Artificial Noise-Aided Secure Relay Communication With Unknown Channel Knowledge of Eavesdropper

Bin Li, *Senior Member, IEEE*, Meiying Zhang, Yue Rong, *Senior Member, IEEE*, and Zhu Han, *Fellow, IEEE*

***Abstract*— In this article, a new relay-aided secure communication system is investigated, where a transmitter sends signals to a destination via an amplify-and-forward (AF) relay in the presence of an eavesdropper. We consider a general system configuration, where the source, relay, destination, and eavesdropper are all equipped with multiple antennas. In the practical scenarios of unknown eavesdropper's channel state information (CSI) and uncertainty of the eavesdropper's location, we aim to maximize the expected value of the system secrecy rate over the presumed distribution of the eavesdropper's channels, by exploiting the artificial noise (AN) transmitted by the source and relay nodes. The system design issue is formulated as a nonconvex stochastic optimization problem with a source transmission power constraint and a nonconvex relay transmission power constraint. A novel computational method is proposed to solve this challenging problem. The new method is developed based on an exact penalty function method together with a parallel stochastic decomposition algorithm. Numerical simulations are performed to study the effectiveness of the proposed scheme at various locations of the eavesdropper. Simulation results show that for most cases, secure communication can be achieved without the CSI knowledge of eavesdropper's channels, and the achievable secrecy rate follows the trend of a benchmark system where the eavesdropper's full CSI is available. In particular, the achievable system secrecy rate increases with the number of antennas at the legitimate users. Moreover, the optimal power allocated for the transmission of the AN increases with the system signal-to-noise ratio. The proposed computational method achieves a higher system secrecy rate than a conventional penalty function based approach.**

Manuscript received July 12, 2020; revised November 2, 2020; accepted December 25, 2020. Date of publication January 8, 2021; date of current version May 10, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 62071317; and in part by the NSF under Grant EARS-1839818, Grant CNS1717454, Grant CNS-1731424, and Grant CNS-1702850. The associate editor coordinating the review of this article and approving it for publication was S. Dey. *(Corresponding author: Yue Rong.)*

Bin Li is with the School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065, China (e-mail: bin.li@scu.edu.cn).

Meiying Zhang is with the College of Electrical Engineering, Sichuan University, Chengdu 610065, China (e-mail: mei.ying.zhang@hotmail.com).

Yue Rong is with the School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Perth, WA 6102, Australia (e-mail: y.rong@curtin.edu.au).

Zhu Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea (e-mail: zhan2@uh.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2020.3047926.

Digital Object Identifier 10.1109/TWC.2020.3047926

***Index Terms*— Secure communication, amplify-and-forward relay, artificial noise.**

I. Introduction

A. Background

With the growing popularity of mobile Internet, providing secure communication services has become a critical issue for system operators and designers. Traditionally, security in wireless communication networks is mainly realized by cryptographic techniques applied to the upper layers of the communication protocol stack utilizing secret keys. However, these techniques have a major drawback that secret keys are often vulnerable to malicious attacks from eavesdroppers.

To improve the security of wireless communication, physical layer security [1] has attracted much research interest recently. Physical layer security technologies prevent smart devices and Internet of Things (IoT) from potential attacks of eavesdroppers. Probabilistic characteristics of the achievable secrecy rates and average secrecy rates were presented in [2]. Physical layer security in probabilistic caching was analyzed in [3]. Three secrecy metrics for secure transmission over quasi-static fading channels were proposed in [4]. Various techniques such as relay-aided security [5] and artificial noise (AN)-aided security [6] have been proposed to increase the system secrecy capacity.

B. Literature Review

Cooperative relaying is an emerging physical layer security approach. It is shown in [5] that relaying is capable of improving the level of physical layer security. This discovery has led to further research efforts devoted to investigating the security of relay-aided communications from the physical layer perspective [7]. A wireless relay can adopt either the amplify-and-forward (AF) or the decode-and-forward (DF) strategy for forwarding source messages. For the DF relaying, the optimal weights that achieve the maximum secrecy capacity were derived in [8] and [9]. The achievable secrecy rate of an AF relay network in the presence of direct links to the destination and eavesdropper was characterized in [10].

Cooperative jamming is a promising technology for improving information secrecy at the physical layer [11].

Cooperative jamming can reduce the probability of being eavesdropped by sending jamming signals, while maintaining the fine reception of intended information. Hence it can efficiently improve the system secrecy performance [12]. Jamming signals can be transmitted at the destination to degrade the eavesdropper's channel [13]. A destination-aided secure transmission scheme was proposed in [14] where the destination broadcasts jamming signals concurrently with the transmission of the confidential messages to protect the confidential messages.

Jamming signals can also be transmitted from the source in the form of temporal AN [15]. Precoding methods can be applied to send the AN to the null space of the legitimate link to enhance the secrecy performance [16]. The optimal power allocation between the information and jamming signals for a DF relay system was derived in [17]. A source-based jamming strategy was proposed in [18] for a dual-hop AF cooperative network. The optimal distribution of the jamming signals from the attacker's point of view was studied in [19]. In [20], a relay-jammer scheme was proposed to improve the primary user's secrecy in cognitive radio networks. An AN-aided precoding strategy employing a full-duplex relay with imperfect channel state information (CSI) was investigated in [21] to enhance the secrecy performance of wireless communication systems. In [22], an AN-aided secure on-off transmission scheme in a wiretap channel was developed. The secrecy capacity of an AF relay system was studied in two scenarios in [23] with the AN added either at the source node or the relay node, and the optimal power allocation between the confidential signal and the jamming signal is derived to maximize the secrecy rate. Wireless-powered jammers were investigated in [24], [25]. Recently, the presence of randomly distributed non-colluding eavesdroppers has been investigated in multi-antenna DF relay wiretap channels [26].

C. Contributions

In this article, we consider secure wireless communications between a pair of multi-antenna source and destination nodes aided by a multi-antenna AF relay node, subjecting to individual transmission power constraints at the source node and the relay node. This setup is applicable to the scenario where a powerful relay station assists secure communications between two wireless terminals.

It is worth noting that in [16], [20], [27], the knowledge on the CSI of the eavesdropper's channel is required. However, in practice, as there is no cooperation between the legitimate users and the eavesdropper, we cannot obtain the eavesdropper's CSI, and the location of the eavesdropper can change over time. To prevent confidential information leakage under the conditions of unknown eavesdropper's CSI and location, the source node transmits AN along with the information signals. Meanwhile, the half-duplex AF relay transmits its own AN jamming signal while forwarding the signal-of-interest to the destination node.

Considering such a practical system setup, we aim at maximizing the expected value of the system secrecy rate over the presumed distribution of the eavesdropper's channels under the source and relay transmission power constraints. Note that in practice, there may be mismatch between the actual and presumed distributions. To the best of our knowledge, this problem has not been studied before for the system setup given above. To maximize the signal power received at the destination node, the source beamforming vector and the relay precoding matrix are chosen to match the strongest subchannels in the source-relay and relay-destination links. Moreover, the direct source-destination link is considered in the system optimization. We show that the system design problem can be formulated as a stochastic programming problem with a nonconvex objective function, a source transmission power constraint and a nonconvex relay transmission power constraint. There are two difficulties in this challenging problem. Firstly, both the objective function and the constraints are nonconvex. Secondly, the distribution of the eavesdropper's CSI is unknown.

To overcome these difficulties, a novel computational method is developed in this article. Firstly, an exact penalty function method is utilized to append the nonconvex relay transmission power constraint to the objective function, which leads to a simpler stochastic optimization problem with a nonconvex objective function and convex constraints. By observing the structure of the resulted problem, a parallel stochastic decomposition algorithm is introduced to tackle the second challenge. The result of the optimization problem provides the optimal power allocation at the source and relay nodes for transmitting the AN and the signal-of-interest to maximize the expected value of the system secrecy rate.

Numerical simulations are performed to study the effectiveness of the proposed scheme at various locations of the eavesdropper. Simulation results show that for most cases, secure communication can be achieved without the CSI knowledge of eavesdropper's channels, and the achievable system secrecy rate follows the trend of a benchmark system where the eavesdropper's full CSI is available. In particular, the achievable secrecy rate increases with the number of antennas at the legitimate users. Moreover, we find out that to maximize the system secrecy rate, the power allocated for the transmission of AN increases with the signal-to-noise ratio (SNR) at both the source and relay nodes. The proposed method consistently achieves a higher system secrecy rate than a conventional penalty function based approach. In many cases, the proposed method also yields a higher secrecy rate than a beamforming algorithm where the eavesdropper's CSI is known, but without applying the AN at the relay node.

Based on our best knowledge, currently there is no wireless communication standard explicitly incorporating physical layer security technologies. Nevertheless, the scheme proposed in our paper can be applied to general two-hop AF relay communication systems, where the source, relay, and destination nodes are equipped with multiple antennas.

D. Structure

The rest of the paper is organized as follows. The model of an AF relay-aided secure communication system with AN from both the source and the relay nodes is presented in Section II. The secrecy rate maximization problem is also

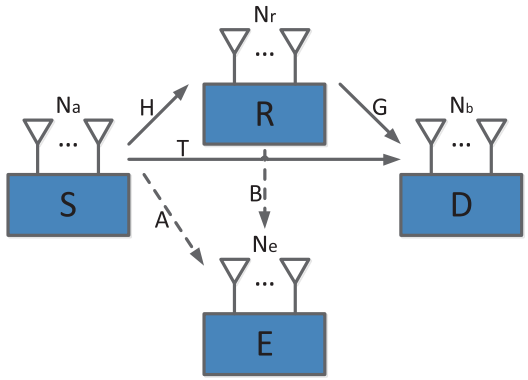


Fig. 1. Block diagram of an AF MIMO relay communication system in the presence of an eavesdropper.

formulated in Section II. An efficient algorithm to solve the optimization problem is developed in Section III. Numerical examples are presented in Section IV to demonstrate the performance of the proposed scheme. Finally, we conclude this article in Section V.

II. System Model

We consider a relay-aided secure communication system in Fig. 1, where the source node sends signals to the destination node via a half-duplex AF multiple-input multiple-output (MIMO) relay node, and there is an eavesdropper in the system.[1] We assume that the number of antennas at the source, relay, destination, and eavesdropper are N_a, N_r, N_b, and N_e, respectively. The channel between the source and relay nodes is $\mathbf{H}$, between the relay and destination nodes is $\mathbf{G}$. The direct link between the source and destination nodes is $\mathbf{T}$, and we denote the channels from the source and relay nodes to the eavesdropper as $\mathbf{A}$ and $\mathbf{B}$, respectively. We assume that channels $\mathbf{H}$, $\mathbf{G}$, and $\mathbf{T}$ are known. However, the eavesdropper's channels $\mathbf{A}$ and $\mathbf{B}$ are unknown, as in practice there is no coordination between the source and relay nodes and the eavesdropper to obtain the knowledge of $\mathbf{A}$ and $\mathbf{B}$. To enable secure communication without the eavesdropper's CSI, we assume that there is no eavesdropper within a certain distance to the source node.[2]

Let us introduce the singular value decomposition (SVD) of $\mathbf{H}$ and $\mathbf{G}$ as $\mathbf{H} = \mathbf{U}_h\mathbf{\Lambda}_h^{\frac{1}{2}}\mathbf{V}_h^H$ and $\mathbf{G} = \mathbf{U}_g\mathbf{\Lambda}_g^{\frac{1}{2}}\mathbf{V}_g^H$, respectively, where $(\cdot)^H$ denotes the matrix and vector Hermitian transpose, and the diagonal elements of $\mathbf{\Lambda}_h$ and $\mathbf{\Lambda}_g$ are sorted in a descending order. The communication between the source and destination nodes is completed in two time slots. At the first time slot, the source node transmits the following signal vector to the relay and destination nodes

$$\mathbf{x}_a = \sqrt{p}\mathbf{v}_{h,1}s + \sqrt{\frac{P_a - p}{N_a - 1}}\mathbf{V}_{h,\bar{1}}\mathbf{w}_a \tag{1}$$

where $\mathbf{V}_h = [\mathbf{v}_{h,1}, \mathbf{V}_{h,\bar{1}}]$, P_a is the transmission power available at the source node, s is the signal-of-interest with $E[|s|^2] = 1$, and $\mathbf{w}_a$ is the AN vector with $E[\mathbf{w}_a\mathbf{w}_a^H] = \mathbf{I}_{N_a-1}$. Here $E[\cdot]$ stands for the statistical expectation, $\mathbf{I}_n$ denotes the $n \times n$ identity matrix, the subscript "1" in $\mathbf{v}_{h,1}$ denotes the singular vector with the largest singular value, and "$\bar{1}$" in $\mathbf{V}_{h,\bar{1}}$ denotes a matrix containing all the other eigenvectors.

It can be seen from (1) that $\mathrm{tr}(E[\mathbf{x}_a\mathbf{x}_a^H]) = P_a$, where $\mathrm{tr}(\cdot)$ stands for the matrix trace, and the amount of power spent on transmitting s and $\mathbf{w}_a$ is p and $P_a - p$, respectively. In (1), $\mathbf{v}_{h,1}$ is used to maximize the gain of sending s through $\mathbf{H}$, and $\mathbf{V}_{h,\bar{1}}$ is applied such that $\mathbf{w}_a$ does not affect the transmission of s to the relay node, as $\mathbf{w}_a$ is transmitted through a channel that is orthogonal to that of s. The received signal vector at the relay node is given by

$$\mathbf{y}_r = \mathbf{H}\mathbf{x}_a + \mathbf{n}_r \tag{2}$$

where $\mathbf{n}_r$ is the additive noise vector at the relay node with $E[\mathbf{n}_r\mathbf{n}_r^H] = \sigma_r^2\mathbf{I}_{N_r}$ and σ_r^2 is the noise variance. Substituting (1) into (2), we have

$$\mathbf{y}_r = \sqrt{\lambda_{h,1}p}\mathbf{u}_{h,1}s + \sqrt{\frac{P_a - p}{N_a - 1}}\mathbf{U}_{h,\bar{1}}\mathbf{\Lambda}_{h,\bar{1}}^{\frac{1}{2}}\mathbf{w}_a + \mathbf{n}_r \tag{3}$$

where $\mathbf{U}_h = [\mathbf{u}_{h,1}, \mathbf{U}_{h,\bar{1}}]$ and $\mathbf{\Lambda}_h = \mathrm{diag}[\lambda_{h,1}, \mathbf{\Lambda}_{h,\bar{1}}]$. Here $\mathrm{diag}[\cdot]$ stands for a diagonal matrix.

The received signal vector at the destination node at the first time slot is given by

$$\mathbf{y}_{b,1} = \mathbf{T}\mathbf{x}_a + \mathbf{n}_{b,1} \tag{4}$$

where $\mathbf{n}_{b,1}$ is the additive noise vector at the destination node at the first time slot with $E[\mathbf{n}_{b,1}\mathbf{n}_{b,1}^H] = \sigma_b^2\mathbf{I}_{N_b}$ and σ_b^2 is the noise variance at the destination node. By substituting (1) into (4), we obtain

$$\mathbf{y}_{b,1} = \sqrt{p}\mathbf{T}\mathbf{v}_{h,1}s + \sqrt{\frac{P_a - p}{N_a - 1}}\mathbf{T}\mathbf{V}_{h,\bar{1}}\mathbf{w}_a + \mathbf{n}_{b,1}. \tag{5}$$

At the second time slot, the relay node linearly precodes $\mathbf{y}_r$ with $\mathbf{F}$, superimposes its own AN jamming vector $\mathbf{w}_r$, and transmits the following signal vector to the destination node

$$\mathbf{x}_r = \mathbf{F}\mathbf{y}_r + \sqrt{\beta}\mathbf{V}_{g,\bar{1}}\mathbf{w}_r \tag{6}$$

where $\mathbf{V}_g = [\mathbf{v}_{g,1}, \mathbf{V}_{g,\bar{1}}]$ and $E[\mathbf{w}_r\mathbf{w}_r^H] = \mathbf{I}_{N_r-1}$. Here we choose $\mathbf{F} = \sqrt{\alpha}\mathbf{v}_{g,1}\mathbf{u}_{h,1}^H$ such that the power of s is maximized at the destination node, as s is transmitted through the strongest subchannel of $\mathbf{G}$. Furthermore, $\mathbf{V}_{g,\bar{1}}$ is used such that $\mathbf{w}_r$ does not interfere with the transmission of $\mathbf{y}_r$ to the destination node, as $\mathbf{w}_r$ is transmitted through a channel which is orthogonal to that of $\mathbf{y}_r$. Substituting (3) into (6), we have

$$\mathbf{x}_r = \sqrt{\alpha\lambda_{h,1}p}\mathbf{v}_{g,1}s + \sqrt{\alpha}\mathbf{v}_{g,1}\mathbf{u}_{h,1}^H\mathbf{n}_r + \sqrt{\beta}\mathbf{V}_{g,\bar{1}}\mathbf{w}_r. \tag{7}$$

[1]For a clear demonstration of the proposed approach, one relay node and one eavesdropper are considered here. We would like to note that the proposed approach can be applied to systems with multiple relay nodes and multiple eavesdroppers.

[2]When the CSI of the eavesdropper is not available, secure communication may be impossible to achieve if the eavesdropper's channel is much stronger than that of the relay and destination nodes, which occurs when the eavesdropper is located much closer to the source node than the legitimate users. This can be seen from the simulation results in Fig. 10 where for certain source-eavesdropper distances, except for the scheme with the full CSI, the other three systems considered cannot achieve secure communication, due to a much weaker channel of the legitimate users. To exclude this situation, we assume that there is no eavesdropper within a certain distance to the source node.

From (7), we can see that the amount of power spent on transmitting the AN vector $\mathbf{w}_r$ is $\beta(N_r-1)$, and the total amount of transmission power consumed by the relay node is

$$\mathrm{tr}(E[\mathbf{x}_r\mathbf{x}_r^H]) = \alpha\lambda_{h,1}p + \alpha\sigma_r^2 + \beta(N_r-1). \tag{8}$$

The received signal vector at the destination node at the second time slot can be written as

$$\mathbf{y}_{b,2} = \mathbf{G}\mathbf{x}_r + \mathbf{n}_{b,2} \tag{9}$$

where $\mathbf{n}_{b,2}$ is the additive noise vector at the destination node at the second time slot with $E[\mathbf{n}_{b,2}\mathbf{n}_{b,2}^H] = \sigma_b^2\mathbf{I}_{N_b}$. By substituting (7) into (9), we obtain

$$\begin{aligned}\mathbf{y}_{b,2} = \sqrt{\alpha p\lambda_{h,1}\lambda_{g,1}}\mathbf{u}_{g,1}s + \sqrt{\alpha\lambda_{g,1}}\mathbf{u}_{g,1}\mathbf{u}_{h,1}^H\mathbf{n}_r \\ + \sqrt{\beta}\mathbf{U}_{g,\bar{1}}\mathbf{\Lambda}_{g,\bar{1}}^{\frac{1}{2}}\mathbf{w}_r + \mathbf{n}_{b,2}\end{aligned} \tag{10}$$

where $\mathbf{U}_g = [\mathbf{u}_{g,1}, \mathbf{U}_{g,\bar{1}}]$ and $\mathbf{\Lambda}_g = \mathrm{diag}[\lambda_{g,1}, \mathbf{\Lambda}_{g,\bar{1}}]$, From (5) and (9), the received signals at the destination node over two time slots are given by

$$\begin{aligned}\mathbf{y}_b &= \begin{pmatrix}\mathbf{y}_{b,2}\\ \mathbf{y}_{b,1}\end{pmatrix} \\ &= \begin{pmatrix}\sqrt{\alpha p\lambda_{h,1}\lambda_{g,1}}\mathbf{u}_{g,1}\\ \sqrt{p}\,\mathbf{T}\mathbf{v}_{h,1}\end{pmatrix}s \\ &\quad + \begin{pmatrix}\sqrt{\alpha\lambda_{g,1}}\mathbf{u}_{g,1}\mathbf{u}_{h,1}^H\mathbf{n}_r + \sqrt{\beta}\mathbf{U}_{g,\bar{1}}\mathbf{\Lambda}_{g,\bar{1}}^{\frac{1}{2}}\mathbf{w}_r + \mathbf{n}_{b,2}\\ \sqrt{\frac{P_a-p}{N_a-1}}\mathbf{T}\mathbf{V}_{h,\bar{1}}\mathbf{w}_a + \mathbf{n}_{b,1}\end{pmatrix}.\end{aligned} \tag{11}$$

Using the maximal ratio combining (MRC) technique, the SNR at the destination node can be obtained from (11) as

$$\begin{aligned}\mathrm{SNR}_b = \frac{\alpha p\lambda_{h,1}\lambda_{g,1}}{\alpha\lambda_{g,1}\sigma_r^2 + \sigma_b^2} \\ + p\mathbf{v}_{h,1}^H\mathbf{T}^H\left(\frac{P_a-p}{N_a-1}\mathbf{T}\mathbf{V}_{h,\bar{1}}\mathbf{V}_{h,\bar{1}}^H\mathbf{T}^H + \sigma_b^2\mathbf{I}_{N_b}\right)^{-1}\mathbf{T}\mathbf{v}_{h,1}\end{aligned} \tag{12}$$

where $(\cdot)^{-1}$ denotes the matrix inversion.

From (1), the received signal vector at the eavesdropper at the first time slot is given by

$$\begin{aligned}\mathbf{y}_{e,1} &= \mathbf{A}\mathbf{x}_a + \mathbf{n}_{e,1} \\ &= \sqrt{p}\mathbf{A}\mathbf{v}_{h,1}s + \sqrt{\frac{P_a-p}{N_a-1}}\mathbf{A}\mathbf{V}_{h,\bar{1}}\mathbf{w}_a + \mathbf{n}_{e,1}\end{aligned} \tag{13}$$

where $\mathbf{n}_{e,1}$ is the additive noise vector at the eavesdropper at the first time slot with $E[\mathbf{n}_{e,1}\mathbf{n}_{e,1}^H] = \sigma_e^2\mathbf{I}_{N_e}$ and σ_e^2 is the noise variance at the eavesdropper. The received signal vector at the eavesdropper at the second time slot can be obtained from (7) as

$$\begin{aligned}\mathbf{y}_{e,2} &= \mathbf{B}\mathbf{x}_r + \mathbf{n}_{e,2} \\ &= \sqrt{\alpha\lambda_{h,1}p}\mathbf{B}\mathbf{v}_{g,1}s + \sqrt{\alpha}\mathbf{B}\mathbf{v}_{g,1}\mathbf{u}_{h,1}^H\mathbf{n}_r \\ &\quad + \sqrt{\beta}\mathbf{B}\mathbf{V}_{g,\bar{1}}\mathbf{w}_r + \mathbf{n}_{e,2}\end{aligned} \tag{14}$$

where $\mathbf{n}_{e,2}$ is the additive noise vector at the eavesdropper at the second time slot with $E[\mathbf{n}_{e,2}\mathbf{n}_{e,2}^H] = \sigma_e^2\mathbf{I}_{N_e}$. From (13) and (14), the received signals at the eavesdropper over two time slots are given by

$$\mathbf{y}_e = \begin{pmatrix}\mathbf{y}_{e,1}\\ \mathbf{y}_{e,2}\end{pmatrix} = \mathbf{h}s + \mathbf{v} \tag{15}$$

where $\mathbf{h}$ is the equivalent channel and $\mathbf{v}$ is the equivalent noise vector given by

$$\mathbf{h} = \begin{pmatrix}\sqrt{p}\mathbf{A}\mathbf{v}_{h,1}\\ \sqrt{\alpha\lambda_{h,1}p}\mathbf{B}\mathbf{v}_{g,1}\end{pmatrix}, \tag{16}$$

$$\mathbf{v} = \begin{pmatrix}\sqrt{\frac{P_a-p}{N_a-1}}\mathbf{A}\mathbf{V}_{h,\bar{1}}\mathbf{w}_a + \mathbf{n}_{e,1}\\ \sqrt{\alpha}\mathbf{B}\mathbf{v}_{g,1}\mathbf{u}_{h,1}^H\mathbf{n}_r + \sqrt{\beta}\mathbf{B}\mathbf{V}_{g,\bar{1}}\mathbf{w}_r + \mathbf{n}_{e,2}\end{pmatrix}. \tag{17}$$

Based on (15), the SNR at the eavesdropper with an MRC receiver is given by

$$\begin{aligned}\mathrm{SNR}_e &= \mathbf{h}^H\mathbf{R}_v^{-1}\mathbf{h} \\ &= p\mathbf{v}_{h,1}^H\mathbf{A}^H\left(\frac{P_a-p}{N_a-1}\mathbf{A}\mathbf{V}_{h,\bar{1}}\mathbf{V}_{h,\bar{1}}^H\mathbf{A}^H + \sigma_e^2\mathbf{I}_{N_e}\right)^{-1}\mathbf{A}\mathbf{v}_{h,1} \\ &\quad + \alpha\lambda_{h,1}p\mathbf{v}_{g,1}^H\mathbf{B}^H\Big(\beta\mathbf{B}\mathbf{V}_{g,\bar{1}}\mathbf{V}_{g,\bar{1}}^H\mathbf{B}^H \\ &\quad + \alpha\sigma_r^2\mathbf{B}\mathbf{v}_{g,1}\mathbf{v}_{g,1}^H\mathbf{B}^H + \sigma_e^2\mathbf{I}_{N_e}\Big)^{-1}\mathbf{B}\mathbf{v}_{g,1}\end{aligned} \tag{18}$$

where

$$\begin{aligned}\mathbf{R}_v &= E[\mathbf{v}\mathbf{v}^H] \\ &= \mathrm{diag}\Big[\frac{P_a-p}{N_a-1}\mathbf{A}\mathbf{V}_{h,\bar{1}}\mathbf{V}_{h,\bar{1}}^H\mathbf{A}^H + \sigma_e^2\mathbf{I}_{N_e}, \\ &\quad \beta\mathbf{B}\mathbf{V}_{g,\bar{1}}\mathbf{V}_{g,\bar{1}}^H\mathbf{B}^H + \alpha\sigma_r^2\mathbf{B}\mathbf{v}_{g,1}\mathbf{v}_{g,1}^H\mathbf{B}^H + \sigma_e^2\mathbf{I}_{N_e}\Big]\end{aligned} \tag{19}$$

is the covariance matrix of $\mathbf{v}$ in (17). From (8), (12), and (18), the optimization problem which maximizes the expected value of the secrecy capacity subjecting to the transmission power constraints at the source and relay nodes can be formulated as

$$\max_{p,\alpha,\beta}\ \frac{1}{2}E_{\mathbf{A},\mathbf{B}}\{[\log(1+\mathrm{SNR}_b) - \log(1+\mathrm{SNR}_e)]^+\} \tag{20}$$

$$\text{s.t. } 0 < p < P_a, \quad 0 < \alpha, \quad 0 < \beta \tag{21}$$

$$\alpha\lambda_{h,1}p + \alpha\sigma_r^2 + \beta(N_r-1) \le P_r \tag{22}$$

where a factor $1/2$ is included to consider that two time slots are used in transmission, P_r is the transmission power available at the relay node, for a real number x, $[x]^+ = \max\{x, 0\}$, $E_{\mathbf{A},\mathbf{B}}\{\cdot\}$ denotes the expectation over the presumed distribution of channels $\mathbf{A}$ and $\mathbf{B}$, since the exact value of $\mathbf{A}$ and $\mathbf{B}$ is unknown. Note that as SNR_e is a complicated function of $\mathbf{A}$ and $\mathbf{B}$, a closed-form expression of the expectation in (20) cannot be obtained.

Remark 1: It is worth noting that as the eavesdropper is assumed to be capable of performing the MRC operation (18) on the received signals, (20) is a pessimistic value of the achievable secrecy rate. In practice, it is difficult for the eavesdropper to obtain the knowledge of $\mathbf{V}_h$ and $\mathbf{V}_g$. Thus, the SNR achievable at the eavesdropper is lower than (18).

Remark 2: A single data stream is used mainly to demonstrate the key idea of the algorithm proposed in this article. In general, the data rates from the source node to both the destination node and the eavesdropper increase with the number of data streams.

III. Proposed Algorithm

The problem in (20)-(22) is a stochastic optimization problem with a nonconvex objective function in (20) and a nonconvex constraint in (22). There are two challenges in this problem. First, this optimization problem has a nonconvex objective function and a nonconvex constraint. The second challenge is that the distributions of $\mathbf{A}$ and $\mathbf{B}$ in (20) are unknown. In this section, we develop a novel computational method to solve this challenging problem. More specifically, for tackling the first challenge, we construct an exact penalty function from the constraint in (22) to augment the objective function in (20). This leads to a simpler stochastic optimization problem with nonconvex objective function and convex constraints. Then a parallel stochastic successive convex approximation-based algorithm [28] is introduced to solve the augmented problem by observing its structure. Since this approach does not require the exact distribution information of $\mathbf{A}$ and $\mathbf{B}$ in (20), the second challenge is tackled in this step.

By applying the inequality of $E\{[x]^+\} \geq [E\{x\}]^+$, it follows that

$$-E_{\mathbf{A},\mathbf{B}}\{[\log(1+\mathrm{SNR}_b)-\log(1+\mathrm{SNR}_e)]^+\} \leq -[\log(1+\mathrm{SNR}_b)-E_{\mathbf{A},\mathbf{B}}\{\log(1+\mathrm{SNR}_e)\}]^+. \quad (23)$$

Using the upper bound (23), the problem in (20)-(22) is converted to the following problem

$$\min_{p,\alpha,\beta} -\log(1+\mathrm{SNR}_b)+E_{\mathbf{A},\mathbf{B}}\{\log(1+\mathrm{SNR}_e)\} \quad (24)$$

$$\text{s.t. } 0<p<P_a, \quad 0<\alpha, \quad 0<\beta \quad (25)$$

$$\alpha\lambda_{h,1}p+\alpha\sigma_r^2+\beta(N_r-1)\leq P_r. \quad (26)$$

Note that for simplicity, the factor $1/2$ in (20) is omitted in (24) and the derivations later on. By utilizing an exact penalty function method introduced in [29] and [30], the problem (24)-(26) is converted to the following problem with an augmented objective function

$$\min_{\boldsymbol{\xi}} -\log(1+\mathrm{SNR}_b)+E_{\mathbf{A},\mathbf{B}}\{\log(1+\mathrm{SNR}_e)\} + \epsilon^{-\eta}\Delta(p,\alpha,\beta,\epsilon)+\delta\epsilon^{\theta} \quad (27)$$

$$\text{s.t. } 0<p<P_a, \quad 0<\alpha, \ 0<\beta \quad (28)$$

$$0<\epsilon \quad (29)$$

where $\boldsymbol{\xi}=[p,\alpha,\beta,\epsilon]^T$ is the vector containing all optimization variables, $\delta>0$ is a penalty parameter, $\eta>0$ and $\theta>2$ are fixed constants. The penalty function $\epsilon^{-\eta}\Delta(p,\alpha,\beta,\epsilon)+\delta\epsilon^{\theta}$ in (27) includes the violation function $\Delta(p,\alpha,\beta,\epsilon)$ for the constraint in (22) defined by

$$\Delta(p,\alpha,\beta,\epsilon) = [\max\{0,\alpha\lambda_{h,1}p+\alpha\sigma_r^2+\beta(N_r-1)-P_r-\epsilon^{\mu}W\}]^2 \quad (30)$$

where $\mu>0$ and $0<W<1$ are fixed constants. Different from the conventional penalty function method, the penalty function in (27) also includes an additional penalty term $\delta\epsilon^{\theta}$. In addition, ϵ is a decision variable and the violation function (30) has an extra relaxed term $-\epsilon^{\mu}W$.

Remark 3: One of the key differences between the exact penalty function method and the conventional method is that for the proposed algorithm, there are two penalty factors (ϵ and δ) in the penalty function. One of the penalty factors ϵ is a variable to be optimized in the problem (27)-(29), while in the conventional method, there is only one fixed penalty factor. Moreover, for the exact penalty function method, the optimal solution can be obtained as the other penalty factor δ is set as a finite number. In contrast, the conventional method cannot obtain an optimal solution with a finite penalty factor.

The idea of the exact penalty function can be interpreted as follows. It can be shown from [29] and [30] that if δ is sufficiently large, the solutions of p, α, β from the problem in (27)-(29) are equal to the solutions of the problem in (24)-(26). Intuitively, during the process of minimizing the objective function in (27), since δ is fixed and ϵ is a decision variable, ϵ^{θ} should be reduced. As a consequence, $\epsilon^{-\eta}$ will increase, and hence the value of the constraint violation function $\Delta(p,\alpha,\beta,\epsilon)$ is reduced, leading to the satisfaction of the following constraint

$$\alpha\lambda_{h,1}p+\alpha\sigma_r^2+\beta(N_r-1)-P_r-\epsilon^{\mu}W\leq 0.$$

Remark 4: An optimal solution of the problem in (27)-(29) is also an optimum of (24)-(26) according to [29] and [30].

Note that the problem in (27)-(29) is a stochastic optimization problem with convex constraints, which is much simpler than the problem in (20)-(22). By taking advantage of this structure, a parallel stochastic successive convex approximation-based algorithm [28] is introduced to solve the problem in (27)-(29). In particular, the variables in $\boldsymbol{\xi}$ of the problem in (27)-(29) are optimized in an iterative fashion. More specifically, in the tth iteration, a random realization of $\mathbf{C}^{(t)} \triangleq [\mathbf{A}^{(t)},\mathbf{B}^{(t)}]$ is taken following the presumed distribution, and $\boldsymbol{\xi}$ is updated as

$$\boldsymbol{\xi}^{(t+1)}=(1-\gamma^{(t+1)})\boldsymbol{\xi}^{(t)}+\gamma^{(t+1)}\hat{\boldsymbol{\xi}}^{(t)}(\mathbf{C}^{(t)}) \quad (31)$$

where the superscript (t) denotes the variables at the tth iteration, $0<\gamma^{(t)}\leq 1$ is a sequence to be chosen, and $\hat{\boldsymbol{\xi}}^{(t)}(\mathbf{C}^{(t)})$ is the solution to the problem of minimizing surrogate functions [28] as

$$\hat{\xi}_i^{(t)}(\mathbf{C}^{(t)})=\arg\min_{\xi_i\in\Xi_i}\hat{f}_i(\xi_i;\mathbf{C}^{(t)}), \quad i=1,\ldots,4 \quad (32)$$

where $\hat{\xi}_i^{(t)}(\mathbf{C}^{(t)})$ and ξ_i are the ith entry of $\hat{\boldsymbol{\xi}}^{(t)}(\mathbf{C}^{(t)})$ and $\boldsymbol{\xi}$, respectively, $\hat{f}_i(\xi_i;\mathbf{C}^{(t)})$ is the surrogate function of the variable ξ_i as explained below, and Ξ_i is the feasible region of ξ_i.

Remark 5: As the actual distribution of $\mathbf{A}$ and $\mathbf{B}$ is unknown, we assume the channel from the source and relay nodes to the eavesdropper has independent and identically distributed (i.i.d.) Rayleigh fading, which is commonly used in practice. In particular, we assume that the source-eavesdropper distance is similar to the source-destination distance, and the distance from the relay to the eavesdropper is similar to the relay-destination distance. Thus, entries in $\mathbf{A}$ and $\mathbf{B}$ are presumed to have distributions of $\mathcal{CN}(0,\sigma_{SD}^2)$ and $\mathcal{CN}(0,\sigma_{RD}^2)$, respectively, where σ_{SD}^2 and σ_{RD}^2 (known) are the variance

(large-scale path loss) of entries in $\mathbf{T}$ and $\mathbf{G}$, respectively. Apparently, there can be mismatch between the presumed and actual distribution of $\mathbf{A}$ and $\mathbf{B}$, which is studied in Section IV.

For each ξ_i, we denote the objective function in (27) without the expectation operator as $f_{i,s}(\xi_i)$, when all other variables are fixed. Furthermore, we write $f_{i,s}(\xi_i) = f_{i,c}(\xi_i) + f_{i,n}(\xi_i)$, $i = 1, \ldots, 4$, where $f_{i,c}(\xi_i)$ and $f_{i,n}(\xi_i)$ are the convex and nonconvex parts of $f_{i,s}(\xi_i)$ with respect to ξ_i, respectively. Then the surrogate function of ξ_i at the tth iteration is given by

$$\hat{f}_i(\xi_i; \mathbf{C}^{(t)}) = \rho^{(t)} f_{i,c}(\xi_i) + \rho^{(t)}(\xi_i - \xi_i^{(t)}) g_{i,n}(\xi_i^{(t)}; \mathbf{C}^{(t)}) + (1-\rho^{(t)})(\xi_i - \xi_i^{(t)}) q_i^{(t-1)} + \tau(\xi_i - \xi_i^{(t)})^2 \quad (33)$$

where $0 < \rho^{(t)} \leq 1$ is a sequence to be chosen, $g_{i,n}(\xi_i^{(t)}; \mathbf{C}^{(t)})$ is the gradient of $f_{i,n}(\xi_i)$ at $\xi_i^{(t)}$ under the channel realization $\mathbf{C}^{(t)}$, $\tau > 0$ is a constant, and $q_i^{(t-1)}$ is an accumulation number updated recursively according to

$$q_i^{(t)} = (1 - \rho^{(t)}) q_i^{(t-1)} + \rho^{(t)} g_{i,s}(\xi_i^{(t)}; \mathbf{C}^{(t)}). \quad (34)$$

Here $g_{i,s}(\xi_i^{(t)}; \mathbf{C}^{(t)})$ is the gradient of $f_{i,s}(\xi_i)$ at $\xi_i^{(t)}$ under the channel realization $\mathbf{C}^{(t)}$. It can be seen from (33) that the surrogate function preserves the convex part of the objective function in (27), while linearizes its nonconvex part. In fact, (33) can be viewed as an incremental sample estimate of the objective function in (27) with respect to the variable ξ_i.

The details of the surrogate functions for the variables p, α, β, ϵ, and the accumulation numbers q_i, $i = 1, \ldots, 4$, are listed in the Appendix. As the surrogate functions in (33) are convex with respect to ξ_i, they can be efficiently optimized, for example, through the golden section search method [31]. Note that as the eavesdropper's channels $\mathbf{A}$ and $\mathbf{B}$ are unknown, they are randomly realized in each iteration. From (31) and (33), we can see that at each iteration, the variables are updated as a function of all channel realizations up to this iteration. Moreover, based on (34), the estimation of the gradient of the objective function becomes increasingly accurate as t increases.

The procedure of the proposed algorithm is listed in Algorithm 1, where $\|\cdot\|$ denotes the Euclidean norm of a vector. It can be shown similar to [28] that Algorithm 1 converges to a stationary point, if sequences $\rho^{(t)}$ and $\gamma^{(t)}$ satisfy a diminishing step size rule as

$$\rho^{(t)} = t^{-\pi_1}, \quad \gamma^{(t)} = t^{-\pi_2}, \quad 0.5 < \pi_1 < \pi_2 \leq 1. \quad (35)$$

It can be seen from Algorithm 1 that as the value of the penalty parameter δ increases in the outer loop, the penalty for violating the power constraint in (22) is heavier than that of conventional penalty function methods. After the completion of the optimization process, for a certain realization of $\mathbf{A}$ and $\mathbf{B}$, if $\log(1 + \mathrm{SNR}_b) > \log(1 + \mathrm{SNR}_e)$, then the secrecy rate is nonzero. Otherwise, the capacity of the eavesdropper's channel is larger than that of the destination node. In this case, the system secrecy rate is zero.

We can see that most of computations in Algorithm 1 are spent on Step 6 to update $\boldsymbol{\xi}^{(t)}$ following (31)-(34), which has a complexity order of $\mathcal{O}(N_e^3 + N_b^3)$ based on equations in the Appendix. Thus, the complexity of the proposed algorithm is $\mathcal{O}(\kappa(N_e^3 + N_b^3))$, where κ is the number of iterations till convergence.

Algorithm 1 Solving the Problem (27)-(29) Through the Proposed Exact Penalty Function Method Combined With the Parallel Stochastic Decomposition Algorithm

Input: $\mathbf{H}$, $\mathbf{G}$, $\mathbf{T}$, P_a, P_r, σ_r^2, σ_b^2, σ_e^2, ε, η, θ, μ, W, π_1, π_2, τ.

Output: p^*, α^*, β^*, ϵ^*.

Initialization: $\delta = 10$.

1: **repeat**
2: Set $t = 0$ and $\boldsymbol{\xi}^{(0)} = \mathbf{0}$.
3: **repeat**
4: Set $t := t + 1$.
5: Set $\rho^{(t)}$ and $\gamma^{(t)}$ according to (35).
6: Choose a random realization of $\mathbf{A}^{(t)}$, $\mathbf{B}^{(t)}$, and update $\boldsymbol{\xi}^{(t)}$ following (31)-(34).
7: **until** $\|\boldsymbol{\xi}^{(t)} - \boldsymbol{\xi}^{(t-1)}\|^2 < \varepsilon$
8: **if** $\alpha^{(t)}\lambda_{h,1}p^{(t)} + \alpha^{(t)}\sigma_r^2 + \beta^{(t)}(N_r - 1) - P_r \leq 0$ **then**
9: Set $p^* = p^{(t)}$, $\alpha^* = \alpha^{(t)}$, $\beta^* = \beta^{(t)}$, $\epsilon^* = \epsilon^{(t)}$.
10: Set $c = 1$.
11: **else**
12: Set $\delta := 10\delta$.
13: **end if**
14: **until** $\delta > 10^8$ or $c = 1$.

IV. Simulations

In this section, we study the performance of the proposed algorithm through numerical simulations. We compare the exact penalty function method with a conventional penalty function approach which solves the following optimization problem using the parallel stochastic decomposition algorithm

$$\min_{p,\alpha,\beta} \frac{1}{2}[-\log(1 + \mathrm{SNR}_b) + E_{\mathbf{A},\mathbf{B}}\{\log(1 + \mathrm{SNR}_e)\}] + \phi\Omega(p, \alpha, \beta) \quad (36)$$

$$\text{s.t. } 0 < p < P_a, \quad 0 < \alpha, \quad 0 < \beta \quad (37)$$

where $\phi > 0$ is a fixed penalty factor and $\Omega(p, \alpha, \beta) = [\max\{0, \alpha\lambda_{h,1}p + \alpha\sigma_r^2 + \beta(N_r - 1) - P_r\}]^2$. It is worth noting that one of the key differences between the proposed algorithm and the conventional method is that for the proposed algorithm, the penalty factor ϵ is a variable to be optimized in the problem in (27)-(29), while in the conventional method in (36)-(37), the penalty factor ϕ is fixed.

The performance of the proposed algorithm is also compared with the following two benchmark systems:

- The system where the full CSI of the source-eavesdropper channel $\mathbf{A}$ and the relay-eavesdropper channel $\mathbf{B}$ is available. Obviously, the secrecy rate achieved by this system is an upper-bound of that of the system in this article. Hereafter, we denote this system as "Full CSI".
- The conventional beamforming scheme, where the source beamforming vector is $c_1\mathbf{v}_{h,1}$, the source artificial noise vector is $c_2\mathbf{V}_{h,\bar{1}}\mathbf{w}_a$, and the relay beamforming matrix is $c_3\mathbf{v}_{g,1}\mathbf{u}_{h,1}^H$. Moreover, the relay node does not transmit any artificial noise ($\beta = 0$), and c_1, c_2, c_3 are optimized

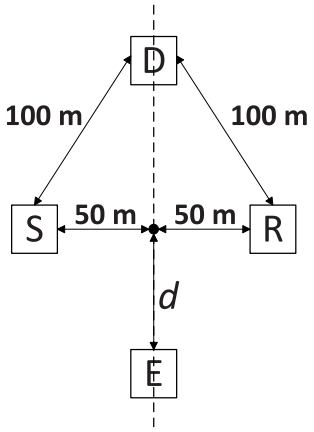


Fig. 2. Locations of the source, relay, destination, and eavesdropper.

with the full CSI knowledge of $\mathbf{A}$ and $\mathbf{B}$. This system is denoted as "Beamforming".

We consider a scenario illustrated in Fig. 2 where the source, relay, and destination nodes are located on the vertices of an equilateral triangle, while the position of the eavesdropper varies along the height line of this triangle. This scenario allows us to investigate the achievable secrecy rate with respect to the location of the eavesdropper. In particular, the distance between node i and node j ($i \neq j$ and $i, j \in \{S, R, D, E\}$) is denoted as D_{ij}. We choose $D_{SR} = D_{RD} = D_{SD} = 100$ m. Thus, from Fig. 2, we have $D_{SE} = D_{RE} = \sqrt{d^2 + 50^2}$ m. The channel matrices are modeled as $\mathbf{H} = D_{SR}^{-\zeta/2}\bar{\mathbf{H}}$, $\mathbf{G} = D_{RD}^{-\zeta/2}\bar{\mathbf{G}}$, $\mathbf{T} = D_{SD}^{-\zeta/2}\bar{\mathbf{T}}$, $\mathbf{A} = D_{SE}^{-\zeta/2}\bar{\mathbf{A}}$, and $\mathbf{B} = D_{RE}^{-\zeta/2}\bar{\mathbf{B}}$, where ζ is the path loss exponent, $\bar{\mathbf{H}}$, $\bar{\mathbf{G}}$, $\bar{\mathbf{T}}$, $\bar{\mathbf{A}}$, and $\bar{\mathbf{B}}$ denote the small-scale Rayleigh fading. Except for the last simulation example, $\bar{\mathbf{H}}$, $\bar{\mathbf{G}}$, $\bar{\mathbf{T}}$, $\bar{\mathbf{A}}$, and $\bar{\mathbf{B}}$ have i.i.d. complex Gaussian entries with zero-mean and variances of $1/N_a$, $1/N_r$, $1/N_a$, $1/N_a$, and $1/N_r$, respectively. The SNRs of the source-relay and relay-destination links are defined as $\text{SNR}_{SR} = P_a/(D_{SR}^{\zeta}\sigma_r^2)$ and $\text{SNR}_{RD} = P_r/(D_{RD}^{\zeta}\sigma_b^2)$, respectively. In the simulations, we choose $\zeta = 3$ and set $\text{SNR}_{SR} = \text{SNR}_{RD} = \text{SNR}$.

As stated in Remark 5, for the proposed algorithm and the conventional penalty function approach (36)-(37), since under the simulation setup $\sigma_{SD}^2 = D_{SD}^{-\zeta}/N_a$ and $\sigma_{RD}^2 = D_{RD}^{-\zeta}/N_r$, the distributions of the entries in $\mathbf{A}$ and $\mathbf{B}$ are assumed to be

$$\mathcal{CN}(0, D_{SD}^{-\zeta}/N_a), \quad \mathcal{CN}(0, D_{RD}^{-\zeta}/N_r) \tag{38}$$

respectively. For all simulation examples, we choose $N_e = 3$, $\pi_1 = 0.7$, $\pi_2 = 0.71$, $\varepsilon = 10^{-9}$, and $N_a = N_r = N_b = N$. All the numerical simulation results are averaged over $1,000$ independent channel realizations.

A. Example 1: System Secrecy Rate Versus SNR *With Various Number of Antennas*

In the first example, we investigate the system secrecy rate at various N. We set $d = 50\sqrt{3}$ m, and thus, $D_{SE} = D_{RE} = 100$ m. The secrecy rates achieved by the four systems studied versus SNR are shown in Fig. 3 with $N = 3$. It can be seen from Fig. 3 that the secrecy rate of

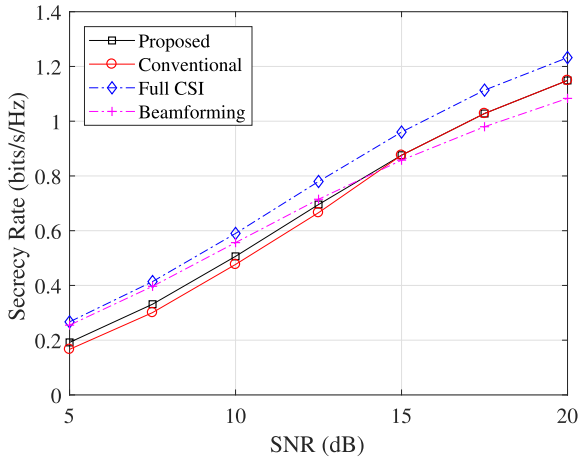


Fig. 3. Example 1: System secrecy rate versus SNR, $N = 3$.

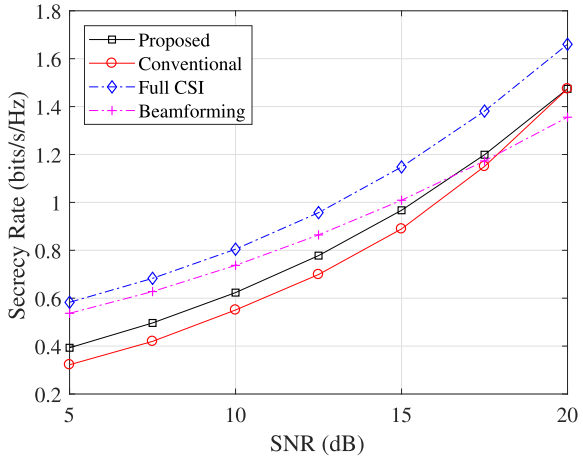


Fig. 4. Example 1: System secrecy rate versus SNR, $N = 5$.

the proposed algorithm increases with SNR. This indicates that using the proposed algorithm, secure communication is achievable even without the knowledge of the eavesdropper's channels. We can also observe from Fig. 3 that the secrecy rate gap between the two full CSI systems increases with SNR. At high SNRs, the proposed algorithm has a higher secrecy rate than the beamforming system with full CSI and $\beta = 0$. This indicates the importance of employing jamming signal at the relay node to confuse the eavesdropper.

Fig. 4 shows the secrecy rates achieved by the four systems tested with $N = 5$. From Figs. 3 and 4, we can see that the system secrecy rate increases with the number of antennas at the legitimate users. Figs. 3 and 4 demonstrate that at low to medium SNRs, the proposed exact penalty function method has a higher system secrecy rate than the conventional penalty function approach. Moreover, the gap between these two methods increases with N.

B. Example 2: Power Allocation for the Transmission of the AN Versus SNR

In the second example, we study the power allocated for the transmission of the AN at the source node (i.e., $P_a - p$

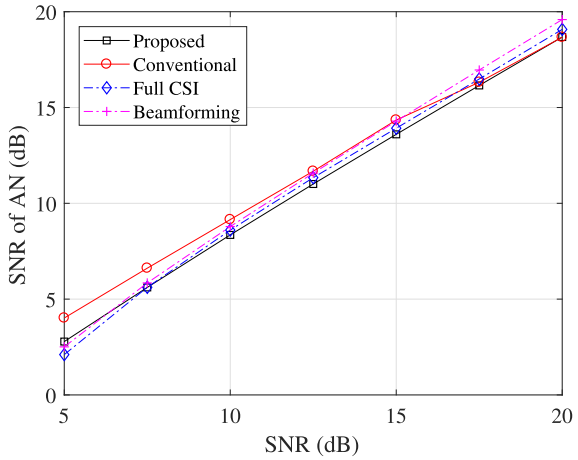


Fig. 5. Example 2: Power allocation for the AN at the source node versus SNR, $N = 3$.

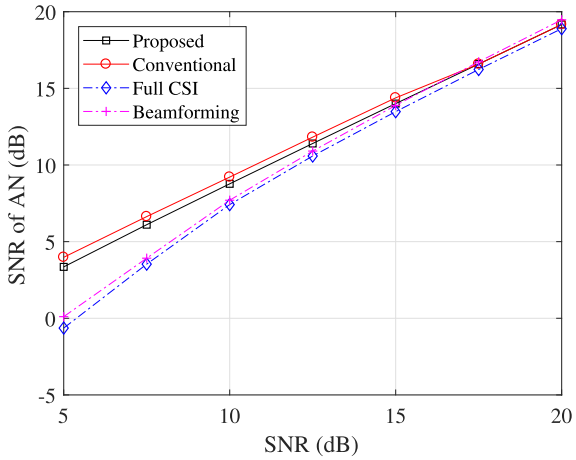


Fig. 7. Example 2: Power allocation for the AN at the source node versus SNR, $N = 5$.

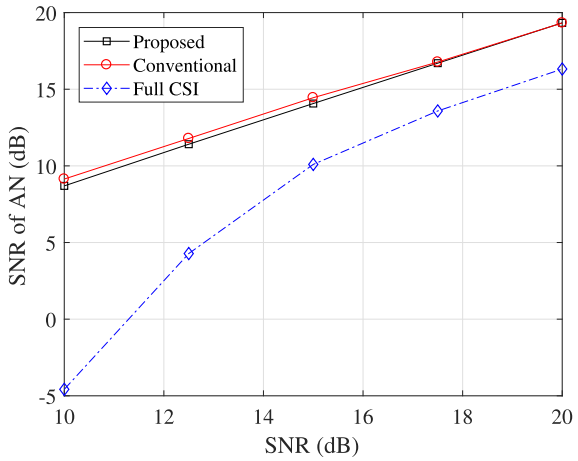


Fig. 6. Example 2: Power allocation for the AN at the relay node versus SNR, $N = 3$.

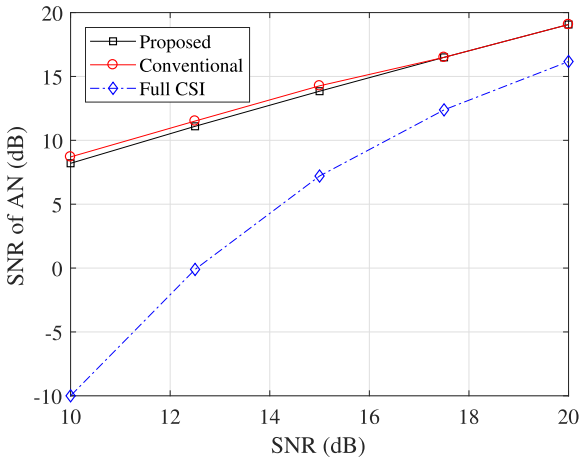


Fig. 8. Example 2: Power allocation for the AN at the relay node versus SNR, $N = 5$.

in (1)) and the relay node (which is $\beta(N_r - 1)$ in (8)). In this example, we set $d = 50\sqrt{3}$ m, so $D_{SE} = D_{RE} = 100$ m. Fig. 5 shows the AN power allocation by the four algorithms tested at the source node versus SNR with $N = 3$. It can be seen from Fig. 5 that for all four algorithms, the power allocated for transmitting the AN at the source node increases with SNR.

We can also observe from Fig. 5 that in the low and medium SNR region, the conventional penalty function approach allocates more power for transmitting the AN compared with the other three schemes. This is the reason for the lower system secrecy rate achieved by the conventional approach. Similarly, at high SNRs, the power of AN allocated by the beamforming system with the full CSI and $\beta = 0$ is higher than the other schemes, which leads to a lower secrecy rate of this system.

Fig. 6 shows the AN power allocation at the relay node versus SNR for $N = 3$. The beamforming system is not shown in this figure, as for this algorithm the AN power is always zero at the relay node. Similar to Fig. 5, we find out that to maximize the system secrecy rate, the power allocated for transmitting the AN increases with SNR at the relay node. It can be seen that compared with the other two algorithms, the full CSI system allocates much less power at the relay node for the AN. This is due to the benefit of the eavesdropper's CSI in the full CSI system, which requires weaker AN at the relay node.

The power of AN allocated by the four algorithms at the source node is shown in Fig. 7 for $N = 5$, where similar trends can be observed as Fig. 5. Interestingly, by comparing Fig. 5 with Fig. 7, we can see that the gap between the two systems with the eavesdropper's CSI and two systems without CSI increases at the low SNR range in terms of the power of the AN. This explains the bigger secrecy rate gap between the corresponding algorithms in Fig. 4 ($N = 5$) compared with Fig. 3 ($N = 3$).

Fig. 8 illustrates the AN power at the relay node for $N = 5$. By comparing Fig. 8 with Fig. 6, we can see that by increasing the number of antennas at the legitimate users, less power at the relay node is allocated for transmitting the AN when the eavesdropper's CSI is known. However, for the proposed algorithm, the AN power allocation remains almost same when N is increased from 3 to 5.

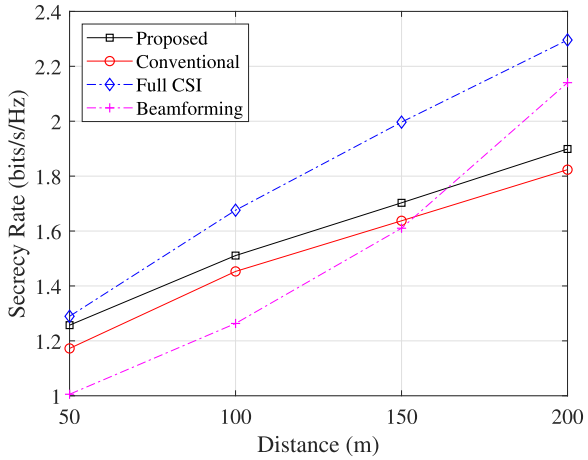


Fig. 9. Example 3: System secrecy rate versus d, SNR $= 15$ dB, $N = 5$.

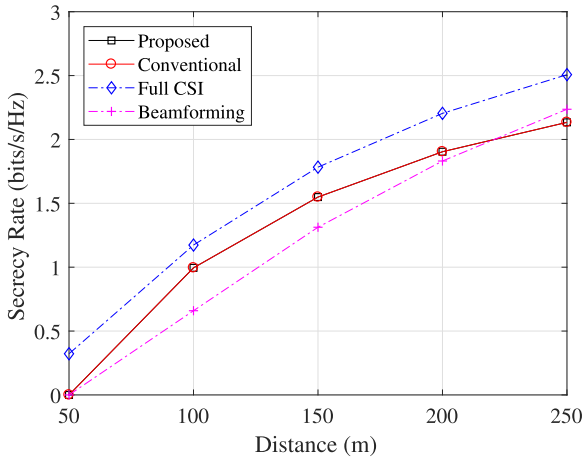


Fig. 10. Example 3: System secrecy rate versus d, SNR $= 15$ dB, $N = 3$.

C. Example 3: System Secrecy Rate Versus the Distance Between the Eavesdropper and the Source Node

In the third example, we investigate the system secrecy rate at various locations of the eavesdropper as in Fig. 2. We set SNR $=$ 15 dB and change d between 50 m and 200 m to simulate the variation of the eavesdropper's location. Note that the position of the eavesdropper and the CSI of $\mathbf{A}$ and $\mathbf{B}$ are unknown to the source and relay nodes for the proposed exact penalty algorithm and the conventional penalty approach.

Figs. 9 and 10 show the system secrecy rate versus d with $N = 5$ and $N = 3$, respectively. We can observe from these two figures that for all four systems tested, the secrecy rate increases with d. This is due to the fact that the eavesdropper's channel becomes weaker when d increases. It can be seen from Fig. 9 that the secrecy rate of the proposed scheme is always higher than that of the conventional approach at the range of d tested. Interestingly, despite unknown $\mathbf{A}$ and $\mathbf{B}$, the proposed scheme can achieve secure communication even at $d = 50$ m, where $D_{SE} < D_{SD}$ and $D_{RE} < D_{RD}$. This is mainly contributed by a larger number of antennas at the legitimate users compared with the eavesdropper, which not only increases the decoding capability of the destination node, but also enhances the capability of the source and relay nodes in transmitting the AN to confuse the eavesdropper. While from Fig. 10 we can see that except for the system with full CSI, the other three schemes cannot achieve secured communication, due to a weaker channel of the legitimate users. Figs. 9 and 10 also demonstrate the importance of transmitting the AN ($\beta \neq 0$) at short and medium range of d.

D. Example 4: System Secrecy Rate Versus SNR With Channel Fading Type Mismatch

In the last simulation example, we study the performance of the proposed approach for the scenario where the actual fading type of the eavesdropper's channels is different to the presumed one. The actual channel matrices $\mathbf{A}$ and $\mathbf{B}$ satisfy the Gaussian-Kronecker model as $\mathbf{A} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_a \otimes \mathbf{\Psi}_a)$ and $\mathbf{B} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_b \otimes \mathbf{\Psi}_b)$, where $\otimes$ denotes the matrix Kronecker product, $\mathbf{\Sigma}_a$ and $\mathbf{\Psi}_a$ are the row and column covariance matrices of $\mathbf{A}$, respectively, and similarly, $\mathbf{\Sigma}_b$ and $\mathbf{\Psi}_b$ are the row and column covariance matrices of $\mathbf{B}$, respectively. In other words, the entries of $\mathbf{A}$ and $\mathbf{B}$ have correlated Rayleigh fading, which is different from the i.i.d. Rayleigh fading assumption in (38).

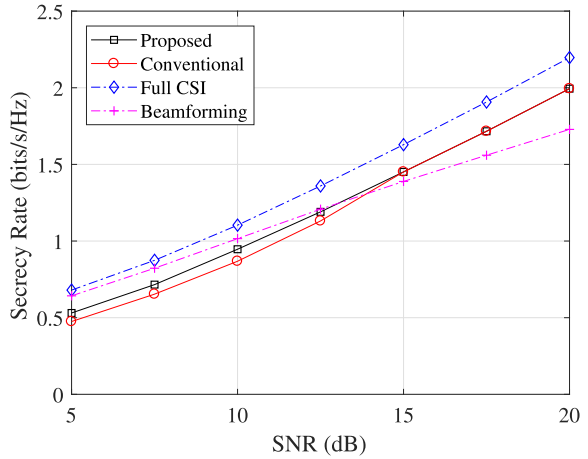


Fig. 11. Example 4: System secrecy rate versus SNR, $N = 5$.

The four covariance matrices are modeled as

$$[\mathbf{\Sigma}_a]_{m,n} = \frac{\sigma_a^{|m-n|}}{D_{SE}^{\zeta} N_a}, \quad m,n = 1,\cdots,N_e$$
$$[\mathbf{\Psi}_a]_{m,n} = \psi_a^{|m-n|}, \quad m,n = 1,\cdots,N_a$$
$$[\mathbf{\Sigma}_b]_{m,n} = \frac{\sigma_b^{|m-n|}}{D_{RE}^{\zeta} N_r}, \quad m,n = 1,\cdots,N_e$$
$$[\mathbf{\Psi}_b]_{m,n} = \psi_b^{|m-n|}, \quad m,n = 1,\cdots,N_b$$

where $|\cdot|$ denotes the absolute value and we set $d = 50\sqrt{3}$ m and $\sigma_a = \psi_a = \sigma_b = \psi_b = 0.7$. Fig. 11 illustrates the secrecy rate of the four systems with $N = 5$. It can be seen that despite the mismatch between the actual and presumed fading type of the eavesdropper's channels, the proposed algorithm achieves a good performance. Compared with Fig. 4, we can observe that the systems in this example have a higher secrecy rate. This is due to the correlation in the eavesdropper's channels,

which reduces the rate from the source and relay nodes to the eavesdropper.

V. Conclusion

A practical relay-aided secure communication scheme has been investigated in this article, where the knowledge of the eavesdropper's channels is unavailable. The system design was formulated as a stochastic programming problem with a nonconvex objective function and nonconvex constraints. To solve this challenging problem, a novel computational method has been developed to optimize the power allocation at the source and relay nodes for the transmission of the AN and the signal-of-interest. Firstly, an exact penalty function method was adopted to append the nonconvex constraint into the objective function, which leads to a simpler optimization problem with only convex constraints. Then, a parallel stochastic decomposition method was introduced to solve the resulted problem. Numerical simulations have shown that the proposed system is capable of ensuring the secrecy of relay-aided transmission without the knowledge of the eavesdropper's channel state information. Moreover, the proposed scheme achieves a higher secrecy rate than a conventional penalty function approach.

Appendix

In the Appendix, we provide details of the surrogate functions in (33) and the incremental numbers in (34) for all variables in $\boldsymbol{\xi}$. Firstly, we rewrite the objective function in (27) without the expectation operation as

$$\begin{aligned}&-\log(1+\mathrm{SNR}_b)+\log(1+\mathrm{SNR}_e)+\epsilon^{-\eta}\Delta(\boldsymbol{\xi})+\delta\epsilon^{\theta}\\&\quad=-\log\left|p\mathbf{g}_{\alpha}\mathbf{g}_{\alpha}^{H}+\mathrm{diag}[\alpha\lambda_{g,1}\sigma_r^2+\sigma_b^2,\mathbf{R}_{3,p}]\right|\\&\qquad+\log(\alpha\lambda_{g,1}\sigma_r^2+\sigma_b^2)+\log|\mathbf{R}_{3,p}|+\log|\mathbf{h}\mathbf{h}^{H}+\mathbf{R}_v|\\&\qquad-\log|\mathbf{R}_v|+\epsilon^{-\eta}\Delta(\boldsymbol{\xi})+\delta\epsilon^{\theta}\end{aligned}\tag{39}$$

where $|\cdot|$ denotes the matrix determinant, $\mathbf{h}$ and $\mathbf{R}_v$ are defined in (16) and (19), respectively, and

$$\begin{aligned}\mathbf{R}_{3,p}&=\sigma_b^2\mathbf{I}_{N_b}+(P_a-p)\mathbf{X}_4\\\mathbf{X}_4&=\frac{1}{N_a-1}\mathbf{T}\mathbf{V}_{h,\bar{1}}\mathbf{V}_{h,\bar{1}}^{H}\mathbf{T}^{H}\\\mathbf{g}_{\alpha}&=[\sqrt{\alpha\lambda_{h,1}\lambda_{g,1}},(\mathbf{T}\mathbf{v}_{h,1})^{H}]^{H}.\end{aligned}$$

The surrogate functions consist of the convex part of (39) with respect to a particular variable and the linearization of the nonconvex part.

Using (39), the surrogate function for p is given below

$$\begin{aligned}&\hat{f}_1(p;\mathbf{C}^{(t)})\\&=\rho^{(t)}\Big[-\log\left|p\mathbf{g}_{\alpha}\mathbf{g}_{\alpha}^{H}+\mathrm{diag}[\alpha^{(t)}\lambda_{g,1}\sigma_r^2+\sigma_b^2,\mathbf{R}_{3,p}]\right|\\&\quad-\log|(P_a-p)\mathbf{X}_1+\sigma_e^2\mathbf{I}_{N_e}|+(\epsilon^{(t)})^{-\eta}\Delta(p,\alpha^{(t)},\beta^{(t)},\epsilon^{(t)})\Big]\\&\quad+\rho^{(t)}(p-p^{(t)})\big[\mathrm{tr}(\mathbf{R}^{-1}(\mathbf{h}_t\mathbf{h}_t^{H}-\mathrm{diag}[\mathbf{X}_1,\mathbf{0}]))\\&\quad-\mathrm{tr}(\mathbf{R}_3^{-1}\mathbf{X}_4)\big]\\&\quad+(1-\rho^{(t)})(p-p^{(t)})q_1^{(t-1)}+\tau(p-p^{(t)})^2\end{aligned}$$

where

$$\begin{aligned}\mathbf{X}_1&=\frac{1}{N_a-1}\mathbf{A}^{(t)}\mathbf{V}_{h,\bar{1}}\mathbf{V}_{h,\bar{1}}^{H}(\mathbf{A}^{(t)})^{H}\\\mathbf{X}_2&=\mathbf{B}^{(t)}\mathbf{V}_{g,\bar{1}}\mathbf{V}_{g,\bar{1}}^{H}(\mathbf{B}^{(t)})^{H}\\\mathbf{X}_3&=\sigma_r^2\mathbf{B}^{(t)}\mathbf{v}_{g,1}\mathbf{v}_{g,1}^{H}(\mathbf{B}^{(t)})^{H}\\\mathbf{R}_3&=\sigma_b^2\mathbf{I}_{N_b}+(P_a-p^{(t)})\mathbf{X}_4\\\mathbf{R}&=\mathrm{diag}[\mathbf{R}_1,\mathbf{R}_2]+p^{(t)}\mathbf{h}_t\mathbf{h}_t^{H}\\\mathbf{R}_1&=\sigma_e^2\mathbf{I}_{N_e}+(P_a-p^{(t)})\mathbf{X}_1\\\mathbf{R}_2&=\sigma_e^2\mathbf{I}_{N_e}+\beta^{(t)}\mathbf{X}_2+\alpha^{(t)}\mathbf{X}_3\\\mathbf{h}_t&=\Big[\big(\mathbf{A}^{(t)}\mathbf{v}_{h,1}\big)^{T},\sqrt{\alpha^{(t)}\lambda_{h,1}}\big(\mathbf{B}^{(t)}\mathbf{v}_{g,1}\big)^{T}\Big]^{T}.\end{aligned}$$

Secondly, the surrogate function for α is given by

$$\begin{aligned}&\hat{f}_2(\alpha;\mathbf{C}^{(t)})\\&=\rho^{(t)}\big[-\log(\alpha p^{(t)}\lambda_{h,1}\lambda_{g,1}+v(\alpha\lambda_{g,1}\sigma_r^2+\sigma_b^2))\\&\quad-\log|\mathbf{R}_{2,\alpha}|+(\epsilon^{(t)})^{-\eta}\Delta(p^{(t)},\alpha,\beta^{(t)},\epsilon^{(t)})\big]\\&\quad+\rho^{(t)}(\alpha-\alpha^{(t)})[\lambda_{g,1}\sigma_r^2/(\alpha^{(t)}\lambda_{g,1}\sigma_r^2+\sigma_b^2)+\mathrm{tr}(\mathbf{R}^{-1}\mathbf{M})]\\&\quad+(1-\rho^{(t)})(\alpha-\alpha^{(t)})q_2^{(t-1)}+\tau(\alpha-\alpha^{(t)})^2\end{aligned}$$

where

$$\begin{aligned}v&=1+p^{(t)}\mathbf{v}_{h,1}^{H}\mathbf{T}^{H}\Big((P_a-p^{(t)})\mathbf{X}_4+\sigma_b^2\mathbf{I}_{N_b}\Big)^{-1}\mathbf{T}\mathbf{v}_{h,1}\\\mathbf{R}_{2,\alpha}&=\sigma_e^2\mathbf{I}_{N_e}+\beta^{(t)}\mathbf{X}_2+\alpha\mathbf{X}_3\\\mathbf{M}&=\begin{pmatrix}\mathbf{0}&\mathbf{X}_6\\\mathbf{X}_6^{H}&\mathbf{X}_5\end{pmatrix}\\\mathbf{X}_5&=\mathbf{X}_3+p^{(t)}\lambda_{h,1}\mathbf{B}^{(t)}\mathbf{v}_{g,1}\mathbf{v}_{g,1}^{H}(\mathbf{B}^{(t)})^{H}\\\mathbf{X}_6&=\frac{p^{(t)}}{2}\sqrt{\frac{\lambda_{h,1}}{\alpha^{(t)}}}\mathbf{A}^{(t)}\mathbf{v}_{h,1}\mathbf{v}_{g,1}^{H}(\mathbf{B}^{(t)})^{H}.\end{aligned}$$

Thirdly, for the variable β, the surrogate function is given by

$$\begin{aligned}&\hat{f}_3(\beta;\mathbf{C}^{(t)})\\&\quad=\rho^{(t)}\big[-\log|\mathbf{R}_{2,\beta}|+(\epsilon^{(t)})^{-\eta}\Delta(p^{(t)},\alpha^{(t)},\beta,\epsilon^{(t)})\big]\\&\qquad+\rho^{(t)}(\beta-\beta^{(t)})\mathrm{tr}(\mathbf{R}^{-1}\mathrm{diag}[\mathbf{0},\mathbf{X}_2])\\&\qquad+(1-\rho^{(t)})(\beta-\beta^{(t)})q_3^{(t-1)}+\tau(\beta-\beta^{(t)})^2\end{aligned}$$

where $\mathbf{R}_{2,\beta}=\sigma_e^2\mathbf{I}_{N_e}+\beta\mathbf{X}_2+\alpha^{(t)}\mathbf{X}_3$. Finally, the surrogate function for ϵ can be written as

$$\begin{aligned}\hat{f}_4(\epsilon;\mathbf{C}^{(t)})=\rho^{(t)}\big[&\delta\epsilon^{\theta}+(\epsilon-\epsilon^{(t)})\big(-\eta(\epsilon^{(t)})^{-\eta-1}\Delta^{(t)}\\&-2\mu(\epsilon^{(t)})^{\mu-\eta-1}W(\Delta^{(t)})^{\frac{1}{2}}\big)\big]\\&+(1-\rho^{(t)})(\epsilon-\epsilon^{(t)})q_4^{(t-1)}+\tau(\epsilon-\epsilon^{(t)})^2.\end{aligned}$$

where $\Delta^{(t)}\triangleq\Delta(p^{(t)},\alpha^{(t)},\beta^{(t)},\epsilon^{(t)})$.

For each variable in $\boldsymbol{\xi}$, the accumulation numbers are updated recursively according to (34) based on the gradient of (39). Firstly, the accumulation number for p can be written as

$$\begin{aligned}q_1^{(t)}=(1-\rho^{(t)})q_1^{(t-1)}+\rho^{(t)}\big[&\mathrm{tr}(\mathbf{R}^{-1}(\mathbf{h}_t\mathbf{h}_t^{H}-\mathrm{diag}[\mathbf{X}_1,\mathbf{0}]))\\&-\mathrm{tr}(\mathbf{R}_3^{-1}\mathbf{X}_4)+\mathrm{tr}(((P_a-p^{(t)})\mathbf{X}_1+\sigma_e^2\mathbf{I}_{N_e})^{-1}\mathbf{X}_1)\\&-\mathrm{tr}(\mathbf{Z}^{-1}(\mathbf{g}\mathbf{g}^{H}+\mathrm{diag}[0,-\mathbf{X}_4]))\\&+2(\epsilon^{(t)})^{-\eta}(\Delta^{(t)})^{\frac{1}{2}}\alpha^{(t)}\lambda_{h,1}\big]\end{aligned}$$

where

$$\mathbf{g} = [\sqrt{\alpha^{(t)}\lambda_{h,1}\lambda_{g,1}}, (\mathbf{T}\mathbf{v}_{h,1})^H]^H$$
$$\mathbf{Z} = p^{(t)}\mathbf{g}\mathbf{g}^H + \mathrm{diag}[\alpha^{(t)}\lambda_{g,1}\sigma_r^2 + \sigma_b^2, \mathbf{R}_3].$$

Secondly, for variable α, we have

$$q_2^{(t)} = (1-\rho^{(t)})q_2^{(t-1)} + \rho^{(t)}\big[\lambda_{g,1}\sigma_r^2/(\alpha^{(t)}\lambda_{g,1}\sigma_r^2 + \sigma_b^2) + \mathrm{tr}(\mathbf{R}^{-1}\mathbf{M}) - \mathrm{tr}(\mathbf{R}_2^{-1}\mathbf{X}_3) - k/(k\alpha^{(t)} + v\sigma_b^2) + 2(\epsilon^{(t)})^{-\eta}(\Delta^{(t)})^{\frac{1}{2}}(\lambda_{h,1}p^{(t)} + \sigma_r^2)\big]$$

where $k = p^{(t)}\lambda_{h,1}\lambda_{g,1} + v\lambda_{g,1}\sigma_r^2$. Thirdly, the accumulation number for β is given by

$$q_3^{(t)} = (1-\rho^{(t)})q_3^{(t-1)} + \rho^{(t)}\big[\mathrm{tr}(\mathbf{R}^{-1}\mathrm{diag}[\mathbf{0}, \mathbf{X}_2]) - \mathrm{tr}(\mathbf{R}_2^{-1}\mathbf{X}_2) + 2(\epsilon^{(t)})^{-\eta}(\Delta^{(t)})^{\frac{1}{2}}(N_r - 1)\big].$$

Finally, for ϵ, there is

$$q_4^{(t)} = (1-\rho^{(t)})q_4^{(t-1)} + \rho^{(t)}\big[-\eta(\epsilon^{(t)})^{-\eta-1}\Delta^{(t)} - 2\mu(\epsilon^{(t)})^{\mu-\eta-1}W(\Delta^{(t)})^{\frac{1}{2}} + \delta\theta(\epsilon^{(t)})^{\theta-1}\big].$$

REFERENCES

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

[2] J. Liu, J. Dai, Y. Shi, W. Sun, and N. Kato, "On physical layer security in finite-area wireless networks: An analysis framework," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2017, pp. 1–5.

[3] S. Zhang, W. Sun, J. Liu, and K. Nei, "Physical layer security in large–scale probabilistic caching: Analysis and optimization," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1484–1487, Sep. 2019.

[4] B. He, X. Zhou, and A. L. Swindlehurst, "On secrecy metrics for physical layer security over quasi-static fading channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6913–6924, Oct. 2016.

[5] L. Lai and H. El Gamal, "The relay–eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4005–4019, Sep. 2008.

[6] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[7] H.-M. Wang and X.-G. Xia, "Enhancing wireless secrecy via cooperation: Signal design and optimization," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 47–53, Dec. 2015.

[8] 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.

[9] J. Li, A. P. Petropulu, and S. Weber, "On cooperative relaying schemes for wireless physical layer security," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4985–4997, Oct. 2011.

[10] A. M. Akhtar, A. Behnad, and X. Wang, "On the secrecy rate achievability in dual-hop amplify-and-forward relay networks," *IEEE Wireless Commun. Lett.*, vol. 3, no. 5, pp. 493–496, Oct. 2014.

[11] M. Li, Y. Zhang, L. Wang, M. Song, and Z. Han, "Incentive design for collaborative jamming using contract theory in physical layer security," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Jul. 2016, pp. 1–6.

[12] L. Fan, R. Zhao, F.-K. Gong, N. Yang, and G. K. Karagiannidis, "Secure multiple amplify-and-forward relaying over correlated fading channels," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2811–2820, Jul. 2017.

[13] S.-I. Kim, I.-M. Kim, and J. Heo, "Secure transmission for multi-user relay networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3724–3737, Jul. 2015.

[14] K.-H. Park, T. Wang, and M.-S. Alouini, "On the jamming power allocation for secure amplify-and-forward relaying via cooperative jamming," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1741–1750, Sep. 2013.

[15] M. F. Marzban, R. Chabaan, N. Al-Dhahir, and A. El Shafie, "Securing OFDM-based wireless links using temporal artificial-noise injection," in *Proc. 15th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2018, pp. 1–6.

[16] J. Huang and A. L. Swindlehurst, "Cooperative jamming for secure communications in MIMO relay networks," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4871–4884, Oct. 2011.

[17] H. Deng, H.-M. Wang, W. Guo, and W. Wang, "Secrecy transmission with a helper: To relay or to jam," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 2, pp. 293–307, Feb. 2015.

[18] P. M. Shemi, M. G. Jibukumar, and M. A. Ali, "Artificial noise aided secrecy enhancement in amplify-and-forward relay networks," in *Proc. 5th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2018, pp. 192–196.

[19] X. Zhou, M. Qiu, S.-C. Lin, and Y.-W.-P. Hong, "On the jamming power allocation and signal design in DF relay networks," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2013, pp. 1268–1272.

[20] N. Zhang, N. Lu, N. Cheng, J. W. Mark, and X. S. Shen, "Cooperative spectrum access towards secure information transfer for CRNs," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2453–2464, Nov. 2013.

[21] Y. Li, R. Zhao, Y. Wang, G. Pan, and C. Li, "Artificial noise aided precoding with imperfect CSI in full-duplex relaying secure communications," *IEEE Access*, vol. 6, pp. 44107–44119, 2018.

[22] J. Hu, Y. Cai, N. Yang, X. Zhou, and W. Yang, "Artificial-noise-aided secure transmission scheme with limited training and feedback overhead," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 193–205, Jan. 2017.

[23] N. R. Zhou, Z. J. Kang, and X. R. Liang, "Secure cooperative communication via artificial noise for wireless two-hop relaying networks," *Wireless Pers. Commun.*, vol. 82, no. 3, pp. 1759–1771, Jun. 2015.

[24] W. Liu, X. Zhou, and S. Durrani, "Wireless-powered friendly jammer for physical layer security," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2015, pp. 1–5.

[25] W. Liu, X. Zhou, S. Durrani, and P. Popovski, "Secure communication with a wireless-powered friendly jammer," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 401–415, Jan. 2016.

[26] C. Liu, N. Yang, R. Malaney, and J. Yuan, "Artificial-noise-aided transmission in multi-antenna relay wiretap channels with spatially random eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7444–7456, Nov. 2016.

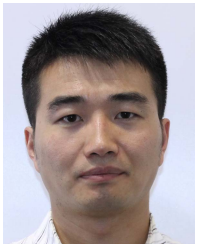
[27] O. Cepheli, S. Tedik, and G. K. Kurt, "A high data rate wireless communication system with improved secrecy: Full duplex beamforming," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 1075–1078, Jun. 2014.

[28] Y. Yang, G. Scutari, D. P. Palomar, and M. Pesavento, "A parallel decomposition method for nonconvex stochastic multi-agent optimization problems," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2949–2964, Jun. 2016.

[29] C. Yu, K. L. Teo, L. Zhang, and Y. Bai, "A new exact penalty function method for continuous inequality constrained optimization problems," *J. Ind. Manage. Optim.*, vol. 6, no. 4, pp. 895–910, Oct. 2010.

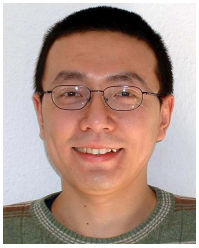
[30] B. Li, C. J. Yu, K. L. Teo, and G. R. Duan, "An exact penalty function method for continuous inequality constrained optimal control problem," *J. Optim. Theory Appl.*, vol. 151, no. 2, pp. 260–291, Nov. 2011.

[31] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*. New York, NY, USA: Spring Street, 2007.



Bin Li (Senior Member, IEEE) received the bachelor's degree in automation and the master's degree in control science and engineering from the Harbin Institute of Technology, China, in 2005 and 2008, respectively, and the Ph.D. degree in mathematics and statistics from Curtin University, Australia, in 2011. From 2012 to 2014, he was a Research Associate with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Australia. From 2014 to 2017, he was a Research Fellow with the Department of Mathematics and Statistics, Curtin University. He is currently a Research Professor with the School of Aeronautics and Astronautics, Sichuan University, China. His research interests include signal processing, wireless communications, optimization, and optimal control.

Meiying Zhang received the bachelor's degree in electrical engineering and automation from the University of Jinan, China, in 2014. She is currently pursuing the master's degree in control theory and control engineering with the College of Electrical Engineering, Sichuan University, China. Her research interests include signal processing and wireless communications.



Yue Rong (Senior Member, IEEE) received the Ph.D. degree *(summa cum laude)* in electrical engineering from the Darmstadt University of Technology, Darmstadt, Germany, in 2005.

He was a Post-Doctoral Researcher with the Department of Electrical Engineering, University of California at Riverside, Riverside, from February 2006 to November 2007. Since December 2007, he has been with Curtin University, Bentley, Australia, where he is currently a Professor. His research interests include signal processing for communications, wireless communications, underwater acoustic communications, underwater optical wireless communications, applications of linear algebra and optimization methods, and statistical and array signal processing. He has published more than 180 journals and conference papers in these areas. He was a TPC Member of the IEEE ICC, IEEE GlobalSIP, EUSIPCO, IEEE ICCC, WCSP, IWCMC, and ChinaCom. He was a recipient of the Best Paper Award at the 2011 International Conference on Wireless Communications and Signal Processing, the Best Paper Award at the 2010 Asia–Pacific Conference on Communications, and the Young Researcher of the Year Award of the Faculty of Science and Engineering at Curtin University in 2010. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2014 to 2018, an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS from 2012 to 2014, and a Guest Editor of the Special Issue on Theories and Methods for Advanced Wireless Relays of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is a Senior Area Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Zhu Han (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University in 1997 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was a Research and Development Engineer with JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, ID. He is currently the John and Rebecca Moores Professor of the Electrical and Computer Engineering Department and the Computer Science Department, University of Houston, TX. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. He was an IEEE Communications Society Distinguished Lecturer from 2015 to 2018, an AAAS fellow since 2019, and ACM Distinguished Member since 2019. He received the NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, the IEEE Leonard G. Abraham Prize in the field of Communications Systems (Best Paper Award in IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS) in 2016, and several best paper awards in IEEE conferences. He is the 1% highly cited researcher since 2017 according to Web of Science. He is also the winner of the 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks."