Power Allocation and Time-Domain Artificial Noise Design for Wiretap OFDM with Discrete Inputs

Haohao Qin, Yin Sun, Member, IEEE, Tsung-Hui Chang, Member, IEEE, Xiang Chen, Member, IEEE, Chong-Yung Chi, Senior Member, IEEE, Ming Zhao, Member, IEEE, and Jing Wang, Member, IEEE

Abstract—Optimal power allocation for orthogonal frequency division multiplexing (OFDM) wiretap channels with Gaussian channel inputs has already been studied in some previous works from an information theoretical viewpoint. However, these results are not sufficient for practical system designs. One reason is that discrete channel inputs, such as quadrature amplitude modulation (QAM) signals, instead of Gaussian channel inputs, are deployed in current practical wireless systems to maintain moderate peak transmission power and receiver complexity. In this paper, we investigate the power allocation and artificial noise design for OFDM wiretap channels with discrete channel inputs. We first prove that the secrecy rate function for discrete channel inputs is nonconcave with respect to the transmission power. To resolve the corresponding nonconvex secrecy rate maximization problem, we develop a low-complexity power allocation algorithm which yields a duality gap diminishing in the order of $O(1/\sqrt{N})$, where $N$ is the number of subcarriers of OFDM. We then show that independent frequency-domain artificial noise cannot improve the secrecy rate of single-antenna wiretap channels. Towards this end, we propose a novel time-domain artificial noise design which exploits temporal degrees of freedom provided by the cyclic prefix of OFDM systems to jam the eavesdropper and boost the secrecy rate even with a single antenna at the transmitter. Numerical results are provided to illustrate the performance of the proposed design schemes.

Index Terms—Artificial noise, wiretap OFDM, power allocation, discrete channel inputs, secrecy rate.

I. INTRODUCTION

Security has become increasingly important for wireless networks due to the proliferation of privacy-sensitive wireless services. Traditionally, wireless information security is handled by cryptographic protocols in media access control (MAC) and higher layers [1]. However, these techniques face severe challenges due to the rapid developments of encryption breaking algorithms and super-computers [2]. Recently, various physical-layer techniques have been proposed to realize perfect secrecy in wireless networks [3], [4].

The fundamentals of physical-layer security techniques were laid in [5]–[7]. These works studied the maximum data rate for secrecy communications, i.e., the secrecy capacity, for a wiretap channel in which an eavesdropper (Eve) intends to wiretap the secrecy communications from a transmitter (Alice) to a legitimate receiver (Bob). Recently, various techniques, such as power allocation, beamforming and training schemes, have been developed to maximize the secrecy capacity of wiretap channels, e.g., [8]–[20]. One effective technique is initiatively transmitting artificial noise to jam the eavesdropper if the transmitter is equipped with multiple antennas, e.g., [15]–[19]. Besides, advanced techniques based on secrecy-key agreement were also studied in [21]–[25] to improve the security of wireless networks.

A common assumption of these studies is that the transmitted signal has a Gaussian distribution. However, Gaussian inputs are hardly used in practice due to its infinite peak power and its excessive detection complexities. Instead, discrete inputs such as Phase Shift Keying (PSK) and Quadrature Amplitude Modulation (QAM) (see Fig. 1(a)) are prevalent in practical digital communication systems [26]. Furthermore, most existing artificial noise designs rely heavily on the spatial degrees of freedom provided by multiple transmit antennas, which are not available in single-antenna wiretap channels.

In this paper, we consider an OFDM wiretap channel with discrete channel inputs, where each node is equipped with a single antenna and Alice has perfect knowledge of the channel state information (CSI) for the wireless links to Bob and Eve. We intend to answer the following questions: What are the differences between the secrecy rate functions corresponding to Gaussian and discrete channel inputs? Do these differences introduce additional difficulty in solving the power allocation problem of the OFDM wiretap channel with discrete channel inputs? Can we make use of artificial noise to improve the secrecy rate of an OFDM wiretap channel when the transmitter is equipped with only one antenna? To address these questions, we first study the convexity of the secrecy rate function with discrete channel inputs, and then develop power allocation and time-domain artificial noise designs to maximize the secrecy rate of OFDM wiretap channels. The main contributions of this paper are summarized as follows:
We prove that the secrecy rate of any discrete channel inputs with a finite number of possible values (or more generally with a finite entropy) is a nonconcave function with respect to the transmit power (Proposition 1). This is in contrast to the case of Gaussian channel inputs, where the associated secrecy rate is concave and the optimal power allocation has a closed-form solution [8]–[10].

A low-complexity power allocation algorithm based on Lagrange dual optimization is then proposed for discrete channel inputs. We show that the duality gap of the proposed algorithm diminishes asymptotically in the order of $O(1/\sqrt{N})$ as $N$ increases, where $N$ is the number of subcarriers of OFDM (Proposition 2).

We show that simply inserting independent artificial noise in the frequency domain cannot improve the secrecy rate of single-antenna wiretap channels (Proposition 3). To resolve this problem, we propose a time-domain artificial noise design for the considered single-antenna OFDM wiretap channel which exploits temporal degrees of freedom provided by the cyclic prefix of OFDM systems to jam the eavesdropper. To the best of our knowledge, this is the first time-domain artificial noise design for OFDM wiretap channels.

Finally, we jointly optimize the subcarrier power allocation and the covariance matrix of the time-domain artificial noise to improve the secrecy rate. Successive convex approximation methods are proposed to handle the joint design problem efficiently. Numerical results are presented to show that the proposed artificial noise schemes can considerably boost the secrecy rate.

There are several related works published recently. For example, linear precoding was studied for multiple-input power allocation algorithm without using artificial noise is [27], [28], where the solution is locally optimal. In [29], the OFDM wiretap channel was treated as a special instance of the MIMO wiretap channel and its achievable secrecy rates were studied under both Gaussian inputs and rectangular QAM constellations through asymptotic high/low SNR analysis and numerical evaluations. In contrast, we consider a broader discrete channel input setting along with more general analytical results in Propositions 1–3. Artificial noise design was studied in [30] for single antenna wiretap channel with discrete inputs (without OFDM), by assuming an AWGN channel to the Bob and a fast fading channel to the Eve. Our work considers quasi-static fading channels to both Bob and Eve, and respectively studies frequency domain and time domain artificial noise designs.

The remainder of this paper is organized as follows. In Section II, we present the system model and the power allocation problem, and then prove the nonconcavity of the secrecy rate under discrete channel inputs. In Section III, a power allocation algorithm without using artificial noise is presented for handling the secrecy rate maximization problem. Section IV presents the time-domain artificial noise design and two artificial noise aided power allocation algorithms (one for discrete inputs and the other for Gaussian inputs). Finally, some conclusions are drawn in Section V.

Notation: $\mathbb{C}$, $\mathbb{C}^n$ and $\mathbb{C}^{m \times n}$ denote the set of complex numbers, $n$-vectors, $m \times n$ matrices, respectively. Bold upper-case letters denote matrices and bold lowercase letters denote column vectors. $\mathbf{I}_N$ denotes an $N \times N$ identity matrix. $\mathbf{A} \succeq \mathbf{0}$ denotes that the matrix $\mathbf{A}$ is a positive semi-definite matrix. $\text{tr}(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$. $\mathbf{p} \succeq \mathbf{0}$ means that each component of vector $\mathbf{p}$ is nonnegative. $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes that $\mathbf{x}$ is a complex Gaussian random vector with zero mean and covariance matrix $\mathbf{\Sigma}$. $\mathbb{E}[\mathbf{x}]$ represents the expectation of the random variable $\mathbf{x}$, and $\mathbb{E}[\mathbf{x}|\mathbf{y}]$ denotes the conditional expectation of $\mathbf{x}$ given $\mathbf{y}$. Function $\mathcal{H}(\mathbf{x})$ denotes the entropy of random variable $\mathbf{x}$, and $\mathcal{I}(\mathbf{x};\mathbf{y})$ denotes the mutual information between random variables $\mathbf{x}$ and $\mathbf{y}$. $\text{Diag}(x_1, x_2, \ldots, x_N)$ denotes the $N \times N$ diagonal matrix whose diagonal elements are $x_1, x_2, \ldots, x_N$. $[\mathbf{x}]^\dagger \triangleq \max\{x, 0\}$, and $f'(x_0)$ denotes the first derivative of $f(x)$ at the point $x_0$.

II. SYSTEM MODEL AND POWER ALLOCATION PROBLEM

Consider a single-antenna OFDM wiretap channel with $N$ subcarriers. Let $H_i \in \mathbb{C}$ and $G_i \in \mathbb{C}$ represent the complex channel coefficients of the $i$th subcarrier from the transmitter to the legitimate receiver and to the eavesdropper, respectively (see Fig. 1(b)). The received signals of the legitimate receiver and the eavesdropper can be expressed as

$$y_i = H_i \sqrt{p_i} s_i + w_i, \quad i = 1, \ldots, N, \quad (1a)$$
$$z_i = G_i \sqrt{p_i} s_i + v_i, \quad i = 1, \ldots, N, \quad (1b)$$

respectively, where $s_i \in \mathbb{C}$ is the normalized channel input signal with zero mean and unity variance; $p_i \geq 0$ denotes the power of the $i$th subcarrier; $w_i \in \mathbb{C}$ and $v_i \in \mathbb{C}$ are independent zero-mean circularly symmetric complex Gaussian noises with unity variance at the legitimate receiver and the eavesdropper, respectively. The channel coefficient $H_i$ is known to the transmitter and the legitimate receiver, and $G_i$ is known to the transmitter and the eavesdropper, which are satisfied if all the three nodes are within the same wireless system. The channel inputs $\{s_i\}$ are statistically independent and identically distributed (i.i.d.) Gaussian signals or some practical discrete signals, e.g., QAM (see Fig. 1(a)).

The secrecy rate associated with the signal model in (1) can be shown to be [8]

$$R_s(\mathbf{p}) = \frac{1}{N} \sum_{i=1}^{N} \mathcal{I}(s_i; H_i \sqrt{p_i} s_i + w_i) - \mathcal{I}(s_i; G_i \sqrt{p_i} s_i + v_i),$$

(2)

where $\mathbf{p} = [p_1, p_2, \ldots, p_N]^T$ contains all the subcarrier power variables.

Let us consider the following secrecy rate maximization problem:

$$R^* = \max_{\mathbf{p} \succeq \mathbf{0}} R_s(\mathbf{p}) \quad \text{(3a)}$$
$$s.t. \quad \frac{1}{N} \sum_{i=1}^{N} p_i \leq P, \quad \text{(3b)}$$

where $P$ in (3b) is the allowed maximal transmit power, and $R^*$ denotes the maximum achievable secrecy rate.
Proposition 1 Consider the following secrecy rate function:

\[ R(p) = \frac{1}{N} \sum_{i=1}^{N} \left[ \log_2 (1 + |H_i|^2 p_i) - \log_2 (1 + |G_i|^2 p_i) \right], \]  

where \( s \) has zero mean and unity variance, and \( w \) and \( v \) are circularly symmetric complex Gaussian random variables with zero mean and unity variance. Suppose that \( s \) has a finite entropy, i.e., \( \mathcal{H}(s) < \infty \). Then, if \( |H| > |G|, R(p) \geq 0 \) and \( R(p) \) is nonconcave in \( p \); otherwise, \( R(p) = 0 \) for all \( p \geq 0 \).

**Proof:** We first show that \( R(0) = \lim_{p \to \infty} R(p) = 0 \). When \( p = 0 \), one can easily show that \( I(s; w) = I(s; v) = 0 \), because \( s \) is statistically independent of \( w \) and \( v \). Thus, \( R(0) = 0 \). Since \( s \) has a finite entropy, it must be a discrete random variable. According to Lemma 6 of [34], we have

\[ \lim_{p \to \infty} I(s; H\sqrt{p}s + w) = \lim_{p \to \infty} I(s; G\sqrt{p}s + v) = \mathcal{H}(s) < \infty. \]  

Therefore,

\[ \lim_{p \to \infty} R(p) = 0. \]  

Next, let us show that, when \( |H| > |G| \), there exists a \( \hat{p} \in (0, \infty) \) such that \( R(\hat{p}) > 0 \). According to [34], the gradient of \( I(s; H\sqrt{p}s + w) \) is given by

\[ \frac{\partial I(s; H\sqrt{p}s + w)}{\partial p} = |H|^2 \text{mmse} (|H|^2 p) \geq 0, \]  

where

\[ \text{mmse}(|H|^2 p) \triangleq E \left[ |s - E(s)H\sqrt{p}s + w|^2 \right]. \]  

is the minimum mean square error (MMSE) of estimating \( s \) with the received signal \( y = H\sqrt{p}s + w \). When \( p \) equals zero, \( \text{mmse}(|H|^2 p) \) in (10) attains its maximum value, i.e., \( \text{mmse}(0) = E |s|^2 = 1 \). Thus, by (6) and (9), it can be seen that \( R(0) = |H|^2 - |G|^2 > 0 \), which implies that there must exist a positive \( \hat{p} \) such that \( R(\hat{p}) > 0 \).

Since \( R(p) \) is continuous and differentiable [34], by the Lagrange’s mean value theorem [35] and the fact that \( R(\hat{p}) > \lim_{p \to \infty} R(p) = 0 \), it can be inferred that there must exist a point \( \hat{p} \in [\hat{p}, \infty) \) satisfying \( R(\hat{p}) < \infty \) and \( R'(\hat{p}) < 0 \).

Now suppose that \( R(p) \) is a concave function on \( p \in [0, \infty) \). Then the following inequality

\[ R(p) \leq R(\hat{p}) + R'(\hat{p})(p - \hat{p}) \]  

must hold for any \( p \in [0, \infty) \). By letting \( p \to \infty \) in (11), we obtain \( \lim_{p \to \infty} R(p) = -\infty \) since \( R'(\hat{p}) < 0 \), which however contradicts the fact of \( \lim_{p \to \infty} R(p) = 0 \). Therefore, the concavity assumption for \( R(p) \) is not true.

When \( |H| \leq |G|, I(s; H\sqrt{p}s + w) \leq I(s; G\sqrt{p}s + v) \) for any \( p \geq 0 \). Hence, in this case, \( R(p) = 0 \) for all \( p \geq 0 \). Proposition 1 is thus proved.

Proposition 1 implies that finding an optimal power allocation for maximizing \( R_s(p) \) is a non-trivial design problem, as will be discussed shortly. By Proposition 1, the secrecy rate function is not a concave function for any channel input distribution with a finite entropy. Basically two types of inputs have finite entropy, including discrete distributions with a finite number of distinct values and some of the discrete distributions with a countable number of distinct values\(^1\). On the other hand, distributions with an infinite entropy basically include all

\(^1\)A discrete distribution with a countable number of values may either have a finite entropy or an infinite entropy [36, page 48].
Fig. 2: (a) Mutual information and (b) secrecy rate for a Gaussian channel input. (c) Mutual information and (d) secrecy rate for a QPSK channel input. Here $|H| = 1$ and $|G| = 0.5$.

Let us provide some numerical results to illustrate Proposition 1. Figure 2 shows the mutual information $I(s; H\sqrt{ps} + w)$, $I(s; G\sqrt{ps} + v)$, and the secrecy rate $R(p)$ for Gaussian and QPSK channel inputs. One can observe from Fig. 2(a) and Fig. 2(b) that $I(s; H\sqrt{ps} + w)$, $I(s; G\sqrt{ps} + v)$ and $R(p)$ are all concave when the inputs are Gaussian. From Fig. 2(c), $I(s; H\sqrt{ps} + w)$ and $I(s; G\sqrt{ps} + v)$ are also concave for QPSK channel inputs; however, the associated $R(p)$ is obviously not concave, as shown in Fig. 2(d).

We further claim that the secrecy rate $R(p)$ is not even an quasi-concave function for certain channel input distributions, by showing a counter example that the secrecy rate may have multiple peaks. Figure 3 further illustrates the secrecy rate $R(p)$ in the logarithmic scale of $p$ for a non-standard 4-PAM channel input $s$, where the probability mass function of $s$ is given by

$$P_s \sim \begin{bmatrix}
-51q & -50q & 50q & 51q \\
0.25 & 0.25 & 0.25 & 0.25
\end{bmatrix},$$

(12)

where $q$ is a normalization parameter used to maintain the unity variance of $s$. It is interesting to see from Fig. 3 that the secrecy rate $R(p)$ has two peaks. Hence, $R(p)$ is neither concave nor quasi-concave. When $p < -10$dB, both Bob and Eve cannot identify the constellation points. Since the constellation of $s$ in (12) has two groups (i.e., the group of $\{-51q, -50q\}$ and the group $\{51q, 50q\}$), Bob and Eve start to identify the two groups as $p$ increases. Since $|H| > |G|$, $I(s; H\sqrt{ps} + w)$ is larger than $I(s; G\sqrt{ps} + v)$, and so $R(p) > 0$. When $p \approx 18$dB, both Bob and Eve can identify the two groups of $s$, and $I(s; H\sqrt{ps} + w) \approx I(s; G\sqrt{ps} + v) \approx 1$ bits/s/Hz, and so $R(p)$ decreases to nearly 0. When $p \geq 20$dB, Bob and Eve start to identify each constellation point. Since Bob has better channel quality, $R(p)$ increases with $p$ again. This example also shows that the conjecture in [32], which claims the secrecy rate under discrete finite constellations has a single maximum, is not true for some discrete channel inputs.

III. PROPOSED POWER ALLOCATION ALGORITHM

We now present a computationally efficient Lagrange dual optimization algorithm to handle the nonconvex problem (3). We will show that the proposed algorithm yields a power allocation solution for which the duality gap decreases with $N$ in the order of $O(1/\sqrt{N})$, provided that the channel has a finite delay spread. This mild condition is satisfied in practical OFDM systems.

A. Asymptotic Optimal Power Allocation by Dual Optimization

The Lagrangian of problem (3) is given by

$$L(p, \lambda) = \frac{1}{N} \sum_{i=1}^{N} [I(s_i; H_i\sqrt{ps_i} + w_i) - I(s_i; G_i\sqrt{ps_i} + v_i)] +$$

$$+ \lambda \left( P - \frac{1}{N} \sum_{i=1}^{N} p_i \right),$$

(13)

where $\lambda \geq 0$ is the dual variable associated with the constraint (3b). The dual problem is given by

$$D^* = \min_{\lambda \geq 0} d(\lambda),$$

(14)
TABLE I

| Algorithm 1: Proposed power allocation scheme for discrete inputs. |
|---|
| **Given** |
| $\lambda_0 \geq \lambda_i = 0$, convergence tolerance $\varepsilon$ |
| **repeat** |
| step 1: update $\lambda = \frac{1}{2}(\lambda_i + \lambda_h)$ |
| step 2: obtain $(p_i)_{i=1}^{N}$ by solving problem (17) |
| step 3: if $\frac{1}{N} \sum_{i=1}^{N} p_i < P$, then update $\lambda_h = \lambda$, else update $\lambda_l = \lambda$ |
| **until** |
| $\lambda_h - \lambda_l < \varepsilon$ |
| **output** |
| output $\lambda^* = \lambda_h$. |

where $D^*$ denotes the optimal dual objective value, and $d(\lambda)$ is the dual function given by

$$d(\lambda) = \max_{p \geq 0} L(p, \lambda).$$

(15)

The Lagrange dual method first solves problem (15) for a given dual variable $\lambda$. According to (13), problem (15) can be decomposed into $N$ separate subproblems, i.e.,

$$d(\lambda) = \frac{1}{N} \sum_{i=1}^{N} B(p_i; \lambda, H_i, G_i) + \lambda P,$$

(16)

where

$$B(p_i; \lambda, H_i, G_i) = \max_{p_i \geq 0} \left[ I(s_i; H_i \sqrt{p_i s_i + w_i}) - I(s_i; G_i \sqrt{p_i s_i + v_i}) \right] - \lambda p_i$$

(17)

is a one-dimensional power allocation subproblem for subcarrier $i$, which can be efficiently solved by simple line search [37].

Notice that $d(\lambda)$ is a convex function [37] and its subgradient can be easily seen, from (13) and (16), to be

$$\nabla d(\lambda) = P - \frac{1}{N} \sum_{i=1}^{N} p_i^*,$$

where $p_i^*$ is the optimal solution to (17). Therefore, the dual variable $\lambda$ can be efficiently updated using the bisection method [37]. If $P - \frac{1}{N} \sum_{i=1}^{N} p_i^* > 0$, then the subgradient $\nabla d(\lambda) > 0$, and thus we decrease $\lambda$ in the bisection method for finding $D^*$ given by (14); otherwise we increase $\lambda$. The resulting power allocation algorithm, called Algorithm 1, for finding the desired $p^*$ for problem (3) is summarized in Table I. It is important to note from (8) that the secrecy rate is not an increasing function of the transmit power. Therefore, when the total power $P$ is large enough, the optimal total transmit power $\frac{1}{N} \sum_{i=1}^{N} p_i^*$ by solving (17) can be strictly lower than the available transmit power $P$, and the optimal dual variable is $\lambda^* = 0$ by the Karush-Kuhn-Tucker (KKT) conditions [37]. It is worthwhile to mention that Algorithm 1 can also be applied to the case of Gaussian inputs, where the per-subcarrier power allocation subproblem (17) has a closed-form solution exactly the same as (5) [8]–[10].

Since the primal problem (3) is nonconvex, there is a duality gap between the optimal $R^*$ and optimal $D^*$, i.e., $D^* - R^* > 0$, where $R^*$ and $D^*$ are defined in (3) and (14), respectively. Under a Lipschitz continuity assumption on the channel coefficients [38], it can be shown that the duality gap between $D^*$ and $R^*$ diminishes with $N$ in the order of $O(1/\sqrt{N})$. Prior to presenting such result, some quantities used in the discrete frequency domain and the corresponding quantities in the continuous frequency domain need to be reviewed. Let $H(f)$ and $G(f)$ denote the frequency domain channel responses of the legitimate receiver and the eavesdropper, respectively. Owing to uniform sampling in Discrete Fourier transform (DFT) over the normalized frequency interval $0 \leq f \leq 1$, we have

$$H(i/N) = H_i, \quad G(i/N) = G_i, \quad 1 \leq i \leq N.$$  (18)

In general, the time-domain channel has a finite delay spread (i.e., it is nonzero only on a finite time interval), leading to the fact that any order derivatives of its frequency response exist and are bounded [39, pp. 94–96]. In particular, the first derivatives of $H(f)$ and $G(f)$ exist and are bounded. Therefore, there exist some $L_H, L_G > 0$ such that the following Lipschitz continuous conditions hold for all $f, f' \in [0, 1]$:

$$|H(f) - H(f')| \leq L_H |f - f'|, \quad |G(f) - G(f')| \leq L_G |f - f'|.$$  (19)

We can show the following proposition:

**Proposition 2** Suppose that the channel coefficients $H(f)$ and $G(f)$ are Lipschitz continuous satisfying (19). Then

$$0 \leq D^* - R^* \leq O \left( \frac{1}{\sqrt{N}} \right),$$  (20)

where $R^*$ and $D^*$ are defined in problem (3) and problem (14), respectively.

**Proof:** In accordance with [38, Theorem 2], for proving (20) it is sufficient to show that there exists a constant $L_R > 0$ such that the difference between the secrecy rates on any two frequencies, $R_s(f, p)$ and $R_s(f', p')$, is bounded, i.e.,

$$|R_s(f, p) - R_s(f', p')| \leq L_R (|f - f'| + |p - p'|),$$  (21)

for any $f, f' \in [0, 1]$ and $p, p' \geq 0$, namely, that the secrecy rate $R_s(f, p)$ is Lipschitz continuous. The proof of (21) is presented in Appendix I.

**B. Numerical Results**

We now provide some numerical results to illustrate the efficacy of the proposed power allocation scheme (Algorithm 1 in Table I). Suppose that the OFDM system has $N = 64$ subcarriers. The length of the cyclic prefix (CP) is $N_{cp} = 16$ channel samples. Each channel realization is composed of 8 i.i.d. Rayleigh fading paths with the maximum time delay of 7 samples and the channel power normalized to unity. The numerical results are obtained by averaging over 300 channel realizations for each signal-to-noise ratio (SNR) value, where $SNR = P$ (which is the ratio of the available transmit power to AWGN power $E(|w_i|^2) = E(|v_i|^2) = 1$) per subcarrier. The proposed power allocation algorithm is compared with two existing algorithms. The first one is the equal power allocation scheme, denoted by “equal PA”, which allocates the total power uniformly over all the subcarriers that satisfy $|H_i| > |G_i|$; the second one is the power allocation scheme for Gaussian channel inputs, which is given by (5), denoted by “PA of (5)”.
where \( \hat{a}_i \in \mathbb{C} \) is the artificial noise term added to the \( i \)th subcarrier. It is assumed that \( \hat{a}_i \) are statistically independent across the subcarriers and \( \hat{a}_i \sim \mathcal{CN}(0, \sigma^2_{a,i}) \), where \( \sigma^2_{a,i} \geq 0 \) is the artificial noise power for subcarrier \( i \). According to [8], the secrecy rate for the signal model (22) is given by

\[
R_{AN} = \frac{1}{N} \sum_{i=1}^{N} \left[ I(s_i; H_i(\sqrt{p_is_i + \hat{a}_i}) + w_i) - I(s_i; G_i(\sqrt{p_is_i + \hat{a}_i}) + v_i) \right].
\]  

(23)
In order to maximize $R_s^{AN}$ in (23), power parameters $\{p_i\}$ and $\{\sigma_{a,i}^2\}$ should be jointly optimized, which can be formulated as the following optimization problem:

$$\max_{\{p_i\}^N_{i=1},\{\sigma_{a,i}^2\}^N_{i=1}} R_s^{AN}$$

s.t. $\frac{1}{N} \sum_{i=1}^{N} (p_i + \sigma_{a,i}^2) \leq P$ \hspace{1cm} (24a)

$$\sigma_{a,i}^2 \geq 0, \ p_i \geq 0, \ i = 1, 2, ..., N. \hspace{1cm} (24b)$$

The optimal solution for artificial noise power parameters of (24) is given in the following proposition.

**Proposition 3** For any given distribution of $\{s_i\}$, the optimal artificial noise power $\sigma_{a,i}^2$ to problem (24) is equal to zero for all $i$.

For the case that $\{s_i\}$ are Gaussian, Proposition 3 can be simply proved by showing that, when $|H_i|^2 > |G_i|^2$, the secrecy rate $R_s^{AN}$ is strictly decreasing in $\sigma_{a,i}^2$ for any $p_i \geq 0$. The proof for the case of arbitrarily distributed $\{s_i\}$ is more complicated, because $R_s^{AN}$ may be non-monotonic in $\sigma_{a,i}^2$. The proof details for general distributions of $\{s_i\}$ are presented in Appendix II.

Proposition 3 implies that adding artificial noise in the frequency domain is not effective since it only degrades the achievable secrecy rate. Next let us turn to the design of artificial noise in the time domain which is able to boost the secrecy rate.

### B. Proposed Time-domain Artificial Noise Design

The proposed wiretap OFDM system with a time-domain artificial noise is illustrated in Fig. 6(b). As a standard OFDM system, at first the frequency domain signal $P^{1/2} s$ is transformed to the time domain by inverse fast Fourier transform (IFFT) and then the cyclic prefix (CP) is inserted. Then, an artificial noise term is added to this time-domain signal before transmission. The receiver will discard the CP and then transform the remaining signal to the frequency domain by FFT. All of these operations can be expressed by linear matrix operations [40].

Let $F$ and $F^H$ denote the $N \times N$ FFT and IFFT matrices, and let $N_{cp}$ denote the length of CP. The matrices for CP insertion and removal are represented by $T^{cp} = [\tilde{E}_{2N_{cp}} F]_T$ and $R^{cp} = [0_{N-1 \times N_{cp}}, I_N]$, respectively, where $\tilde{E}_{N_{cp}}$ contains the last $N_{cp}$ rows of the $N \times N$ identity matrix $I_N$. Let $[h(0), h(1), ..., h(L)]$ and $[g(0), g(1), ..., g(L)]$ represent the time-domain channel impulse responses from the transmitter to the legitimate receiver and the eavesdropper, respectively, where $L < N_{cp}$ is the maximum delay spread. Then, following the system block diagram in Fig. 6(b), the received signals of the legitimate receiver and the eavesdropper can be expressed as [40]

$$y = FR^{cp}H_0(T^{cp}F^HP^{1/2} s + a) + w = HD^{1/2} s + FR^{cp}H_0a + w,$$ \hspace{1cm} (25a)

$$z = FR^{cp}G_0(T^{cp}F^HP^{1/2} s + a) + v = GP^{1/2} s + FR^{cp}G_0a + v,$$ \hspace{1cm} (25b)

where $a \in \mathbb{C}^{N_{cp}}$ is a zero-mean complex Gaussian random vector, $H_0 \in \mathbb{C}^{(N+N_{cp}) \times (N+N_{cp})}$ is a Toeplitz channel matrix given by

$$H_0 = \begin{bmatrix}
  h(0) & 0 & 0 & \cdots & 0 \\
  \vdots & h(0) & 0 & \cdots & 0 \\
  h(L) & \cdots & \cdots & \cdots & 0 \\
  \vdots & \vdots & \ddots & \ddots & \vdots \\
  0 & \cdots & \cdots & h(L) & h(0)
\end{bmatrix},$$

and so is the matrix $G_0$, i.e., by replacing $h(l)$ with $g(l), l = 0, 1, ..., L$, in $H_0$. Moreover, in (25), $H = FR^{cp}H_0T^{cp}F^H = \text{Diag}(H_1, H_2, ..., H_N)$ and $G = FR^{cp}G_0T^{cp}F^H = \text{Diag}(G_1, G_2, ..., G_N)$, in which $H_i$ and $G_i$ are the frequency matrices.
responses of the legitimate receiver’s channel and the eavesdropper’s channel, respectively, and \(w \sim CN(0, I_N)\) and \(v \sim CN(0, I_N)\) are the corresponding noise vectors at the legitimate receiver and the eavesdropper, respectively.

The received signal vectors given in (25) can be equivalently written as

\[
y_i = H_i \sqrt{p_i} s_i + f_i^T R_{\text{cp}} H_0 a + w_i, \quad i = 1, 2, ..., N, \quad (26a)
\]

\[
z_i = G_i \sqrt{p_i} s_i + f_i^T R_{\text{cp}} G_0 a + v_i, \quad i = 1, 2, ..., N, \quad (26b)
\]

respectively, where \(f_i^T \in \mathbb{C}^N\) is the \(i\)th row of the FFT matrix \(F\). Note that this corresponds to a simplified model of secure design in MIMO secrecy networks with a cooperative jammer [41].

In order not to interfere with the legitimate receiver, the artificial noise is fully laid in the null space of the channel of the legitimate receiver. Specifically, let us write

\[
a = U d, \quad \text{(27)}
\]

where \(U\) is a semi-unitary matrix whose column vectors span the null space of \(R_{\text{cp}} H_0\), i.e.,

\[
R_{\text{cp}} H_0 U = 0, \quad U^H U = I_{N_c p}, \quad \text{(28)}
\]

and \(d \sim CN(0, \Sigma_d)\) in which \(\Sigma_d\) is the covariance matrix of the artificial noise vector \(d\) to be determined. As the dimension of \(R_{\text{cp}} H_0\) is \(N \times (N_c p)\), the dimension of \(U\) is \((N_c p) \times N_c p\) and the dimension of \(d\) is \(N_c p\).

Now, by (27) and (28), the received signal in (26) reduces to

\[
y_i = H_i \sqrt{p_i} s_i + w_i, \quad i = 1, 2, ..., N, \quad (29a)
\]

\[
z_i = G_i \sqrt{p_i} s_i + f_i^T R_{\text{cp}} G_0 U d + v_i, \quad i = 1, 2, ..., N. \quad (29b)
\]

In general, the receiver detects the information in a subcarrier-by-subcarrier manner, and the secrecy rate achieved by (29) is given by [8]

\[
\frac{1}{N} \sum_{i=1}^{N} \left[ \mathcal{I}(s_i; H_i \sqrt{p_i} s_i + w_i) - \mathcal{I}(s_i; \frac{G_i \sqrt{p_i} s_i}{\sqrt{b_i^H \Sigma_d b_i + 1}} + \hat{v}_i) \right]^+, \quad \text{(30)}
\]

where \(b_i^H \triangleq f_i^T R_{\text{cp}} C_0 U d\), and \(\hat{v}_i \sim CN(0, 1)\). The joint power allocation and artificial noise design problem can be formulated as

\[
\max_{\tilde{p}_i, \tilde{\Sigma}_d, \Sigma_d} \frac{1}{N} \sum_{i=1}^{N} \left[ \mathcal{I}(s_i; H_i \sqrt{p_i} s_i + w_i) - \mathcal{I}(s_i; \frac{G_i \sqrt{p_i} s_i}{\sqrt{b_i^H \tilde{\Sigma}_d b_i + 1}} + \hat{v}_i) \right]^+ \quad \text{(31a)}
\]

subject to

\[
\begin{align*}
\frac{1}{N} \left( \sum_{i=1}^{N} \tilde{p}_i + \text{tr}(\tilde{\Sigma}_d) \right) & \leq P \\
\tilde{\Sigma}_d & \succeq 0, \quad \tilde{p}_i \geq 0, \quad i = 1, \ldots, N. \quad \text{(31b)}
\end{align*}
\]

The Lagrange dual optimization to Problem (31) through Successive Convex Approximation

Problem (31) is nonconvex and difficult to handle. Again, as in Section III, we consider Lagrange dual optimization to solve problem (31). The Lagrange of problem (31) is given by (32) at the bottom of the page, where \(\lambda\) is the Lagrange dual variable associated with constraint (31b). The associated dual function is defined as

\[
d(\lambda) = \max_{\tilde{p}, \tilde{\Sigma}_d \succeq 0} L(\tilde{p}, \tilde{\Sigma}_d, \lambda). \quad \text{(33)}
\]

The bisection method used in Section III-A can also be applied to the following dual problem

\[
\min_{\lambda \geq 0} d(\lambda). \quad \text{(34)}
\]

However, solving problem (33) is still challenging since the Lagrangian is not concave in \((\tilde{p}, \tilde{\Sigma}_d)\). We use the coordinate descent method [42] to handle problem (33), that tries to maximize the Lagrangian by updating variable \(\tilde{p}\) and \(\tilde{\Sigma}_d\) in an alternating fashion, to be presented next.

1) Update of \(\tilde{p}\) with fixed \(\tilde{\Sigma}_d\): With \(\tilde{\Sigma}_d\) fixed, the optimal \(\tilde{p}\) to problem (33) can be obtained by solving the one-dimensional problems (35) given at the bottom of the page.

2) Update of \(\tilde{\Sigma}_d\) with fixed \(\tilde{p}\): Because the solution \(\tilde{p}^*\) given by (35) would yield nonnegative \(\{I(s_i; H_i \sqrt{p_i} s_i + w_i) - \mathcal{I}(s_i; \frac{G_i \sqrt{p_i} s_i}{\sqrt{b_i^H \Sigma_d b_i + 1}} + \hat{v}_i), \forall i\}\), with \(\tilde{p}\) fixed, problem (33) is equivalent to the following minimization problem:

\[
\min_{\Sigma_d \succeq 0} \sum_{i=1}^{N} \mathcal{I}(s_i; \frac{G_i}{\sqrt{b_i^H \Sigma_d b_i + 1}} \sqrt{p_i} s_i + \hat{v}_i) + \lambda \text{tr}(\Sigma_d). \quad \text{(36)}
\]

For ease of presentation, let us define

\[
T_i(\Sigma_d) \triangleq \mathcal{I}(s_i; \frac{G_i}{\sqrt{b_i^H \Sigma_d b_i + 1}} \sqrt{p_i} s_i + \hat{v}_i), \quad i = 1, 2, ..., N. \quad \text{(37)}
\]

Next, we apply a successive convex approximation method which guarantees to yield a stationary point of problem (36).

Consider the first-order approximation to \(T_i(\Sigma_d)\). Let \(t_i = (b_i^H \Sigma_d b_i + 1)^{-1}\). Then \(T_i(\Sigma_d)\) becomes a function of \(t_i\), i.e., \(T_i(t_i)\), and its first derivative with respect to \(t_i\) is given by

\[
T_i'(t_i) = |G_i|^2 p_i \text{mmse}(|G_i|^2 p_i t_i). \quad \text{(by (9))} \quad \text{(38)}
\]

Then it is readily to see that the first-order approximation of \(T_i(t_i)\) at the point \(\hat{t}_i = (b_i^H \tilde{\Sigma}_d b_i + 1)^{-1}\), where \(\tilde{\Sigma}_d\) is the one obtained in the previous iteration, is given by

\[
\hat{T}_i(t_i) = T_i(\hat{t}_i) + |G_i|^2 p_i \text{mmse}(|G_i|^2 p_i \hat{t}_i) (t_i - \hat{t}_i). \quad \text{(39)}
\]
TABLE II

| Algorithm: Algorithm for solving problem (36) |
|---|
| **Given**: \( \Sigma_d \) |
| **repeat**: |
| step 1: solve problem (43) by CVX and obtain optimal \( \Sigma_d \) |
| step 2: set \( \Sigma_f = \Sigma_d \) |
| **until**: a specified convergence criterion is satisfied. |

So we come up with the following first-order approximation to problem (36):

\[
\min_{\Sigma_d \geq 0, (t_i)_{i=1}^N} \sum_{i=1}^N T_i(t_i) + \lambda \text{tr}(\Sigma_d) \tag{40}
\]

\[\text{s.t.} \quad t_i = (b_i^H \Sigma_d b_i + 1)^{-1}, \quad i = 1, 2, ..., N.\]

Next, we show that the approximated problem (40) can be reformulated as a convex semi-definite program (SDP). Omitting all the constant terms, problem (40) can be equivalently formulated as

\[
\min_{\Sigma_d \geq 0, (t_i)_{i=1}^N} \sum_{i=1}^N |G_i|^2 p_i \left[ \text{mmse}(\frac{|G_i|^2 p_i}{b_i^H \Sigma_d b_i + 1}) \right] t_i + \lambda \text{tr}(\Sigma_d) \tag{41}
\]

As MMSE is nonnegative [43], problem (41) is equivalent to the following problem:

\[
\min_{\Sigma_d \geq 0, (t_i)_{i=1}^N} \sum_{i=1}^N |G_i|^2 p_i \left[ \text{mmse}(\frac{|G_i|^2 p_i}{b_i^H \Sigma_d b_i + 1}) \right] t_i + \lambda \text{tr}(\Sigma_d) \tag{42}
\]

\[\text{s.t.} \quad t_i \geq (b_i^H \Sigma_d b_i + 1)^{-1}, \quad i = 1, 2, ..., N.\]

Finally, by applying Shur complement [37], problem (42) can be recast as

\[
\min_{\Sigma_d \geq 0, (t_i)_{i=1}^N} \sum_{i=1}^N |G_i|^2 p_i \left[ \text{mmse}(\frac{|G_i|^2 p_i}{b_i^H \Sigma_d b_i + 1}) \right] t_i + \lambda \text{tr}(\Sigma_d) \tag{43}
\]

\[\text{s.t.} \quad \left[ b_i^H \Sigma_d b_i + 1 \begin{array}{c} 1 \\ 1 \end{array} \right] \succeq 0, \quad i = 1, 2, ..., N.\]

Problem (43) is a standard SDP, which can be solved by convex solvers such as CVX [44] or SeDuMi [45].

We now show the convergence of the proposed successive convex approximation algorithm. The first-order convex approximation of the function \( T_i(t_i) \), i.e., \( \tilde{T}_i(t_i) \) in (39), satisfies the following three relations:

- \( \tilde{T}_i(t_i) \geq T_i(t_i), \forall t_i, \) since \( T_i(t_i) \) is concave with respect to \( t_i \) [43];
- \( \tilde{T}_i(t_i) = T_i(t_i) \);
- \( \tilde{T}_i(t_i) = T_i(t_i) \).

Hence, according to [46], the proposed successive convex approximation algorithm will converge to a point satisfying KKT conditions of the original problem (36), thereby leading to a stationary local maximum of problem (36). This successive convex approximation method for solving problem (36), called SCA algorithm, is summarized in Table II, and the artificial noise approximation algorithm will converge to a point satisfying KKT conditions of the original problem (36), thereby leading to a feasible solution of problem (36) for this case also reduces to a simple convex optimization problem as follows:

\[
\text{Problem (45) for Gaussian inputs.}
\]

\[\text{Algorithm 3: Proposed artificial noise aided power allocation scheme for Gaussian inputs.} \]

\[\text{Given:} \lambda_h \geq \lambda_i = 0, \{p_i\}_{i=1}^N \]

\[\text{repeat:} \]

step 1: update \( \lambda = \frac{1}{2}(\lambda_i + \lambda_h) \)

step 2: repeat: by solving problem (45) using CVX and then obtain \( \{p_i\}_{i=1}^N \) by (44)

\[\text{until:} \lambda \text{meets a specified convergence criterion} \]

step 3: if \( \frac{1}{N} \left( \text{tr}(\Sigma_d) + \sum_{i=1}^N p_i \right) < P \), then update \( \lambda_h = \lambda \),

else update \( \lambda_i = \lambda \)

\[\text{until:} \lambda \text{meets a specified convergence criterion.} \]

D. Gaussian Channel Inputs

The proposed artificial noise aided power allocation algorithm can be significantly simplified when all \( s_i \) are Gaussian. In this case, the optimal solution \( p_i^* \) in (35) is simply given by

\[
p_i^* = \begin{cases} 
\frac{1}{2} |H_i|^2 |G_i|^2 \left( \sqrt{C^2 - 4 |H_i|^2 |G_i|^2} + \frac{\lambda + |G_i|^2 + |H_i|^2 - C}{\lambda} \right), & \text{if } |H_i|^2 - |G_i|^2 > \lambda \\
0, & \text{otherwise,}
\end{cases}
\]

\[\text{where } \hat{G}_i = \frac{-G_i}{\sqrt{b_i^H \Sigma_d b_i + 1}} \text{ and } \hat{C} = |H_i|^2 + |G_i|^2.\]

Problem (36) for this case also reduces to a simple convex optimization problem as follows:

\[
\Sigma_d^* = \arg \min_{\Sigma_d \geq 0} \sum_{i=1}^N \log_2 \left( 1 + \frac{|G_i|^2 p_i}{b_i^H \Sigma_d b_i + 1} \right) + \lambda \text{tr}(\Sigma_d). \tag{45}
\]

The resulting algorithm, Algorithm 3, for Gaussian inputs is given in Table IV.

E. Numerical Results

We now show some numerical results to compare the performance of the proposed artificial noise aided power allocation algorithms (Algorithm 2 in Table III for discrete inputs and Algorithm 3 in Table IV for Gaussian inputs) with the power allocation algorithm without using artificial noise (Algorithm 1 in Table I) and the classical/mercury water-filling strategy with no eavesdropper (Eve) [47]. The OFDM system with \( N = 64 \) subcarriers and CP length...
and Algorithm 3 using artificial noise for Gaussian inputs. The duality gap for Algorithm 1 was shown to decrease with $N$ (the number of subcarriers of OFDM) in the order $O(1/\sqrt{N})$. Numerical results were also provided to demonstrate that Algorithm 1 significantly outperforms some existing schemes though artificial noise is not used, and that Algorithm 2 (for discrete inputs) and Algorithm 3 (for Gaussian inputs) can further upgrade the secrecy rate performance compared to Algorithm 1 due to the use of the time-domain artificial noise.

As a future research, it is of interest to consider the scenario where the CSI of the eavesdropper is unknown or partially known to the transmitter [48]. Moreover, the scenario that instead of throwing away the cyclic prefix in standard OFDM systems, the eavesdropper can employ a more effective receiver, is worthy of further investigation. In our work, the artificial noise is only placed in the null space of the legitimate receiver, is worthy of further investigation. In our work, the artificial noise design by placing the artificial noise in a suitable subspace depending on both the legitimate receiver’s and the eavesdropper’s channels is also left as a future study.

APPENDIX I: PROOF OF (21)

In view of the fact that $\{s_i\}$ are i.i.d., and $\{w_i\}$ and $\{v_i\}$ are also identically distributed, let us write the secrecy rate on the frequency $f$ as

$$R_s(f, p) \triangleq \mathbb{E}[\mathbb{I}(s; H(f)\sqrt{p_s+n}) - \mathbb{I}(s; G(f)\sqrt{p_s+n})]^2,$$ (46)

where $s$ has the same distribution as $s_i$, and $n$ has the same distribution as $w_i$ and $v_i$. According to Proposition 1, we only need to consider the frequencies with $H(f) > G(f)$.

Since $H(f)$ and $G(f)$ are Lipschitz continuous, they are also uniformly bounded, such that

$$|H(f)| \leq M_H, \quad |G(f)| \leq M_G, \quad \forall f \in [0, 1]$$ (47)

where $M_H > 0$ and $M_G > 0$ are constants. By (9), (10), (47) and $\text{mmse}(|H(f)|^2) = E[|s|^2] = 1$ [34], we have that

$$\partial R_s(f, p) \over\partial p = H(f)^2 \text{mmse}(|H(f)|^2) p - G(f)^2 \text{mmse}(|G(f)|^2) p$$ (48a)

$$\leq |H(f)|^2 \leq M_H^2.$$ (48c)

According to the Lagrange’s mean value theorem [35], (48) implies that

$$|R_s(f, p) - R_s(f', p')| \leq M_H^2 |p - p'|,$$ (49)

for any two $p, p' \geq 0$. On the other hand, using (19), (47), $\text{mmse}(|H(f)|^2) \leq 1$, and

$$2pH \text{mmse}(|H|^2) \leq 2pH \text{mmse}(|H|^2)$$ (50)

we can attain an upper bound of $|R_s(f, p) - R_s(f', p)|$, i.e.,

$$|R_s(f, p) - R_s(f', p)| \leq \mathbb{E}[\mathbb{I}(s; H(f)\sqrt{p_s+n}) - \mathbb{I}(s; H(f')\sqrt{p_s+n})] + \mathbb{E}[\mathbb{I}(s; G(f')\sqrt{p_s+n}) - \mathbb{I}(s; G(f)\sqrt{p_s+n})]$$

$$\leq 2pM_H |H(f) - H(f')| + 2pM_G |G(f) - G(f')|$$

$$\leq (2pM_H L_H + 2pM_G L_G) |f - f'|.$$ (51)

V. CONCLUSIONS

We have presented three efficient transmission schemes for OFDM wiretap channels. The use of time-domain artificial noise was also considered for jamming the eavesdropper. These schemes were designed by maximizing secrecy rate under the constraint of total transmit power, and the Lagrange dual optimization method was used to resolve the associated nonconvex optimization problems. The proposed schemes consist of Algorithm 1 without using artificial noise for discrete inputs, Algorithm 2 using artificial noise for discrete inputs, and Algorithm 3 using artificial noise for Gaussian inputs. The performance of the scheme with no eavesdropper (denoted by “•”), Algorithm 1 in Table I without using artificial noise (denoted by “×”), and transmission rate with no eavesdropper achieved by classical/mercury water-filling strategy (denoted by “○”).

$N_{cp} = 16$ samples is considered. The channel coefficients $\{H_i\}$ and $\{G_i\}$ are generated by following the same procedure in Section III-B. The numerical results are averaged over 300 channel realizations with different values of $SNR = P$ (where artificial noise power is included in the total transmit power $P$). Figure 7 shows the secrecy rate performance of the three transmission schemes. One can observe from this figure that Algorithm 2 and Algorithm 3 (denoted by “▶”) using artificial noise outperform Algorithm 1 (denoted by “×”) without using artificial noise for all SNR values, and the amount of performance improvement is larger for higher SNR, justifying the effectiveness of the artificial noise design. The performance of the scheme with no eavesdropper (denoted by “○”) is expectantly better than either of Algorithm 2 and Algorithm 3 using artificial noise. However, the amount of performance difference for both finite constellations of QPSK and 16QAM is smaller for higher SNR, while it tends to saturate for higher SNR for the case of Gaussian inputs. These numerical results have substantiated the efficacy of Algorithm 2 and Algorithm 3 using artificial noise.
By (49) and (51), we obtain that
\[ |R_s(f,p) - R_s(f',p')| = |R_s(f,p) - R_s(f,p') + R_s(f,p') - R_s(f',p')| \]
\[ \leq |R_s(f,p) - R_s(f,p')| + |R_s(f,p') - R_s(f',p')| \]
\[ \leq M_H^2 |p - p'| + (2M_H L_H + 2p M_G L_G) |f - f'| \]
\[ \leq \max\{M_H^2, 2p M_H L_H + 2p M_G L_G\} \left( |p - p'| + |f - f'| \right), \]  
(52)

which is the same as (21) in which
\[ L_R = \max\{M_H^2, 2p M_H L_H + 2p M_G L_G\}. \]

**APPENDIX II: PROOF OF PROPOSITION 3**

The objective function of problem (24) can be rewritten as follows:
\[ R_s^{AN} = \frac{1}{N} \sum_{i=1}^{N} \left[ \mathcal{I}\left(s_i; \frac{H_s \sqrt{p_i}}{\sqrt{|H_s|^2 \sigma_{a,i}^2 + 1}} s_i + \tilde{w}_i \right) \right. \]
\[ - \mathcal{I}\left(s_i; \frac{G \sqrt{p_i}}{\sqrt{|G|^2 \sigma_{a,i}^2 + 1}} s_i + \tilde{v}_i \right) \]  
(53)

where \( \tilde{w}_i = (H_s \tilde{a}_i + w_i) / \sqrt{|H_s|^2 \sigma_{a,i}^2 + 1} \) and \( \tilde{v}_i = (G \tilde{a}_i + v_i) / \sqrt{|G|^2 \sigma_{a,i}^2 + 1} \) are Gaussian with zero mean and unity variance. The Lagrangian of problem (24) is given by
\[ R_s^{AN} + \lambda \left( P - \frac{1}{N} \sum_{i=1}^{N} p_i - \frac{1}{N} \sum_{i=1}^{N} \sigma_{a,i}^2 \right) + \sum_{i=1}^{N} \mu_i \sigma_{a,i}^2 + \sum_{i=1}^{N} \eta_i p_i, \]
where \( \lambda \geq 0 \) and \( \{ \mu_i \geq 0, \eta_i \geq 0 \} \) are the Lagrange dual variables associated with the constraints (24b) and (24c), respectively. Then the optimal solutions of problem (24), denoted by \( \{ p_i^* \} \) and \( \{ \sigma_{a,i}^{2*} \} \), must satisfy the following KKT necessary conditions [42]:
\[ \frac{\partial R_s^{AN}}{\partial p_i} = \frac{1}{N} \lambda + \eta_i = 0, \quad i = 1, 2, \ldots, N, \]  
(54a)
\[ \frac{\partial R_s^{AN}}{\partial \sigma_{a,i}^2} = \frac{1}{N} \lambda + \mu_i = 0, \quad i = 1, 2, \ldots, N, \]  
(54b)
\[ \lambda \left( P - \frac{1}{N} \sum_{i=1}^{N} p_i^* - \frac{1}{N} \sum_{i=1}^{N} \sigma_{a,i}^{2*} \right) = 0, \]  
(54c)
\[ \mu_i \sigma_{a,i}^{2*} = 0, \quad \eta_i p_i^* = 0, \quad i = 1, 2, \ldots, N, \]  
(54d)
\[ \lambda \geq 0, \quad \mu_i \geq 0, \quad \eta_i \geq 0, \quad i = 1, 2, \ldots, N, \]  
(54e)
\[ \sigma_{a,i}^{2*} \geq 0, \quad p_i^* \geq 0. \]  
(54f)

If \(|H_s| \leq |G_s|\) for some \( i \), then \(|H_s|^2 p_i^* / (|H_s|^2 \sigma_{a,i}^{2*} + 1) \leq |G_s|^2 p_i^* / (|G_s|^2 \sigma_{a,i}^{2*} + 1)\), and thus
\[ \mathcal{I}\left(s_i; \frac{H_s \sqrt{p_i}}{\sqrt{|H_s|^2 \sigma_{a,i}^2 + 1}} s_i + \tilde{w}_i \right) - \mathcal{I}\left(s_i; \frac{G \sqrt{p_i}}{\sqrt{|G|^2 \sigma_{a,i}^2 + 1}} s_i + \tilde{v}_i \right) = 0 \]  
(55)

regardless of the values of \( p_i^* \) and \( \sigma_{a,i}^{2*} \). Therefore, we can assume that \(|H_s| > |G_s| > 0\) for all \( i = 1, \ldots, N \), without loss of generality.

We use contradiction to prove the result. Suppose that \( \sigma_{a,i}^{2*} > 0 \) for some \( i \). Then it is necessary that \( p_i^* > 0 \); otherwise we end up with (55) again, and thus \( \sigma_{a,i}^{2*} > 0 \) is also a feasible solution. With \( \sigma_{a,i}^{2*} > 0 \) and \( p_i^* > 0 \), we have \( \mu_i = 0 \) and \( \eta_i = 0 \) by (54d). By (54a), (54b) and (54e), we further obtain
\[ \frac{\partial R_s^{AN}}{\partial p_i} = \frac{p_i^*}{(\sigma_{a,i}^{2*})^2 + 1} \geq 0, \]  
(57)

By [49, Equation (22)] and by using the chain rule, one can show that
\[ \frac{\partial R_s^{AN}}{\partial \sigma_{a,i}^2} = \frac{1}{N} \left[ \frac{|H_s|^2 \mathcal{E}_{b,i}(p_i^*, \sigma_{a,i}^{2*})^2}{(|G_s|^2 \sigma_{a,i}^{2*} + 1)^2} \right] \geq 0, \]  
(58)

where \( \mathcal{E}_{b,i}(p_i^*, \sigma_{a,i}^{2*}) \), \( \mathcal{E}_{c,i}(p_i^*, \sigma_{a,i}^{2*}) \), similar to (10), are given by
\[ \mathcal{E}_{b,i}(p_i^*, \sigma_{a,i}^{2*}) = E\left[ |s_i - E[s_i | H_s \sqrt{p_i} s_i + H_s \tilde{a}_i + w_i]|^2 \right] \geq 0, \]  
(59)

Combining \(|H_s|^2 p_i^* / (|H_s|^2 \sigma_{a,i}^{2*} + 1) > |G_s|^2 p_i^* / (|G_s|^2 \sigma_{a,i}^{2*} + 1)| = 0 \) (due to \(|H_s| > |G_s|\)) and (58) yields
\[ \frac{|H_s|^2 p_i^*}{(|H_s|^2 \sigma_{a,i}^{2*} + 1)} \geq \frac{|G_s|^2 p_i^*}{(|G_s|^2 \sigma_{a,i}^{2*} + 1)} \mathcal{E}_{b,i}(p_i^*, \sigma_{a,i}^{2*}), \]  
(60)

which, however, contradicts with (59). Therefore, (56) is not true and \( \sigma_{a,i}^{2*} = 0 \) for all \( i \). By this, the asserted statement is proved.

**ACKNOWLEDGMENT**

The authors would like to thank Fei He, Nicola Laurenti, Wei-Chiang Li, Kun-Yu Wang, and Shidong Zhou for valuable discussions to improve the quality of this paper.

**REFERENCES**

[1] J. L. Massey, “An introduction to contemporary cryptology,” Proc. IEEE, vol. 76, no. 5, pp. 533–549, May 1988.
[2] R. K. Nichols and P. C. Leekas, Wireless Security: Models, Threats, and Solutions, 3rd ed. McGraw-Hill Professional, 2004.
[3] Y. Liang, H. V. Poor, and S. Shamai (Shitz), “Information theoretic bounds,” Found. Trends Commun. Inf. Theory, vol. 5, pp. 355–580, 2008.
[4] R. Liu and W. Trappe, Securing Wireless Communications at the Physical Layer. Springer, 2010.
[5] A. Wyner, “The wire-tap channel,” Bell Syst. Tech. J., vol. 54, no. 8, pp. 1355–1387, Jan. 1975.
[6] S. Leung-Yan-Cheong and M. Hellman, “The Gaussian wire-tap channel,” IEEE Trans. Inf. Theory, vol. 24, no. 4, pp. 451–456, July 1978.
[7] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” IEEE Trans. Inf. Theory, vol. 24, no. 3, pp. 339–348, May 1978.
Z. Li, R. Yates, and W. Trappe, “Secrecy capacity of independent parallel channels,” in Proc. 44th Annu. Allerton Conf., Allerton House, Illinois, Sept. 27-29, 2006, pp. 841–848.

Y. Liang and H. V. Poor, “Secure communication over fading channels,” IEEE Trans. Inf. Theory, vol. 54, no. 6, pp. 2470–2492, June 2008.

E. Jorswieck and A. Wolf, “Resource allocation for the wire-tap multi-carrier broadcast channel,” in Proc. International Workshop Multiple Access Commun. Conf., Petersburg, Russia, June 16-19, 2008, pp. 1–6.

Z. Li, W. Trappe, and R. Yates, “Secret communication via multi-antenna transmission,” in Proc. Conf. Inf. Science and Syst., Mar. 14-16, 2007, pp. 905–910.

T. Liu and S. Shamai (Shitz), “A note on the secrecy capacity of the multiple-antenna wiretap channel,” IEEE Trans. Inf. Theory, vol. 55, no. 6, pp. 2547–2553, June 2009.

A. Khisti and G. Wornell, “Secure transmission with multiple antennas I: The MISOME wiretap channel,” IEEE Trans. Inf. Theory, vol. 56, no. 7, pp. 3088–3104, July 2010.

——, “Secure transmission with multiple antennas II: The MIMOME wiretap channel,” IEEE Trans. Inf. Theory, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

S. Goel and R. Negi, “Guaranteeing secrecy using artificial noise,” IEEE Trans. Wireless Commun., vol. 7, pp. 2180–2189, June 2008.

W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, “QoS-based transmit beamforming in the presence of eavesdroppers: An artificial-noise-sided approach,” IEEE Trans. Signal Process., vol. 59, no. 3, pp. 1202–1216, Mar. 2011.

T.-H. Chang, W.-C. Chiang, Y.-W. Peter Hong, and C.-Y. Chi, “Training sequence design for discriminatory channel estimation in wireless MIMO systems,” IEEE Trans. Signal Process., vol. 58, no. 12, pp. 6223–6237, Dec. 2010.

Q. Li and W.-K. Ma, “Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming,” IEEE Trans. Signal Process., vol. 59, no. 8, pp. 3799–3812, Aug. 2011.

Q. Li, W.-K. Ma, and A. M.-C. So, “Safe convex approximation to outage-based MISO secrecy rate optimization under imperfect CSI and with artificial noise,” in Proc. Asilomar Conf. Signal, Syst., and Comp., Pacific Grove, Nov. 2011.

X. He, A. Khisti, and A. Yener, “MIMO multiple access channel with an arbitrarily varying eavesdropper.” [Online]. Available: http://arxiv.org/pdf/1203.1376v1

M. Bloch, J. Barros, and M. R. Rodrigues, “Wireless information-theoretic security,” IEEE Trans. Inf. Theory: Special Issue on Information-Theoretic Security, vol. 54, no. 6, pp. 2515–2534, Dec. 2008.

L. Lai, H. E. Gamal, and H. V. Poor, “The wiretap channel with feedback: Encryption over the channel,” IEEE Trans. Inf. Theory, vol. 54, no. 11, pp. 381–392, Nov. 2008.

J. W. Wallace and R. K. Sharma, “Automatic secret keys from reciprocal MIMO wireless channels: Measurement and analysis,” IEEE Trans. Infom. Forensics and Security, vol. 5, no. 3, pp. 381–392, 2010.

M. Bloch and J. Barros, Physical-Layer Security: From Information Theory to Security Engineering. Cambridge University Press, July 2011.

L. Lai, Y. Liang, and H. V. Poor, “A unified framework for key agreement over wireless fading channels,” IEEE Trans. Inform. Forensics and Security, vol. 7, no. 2, pp. 480–490, Apr. 2012.

G. D. Raghava and B. S. DiRajan, “Secrecy capacity of the Gaussian wire-tap channel with finite complex constellation input.” [Online]. Available: http://arxiv.org/abs/1010.1163v1