Zero-forcing Beamforming for Jamming Aided Covert Communication in MIMO Systems

He Zhu[1], and Huihui Wu[2]
[1]School of Systems Information Science, Future University Hakodate, 116-2
Kamedanakano-cho, Hakodate, Hokkaido, 041-8655, Japan
[2]Department of Automation, Tsinghua University, 100084, Beijing, China

This paper investigates covert wireless communication in a multi-input multi-output (MIMO) system, where a multi-antenna transmitter transmits covert signal to a multi-antenna receiver with the assistance of a multi-antenna jammer employing zero-forcing (ZF) beamforming to avoid detection by a warden equipped with single antenna. For the given scenario, we first provide theoretical modeling for the minimum detection error probability at warden and covert rate from transmitter to receiver. We then explore the optimal beamformer design of the jammer to maximize the covert rate, subjecting to the constraints of covertness, upper bounds of transmitter and jammer transmit power, and beamformer on jammer. We further devise an iterative algorithm to determine the related optimal precoding matrix for covert rate maximization. Finally, we provide extensive numerical results to illustrate the impact of system parameters on covert performance.

***Index Terms*—Covert communication, multi-input multi-output, zero-forcing beamforming, jamming**

I. Introduction

A. Background

SECURITY of private information transmission has always been considered a critical issue in wireless communication networks. However, due to the broadcast nature of wireless communications, the open communications environment makes wireless transmissions more vulnerable to malicious attackers [1]. Traditionally, to protect the private information security, the researchers employ cryptography to prevent unauthorized disclosure and modification of data [2], [3] or employ physical layer security (PLS) to avoid the information leakage [4], [5]. However, cryptography or PLS is not secure enough since both of them only prevent eavesdropping but do not hide the existence of transmission. Some wireless scenarios require higher security, such as submarines, unmanned aerial vehicles (UAVs), government, etc. In this regard, covert communication is emerging.

Covert communication is a technique which hides the existence of wireless transmission with a Low Probability of Detection (LPD). The objective of covert communication is to reliably deliver information from a legitimate transmitter to its intended receiver while minimizing the risk of detection by a vigilant warden. Here, warden typically refers to any entity or system that might attempt to intercept or detect the communication. To make this possible, the transmitter exploit the inherent uncertainty or interference of the channel such that the legitimate users can correctly decode the information, while the warden can not detect the communication from their observations.

B. Related Works

Covert communication owes its origins to the pioneering work of [6], where the authors proved that the theoretical capacity limit of covert communication is subjected to a square root law (SRL). The SRL states that in a basic three-node model scenario, operating in an Additive White Gaussian Noise (AWGN) channel, it is possible to reliably and covertly transmit $O(\sqrt{n})$ bits in n channels, where $O(f(n))$ represents the asymptotic upper bound. In addition, when $n \to \infty$, the transmitted signal per channel equals 0, i.e., $\lim_{n\to\infty} \frac{O(\sqrt{n})}{n} \to 0$. This result can be extended to other channel models, e.g., Discrete Memoryless Channels (DMCs), Binary Symmetric Channels (BSCs), Multiple Access Channels (MACs), and Poisson channels [7]–[10].

To escape the predicament of a zero rate, recently, some research efforts have demonstrated that the positive covert rate can be achieved when the warden has uncertainty regarding the channel state information (CSI) from the transmitter to itself (e.g., noise uncertainty, channel estimation uncertainty, transmission time slot uncertainty, and position uncertainty). These uncertainty information can confuse the warden and make it impossible to distinguish between the power of transmission signal and the background noise. The work in [11] considered a large-scale wireless network where the transmitters establish a stable Poisson Point Process (PPP), then the warden is subjected to aggregate interference by other transmitters. The works in [12], [13] employ cooperative jamming techniques to prevent the warden from detecting the communication signal. The works in [14], [15] utilize artificial noise (AN) injection to confuse the warden and also offer guidelines to determine the optimal AN power for enhancing covertness. Besides, the work in [16] employs an additional non-covert channel (i.e., a public channel) as a cover and achieves a constant covert rate. The work in [17] considers a movable transmitter (i.e., unmanned aerial vehicle) to escape detection and proposes a joint optimization of UAV trajectory and transmission power to enhance covert performance. Furthermore, the work in [18] adopts power allocation method to achieve a positive covert

Manuscript received December 4, 2023; revised January 4, 2024. Corresponding author: He Zhu (email: zhuhe2023@163.com).

rate.

Recently, some studies have shown that the multi-antenna technique can significantly improve covert communication performance. In the notable work [19], the authors establish that in the MIMO AWGN channel system, when the detector possesses bounded spectral norm channel state information, the system can achieve a maximum number of covert bits $O(\sqrt{n})$. The result verifies that the SRL is still satisfied with covert MIMO system. In work [20], the authors consider the presence of secret codebook scenario and utilize KL divergence constraint to analyze the covert capacity in MIMO AWGN channels. The work [21] focuses on enhancing noise uncertainty by using the multi-antenna technique, thereby achieving a positive transmission rate. The work of [22] considers a scenario in which a multi-antenna jammer assists the transmission and then optimizes the transmit power and target rate to maximize the covert rate. In work [23], the 3D beamformer and an iterative algorithm are proposed to obtain the optimal beamformer matrix to maximize the covert rate.

C. Motivations

Although the previous works provide a comprehensive understanding of utilizing the uncertainty of CSI to achieve a positive covert rate and demonstrate that multi-antenna can enhance the covert performance. However, there are still research gaps in covert MIMO communication: (1) how to model the covert system to achieve a positive covert rate in the MIMO scenario? (2) how to design the beamforming vector to benefit the communication covertness? To address these problems, this paper, for the first time, focuses on the study of beamforming with ZF beamforming for covert communication in MIMO systems. Our main contributions are summarized as follows:

- We investigate the covert communication in the MIMO system, where a multi-antenna transmitter Alice transmits a covert message to a multi-antenna receiver Bob with the help of a multi-antenna jammer, subjecting to the detection from a single-antenna warden Willie. The jammer employs the ZF beamforming technique so that it can confuse Willie's detection and ensure reliable covert transmission from Alice to Bob.
- Under this model, we apply hypothesis testing, statistics theories, and matrix theories to develop a theoretical framework for covert performance modeling in terms of detection error probability and achievable covert rate. We further develop an iterative algorithm to find the optimal beamforming matrix.
- We provide extensive numerical results to illustrate the potential performance enhancement from adopting the beamforming technique in covert MIMO communication.

The reminder of this paper is organized as follows. Section II introduces the system models including network model, channel model, transmission model and detection model. Section III formulates covert performance in the considered MIMO system and a optimal problem to achieve the maximal covert rate is developed. We present the numerical results in Section IV, and make a conclusion in Section V.

Throughout this paper, we use the following notations. The lower-case and bold lower-case letters denote scalars and vectors (e.g., x, and $\mathbf{x}$). The bold upper-case letters denote matrices (e.g., $\mathbf{X}$). Let $||\mathbf{X}||$, $\mathbf{X}^H$, $Tr(\mathbf{X})$, $det(\mathbf{X})$ denote the Frobenius norm, Hermite, trace, determinant of matrix $\mathbf{X}$, respectively. Denote $\mathbf{0}$ as the null vector and $\mathbf{I}$ as the identity matrix. The notation $\mathbb{E}(\cdot)$ denotes the expectation operator.

II. System Model

A. Network Model

As shown in Fig.1, we consider a four-node MIMO system consisting of a transmitter (Alice), a receiver (Bob), a warden (Willie), and a jammer. Under the considered system, Alice intends to send a covert signal to Bob via the help of a jammer against the detection of Willie, who passively observes the transmission and distinguishes whether there is a transmission between Alice and Bob or not. It is assumed that Willie is equipped with a single antenna, Alice, Bob, and jammer are all equipped with multiple antennas where the number of antennas equipped in Alice, Bob, and jammer are denoted as N_a, N_b, and N_j, respectively. To achieve covert communication between Alice and Bob, the jammer continuously injects a jamming signal to confuse Willie. We assume that the location of Willie is unchanged, and the jammer employs the zero-forcing beamforming, which means the main beam of the jammer focuses on Willie, and Bob does not receive any jamming signal from the jammer.

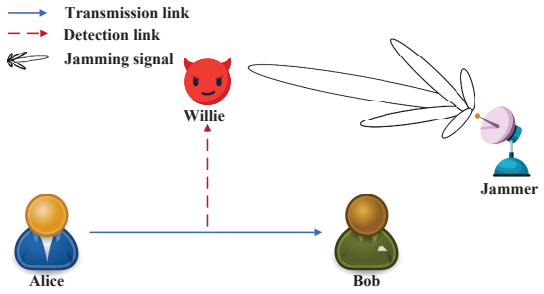

Fig. 1. System model.

B. Channel Model

All the wireless links are modeled as a quasi-static Rayleigh fading channel model, where the channel fading coefficients remain constant in one time slot, but vary independently and randomly from one time slot to the next. The channel matrixes from Alice to Bob and jammer to Bob are denoted as $\mathbf{H}_{ab} \sim \mathcal{CN}(\mathbf{0}_{N_a \times N_b}, \mathbf{C}_{ab})$, $\mathbf{H}_{jb} \sim \mathcal{CN}(\mathbf{0}_{N_j \times N_b}, \mathbf{C}_{jb})$, respectively. Here, $\mathbf{0}$ represents the zero matrix, and $\mathbf{C}_{ab}$ and $\mathbf{C}_{jb}$ denote the positive definite channel covariance matrix from Alice to Bob and jammer to Bob, respectively. The channel matrixes between Alice and Willie, and between jammer and Willie, are denoted as $\mathbf{H}_{aw} \sim \mathcal{CN}(\mathbf{0}_{N_a \times 1}, \mathbf{C}_{aw})$, $\mathbf{H}_{jw} \sim \mathcal{CN}(\mathbf{0}_{N_j \times 1}, \mathbf{C}_{jw})$, respectively, where $\mathbf{C}_{aw}$ and $\mathbf{C}_{jw}$ are the channel covariance matrixes between Alice and Willie, and between jammer and Willie. Following the common

assumptions [23], we assume that the instantaneous value of $\mathbf{H}_{ab}$ and $\mathbf{H}_{jb}$ are available to Alice, Bob and jammer, but the statistical characterizations of $\mathbf{H}_{aw}$ and $\mathbf{H}_{jw}$ are known to Willie.

C. Transmission Model

Without loss of generality, the transmission process is divided into N time slots, and n symbols are transmitted in each time slot. In time slot t ($t \in \{1, 2, \cdots, N\}$), the transmitted covert signal at Alice is denoted as $\mathbf{s}_a = [s_a(1), s_a(2), \ldots, s_a(n)]$, where $s_a(i)$ is the i-th transmitted symbol from Alice which satisfies $\mathbb{E}\{|s_a(n)|^2\} = 1$. The associated beamforming vector is denoted as $\mathbf{G} \in \mathbb{C}^{N_a \times 1}$. Let $\mathbf{s}_j = [s_j(1), s_j(2), \ldots, s_j(n)]$ denote the jammer's signal in the t-th time slot, where $s_j(i)$ is the i-th transmitted symbol from jammer with the covariance $\mathbb{E}\{|s_j(i)|^2\} = 1$. The associated jammer beamforming vector is denoted as $\mathbf{F} \in \mathbb{C}^{N_j \times 1}$. Then the received signal $\mathbf{y}_b$ at Bob can be given as

$$\mathbf{y}_b(i) = \sqrt{P_a}\mathbf{H}_{ab}\mathbf{G}\mathbf{s}_a(i) + \mathbf{H}_{jb}\mathbf{F}\mathbf{s}_j(i) + \mathbf{n}_b(i), \quad (1)$$

where P_a is the transmit power of Alice, which is subjected to the maximum transmit power constraint of P_a^{max}; $\mathbf{n}_b(i)$ is the zero-mean white Gaussian noise at B with variance $\sigma_b^2\mathbf{I}_{N_b}$, i.e., $\mathbf{n}_{ab}(i) \sim (0, \sigma_{ab}^2\mathbf{I}_{N_b})$.

Note that the ZF beamforming scheme is practical and simple since it is constructed based on linear beamforming. The beamforming vector $\mathbf{F}$ and $\mathbf{G}$ are designed as a low-rank matrix, which brings a low computational cost [24].

D. Detection Model

Regarding the detection of covert communication, we assume that Willie employs the energy detection method and uses a radiometer to detect the signals [25], [26]. Under this assumption, Willie attempts to distinguish the signals between noise (i.e., background noise and jamming signal) and transmission signals to determine whether Alice is transmitting signals or not. To this end, Willie needs to conduct a hypothesis test based on its received signals y_w, enabling the formulation of a binary decision. In this test, the null hypothesis $\mathcal{H}_0$ means that there is no communication between Alice and Bob in the time slot, while the alternative hypothesis $\mathcal{H}_1$ indicates that Alice and Bob are communicating with each other.

Following the hypothesis test theory in signal detection, the received signals $\mathrm{y}_w(i)$ at Willie for the i-th symbol can be given by

$$\mathrm{y}_w(i) = \begin{cases} \mathbf{H}_{jw}^H\mathbf{F}s_j(i) + n_w(i), & \mathcal{H}_0, \\ \sqrt{P_a}\mathbf{H}_{aw}^H\mathbf{G}s_a(i) + \mathbf{H}_{jw}^H\mathbf{F}s_j(i) + n_w(i), & \mathcal{H}_1, \end{cases} \quad (2)$$

where $n_{jw}(i)$ is the AWGN signal at Willie with zero mean and variance σ_w^2, i.e., $n_w(i) \sim \mathcal{CN}(0, \sigma_w^2)$.

In general, the hypothesis test introduces two types of detection errors, i.e., the false alarm (FA) and the miss detection (MD). FA represents Willie erroneously judges the presence of communication between Alice and Bob, even when there is no actual communication between the transceivers. MD represents that the communication between Alice and Bob exists, but Willie fails to detect it. Thus, based on the received signal of the two statuses at Willie in (2), the decision rule can be given by

$$P_w \triangleq \frac{1}{n}\sum_{i=1}^{n} |\ \mathrm{y}_w(i)\ |^2 \underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}} \tau, \quad (3)$$

where P_w denotes the average received power at Willie in the t-th time slot; τ represents the detection threshold; $\mathcal{D}_0$ and $\mathcal{D}_1$ represent the decision results, when $P_w \geq \tau$, Alice transmits the signals which is denoted by $\mathcal{D}_1$, when $P_w < \tau$, Alice keeps silence which is denoted by $\mathcal{D}_0$.

The probability of FA $\mathcal{P}_{FA}$ and the probability of MD $\mathcal{P}_{MD}$ can be determined as

$$\mathcal{P}_{FA} = \mathbb{P}(\mathcal{D}_1|\mathcal{H}_0) = \mathbb{P}(P_w \geq \tau|\mathcal{H}_0), \quad (4)$$

and

$$\mathcal{P}_{MD} = \mathbb{P}(\mathcal{D}_0|\mathcal{H}_1) = \mathbb{P}(P_w < \tau|\mathcal{H}_1). \quad (5)$$

Let ξ represents the total detection error probability of Willie, then ξ can be given by

$$\xi = \mathcal{P}_{FA} + \mathcal{P}_{MD}. \quad (6)$$

We define that the transmission can achieve covert communication when the detection error probability $\xi \geq 1 - \epsilon$ for any $\epsilon > 0$, where ϵ is the covert constraint. When $\xi = 0$, it indicates that Willie can completely and correctly detect covert signal transmitted by Alice; when ξ approaches 1, it indicates that the reliability of the detection at Willie decreases. The metrice ξ can intuitively represent the covert probability of the system. It is apparent that as the sum of the two types of errors in the hypothesis test increases, we will have a higher likelihood of successfully hiding the communication between Alice and Bob.

III. Covert Performance Analysis

In this section, we provide the analysis for two kinds of matrices to access the covert performance for the considered model, which is depicted in Fig. 1. We first derive the detection error probability at Willie. Subsequently, we provide the theoretical analysis for covert rate of the system.

A. Detection Error Probability

To analysis the total detection error probability ξ at Willie, we first derive $\mathcal{P}_{FA}$ and $\mathcal{P}_{MD}$, respectively.

False alarm (FA) case: Recall that when Alice does not send the covert signal, Willie only receives jamming signals and background noise. In this case, if $P_w \geq \tau$, Willie still accepts the null hypothesis $\mathcal{H}_0$, then FA occurs. As the same assumption in [27], Alice employs Gaussian codebook and jammer uses Gaussian signal; then $\mathrm{y}_w|\mathbf{H}_{jw} \sim \mathcal{CN}(0, |\mathbf{H}_{jw}^H\mathbf{F}|^2 + \sigma_w^2)$. According to (3), $\sum_{i=1}^{n} |\mathrm{y}_w|^2|\mathbf{H}_{jw} \sim \mathcal{CN}(0, |\mathbf{H}_{jw}^H\mathbf{F}|^2 + \sigma_w^2)\chi_{2n}^2$, where χ_{2n}^2 is the chi-squared random variable with $2n$ degrees of freedom [16]. According to the Strong Law of Large Numbers, $\frac{\chi_{2n}^2}{n}$ converges to 1,

and based on Lebesgue's Dominated Convergence Theorem, we replace $\frac{\chi^2_{2n}}{n}$ with 1 when $n \to \infty$. By substituting (3) into (4), the probability of FA $\mathcal{P}_{FA}$ is given by

$$\begin{aligned}\mathcal{P}_{FA} &= \mathbb{P}\left(P_w \geq \tau | \mathcal{H}_0\right) \\ &= \mathbb{P}\left(\left(\mid \mathbf{H}_{jw}^H \mathbf{F} \mid^2 + \sigma_w^2\right)\frac{\chi^2_{2n}}{n} \geq \tau | \mathcal{H}_0\right) \\ &= \mathbb{P}\left(\mid \mathbf{H}_{jw}^H \mathbf{F} \mid^2 + \sigma_w^2 \geq \tau | \mathcal{H}_0\right).\end{aligned} \tag{7}$$

Let $J_w = \mid \mathbf{H}_{jw}^H \mathbf{F} \mid^2$ denote the channel gain from jammer to Willie, which can be modeled as a random variable following an exponential distribution [27], Then, the probability density function (PDF) of J_w can be given by

$$f_{J_w}(x) = \frac{e^{-\frac{x}{\mathbf{F}^H \mathbf{C}_{jw} \mathbf{F}}}}{\mathbf{F}^H \mathbf{C}_{jw} \mathbf{F}}. \tag{8}$$

Substituting (8) into (7), $\mathcal{P}_{FA}$ can be given as

$$\mathcal{P}_{FA} = \begin{cases} e^{-\frac{\tau - \sigma_w^2}{\mu_1}}, & \tau > \sigma_w^2, \\ 1, & \text{otherwise}, \end{cases} \tag{9}$$

where $\mu_1 = \mathbf{F}^H \mathbf{C}_{jw} \mathbf{F}$.

Miss detection (MD) case: When Alice sends the covert signal, Willie receives covert signals, jamming signals, and background noise. In this case, if $P_w < \tau$, Willie still accepts the alternative hypothesis $\mathcal{H}_1$, leading to a MD event. Same as the analysis in FA case, let $A_w = |\mathbf{H}_{aw}^H \mathbf{G}|^2$ denote the channel gain from Alice to Willie, then $\mathrm{y}_w | \mathbf{H}_{jw} \sim \mathcal{CN}(0, P_a A_w + J_w + \sigma_w^2)$, and $\sum_{i=1}^{n} |\mathrm{y}_w|^2 | \mathbf{H}_{jw} \sim \mathcal{CN}(0, P_a A_w + J_w + \sigma_w^2)\chi^2_{2n}$. The PDF of the power received from Alice and jammer can be given as

$$f_{P_a A_w + J_w}(x) = \frac{e^{-\frac{x}{P_a \mathbf{G}^H \mathbf{C}_{aw} \mathbf{G}}} - e^{-\frac{x}{\mathbf{F}^H \mathbf{C}_{jw} \mathbf{F}}}}{P_a \mathbf{G}^H \mathbf{C}_{aw} \mathbf{G} - \mathbf{F}^H \mathbf{C}_{jw} \mathbf{F}}. \tag{10}$$

By substituting (3) and (10) into (5), the probability of MD $\mathcal{P}_{MD}$ is given by

$$\begin{aligned}\mathcal{P}_{MD} &= \mathbb{P}\left(P_w < \tau | \mathcal{H}_1\right) = \mathbb{P}\left(P_a A_w + J_w + \sigma_w^2 < \tau | \mathcal{H}_1\right) \\ &= \int_0^{\tau - \sigma_w^2} \frac{e^{-\frac{x}{P_a \mu_2}} - e^{-\frac{x}{\mu_1}}}{P_a \mu_2 - \mu_1} dx \\ &= \begin{cases} 1 - \frac{P_a \mu_2}{P_a \mu_2 - \mu_1} e^{-\frac{\tau - \sigma_w^2}{P_a \mu_2}} + \frac{\mu_1}{P_a \mu_2 - \mu_1} e^{-\frac{\tau - \sigma_w^2}{\mu_1}}, & \tau > \sigma_w^2, \\ 0, & \text{otherwise}, \end{cases}\end{aligned} \tag{11}$$

where $\mu_2 = \mid \mathbf{H}_{aw}^H \mathbf{G} \mid^2$.

Substituting (9) and (11) into (6), the total detection error probability ξ of Willie is given by

$$\xi = \begin{cases} 1 - \frac{P_a \mu_2}{P_a \mu_2 - \mu_1} \left(e^{-\frac{\tau - \sigma_w^2}{P_a \mu_2}} - e^{-\frac{\tau - \sigma_w^2}{\mu_1}} \right), & \tau > \sigma_w^2, \\ 1, & \text{otherwise}. \end{cases} \tag{12}$$

According to (12), the ξ is influenced by τ. In evaluating the covert performance, we consider the worst-case scenario for the system, where Willie deploys an optimal detection threshold τ^* to minimize its detection error probability. Then, we have the following theorem regarding the minimum detection error probability.

Theorem 1. *For the concerned MIMO system with transmit power P_a at Alice, the beamforming vector $\mathbf{G}$ at Alice and $\mathbf{F}$ at jammer, the channel covariance matrices $\mathbf{C}_{aw}$ between Alice and Willie and $\mathbf{C}_{jw}$ between jammer and Willie, the minimum detection error probability ξ^* at Willie is determined as*

$$\xi^* = 1 - \frac{P_a \mu_2}{P_a \mu_2 - \mu_1} \left(\left(\frac{P_a \mu_2}{\mu_1}\right)^{-\frac{\mu_1}{P_a \mu_2 - \mu_1}} - \left(\frac{P_a \mu_2}{\mu_1}\right)^{-\frac{P_a \mu_2}{P_a \mu_2 - \mu_1}} \right). \tag{13}$$

Proof. According to (12), we can observe that when $\tau \leq \sigma_w^2$, $\xi = 1$. It indicates that τ is too small to distinguish the transmitted signals and the noise, thus Willie cannot detect the communication between Alice and Bob absolutely. When $\tau > \sigma_w^2$, we calculate the first derivative of ξ with respect to τ, it can be determined as

$$\frac{\partial \xi}{\partial \tau} = \frac{1}{P_a \mu_2 - \mu_1} e^{-\frac{\tau - \sigma_w^2}{P_a \mu_2}} - \frac{P_a \mu_2}{\mu_1 (P_a \mu_2 - \mu_1)} e^{-\frac{\tau - \sigma_w^2}{\mu_1}}, \tag{14}$$

Let $\frac{\partial \xi}{\partial \tau} = 0$, then the optimal detection threshold τ^* can be given by

$$\tau^* = \frac{P_a \mu_1 \mu_2}{P_a \mu_2 - \mu_1} \ln \frac{P_a \mu_2}{\mu_1} + \sigma_w^2. \tag{15}$$

Then, we can derive that ξ decreases with τ when $\tau \in (0, \tau^*)$, and ξ increases with τ when $\tau \in (\tau^*, \infty)$. By substituting τ^* in (15) into (12), the minimum detection error probability ξ^* can be determined as

$$\begin{aligned}\xi^* &= 1 - \frac{P_a \mu_2}{P_a \mu_2 - \mu_1} \left(e^{-\frac{\tau^* - \sigma_w^2}{P_a \mu_2}} - e^{-\frac{\tau^* - \sigma_w^2}{\mu_1}} \right) \\ &= 1 - \frac{P_a \mu_2}{P_a \mu_2 - \mu_1} \left(e^{-\frac{\mu_1}{P_a \mu_2 - \mu_1} \ln \frac{P_a \mu_2}{\mu_1}} - e^{-\frac{P_a \mu_2}{P_a \mu_2 - \mu_1} \ln \frac{P_a \mu_2}{\mu_1}} \right).\end{aligned} \tag{16}$$

■

B. Covert Rate

To measure the covert performance of the system, we calculate the achievable rate from Alice to Bob based on the instantaneous information. We then formulate the optimization problem of ZF beamforming on jammer to maximize the achievable covert rate at Bob. It is important to note that in this model, we assume that the location of Willie is known.

According to (1), the achievable rate R_b from Alice to Bob can be given as [28]

$$R_b = \log_2 \det \left(\mathbf{I} + \frac{P_a \mathbf{G} \mathbf{H}_{ab}^H \mathbf{H}_{ab} \mathbf{G}^H}{\mathbf{F} \mathbf{H}_{jb}^H \mathbf{H}_{jb} \mathbf{F}^H + \sigma_{ab}^2 \mathbf{I}} \right). \tag{17}$$

Then, the optimization problem for maximizing the covert rate R_b from Alice to Bob, subjecting to covertness, the upper bound of transmit power at Alice and jammer, and beamforming constraints, can be mathematically determined

by

$$\max_{\mathbf{F}} \quad R_b, \tag{18a}$$
$$s.t. \quad \xi^* \geq 1-\epsilon, \tag{18b}$$
$$P_a \leq P_a^{max}, \tag{18c}$$
$$||\mathbf{F}||^2 \leq P_j^{max}, \tag{18d}$$
$$\mathrm{Tr}(\mathbf{F}\mathbf{H}_{jb}^H\mathbf{H}_{jb}\mathbf{F}^H) = 0, \tag{18e}$$

where $\mathrm{Tr}(\cdot)$ is the trace operator. Constraint (18b) is the covertness constraint, which ensures the transmission between Alice and Bob is covert for the worst-case of the system. Constraint (18c) and (18d) represent the upper bound of transmit power, which limits the maximum transmit power of Alice and the jammer, respectively. Constraint (18e) corresponds to the ZF beamforming of jammer, ensuring null space beamforming for Bob. In this paper, the null space means that Bob can eliminate the interference by the jamming [29], mathematically, the beamforming vector $\mathbf{F}$ at jammer has to be orthogonal to the channel matrix $\mathbf{H}_{jb}$ from jammer to Bob, i.e., $\mid \mathbf{F}^H\mathbf{H}_{jb} \mid^2 = 0$. As shown in [30], $\mid \mathbf{F}^H\mathbf{H}_{jb} \mid^2 = 0$ is equivalent to the constraint (18e).

Notice that the proposed optimization problem (18) is non-convex due to the non-convexity of the objective function (18a) and the constraint (18b). Hence, we can not solve this problem using the existing convex optimization methods. To address this problem, we transform the non-convexity (18a) and (18b) into a series of convex problems. We first focus on the objective function (18a). By substituting (17) and (18e) into (18a), the objective function (18a) can be rewritten as

$$\max_{\mathbf{F}} \quad \log_2\left(1+\frac{P_a\sigma_{ab}^2\mathrm{Tr}(\mathbf{H}_{ab})}{\sigma_{ab}^2}\right). \tag{19}$$

We then transform non-convexity constraints into convex constraints. Considering the worst-case of the system, we assume Willie can make the decision with the optimal detection threshold. By substituting (16) into (18b), the covertness constraint (18b) can be rewritten as

$$\mu_1 \ln\left(\frac{\mu_1}{P_a\mu_2}\right) - \ln\epsilon(P_a\mu_2-\mu_1) \leq 0. \tag{20}$$

Without loss of generality, following the assumption in [30], we consider the case where the antennas are separated. Consequently, the covariance matrix $\mathbf{C}_{jw}$ and $\mathbf{C}_{aw}$ are assumed to be diagonal. Based on this assumption, we can rewrite $\mathbf{F}^H\mathbf{C}_{jw}\mathbf{F}$ as

$$\begin{aligned}\mathbf{F}^H\mathbf{C}_{jw}\mathbf{F} &= \mathrm{Tr}(\mathbf{F}^H\mathbf{C}_{jw}\mathbf{F}) = \mathrm{Tr}(\mathbf{F}(\mathbf{F}^H\mathbf{C}_{jw}) = \sigma_{jw}^2\mathrm{Tr}(\mathbf{F}\mathbf{F}^H\mathbf{I}) \\ &= \sigma_{jw}^2\mathrm{Tr}(\mathbf{F}\mathbf{F}^H) = \sigma_{jw}^2\mathrm{Tr}(\mathbf{W}_j),\end{aligned} \tag{21}$$

where $\mathbf{W}_j = \mathbf{F}\mathbf{F}^H$. Similar to (21), $\mathbf{G}^H\mathbf{C}_{aw}\mathbf{G}$ can be given by

$$\mathbf{G}^H\mathbf{C}_{aw}\mathbf{G} = \sigma_{aw}^2\mathrm{Tr}(\mathbf{W}_a), \tag{22}$$

where $\mathbf{W}_a = \mathbf{G}\mathbf{G}^H$.

Then, by substituting (21) and (22) into (20), (18b) can be converted to linear constraints. Thus, the optimization problem can be rewritten as

$$\max_{\mathbf{F}} \quad \log_2\left(1+\frac{P_a\sigma_{ab}^2\mathrm{Tr}(\mathbf{H}_{ab})}{\sigma_{jb}^2\mathrm{Tr}(\mathbf{W}_j\mathbf{H}_{jb})+\sigma_{ab}^2}\right), \tag{23a}$$
$$s.t. \quad P_a\sigma_{aw}^2\mathrm{Tr}(\mathbf{W}_a) - \sigma_{jw}^2\mathrm{Tr}(\mathbf{W}_j) < \epsilon_1, \tag{23b}$$
$$\sigma_{jw}^2\mathrm{Tr}(\mathbf{W}_j)\ln\left(\frac{\sigma_{jw}^2\mathrm{Tr}(\mathbf{W}_j)}{P_a\sigma_{aw}^2\mathrm{Tr}(\mathbf{W}_a)}\right) - \epsilon_1\ln\epsilon \leq 0, \tag{23c}$$
$$P_a \leq P_a^{max}, \tag{23d}$$
$$\mathrm{Tr}(\mathbf{W}_j) \leq P_j^{max}, \tag{23e}$$
$$\mathrm{Tr}(\mathbf{W}_j\mathbf{H}_{jb}) = 0. \tag{23f}$$

Finally, (23) is a convex problem because the objective function and the constraints are all convex. We then adopt an iterative algorithm to find the optimal beamformer $\mathbf{W}_j^*$. The details of the algorithm are summarized in Algorithm 1. When we obtain the $\mathbf{W}_j^*$, we can derive the beamforming vector $\mathbf{F}$ at jammer by a Gaussian randomization procedure [31].

Algorithm 1 Iterative Algorithm for Optimal Beamforming Matrix

Input: Channel matrixes;
Output: Optimal beamforming matrix $\mathbf{W}_j$;
1: Initialize the transmit power P_a, the tolerance ϕ, the iteration index $v=0$, the maximum number of iterations v_{max};
2: **repeat**
3: Solving (17) to obtain the corresponding $\mathbf{W}_j$;
4: $v = v+1$;
5: $\mathbf{W}_j(v+1) = \mathbf{W}_j(v)$;
6: **until** $\mid R_b(v+1) - R_b(v) \mid \leq \phi$ or $v \geq v_{max}$;
7: $\mathbf{W}_j^* \leftarrow \mathbf{W}_j$;
8: **return** $\mathbf{W}_j^*$;

IV. Numerical Results and Discussions

In this section, we present the numerical results to validate the covert performance of the system and investigate the impacts of various system parameters under the covertness constraint. Unless otherwise stated, in the numerical results, the related parameters are set as $\sigma_{ab}^2 = \sigma_{aw}^2 = \sigma_{jw}^2 = 1W$.

To validate the theoretical results of the detection error probability in (12), we conduct a simulation utilizing Monte-Carlo. The duration of each simulation is set to be 10000 time slots. In addition we set the noise variance as 1W, $P_a = 2W$, $N_a = N_j = 2$. We count the number of the FA N_{FA} and MD N_{MD} events, respectively. Then the simulated detection error probability (SDEP) is calculated as

$$\mathrm{SDEP} = \frac{N_{FA}+N_{MD}}{10000} * 100\%. \tag{24}$$

Fig.2 shows that the theoretical results match the simulation results, indicating that our theoretical results can accurately capture the behavior of detection error probability.

Fig.2 also illustrates the impact of detection threshold τ on detection error probability ξ with the setting of $P_a = 2W$. We can see from Fig.2 that as τ increases, ξ first decreases

and then increases. This is because when τ is small, the false alarm error dominates ξ, and we can see from (9) that $\mathcal{P}_{FA}$ is a decreasing function, so ξ decreases as τ increases. As τ increases further, the missed detection error then dominates ξ. As shown in (11), $\mathcal{P}_{MD}$ is a increasing function, so ξ increases. Additionally, the figure also demonstrates that there exists a minimum detection error probability corresponding to the worst-case scenario of the system.

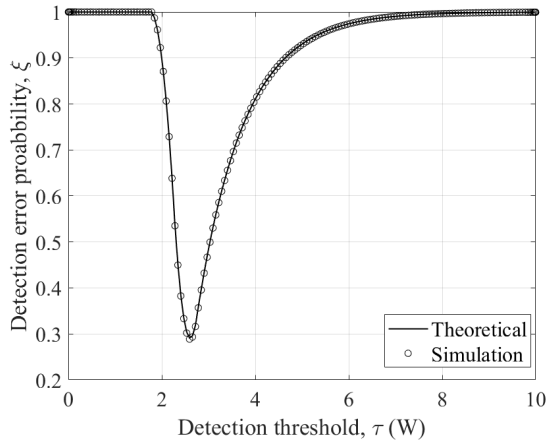

Fig. 2. Detection error probability ξ vs. Detection threshold τ.

We then investigate the impact of transmit power P_a on the detection error probability ξ, as shown in Fig. 3. The figure illustrates how the detection error probability ξ varies with the transmit power P_a for the different detection threshold settings of $\tau = \{3W, 5W, 7W\}$. It is observed that ξ monotonically decreases as P_a increases. This is because a higher transmit power results in a higher probability of detection by Willie. This observation emphasizes the characteristic of low transmit power in covert wireless communication.

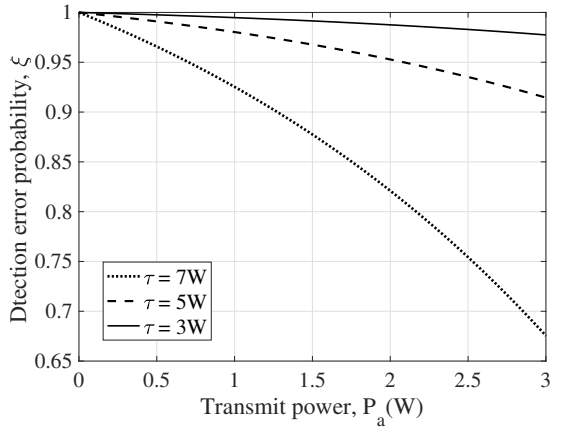

Fig. 3. Detection error probability ξ vs. Transmit power at A P_a.

To explore the impact of the covertness requirement ϵ on the achievable covert rate R_b of the system, we present in Fig.4 the variation of R_b with respect to ϵ for different settings of $(N_a, N_b) = \{(2,4),(4,4),(8,4)\}$ and $P^{max} = \{8W, 10W\}$. As depicted in the figure, we observe that as ϵ increases, the

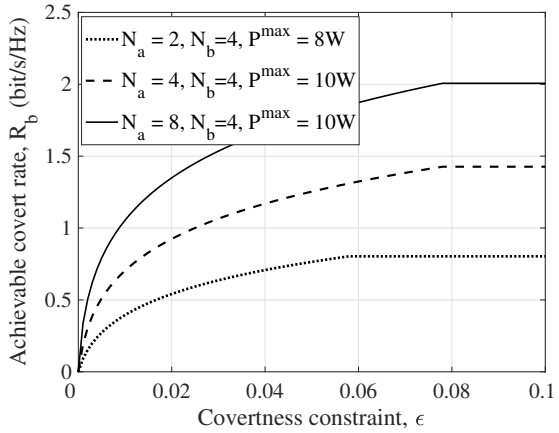

Fig. 4. Achievable covert rate R_b vs. Covertness constraint ϵ.

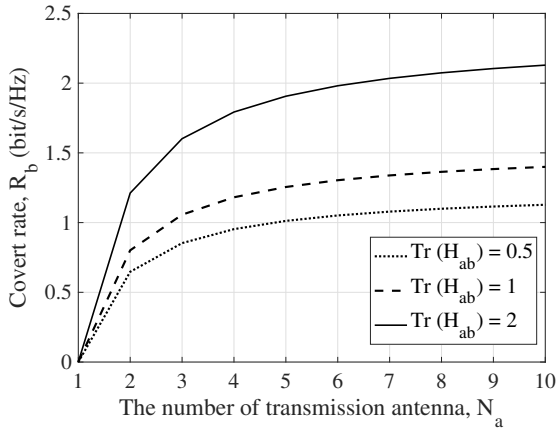

Fig. 5. Achievable covert rate R_b vs. The number of transmission antenna N_a.

covert rate R_b increases and then remains unchanged. This phenomenon can be attributed to the fact that as ϵ becomes larger, Alice can transmit messages using a larger power while keeping covertness, leading to a larger R_b. Thus, a less strict covertness constraint leads to a higher covert rate. Since the transmit power cannot be larger than the transmit power constraint, the R_b keeps unchanged when ϵ increases further. Furthermore, we note that for each fixed covertness constraint and transmit power constraint, R_b increases with the increases of the numbers of transceiver antennas (N_a, N_b). This is because the larger the number of antennas is, the higher spatial multiplexing gains are achieved. This phenomenon indicates that MIMO with beamforming technique has a significant impact on covert performance.

Last, we illustrate in Fig.5 now covert rate R_b varies with the number of transmission antenna N_a under the settings of $P_a = 1.5W$, $\epsilon = 0.01$ and $\text{Tr}(\mathbf{H}_{ab}) = \{0.5, 1, 2\}$. We can see from Fig.5 that as N_a increases, R_b increases. This is due to the reason that with more antennas, the receiver can exploit more spatial multiplexing gains. Since the bandwidth and transmit power constraint, the covert rate will remain unchanged even N_a increases. We can also see from Fig.5

that for a given setting of N_a, R_b increases as the trace of channel gain $\mathrm{Tr}(\mathbf{H}_{ab})$ from Alice to Bob increases. This is mainly due to the reason that a larger channel gain leads to a stronger received signal at the receiver, also a larger signal-to-interference-plus-noise ratio (SINR), resulting in a larger covert rate.

V. Conclusion

This paper investigates the covert performance in a MIMO system by employing a ZF beamforming jammer. The theoretical expressions for the minimum detection error probability are established, and an analysis is conducted to determine the maximum achievable covert rate under covertness, upper bound transmit power, and beamforming constraints imposed on the system. The numerical results reveal the validity of the designed beamforming matrix. The covert rate of the system shows an increasing trend as the number of antenna increases, which indicates the benefits of utilizing a large-scale MIMO configuration to enhance covert performance.

References

[1] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, Sept. 2016.

[2] A. Hamlin, N. Schear, E. Shen, M. Varia, S. Yakoubov, and A. Yerukhimovich, *Cryptography for big data security*. Taylor & Francis LLC, CRC Press, 2016.

[3] N. Kaaniche and M. Laurent, "Data security and privacy preservation in cloud storage environments based on cryptographic mechanisms," *Computer Communications*, vol. 111, pp. 120–141, Oct. 2017.

[4] Y. Cao, N. Zhao, Y. Chen, M. Jin, Z. Ding, Y. Li, and F. R. Yu, "Secure transmission via beamforming optimization for NOMA networks," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 193–199, Feb. 2019.

[5] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2637–2652, Nov. 2020.

[6] 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, Sept. 2013.

[7] G. Frèche, M. R. Bloch, and M. Barret, "Polar codes for covert communications over asynchronous discrete memoryless channels," *Entropy*, vol. 20, no. 1, p. 3, Dec. 2017.

[8] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: Hiding messages in noise," in *2013 IEEE International Symposium on Information Theory*. IEEE, Jul. 2013, pp. 2945–2949.

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

[10] L. Wang, "Covert communication over the poisson channel," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 23–31, Jan. 2021.

[11] Z. Liu, J. Liu, Y. Zeng, J. Ma, and Q. Huang, "On covert communication with interference uncertainty," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–6.

[12] Y. Wang, S. Yan, W. Yang, and Y. Cai, "Covert communications with constrained age of information," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 368–372, Feb. 2020.

[13] B. He, S. Yan, X. Zhou, and H. Jafarkhani, "Covert wireless communication with a poisson field of interferers," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6005–6017, Sept. 2018.

[14] J. Hu, S. Yan, X. Zhou, F. Shu, and J. Li, "Covert wireless communications with channel inversion power control in rayleigh fading," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12 135–12 149, Dec. 2019.

[15] L. Yang, W. Yang, S. Xu, L. Tang, and Z. He, "Achieving covert wireless communications using a full-duplex multi-antenna receiver," in *2019 IEEE 5th International Conference on Computer and Communications (ICCC)*. IEEE, Dec. 2019, pp. 912–916.

[16] K. Shahzad, X. Zhou, and S. Yan, "Covert communication in fading channels under channel uncertainty," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. IEEE, Jun. 2017, pp. 1–5.

[17] X. Zhou, S. Yan, J. Hu, J. Sun, J. Li, and F. Shu, "Joint optimization of a UAV's trajectory and transmit power for covert communications," *IEEE Transactions on Signal Processing*, vol. 67, no. 16, pp. 4276–4290, Aug. 2019.

[18] Y. Jiang, L. Wang, H. Zhao, and H.-H. Chen, "Covert communications in D2D underlaying cellular networks with power domain NOMA," *IEEE Systems Journal*, vol. 14, no. 3, pp. 3717–3728, Sept. 2020.

[19] 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, Oct. 2017, pp. 1–9.

[20] S.-Y. Wang and M. R. Bloch, "Covert MIMO communications under variational distance constraint," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4605–4620, Sept. 2021.

[21] W. Xiang, J. Wang, S. Xiao, and W. Tang, "Achieving constant rate covert communication via multiple antennas," in *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*. IEEE, Jun. 2022, pp. 1–6.

[22] O. Shmuel, A. Cohen, and O. Gurewitz, "Multi-antenna jamming in covert communication," *IEEE Transactions on Communications*, vol. 69, no. 7, pp. 4644–4658, Jul. 2021.

[23] M. Forouzesh, 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*, vol. 69, no. 8, pp. 8568–8576, Aug. 2020.

[24] C. Liu, G. Geraci, N. Yang, J. Yuan, and R. Malaney, "Beamforming for MIMO gaussian wiretap channels with imperfect channel state information," in *2013 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2013, pp. 3253–3258.

[25] J. Hu, K. Shahzad, S. Yan, X. Zhou, F. Shu, and J. Li, "Covert communications with a full-duplex receiver over wireless fading channels," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, May 2018, pp. 1–6.

[26] J. Bai, J. He, Y. Chen, Y. Shen, and X. Jiang, "On covert communication performance with outdated csi in wireless greedy relay systems," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 2920–2935, Aug. 2022.

[27] D. B. Cheikh, J.-M. Kelif, M. Coupechoux, and P. Godlewski, "Multicellular zero forcing precoding performance in rayleigh and shadow fading," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*. IEEE, May 2011, pp. 1–5.

[28] G. Tsoulos, *MIMO system technology for wireless communications*. CRC press, 2018.

[29] Y. Yang, C. Sun, H. Zhao, H. Long, and W. Wang, "Algorithms for secrecy guarantee with null space beamforming in two-way relay networks," *IEEE transactions on signal processing*, vol. 62, no. 8, pp. 2111–2126, Jan. 2014.

[30] S. Ma, Y. Zhang, H. Li, S. Lu, N. Al-Dhahir, S. Zhang, and S. Li, "Robust beamforming design for covert communications," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 3026–3038, Apr. 2021.

[31] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, May 2010.