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 signals 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

Manuscript received May 12, 2012; revised August 18, 2012 and November 26, 2012. The associate editor coordinating the review of this paper and approving it for publication was Sofiène Affes.

H. Qin, X. Chen (corresponding author), M. Zhao, and J. Wang are with the State Key Laboratory on Microwave and Digital Communications, Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, BJ 100084, P. R. China (e-mail: haohaoqin07, chenxiang98@gmail.com). Xiang Chen is also with the Aerospace Center, Tsinghua.

Y. Sun is with the Department of Electrical and Computer Engineering, the Ohio State University, Columbus, OH, 43210, USA (e-mail: sunyin02@gmail.com).

T.-H. Chang is with the Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan (e-mail: tsunghui.chang@ieee.org).

C.-Y. Chi is with the Institute of Communications Engineering and the Department of Electrical Engineering, National Tsinghua University, Hsinchu, Taiwan, 30013 (e-mail: cychi@ee.nthu.edu.tw).

This work is supported in part by the National Basic Research Program of China (2012CB316002), the National S&T Major Project (2011ZX03004-004), the National Natural Science Foundation of China (61132002), Tsinghua Research Funding-No.2010THZ02-3, the National Science Council, Taiwan, under Grant NSC99-2221-E007-052-MY3, the National Science Council, Taiwan, under Grant NSC101-2218-E-011-043, the International S&T Cooperation Program (2012DFG12010), and Ericsson Company.

Digital Object Identifier 10.1109/TCOMM.2013.01.xxxxx

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 multiple-output (MIMO) wiretap channels with discrete inputs [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 uppercase 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. $\mathrm{tr}(\mathbf{A})$ 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 $\mathbf{\Sigma}$. $\mathbb{E}[x]$ represents the expectation of the random variable x, and $\mathbb{E}[x|y]$ denotes the conditional expectation of x given y. Function $\mathcal{H}(x)$ denotes the entropy of random variable x, and $\mathcal{I}(x;y)$ denotes the mutual information between random variables x and y. $\mathrm{Diag}(x_1, x_2, ..., x_N)$ denotes the $N \times N$ diagonal matrix whose diagonal elements are $x_1, x_2, ..., x_N$. $[x]^+ \triangleq \max\{x, 0\}$, 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 ith 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

$$y_i = H_i\sqrt{p_i}s_i + w_i, \qquad i = 1, \ldots, N, \tag{1a}$$

$$z_i = G_i\sqrt{p_i}s_i + v_i, \qquad i = 1, \ldots, N, \tag{1b}$$

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 ith 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]

$$R_s(\mathbf{p}) = \frac{1}{N}\sum_{i=1}^{N}[\mathcal{I}(s_i; H_i\sqrt{p_i}s_i + w_i) - \mathcal{I}(s_i; G_i\sqrt{p_i}s_i + v_i)]^+, \tag{2}$$

where $\mathbf{p} = [p_1, p_2, ..., p_N]^T$ contains all the subcarrier power variables. Let us consider the following secrecy rate maximization problem:

$$R^\star = \max_{\mathbf{p}\succeq\mathbf{0}} \quad R_s(\mathbf{p}) \tag{3a}$$

$$\text{s.t.} \quad \frac{1}{N}\sum_{i=1}^{N} p_i \leq P, \tag{3b}$$

where P in (3b) is the allowed maximal transmit power, and $R^\star$ denotes the maximum achievable secrecy rate.

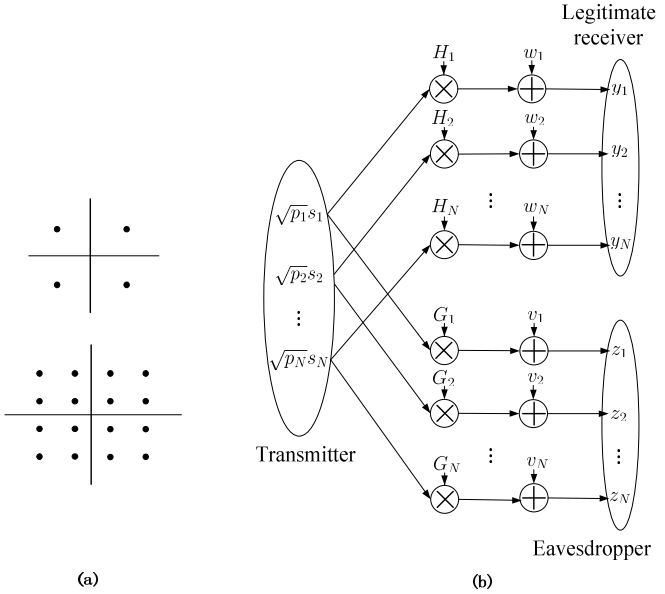


Fig. 1. (a) Finite discrete inputs: QPSK and 16QAM, (b) OFDM wiretap channel.

A. Gaussian Channel Inputs

We first briefly review the case when $\{s_i\}$ are Gaussian signals. In this case, the secrecy rate $R_s(\mathbf{p})$ in (2) can be easily reduced to

$$R_s(\mathbf{p}) = \frac{1}{N}\sum_{i=1}^{N}[\ \log_2(1+|H_i|^2p_i) - \log_2(1+|G_i|^2p_i)\ ]^+, \tag{4}$$

which is known to be concave in $\mathbf{p}$. Therefore, the power allocation problem (3) is a convex problem which can be solved efficiently thanks to the following closed-form solution [8]–[10]:

$$p_i^\star = \begin{cases} \frac{1}{2|H_i|^2|G_i|^2}\left[\sqrt{C_i^2 - \frac{4|H_i|^2|G_i|^2(\lambda+|G_i|^2-|H_i|^2)}{\lambda}} - C_i\right], \\ \qquad\qquad \text{if } |H_i|^2-|G_i|^2>\lambda \\ 0, \qquad\qquad \text{otherwise,} \end{cases} \tag{5}$$

where $C_i = |H_i|^2+|G_i|^2$, and$\lambda \geq 0$ is the Lagrange multiplier associated with the total power constraint (3b), and should be chosen such that $\{p_i^\star\}$ in (5) satisfies $\frac{1}{N}\sum_{i=1}^{N} p_i^\star = P$. The reader can refer to [8]–[10] for more details.

B. Discrete Channel Inputs

In [28], [30]–[33], it was observed from computer simulations that the secrecy rate $R_s(\mathbf{p})$ in (4) is nonconcave in $\mathbf{p}$ for some discrete constellations, such as QPSK and 16QAM. One can infer from Theorem 1 in [32] that $R_s(\mathbf{p})$ is nonconcave for any uniformly distributed discrete inputs. We now prove that $R_s(\mathbf{p})$ is nonconcave for any discrete channel input distribution with a finite number of possible values:

Proposition 1 *Consider the following secrecy rate function:*

$$R(p) \triangleq [\mathcal{I}(s; H\sqrt{p}s+w) - \mathcal{I}(s; G\sqrt{p}s+v)]^+, \tag{6}$$

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 $\mathcal{I}(s; w) = \mathcal{I}(s; v) = 0$, 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

$$\lim_{p\to\infty} \mathcal{I}(s; H\sqrt{p}s+w) = \lim_{p\to\infty} \mathcal{I}(s; G\sqrt{p}s+v) = \mathcal{H}(s) < \infty. \tag{7}$$

Therefore,

$$\lim_{p\to\infty} R(p) = 0. \tag{8}$$

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 $\mathcal{I}(s; H\sqrt{p}s+w)$ is given by

$$\frac{\partial \mathcal{I}(s; H\sqrt{p}s+w)}{\partial p} = |H|^2 \mathrm{mmse}\left(|H|^2 p\right) \geq 0, \tag{9}$$

where

$$\mathrm{mmse}(|H|^2p) \triangleq \mathbb{E}\left[|s - \mathbb{E}(s|H\sqrt{p}s+w)|^2\right] \tag{10}$$

is the minimum mean square error (MMSE) of estimating s with the received signal $y = H\sqrt{p}s + w$. When p equals zero, $\mathrm{mmse}(|H|^2p)$ in (10) attains its maximum value, i.e., $\mathrm{mmse}(0) = E[|s|^2] = 1$. 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 $\tilde{p} \in [\hat{p}, \infty)$ satisfying $R(\tilde{p}) < \infty$ and $R'(\tilde{p}) < 0$.

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

$$R(p) \leq R(\tilde{p}) + R'(\tilde{p})(p-\tilde{p}) \tag{11}$$

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 $R'(\tilde{p}) < 0$, which however contradicts the fact of $\lim_{p\to\infty} R(p) = 0$. Therefore, the concavity assumption for $R(p)$ is not true.

When $|H| \leq |G|$, $\mathcal{I}(s; H\sqrt{p}s+w) \leq \mathcal{I}(s; G\sqrt{p}s+v)$ 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 $R_s(\mathbf{p})$ 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 continuous distributions and some of the discrete distributions

[1]A discrete distribution with a countable number of values may either have a finite entropy or an infinite entropy [36, page 48].

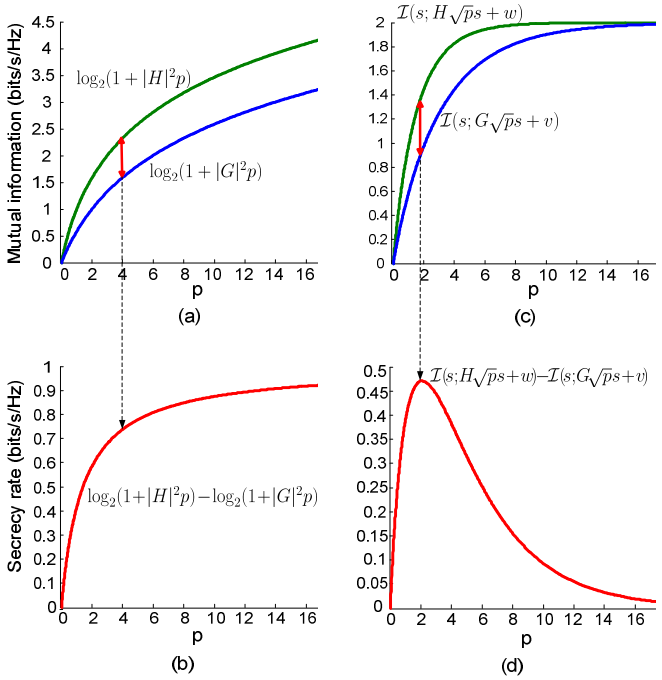


Fig. 2. (a) Mutual information and (b) secrecy rate for a Gaussian channel input. (c) Mutual information and (d) secrecy rate for a QPSK channel input. Here $|H|=1$ and $|G|=0.5$.

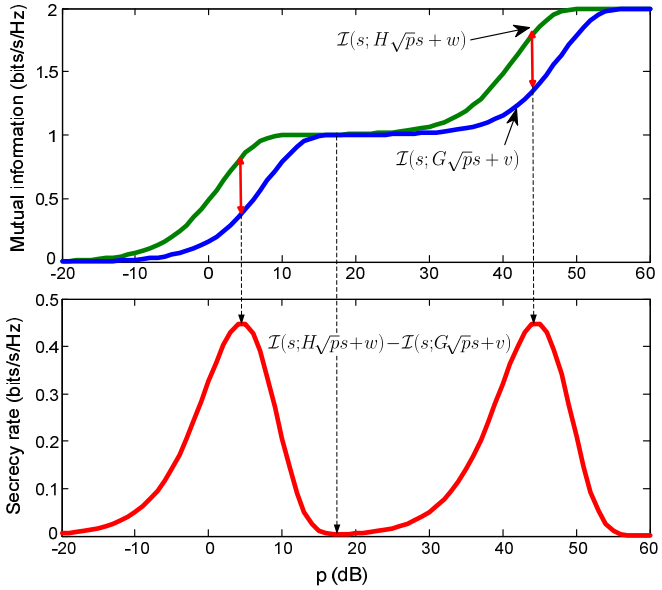


Fig. 3. Secrecy rate for a non-standard 4-PAM channel input with probability mass function given by (12). Here $|H|=1$ and $|G|=0.5$.

with a countable number of values, and this type of channel inputs may have a concave secrecy rate function.

Let us provide some numerical results to illustrate Proposition 1. Figure 2 shows the mutual information $\mathcal{I}(s;H\sqrt{p}s+w)$, $\mathcal{I}(s;G\sqrt{p}s+v)$, and the secrecy rate $R(p)$ for Gaussian and QPSK channel inputs. One can observe from Fig. 2(a) and Fig. 2(b) that $\mathcal{I}(s;H\sqrt{p}s+w)$, $\mathcal{I}(s;G\sqrt{p}s+v)$ and $R(p)$ are all concave when the inputs are Gaussian. From Fig. 2(c), $\mathcal{I}(s;H\sqrt{p}s+w)$ and $\mathcal{I}(s;G\sqrt{p}s+v)$ 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 [2], where the probability mass function of s is given by

$$P_s \sim \begin{bmatrix} -51q & -50q & 50q & 51q \\ 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix}, \tag{12}$$

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 $p < -10\text{dB}$, 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 be able to identify the two groups as p increases. Since $|H|>|G|$, $I(s;H\sqrt{p}s+w)$ is larger than $I(s;G\sqrt{p}s+v)$, and so $R(p)>0$. When $p\approx 18\text{dB}$, both Bob and Eve can identify the two groups of s, and $I(s;H\sqrt{p}s+w)\approx I(s;G\sqrt{p}s+v)\approx 1$ bits/s/Hz, and so $R(p)$ decreases to nearly 0. When $p\geq 20\text{dB}$, 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

$$L(\mathbf{p},\lambda) = \frac{1}{N}\sum_{i=1}^{N}[\mathcal{I}(s_i;H_i\sqrt{p_i}s_i+w_i)-\mathcal{I}(s_i;G_i\sqrt{p_i}s_i+v_i)]^+ + \lambda\left(P-\frac{1}{N}\sum_{i=1}^{N}p_i\right), \tag{13}$$

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

$$D^\star = \min_{\lambda\geq 0}\ d(\lambda), \tag{14}$$

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

$$d(\lambda) \triangleq \max_{\mathbf{p}\succeq 0}\ L(\mathbf{p},\lambda). \tag{15}$$

[2]While the mutual information $\mathcal{I}(s;H\sqrt{p}s+w)$ and $\mathcal{I}(s;G\sqrt{p}s+v)$ are concave functions in p, they appear to be nonconcave in logarithmic scale.

TABLE I

Algorithm 1: Proposed power allocation scheme for discrete inputs.	
Given:	$\lambda_h \geq \lambda_l = 0$, convergence tolerance ε
repeat:	
step 1:	update $\lambda = \frac{1}{2}(\lambda_l + \lambda_h)$
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 $\lambda^\star = \lambda_l$.

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

$$d(\lambda) = \frac{1}{N}\sum_{i=1}^{N} B(p_i;\lambda,H_i,G_i) + \lambda P, \tag{16}$$

where

$$\begin{aligned} &B(p_i;\lambda,H_i,G_i) \\ &= \max_{p_i\geq 0}\left[\mathcal{I}(s_i;H_i\sqrt{p_i}s_i+w_i)-\mathcal{I}(s_i;G_i\sqrt{p_i}s_i+v_i)\right]^+ - \lambda p_i \end{aligned} \tag{17}$$

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 (16) and (17), to be

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

where $p_i^\star$ is the optimal solution to (17). Therefore, the dual variable λ can be efficiently updated using the bisection method [37]. If $P - \frac{1}{N}\sum_{i=1}^N p_i^\star > 0$, then the subgradient $\nabla d(\lambda) > 0$, and thus we decrease λ in the bisection method for finding $D^\star$ given by (14); otherwise we increase λ. The resulting power allocation algorithm, called Algorithm 1, for finding the desired $\mathbf{p}^\star$ 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 $\frac{1}{N}\sum_{i=1}^N p_i^\star$ by solving (17) can be strictly lower than the available transmit power P, and the optimal dual variable is $\lambda^\star = 0$ 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 $R^\star$ and optimal $D^\star$, i.e., $D^\star - R^\star > 0$, where $R^\star$ and $D^\star$ 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 $D^\star$ and $R^\star$ 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

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

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 $f, f' \in [0,1]$:

$$\left|H(f)–H(f')\right| \leq L_H|f-f'|,\ \left|G(f)–G(f')\right| \leq L_G|f-f'| \tag{19}$$

We can show the following proposition:

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

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

where $R^\star$ and $D^\star$ 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.,

$$|R_s(f,p) - R_s(f',p')| \leq L_R\left(|f-f'| + |p-p'|\right), \tag{21}$$

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 $\mathrm{SNR} = P$ (which is the ratio of the available transmit power to AWGN power ($\mathbb{E}[|w_i|^2] = \mathbb{E}[|v_i|^2] = 1$) 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)".

Figure 4 illustrates the achievable secrecy rates of the three power allocation schemes for different channel inputs. It can be observed from this figure that Algorithm 1 (denoted by "+") achieves the highest secrecy rate (although it yields the same optimal solution (5) for Gaussian channel inputs). One

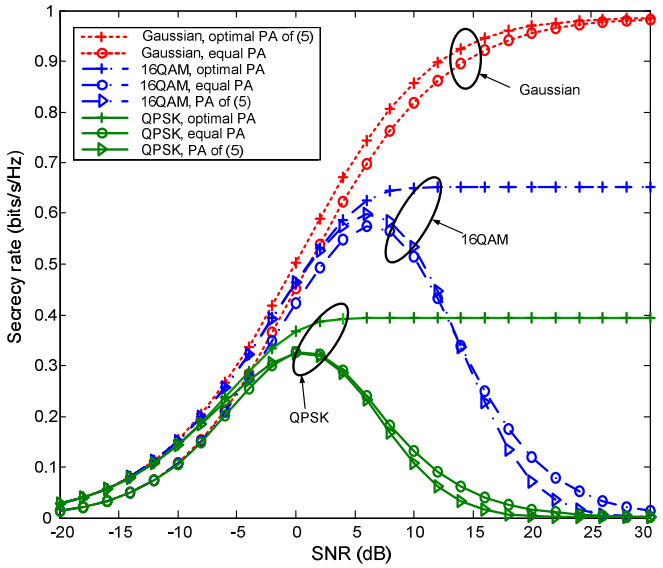


Fig. 4. Secrecy rates achieved by Algorithm 1 (in Table I) denoted by "+", equal power allocation scheme denoted by "○", and power allocation scheme given by (5) denoted by "▷".

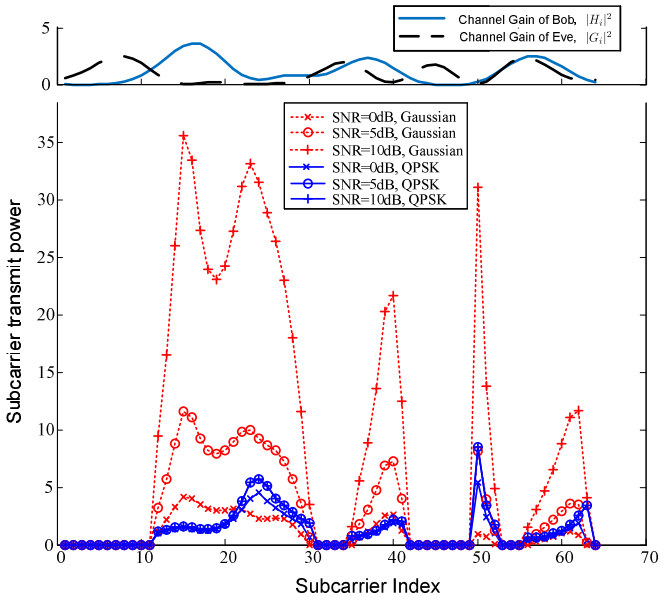


Fig. 5. Power allocation results obtained using Algorithm 1 for the case of Gaussian inputs and the case of QPSK inputs, where $\mathrm{SNR} = 0\mathrm{dB}$; $5\mathrm{dB}$; $10\mathrm{dB}$ and $N = 64$.

can observe, from Fig. 4, that under discrete inputs (QPSK and 16QAM), both the equal PA scheme (denoted by "○") and the PA scheme of (5) (denoted by "▷") do not perform well—their secrecy rates even drop to zero in the high SNR regime. The secrecy rate utilized by Algorithm 1 increases with SNR, but saturates (rather than drops to zero) in the high SNR regime, implying that there exists a power threshold for which the secrecy rate reaches the maximum, and the power exceeding the threshold will not improve the secrecy rate any more. The reason for this is that the secrecy rate is not an increasing function of the transmit power due to (8). Similar observations were also reported in [28], [31].

Figure 5 illustrates the power allocation results obtained by Algorithm 1 for Gaussian and QPSK channel inputs. Three SNR values are considered, i.e., $\mathrm{SNR} = 0\mathrm{dB}$, $5\mathrm{dB}$, $10\mathrm{dB}$. It can be observed, from Fig. 5, that the transmit power is only allocated to the subcarriers on which Bob has larger channel gain than Eve. As previously mentioned, for discrete channel inputs, it may not be true that the total power constraint (3b) is active for the power allocation result obtained by Algorithm 1. Actually, when P is larger than a threshold, the power allocation result remains unchanged. As shown in Fig. 5, when the channel inputs are QPSK, the power allocation result associated with $\mathrm{SNR} = 5\mathrm{dB}$ almost coincides with that associated with $\mathrm{SNR} = 10\mathrm{dB}$. This also demonstrates that the total transmit power constraint (3b) is inactive when $\mathrm{SNR} = 10\mathrm{dB}$ for QPSK inputs.

IV. Joint Signal and Artificial Noise Design

As presented in the previous section, the secrecy rate obtained by Algorithm 1 reaches the maximum value and saturates in the high SNR regime. In other words, the transmitter in the high SNR regime will not consume all the available transmit power. This motivates the use of the remaining transmit power for generating artificial noise to jam the eavesdropper, thus further increasing the secrecy rate. In the subsequent subsections, the independent frequency-domain artificial noise design is shown to be ineffective first. Then a novel time-domain artificial noise scheme is proposed to exploit temporal degrees of freedom provided by the cyclic prefix of OFDM systems to jam the eavesdropper and boost the secrecy rate.

A. Ineffectiveness of Independent Frequency-Domain Artificial Noise Design

Let us first consider a naive strategy by adding artificial noise to all the subcarriers in the frequency domain, as illustrated in Fig. 6(a), where $\mathbf{P} = \mathrm{Diag}(p_1, p_2, ..., p_N)$ and $\mathbf{s} = [s_1, s_2, ..., s_N]^T$. The corresponding received signals of the legitimate receiver and the eavesdropper on the ith subcarrier can be expressed as

$$y_i = H_i(\sqrt{p_i}s_i + \hat{a}_i) + w_i, \qquad i = 1, \ldots, N, \tag{22a}$$

$$z_i = G_i(\sqrt{p_i}s_i + \hat{a}_i) + v_i, \qquad i = 1, \ldots, N, \tag{22b}$$

where $\hat{a}_i \in \mathbb{C}$ is the artificial noise term added to the ith subcarrier. It is assumed that $\hat{a}_i$ are statistically independent across the subcarriers and $\hat{a}_i \sim \mathcal{CN}(0, \sigma_{a,i}^2)$, where $\sigma_{a,i}^2 \geq 0$ is the artificial noise power for subcarrier i. According to [8], the secrecy rate for the signal model (22) is given by

$$R_s^{AN} = \frac{1}{N}\sum_{i=1}^{N}\Big[\mathcal{I}\big(s_i; H_i(\sqrt{p_i}s_i + \hat{a}_i) + w_i\big) - \mathcal{I}\big(s_i; G_i(\sqrt{p_i}s_i + \hat{a}_i) + v_i\big)\Big]^+. \tag{23}$$

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:

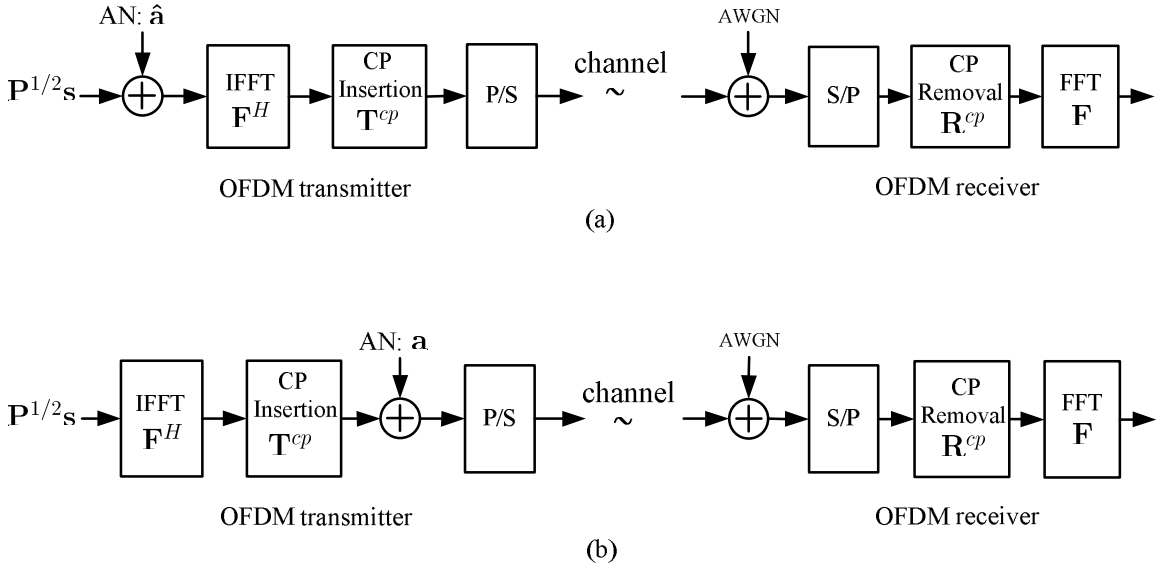


Fig. 6. (a) System model of OFDM transceiver with artificial noise (AN) added in the frequency domain. (b) System model of OFDM transceiver with artificial noise added in the time domain.

$$\max_{\{p_i\}_{i=1}^N,\{\sigma_{a,i}^2\}_{i=1}^N} \quad R_s^{AN} \tag{24a}$$

$$\text{s.t.} \quad \frac{1}{N}\sum_{i=1}^{N}(p_i+\sigma_{a,i}^2) \le P \tag{24b}$$

$$\sigma_{a,i}^2 \ge 0,\ p_i \ge 0,\ i=1,2,...,N. \tag{24c}$$

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 $p_i \ge 0$. 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 $\mathbf{P}^{1/2}\mathbf{s}$ 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 $\mathbf{F}$ and $\mathbf{F}^H$ 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 $\mathbf{T}^{cp} = [\tilde{\mathbf{E}}_{N_{cp}\times N}^T\ \mathbf{I}_N]^T$ and $\mathbf{R}^{cp} = [\mathbf{0}_{N\times N_{cp}}\ \mathbf{I}_N]$, respectively, where $\tilde{\mathbf{E}}_{N_{cp}\times N}$ contains the last N_{cp} rows of the $N \times N$ identity matrix $\mathbf{I}_N$. 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]

$$\begin{aligned}\mathbf{y} &= \mathbf{F}\mathbf{R}^{cp}\mathbf{H}_0(\mathbf{T}^{cp}\mathbf{F}^H\mathbf{P}^{1/2}\mathbf{s}+\mathbf{a})+\mathbf{w}\\ &= \mathbf{H}\mathbf{P}^{1/2}\mathbf{s}+\mathbf{F}\mathbf{R}^{cp}\mathbf{H}_0\mathbf{a}+\mathbf{w},\end{aligned} \tag{25a}$$

$$\begin{aligned}\mathbf{z} &= \mathbf{F}\mathbf{R}^{cp}\mathbf{G}_0(\mathbf{T}^{cp}\mathbf{F}^H\mathbf{P}^{1/2}\mathbf{s}+\mathbf{a})+\mathbf{v}\\ &= \mathbf{G}\mathbf{P}^{1/2}\mathbf{s}+\mathbf{F}\mathbf{R}^{cp}\mathbf{G}_0\mathbf{a}+\mathbf{v},\end{aligned} \tag{25b}$$

where $\mathbf{a} \in \mathbb{C}^{N+N_{cp}}$ is a zero-mean complex Gaussian random vector, $\mathbf{H}_0 \in \mathbb{C}^{(N+N_{cp})\times(N+N_{cp})}$ is a Toeplitz channel matrix given by

$$\mathbf{H}_0 = \begin{bmatrix} h(0) & 0 & 0 & \cdots & 0 \\ \vdots & h(0) & 0 & \cdots & 0 \\ h(L) & \cdots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \cdots & \ddots & 0 \\ 0 & \cdots & h(L) & \cdots & h(0) \end{bmatrix},$$

and so is the matrix $\mathbf{G}_0$, i.e., by replacing $h(l)$ with $g(l), l = 0, 1, ..., L$, in $\mathbf{H}_0$. Moreover, in (25), $\mathbf{H} = \mathbf{F}\mathbf{R}^{cp}\mathbf{H}_0\mathbf{T}^{cp}\mathbf{F}^H = \mathrm{Diag}(H_1, H_2, ..., H_N)$ and $\mathbf{G} = \mathbf{F}\mathbf{R}^{cp}\mathbf{G}_0\mathbf{T}^{cp}\mathbf{F}^H = \mathrm{Diag}(G_1, G_2, ..., G_N)$, in which H_i and G_i are the frequency responses of the legitimate receiver's channel and the eavesdropper's channel, respectively, and $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ and $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ are the corresponding noise vectors at the legitimate receiver and the eavesdropper, respectively.

$$L(\mathbf{p},\mathbf{\Sigma}_d,\lambda)=\tfrac{1}{N}\sum_{i=1}^{N}\left[\mathcal{I}(s_i;H_i\sqrt{p_i}s_i+w_i)-\mathcal{I}(s_i;\tfrac{G_i\sqrt{p_i}s_i}{\sqrt{\mathbf{b}_i^H\mathbf{\Sigma}_d\mathbf{b}_i+1}}+\hat{v}_i)\right]^+ +\lambda\left(P-\tfrac{1}{N}\sum_{i=1}^{N}p_i-\tfrac{1}{N}\mathrm{tr}(\mathbf{\Sigma}_d)\right) \tag{32}$$

$$p_i^\star=\arg\max_{p_i>0}\ [\mathcal{I}(s_i;H_i\sqrt{p_i}s_i+w_i)-\mathcal{I}(s_i;\frac{G_i\sqrt{p_i}s_i}{\sqrt{\mathbf{b}_i^H\mathbf{\Sigma}_d\mathbf{b}_i+1}}+\hat{v}_i)]^+-\lambda p_i,\ i=1,2,...,N. \tag{35}$$

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

$$y_i=H_i\sqrt{p_i}s_i+\mathbf{f}_i^T\mathbf{R}^{cp}\mathbf{H}_0\mathbf{a}+w_i,\quad i=1,2,...,N, \tag{26a}$$

$$z_i=G_i\sqrt{p_i}s_i+\mathbf{f}_i^T\mathbf{R}^{cp}\mathbf{G}_0\mathbf{a}+v_i,\quad i=1,2,...,N, \tag{26b}$$

respectively, where $\mathbf{f}_i^T\in\mathbb{C}^N$ is the ith row of the FFT matrix $\mathbf{F}$. 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, we let

$$\mathbf{a}=\mathbf{U}\mathbf{d}, \tag{27}$$

where $\mathbf{U}$ is a semi-unitary matrix whose column vectors span the null space of $\mathbf{R}^{cp}\mathbf{H}_0$, i.e.,

$$\mathbf{R}^{cp}\mathbf{H}_0\mathbf{U}=\mathbf{0},\quad \mathbf{U}^H\mathbf{U}=\mathbf{I}_{N_{cp}}, \tag{28}$$

and $\mathbf{d}\sim\mathcal{CN}(\mathbf{0},\mathbf{\Sigma}_d)$ in which $\mathbf{\Sigma}_d$ is the covariance matrix of the artificial noise vector $\mathbf{d}$ to be determined. As the dimension of $\mathbf{R}^{cp}\mathbf{H}_0$ is $N\times(N+N_{cp})$, the dimension of $\mathbf{U}$ is $(N+N_{cp})\times N_{cp}$ and the dimension of $\mathbf{d}$ is N_{cp}.

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

$$y_i=H_i\sqrt{p_i}s_i+w_i,\ i=1,2,...,N, \tag{29a}$$

$$z_i=G_i\sqrt{p_i}s_i+\mathbf{f}_i^T\mathbf{R}^{cp}\mathbf{G}_0\mathbf{U}\mathbf{d}+v_i,\ i=1,2,...,N. \tag{29b}$$

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

$$\frac{1}{N}\sum_{i=1}^{N}\left[\mathcal{I}(s_i;H_i\sqrt{p_i}s_i+w_i)-\mathcal{I}(s_i;\frac{G_i\sqrt{p_i}s_i}{\sqrt{\mathbf{b}_i^H\mathbf{\Sigma}_d\mathbf{b}_i+1}}+\hat{v}_i)\right]^+, \tag{30}$$

where $\mathbf{b}_i^H\triangleq\mathbf{f}_i^T\mathbf{R}^{cp}\mathbf{G}_0\mathbf{U}$, and $\hat{v}_i\sim\mathcal{CN}(0,1)$. The joint power allocation and artificial noise design problem can be formulated as

$$\max_{\{p_i\}_{i=1}^N,\mathbf{\Sigma}_d}\ \frac{1}{N}\sum_{i=1}^{N}\left[\mathcal{I}(s_i;H_i\sqrt{p_i}s_i+w_i)-\mathcal{I}(s_i;\frac{G_i\sqrt{p_i}s_i}{\sqrt{\mathbf{b}_i^H\mathbf{\Sigma}_d\mathbf{b}_i+1}}+\hat{v}_i)\right]^+ \tag{31a}$$

$$\text{s.t.}\quad \frac{1}{N}\left(\sum_{i=1}^{N}p_i+\mathrm{tr}(\mathbf{\Sigma}_d)\right)\le P \tag{31b}$$

$$\mathbf{\Sigma}_d\succeq 0,\quad p_i\ge 0,\quad i=1,\ldots,N. \tag{31c}$$

C. 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 top of the page, where λ is the Lagrange dual variable associated with constraint (31b). The associated dual function is defined as

$$d(\lambda)=\max_{\mathbf{p}\succeq 0,\mathbf{\Sigma}_d\succeq 0}L(\mathbf{p},\mathbf{\Sigma}_d,\lambda). \tag{33}$$

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

$$\min_{\lambda\ge 0}\ d(\lambda). \tag{34}$$

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

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

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

$$\min_{\mathbf{\Sigma}_d\succeq 0}\ \sum_{i=1}^{N}\mathcal{I}(s_i;\frac{G_i}{\sqrt{\mathbf{b}_i^H\mathbf{\Sigma}_d\mathbf{b}_i+1}}\sqrt{p_i}s_i+\hat{v}_i)+\lambda\mathrm{tr}(\mathbf{\Sigma}_d). \tag{36}$$

For ease of presentation, let us define

$$T_i(\mathbf{\Sigma}_d)\triangleq\mathcal{I}(s_i;\frac{G_i}{\sqrt{\mathbf{b}_i^H\mathbf{\Sigma}_d\mathbf{b}_i+1}}\sqrt{p_i}s_i+\hat{v}_i),\ i=1,2,...,N. \tag{37}$$

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

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

$$T_i'(t_i)=|G_i|^2p_i\mathrm{mmse}(|G_i|^2p_it_i).\qquad(\text{by (9)}) \tag{38}$$

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

$$\tilde{T}_i(t_i)=T_i(\bar{t}_i)+|G_i|^2p_i[\mathrm{mmse}(|G_i|^2p_i\bar{t}_i)]\,(t_i-\bar{t}_i)\,. \tag{39}$$

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

$$\begin{array}{ll}\min\limits_{\mathbf{\Sigma}_d\succeq 0,\{t_i\}_{i=1}^N} & \sum\limits_{i=1}^{N}\tilde{T}_i(t_i)+\lambda\mathrm{tr}(\mathbf{\Sigma}_d)\\ \text{s.t.} & t_i=(\mathbf{b}_i^H\mathbf{\Sigma}_d\mathbf{b}_i+1)^{-1},\quad i=1,2,...,N.\end{array} \tag{40}$$

TABLE II

SCA Algorithm: Algorithm for solving problem (36).

Given: $\mathbf{\Sigma}_d$
repeat:
step 1: solve problem (43) by CVX and obtain optimal $\mathbf{\Sigma}_d$
step 2: set $\bar{\mathbf{\Sigma}}_d = \mathbf{\Sigma}_d$
until: a specified convergence criterion is satisfied.

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

$$\begin{array}{cc} \min\limits_{\mathbf{\Sigma}_d \succeq 0, \{t_i\}_{i=1}^N} & \sum\limits_{i=1}^N |G_i|^2 p_i \left[\mathrm{mmse}(\frac{|G_i|^2 p_i}{\mathbf{b}_i^H \mathbf{\Sigma}_d \mathbf{b}_i + 1})\right] t_i + \lambda \mathrm{tr}(\mathbf{\Sigma}_d) \\ \text{s.t.} & t_i = (\mathbf{b}_i^H \mathbf{\Sigma}_d \mathbf{b}_i + 1)^{-1},\ i = 1, 2, ..., N. \end{array} \tag{41}$$

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

$$\begin{array}{cc} \min\limits_{\mathbf{\Sigma}_d \succeq 0, \{t_i\}_{i=1}^N} & \sum\limits_{i=1}^N |G_i|^2 p_i \left[\mathrm{mmse}(\frac{|G_i|^2 p_i}{\mathbf{b}_i^H \mathbf{\Sigma}_d \mathbf{b}_i + 1})\right] t_i + \lambda \mathrm{tr}(\mathbf{\Sigma}_d) \\ \text{s.t.} & t_i \geq (\mathbf{b}_i^H \mathbf{\Sigma}_d \mathbf{b}_i + 1)^{-1}, \quad i = 1, 2, ..., N. \end{array} \tag{42}$$

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

$$\begin{array}{cc} \min\limits_{\mathbf{\Sigma}_d \succeq 0, \{t_i\}_{i=1}^N} & \sum\limits_{i=1}^N |G_i|^2 p_i \left[\mathrm{mmse}(\frac{|G_i|^2 p_i}{\mathbf{b}_i^H \mathbf{\Sigma}_d \mathbf{b}_i + 1})\right] t_i + \lambda \mathrm{tr}(\mathbf{\Sigma}_d) \\ \text{s.t.} & \begin{bmatrix} \mathbf{b}_i^H \mathbf{\Sigma}_d \mathbf{b}_i + 1 & 1 \\ 1 & t_i \end{bmatrix} \succeq \mathbf{0}, \quad i = 1, 2, ..., N. \end{array} \tag{43}$$

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 respective to t_i [43];
- $\tilde{T}_i(\bar{t}_i) = T_i(\bar{t}_i)$;
- $\tilde{T}_i'(\bar{t}_i) = T_i'(\bar{t}_i)$.

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 aided power allocation algorithm, called Algorithm 2, for solving (31) is summarized in Table III.

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 $p_i^\star$ in (35) is simply given by

$$p_i^\star = \begin{cases} \frac{1}{2|H_i|^2|\hat{G}_i|^2}\left(\sqrt{\hat{C}^2 - 4|H_i|^2|\hat{G}_i|^2 \frac{\lambda + |\hat{G}_i|^2 - |H_i|^2}{\lambda}} - \hat{C}\right), \\ \qquad\qquad \text{if} \quad |H_i|^2 - |\hat{G}_i|^2 > \lambda \\ 0, \qquad\qquad \text{otherwise}, \end{cases} \tag{44}$$

where $\hat{G}_i = \frac{G_i}{\sqrt{\mathbf{b}_i^H \mathbf{\Sigma}_d \mathbf{b}_i + 1}}$ and $\hat{C} = |H_i|^2 + |\hat{G}_i|^2$. Problem (36) for this case also reduces to a simple convex optimization problem as follows:

$$\mathbf{\Sigma}_d^\star = \arg \min_{\mathbf{\Sigma}_d \succeq 0} \sum_{i=1}^N \log_2 \left(1 + \frac{|G_i|^2 p_i}{\mathbf{b}_i^H \mathbf{\Sigma}_d \mathbf{b}_i + 1}\right) + \lambda \mathrm{tr}(\mathbf{\Sigma}_d). \tag{45}$$

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

TABLE III

Algorithm 2: Proposed artificial noise aided power allocation scheme for discrete inputs.

Given: $\lambda_h \geq \lambda_l = 0$, $\{p_i\}_{i=1}^N$
repeat:
step 1: update $\lambda = \frac{1}{2}(\lambda_l + \lambda_h)$
step 2: repeat:
obtain the optimal $\mathbf{\Sigma}_d$ using SCA Algorithm in Table II and then obtain $\{p_i\}_{i=1}^N$ by solving (35)
until: $L(\mathbf{p}, \mathbf{\Sigma}_d, \lambda)$ meets a specified convergence criterion
step 3: if $\frac{1}{N}\left(\mathrm{tr}(\mathbf{\Sigma}_d) + \sum\limits_{i=1}^N p_i\right) < P$, then update $\lambda_h = \lambda$,
else update $\lambda_l = \lambda$
until: λ meets a specified convergence criterion.

TABLE IV

Algorithm 3: Proposed artificial noise aided power allocation scheme for Gaussian inputs.

Given: $\lambda_h \geq \lambda_l = 0$, $\{p_i\}_{i=1}^N$
repeat:
step 1: update $\lambda = \frac{1}{2}(\lambda_l + \lambda_h)$
step 2: repeat:
by solving problem (45) using CVX and then obtain $\{p_i\}_{i=1}^N$ by (44)
until: $L(\mathbf{p}, \mathbf{\Sigma}_d, \lambda)$ meets a specified convergence criterion
step 3: if $\frac{1}{N}\left(\mathrm{tr}(\mathbf{\Sigma}_d) + \sum\limits_{i=1}^N p_i\right) < P$, then update $\lambda_h = \lambda$,
else update $\lambda_l = \lambda$
until: λ meets a specified convergence criterion.

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 III 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 $N_{cp} = 16$ 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 $\mathrm{SNR} = P$ (where artificial noise power is included in the total transmit

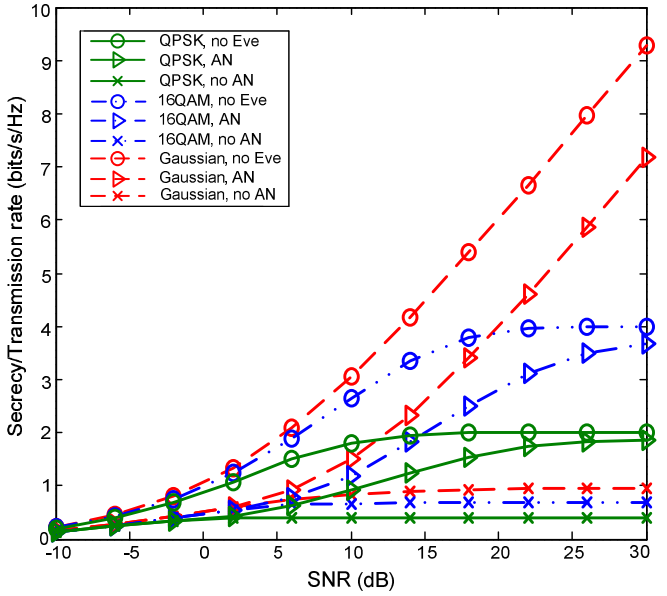


Fig. 7. Secrecy rates achieved by Algorithm 2 in Table III using artificial noise (AN) for discrete inputs (denoted by "▷"), Algorithm 3 in Table IV using artificial noise for Gaussian inputs (also denoted by "▷"), Algorithm 1 in Table I without using artificial noise (denoted by "×"), and transmission rate with no eavesdropper achieved by classical/mercury water-filling strategy (denoted by "○").

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.

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 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's channel. 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

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

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

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

$$|H(f)| \le M_H, \quad |G(f)| \le M_G, \quad \forall f \in [0,1] \tag{47}$$

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

$$\frac{\partial R_s(f,p)}{\partial p} \tag{48a}$$
$$= |H(f)|^2 \mathrm{mmse}(|H(f)|^2 p) - |G(f)|^2 \mathrm{mmse}(|G(f)|^2 p) \tag{48b}$$
$$\le |H(f)|^2 \le M_H^2. \tag{48c}$$

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

$$|R_s(f,p) - R_s(f,p')| \le M_H^2 |p-p'|, \tag{49}$$

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

$$\frac{\partial \mathcal{I}(s; H\sqrt{p}s+n)}{\partial |H|^2 p} = 2p|H|\mathrm{mmse}(|H|^2 p), \tag{50}$$

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

$$\begin{aligned} &|R_s(f,p) - R_s(f',p)| \\ &\le |\mathcal{I}(s;H(f)\sqrt{p}s+n) - \mathcal{I}(s;H(f')\sqrt{p}s+n)| \\ &\quad + |\mathcal{I}(s;G(f)\sqrt{p}s+n) - \mathcal{I}(s;G(f')\sqrt{p}s+n)| \\ &\le 2pM_H \big|H(f) - H(f')\big| + 2pM_G \big|G(f) - G(f')\big| \\ &\le (2pM_H L_H + 2pM_G L_G)\,|f - f'|. \end{aligned} \tag{51}$$

By (49) and (51), we obtain that

$$\begin{aligned} &|R_s(f,p) - R_s(f',p')| \\ =&|R_s(f,p) - R_s(f,p') + R_s(f,p') - R_s(f',p')| \\ \le&|R_s(f,p) - R_s(f,p')| + |R_s(f,p') - R_s(f',p')| \\ \le&M_H^2|p-p'| + (2pM_H L_H + 2pM_G L_G)|f-f'| \end{aligned}$$

$$\leq \max\{M_H^2, 2pM_HL_H + 2pM_GL_G\}\Big(|p-p'|+|f-f'|\Big), \tag{52}$$

which is the same as (21) in which

$$L_R = \max\{M_H^2, 2pM_HL_H + 2pM_GL_G\}.$$

Appendix II: Proof of Proposition 3

The objective function of problem (24) can be rewritten as follows:

$$R_s^{AN} = \frac{1}{N}\sum_{i=1}^{N}\Big[\mathcal{I}\big(s_i; \frac{H_i\sqrt{p_i}}{\sqrt{|H_i|^2\sigma_{a,i}^2+1}}s_i+\tilde{w}_i\big) - \mathcal{I}\big(s_i; \frac{G_i\sqrt{p_i}}{\sqrt{|G_i|^2\sigma_{a,i}^2+1}}s_i+\tilde{v}_i\big)\Big]^+ \tag{53}$$

where $\tilde{w}_i = (H_i\hat{a}_i + w_i)/\sqrt{|H_i|^2\sigma_{a,i}^2+1}$ and $\tilde{v}_i = (G_i\hat{a}_i + v_i)/\sqrt{|G_i|^2\sigma_{a,i}^2+1}$ 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 $\{p_i^\star\}$ and $\{\sigma_a^{2\star}\}$, must satisfy the following KKT necessary conditions [42]:

$$\frac{\partial R_s^{AN}}{\partial p_i^\star} - \frac{1}{N}\lambda + \eta_i = 0, \quad i = 1,2,...,N, \tag{54a}$$

$$\frac{\partial R_s^{AN}}{\partial \sigma_{a,i}^{2\star}} - \frac{1}{N}\lambda + \mu_i = 0, \quad i = 1,2,...,N, \tag{54b}$$

$$\lambda\left(P - \frac{1}{N}\sum_{i=1}^{N}p_i^\star - \frac{1}{N}\sum_{i=1}^{N}\sigma_{a,i}^{2\star}\right) = 0, \tag{54c}$$

$$\mu_i\sigma_{a,i}^{2\star} = 0, \quad \eta_i p_i^\star = 0, \quad i = 1,2,...,N, \tag{54d}$$

$$\lambda \geq 0,\ \mu_i \geq 0,\ \eta_i \geq 0,\ i=1,2,...,N, \tag{54e}$$

$$\sigma_{a,i}^{2\star} \geq 0,\ p_i^\star \geq 0. \tag{54f}$$

If $|H_i| \leq |G_i|$ for some i, then $\frac{|H_i|^2p_i}{|H_i|^2\sigma_{a,i}^2+1} \leq \frac{|G_i|^2p_i}{|G_i|^2\sigma_{a,i}^2+1}$, and thus

$$\left[\mathcal{I}\left(s_i; \frac{H_i\sqrt{p_i}}{\sqrt{|H_i|^2\sigma_{a,i}^2+1}}s_i+\tilde{w}_i\right) - \mathcal{I}\left(s_i; \frac{G_i\sqrt{p_i}}{\sqrt{|G_i|^2\sigma_{a,i}^2+1}}s_i+\tilde{v}_i\right)\right]^+ = 0 \tag{55}$$

regardless of the values of $p_i^\star$ and $\sigma_{a,i}^{2\star}$. Therefore, we can assume that $|H_i| > |G_i| > 0$ for all $i = 1,\ldots,N$, without loss of generality[3].

We use contradiction to prove the result. Suppose that

$$\sigma_{a,i}^{2\star} > 0 \tag{56}$$

for some i. Then it is necessary that $p_i^\star > 0$; otherwise we end up with (55) again, and thus $\sigma_{a,i}^{2\star} = 0$ is also a feasible solution. With $\sigma_{a,i}^{2\star} > 0$ and $p_i^\star > 0$, we have $\mu_i = 0$ and $\eta_i = 0$ by (54d). By (54a), (54b) and (54e), we further obtain

$$\frac{\partial R_s^{AN}}{\partial p_i^\star} = \frac{\partial R_s^{AN}}{\partial \sigma_{a,i}^{2\star}} = \frac{1}{N}\lambda \geq 0. \tag{57}$$

By [49, Equation (22)] and by using the chain rule, one can show that

$$\frac{\partial R_s^{AN}}{\partial p_i^\star} = \frac{1}{N}\left[\frac{|H_i|^2\mathcal{E}_{b,i}(p_i^\star,\sigma_{a,i}^{2\star})}{|H_i|^2\sigma_{a,i}^{2\star}+1} - \frac{|G_i|^2\mathcal{E}_{e,i}(p_i^\star,\sigma_{a,i}^{2\star})}{|G_i|^2\sigma_{a,i}^{2\star}+1}\right] \geq 0, \tag{58}$$

$$\frac{\partial R_s^{AN}}{\partial \sigma_{a,i}^{2\star}} = \frac{1}{N}\left[\frac{|G_i|^4p_i^\star\mathcal{E}_{e,i}(p_i^\star,\sigma_{a,i}^{2\star})}{(|G_i|^2\sigma_{a,i}^{2\star}+1)^2} - \frac{|H_i|^4p_i^\star\mathcal{E}_{b,i}(p_i^\star,\sigma_{a,i}^{2\star})}{(|H_i|^2\sigma_{a,i}^{2\star}+1)^2}\right] \geq 0, \tag{59}$$

where $\mathcal{E}_{b,i}(p_i^\star,\sigma_{a,i}^{2\star})$ and $\mathcal{E}_{e,i}(p_i^\star,\sigma_{a,i}^{2\star})$, similar to (10), are given by

$$\mathcal{E}_{b,i}(p_i^\star,\sigma_{a,i}^{2\star}) = E\Big[\Big|s_i - E\big[s_i|H_i\sqrt{p_i^\star}s_i + H_i\hat{a}_i + w_i\big]\Big|^2\Big] \geq 0,$$

$$\mathcal{E}_{e,i}(p_i^\star,\sigma_{a,i}^{2\star}) = E\Big[\Big|s_i - E\big[s_i|G_i\sqrt{p_i^\star}s_i + G_i\hat{a}_i + v_i\big]\Big|^2\Big] \geq 0.$$

Combining $\frac{|H_i|^2p_i^\star}{|H_i|^2\sigma_{a,i}^{2\star}+1} > \frac{|G_i|^2p_i^\star}{|G_i|^2\sigma_{a,i}^{2\star}+1} > 0$ (due to $|H_i| > |G_i|$) and (58) yields

$$\frac{|H_i|^2p_i^\star}{(|H_i|^2\sigma_{a,i}^{2\star}+1)}\frac{|H_i|^2}{(|H_i|^2\sigma_{a,i}^{2\star}+1)}\mathcal{E}_{b,i}(p_i^\star,\sigma_{a,i}^{2\star}) > \frac{|G_i|^2p_i^\star}{(|G_i|^2\sigma_{a,i}^{2\star}+1)}\frac{|G_i|^2}{(|G_i|^2\sigma_{a,i}^{2\star}+1)}\mathcal{E}_{e,i}(p_i^\star,\sigma_{a,i}^{2\star}), \tag{60}$$

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

[3]As $|G_i| = 0$ means no eavesdropper, there is no need to use artificial noise, i.e., $\sigma_{a,i}^{2\star} = 0$.

Acknowledgment

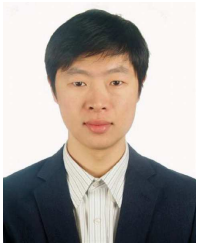
The authors would like to thank Fei He, Nicola Laurenti, Wei-Chiang Li, Kun-Yu Wang, and Shidong Zhou for valuable discussions to improve the quality of this paper.

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. Lekkas, *Wireless Security: Models, Threats, and Solutions*, 3rd Ed. McGraw-Hill Professional, 2004.
[3] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *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.
[8] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," in *Proc. 44th Allerton Conf.*, Sep. 2006, pp. 841–848.
[9] Y. Liang and H. V. Poor, "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, June 2008.
[10] E. Jorswieck and A. Wolf, "Resource allocation for the wire-tap multi-carrier broadcast channel," in *Proc. Int. Workshop Multiple Access Commun. Conf.*, June 2008, pp. 1–6.
[11] Z. Li, W. Trappe, and R. Yates, "Secret communication via multi-antenna transmission," in *Proc. Conf. Inf. Science Syst.*, Mar. 2007, pp. 905–910.

[12] 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.
[13] 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.
[14] ——, "Secure transmission with multiple antennas II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.
[15] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, June 2008.
[16] 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-aided approach," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1202–1216, Mar. 2011.
[17] 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.
[18] 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.
[19] 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.*, Nov. 2011.
[20] X. He, A. Khisti, and A. Yener, "MIMO multiple access channel with an arbitrarily varying eavesdropper," [Online]. Available: http://arxiv.org/pdf/1203.1376v1
[21] M. Bloch, J. Barros, and M. R. D. Rodrigues, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory, Special Issue on Information-Theoretic Security*, vol. 54, no. 6, pp. 2515–2534, Dec. 2008.
[22] 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.
[23] J. W. Wallace and R. K. Sharma, "Automatic secret keys from reciprocal MIMO wireless channels: Measurement and analysis," *IEEE Trans. Inf. Forens. Security*, vol. 5, no. 3, pp. 381–392, 2010.
[24] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.
[25] L. Lai, Y. Liang, and H. V. Poor, "A unified framework for key agreement over wireless fading channels," *IEEE Trans. Inf. Forens. Security*, vol. 7, no. 2, pp. 480–490, Apr. 2012.
[26] S. G. Wilson, *Digital Modulation and Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[27] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2599–2612, July 2012.
[28] S. Bashar, Z. Ding, and C. Xiao, "On secrecy rate analysis of MIMO wiretap channels driven by finite-alphabet input," *IEEE Trans. Commun.*, vol. 60, pp. 3816–3825, Dec. 2012.
[29] F. Renna, N. Laurenti, and H. V. Poor, "Physical-layer secrecy for OFDM transmissions over fading channels," *IEEE Trans. Inf. Forens. Security*, vol. 7, no. 4, pp. 1354–1367, Aug. 2012.
[30] Z. Li, R. Yates, and W. Trappe, "Achieving secret communication for fast Rayliegh fading channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2792–2799, Sep. 2010.
[31] M. Rodrigues, A. S. Baruch, and M. Bloch, "On Gaussian wiretap channels with M-PAM inputs," in *Proc. European Wireless Conf.*, Apr. 2010, pp. 774–781.
[32] 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
[33] S. Bashar, Z. Ding, and C. Xiao, "On the secrecy rate of multi-antenna wiretap channel under finite-alphabet input," *IEEE Commun. Lett.*, vol. 15, pp. 527–529, May 2011.
[34] D. Guo, S. Shamai (Shitz), and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1283, Apr. 2005.
[35] H. Jeffreys and B. S. Jeffreys, *Methods of Mathematical Physics*, 3rd Ed. Cambridge, UK: Cambridge University Press, 1988.
[36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ: John Wiley & Sons, Inc., 2006.
[37] L. Boyd and S. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
[38] Z. Luo and S. Zhang, "Duality gap estimation and polynomial time approximation for optimal spectrum management," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2675–2689, July 2009.
[39] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill Book Company, 1977.
[40] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 29–48, May 2000.
[41] S. A. A. Fakoorian and A. L. Swindlehurst, "Solutions for the MIMO Gaussian wiretap channel with a cooperative jammer," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 381–392, Oct. 2011.
[42] D. P. Bertsekas, *Nonliear Programming*. Belmont, MA: Athena Scientific, 1999.
[43] D. Guo, Y. Wu, S. Shamai (Shitz), and S. Verdú, "Estimation in Gaussian noise: Properties of the minimum mean-square error," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2371–2385, Apr. 2011.
[44] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," June 2009, [Online]. Available: http://stanford.edu/~boyd/cvx
[45] J. Strum, "Using SeDuMi 1.02: A MATLAB toolbox for optimization over symmetric cones," *Opt. Methods Software*, vol. 11–12, pp. 625–653, July 1999.
[46] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical program," *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.
[47] A. Lozano, A. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.
[48] A. Garnaev and W. Trappe, "An eavesdropping game with SINR as an objective function," in *Proc. Int. Conf. Security Privacy Commun. Netw.*, Sep. 2009.
[49] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, Jan. 2006.

Haohao Qin (S'13) received her B.S. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 2008, and is currently pursuing the Ph.D. degree in the Wireless and Mobile Communication Technology R&D Center, Research Institute of Information Technology, Tsinghua University, Beijing, China. From April 2011 to October 2011, she was a visiting research student at the Wireless Communications & Signal Processing (WCSP) Lab of National Tsing Hua University, Hsinchu, Taiwan, China. Her research interests include signal processing, wireless communication, and optimization.



Yin Sun (S'08-M'11) received the B. Eng. degree and Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2006 and 2011, respectively.

He is currently a Post-doctoral Researcher at the Ohio State University. His research interests include probability theory, optimization, information theory, and wireless communications.

Dr. Sun received the Tsinghua University Outstanding Doctoral Dissertation Award in 2011.

Tsung-Hui Chang (S'07-M'08) received the B.S. degree in electrical engineering and the Ph.D. degree in communications engineering from National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2003 and 2008, respectively. Since September 2012, he has been with the Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan, as an Assistant Professor. Before joining NTUST, he held research positions with NTHU (2008–2011) and the University of California at Davis, CA (2011–2012). He was also a visiting scholar of the University of Minnesota, Twin Cities, MN, and the Chinese University of Hong Kong. His research interests vary widely, from signal processing problems in wireless communications and smart grid, to convex optimization methods and their applications.

Xiang Chen (S'02-M'07) received both the B.E. and Ph.D. degrees from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2002 and 2008, respectively.

From July 2008 to July 2012, he was with the Wireless and Mobile Communication Technology R&D Center (Wireless Center), Research Institute of Information Technology at Tsinghua University. Since August 2012, he has been serving as an assistant researcher at the Aerosapce Center, School of Aerospace, Tsinghua University, Beijing, China. From July 2005 to August 2005, he was an internship student at the Audio Signal Group of Multimedia Laboratories, NTT DoCoMo R&D, in YRP, Japan. From September 2006 to April 2007, he was a visiting research student at the Wireless Communications & Signal Processing (WCSP) Lab of National Tsing Hua University, Hsinchu, Taiwan. Dr. Chen's research interests mainly focus on statistical signal processing, digital signal processing, software radio, and wireless communications.

Chong-Yung Chi (S'83-M'83-SM'89) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, California, in 1983.

From 1983 to 1988, he was with the Jet Propulsion Laboratory, Pasadena, California. He has been a Professor with the Department of Electrical Engineering at National Tsing Hua University, Hsinchu, Taiwan, since 1989 and the Institute of Communications Engineering (ICE) since 1999 (he was also the Chairman of ICE from 2002–2005). He has published more than 200 technical papers, including more than 70 journal papers (mostly in the IEEE TRANSACTIONS ON SIGNAL PROCESSING), two book chapters, and more than 130 peer-reviewed conference papers, as well as a graduate-level textbook, *Blind Equalization and System Identification*, Springer-Verlag, 2006. His current research interests include signal processing for wireless communications, convex analysis and optimization for blind source separation, and biomedical and hyperspectral image analysis.

Dr. Chi is a senior member of IEEE. He has been a Technical Program Committee member for many IEEE sponsored and co-sponsored workshops, symposia, and conferences on signal processing and wireless communications, including Co-organizer and General Co-chairman of the 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Co-Chair of Signal Processing for the Communications (SPC) Symposium, ChinaCOM (2008), and Lead Co-Chair of the SPC Symposium, ChinaCOM (2009). He was an Associate Editor (AE) of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (5/2001–4/2006), the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II (1/2006–12/2007), the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I (1/2008–12/2009), and AE of IEEE SIGNAL PROCESSING LETTERS (6/2006–5/2010). He has been a member of the Editorial Board of the *EURASIP Signal Processing Journal* (6/2005–5/2008) as well as an editor (7/2003–12/2005), and was a Guest Editor (2006) of the *EURASIP Journal on Applied Signal Processing*. He was also a member of the IEEE Signal Processing Committee on "Signal Processing Theory and Methods" (2005–2010). Currently, he is a member of the IEEE Signal Processing Committee on "Signal Processing for Communications and Networking," a member of the IEEE Signal Processing Committee on "Sensor Array and Multichannel," and an AE of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Ming Zhao (M'98) received the B.S. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1993 and 1998, respectively. He has been on the faculty of Tsinghua University since 1998. He is currently a professor. His research interests are in the area of wireless digital communications, including information theory, channel coding, multi-user detection, statistical signal processing, estimation theory, spread-spectrum communications, and multiple antenna systems.

Jing Wang (M'99) received the B.S. and M.S. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1983 and 1986, respectively. He has been on the faculty of Tsinghua University since 1986. He currently is a Professor and the Vice-Dean of the Tsinghua National Laboratory for Information Science and Technology. His research interests are in the area of wireless digital communications, including modulation, channel coding, multiuser detection, and 2D RAKE receivers. He has published more than 100 conference and journal papers. He is a member of the Technical Group of the China 3G Mobile Communication R&D Project, and he serves as an expert of communication technology in the National 863 Program. Mr. Wang is also a member of the Radio Communication Committee of the Chinese Institute of Communications and a senior member of the Chinese Institute of Electronics.