IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 10, 2017, accepted February 4, 2017, date of publication February 23, 2017, date of current version April 24, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2671878

Secure Transmission With Aid of a Helper for MIMOME Network Having Finite Alphabet Inputs

KUO CAO[1], (Student Member, IEEE), YONGPENG WU[2], (Member, IEEE), YUEMING CAI[1], (Senior Member, IEEE), AND WEIWEI YANG[1], (Member, IEEE)

[1]College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China
[2]Institute for Communications Engineering, Technical University of Munich, 80333 Munich, Germany

Corresponding author: K. Cao (caokuo90@sina.cn)

This work was supported by the National Natural Science Foundation of China under Project 61371122 and Project 61471393.

ABSTRACT This paper considers secure transmission with the aid of a helper for finite alphabet signals in multiple-input–multiple-output multiple antenna eavesdropper networks, where the helper transmits a jamming signal along with the confidential message sent by the source node to confuse the eavesdropper. For the scenario in which the only statistical channel-state-information (CSI) of the eavesdropper links is available at the transmitter, the ergodic secrecy rate lacks closed-form expression, and the evaluation of the ergodic secrecy rate is computationally prohibitive. To address this problem, an accurate approximation of the ergodic secrecy rate is proposed to reduce the computational complexity. Utilizing this approximation of the ergodic secrecy rate, the joint optimization of precoding design and power allocation between the source and the helper is investigated to improve the ergodic secrecy rate. Furthermore, to achieve a tradeoff between computational complexity and performance, low-complexity schemes without iteration are proposed based on the analysis at extreme signal-to-noise ratio (SNR). In the low SNR regime, we prove that it is optimal to transmit confidential messages with full power, and the beamforming design is the first-order optimal precoder. At high SNR, we transform the problem of precoding design into a semi-definite programming problem, which can be efficiently solved by the interior-point method.

INDEX TERMS Ergodic secrecy rate, finite alphabet inputs, precoding design, jamming signal, power allocation.

I. INTRODUCTION

Information secrecy has been a fundamental problem in wireless networks due to the inherent openness of the wireless propagation channel. Traditionally, security is guaranteed by the key-based cryptographic technologies [1] above the physical layer. Nevertheless, it is becoming more difficult to implement the secret key distribution and management with the rapidly increasing number of wireless devices. Physical layer security, which secures the transmission without need of secret key, has attracted tremendous attention recently. The seminal work [2] on physical layer security was pioneered by Wyner, who introduced the wiretap channel and defined the concept of secrecy capacity, which is the maximum transmission rate from source to its intended receiver while the eavesdropper knows no information about confidential messages. With great potential to enhance legitimate user's reception and weaken the eavesdropper's reception, multiple-input multiple-output (MIMO) technique plays a significant role in improving the physical layer security. The secrecy capacity of MIMO network has been fully investigated in [3]–[8], and the results in [7] and [8] show that the capacity is achieved by Gaussian wiretap codes.

Though the secrecy capacity can be achieved when the signal is Gaussian distributed in the additive white Gaussian noise (AWGN) channel, the Gaussian codebook has never been realized in practical system for various reasons. The finite alphabet inputs, such as pulse amplitude modulation (PAM), phase shift keying (PSK) modulation, and quadrature amplitude modulation (QAM), serves as an efficient substitution which has been widely applied in practical system. The finite alphabet inputs depart significantly from Gaussian input, therefore, the linear precoding design with finite alphabet inputs has attracted great research interests [9]–[15]. The relation between mutual information and minimum mean-square error (MMSE) had been revealed in [9] and [10]. Utilizing this relationship, a power allocation based on generalized singular value decomposition (GSVD) was developed to secure the transmission [11]. Furthermore, necessary conditions for the optimal precoding matrix were presented in [12] and an iterative algorithm based on gradient

method was proposed to maximize the achievable secrecy rate. Especially, a jamming scheme, which utilizes additional power for transmitting jamming signals in the null space of the main channel, was proposed to enhance the ergodic secrecy rate [12]. The same idea was also shown in multiple-input-single-output-single antenna eavesdropper (MISOSE) wiretap networks [13]. The work of [14] studied the precoding design for cognitive radio network with statistical CSI. A global optimization approach based on the branch-and-bound method was developed to maximize an accurate approximation of the ergodic secrecy rate [14]. Based on channel reciprocal, the strict positive secrecy rate is realized in Rayleigh flat fading channel by utilizing M-ary PSK and orthogonal space-time block code [15].

Recently, some researches have considered using friendly helper to transmit jamming signals to deteriorate the eavesdropper's reception [16]–[19], which serve as an effective means to further enhance the achievable secrecy rate. In [16], the closed-form structure of the artificial noise covariance matrix was obtained to guarantee the secrecy rate larger or at least equal to the secrecy capacity of MIMO wiretap channel with no jamming signal. By using the relay to transmit jamming signal, the authors in [17] proposed a null-steering jamming scheme with optimal power allocation to improve the achievable secrecy rate of cooperative network. Furthermore, the optimal transmit weights for jamming signal with individual power constraint were provided in [18]. Considering the case where only partial information of CSI for the eavesdropper links is available at transmitter, J. Huang *et* al. [19] studied robust transmission design for multiple-input single-output (MISO) wiretap channel with aid of a helper. Based on the worst-case secrecy rate maximization, robust transmit covariance matrices and the optimal power allocation between the source and the helper were derived in [19] under the global power constraint. However, scarce works have considered secure transmission with a friendly helper for finite alphabet signals.

In this paper, we investigate secure transmission design for MIMO wiretap channel with a cooperative helper under finite alphabet inputs. We assume that the source and the helper have perfect instantaneous CSI of the desired user and only statistical CSI of the eavesdropper. To secure the transmission, the helper utilizes additional degrees of spatial freedom to transmit jamming signal to confuse the eavesdropper. Considering the global power constraint, this paper studies the joint optimization of precoding design and power allocation between the source and the helper for maximizing the approximated ergodic secrecy rate. Furthermore, low complexity suboptimal schemes are also provided based on the analysis at extreme SNR. In particular, our contributions are summarized as follows:

- Our work considers secure transmission for MIMO wiretap channels with aid of an external helper. For the scenario where only statistical CSI of eavesdropper's link is available, the evaluation of the ergodic secrecy rate for finite alphabet inputs is computationally prohibitive. To address this issue, an accurate approximation of the ergodic secrecy rate is provided.
- Considering the global power constraint, we derive the gradient of the approximated ergodic secrecy rate with respect to the precoding matrix and necessary conditions for the optimal power allocation between the source and the helper. Based on these results, a joint optimization algorithm is developed to find the optimal precoding matrix and power allocation between the source and the helper, which maximize the approximated ergodic secrecy rate.
- To further reduce the computational complexity, low-complexity schemes are proposed based on the analysis of the ergodic secrecy rate at extreme SNR. At low SNR, it is proved that to transmit confidential message with full power is the first order optimal. On the other hand, the precoding design for confidential message at high SNR can be transformed into a SDP problem, which can be solved efficiently by the standard optimization method. Numerical results show that our proposed schemes significantly improve the ergodic secrecy rate.

The rest of the work is organized as follows. Section II illustrates the system model. The approximated ergodic secrecy rate and problem formulation are introduced in Section III. Section IV studies the joint optimization problem in detail. Suboptimal schemes at extreme SNR are investigated in Section V. Section VI depicts the numerical results, and Section VII summarizes the main results.

Notation: Boldface uppercase letters, boldface lowercase letters, and italics denote matrices, column vectors, and scalars, respectively. The superscripts $(.)^T$, $(.)^*$ and $(.)^H$ stand for transpose, conjugate and conjugate transpose operations respectively. $||.||$ designates norm of vector; $Tr(.)$ and det(.) denote the trace and determinant of a matrix, respectively; $vec(.)$ and $\otimes$ denote vectorization and Kronecker product operation, respectively; $\mathbb{E}(.)$ denotes the statistical expectation with respect to its variable; $\mathbf{I}_N$ denotes an $N \times N$ identity matrix; $\mathcal{CN}(\mu, \sigma^2)$ denotes the circularly symmetric, complex Gaussian distribution with mean μ and variance σ^2; $\log_2(.)$ denotes the base two logarithm.

II. SYSTEM MODEL

Consider a MIMOME network with a source node (Alice), a destination (Bob), a helper (Helper), and an eavesdropper (Eve), as shown in Fig. 1. The number of antennas equipped at Alice, Bob, Helper, and Eve are N_a, N_b, N_h and N_e, respectively, where $N_h > N_b$. It is assumed that the instantaneous CSIs of links from Alice and Helper to Bob are available at Alice and Helper, but they only know the statistical distribution of their links to Eve [12], [19], [20].

While Alice sends its signal to its intended receiver, Helper transmits jamming signal $\sqrt{P_J}\mathbf{u}z$ to secure the transmission, where P_J is the power budget for jamming signal; $\mathbf{u} \in \mathbb{C}^{N_h \times 1}$ is the weighted vector for jamming signal, which satisfies $\mathbf{u}^H\mathbf{u} \leq 1$; z is a PSK signal, which is denoted

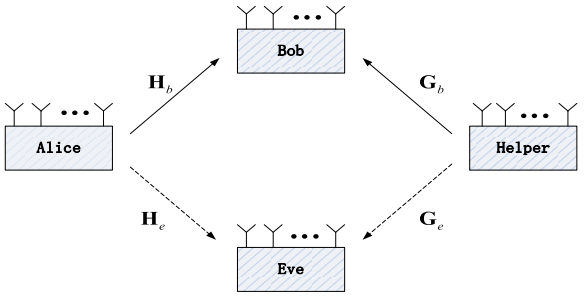

FIGURE 1. System model for MIMO wiretap channel with a helper.

as $z = e^{j\theta}$, where θ is a random variable within $[0, 2\pi]$. Then the received signal vectors at Bob, $\mathbf{y}_d$, and Eve, $\mathbf{y}_e$, are given by

$$\mathbf{y}_b = \sqrt{P_S}\mathbf{H}_b\mathbf{B}\mathbf{x} + \sqrt{P_J}\mathbf{G}_b\mathbf{u}z + \mathbf{n}_b, \tag{1}$$
$$\mathbf{y}_e = \sqrt{P_S}\mathbf{H}_e\mathbf{B}\mathbf{x} + \sqrt{P_J}\mathbf{G}_e\mathbf{u}z + \mathbf{n}_e, \tag{2}$$

where $\mathbf{H}_b \in \mathbb{C}^{N_b\times N_a}$ and $\mathbf{G}_b \in \mathbb{C}^{N_b\times N_h}$ are channel matrices from Alice and Helper to Bob respectively, $\mathbf{H}_e \in \mathbb{C}^{N_e\times N_a}$ and $\mathbf{G}_e \in \mathbb{C}^{N_e\times N_h}$ are channel matrices from Alice and Helper to Eve respectively; $\mathbf{n}_b \sim \mathcal{CN}(\mathbf{0}, \sigma_b^2\mathbf{I}_{N_b})$ and $\mathbf{n}_e \sim \mathcal{CN}(\mathbf{0}, \sigma_e^2\mathbf{I}_{N_e})$ are thermal noises at Bob and Eve, respectively, in which $\sigma_e^2 = \alpha\sigma_b^2$; P_S is the power budget for confidential message, $\mathbf{B} \in \mathbb{C}^{N_a\times N_a}$ is the precoding matrix for confidential meassage, which satisfies $\mathrm{Tr}\left(\mathbf{B}^H\mathbf{B}\right) \le 1$; $\mathbf{x} \in \mathbb{C}^{N_a\times 1}$ is the confidential signal vector, each element of which is drawn from equiprobable constellation set with cardinality M, such as PSK, QAM, with unit covariance, i.e., $\mathbb{E}_{\mathbf{x}}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_{N_a}$. Note that the case with global power constraint is considered, where P_S and P_J satisfy the power constraint $P_S + P_J \le P_0$, in which P_0 is the total power budget.

For channels from Alice and Helper to Eve, only statistical distribution are available at Alice and Helper. We model $\mathbf{H}_e$ and $\mathbf{G}_e$ as the doubly correlated fading MIMO channels [21], [22], which are given by

$$\mathbf{H}_e = \Psi_h^{1/2}\mathbf{H}_\mathbf{w}\Phi_h^{1/2}, \tag{3}$$
$$\mathbf{G}_e = \Psi_g^{1/2}\mathbf{G}_\mathbf{w}\Phi_g^{1/2}, \tag{4}$$

where $\mathbf{H}_\mathbf{w} \in \mathbb{C}^{N_e\times N_a}$ and $\mathbf{G}_\mathbf{w} \in \mathbb{C}^{N_e\times N_h}$ are complex random matrices, each element of which follows $\mathcal{CN}(0, 1)$; $\Psi_h \in \mathbb{C}^{N_e\times N_e}$ and $\Psi_g \in \mathbb{C}^{N_e\times N_e}$ represent receive positive correlation matrices; $\Phi_h \in \mathbb{C}^{N_a\times N_a}$ and $\Phi_g \in \mathbb{C}^{N_h\times N_h}$ denote transmit positive correlation matrices.

To simplify the problem, we adopt zero-force constraint on the jamming signal, that is, the jamming signal is transmitted in the null space of the channel $\mathbf{G}_b$. Denote $\mathbf{u} = \mathbf{V}_b\mathbf{f}$, where $\mathbf{V}_b \in \mathbb{C}^{N_h\times(N_h-N_b)}$ is an orthonormal basis of the null space of the channel $\mathbf{G}_b$, $\mathbf{f} \in \mathbb{C}^{(N_h-N_b)\times 1}$ fulfills $\mathbf{f}^H\mathbf{f} \le 1$. Therefore, the received signal at Bob and Eve can be equivalently written as

$$\mathbf{y}_b = \sqrt{P_S}\mathbf{H}_b\mathbf{B}\mathbf{x} + \mathbf{n}_b, \tag{5}$$
$$\mathbf{y}_e = \sqrt{P_S}\mathbf{H}_e\mathbf{B}\mathbf{x} + \sqrt{P_J}\mathbf{G}_e\mathbf{V}_b\mathbf{f}z + \mathbf{n}_e. \tag{6}$$

For simplicity, the signal received at Bob and Eve can be normalized as follows

$$\bar{\mathbf{y}}_b = \sqrt{\frac{P_S}{\sigma_b^2}}\mathbf{H}_b\mathbf{B}\mathbf{x} + \bar{\mathbf{n}}_b, \tag{7}$$
$$\bar{\mathbf{y}}_e = \sqrt{\frac{P_S}{\alpha\sigma_b^2}}\mathbf{R}_e^{-\frac{1}{2}}\mathbf{H}_e\mathbf{B}\mathbf{x} + \bar{\mathbf{n}}_e, \tag{8}$$

where $\bar{\mathbf{y}}_b = \sigma_b^{-1}\mathbf{y}_b$, $\bar{\mathbf{y}}_e = \left(\alpha\sigma_b^2\mathbf{R}_e\right)^{-\frac{1}{2}}\mathbf{y}_e$, $\bar{\mathbf{n}}_b \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_b})$ and $\bar{\mathbf{n}}_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_e})$ are effective noises at Bob and Eve, respectively, and $\mathbf{R}_e$ is given by

$$\begin{aligned}\mathbf{R}_e &= \frac{\mathbb{E}_{\mathbf{G}_e,\mathbf{n}_e}\left\{(\sqrt{P_J}\mathbf{G}_e\mathbf{u}z + \mathbf{n}_e)(\sqrt{P_J}\mathbf{G}_e\mathbf{u}z + \mathbf{n}_e)^H\right\}}{\alpha\sigma_b^2} \\ &= \frac{P_J}{\alpha\sigma_b^2}\mathrm{Tr}\left(\mathbf{V}_b^H\Phi_g\mathbf{V}_b\mathbf{f}\mathbf{f}^H\right)\Psi_g + \mathbf{I}_{N_e}.\end{aligned} \tag{9}$$

III. THE APPROXIMATED ERGODIC SECRECY RATE AND PROBLEM FORMULATION

Based on (7) and (8), the expressions of mutual information are given by [10]

$$\begin{aligned}I\left(\mathbf{x}; \bar{\mathbf{y}}_b\right) &= N_a\log_2 M \\ &\quad - \frac{1}{M^{N_a}}\sum_{m=1}^{M^{N_a}}\mathbb{E}_{\bar{\mathbf{n}}_b}\left\{\log_2\sum_{n=1}^{M^{N_a}}\exp\left(-f_{b,mn}\right)\right\},\end{aligned} \tag{10}$$

$$\begin{aligned}\mathbb{E}_{\mathbf{H}_e}\left\{I\left(\mathbf{x}; \bar{\mathbf{y}}_e\right)\right\} &= N_a\log_2 M \\ &\quad - \frac{1}{M^{N_a}}\sum_{m=1}^{M^{N_a}}\mathbb{E}_{\mathbf{H}_e}\mathbb{E}_{\bar{\mathbf{n}}_e}\left\{\log_2\sum_{n=1}^{M^{N_a}}\exp\left(-f_{e,mn}\right)\right\},\end{aligned} \tag{11}$$

where $f_{b,mn} = \left\|\sqrt{\frac{P_S}{\sigma_b^2}}\mathbf{H}_b\mathbf{B}\mathbf{e}_{mn} + \bar{\mathbf{n}}_b\right\|^2 - \|\bar{\mathbf{n}}_b\|^2$, $f_{e,mn} = \left\|\sqrt{\frac{P_S}{\alpha\sigma_b^2}}\mathbf{R}_e^{-\frac{1}{2}}\mathbf{H}_e\mathbf{B}\mathbf{e}_{mn} + \bar{\mathbf{n}}_e\right\|^2 - \|\bar{\mathbf{n}}_e\|^2$, $\mathbf{e}_{mn} = \mathbf{x}_m - \mathbf{x}_n$. Both $\mathbf{x}_m$ and $\mathbf{x}_k$ contain N_a symbols, which are taken from the equiprobable constellation set with cardinality M.

Combining (10) and (11), the ergodic secrecy rate is given by

$$C_{\mathrm{erg}} = \max\left\{I\left(\mathbf{x}; \bar{\mathbf{y}}_b\right) - \mathbb{E}_{\mathbf{H}_e}\left\{I\left(\mathbf{x}; \bar{\mathbf{y}}_e\right)\right\}; 0\right\}. \tag{12}$$

As can be seen from (10), (11) and (12), $I\left(\mathbf{x}; \bar{\mathbf{y}}_b\right)$ refers to calculate expectation over the effective noise vector $\bar{\mathbf{n}}_b$; Besides, the calculation of $\mathbb{E}_{\mathbf{H}_e}\left\{I\left(\mathbf{x}; \bar{\mathbf{y}}_e\right)\right\}$ needs to compute expectations over $\mathbf{H}_e$ and $\bar{\mathbf{n}}_e$. These expections refer to compute $2N_b + 2N_e\left(N_a + 1\right)$ integrals from $-\infty$ to $+\infty$ for calculating C_{erg} once, which leads to that the ergodic secrecy rate C_{erg} is difficult to evaluate. Though Monte Carlo random sampling can be utilized to estimate the ergodic secrecy rate, it is still time-consuming and complicated. To reduce the computational complexity, $I\left(\mathbf{x}; \bar{\mathbf{y}}_b\right)$ can be estimated by an accurate approximation $I_A\left(\mathbf{x}; \bar{\mathbf{y}}_b\right)$, which is given

by [23],

$$I_A(\mathbf{x};\bar{\mathbf{y}}_b) = N_a \log_2 M - \frac{1}{M^{N_a}} \sum_{m=1}^{M^{N_a}} \times \log_2 \sum_{n=1}^{M^{N_a}} \exp\left(-\frac{P_S \mathbf{e}_{mn}^H \mathbf{B}^H \mathbf{H}_b^H \mathbf{H}_b \mathbf{B} \mathbf{e}_{mn}}{2\sigma_b^2}\right). \quad (13)$$

On the other hand, we define $I_A(\mathbf{x};\bar{\mathbf{y}}_e)$ as the approximation of $\mathbb{E}_{\mathbf{H}_e}\{I(\mathbf{x};\bar{\mathbf{y}}_e)\}$, which is derived in [24],

$$\begin{aligned} &I_A(\mathbf{x};\bar{\mathbf{y}}_e) \\ &= N_a \log_2 M - \frac{1}{M^{N_a}} \sum_{m=1}^{M^{N_a}} \log_2 \sum_{n=1}^{M^{N_a}} \prod_{q=1}^{N_e} \left(1 + r_q g_{mn}\right)^{-1} \\ &= N_a \log_2 M - \frac{1}{M^{N_a}} \sum_{m=1}^{M^{N_a}} \log_2 \sum_{n=1}^{M^{N_a}} \frac{1}{\det(\mathbf{W}_{mn})}, \end{aligned} \quad (14)$$

with

$$\mathbf{W}_{mn} = \mathbf{I}_{N_e} + g_{mn} \Psi_h^{H/2} \mathbf{R}_e^{-1} \Psi_h^{1/2} \quad (15)$$

where r_q is the q-th eigenvalue of the matrix $\Psi_h^{H/2} \mathbf{R}_e^{-1} \Psi_h^{1/2}$, $g_{mn} = \frac{P_S}{2\alpha\sigma_b^2} \mathbf{e}_{mn}^H \mathbf{B}^H \Phi_h \mathbf{B} \mathbf{e}_{mn}$. By using $I_A(\mathbf{x};\bar{\mathbf{y}}_b)$ and $I_A(\mathbf{x};\bar{\mathbf{y}}_e)$ as alternatives, we derive an accurate approximation of the ergodic secrecy rate, which is given by

$$C_{\text{erg,A}} = \max\left\{I_A(\mathbf{x};\bar{\mathbf{y}}_b) - I_A(\mathbf{x};\bar{\mathbf{y}}_e); 0\right\}. \quad (16)$$

By replacing the ergodic secrecy rate with the approximated ergodic secrecy rate (16), we turn to maximize the approximated ergodic secrecy rate, which can be formulated as

$$\begin{aligned} \max_{\mathbf{f},\mathbf{B},(P_S,P_J)} \quad & C_{\text{erg,A}} \\ \text{s.t. } & \mathbf{f}^H \mathbf{f} \le 1, \\ & \text{Tr}\left(\mathbf{B}^H \mathbf{B}\right) \le 1, \\ & P_S + P_J \le P_0, \quad P_S > 0, \quad P_J \ge 0. \end{aligned} \quad (17)$$

As can be seen from (17), our design objectives consist of the optimization over $\mathbf{f}$ and $\mathbf{B}$, and the power allocation between P_S and P_J.

IV. JOINT OPTIMIZATION OF PRECODING DESIGN AND POWER ALLOCATION

Denote $\omega = \text{Tr}\left(\mathbf{V}_b^H \Phi_g \mathbf{V}_b \mathbf{f}\mathbf{f}^H\right)$. From (9), (13), (14), we know that $C_{erg,A}$ depends on $\mathbf{f}$ through ω. In the following proposition, we prove that $C_{\text{erg,A}}$ is a monotone increasing function of the variable ω.

Proposition 1: The gradient of $C_{\text{erg,A}}$ with respect to ω is given by

$$\nabla_\omega C_{\text{erg,A}} = \frac{P_S P_J \log_2 e}{2\alpha^2 \sigma_b^4} \text{Tr}\left(\mathbf{B}^H \Phi_h \mathbf{B} \Pi_e\right) \quad (18)$$

where

$$\Pi_e = \frac{1}{M^{N_a}} \sum_{m=1}^{M^{N_a}} \frac{\sum_{n=1}^{M^{N_a}} \frac{\text{Tr}\left(\mathbf{W}_{mn}^{-1} \Psi_h^{H/2} \mathbf{R}_e^{-1} \Psi_g \mathbf{R}_e^{-1} \Psi_h^{1/2}\right)}{\det(\mathbf{W}_{mn})} \mathbf{e}_{mn} \mathbf{e}_{mn}^H}{\sum_{n=1}^{M^{N_a}} \frac{1}{\det(\mathbf{W}_{mn})}}. \quad (19)$$

Note that Π_e is a positive definite matrix, we have $\nabla_\omega C_{\text{erg,A}} \ge 0$. Therefore, $C_{\text{erg,A}}$ is the monotone increasing function of ω.

Proof: See Appendix A.

Since that $C_{\text{erg,A}}$ is the monotone increasing function of ω, the problem of maximizing $C_{\text{erg,A}}$ over $\mathbf{f}$ is equivalent to the following problem

$$\begin{aligned} \max_{\mathbf{f}} \quad & \omega \\ \text{s.t. } & \mathbf{f}^H \mathbf{f} \le 1. \end{aligned} \quad (20)$$

Denote the maximum of ω as $\omega_{\max}$, which is the largest eigenvalue of the matrix $\left(\mathbf{V}_b^H \Phi_g \mathbf{V}_b\right)$. The optimal $\mathbf{f}$ is the unit-norm eigenvector of the matrix $\left(\mathbf{V}_b^H \Phi_g \mathbf{V}_b\right)$ corresponding to its largest eigenvalue $\omega_{\max}$.

With the optimal $\mathbf{f}$, $\mathbf{R}_e = \frac{P_J}{\alpha\sigma_b^2} \omega_{\max} \Psi_g + \mathbf{I}_{N_e}$, the problem (17) is identical to the following problem,

$$\begin{aligned} \max_{\mathbf{B},(P_S,P_J)} \quad & C_{\text{erg,A}} \\ \text{s.t. } & \text{Tr}\left(\mathbf{B}^H \mathbf{B}\right) \le 1, \\ & P_S + P_J \le P_0, \quad P_S > 0, \quad P_J \ge 0. \end{aligned} \quad (21)$$

It is obvious that $C_{\text{erg,A}}$ is now dependent only on the precoding matrix $\mathbf{B}$ and the power allocation between P_S and P_J. However, it is difficult to deal with the problem (21) in one step since that $\mathbf{B}$ and (P_S, P_J) are coupled. To address this problem, we develop a two-step optimization algorithm to find the optimal precoding matrix and power allocation between the source and the helper. Firstly, we maximize $C_{\text{erg,A}}$ through the precoding matrix $\mathbf{B}$ while fixing P_S and P_J. Secondly, we maximize $C_{\text{erg,A}}$ via power allocation between P_S and P_J while fixing the precoding matrix $\mathbf{B}$. Finally, the solution of the original problem (21) can be found by repeating the above two steps.

A. PRECODING DESIGN

First, we consider maximizing $C_{\text{erg,A}}$ through the precoding matrix $\mathbf{B}$ for a given P_S and P_J, which is formulated as

$$\begin{aligned} \max_{\mathbf{B}} \quad & C_{\text{erg,A}} \\ \text{s.t. } & \text{Tr}\left(\mathbf{B}^H \mathbf{B}\right) \le 1. \end{aligned} \quad (22)$$

Using the complex-valued matrix differentiation technique [25], the gradient of $C_{\text{erg,A}}$ with respect to $\mathbf{B}$ is obtained as follows

$$\nabla_{\mathbf{B}} C_{\text{erg,A}} = \frac{P_S \log_2 e}{2\sigma_b^2} \mathbf{H}_b^H \mathbf{H}_b \mathbf{B} \Omega_b - \frac{P_S \log_2 e}{2\alpha\sigma_b^2} \Phi_h \mathbf{B} \Omega_e, \quad (23)$$

with

$$\Omega_b = \frac{1}{M^{N_a}} \sum_{m=1}^{M^{N_a}} \frac{\sum_{n=1}^{M^{N_a}} \exp\left(-\frac{P_S \mathbf{e}_{mn}^H \mathbf{B}^H \mathbf{H}_b^H \mathbf{H}_b \mathbf{B} \mathbf{e}_{mn}}{2\sigma_b^2}\right) \mathbf{e}_{mn}\mathbf{e}_{mn}^H}{\sum_{n=1}^{M^{N_a}} \exp\left(-\frac{P_S \mathbf{e}_{mn}^H \mathbf{B}^H \mathbf{H}_b^H \mathbf{H}_b \mathbf{B} \mathbf{e}_{mn}}{2\sigma_b^2}\right)}, \tag{24}$$

and

$$\Omega_e = \frac{1}{M^{N_a}} \sum_{m=1}^{M^{N_a}} \frac{\sum_{n=1}^{M^{N_a}} \frac{\mathrm{Tr}\left(\mathbf{W}_{mn}^{-1} \Psi_h^{H/2} \mathbf{R}_e^{-1} \Psi_h^{1/2}\right)}{\det(\mathbf{W}_{mn})} \mathbf{e}_{mn}\mathbf{e}_{mn}^H}{\sum_{n=1}^{M^{N_a}} \frac{1}{\det(\mathbf{W}_{mn})}}. \tag{25}$$

Utilizing the gradient of $C_{\mathrm{erg,A}}$ with respect to $\mathbf{B}$, the optimal precoding matrix $\mathbf{B}$ to the problem (22) can be iteratively found using the gradient method. The precoding matrix $\mathbf{B}$ is updated as

$$\mathbf{B}(l+1) = \mathbf{B}(l) + \gamma \nabla_{\mathbf{B}} C_{\mathrm{erg,A}}\big|_{\mathbf{B}=\mathbf{B}(l)}, \tag{26}$$

where γ is a positive step size. When the obtained $\mathbf{B}(l+1)$ satisfies $\mathrm{Tr}\left\{(\mathbf{B}(l+1))^H \mathbf{B}(l+1)\right\} > 1$, it is reasonable to project $\mathbf{B}(l+1)$ into the feasible set through a normalization step: $\mathbf{B}(l+1) = \mathbf{B}^{(l+1)}/\sqrt{\mathrm{Tr}\left\{(\mathbf{B}(l+1))^H \mathbf{B}(l+1)\right\}}$.

B. POWER ALLOCATION

By fixing the precoding matrix $\mathbf{B}$, the problem (21) is identical to the power allocation between P_S and P_J for maximizing the approximated ergodic secrecy rate, which is written as

$$\begin{aligned} &\max_{(P_S,P_J)} \ C_{\mathrm{erg,A}} \\ &\text{s.t. } P_S + P_J \le P_0, \\ &\quad P_S > 0, \quad P_J \ge 0. \end{aligned} \tag{27}$$

Based on Karush-Kuhn-Tucker (KKT) analysis, the optimal power allocation to the problem (27) is obtained as follows.

Proposition 2: The optimal power allocation of the problem (27) is the solution to following equations

$$\begin{aligned} &\mathrm{Tr}\left(\mathbf{B}^H \mathbf{H}_b^H \mathbf{H}_b \mathbf{B} \Omega_b\right) \\ &\quad = \mathrm{Tr}\left(\mathbf{B}^H \Phi_h \mathbf{B}\left(\frac{1}{\alpha}\Omega_e + \frac{P_S \omega_{\max}}{\alpha^2 \sigma_b^2} \Pi_e\right)\right), \end{aligned} \tag{28}$$

$$P_s + P_J = P_0. \tag{29}$$

Note that if there is no solution to above equations, then the optimal power allocation is given by

$$(P_S, P_J) = (P_0, 0). \tag{30}$$

Proof: See Appendix B.

Consequently, a joint optimization algorithm can be developed to solve the problem (21) by combing the precoding design and power allocation. The details of the joint optimization algorithm are summarized in Algorithm 1 as follows

Algorithm 1 Joint Optimization Algorithm for Maximizing the Approximated Ergodic Secrecy Rate

1: Initialization: given a feasible $(P_S^{(1)}, P_J^{(1)})$, $\mathbf{B}^{(1)}$, and set $k = 1$.
2: Substituting $(P_S^{(k)}, P_J^{(k)})$ into problem (22), solve the problem (22) by using the gradient method and obtain the solution $\mathbf{B}^{(k+1)}$.
3: Substituting $\mathbf{B}^{(k+1)}$ into (28), obtain $(P_S^{(k+1)}, P_J^{(k+1)})$ by deriving the solution to equations (28) and (29). If there is no solution to above equations, then set $(P_S^{(k+1)}, P_J^{(k+1)}) = (P_0, 0)$.
4: $k = k + 1$, go to step 2 until convergence or the predefined number of iterations has been reached.

Note that the joint optimization algorithm is convergent due to that the sequence generated by the joint optimization algorithm is increasing and upper-bounded. To avoid the algorithm being stopped at local optimum, multiple initial points should be used and the one with the maximum is chosen as the final solution. Since that the performance of the joint optimization algorithm depends heavily on initial points, it motives us to design low-complexity and high-performance schemes without iteration.

V. LOW-COMPLEXITY PRECODING DESIGN AND POWER ALLOCATION

In this section, we resort to low-complexity scheme designs without iteration in the extreme SNR regime.

A. SUBOPTIMAL SCHEME DESIGN AT LOW SNR

In the low SNR regime, i.e., $P_0/\sigma_b^2 \to 0$, the precoding design and the power allocation between the source and the helper are given as follows.

Proposition 3: The first order optimal power allocation between the source and the helper at low SNR is the same with (30) and the precoding matrix $\mathbf{B}$ is given by

$$\mathbf{B} = \phi(\lambda_{\max}) \left[\mathbf{v}_{\max} \quad \mathbf{0}_{N_a \times (N_a - 1)}\right], \tag{31}$$

where $\mathbf{v}_{\max}$ is the unit-norm eigenvector of the matrix $\left(\mathbf{H}_b^H \mathbf{H}_b - \frac{\mathrm{Tr}(\Psi_h)}{\alpha} \Phi_h\right)$ corresponding to its largest eigenvalue $\lambda_{\max}$; $\phi(\cdot)$ is the Heaviside step function where $\phi(\lambda) = 1$ if $\lambda > 0$, otherwise $\phi(\lambda) = 0$.

Proof: See Appendix C.

B. SUBOPTIMAL SCHEME DESIGN AT HIGH SNR

In the high SNR regime, i.e., $P_0/\sigma_b^2 \to \infty$, it is reasonable to allocate appropriate power with jamming signal to deteriorate the signal-to-noise-plus-interference ratio (SINR) performance of Eve. However, it is difficult to obtain the optimal $\mathbf{B}$ and power allocation between P_S and P_J since that $\mathbf{B}$ and (P_S, P_J) are coupled. To decouple this problem, we utilize the precoding matrix $\mathbf{B}$ to enhance $I(\mathbf{x}; \bar{\mathbf{y}}_d)$ and decrease the average SINR performance of Eve concurrently.

TABLE 1. The number of matrix multiplication steps for calculating the mutual information once based on Monte Carlo Sampling method.

$I(\mathbf{x};\bar{\mathbf{y}}_b)$	$I_A(\mathbf{x};\bar{\mathbf{y}}_b)$	$\mathbb{E}_{\mathbf{H}_e}\{I(\mathbf{x};\bar{\mathbf{y}}_e)\}$	$I_A(\mathbf{x};\bar{\mathbf{y}}_e)$
$\mathcal{O}\left(N_{\text{noise}}M^{2N_a}\right)$	$\mathcal{O}\left(M^{2N_a}\right)$	$\mathcal{O}\left(N_{\text{noise}}N_{\text{channel}}M^{2N_a}\right)$	$\mathcal{O}\left(M^{2N_a}\right)$

The average SINR of Eve is denoted as SINR_E, which is given by

$$\begin{aligned}\text{SINR}_E &= \frac{P_S}{\alpha\sigma_b^2}\mathbb{E}_{\mathbf{H}_e}\left\{\text{Tr}\left(\mathbf{B}^H\mathbf{H}_e^H\mathbf{R}_e^{-1}\mathbf{H}_e\mathbf{B}\right)\right\}\\ &= \frac{P_S}{\alpha\sigma_b^2}\text{Tr}\left(\Psi_h\mathbf{R}_e^{-1}\right)\text{Tr}\left(\mathbf{B}^H\Phi_h\mathbf{B}\right). \qquad (32)\end{aligned}$$

As pointed out in [23], the problem of maximizing $I(\mathbf{x};\bar{\mathbf{y}}_d)$ in the high SNR regime is identical to the problem of maximizing the minimum distance of input vectors. Here we denote $d_{\min}$ as the minimum distance of input vectors of Bob, which is given by,

$$d_{\min} = \min_{m\neq n}\frac{P_S}{\sigma_b^2}\mathbf{e}_{mn}^H\mathbf{B}^H\mathbf{H}_b^H\mathbf{H}_b\mathbf{B}\mathbf{e}_{mn}. \qquad (33)$$

To enhance $I(\mathbf{x};\bar{\mathbf{y}}_d)$ and decrease SINR_E at high SNR concurrently, it is reasonable to maximize the ratio between $d_{\min}$ and SINR_E through $\mathbf{B}$, which can be formulated as

$$\begin{aligned}\max_{\mathbf{B}}\ &\min_{m\neq n}\frac{\alpha}{\text{Tr}\left(\Psi_h\mathbf{R}_e^{-1}\right)}\frac{\mathbf{e}_{mn}^H\mathbf{B}^H\mathbf{H}_b^H\mathbf{H}_b\mathbf{B}\mathbf{e}_{mn}}{\text{Tr}\left(\mathbf{B}^H\Phi_h\mathbf{B}\right)}\\ &\text{s.t. Tr}\left(\mathbf{B}^H\mathbf{B}\right)\leq 1. \qquad (34)\end{aligned}$$

Note that $\frac{\alpha}{\text{Tr}\left(\Psi_h\mathbf{R}_e^{-1}\right)}$ is irrelevant with the precoding matrix $\mathbf{B}$, the problem (34) can be equivalent to the following problem

$$\begin{aligned}\max_{\mathbf{B}}\ &\min_{m\neq n}\frac{\mathbf{e}_{mn}^H\mathbf{B}^H\mathbf{H}_b^H\mathbf{H}_b\mathbf{B}\mathbf{e}_{mn}}{\text{Tr}\left(\mathbf{B}^H\Phi_h\mathbf{B}\right)}\\ &\text{s.t. Tr}\left(\mathbf{B}^H\mathbf{B}\right)\leq 1. \qquad (35)\end{aligned}$$

Denote $\mathbf{b}=\text{vec}(\mathbf{B})$, $\mathbf{p}=\mathbf{D}^{1/2}\mathbf{b}$, $\mathbf{D}=(\mathbf{I}_{N_a}\otimes\Phi_h)$ and $\mathbf{q}=\mathbf{p}/||\mathbf{p}||$, then the problem (35) is identical to the following problem.

Proposition 4: The problem (35) can be transformed into the following problem

$$\begin{aligned}&\min_{\mathbf{Q}}\ -t\\ &\text{s.t. Tr}\left\{\mathbf{D}^{-\frac{H}{2}}\left(\mathbf{E}_{mn}^T\otimes\mathbf{H}_b^H\mathbf{H}_b\right)\mathbf{D}^{-\frac{1}{2}}\mathbf{Q}\right\}\geq t \quad m<n,\\ &\quad\text{Tr}(\mathbf{Q})=1,\\ &\quad\mathbf{Q}\succeq\mathbf{0},\quad \text{rank}(\mathbf{Q})=1, \qquad (36)\end{aligned}$$

where $\mathbf{E}_{mn}=\mathbf{e}_{mn}\mathbf{e}_{mn}^H$, $\mathbf{Q}=\mathbf{q}\mathbf{q}^H$, $\mathbf{Q}\succeq\mathbf{0}$ represents that $\mathbf{Q}$ is symmetric positive semi-definite.

Proof: See Appendix D.

Dropping the rank-one constraint, the problem (36) can be formulated as

$$\begin{aligned}&\min_{\mathbf{Q}}\ -t\\ &\text{s.t. Tr}\left\{\mathbf{D}^{-\frac{H}{2}}\left[\mathbf{E}_{mn}^T\otimes\mathbf{H}_b^H\mathbf{H}_b\right]\mathbf{D}^{-\frac{1}{2}}\mathbf{Q}\right\}\geq t \quad m<n,\\ &\quad\text{Tr}(\mathbf{Q})=1,\quad \mathbf{Q}\succeq\mathbf{0}, \qquad (37)\end{aligned}$$

which is a typical SDP problem and can be efficiently solved by using interior-point method [26]. Note that the solution $\mathbf{Q}$ obtained by solving SDP problem (37) is generally not of rank one. If the obtained $\mathbf{Q}$ is of rank one, then its principal eigenvector is the optimal solution to the original problem (36). Otherwise, randomization techniques [27] can be utilized to provide an near-optimal solution of the problem (36). When $\mathbf{q}$ is obtained, we have

$$\mathbf{b}=\frac{\mathbf{D}^{-1/2}\mathbf{q}}{\sqrt{\mathbf{q}^H\mathbf{D}^{-1}\mathbf{q}}}, \qquad (38)$$

$$\mathbf{B}=\text{reshape}(\mathbf{b},N_a,N_a). \qquad (39)$$

Finally, substituting the precoding matrix $\mathbf{B}$ into (28), the power allocation between P_S and P_J can be obtained by deriving the solution of equations (28) and (29).

VI. NUMERICAL RESULTS

In this section, numerical examples are provided to illustrate the efficacy of the proposed schemes. The thermal noise variance and α are set as $\sigma_b^2=1$ and $\alpha=1$, respectively. Channel matrices $\mathbf{H}_e$ and $\mathbf{G}_e$ follow the exponential correlation model, where the (a,b)-th element of the correlation matrice is given by [14], [24]

$$[\mathbf{T}(\rho)]_{a,b}=\rho^{|a-b|},\quad \rho\in[0,1). \qquad (40)$$

The channel statistics for $\mathbf{H}_e$ and $\mathbf{G}_e$ are given by $\mathbf{\Psi}_h=\mathbf{T}(0.95)$, $\mathbf{\Phi}_h=\mathbf{T}(0.5)$, $\mathbf{\Psi}_g=\mathbf{T}(0.8)$ and $\mathbf{\Phi}_g=\mathbf{T}(0.6)$, respectively. We denote N_{noise} and N_{channel} as the number of sample points used for calculating expectations over effective noises $(\bar{\mathbf{n}}_b,\bar{\mathbf{n}}_e)$ and random matrices $(\mathbf{H}_\mathbf{w},\mathbf{G}_\mathbf{w})$, respectively. In our simulations, we set $N_{\text{noise}}=1000$ and $N_{\text{channel}}=500$. Besides, channel matrices $\mathbf{H}_b$ and $\mathbf{G}_b$ are generated as independent identical distributed complex Gaussian random variable with zero mean and unit variance, and the simulation results are calculated based on the average of 500 channel realizations of $\mathbf{H}_b$ and $\mathbf{G}_b$.

Table 1 lists the number of matrix multiplication steps for calculating the mutual information once based on Monte Carlo sampling method. From Table 1, we obtain that the computational complexities of evaluating the ergodic secrecy rate C_{erg} and its approximation $C_{\text{erg,A}}$ are $\mathcal{O}\left(N_{\text{noise}}N_{\text{channel}}M^{2N_a}\right)$ and $\mathcal{O}\left(M^{2N_a}\right)$,

respectively. Therefore, employing the approximation $C_{\text{erg,A}}$ significantly reduces the computational complexity compared to maximizing the ergodic secrecy rate C_{erg} directly.

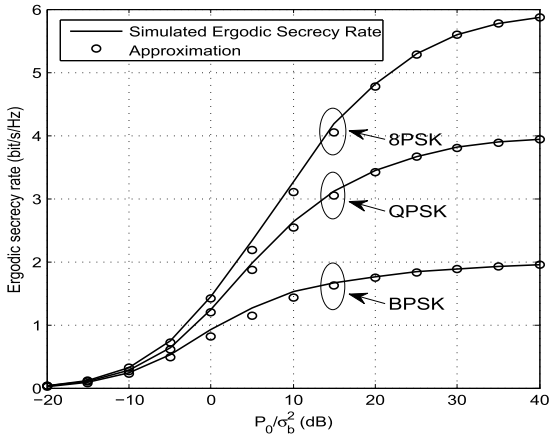

FIGURE 2. Ergodic secrecy rate of the joint optimization algorithm for different modulations when $N_a = 2$. $N_h = 3$, $N_b = 2$ and $N_e = 2$.

Fig. 2 presents the relationship between the simulated ergodic secrecy rate C_{erg} obtained by Monte Carlo sampling method and its approximation $C_{\text{erg,A}}$. As shown in Fig. 2, the approximation $C_{\text{erg,A}}$ exhibits excellent agreement with the simulated ergodic secrecy rate C_{erg} in the low and high total power regime, and it closes to the simulated ergodic secrecy rate C_{erg} in medium total power regime. Thus, $C_{\text{erg,A}}$ provides a very good approximation to the ergodic secrecy rate C_{erg} for the whole total power regime.

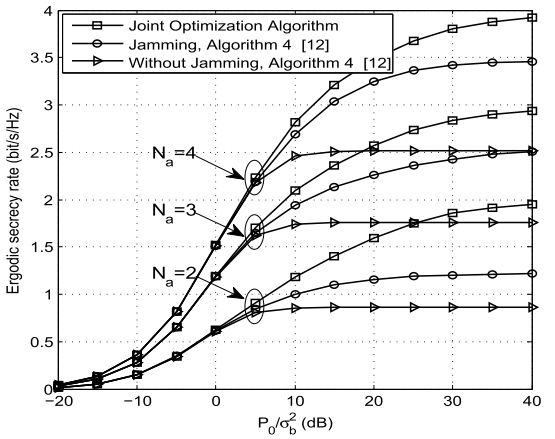

FIGURE 3. Ergodic secrecy rate comparison with different schemes and different number of antennas equipped at Alice under BPSK input when $N_h = 3$, $N_b = 1$ and $N_e = 2$.

Fig. 3 compares the ergodic secrecy rate with different schemes and different number of antennas equipped at Alice under BPSK input when $N_h = 3$, $N_b = 1$ and $N_e = 2$. Based on Fig. 3, we observe that increasing the number of antennas equipped at Alice improves the ergodic secrecy rate. In the low total power regime, the ergodic secrecy rate

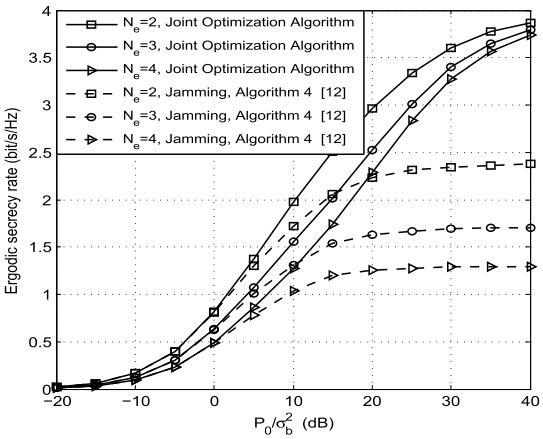

FIGURE 4. Ergodic secrecy rate comparison with different schemes and different number of antennas equipped at Eve under QPSK input when $N_a = 2$ $N_h = 2$, and $N_b = 1$.

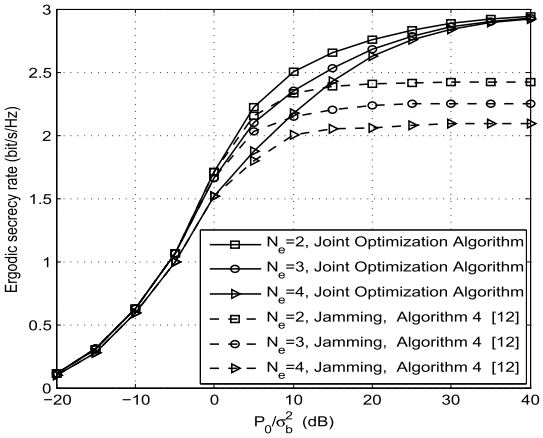

FIGURE 5. Ergodic secrecy rate comparison with different schemes and different number of antennas equipped at Eve under BPSK input when $N_a = 3$, $N_h = 3$, and $N_b = 2$.

achieved by the scheme with jamming is the same with that of Algorithm 4 [12] without jamming, which indicates that the optimal power allocation at low total power is to allocate full power for transmitting confidential message. At medium and high total power, the scheme with jamming further improves the ergodic secrecy rate compared to the scheme without jamming, it is because that the transmission rate of Eve is suppressed by the jamming signal. Besides, the joint optimization algorithm outperforms Algorithm 4 [12] and approaches $N_a \log_2 M$ bit/s/Hz at high total power.

In Fig. 4 and Fig. 5, we compare the ergodic secrecy rate with different schemes and different number of antennas equipped at Eve for various settings. It is obvious that the ergodic secrecy rate of all schemes decreases with the increasing number of antennas equipped at Eve. However, from Fig. 4 and Fig. 5, we observe that the joint optimization algorithm restores the ergodic secrecy rate to close to $N_a \log_2 M$ bit/s/Hz at high total power regardless of the number of antennas equipped at Eve.

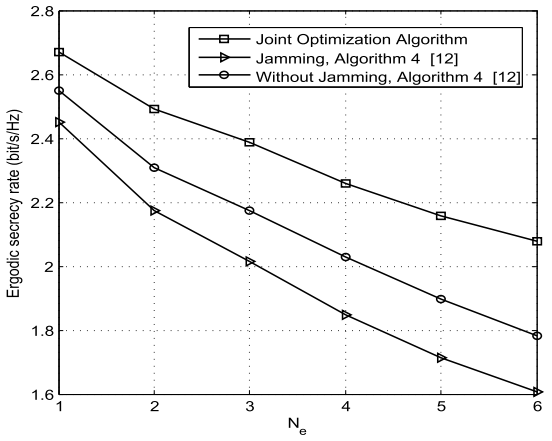

FIGURE 6. Ergodic secrecy rate of different schemes versus different number of antennas equipped at Eve under BPSK input when $N_a = 3$ $N_b = 2$, $N_h = 3$ and $P_0/\sigma_b^2 = 10$ dB.

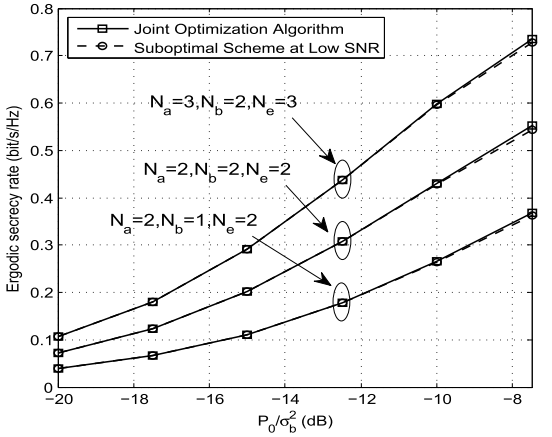

FIGURE 7. Ergodic secrecy rate comparison with different schemes and various settings in the low total power regime under BPSK input when $N_h = 3$.

Fig. 6 depicts the ergodic secrecy rate performance for different number of antennas equipped at Eve under BPSK input when $P_0/\sigma_b^2 = 10$ dB. With an appropriate power allocation between the source and the helper, the ergodic secrecy rate obtained by the joint optimization algorithm is higher than that of Algorithm 4 [12] for arbitrary number of antennas equipped at Eve, which is shown in Fig. 6.

Fig. 7 demonstrates the ergodic secrecy rate performance of different schemes in the lower total power regime with BPSK input. As can be seen from Fig. 7, the suboptimal scheme at low SNR and the joint optimization algorithm match exactly in the low total power regime, which suggests that it is optimal to transmit confidential message with full power at low SNR.

Fig. 8 shows the ergodic secrecy rate performance of different schemes for various settings at high total power with BPSK input. In Fig. 8, it can be observed that the ergodic secrecy rate performance of the suboptimal scheme at high SNR is very close to that of the joint optimization algorithm

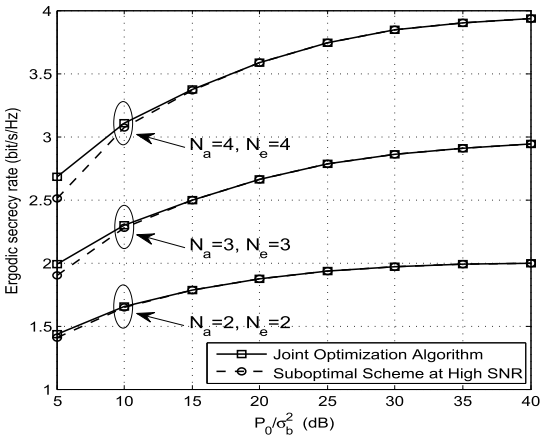

FIGURE 8. Ergodic secrecy rate comparison with different schemes and various settings in the high total power regime under BPSK input when $N_h = 3$ and $N_b = 2$.

in the high total power regime. Furthermore, the suboptimal scheme at high SNR ensures the ergodic secrecy rate to approach $N_a \log_2 M$ bit/s/Hz with the increasing total power. Therefore, the optimization of the precoder at high SNR can be equivalent to a semi-definite programming problem.

VII. CONCLUSIONS

In this paper, we have studied the joint optimization of precoding design and power allocation for MIMOME network with a friendly helper under finite alphabet inputs. Utilizing the statistical CSI of the eavesdropper's channel, an accurate approximation of the ergodic secrecy rate with closed-form expression was proposed in this paper. Besides, the gradient of the approximated ergodic secrecy rate with respect to the precoding matrix, and necessary conditions for the optimal power allocation were provided. Based on these results, a joint iterative algorithm was developed to maximize the approximated ergodic secrecy rate. To further reduce the computational complexity, suboptimal schemes without iteration were proposed at extreme SNR. Moreover, we proved that it is optimal to transmit confidential message with full power at low SNR, and transformed the problem of precoding design for confidential message at high SNR into a SDP problem. Simulation results show that our proposed schemes significantly improve the ergodic secrecy rate of MIMO wiretap channel compared to their conventional counterparts.

APPENDIX A
PROOF OF PROPOSITION 1

$I(\mathbf{x}; \bar{\mathbf{y}}_d)$ is independent with ω, therefore, we have $\nabla_\omega C_{\mathrm{erg,A}} = -\nabla_\omega I_{\mathrm{A}}(\mathbf{x}; \bar{\mathbf{y}}_e)$, where $\nabla_\omega I_{\mathrm{A}}(\mathbf{x}; \bar{\mathbf{y}}_e)$ can be expressed as

$$\nabla_\omega I_{\mathrm{A}}(\mathbf{x}; \bar{\mathbf{y}}_e) = \frac{\log_2 e}{M^{N_a}} \sum_{m=1}^{M^{N_a}} \frac{\sum_{n=1}^{M^{N_a}} \frac{1}{[\det(\mathbf{W}_{mn})]^2} \frac{\partial \det(\mathbf{W}_{mn})}{\partial \omega}}{\sum_{n=1}^{M^{N_a}} \frac{1}{\det(\mathbf{W}_{mn})}}. \quad (41)$$

With $\mathbf{W}_{mn}^H = \mathbf{W}_{mn}$, one can use the chain rule to obtain

$$\frac{\partial \det(\mathbf{W}_{mn})}{\partial \omega} = \mathrm{Tr}\left(\left(\frac{\partial \det(\mathbf{W}_{mn})}{\partial \mathbf{W}_{mn}}\right)^T \frac{\partial \mathbf{W}_{mn}}{\partial \omega}\right), \quad (42)$$

where

$$\frac{\partial \det(\mathbf{W}_{mn})}{\partial \mathbf{W}_{mn}} = \det(\mathbf{W}_{mn})\left(\mathbf{W}_{mn}^{-1}\right)^T, \quad (43)$$

$$\frac{\partial \mathbf{W}_{mn}}{\partial \omega} = -\frac{P_J g_{mn}}{\alpha\sigma_b^2}\Psi_h^{H/2}\mathbf{R}_e^{-1}\Psi_g\mathbf{R}_e^{-1}\Psi_h^{1/2}. \quad (44)$$

Substituting (42) into (41), we have

$$\nabla_\omega I_{\mathrm{A}}(\mathbf{x};\bar{\mathbf{y}}_e) = -\frac{P_S P_J \log_2 e}{2\alpha^2\sigma_b^4}\mathrm{Tr}\left(\mathbf{B}^H\Phi_h\mathbf{B}\Pi_e\right). \quad (45)$$

Keep in mind that $\nabla_\omega C_{\mathrm{erg,A}} = -\nabla_\omega I_{\mathrm{A}}(\mathbf{x};\bar{\mathbf{y}}_e)$, hence, the proof of ***Proposition 1*** is completed.

APPENDIX B
PROOF OF PROPOSITION 2

The problem (27) can be written as Lagrangian construction

$$L = -C_{\mathrm{erg,A}} - \eta(P - P_S - P_J) - \xi P_J. \quad (46)$$

Based on KKT analysis, necessary conditions for the optimal power allocation are given by

$$\nabla_{P_S} C_{\mathrm{erg,A}} = \eta, \quad (47)$$

$$\nabla_{P_J} C_{\mathrm{erg,A}} = \eta - \xi, \quad (48)$$

$$\eta(P - P_S - P_J) = 0, \quad (49)$$

$$\eta \geq 0, \quad (50)$$

$$\xi P_J = 0, \quad (51)$$

$$\xi \geq 0. \quad (52)$$

Note that the optimal power allocation occurs at $P_S + P_J = P_0$, it is because that for any given P_S, the optimal power of jamming signal is $P_J = P_0 - P_S$ which minimizes $I_{\mathrm{A}}(\mathbf{x};\bar{\mathbf{y}}_e)$. Thus, if $\xi > 0$, then $P_J = 0$ and $P_S = P_0$, which is obtained from (51); Otherwise, substituting $\xi = 0$ and (47) into (48), we can obtain that the optimal power allocation is the solution of the following equation,

$$\nabla_{P_S} C_{\mathrm{erg,A}} = \nabla_{P_J} C_{\mathrm{erg,A}}. \quad (53)$$

Following similar steps in Appendix A, we have

$$\nabla_{P_J} C_{\mathrm{erg,A}} = \frac{P_S\omega_{\max}\log_2 e}{2\alpha^2\sigma_b^4}\mathrm{Tr}\left(\mathbf{B}^H\Phi_h\mathbf{B}\Pi_e\right). \quad (54)$$

On the other hand, employing complex-valued matrix differentiation technique [25], one can obtain that

$$\nabla_{P_S} C_{\mathrm{erg,A}} = \frac{\log_2 e}{2\sigma_b^2}\mathrm{Tr}\left(\mathbf{B}^H\mathbf{H}_b^H\mathbf{H}_b\mathbf{B}\Omega_b\right) - \frac{\log_2 e}{2\alpha\sigma_b^2}\mathrm{Tr}\left(\mathbf{B}^H\Phi_h\mathbf{B}\Omega_e\right). \quad (55)$$

Substituting (54) and (55) into (53), the proof of ***Proposition 2*** is completed.

APPENDIX C
PROOF OF PROPOSITION 3

For analysis, in the low SNR regime, $P_0/\sigma_b^2 \to 0$, which can be seen as that P_0 is fixed while $\sigma_b^2 \to \infty$. Using the first order expansion of mutual information [28], we have

$$I(\mathbf{x};\bar{\mathbf{y}}_b) = \frac{P_S\log_2 e}{\sigma_b^2}\mathrm{Tr}\left(\mathbf{B}^H\mathbf{H}_b^H\mathbf{H}_b\mathbf{B}\right) + \mathcal{O}\left(\frac{1}{\sigma_b^4}\right), \quad (56)$$

$$\begin{aligned}&\mathbb{E}_{\mathbf{H}_e}\{I(\mathbf{x};\bar{\mathbf{y}}_e)\}\\ &= \mathbb{E}_{\mathbf{H}_e}\left\{\frac{P_S\log_2 e}{\alpha\sigma_b^2}\mathrm{Tr}\left(\mathbf{B}^H\mathbf{H}_e^H\mathbf{R}_e^{-1}\mathbf{H}_e\mathbf{B}\right) + \mathcal{O}\left(\frac{1}{\sigma_b^4}\right)\right\}\\ &= \frac{P_S\log_2 e}{\alpha\sigma_b^2}\mathrm{Tr}\left(\Psi_h\mathbf{R}_e^{-1}\right)\mathrm{Tr}\left(\mathbf{B}^H\Phi_h\mathbf{B}\right) + \mathcal{O}\left(\frac{1}{\sigma_b^4}\right).\end{aligned} \quad (57)$$

Note that $\mathbf{R}_e$ in the low SNR regime can be approximated as $\mathbf{R}_e = \frac{P_J}{\alpha\sigma_b^2}\omega_{\max}\Psi_g + \mathbf{I}_{N_e} \approx \mathbf{I}_{N_e}$. Then the approximation of the ergodic secrecy rate at low SNR is given by

$$C_{\mathrm{erg}} \approx \frac{P_S\log_2 e}{\sigma_b^2}\mathrm{Tr}\left(\mathbf{B}^H\left(\mathbf{H}_b^H\mathbf{H}_b - \frac{\mathrm{Tr}(\Psi_h)}{\alpha}\Phi_h\right)\mathbf{B}\right). \quad (58)$$

As can be seen from (58), the ergodic secrecy rate is linear with P_S, therefore, it is optimal to set $P_S = P_0$ and $P_J = 0$ in the low SNR regime. Besides, the optimal $\mathbf{B}$, which maximizes the ergodic secrecy rate in (58), is the beamforming design given in (31).

Thus, the proof of ***Proposition 3*** is completed.

APPENDIX D
PROOF OF PROPOSITION 4

By utilizing the matrix equality $Tr\left\{\mathbf{P}^H\mathbf{A}\mathbf{P}\mathbf{B}\right\} = (vec(\mathbf{P}))^H\left(\mathbf{B}^T\otimes\mathbf{A}\right)vec(\mathbf{P})$, we have

$$\begin{aligned}\mathbf{e}_{mn}^H\mathbf{B}^H\mathbf{H}_b^H\mathbf{H}_b\mathbf{B}\mathbf{e}_{mn} &= \mathrm{Tr}\left(\mathbf{B}^H\mathbf{H}_b^H\mathbf{H}_b\mathbf{B}\mathbf{E}_{mn}\right)\\ &= \mathbf{b}^H\left(\mathbf{E}_{mn}^T\otimes\mathbf{H}_b^H\mathbf{H}_b\right)\mathbf{b}\\ &= \mathbf{p}^H\mathbf{D}^{-\frac{H}{2}}\left(\mathbf{E}_{mn}^T\otimes\mathbf{H}_b^H\mathbf{H}_b\right)\mathbf{D}^{-\frac{1}{2}}\mathbf{p}.\end{aligned} \quad (59)$$

Similarly, we have $\mathrm{Tr}\left(\mathbf{B}^H\Phi_h\mathbf{B}\right) = \mathbf{b}^H\mathbf{D}\mathbf{b} = ||\mathbf{p}||^2$, then the problem (35) can be formulated as

$$\begin{aligned}&\max_{\mathbf{q}}\ \min_{m\neq n}\mathbf{q}^H\mathbf{D}^{-\frac{H}{2}}\left(\mathbf{E}_{mn}^T\otimes\mathbf{H}_b^H\mathbf{H}_b\right)\mathbf{D}^{-\frac{1}{2}}\mathbf{q}\\ &\text{s.t. } \mathbf{q}^H\mathbf{q} = 1.\end{aligned} \quad (60)$$

With $\mathbf{Q} = \mathbf{q}\mathbf{q}^H$, the problem (60) is equivalent to the following problem

$$\begin{aligned}&\max_{\mathbf{Q}}\min_{m\neq n}\ \mathrm{Tr}\left\{\mathbf{D}^{-\frac{H}{2}}\left(\mathbf{E}_{mn}^T\otimes\mathbf{H}_b^H\mathbf{H}_b\right)\mathbf{D}^{-\frac{1}{2}}\mathbf{Q}\right\}\\ &\text{s.t. } \mathrm{Tr}(\mathbf{Q}) = 1.\\ &\quad \mathbf{Q}\succeq\mathbf{0}, \mathrm{rank}(\mathbf{Q}) = 1.\end{aligned} \quad (61)$$

Denote $t = \min_{m\neq n}\mathrm{Tr}\left\{\mathbf{D}^{-\frac{H}{2}}\left(\mathbf{E}_{mn}^T\otimes\mathbf{H}_b^H\mathbf{H}_b\right)\mathbf{D}^{-\frac{1}{2}}\mathbf{Q}\right\}$ and consider that $\mathbf{E}_{mn} = \mathbf{E}_{nm}$, the problem (61) can be

transformed as follows

$$\begin{aligned} &\max_{\mathbf{Q}} \ t \\ &\text{s.t. } \mathrm{Tr}\left\{\mathbf{D}^{-\frac{H}{2}}\left(\mathbf{E}_{mn}^{T}\otimes\mathbf{H}_{b}^{H}\mathbf{H}_{b}\right)\mathbf{D}^{-\frac{1}{2}}\mathbf{Q}\right\}\geq t \quad m<n, \\ &\quad \mathrm{Tr}\left(\mathbf{Q}\right)=1, \\ &\quad \mathbf{Q}\geq\mathbf{0}, \mathrm{rank}(\mathbf{Q})=1. \end{aligned} \tag{62}$$

Hence, the proof of ***Proposition 4*** is completed.

REFERENCES

[1] E. D. Silva, A. L. D. Santos, L. C. P. Albini, and M. N. Lima, "Identity-based key management in mobile ad hoc networks: Techniques and applications," *IEEE Wireless Commun.*, vol. 15, no. 5, pp. 46–52, Oct. 2008.

[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[3] 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, Jun. 2009.

[4] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[5] S.-C. Lin and C.-L. Lin, "On secrecy capacity of fast fading MIMOME wiretap channels with statistical CSIT," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3293–3306, Jun. 2014.

[6] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Trans. Inf. Theory.*, vol. 62, no. 7, pp. 3880–3890, Jul. 2016.

[7] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—Part I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory.*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[8] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory.*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[9] 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.

[10] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3301–3314, Jul. 2011.

[11] 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, no. 12, pp. 3816–3825, Dec. 2012.

[12] 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, Jul. 2012.

[13] 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, no. 5, pp. 527–529, May 2011.

[14] W. Zeng, Y. R. Zheng, and C. Xiao, "Multi-antenna secure cognitive radio networks with finite-alphabet inputs: A global optimization approach for precoder design," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 3044–3067, Apr. 2016.

[15] X. Li, R. Fan, X. Ma, J. An, and T. Jiang, "Secure space-time communications over Rayleigh flat fading channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1491–1504, Feb. 2016.

[16] S. 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. 5013–5022, Oct. 2011.

[17] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875–1888, Mar. 2010.

[18] G. Zheng, L. C. Choo, and K. K. Wong, "Optimal cooperative jammming to enhance physical layer security using relays," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1317–1322, Mar. 2011.

[19] J. Huang and A. L. Swindlehurst, "Robust secure transmission in MISO channels based on worst-case optimization," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1696–1707, Apr. 2012.

[20] A. Mukherjee and A. L. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351–361, Jan. 2011.

[21] J.-P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[22] C. Xiao, J. Wu, S. Leong, Y. R. Zheng, and K. Letaief, "A discrete-time model for triply selective MIMO Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1678–1688, Sep. 2004.

[23] W. Zeng, C. Xiao, and J. Lu, "A low-complexity design of linear precoding for MIMO channels with finite-alphabet inputs," *IEEE Commun. Lett.*, vol. 1, no. 1, pp. 38–41, Feb. 2012.

[24] W. Zeng, C. Xiao, M. Wang, and J. Lu, "Linear precoding for finite-alphabet inputs over MIMO fading channels with statistical CSI," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3134–3148, Jun. 2012.

[25] A. Hjørungnes, *Complex-Valued Matrix Derivatives*. New York, NY, USA: Cambridge Univ. Press, 2011.

[26] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[27] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[28] F. Pérez-Cruz, M. R. D. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1070–1084, Mar. 2010.

KUO CAO (S'15) received the B.S. degree in communications engineering from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2013, where he is currently pursuing the Ph.D. degree in communications and information systems. His research interests focus on MIMO systems, cooperative communications, and network security.

YONGPENG WU (S'08–M'13) received the B.S. degree in telecommunications engineering from Wuhan University, Wuhan, China, in 2007, and the Ph.D. degree in communications and signal processing with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China, in 2013. He was the Humboldt Research Fellow and the Senior Research Fellow with the Institute for Digital Communications, Universität Erlangen-Nürnberg, Germany. During his doctoral studies, he has conducted cooperative research with the Department of Electrical Engineering, Missouri University of Science and Technology, USA. He is currently a Senior Research Fellow with the Institute for Communications Engineering, Technical University of Munich, Germany. His research interests include massive MIMO/MIMO systems, physical layer security, signal processing for wireless communications, and multivariate statistical theory. He has been a TPC Member of various conferences, including ICC, VTC, and WCSP. He received the IEEE Student Travel Grants of the IEEE International Conference on Communications (ICC) 2010, the Alexander von Humboldt Fellowship in 2014, and the Travel Grants for IEEE Communication Theory Workshop 2016. He was an Exemplary Reviewer of the IEEE TRANSACTIONS ON COMMUNICATIONS in 2015. He is the Lead Guest Editor for the upcoming special issue Physical Layer Security for 5G Wireless Networks of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is currently an Editor of the IEEE COMMUNICATION LETTER.

YUEMING CAI (M'05–SM'12) received the B.S. degree in physics from Xiamen University, Xiamen, China, in 1982, and the M.S. degree in micro-electronics engineering, and the Ph.D. degree in communications and information systems from Southeast University, Nanjing, China, in 1988 and 1996, respectively. His research interests include MIMO systems, OFDM systems, signal processing in communications, cooperative communications, and wireless sensor networks.

WEIWEI YANG (S'08–M'12) received the B.S., M.S., and Ph.D. degrees from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2003, 2006, and 2011, respectively. His research interests include orthogonal frequency domain multiplexing systems, signal processing in communications, cooperative communications, wireless sensor networks, and network security.