in 1997, and the Ph.D. degree from Southeast University, Nanjing, in 2002. In 2005, he joined the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China, where he is currently a Professor and a Supervisor of Ph.D. and graduate students. From 2009 to 2010, he was a Visiting Post-Doctoral with The University of Texas at Dallas, Richardson, TX, USA. He was with Fujian Agriculture and Forestry University and was awarded with the Minjiang Scholar Chair Professor in 2016 and the Fujian Hundred-Talents Program in Fujian Province in 2018. He serves as the Director of the Institute of Wireless Communications and Sensor Networks. He has published over 200 archival journal papers, including over 50 papers in IEEE journals and over 80 SCI-indexed papers. He holds six Chinese patents. His interest research areas span physical-layer security, wireless covert communications, UAV, wireless networks, wireless location, and array signal processing. He has served as the Session Chair or as a Technical Program Committee Member for various international conferences, such as ICC 2019, IEEE ICCS 2016/2018, ISAPE 2018, IEEE WCSP 2016, IEEE VTC 2016, and so on. He is an Editor of the IEEE ACCESS.

Jun Li (M’09–SM’16) received the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009. In 2009, he was with the Department of Research and Innovation, Alcatel Lucent Shanghai Bell, as a Research Scientist. Since 2015, he has been with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China. His research interests include network information theory, channel coding theory, wireless network coding, and cooperative communications.