Joint Transceiver Design for Secure Downlink Communications Over an Amplify-and-Forward MIMO Relay

Yunlong Cai, *Senior Member, IEEE*, Qingjiang Shi, Benoit Champagne, *Senior Member, IEEE*, and Geoffrey Ye Li, *Fellow, IEEE*

***Abstract*— This paper addresses joint transceiver design for secure downlink communications over a multiple-input multiple-output relay system in the presence of multiple legitimate users and malicious eavesdroppers. Specifically, we jointly optimize the base station (BS) beamforming matrix, the relay station (RS) amplify-and-forward transformation matrix, and the covariance matrix of artificial noise, so as to maximize the system worst-case secrecy rate in the presence of the colluding eavesdroppers under power constraints at the BS and the RS, as well as quality of service constraints for the legitimate users. This problem is very challenging due to the highly coupled design variables in the objective function and constraints. By adopting a series of transformation, we first derive an equivalent problem that is more tractable than the original one. Then, we propose and fully develop a novel algorithm based on the penalty concave-convex procedure (penalty-CCCP) to solve the equivalent problem, where the difficult coupled constraint is penalized into the objective and the resulting nonconvex problem is solved at each iteration by resorting to the CCCP method. It is shown that the proposed joint transceiver design algorithm converges to a stationary solution of the original problem. Finally, our simulation results reveal that the proposed algorithm achieves better performance than other recently proposed transceiver designs.**

***Index Terms*— Beamforming, transceiver design, MIMO relay, physical layer security, penalty-CCCP.**

I. INTRODUCTION

THE explosive growth in the number and types of mobile devices poses a formidable challenge to the design of wireless communication systems, which need to provide a high level of message security in the presence of eavesdroppers who try to retrieve information from ongoing transmissions [1]. As a complement to traditional encryption, physical layer security (PLS) [2]–[6] aims to exploit wireless channel characteristics through signal processing techniques in order to improve the security of wireless transmission. It has become a very active research area lately and is called upon to play a key role in future 5G networks [7], [8].

Using multiple-input multiple-output (MIMO) relays in multiuser multi-antenna cellular systems can further increase the link quality, reliability, and data rate. Among the different available relaying strategies, the amplify-and-forward (AF) scheme is often considered to be the most practical approach due to its low-complexity of operation at the relay station (RS) [9]. However, with the introduction of one or more relay nodes in the transmission, the confidentiality of the transmitted information may be more easily compromised.

To address this issue, a number of schemes have been proposed for MIMO relay systems that offer improved security against eavesdroppers [10]–[24]. In particular, as in [10]–[12], secrecy beamforming and precoding schemes can be divided into two categories, depending on whether the relay is trusted or untrusted. Beamforming and precoding for untrusted MIMO relay systems have been investigated in [13]–[15]. Transceiver design for trusted MIMO relays has received more attention due to its wider spread application [16]–[25]. Ding *et al.* [16] have combined interference alignment with cooperative jamming in order to ensure secure transmission to the legitimate receiver in the presence of an eavesdropper. A generalized singular-value decomposition (GSVD) method has been used by Huang and Swindlehurst [17] to develop a cooperative jamming (CJ) scheme for secure communications with MIMO relays. Besides, secure relay beamforming for simultaneous wireless information and power transfer has been investigated in [18]. The above works in [16]–[18] consider one eavesdropper in the MIMO relay sub-network. When there are multiple eavesdroppers, the studies in [19]–[21] have developed null-space beamforming and jamming techniques to maximize the secrecy capacity of each user. Gao *et al.* [22] have developed an energy-efficient secure beamforming scheme for self-sustainable relay-aided multicast networks. In addition, the investigations in [23] and [24] have proposed robust beamforming algorithms to improve the secrecy performance in the presence of imperfect eavesdroppers' channel state information (CSI). However, there is only

Manuscript received October 12, 2016; revised February 28, 2017 and April 30, 2017; accepted May 15, 2017. Date of publication May 25, 2017; date of current version September 14, 2017. The work of Y. Cai was supported in part by the National Natural Science Foundation of China under Grant 61471319 and the Fundamental Research Funds for the Central Universities. The work of Q. Shi was supported by the National Natural Science Foundation of China under Grants 61671411 and 61631020. The associate editor coordinating the review of this paper and approving it for publication was M. Abdallah. *(Corresponding author: Qingjiang Shi.)*

Y. Cai is with the Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: ylcai@zju.edu.cn).

Q. Shi is with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (e-mail: qing.j.shi@gmail.com).

B. Champagne is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: benoit.champagne@mcgill.ca).

G. Y. Li is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: liye@ece.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2017.2708110

limited work on the joint optimization of the base station (BS) beamforming matrix and the RS transformation matrix for secrecy sum-rate maximization in secure multiuser MIMO relay systems, which could potentially further improve the performance metrics for secure communications. An optimal power allocation scheme has been proposed in [25] to maximize the secure energy efficiency in relay systems.

In this paper, we study the PLS problem of a general MIMO relay system with multiple legitimate users and malicious eavesdroppers, where information leakage in both hops from the BS to the RS and from the RS to the legitimate users is considered. The eavesdroppers are assumed to cooperate with each other for stealing the legitimate users' information. We consider the joint optimization problem for the BS beamforming matrix, the RS AF transformation matrix, and the covariance matrix of artificial noise (AN), where the aim is to maximize the system worst-case secrecy rate in the presence of the colluding eavesdroppers under power constraints at the BS and the RS, as well as quality of service (QoS) constraints for the legitimate users.[1] Since it is very challenging in practice to obtain accurate eavesdroppers' CSI, we employ a norm-bounded error (NBE) model to describe the CSI errors in this paper. We then consider the robust joint transceiver design based on incorporating the CSI errors into the formulation of the optimization problem.

Since the optimized variables are highly coupled in both the objective function and constraints, it is difficult to globally solve the resulting optimization problem. By applying a series of suitable transformations, we first recast this problem into an equivalent but more tractable form. Then, for the resultant problem we propose a new algorithm based on the penalty concave-convex procedure (penalty-CCCP) to handle the highly coupled terms and jointly optimize the transceiver parameters. With the aid of the penalty-CCCP algorithm, the problematic coupled constraints are incorporated into the objective function as a penalty component. The penalized objective is then optimized via the proposed two-tier iterative algorithm. In the inner loop, we resort to the CCCP method [26], [27] to update the optimization variables; while in the outer loop, we adjust the penalty parameter of the penalized cost function.

Within this framework, the main original contributions are summarized as follows:

1) We formulate a new optimization problem, aiming to maximize the system worst-case secrecy rate with respect to the BS beamforming matrix, the RS AF transformation matrix and the AN covariance matrix, subject to transmit power and legitimate users' QoS constraints. The eavesdroppers' CSI errors are also considered in this problem formulation via the NBE model.
2) To solve this challenging optimization problem, which is characterized by highly coupled design variables, we propose and fully develop a novel penalty-CCCP algorithm. It is further shown that this algorithm ensures convergence to a stationary solution of the optimization problem, thereby leading to a robust joint transceiver design for the MIMO relay system in the presence of colluding eavesdroppers.
3) We provide a detailed computational complexity analysis of the newly proposed algorithm along with exhaustive simulation results that demonstrate its advantages over other recently proposed transceiver designs.

[1] In order to guarantee each user's QoS, the legitimate user's SINR is set to be above a certain threshold.

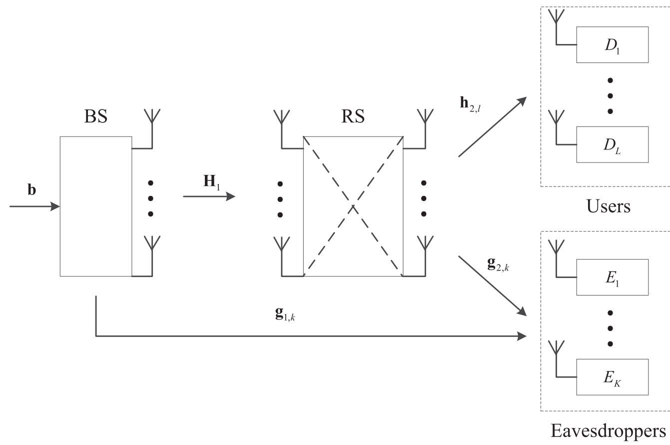

Fig. 1. MIMO relay system in the presence of multiple users and eavesdroppers.

The paper is structured as follows. Section II briefly describes the system model and formulates the optimization problem in mathematical terms. Section III develops the proposed joint transceiver design algorithm and analyzes its complexity and convergence behavior. The simulation results are presented in Section IV and conclusions are drawn in Section V.

Notations: Scalars, vectors and matrices are respectively denoted by lower case, boldface lower case and boldface upper case letters. $\mathbf{I}$ represents an identity matrix and $\mathbf{0}$ denotes an all-zero matrix. For a matrix $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}^*$, $\mathbf{A}^H$ and $\|\mathbf{A}\|$ denote its transpose, conjugate, conjugate transpose and Frobenius norm, respectively. For a square matrix $\mathbf{A}$, $\mathrm{Tr}\{\mathbf{A}\}$ denotes its trace, $\mathbf{A} \succeq \mathbf{0}$ ($\mathbf{A} \preceq \mathbf{0}$) means that $\mathbf{A}$ is positive (negative) semi-definite, and $\mathbf{A} \succ \mathbf{0}$ indicates that $\mathbf{A}$ is positive definite. For a vector $\mathbf{a}$, $\|\mathbf{a}\|$ represents its Euclidean norm. $E\{.\}$ denotes the statistical expectation. $\Re\{.\}$ ($\Im\{.\}$) denotes the real (imaginary) part of a variable. The operator $\mathrm{vec}(\cdot)$ stacks the elements of a matrix in one long column vector. $|\cdot|$ denotes the absolute value of a complex scalar. $\mathbb{C}^{m\times n}$ ($\mathbb{R}^{m\times n}$) denotes the space of $m \times n$ complex (real) matrices.

II. System Model and Problem Formulation

In this section, we first introduce the MIMO relay system model and then mathematically formulate the optimization problem of interest.

A. System Model

Consider a MIMO relay system as depicted in Fig. 1, which consists of one BS, one RS, L legitimate users denoted as node $D_1, \ldots, D_L$, and K eavesdroppers denoted as $E_1, \ldots, E_K$. D_l and $E_k, \forall l \in \mathcal{L} \triangleq \{1, 2, \cdots, L\}, \forall k \in \mathcal{K} \triangleq \{1, 2, \cdots, K\}$ are each equipped with a single antenna while

$$\text{SINR}_l^D = \frac{|\mathbf{h}_{2,l}^H \mathbf{W}\mathbf{H}_1\mathbf{p}_l|^2}{\sum_{l'\neq l}^L |\mathbf{h}_{2,l}^H \mathbf{W}\mathbf{H}_1\mathbf{p}_{l'}|^2 + \sigma_1^2\|\mathbf{h}_{2,l}^H\mathbf{W}\|^2 + \mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l} + \sigma_{2,l}^2} \tag{6}$$

$$\text{SINR}_l^E = \sum_{k=1}^K \left(\frac{|\mathbf{g}_{1,k}^H\mathbf{p}_l|^2}{\sum_{l'\neq l}^L |\mathbf{g}_{1,k}^H\mathbf{p}_{l'}|^2 + \sigma_{1,k}^2} + \frac{|\mathbf{g}_{2,k}^H\mathbf{W}\mathbf{H}_1\mathbf{p}_l|^2}{\sum_{l'\neq l}^L |\mathbf{g}_{2,k}^H\mathbf{W}\mathbf{H}_1\mathbf{p}_{l'}|^2 + \sigma_1^2\|\mathbf{g}_{2,k}^H\mathbf{W}\|^2 + \mathbf{g}_{2,k}^H\mathbf{V}\mathbf{g}_{2,k} + \sigma_{2,k}^2}\right) \tag{9}$$

the BS and the RS are equipped with N_t and N_r antennas, respectively. We assume that there is no direct link between the BS and the legitimate users due to severe attenuation. As in most of the MIMO relaying literature, the transmission is divided into two orthogonal phases. Flat channel fading is assumed throughout, but extension to frequency selective channels is possible.

In the first phase (from BS to RS), the received vector at the RS is given by

$$\mathbf{r}_R = \mathbf{H}_1\mathbf{P}\mathbf{b} + \mathbf{n}_1, \tag{1}$$

where $\mathbf{b} = [b_1, \ldots, b_L]^T$ denotes the $L \times 1$ transmitted symbol vector, whose elements are modeled as independent zero-mean circular complex Gaussian random variables with variance $E\{|b_l|^2\} = 1$, $\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_L] \in \mathbb{C}^{N_t \times L}$ denotes the BS beamforming matrix, $\mathbf{H}_1 \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix between the BS and the RS and $\mathbf{n}_1 \in \mathbb{C}^{N_r \times 1}$ is the additive zero-mean circular complex Gaussian noise vector at the RS with covariance matrix $E\{\mathbf{n}_1\mathbf{n}_1^H\} = \sigma_1^2\mathbf{I}$, where σ_1^2 denotes the noise variance.

In the second phase (from RS to D_l, $\forall l \in \mathcal{L}$), the received vector $\mathbf{r}_R \in \mathbb{C}^{N_r \times 1}$ at the RS is operated by the RS AF transformation matrix $\mathbf{W} \in \mathbb{C}^{N_r \times N_r}$. Moreover, an artificial noise (AN) vector $\mathbf{v} \in \mathbb{C}^{N_r \times 1}$ with zero mean and covariance matrix $\mathbf{V} = E[\mathbf{v}\mathbf{v}^H]$ is sent by the RS to guarantee security. Accordingly, the forwarded signal vector from the RS is given by

$$\mathbf{x}_R = \mathbf{W}(\mathbf{H}_1\mathbf{P}\mathbf{b} + \mathbf{n}_1) + \mathbf{v}. \tag{2}$$

Consequently, the received signal at the lth legitimate user is given by

$$y_{D,l} = \mathbf{h}_{2,l}^H\mathbf{x}_R + n_{2,l} = \mathbf{h}_{2,l}^H\mathbf{W}(\mathbf{H}_1\mathbf{P}\mathbf{b} + \mathbf{n}_1) + \mathbf{h}_{2,l}^H\mathbf{v} + n_{2,l}, \tag{3}$$

where $\mathbf{h}_{2,l} \in \mathbb{C}^{N_r \times 1}$ is the conjugate transpose of the channel vector between the RS and the lth legitimate user, and $n_{2,l}$ denotes the additive noise, modeled as a zero-mean circular complex Gaussian noise, where with variance $E\{|n_{2,l}|^2\} = \sigma_{2,l}^2$.

From (1) and (2), the transmit power of the BS in the first phase and that of the RS in the second phase will be

$$P_B = E\{\|\mathbf{P}\mathbf{b}\|^2\} = \|\mathbf{P}\|^2 \tag{4}$$

and

$$P_R = E\{\|\mathbf{x}_R\|^2\} = \|\mathbf{W}\mathbf{H}_1\mathbf{P}\|^2 + \sigma_1^2\|\mathbf{W}\|^2 + \text{Tr}\{\mathbf{V}\}, \tag{5}$$

respectively. From (3), the received signal-to-interference-plus-noise ratio (SINR) at legitimate user l is given by (6), as shown at the top of this page.

During the transmission, each E_k for $k \in \mathcal{K}$ can overhear signals from both the BS and the RS. Let $\mathbf{g}_{1,k} \in \mathbb{C}^{N_t \times 1}$ and $\mathbf{g}_{2,k} \in \mathbb{C}^{N_r \times 1}$ denote the conjugate transpose BS-E_k and RS-E_k channels, respectively. The signals overheard by E_k from the BS and the RS during the two transmission phases will be

$$y_{1,k} = \mathbf{g}_{1,k}^H\mathbf{P}\mathbf{b} + n_{1,k}, \tag{7}$$

and

$$y_{2,k} = \mathbf{g}_{2,k}^H\mathbf{x}_R + n_{2,k} = \mathbf{g}_{2,k}^H\mathbf{W}(\mathbf{H}_1\mathbf{P}\mathbf{b} + \mathbf{n}_1) + \mathbf{g}_{2,k}^H\mathbf{v} + n_{2,k}, \tag{8}$$

respectively, where $n_{1,k}$ and $n_{2,k}$ denote circular complex Gaussian additive noise terms with zero-mean and variances $\sigma_{1,k}^2$ and $\sigma_{2,k}^2$, respectively. In this work, we consider the worst-case scenario of colluding eavesdroppers, where all the E_k can be treated as a virtual multi-antenna eavesdropper that performs joint processing. From the derivation in Appendix A, and by assuming that the optimal maximum ratio combiner (MRC) scheme is employed, the SINR at the colluding eavesdropper for user l is given by (9), as shown at the top of this page.

B. Problem Formulation

Acquiring eavesdroppers' CSI in a practical wireless network is quite challenging. However, according to [28] we can still estimate the CSI through the local oscillator power inadvertently leaked from the eavesdroppers receiver RF frontend. However, when compared to the legitimate receivers' CSI, the eavesdroppers' CSI is expected to be much less accurate. In this work, we consider the commonly used norm-bounded error (NBE) model [29] to characterize the eavesdroppers' CSI. The true (but unknown) CSI for the BS-E_k and RS-E_k channels can be expressed as follows:

$$\mathbf{g}_{j,k} = \hat{\mathbf{g}}_{j,k} + \triangle\mathbf{g}_{j,k}, \quad k \in \mathcal{K}, \ j \in \{1, 2\} \tag{10}$$

where $\hat{\mathbf{g}}_{j,k}$ denotes the imperfect channel estimate while $\triangle\mathbf{g}_{j,k}$ captures the corresponding uncertainty. Without any statistical knowledge of the latter we assume that it lies in the following bounded region

$$\mathcal{R}_{j,k} = \{\triangle\mathbf{g}_{j,k} : \|\triangle\mathbf{g}_{j,k}\|^2 \leq \varepsilon_{j,k}^2\} \tag{11}$$

where $\varepsilon_{j,k}$ denotes the radius of the uncertainty region. Therefore, the worst-case system secrecy rate [12], [21] can be expressed as follows

$$C(\mathbf{P}, \mathbf{V}, \mathbf{W}) = \frac{1}{2}\sum_{l=1}^L \log\left(1 + \text{SINR}_l^D\right) - \max_{\forall \triangle\mathbf{g}_{j,k} \in \mathcal{R}_{j,k}} \frac{1}{2}\sum_{l=1}^L \log\left(1 + \text{SINR}_l^E\right). \tag{12}$$

$$\max_{\mathbf{P},\mathbf{W},\mathbf{V}\succeq\mathbf{0},\mathbf{X}} \; C(\mathbf{P},\mathbf{V},\mathbf{W},\mathbf{X}) \tag{14a}$$

$$\text{s.t. } \|\mathbf{P}\|^2 \le P_t \tag{14b}$$

$$\|\mathbf{X}\|^2 + \sigma_1^2\|\mathbf{W}\|^2 + \text{Tr}\{\mathbf{V}\} \le P_r, \tag{14c}$$

$$\sum_{l'\neq l}^{L} |\mathbf{h}_{2,l}^H\mathbf{x}_{l'}|^2 + \sigma_1^2\|\mathbf{h}_{2,l}^H\mathbf{W}\|^2 + \mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l} + \sigma_{2,l}^2 \le \frac{|\mathbf{h}_{2,l}^H\mathbf{x}_l|^2}{\gamma_1}, \quad \forall l \in \mathcal{L}, \tag{14d}$$

$$\mathbf{X} = \mathbf{W}\mathbf{H}_1\mathbf{P} \tag{14e}$$

To guarantee the secure transmission, we aim to design the BS beamforming matrix $\mathbf{P}$, the AN covariance matrix $\mathbf{V}$ and the RS AF transformation matrix $\mathbf{W}$ jointly, in order to maximize the worst-case system secrecy rate under the BS and the RS power constraints while keeping the SINRs of the legitimate users above a certain threshold, i.e., QoS constraint. The corresponding optimization problem can be formulated as

$$\max_{\mathbf{P},\mathbf{W},\mathbf{V}\succeq\mathbf{0}} \; C(\mathbf{P},\mathbf{V},\mathbf{W}) \tag{13a}$$

$$\text{s.t. } P_B \le P_t, \quad P_R \le P_r, \tag{13b}$$

$$\text{SINR}_l^D \ge \gamma_l, \quad \forall l \in \mathcal{L}. \tag{13c}$$

As we can see from the inspection of (4)-(6), (9) and (12), the design variables are highly coupled in the objective function and constraints and accordingly, it appears impossible to globally solve problem (13). In the following section, we propose an efficient joint design algorithm that can find a local stationary solution by applying proper convex optimization techniques.

III. Proposed Joint Transceiver Design

In this section, we propose to employ an extended penalty method, referred to as the penalty-CCCP algorithm, which is detailed in Appendix B, to address problem (13). In Subsection A, by applying the penalty technique, we present an alternative formulation to (13) where a penalty term is incorporated into the objective in order to handle a complex constraint. In Subsections B and C, we transform the penalized problem into a more tractable yet equivalent problem and develop an efficient CCCP algorithm to solve it. In particular, in order to approximate the equivalent problem as a convex problem, a locally tight lower bound for the objective function is derived; furthermore, linearization is used to approximate the nonconvex constraints. In Subsection D, we analyze the computational complexity of the proposed transceiver design algorithm.

A. The Penalized Problem

The structure of the relay system generates a highly coupled term, i.e., $\mathbf{W}\mathbf{H}_1\mathbf{P}$, in the objective function and constraints appearing in (13). In order to handle this term in the design, we introduce a new auxiliary variable matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_L] \in \mathbb{C}^{N_r \times L}$ along with equality constraint $\mathbf{X} = \mathbf{W}\mathbf{H}_1\mathbf{P}$. Therefore, problem (13) can be equivalently formulated as (14), as shown at the top of this page, where

$$C(\mathbf{P},\mathbf{V},\mathbf{W},\mathbf{X}) = \frac{1}{2}\sum_{l=1}^{L}\log\left(1+\eta_l^D\right) - \max_{\forall\triangle\mathbf{g}_{j,k}\in\mathcal{R}_{j,k}} \frac{1}{2}\sum_{l=1}^{L}\log\left(1+\eta_l^E\right), \tag{15}$$

$$\eta_l^D = \frac{|\mathbf{h}_{2,l}^H\mathbf{x}_l|^2}{\sum_{l'\neq l}^{L}|\mathbf{h}_{2,l}^H\mathbf{x}_{l'}|^2 + \sigma_1^2\|\mathbf{h}_{2,l}^H\mathbf{W}\|^2 + \mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l} + \sigma_{2,l}^2} \tag{16}$$

and

$$\eta_l^E = \sum_{k=1}^{K}\Bigg(\frac{|\mathbf{g}_{1,k}^H\mathbf{p}_l|^2}{\sum_{l'\neq l}^{L}|\mathbf{g}_{1,k}^H\mathbf{p}_{l'}|^2 + \sigma_{1,k}^2} + \frac{|\mathbf{g}_{2,k}^H\mathbf{x}_l|^2}{\sum_{l'\neq l}^{L}|\mathbf{g}_{2,k}^H\mathbf{x}_{l'}|^2 + \sigma_1^2\|\mathbf{g}_{2,k}^H\mathbf{W}\|^2 + \mathbf{g}_{2,k}^H\mathbf{V}\mathbf{g}_{2,k} + \sigma_{2,k}^2}\Bigg). \tag{17}$$

Note that here for clarity, we have rewritten the SINR expressions in (16) and (17). According to [30], in order to deal with the equality constraint (14e), we can equivalently convert it into a combination of a linear matrix inequality (LMI) and a difference of convex (DC) function equality:

$$\begin{bmatrix} \mathbf{G}_{11} & \mathbf{X} & \mathbf{W}\mathbf{H}_1 \\ \mathbf{X}^H & \mathbf{G}_{22} & \mathbf{P}^H \\ \mathbf{H}_1^H\mathbf{W}^H & \mathbf{P} & \mathbf{I} \end{bmatrix} \succeq \mathbf{0}, \tag{18}$$

$$\text{Tr}\{\mathbf{G}_{11}\} - \text{Tr}\{\mathbf{W}\mathbf{H}_1\mathbf{H}_1^{\text{H}}\mathbf{W}^{\text{H}}\} = 0 \tag{19}$$

where $\mathbf{G}_{11}$ and $\mathbf{G}_{22}$ are introduced as auxiliary variables. Moreover, due to (18), the left-hand-side of (19) must be nonnegative. Making use of (18) and (19), problem (14) can be recast as

$$\max_{\mathbf{P},\mathbf{W},\mathbf{V}\succeq\mathbf{0},\mathbf{X}} \; C(\mathbf{P},\mathbf{V},\mathbf{W},\mathbf{X}) \tag{20a}$$

$$\text{s.t. } (14b) - (14d), (18), \tag{20b}$$

$$\sqrt{\text{Tr}\{\mathbf{G}_{11}\} - \text{Tr}\{\mathbf{W}\mathbf{H}_1\mathbf{H}_1^{\text{H}}\mathbf{W}^{\text{H}}\}} = 0, \tag{20c}$$

where we have replaced (19) with the last equality constraint in order to use our optimization framework, i.e. the penalty-CCCP, proposed in Appendix B.

Note that the equality constraint (20c) complicates the problem significantly, for otherwise we could directly use the standard CCCP method to address the problem. By applying the penalty-CCCP method (see the details in Appendix B),

$$\max_{\substack{\{\alpha_{k,l},\beta_{k,l},\\ \rho_{k,l},\phi_{k,l}\},\\ \mathbf{P},\mathbf{W},\mathbf{V}\succeq\mathbf{0},\mathbf{X}}} \max_{\{u_l,\upsilon_l>0\}} \frac{1}{2}\sum_{l=1}^{L}\Big(\log\ (\upsilon_l)-\upsilon_l e_l(u_l,\mathbf{X},\mathbf{W},\mathbf{V})\Big)-\frac{1}{2}\sum_{l=1}^{L}\log\Big(1+\sum_{k=1}^{K}\alpha_{k,l}+\sum_{k=1}^{K}\rho_{k,l}\Big)-\tau\theta \tag{22a}$$

$$\text{s.t. } (14b)-(14d), (18), (21c), \tag{22b}$$

$$|\mathbf{g}_{1,k}^H\mathbf{p}_l|^2-\frac{(\sigma_{1,k}^2+\sum_{l'\neq l}^{L}|\mathbf{g}_{1,k}^H\mathbf{p}_{l'}|^2)}{\phi_{k,l}}\leq 0, \tag{22c}$$

$$|\mathbf{g}_{2,k}^H\mathbf{x}_l|^2-\frac{(\sigma_1^2\|\mathbf{g}_{2,k}^H\mathbf{W}\|^2+\sum_{l'\neq l}^{L}|\mathbf{g}_{2,k}^H\mathbf{x}_{l'}|^2+\mathbf{g}_{2,k}^H\mathbf{V}\mathbf{g}_{2,k}+\sigma_{2,k}^2)}{\beta_{k,l}}\leq 0, \tag{22d}$$

$$\rho_{k,l}\beta_{k,l}\geq 1, \tag{22e}$$

$$\alpha_{k,l}\phi_{k,l}\geq 1, \tag{22f}$$

$$\forall\triangle\mathbf{g}_{j,k}\in\mathcal{R}_{j,k},\ j\in\{1,2\},\quad \forall l\in\mathcal{L},\ \forall k\in\mathcal{K} \tag{22g}$$

we iteratively solve the following problem:

$$\max_{\mathbf{P},\mathbf{W},\mathbf{V}\succeq\mathbf{0},\mathbf{X},\theta\geq 0} C(\mathbf{P},\mathbf{V},\mathbf{W},\mathbf{X})-\tau\theta \tag{21a}$$

$$\text{s.t. } (14b)-(14d), (18), \tag{21b}$$

$$\text{Tr}\{\mathbf{G}_{11}\}-\text{Tr}\{\mathbf{W}\mathbf{H}_1\mathbf{H}_1^{\rm H}\mathbf{W}^{\rm H}\}\leq\theta, \tag{21c}$$

where $\theta\geq 0$ denotes an auxiliary variable introduced to move the penalty term $\text{Tr}\{\mathbf{G}_{11}\}-\text{Tr}\{\mathbf{W}\mathbf{H}_1\mathbf{H}_1^{\rm H}\mathbf{W}^{\rm H}\}$ into the objective via the inequality (21c) and $\tau>0$ denotes the corresponding penalty parameter that prescribes a high cost for the violation of the constraint. In particular, in the limit $\tau\to\infty$, problems (14) and (21) are equivalent [31].

B. Problem Transformation

We first transform optimization problem in (21) into an equivalent yet more tractable form. Specifically, by introducing a number of auxiliary variables, i.e. $\{u_l,\upsilon_l,\alpha_{k,l},\beta_{k,l},\rho_{k,l},\phi_{k,l}\}$, (21) can be formulated as the following equivalent problem (22), as shown at the top of this page, in the sense that both problems share the same global optimal solutions for $\mathbf{P},\mathbf{W}$ and $\mathbf{V}$ under the given constraints, where

$$e_l(u_l,\mathbf{X},\mathbf{W},\mathbf{V})\triangleq|1-u_l^*\mathbf{h}_{2,l}^H\mathbf{x}_l|^2+\sum_{l'\neq l}|u_l^*\mathbf{h}_{2,l}^H\mathbf{x}_{l'}|^2+|u_l|^2(\sigma_1^2\|\mathbf{h}_{2,l}^H\mathbf{W}\|^2+\sigma_2^2+\mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l}) \tag{23}$$

denotes the mean squared error (MSE) of user l, and u_l denotes the receiver weight of user l. Note that u_l only appears in the objective function (22a); hence, by fixing the remaining variables and taking the gradient with respect to u_l^*, we obtain the minimum mean squared error (MMSE) receiver:

$$\tilde{u}_l=\frac{\mathbf{h}_{2,l}^H\mathbf{x}_l}{\sum_{l'=1}^{L}|\mathbf{h}_{2,l}^H\mathbf{x}_{l'}|^2+\sigma_1^2\|\mathbf{h}_{2,l}^H\mathbf{W}\|^2+\mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l}+\sigma_2^2}. \tag{24}$$

Using (24), the value of the MSE in (23) is now given by

$$\tilde{e}_l=1-\frac{|\mathbf{h}_{2,l}^H\mathbf{x}_l|^2}{\sum_{l'=1}^{L}|\mathbf{h}_{2,l}^H\mathbf{x}_{l'}|^2+\sigma_1^2\|\mathbf{h}_{2,l}^H\mathbf{W}\|^2+\mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l}+\sigma_2^2}. \tag{25}$$

The detailed proof of the equivalence between problems (21) and (22) is presented in Appendix C.

C. Proposed CCCP Algorithm for Solving Problem (22)

In the following, we develop an efficient CCCP algorithm for the optimization of $\mathbf{P}$, $\mathbf{W}$ and $\mathbf{V}$ in problem (22), which by nature is non-convex. In order to approximate this problem as a convex one, we first find a locally tight lower bound of the objective and then linearize the nonconvex constraints with the aid of CCCP concepts, so that the nonconvex problem can be approximated as a convex one.

1) Lower Bound for the Objective Function: Let us treat the variables $\{u_l,\upsilon_l:\forall l\in\mathcal{L}\}$ as the intermediate variables, then for any $\mathbf{X},\mathbf{W},\mathbf{V},\bar{\mathbf{X}},\bar{\mathbf{W}},\bar{\mathbf{V}}$, we can obtain the following inequality for the objective function (i.e., the inner maximization) of problem (22):

$$\begin{aligned}\max_{\{u_l,\upsilon_l\}}\ &\frac{1}{2}\sum_{l=1}^{L}\Big(\log\ (\upsilon_l)-\upsilon_l e_l(u_l,\mathbf{X},\mathbf{W},\mathbf{V})\Big)\\ &-\frac{1}{2}\sum_{l=1}^{L}\log\Big(1+\sum_{k=1}^{K}\alpha_{k,l}+\sum_{k=1}^{K}\rho_{k,l}\Big)-\tau\theta\\ \geq&\frac{1}{2}\sum_{l=1}^{L}\Big(\log\ (\bar{\upsilon}_l)-\bar{\upsilon}_l e_l(\bar{u}_l,\mathbf{X},\mathbf{W},\mathbf{V})\Big)\\ &-\frac{1}{2}\sum_{l=1}^{L}\log\Big(1+\sum_{k=1}^{K}\alpha_{k,l}+\sum_{k=1}^{K}\rho_{k,l}\Big)-\tau\theta,\end{aligned} \tag{26}$$

where $\bar{\upsilon}_l$ and $\bar{u}_l$ are given by

$$\bar{\upsilon}_l=\frac{1}{\bar{e}_l}, \tag{27}$$

$$\bar{e}_l=1-\frac{|\mathbf{h}_{2,l}^H\bar{\mathbf{x}}_l|^2}{\sum_{l'=1}^{L}|\mathbf{h}_{2,l}^H\bar{\mathbf{x}}_{l'}|^2+\sigma_1^2\|\mathbf{h}_{2,l}^H\bar{\mathbf{W}}\|^2+\mathbf{h}_{2,l}^H\bar{\mathbf{V}}\mathbf{h}_{2,l}+\sigma_2^2}, \tag{28}$$

$$\bar{u}_l=\frac{\mathbf{h}_{2,l}^H\bar{\mathbf{x}}_l}{\sum_{l'=1}^{L}|\mathbf{h}_{2,l}^H\bar{\mathbf{x}}_{l'}|^2+\sigma_1^2\|\mathbf{h}_{2,l}^H\bar{\mathbf{W}}\|^2+\mathbf{h}_{2,l}^H\bar{\mathbf{V}}\mathbf{h}_{2,l}+\sigma_2^2}. \tag{29}$$

Since the logarithm function is concave, we can see that the right-hand-side of inequality (26) can be rewritten as

$$\hat{f}_2(\mathbf{s}_1^i, \mathbf{s}_1) = f_2(\mathbf{s}_1^i) + 2\Re\{\nabla f_2^H(\mathbf{s}_1^i)(\mathbf{s}_1 - \mathbf{s}_1^i)\} = f_2(\mathbf{s}_1^i) - \frac{1}{2}\sum_{l=1}^{L} \frac{\sum_{k=1}^{K}\Big((\alpha_{k,l} - \alpha_{k,l}^i) + (\rho_{k,l} - \rho_{k,l}^i)\Big)}{1 + \sum_{k=1}^{K}(\alpha_{k,l}^i + \rho_{k,l}^i)} \tag{33}$$

$$\left\| \left[[\mathbf{h}_{2,l}^H\mathbf{x}_1, \ldots, \mathbf{h}_{2,l}^H\mathbf{x}_{(l-1)}, \mathbf{h}_{2,l}^H\mathbf{x}_{(l+1)}, \ldots, \mathbf{h}_{2,l}^H\mathbf{x}_L], \sigma_1\mathbf{h}_{2,l}^H\mathbf{W}, \pi_{1,l}, \sigma_{2,l} \right] \right\| \leq \frac{\mathbf{h}_{2,l}^H\mathbf{x}_l}{\sqrt{\gamma_1}},\ \Im\{\mathbf{h}_{2,l}^H\mathbf{x}_l\} = 0 \tag{37}$$

the following difference of concave functions with respect to $\mathbf{X}, \mathbf{W}, \mathbf{V}, \theta$ and $\mathbf{s}_1$:

$$f_1(\mathbf{X}, \mathbf{W}, \mathbf{V}, \theta) - f_2(\mathbf{s}_1) \tag{30}$$

where

$$f_1(\mathbf{X}, \mathbf{W}, \mathbf{V}, \theta) = \frac{1}{2}\sum_{l=1}^{L}\Big(\log(\bar{v}_l) - \bar{v}_l e_l(\bar{u}_l, \mathbf{X}, \mathbf{W}, \mathbf{V})\Big) - \tau\theta, \tag{31}$$

$$f_2(\mathbf{s}_1) = \frac{1}{2}\sum_{l=1}^{L}\log\left(1 + \sum_{k=1}^{K}\alpha_{k,l} + \sum_{k=1}^{K}\rho_{k,l}\right) \tag{32}$$

and $\mathbf{s}_1 = [\alpha_{1,1}, \rho_{1,1}, \ldots, \alpha_{K,L}, \rho_{K,L}]^T$. Based on the CCCP concept [26], [27], [34], [35], we approximate the concave function $f_2(\mathbf{s}_1)$ in the ith iteration of the CCCP algorithm by its first order Taylor expansion around the current point $\mathbf{s}_1^i = [\alpha_{1,1}^i, \rho_{1,1}^i, \ldots, \alpha_{K,L}^i, \rho_{K,L}^i]^T$, denoted as (33), as shown at the top of this page. Therefore, using the above results we can obtain a locally tight lower bound for the objective function of problem (22) as follows

$$f_1(\mathbf{X}, \mathbf{W}, \mathbf{V}, \theta) - \hat{f}_2(\mathbf{s}_1^i, \mathbf{s}_1). \tag{34}$$

2) Linearizing the Nonconvex Constraints: Note that (14b), (14c), (22e) and (22f) can be equivalently converted to SOC constraints, where the details are presented in Appendix D. Hence, we focus on solving constraint (14d). By introducing auxiliary variables $\pi_{1,l}$, $\forall l \in \mathcal{L}$, which are subject to $\mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l} \leq \pi_{1,l}^2$, (14d) can be equivalently formulated as follows:

$$\sum_{l' \neq l}^{L}|\mathbf{h}_{2,l}^H\mathbf{x}_{l'}|^2 + \sigma_1^2\|\mathbf{h}_{2,l}^H\mathbf{W}\|^2 + \pi_{1,l}^2 + \sigma_{2,l}^2 \leq \frac{|\mathbf{h}_{2,l}^H\mathbf{x}_l|^2}{\gamma_1} \tag{35}$$

and

$$\mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l} - \pi_{1,l}^2 \leq 0. \tag{36}$$

For (35), since this will not affect the optimization result, we can rotate the phase of $\mathbf{h}_{2,l}^H\mathbf{x}_l$ and then take the square root of both sides of (35) to generate the SOC constraint shown in (37), as shown at the top of this page. Note that (36) is a DC function; similarly by linearizing the subtrahend $\pi_{1,l}^2$ in the ith iteration around the current point $\pi_{1,l}^i$, it yields the following approximated linear constraint:

$$\mathbf{h}_{2,l}^H\mathbf{V}\mathbf{h}_{2,l} - \pi_{1,l}^i(2\pi_{1,l} - \pi_{1,l}^i) \leq 0. \tag{38}$$

Subsequently we deal with constraints (22c) and (22d). It can be seen that (22c) is also a DC function, denoted as

$$f_3(\mathbf{p}_l) - f_4(\mathbf{s}_2) \leq 0, \tag{39}$$

where $f_3(\mathbf{p}_l) = |\mathbf{g}_{1,k}^H\mathbf{p}_l|^2$, $f_4(\mathbf{s}_2) = \frac{\sigma_{1,k}^2 + \sum_{l' \neq l}^{L}|\mathbf{g}_{1,k}^H\mathbf{p}_{l'}|^2}{\phi_{k,l}}$ and $\mathbf{s}_2 = [\mathbf{p}_1^T, \ldots, \mathbf{p}_{(l-1)}^T, \mathbf{p}_{(l+1)}^T, \ldots, \mathbf{p}_L^T, \phi_{k,l}]^T$. By linearizing $f_4(\mathbf{s}_2)$ in the ith iteration around the current point $\mathbf{s}_2^i = [\mathbf{p}_1^{iT}, \ldots, \mathbf{p}_{(l-1)}^{iT}, \mathbf{p}_{(l+1)}^{iT}, \ldots, \mathbf{p}_L^{iT}, \phi_{k,l}^i]^T$, (22c) can be approximated as the following convex constraint:

$$f_3(\mathbf{p}_l) - \hat{f}_4(\mathbf{s}_2^i, \mathbf{s}_2) \leq 0, \tag{40}$$

where

$$\begin{aligned}\hat{f}_4(\mathbf{s}_2^i, \mathbf{s}_2) = {} & 2\Re\Bigg\{\sum_{l' \neq l}^{L}\frac{\mathbf{p}_{l'}^{iH}\mathbf{g}_{1,k}\mathbf{g}_{1,k}^H\mathbf{p}_{l'}}{\phi_{k,l}^i}\Bigg\} - \frac{\sigma_{1,k}^2\phi_{k,l}}{\phi_{k,l}^{i2}} \\ & - \sum_{l' \neq l}^{L}\frac{\mathbf{p}_{l'}^{iH}\mathbf{g}_{1,k}\mathbf{g}_{1,k}^H\mathbf{p}_{l'}^i\phi_{k,l}}{\phi_{k,l}^{i2}} + \frac{2\sigma_{1,k}^2}{\phi_{k,l}^i}.\end{aligned} \tag{41}$$

We note that (40) corresponds to an infinite number of constraints since $\mathbf{g}_{1,k} = \hat{\mathbf{g}}_{1,k} + \triangle\mathbf{g}_{1,k}$ and the inequality must hold $\forall\triangle\mathbf{g}_{1,k} \in \mathcal{R}_{1,k}$. By applying the S-procedure [36], we can equivalently convert (40) into finite convex constraints, as equivalently expressed by the following LMI constraints:

$$\begin{bmatrix} -\mathbf{Y}_{k,l} + \lambda_k\mathbf{I} & -\mathbf{Y}_{k,l}\hat{\mathbf{g}}_{1,k} \\ -\hat{\mathbf{g}}_{1,k}^H\mathbf{Y}_{k,l} & \dfrac{2\sigma_{1,k}^2}{\phi_{k,l}^i} - \dfrac{\sigma_{1,k}^2\phi_{k,l}}{\phi_{k,l}^{i2}} - \hat{\mathbf{g}}_{1,k}^H\mathbf{Y}_{k,l}\hat{\mathbf{g}}_{1,k} - \lambda_k\varepsilon_{1,k}^2 \end{bmatrix} \succeq \mathbf{0}, \tag{42}$$

and

$$\begin{bmatrix} \mathbf{D}_l & \mathbf{p}_l \\ \mathbf{p}_l^H & 1 \end{bmatrix} \succeq \mathbf{0} \tag{43}$$

where $\mathbf{Y}_{k,l} \triangleq \mathbf{D}_l - \sum_{l' \neq l}^{L}\Big(\frac{\mathbf{p}_l\mathbf{p}_l^{iH} + \mathbf{p}_l^i\mathbf{p}_l^H}{\phi_{k,l}^i} - \frac{\mathbf{p}_l^i\mathbf{p}_l^{iH}\phi_{k,l}}{\phi_{k,l}^{2\ i}}\Big)$, $\lambda_k \geq 0$ is a slack variable and $\mathbf{D}_l$ denotes an auxiliary matrix variable. The detailed derivation of (42) and (43) is presented in Appendix E.

In order to transform (22d) into a DC constraint, we introduce two auxiliary variables $\pi_{2,k}$ and $\pi_{3,k}$, which are subject to $\mathbf{g}_{2,k}^H\mathbf{V}\mathbf{g}_{2,k} \geq \pi_{3,k}$ and $\pi_{3,k} \geq \pi_{2,k}^2$, $\forall k \in \mathcal{K}$. Thus (22d) can be formulated as the following equivalent constraints:

$$f_5(\mathbf{s}_3) - f_6(\mathbf{s}_3) \leq 0, \tag{44}$$

$$\pi_{3,k} \geq \pi_{2,k}^2 \tag{45}$$

$$\mathbf{g}_{2,k}^H\mathbf{V}\mathbf{g}_{2,k} \geq \pi_{3,k} \tag{46}$$

where $f_5(\mathbf{s}_3) = |\mathbf{g}_{2,k}^H\mathbf{x}_l|^2$, $f_6(\mathbf{s}_3) = \frac{\sigma_1^2\|\mathbf{g}_{2,k}^H\mathbf{W}\|^2 + \sum_{l' \neq l}^{L}|\mathbf{g}_{2,k}^H\mathbf{x}_{l'}|^2 + \pi_{2,k}^2 + \sigma_{2,k}^2}{\beta_{k,l}}$ and $\mathbf{s}_3 = [\mathbf{g}_{2,k}^H\mathbf{W}, \mathbf{x}_1^T, \ldots, \mathbf{x}_{(l-1)}^T, \mathbf{x}_{(l+1)}^T, \ldots, \mathbf{x}_L^T, \beta_{k,l}, \pi_{2,k}]^T$. It is readily seen that (45) can be converted into the following SOCP constraint:

$$\left\|\left[\pi_{2,k}, \frac{\pi_3 - 1}{2}\right]\right\| \leq \frac{\pi_3 + 1}{2}. \tag{47}$$

Note that since we have $\mathbf{g}_{2,k} = \hat{\mathbf{g}}_{2,k} + \triangle\mathbf{g}_{2,k}$ and the inequality must hold $\forall \triangle\mathbf{g}_{2,k} \in \mathcal{R}_{2,k}$, (46) corresponds to an infinite number of linear constraints. Based on the S-procedure, we can equivalently convert (46) into the following finite LMI constraint:

$$\begin{bmatrix} \mathbf{V} + \omega_k \mathbf{I} & \mathbf{V}\hat{\mathbf{g}}_{2,k} \\ \hat{\mathbf{g}}_{2,k}^H \mathbf{V} & \pi_{3,k} + \hat{\mathbf{g}}_{2,k}^H \mathbf{V}\hat{\mathbf{g}}_{2,k} - \omega_k \varepsilon_{2,k}^2 \end{bmatrix} \succeq \mathbf{0}, \tag{48}$$

where $\omega_k \geq 0$ denotes a slack variable.

In addition, we have that (44) is a DC constraint. Similarly, by linearizing $f_6(\mathbf{s}_3)$ in the ith iteration around the current point $\mathbf{s}_3^i = [\mathbf{g}_{2,k}^H \mathbf{W}^i, \mathbf{x}_1^{iT}, \ldots, \mathbf{x}_{(l-1)}^{iT}, \mathbf{x}_{(l+1)}^{iT}, \ldots, \mathbf{x}_L^{iT}, \beta_{k,l}^i, \pi_{2,k}^i]^T$, (44) can be approximated as the following convex constraint:

$$f_5(\mathbf{s}_3) - \hat{f}_6(\mathbf{s}_3^i, \mathbf{s}_3) \leq 0, \tag{49}$$

where $\hat{f}_6(\mathbf{s}_3^i, \mathbf{s}_3)$ is shown in (50), as shown at the bottom of this page. By following the same approach as in Appendix E, we can equivalently convert (49) into the finite LMI constraint (51), as shown at the bottom of this page, and

$$\begin{bmatrix} \mathbf{F}_l & \mathbf{x}_l \\ \mathbf{x}_l^H & 1 \end{bmatrix} \succeq \mathbf{0} \tag{52}$$

where $\mathbf{U}_{k,l} \triangleq \mathbf{F}_l - \frac{\sigma_1^2 \mathbf{W}^i \mathbf{W}^H + \sigma_1^2 \mathbf{W}\mathbf{W}^{iH}}{\beta_{k,l}^i} + \frac{\sigma_1^2 \mathbf{W}^i \mathbf{W}^{iH} \beta_{k,l}}{\beta_{k,l}^{i2}} - \sum_{l' \neq l}^{L} \left(\frac{\mathbf{x}_{l'} \mathbf{x}_{l'}^{iH} + \mathbf{x}_{l'}^i \mathbf{x}_{l'}^H}{\beta_{k,l}^i} - \frac{\mathbf{x}_{l'}^i \mathbf{x}_{l'}^{iH} \beta_{k,l}}{\beta_{k,l}^{i2}} \right)$, and $\mu_k \geq 0$ denotes a slack variable.

Finally let us tackle constraint (21c). Due to the DC constraint, by linearizing the subtrahend $\mathrm{Tr}\{\mathbf{W}\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}}\mathbf{W}^{\mathrm{H}}\}$ around the current point $\mathbf{W}^i$, (21c) can be approximated as the following linear constraint:

$$\begin{aligned} \mathrm{Tr}\{\mathbf{G}_{11}\} \; & -\big(\mathrm{Tr}\{\mathbf{W}^{\mathrm{i}}\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}}\mathbf{W}^{\mathrm{iH}}\} \\ & + 2\Re\{\mathrm{Tr}\{\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}}\mathbf{W}^{\mathrm{iH}}(\mathbf{W} - \mathbf{W}^{\mathrm{i}})\}\}\big) \leq \theta. \end{aligned} \tag{53}$$

Finally, by following the CCCP concept, problem (22) can be reformulated as a convex optimization problem in the ith iteration shown in (54). This problem can be efficiently solved by the convex programming toolbox CVX [37].

The proposed CCCP algorithm for solving problem (22) is summarized in Table I. Moreover, the overall penalty-CCCP algorithm is summarized in Table II, where $f(\tau^{(n)})$ denotes the value of objective function (34) in the nth iteration of the outer loop. Since we propose applying the CCCP algorithm to address the penalized problem, the proposed transceiver design scheme is referred to as penalty-CCCP algorithm. Based on the proof in Appendix B, we can see that the proposed joint transceiver design algorithm converges to a stationary solution of the original problem.[2]

TABLE I

PROPOSED CCCP ALGORITHM FOR PROBLEM (22)

1. Define the tolerance of accuracy δ_2 and the maximum number of iteration N_{max}. Initialize the algorithm with a feasible point ($\mathbf{P}$, $\mathbf{W}$, $\mathbf{X}$, $\mathbf{V}$) and set $\bar{\mathbf{W}} = \mathbf{W}$, $\bar{\mathbf{X}} = \mathbf{X}$, $\bar{\mathbf{V}} = \mathbf{V}$. Set the iteration number $i = 0$.
2. **Repeat**
   - Update $\{\bar{u}_l\}$ based on (29), $\forall l \in \mathcal{L}$.
   - Update $\{\bar{v}_l\}$ based on (27), $\forall l \in \mathcal{L}$.
   - Update $\{\mathbf{P}, \mathbf{W}, \mathbf{V}, \mathbf{X}, \{\alpha_{k,l}, \beta_{k,l}, \rho_{k,l}, \phi_{k,l}\}, \theta\}$ based on solving problem (54), $\forall l \in \mathcal{L}$, $\forall k \in \mathcal{K}$.
   - Update the iteration number : $i = i + 1$.
3. **Until** the difference successive values of the objective function is less than δ_2 or the maximum number of iterations is reached, i.e., $i > N_{\max}$.

TABLE II

PROPOSED PENALTY-CCCP ALGORITHM FOR TRANSCEIVER DESIGN

1. Define the tolerance of accuracy δ_1. Initialize the algorithm with a feasible point. Set the iteration number $n = 0$. Set $c > 1$ and $\tau^{(0)} > 0$.
2. **Repeat**
   - Apply the proposed CCCP algorithm in Table I to update the optimization variables iteratively.
   - $\tau^{(n+1)} = c\tau^{(n)}$.
   - Update the iteration number : $n = n + 1$.
3. **Until** $\frac{|f(\tau^{(n)}) - f(\tau^{(n-1)})|}{|f(\tau^{(n-1)})|} \leq \delta_1$.

D. Computational Complexity

In this part, we present the computational complexity analysis for the proposed penalty-CCCP transceiver design

[2]In the absence of better alternative, it is readily seen that this proposed algorithm is currently the best choice for solving this optimization problem. Due to the NP-hard nature of the problem, based on the current optimization techniques, it does not seem possible to provide a quantitative analysis of the performance gap between the optimal solution and that obtained with the proposed algorithm. At the present time, providing ¡®stationary point convergence¡¯ is the best we can do in terms of convergence analysis for this kind of constrained optimization problem. The characterization of the performance gap between the optimal and iterative solutions remains an open problem for the future research.

$$\begin{aligned} \hat{f}_6(\mathbf{s}_3^i, \mathbf{s}_3) = 2\Re\left\{ \frac{\sigma_1^2 \mathbf{g}_{2,k}^H \mathbf{W}^i \mathbf{W}^H \mathbf{g}_{2,k}}{\beta_{k,l}^i} \right\} & - \frac{\left(\sigma_1^2 \|\mathbf{g}_{2,k}^H \mathbf{W}^i\|^2 + \sum_{l' \neq l}^{L} |\mathbf{g}_{2,k}^H \mathbf{x}_{l'}^i|^2 + \pi_{2,k}^{i2} + \sigma_{2,k}^2\right)\beta_{k,l}}{\beta_{k,l}^{i2}} \\ + 2\Re\left\{ \sum_{l' \neq l}^{L} \frac{\mathbf{x}_{l'}^{iH} \mathbf{g}_{2,k} \mathbf{g}_{2,k}^H \mathbf{x}_{l'}}{\beta_{k,l}^i} \right\} & + \frac{2\sigma_{2,k}^2 + 2\pi_{2,k}\pi_{2,k}^i}{\beta_{k,l}^i} \end{aligned} \tag{50}$$

$$\begin{bmatrix} -\mathbf{U}_{k,l} + \mu_k \mathbf{I} & -\mathbf{U}_{k,l}\hat{\mathbf{g}}_{2,k} \\ -\hat{\mathbf{g}}_{2,k}^H \mathbf{U}_{k,l} & \frac{2\sigma_{2,k}^2 + 2\pi_{2,k}\pi_{2,k}^i}{\beta_{k,l}^i} - \frac{\left(\pi_{2,k}^{i2} + \sigma_{2,k}^2\right)\beta_{k,l}}{\beta_{k,l}^{i2}} - \hat{\mathbf{g}}_{2,k}^H \mathbf{U}_{k,l} \hat{\mathbf{g}}_{2,k} - \mu_k \varepsilon_{2,k}^2 \end{bmatrix} \succeq \mathbf{0} \tag{51}$$

algorithm. To this end, we apply the same basic elements of complexity analysis as in [38].

Since the update of $\bar{u}_l$ and $\bar{v}_l$ only requires solving closed form solutions, the complexity of the proposed algorithm is dominated by solving problem (54), as shown at the bottom of the next page, $I_1 I_2$ times, where I_1 and I_2 denote the number of iterations for the outer and inner loop, respectively. Problem (54) mainly involves $2L + 3K + 1$ SOC constraints, including $3K$ SOC of dimension three, one SOC of dimension $N_r^2 + N_r L + 2$, L SOC of dimension $L + N_r + 1$, and L SOC of dimension $L + 2$.[3] In addition, problem (54) involves $KL + L$ LMI constraints of size $N_t + 1$, $KL + K + L$ LMI constraints of size $N_r + 1$, and one LMI constraint of size $N_r + L + N_t$. The number of variables is $m = \mathcal{O}(N_t^2 L + N_r^2 L)$. Thus, the complexity of solving problem (54) is roughly given by $I_1 I_2 \mathcal{O}(m\sqrt{(KL+K+L)(N_r+1)+(KL+L)(N_t+1)} \times ((N_r + L + N_t)^3 + m(N_r + L + N_t)^2 + (KL+K+L)(N_r+1)^3 + m(KL+K+L)(N_r+1)^2 + (N_r^2 + N_r L + 2)^2 + (L + N_r + 1)^2 L + m^2))$.

Similarly, the computational complexity of the conventional alternating optimization (AO) CCCP algorithm[4] is $I_1(I_2 \mathcal{O}(m_1\sqrt{KL(N_r+1)+KL(N_t+1)}((N_r+1)^3(KL+2K)+m_1(N_r+1)^3(KL+2K)+(N_t+1)^3(KL+L)+m_1(N_t+1)^2(KL+L)+(N_t L+1)^2+(N_r L+2)^2+(L+N_r+1)^2 L+(L+2)^2 L))+I_3 \mathcal{O}(m_2\sqrt{KL(N_r+1)+KL(N_t+1)}(KL(N_r+1)^3+m_2(N_r+1)^2 KL+(N_t+1)^3 KL+m_2(N_t+1)^2 KL+(N_r^2+1)^2+(L+N_r+1)^2 L)))$, where $m_1 = m_2 = \mathcal{O}(N_t^3)$ and I_1, I_2 and I_3 denote the number of iterations. In particular, when $N_t = N_r = L = K \to \infty$ and $I_1 = I_2 = I_3 = I$, the proposed joint design algorithm has the same order complexity compared with the conventional AO-CCCP algorithm, which is given by $\mathcal{O}(I^2 N_t^{11.5})$. However, the proposed algorithm can be guaranteed to converge to a stationary solution of the original problem in theory.

IV. Simulation Results

In this section, we evaluate the performance of the proposed joint transceiver design algorithm based on penalty-CCCP by means of computer simulations. We assume that the channel coefficients are quasi-static flat-fading. The eavesdroppers' channel error bounds are assumed to be identical for simplicity, i.e., $\varepsilon_{j,k} = \varepsilon, \forall j \in \{1, 2\}, \forall k \in \mathcal{K}$. The number of BS antennas and legitimate users are set to $N_t = L = 3$. The other system parameters are chosen as follows: $P_t = 15$dB, $P_r = 20$dB, $\sigma_1^2 = \sigma_{2,l}^2 = 0$dB, $\gamma_l = 8$dB, $\sigma_{1,k}^2 = \sigma_{2,k}^2 = -20$dB, $\forall l \in \mathcal{L}$, $\forall k \in \mathcal{K}$. The tolerance parameters for all algorithms are chosen as $\delta_1 = \delta_2 = 10^{-3}$ while $N_{max} = 20$. The parameter c in Table II is given by $c = 2$. All the results are obtained by averaging over 200 independent Monte Carlo runs.

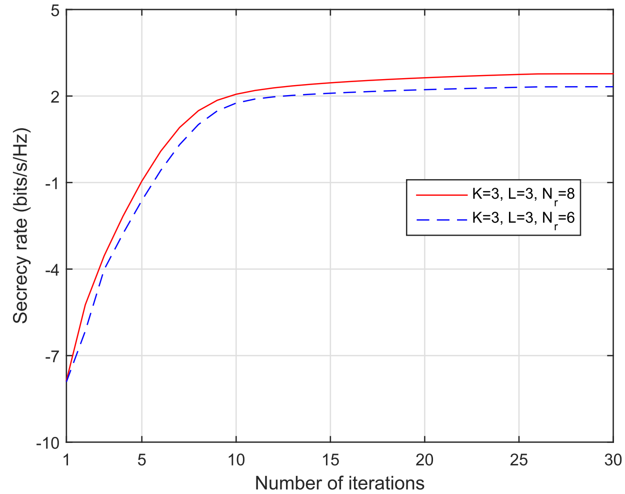

Fig. 2. Secrecy rate versus the number of iterations for the proposed joint transceiver design algorithm ($\varepsilon = 0.1$, $K = 3$).

Fig. 2 shows the system secrecy rate performance versus the number of iterations for the proposed penalty-CCCP based joint transceiver design algorithm when $\varepsilon = 0.1$. In particular, we show the results for $K = 3$ and $N_r \in \{6, 8\}$. It can be observed from this figure that the proposed algorithm converges within a reasonable number of iterations for the different numbers of RS antennas considered. Moreover the proposed algorithm with 8 RS antennas achieves slightly better convergence performance compared to the one with 6 RS antennas. In Fig. 3, we plot the value of penalty term, defined as $Z = \|\mathbf{X} - \mathbf{W}\mathbf{H}_1\mathbf{P}\|^2$, versus the number of iterations. From the figure, we can find that the convergence of the penalty term is very fast, which supports our claim that the proposed penalty-CCCP algorithm can tackle the equality constraint effectively. Furthermore, this convergence rate increases with N_r.

In the next series of simulations, we compare the performance of the proposed joint transceiver design algorithm with AN, with that of its counterpart without AN. We show the results for $\varepsilon = 0.1$ and $\varepsilon = 0.3$. Fig. 4 shows the steady-state secrecy rate performance of the two algorithms versus the number of RS antennas (N_r). From the figure, it is seen that the secrecy rate performance improves as the number of RS antennas increases. Moreover, the proposed joint transceiver design with AN achieves superior performance than the design without AN. In addition, an increase in the size of the eavesdroppers' CSI errors degrade the performance.

[3] Note that the objective function in (54) can be equivalently converted to L SOC constraints of dimension $L + 2$, by following the previous approach.

[4] In this work, we consider the conventional AO-CCCP algorithm for comparison. For each iteration, this algorithm consists of two steps: 1) optimize $\mathbf{P}$ and $\mathbf{V}$ by fixing $\mathbf{W}$; 2) optimize $\mathbf{W}$ by fixing $\mathbf{P}$ and $\mathbf{V}$. The CCCP approach is applied for each step [39], [40].

$$\begin{aligned} \max_{\substack{\mathbf{P},\mathbf{W},\mathbf{X},\mathbf{V}\succeq\mathbf{0},\mathbf{G}_{11},\mathbf{G}_{22},\theta\geq 0,\\ \{\alpha_{k,l},\beta_{k,l},\rho_{k,l},\phi_{k,l},\mathbf{F}_l,\mathbf{D}_l,\\ \mu_k\geq 0,\lambda_k\geq 0,\omega_k\geq 0\}}} \quad & f_1(\mathbf{X},\mathbf{W},\mathbf{V},\theta) - \hat{f}_2(\mathbf{s}_1^i,\mathbf{s}_1) \\ \text{s.t.}\ & (69), (71), (73), (74), (37), (38), (42), (43), \\ & (48), (47), (51), (52), (18), \text{ and } (53), \quad \forall l \in \mathcal{L},\ \forall k \in \mathcal{K}. \end{aligned} \tag{54}$$

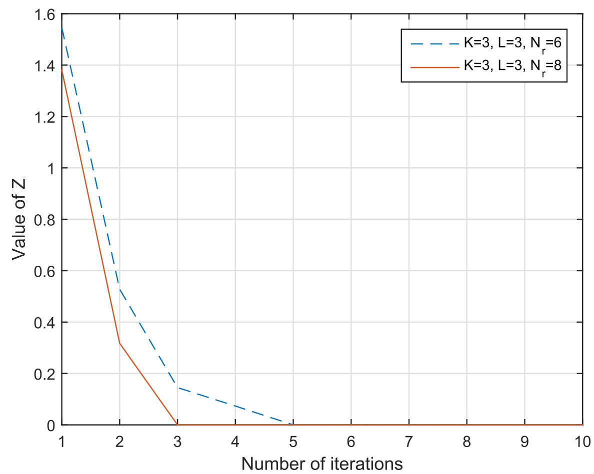

Fig. 3. Penalty term $Z = \|\mathbf{X} - \mathbf{W}\mathbf{H}_1\mathbf{P}\|^2$ versus the number of iterations for the proposed joint transceiver design algorithm ($\varepsilon = 0.1$, $K = 3$).

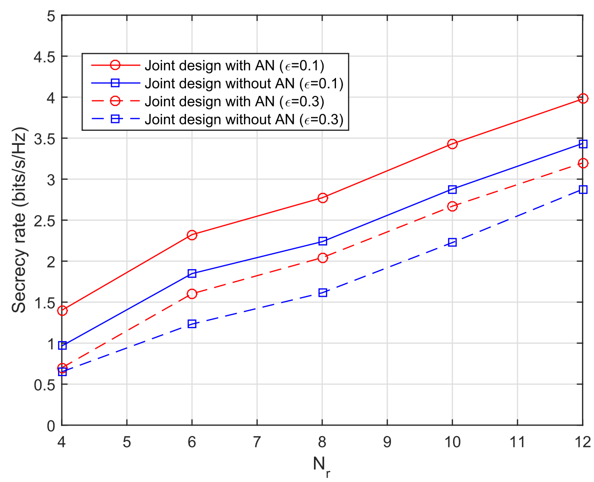

Fig. 4. Secrecy rate versus the number of relay antennas (N_r) for the proposed joint transceiver design algorithm ($K = 3$).

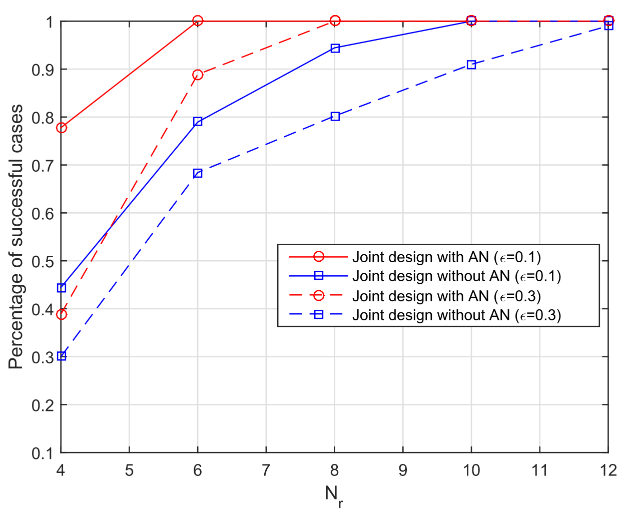

Fig. 5. Percentage of successful cases versus the number of relay antennas (N_r) for the proposed transceiver design algorithm, with and without AN ($K = 3$).

Fig. 5 compares the percentage of successful cases achieved by the proposed joint algorithm with and without AN designs versus the number of RS antennas (N_r), for different values of channel error bound and $K = 3$. In this simulation, a design algorithm is considered unsuccessful for a channel realization if CVX reports an infeasible status. From this figure, it is seen that the rate of success of both algorithms increases with N_r, while it decreases as the eavesdroppers' CSI error bound increases. Moreover, the proposed joint transceiver design algorithm with AN outperforms its counterpart without AN, which shows the effectiveness of the proposed approach.

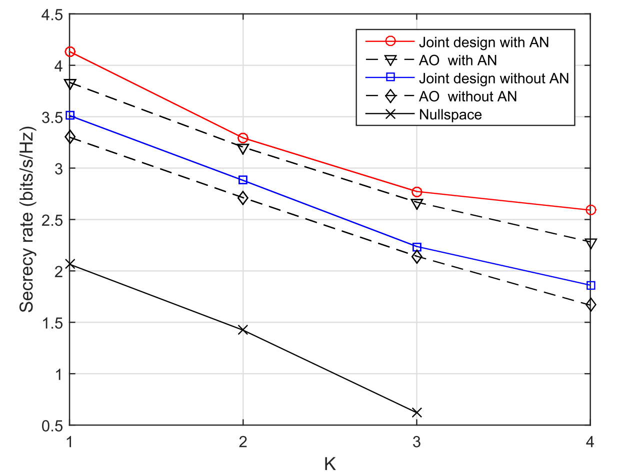

Fig. 6. Secrecy rate versus the number of eavesdroppers for the proposed joint transceiver design algorithm with and without AN, and for the AO and null-space based beamforming algorithms ($N_r = 8$, $\varepsilon = 0.1$).

In the following experiment, we compare the performance of the proposed joint transceiver design algorithm with and without AN, as well as of the conventional AO-CCCP based transceiver design and the existing null-space beamforming approach for MIMO relay systems. For the null-space beamforming scheme, the RS attempts to nullifies eavesdroppers reception by exploiting the null-space of the estimated eavesdroppers' second phase channel matrix in the design of the RS AF transmission matrix [19]–[21]. Therefore, this method is only applicable, when $N_r > 2K$. Fig. 6 depicts the system secrecy rate performance versus the number of eavesdroppers (K) for the analyzed algorithms. In particular, we set $N_r = 8$ and $\varepsilon = 0.1$ in this simulation. From the results, the proposed joint transceiver design algorithm outperforms the conventional alternating optimization based algorithm and the null-space beamforming approach. These results demonstrate the effectiveness of the proposed transceiver design to handle colluding eavesdroppers.

V. Conclusion

In this paper, we have investigated the joint transceiver design problem for secure communications over a MIMO relay with multiple legitimate users and colluding eavesdroppers. To tackle the highly coupled corresponding optimization problem, a novel penalty-CCCP algorithm has been proposed to jointly optimize the BS beamforming matrix, the RS AF transformation matrix and the AN covariance matrix. The eavesdroppers' CSI errors have also been considered in the design. We have analyzed the convergence of the proposed joint transceiver design algorithm and proved that the proposed penalty-CCCP algorithm converges to a stationary solution of the original problem. Our simulation results have shown that the proposed algorithm achieves better performance than the existing transceiver design algorithms. In general, the proposed algorithmic framework, especially the penalty-CCCP concept, can be also applied to solve other highly coupled optimization problems appearing in the design of wireless communication systems and other applications in science and engineering.

Appendix A
Derivation of (9)

For the colluding eavesdroppers, by stacking $2K$ received signals (eq. (7) and eq. (8)) we can express the received data vector as follows

$$\tilde{\mathbf{y}} = \tilde{\mathbf{H}}_l \mathbf{p}_l b_l + \tilde{\mathbf{n}}_l \quad (55)$$

where $\tilde{\mathbf{H}}_l = [\mathbf{g}_{1,1},\ \mathbf{H}_1^H \mathbf{W}^H \mathbf{g}_{2,1},\ \ldots,\ \mathbf{g}_{1,K},\ \mathbf{H}_1^H \mathbf{W}^H \mathbf{g}_{2,K}]^H \in \mathbb{C}^{2K \times N_t}$ denotes the equivalent multi-antenna channel matrix, and $\tilde{\mathbf{n}}_l = [\mathbf{g}_{1,1}^H \sum_{l' \neq l} \mathbf{p}_{l'} b_{l'} + n_{1,1},\ \mathbf{g}_{2,1}^H \mathbf{W}\mathbf{H}_1 \sum_{l' \neq l} \mathbf{p}_{l'} b_{l'} + \mathbf{g}_{2,1}^H \mathbf{W}\mathbf{n}_1 + \mathbf{g}_{2,1}^H \mathbf{v} + n_{2,1},\ \ldots,\ \mathbf{g}_{1,K}^H \sum_{l' \neq l} \mathbf{p}_{l'} b_{l'} + n_{1,K},\ \mathbf{g}_{2,K}^H \mathbf{W}\mathbf{H}_1 \sum_{l' \neq l} \mathbf{p}_{l'} b_{l'} + \mathbf{g}_{2,K}^H \mathbf{W}\mathbf{n}_1 + \mathbf{g}_{2,K}^H \mathbf{v} + n_{2,K}]^T \in \mathbb{C}^{2K \times 1}$ denotes the interference and noise vector. The eavesdropper aims to design a joint receiver to maximize the received SINR. By defining the joint receiver as $\mathbf{f}_l \in \mathbb{C}^{2K \times 1}$, then we formulate the following problem:

$$\max_{\mathbf{f}_l} \Upsilon \quad (56)$$

where $\Upsilon = \frac{|\mathbf{f}_l^H \tilde{\mathbf{p}}_l|^2}{\mathbf{f}_l^H \Sigma_l \mathbf{f}_l}$ denotes the SINR at the colluding eavesdropper for user l. Here, $\tilde{\mathbf{p}}_l = \tilde{\mathbf{H}}_l \mathbf{p}_l$ denotes the equivalent channel vector and

$$\begin{aligned}\Sigma_l = \mathrm{Diag}\{[&\sigma_{1,1}^2 + \mathbf{g}_{1,1}^H \sum_{l' \neq l} \mathbf{p}_{l'} \mathbf{p}_{l'}^H \mathbf{g}_{1,1},\\ &\mathbf{g}_{2,1}^H \mathbf{W}\mathbf{H}_1 \sum_{l' \neq l} \mathbf{p}_{l'} \mathbf{p}_{l'}^H \mathbf{H}_1^H \mathbf{W}^H \mathbf{g}_{2,1} + |\mathbf{g}_{2,1}^H \mathbf{W}|^2 \sigma_1^2\\ &+ \mathbf{g}_{2,1}^H \mathbf{V} \mathbf{g}_{2,1} + \sigma_{2,1}^2,\ \ldots,\ \sigma_{1,K}^2 + \mathbf{g}_{1,K}^H \sum_{l' \neq l} \mathbf{p}_{l'} \mathbf{p}_{l'}^H \mathbf{g}_{1,K},\\ &\mathbf{g}_{2,K}^H \mathbf{W}\mathbf{H}_1 \sum_{l' \neq l} \mathbf{p}_{l'} \mathbf{p}_{l'}^H \mathbf{H}_1^H \mathbf{W}^H \mathbf{g}_{2,K} + |\mathbf{g}_{2,K}^H \mathbf{W}|^2 \sigma_1^2\\ &+ \mathbf{g}_{2,K}^H \mathbf{V} \mathbf{g}_{2,K} + \sigma_{2,K}^2]\}\end{aligned} \quad (57)$$

denotes the noise covariance matrix. Based on the theory of Rayleigh quotient [41], the solution to problem (56) is given by

$$\mathbf{f}_l = \Sigma_l^{-1} \tilde{\mathbf{p}}_l. \quad (58)$$

By using (58) the maximum received SINR of user l can be expressed as (59), as shown at the bottom of this page.

TABLE III

Penalty-CCCP Algorithm for Problem (60)

0. initialize $\boldsymbol{x}^0 \in \mathcal{X}$, $\varrho_0 > 0$, and set $c > 1$, $k = 0$
1. **repeat**
2. $\boldsymbol{x}^{k+1} = CCCP(P_{\varrho_k}, \boldsymbol{x}^k)$
3. $\varrho_{k+1} = c\varrho_k$
4. $k = k + 1$
5. **until** some termination criterion is met

Appendix B
Penalty-CCCP Method

In this appendix, we present our proposed penalty-CCCP method in a general framework and discuss its convergence. Let us consider a general problem:

$$\begin{aligned}(P)\quad &\min_{\mathbf{x} \in X} f(\boldsymbol{x})\\ &\text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0},\\ &\quad\ \ \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}.\end{aligned} \quad (60)$$

where $X \subseteq \mathbb{R}^n$ denotes a closed convex set, $f(\boldsymbol{x})$ is a scalar continuously differentiable function, $\boldsymbol{h}(\boldsymbol{x}) \in \mathbb{R}^p$ is a vector of p continuously differentiable functions, and $\boldsymbol{g}(\boldsymbol{x}) \in \mathbb{R}^q$ is a vector of differentiable but *possibly nonconvex* functions. Note that this problem is different from problem (29) in [42] since here, we do not assume the components functions in $\boldsymbol{g}(\boldsymbol{x})$ to be convex.

When the equality constraints are very difficult to handle, it may be possible to tackle problem (60) using a penalty method [31], i.e., by solving the penalized problem:

$$\begin{aligned}(P_\varrho)\quad &\min_{\boldsymbol{x} \in X} f_\rho(\boldsymbol{x}) \triangleq f(\boldsymbol{x}) + \frac{\varrho}{2}\|\boldsymbol{h}(\boldsymbol{x})\|^2\\ &\text{s.t. } \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}.\end{aligned} \quad (61)$$

where $\varrho > 0$ is a scalar penalty parameter that prescribes a high cost for the violation of the equality constraints. In particular, when $\varrho \to \infty$, solving the above problem yields an identical solution to problem (60) [31]. However, it is still difficult to globally solve problem (P_ϱ) when it is a nonconvex problem. Thus, we propose the penalty-CCCP method summarized in Table III to solve problem (60), where in each iteration problem (P_ϱ) is approximately solved using the CCCP method. The resulting solution is denoted simply as $\boldsymbol{x}^{k+1} = CCCP(P_{\varrho_k}, \boldsymbol{x}^k)$ in Step 2 of Table III.

Regarding the convergence of the penalty-CCCP method, we have the following theorem.

Theorem 1: Let $\boldsymbol{\nu}^k \geq 0$ be the multiplier associated with the inequality constraint set $\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}$ of problem (P_{ρ_k}). Let $\{\boldsymbol{x}^k\}$ be the sequence generated by penalty-CCCP method where

$$\begin{aligned}\mathrm{SINR}_l^E &= \tilde{\mathbf{p}}_k^H \Sigma_k^{-1} \tilde{\mathbf{p}}_k\\ &= \sum_{k=1}^{K} \left(\frac{|\mathbf{g}_{1,k}^H \mathbf{p}_l|^2}{\sum_{l' \neq l}^{L} |\mathbf{g}_{1,k}^H \mathbf{p}_{l'}|^2 + \sigma_{1,k}^2} + \frac{|\mathbf{g}_{2,k}^H \mathbf{W}\mathbf{H}_1 \mathbf{p}_l|^2}{\sum_{l' \neq l}^{L} |\mathbf{g}_{2,k}^H \mathbf{W}\mathbf{H}_1 \mathbf{p}_{l'}|^2 + \sigma_1^2 \|\mathbf{g}_{2,k}^H \mathbf{W}\|^2 + \mathbf{g}_{2,k}^H \mathbf{V} \mathbf{g}_{2,k} + \sigma_{2,k}^2} \right)\end{aligned} \quad (59)$$

the termination condition is

$$\left\|\nabla f_{\varrho_k}(\boldsymbol{x}^k) + \nabla \boldsymbol{g}(\boldsymbol{x}^k)^T \boldsymbol{v}^k\right\| \leq \epsilon_k, \quad \forall k \tag{62}$$

with $\epsilon_k \to 0$ as $k \to \infty$. Suppose that $\boldsymbol{x}^\star$ is a limit point of the sequence $\{\boldsymbol{x}^k\}$ and $\nabla f(\boldsymbol{x}^\star)$ is bounded. In addition, assume that *Robinson's condition*[5] [43, Ch. 3] holds for problem (P) at $\boldsymbol{x}^\star$, i.e.,

$$\left\{\left.\begin{pmatrix} \nabla \boldsymbol{h}(\boldsymbol{x}^\star)\boldsymbol{d}_x \\ \nabla g_1(\boldsymbol{x}^\star)^T\boldsymbol{d}_x - v_1 \\ \vdots \\ \nabla g_q(\boldsymbol{x}^\star)^T\boldsymbol{d}_x - v_q \end{pmatrix}\right| \begin{array}{c} \boldsymbol{d}_x \in \mathbb{R}^n, \\ \boldsymbol{v} \in \mathbb{R}^q, v_\ell \leq 0, \\ \forall \ell \in I_i(\boldsymbol{x}_i^\star),\ \forall\, i \end{array}\right\} = \mathbb{R}^p \times \mathbb{R}^q \tag{63}$$

where $\boldsymbol{v} \triangleq [v_1, \ldots, v_q]^T$, $I_i(\boldsymbol{x}^\star)$ is the i-th index set of active inequality constraints at $\boldsymbol{x}^\star$, i.e.,

$$I_i(\boldsymbol{x}^\star) \triangleq \{\ell \mid g_\ell(\boldsymbol{x}^\star) = 0,\ 0 \leq \ell \leq q\},$$

and $g_\ell(\boldsymbol{x}^\star)$ denotes the ℓ-th component function of $\boldsymbol{g}(\boldsymbol{x}^\star)$. Then $\boldsymbol{x}^\star$ is a KKT point of problem (P).

Proof: Following the proof of Theorem 4.1 in [42], let $\boldsymbol{s}^k \triangleq \nabla f_{\varrho_k}(\boldsymbol{x}^k) + \nabla \boldsymbol{g}(\boldsymbol{x}^k)^T \boldsymbol{v}^k$. Since $\epsilon_k \to 0$ as $k \to 0$, it follows that $\boldsymbol{s}^k$ approaches zero as k goes to infinity. Define $\boldsymbol{\mu}^k \triangleq \varrho_k \boldsymbol{h}(\boldsymbol{x}_k)$. Then we have $\nabla f_{\varrho_k}(\boldsymbol{x}^k) = \nabla f(\boldsymbol{x}^k) + \nabla \boldsymbol{h}(\boldsymbol{x}^k)^T \boldsymbol{\mu}^k$. Using these facts, we obtain

$$\nabla f(\boldsymbol{x}^k) + \nabla \boldsymbol{h}(\boldsymbol{x}^k)^T \boldsymbol{\mu}^k + \nabla \boldsymbol{g}(\boldsymbol{x}^k)^T \boldsymbol{v}^k - \boldsymbol{s}^k = \boldsymbol{0}.$$

Besides, under the assumption of Robinson's condition, it follows [43, Lemma 3.26] that $\boldsymbol{v}_k$'s are bounded. Then using this fact together with Ronbinson's condition, and applying a similar argument as that for the proof of Theorem 4.1 of [42], it can be shown that $\boldsymbol{\mu}^k$ is bounded and thus has a convergent subsequence. Finally, by restricting to subsequence and taking the limit on both sides of the above equality, we can show that there exists $\boldsymbol{\mu}^\star$ and $\boldsymbol{v}^\star \geq 0$ together with $\boldsymbol{x}^\star$ such that

$$\nabla f(\boldsymbol{x}^\star) + \nabla \boldsymbol{h}(\boldsymbol{x}^\star)^T \boldsymbol{\mu}^\star + \nabla \boldsymbol{g}(\boldsymbol{x}^\star)^T \boldsymbol{v}^\star = \boldsymbol{0}.$$

This completes the proof. ■

Remark: The termination condition (62) is used to establish the convergence of the penalty-CCCP algorithm. In practice, however, it is also reasonable to terminate the penalty-CCCP algorithm based on the progress of the objective value $f_\rho(\boldsymbol{x})$, i.e., $\frac{|f_\rho(\boldsymbol{x}^k) - f_\rho(\boldsymbol{x}^{k-1})|}{|f_\rho(\boldsymbol{x}^{k-1})|} \leq \epsilon_k$.

[5]Robinson's condition is a type of constraint qualification condition used for KKT analysis and the assumption is a standard one that is, made in many previous works on constrained optimization, e.g., [32], [33], [43].

APPENDIX C
PROOF OF EQUIVALENT TRANSFORMATION

Firstly, let us focus on the subtrahend in $C(\mathbf{P}, \mathbf{V}, \mathbf{W}, \mathbf{X})$. Due to the monotonicity of the logarithm function, we have (64), as shown at the bottom of this page. Then by introducing auxiliary variables $\{\alpha_{k,l}\}$ and $\{\rho_{k,l}\}$ corresponding to the upper bounds of the eavesdropper's SINR, we can move the maximum functions from (21a) to the constraints. The resultant equivalent optimization problem is given by (65), as shown at the top of the next page. Note that constraints (65c) and (65d) are difficult to handle. In order to transform them to tractable forms, we introduce two additional sets of auxiliary variables $\{\phi_{k,l}\}$ and $\{\beta_{k,l}\}$, which are subject to $\frac{|\mathbf{g}_{1,k}^H \mathbf{p}_l|^2}{\sum_{l' \neq l}^L |\mathbf{g}_{1,k}^H \mathbf{p}_{l'}|^2 + \sigma_{1,k}^2} \leq \frac{1}{\phi_{k,l}} \leq \alpha_{k,l}$ and $\frac{|\mathbf{g}_{2,k}^H \mathbf{x}_l|^2}{\sum_{l' \neq l}^L |\mathbf{g}_{2,k}^H \mathbf{x}_{l'}|^2 + \sigma_1^2 \|\mathbf{g}_{2,k}^H \mathbf{W}\|^2 + \mathbf{g}_{2,k}^H \mathbf{V} \mathbf{g}_{2,k} + \sigma_{2,k}^2} \leq \frac{1}{\beta_{k,l}} \leq \rho_{k,l}$, $\forall l \in \mathcal{L}$, $\forall k \in \mathcal{K}$. Therefore, (65) can be formulated as an equivalent problem shown in (66), as shown at the top of the next page.

In the following, we prove that problems (22) and (66) are equivalent. Note that the optimum u_l^{opt} for minimizing (22) is given by $\tilde{u}_l$ in (24). Then by fixing the other variables it can be seen that the objective function of (22) is convex with respect to v_l. Hence, the optimum v_l^{opt} in (22) can be obtained based on the first order optimality condition, i.e.,

$$v_l^{opt} = \frac{1}{e_l(u_l, \mathbf{X}, \mathbf{W}, \mathbf{V})}. \tag{67}$$

By substituting the optimum $\{u_l^{opt}, v_l^{opt}\}$, $\forall l \in \mathcal{L}$, in (22), we have an equivalent optimization problem shown in (68), as shown at the top of the next page. Finally, by substituting $\tilde{e}_l$ from (25) in (68), we obtain problem (66). This completes the proof.

APPENDIX D
THE SOC CONSTRAINTS

It can be seen that constraint (14b) can be simply converted to

$$\|\text{vec}(\mathbf{P})\| \leq \sqrt{\mathrm{P_t}}. \tag{69}$$

Note that (14c) can be rewritten as the following constraint:

$$\|\mathbf{X}\|^2 + \sigma_1^2\|\mathbf{W}\|^2 + \frac{(P_r - \text{Tr}\{\mathbf{V}\} - 1)^2}{4} \leq \frac{(P_r - \text{Tr}\{\mathbf{V}\} + 1)^2}{4}. \tag{70}$$

$$\begin{aligned} &\max_{\forall \triangle \mathbf{g}_{j,k} \in \mathcal{R}_{j,k}} \frac{1}{2}\sum_{l=1}^{L} \log\left(1 + \eta_l^E\right) \\ &= \frac{1}{2}\sum_{l=1}^{L} \log\Bigg(1 + \sum_{k=1}^{K}\Big(\max_{\forall \triangle \mathbf{g}_{1,k} \in \mathcal{R}_{1,k}} \frac{|\mathbf{g}_{1,k}^H \mathbf{p}_l|^2}{\sum_{l' \neq l}^L |\mathbf{g}_{1,k}^H \mathbf{p}_{l'}|^2 + \sigma_{1,k}^2} \\ &\quad + \max_{\forall \triangle \mathbf{g}_{2,k} \in \mathcal{R}_{2,k}} \frac{|\mathbf{g}_{2,k}^H \mathbf{x}_l|^2}{\sum_{l' \neq l}^L |\mathbf{g}_{2,k}^H \mathbf{x}_{l'}|^2 + \sigma_1^2 \|\mathbf{g}_{2,k}^H \mathbf{W}\|^2 + \mathbf{g}_{2,k}^H \mathbf{V} \mathbf{g}_{2,k} + \sigma_{2,k}^2}\Big)\Bigg) \end{aligned} \tag{64}$$

$$\max_{\{\alpha_{k,l},\rho_{k,l}\},\mathbf{P},\mathbf{W},\mathbf{V}\succeq\mathbf{0},\mathbf{X}} \ \frac{1}{2}\sum_{l=1}^{L}\log\left(1+\eta_l^D\right)-\frac{1}{2}\sum_{l=1}^{L}\log\left(1+\sum_{k=1}^{K}\alpha_{k,l}+\sum_{k=1}^{K}\rho_{k,l}\right)-\tau\theta \tag{65a}$$

$$\text{s.t. } (14b)-(14d),\ (18),\ (21c), \tag{65b}$$

$$\frac{|\mathbf{g}_{1,k}^H\mathbf{p}_l|^2}{\sum_{l'\neq l}^{L}|\mathbf{g}_{1,k}^H\mathbf{p}_{l'}|^2+\sigma_{1,k}^2}\leq\alpha_{k,l}, \tag{65c}$$

$$\frac{|\mathbf{g}_{2,k}^H\mathbf{x}_l|^2}{\sum_{l'\neq l}^{L}|\mathbf{g}_{2,k}^H\mathbf{x}_{l'}|^2+\sigma_1^2\|\mathbf{g}_{2,k}^H\mathbf{W}\|^2+\mathbf{g}_{2,k}^H\mathbf{V}\mathbf{g}_{2,k}+\sigma_{2,k}^2}\leq\rho_{k,l}, \tag{65d}$$

$$\forall\triangle\mathbf{g}_{j,k}\in\mathcal{R}_{j,k},\quad j\in\{1,2\},\ \forall l\in\mathcal{L},\ \forall k\in\mathcal{K} \tag{65e}$$

$$\max_{\substack{\{\alpha_{k,l},\rho_{k,l},\beta_{k,l},\phi_{k,l}\},\\ \mathbf{P},\mathbf{W},\mathbf{V}\succeq\mathbf{0},\mathbf{X}}} \ \frac{1}{2}\sum_{l=1}^{L}\log\left(1+\eta_l^D\right)-\frac{1}{2}\sum_{l=1}^{L}\log\left(1+\sum_{k=1}^{K}\alpha_{k,l}+\sum_{k=1}^{K}\rho_{k,l}\right)-\tau\theta \tag{66a}$$

$$\text{s.t. } (65b), \tag{66b}$$

$$|\mathbf{g}_{2,k}^H\mathbf{x}_l|^2-\frac{(\sigma_1^2\|\mathbf{g}_{2,k}^H\mathbf{W}\|^2+\sum_{l'\neq l}^{L}|\mathbf{g}_{2,k}^H\mathbf{x}_{l'}|^2+\mathbf{g}_{2,k}^H\mathbf{V}\mathbf{g}_{2,k}+\sigma_{2,k}^2)}{\beta_{k,l}}\leq 0, \tag{66c}$$

$$|\mathbf{g}_{1,k}^H\mathbf{p}_l|^2-\frac{(\sigma_{1,k}^2+\sum_{l'\neq l}^{L}|\mathbf{g}_{1,k}^H\mathbf{p}_{l'}|^2)}{\phi_{k,l}}\leq 0, \tag{66d}$$

$$\rho_{k,l}\beta_{k,l}\geq 1,\quad \alpha_{k,l}\phi_{k,l}\geq 1, \tag{66e}$$

$$\forall\triangle\mathbf{g}_{j,k}\in\mathcal{R}_{j,k},\quad j\in\{1,2\},\ \forall l\in\mathcal{L},\ \forall k\in\mathcal{K} \tag{66f}$$

$$\max_{\substack{\{\alpha_{k,l},\beta_{k,l},\rho_{k,l},\phi_{k,l}\},\\ \mathbf{P},\mathbf{W},\mathbf{V}\succeq\mathbf{0},\mathbf{X}}} \ \frac{1}{2}\sum_{l=1}^{L}\log\left(\frac{1}{\tilde{e}_l}\right)-\frac{1}{2}\sum_{l=1}^{L}\log\left(1+\sum_{k=1}^{K}\alpha_{k,l}+\sum_{k=1}^{K}\rho_{k,l}\right)-\tau\theta, \tag{68a}$$

$$\text{s.t. } (22b)-(22g). \tag{68b}$$

$$\begin{bmatrix} -\tilde{\mathbf{Y}}_{k,l}+\lambda_k\mathbf{I} & -\tilde{\mathbf{Y}}_{k,l}\hat{\mathbf{g}}_{1,k} \\ -\hat{\mathbf{g}}_{1,k}^H\tilde{\mathbf{Y}}_{k,l} & \frac{2\sigma_{1,k}^2}{\phi_{k,l}^i}-\frac{\sigma_{1,k}^2\phi_{k,l}}{\phi_{k,l}^{i2}}-\hat{\mathbf{g}}_{1,k}^H\tilde{\mathbf{Y}}_{k,l}\hat{\mathbf{g}}_{1,k}-\lambda_k\varepsilon_{1,k}^2 \end{bmatrix}\succeq\mathbf{0}, \tag{75}$$

$$\begin{bmatrix} -\mathbf{Y}_{k,l}+\lambda_k\mathbf{I} & -\mathbf{Y}_{k,l}\hat{\mathbf{g}}_{1,k} \\ -\hat{\mathbf{g}}_{1,k}^H\mathbf{Y}_{k,l} & \frac{2\sigma_{1,k}^2}{\phi_{k,l}^i}-\frac{\sigma_{1,k}^2\phi_{k,l}}{\phi_{k,l}^{i2}}-\hat{\mathbf{g}}_{1,k}^H\mathbf{Y}_{k,l}\hat{\mathbf{g}}_{1,k}-\lambda_k\varepsilon_{1,k}^2 \end{bmatrix}\succeq\mathbf{0}, \tag{76}$$

By taking the square root of (70), (14c) can be converted to the following SOC constraint:

$$\left\|\left[\text{vec}(\mathbf{X})^{\mathrm{T}},\sigma_1\text{vec}(\mathbf{W})^{\mathrm{T}},\frac{(P_r-\text{Tr}\{\mathbf{V}\}-1)}{2}\right]\right\| \leq\frac{(P_r-\text{Tr}\{\mathbf{V}\}+1)}{2}. \tag{71}$$

Then, let us focus on (22e) and (22f). Note that (22e) can be rewritten as

$$1+\frac{(\rho_{k,l}-\beta_{k,l})^2}{4}\leq\frac{(\rho_{k,l}+\beta_{k,l})^2}{4}. \tag{72}$$

By taking the square root of both sides of (72), (22e) can be converted to the following SOC constraint:

$$\left\|\left[1,\frac{(\rho_{k,l}-\beta_{k,l})}{2}\right]\right\|\leq\frac{(\rho_{k,l}+\beta_{k,l})}{2}. \tag{73}$$

By following the same approach, (22f) can be equivalently transformed into the following SOC constraint:

$$\left\|\left[1,\frac{(\alpha_{k,l}-\phi_{k,l})}{2}\right]\right\|\leq\frac{(\alpha_{k,l}+\phi_{k,l})}{2}. \tag{74}$$

APPENDIX E
THE LMI CONSTRAINTS

Let us equivalently convert (40) into (75) based on the S-procedure, where $\tilde{\mathbf{Y}}_{k,l}\triangleq\mathbf{p}_l\mathbf{p}_l^H-\sum_{l'\neq l}^{L}\left(2\Re\{\frac{\mathbf{p}_l\mathbf{p}_l^{iH}}{\phi_{k,l}^i}\}-\frac{\mathbf{p}_l^i\mathbf{p}_l^{iH}\phi_{k,l}}{\phi_{k,l}^{2\ i}}\right)$. We note that (75), as shown at the top of this page, is not LMI due to the second order term $\mathbf{p}_l\mathbf{p}_l^H$. To tackle this problem, we introduce a new auxiliary matrix variable $\mathbf{D}_l$, $\forall l\in\mathcal{L}$, which is subject to $\mathbf{D}_l\succeq\mathbf{p}_l\mathbf{p}_l^H$. Therefore,

(75) can be equivalently expressed as (76), as shown at the top of the previous page, and

$$\mathbf{D}_l \succeq \mathbf{p}_l \mathbf{p}_l^H, \tag{77}$$

where $\mathbf{Y}_{k,l} \triangleq \mathbf{D}_l - \sum_{l' \neq l}^{L} \left(2\Re\left\{\frac{\mathbf{p}_l \mathbf{p}_l^{iH}}{\phi_{k,l}^i}\right\} - \frac{\mathbf{p}_l^i \mathbf{p}_l^{iH} \phi_{k,l}}{\phi_{k,l}^{2\;i}}\right)$. It is readily seen that (76) now corresponds to an LMI constraint. Furthermore, by applying the Schur complement [36], (77) can be transformed into the following LMI constraint:

$$\begin{bmatrix} \mathbf{D}_l & \mathbf{p}_l \\ \mathbf{p}_l^H & 1 \end{bmatrix} \succeq \mathbf{0}. \tag{78}$$

REFERENCES

[1] Y. Liang, H. V. Poor, and S. Shamai, "Information theoretical security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, 2008.

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

[3] R. Liu and W. Trappe, *Securing Wireless Communications at the Physical Layer*. Norwell, MA, USA: Springer, 2009.

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

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

[6] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[7] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.

[8] T. Lv, H. Gao, and S. Yang, "Secrecy transmit beamforming for heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1154–1170, Jun. 2015.

[9] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[10] Y.-W. Hong, P.-C. Lan, and P.-C. J. Kuo, "Enhancing physical-layer secrecy in multiantenna wireless systems: An overview of signal processing approaches," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 29–40, Sep. 2013.

[11] X. Chen, C. Zhong, C. Yuen, and H.-H. Chen, "Multi-antenna relay aided wireless physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 40–46, Dec. 2015.

[12] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, Aug. 2014.

[13] C. Jeong, I.-M. Kim, and D. Kim, "Joint secure beamforming design at the source and the relay for an amplify-and-forward MIMO untrusted relay system," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 310–325, Jan. 2012.

[14] J. Mo, M. Tao, Y. Liu, and R. Wang, "Secure beamforming for MIMO two-way communications with an untrusted relay," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2185–2199, May 2014.

[15] C. Thai, J. Lee, and T. Q. S. Quek, "Physical-layer secret key generation with colluding untrusted relays," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1517–1530, Feb. 2016.

[16] Z. Ding, M. Peng, and H.-H. Chen, "A general relaying transmission protocol for MIMO secrecy communications," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3461–3471, Nov. 2012.

[17] 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.

[18] Q. Li, Q. Zhang, and J. Qin, "Secure relay beamforming for simultaneous wireless information and power transfer in nonregenerative relay networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2462–2467, Jun. 2014.

[19] C. Wang, H.-M. Wang, D. W. K. Ng, X.-G. Xia, and C. Liu, "Joint beamforming and power allocation for secrecy in peer-to-peer relay networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3280–3293, Jun. 2015.

[20] H.-M. Wang, F. Liu, and X.-G. Xia, "Joint source-relay precoding and power allocation for secure amplify-and-forward MIMO relay networks," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 8, pp. 1240–1250, Aug. 2014.

[21] 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.

[22] H. Gao, T. Lv, W. Wang, and N. C. Beaulieu, "Energy-efficient and secure beamforming for self-sustainable relay-aided multicast networks," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1509–1513, Nov. 2016.

[23] C. Wang and H.-M. Wang, "Robust joint beamforming and jamming for secure AF networks: Low-complexity design," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 2192–2198, May 2014.

[24] C. Zhang, H. Gao, H.-J. Liu, and T. Lv, "Robust beamforming and jamming for secure AF relay networks with multiple eavesdroppers," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Oct. 2014, pp. 495–500.

[25] D. Wang, B. Bai, W. Chen, and Z. Han, "Energy efficient secure communication over decode-and-forward relay channels," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 892–905, Mar. 2015.

[26] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Comput.*, vol. 15, no. 4, pp. 915–936, Apr. 2003.

[27] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.

[28] A. Mukherjee and A. L. Swindlehurst, "Detecting passive eavesdroppers in the MIMO wiretap channel," in *Proc. IEEE ICASSP*, Kyoto, Japan, Mar. 2012, pp. 2809–2812.

[29] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5473–5484, Nov. 2011.

[30] P. Apkarian and H. D. Tuan, "Robust control via concave minimization local and global algorithms," *IEEE Trans. Autom. Control*, vol. 45, no. 2, pp. 299–305, Feb. 2000.

[31] D. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 1999.

[32] Z. Lu and Y. Zhang, "Sparse approximation via penalty decomposition methods," *SIAM J. Optim.*, vol. 23, no. 4, pp. 2448–2478, 2013.

[33] Z. Lu and Y. Zhang, "Penalty decomposition methods for rank minimization," *Optim. Methods Softw.*, vol. 30, no. 3, pp. 531–558, May 2015.

[34] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proc. F, Commun., Radar Signal Process.*, vol. 130, no. 1, pp. 11–16, Feb. 1983.

[35] Y. Cheng and M. Pesavento, "Joint optimization of source power allocation and distributed relay beamforming in multiuser peer-to-peer relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2962–2973, Jun. 2012.

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

[37] M. Grant and S. Boyd. (Sep. 2013). *CVX: Matlab Software for Disciplined Convex Programming, Version 2.0 Beta*. [Online]. Available: http://cvxr.com/cvx

[38] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.

[39] Y. Cai, M.-M. Zhao, Q. Shi, B. Champagne, and M.-J. Zhao, "Joint transceiver design algorithms for multiuser MISO relay systems with energy harvesting," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4147–4164, Oct. 2016.

[40] M.-M. Zhao, Y. Cai, Q. Shi, M. Hong, and B. Champagne, "Joint transceiver designs for full-duplex K-pair MIMO interference channel with SWIPT," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 890–905, Feb. 2017.

[41] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[42] Q. Shi, M. Hong, X. Gao, E. Song, Y. Cai, and W. Xu, "Joint source-relay design for full-duplex MIMO AF relay systems," *IEEE Trans. Signal Process.*, vol. 62, no. 23, pp. 6118–6131, Dec. 2016.

[43] A. Ruszczynski, *Nonlinear Optimization*. Princeton, NJ, USA: Princeton Univ. Press, 2011.

Yunlong Cai (S'07–M'10–SM'16) received the B.S. degree in computer science from Beijing Jiaotong University, Beijing, China, in 2004, the M.Sc. degree in electronic engineering from the University of Surrey, Guildford, U.K., in 2006, and the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2010. From 2010 to 2011, he was a Post-Doctoral Fellow with the Electronics and Communications Laboratory, Conservatoire National des Arts et Metiers, Paris, France. Since 2011, he has been with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, where he is currently an Associate Professor. From 2016 to 2017, he was a Visiting Scholar with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His research interests include transceiver design for multiple-antenna systems, sensor array processing, adaptive filtering, full-duplex communications, cooperative and relay communications, and wireless information and energy transfer.

Qingjiang Shi received the Ph.D. degree in communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. From 2009 to 2010, he visited Prof. Z.-Q. (Tom) Luo's research group at the University of Minnesota, Twin Cities. In 2011, he was a Research Scientist with the Research and Innovation Center (Bell Labs China), Alcatel-Lucent, Shanghai China. He was an Associate Professor with the School of Information and Science Technology, Zhejiang Sci-Tech University, Hangzhou, China. He is currently a professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China. His current research interests lie in algorithm design for signal processing in advanced MIMO, cooperative communication, physical layer security, energy-efficient communication, wireless information, and power transfer.

Dr. Shi received the National Excellent Doctoral Dissertation Nomination Award in 2013, the Shanghai Excellent Doctoral Dissertation Award in 2012, and the Best Paper Award from the IEEE PIMRC 2009 Conference.

Benoit Champagne (S'87–M'89–SM'03) received the B.Ing. degree in engineering physics from the École Polytechnique de Montréal in 1983, the M.Sc. degree in physics from the Université de Montréal in 1985, and the Ph.D. degree in electrical engineering from the University of Toronto in 1990. From 1990 to 1999, he was an Assistant Professor and then an Associate Professor with INRS-Telecommunications, Université du Quebec, Montréal. In1999, he joined McGill University, Montreal, where he served as an Associate Chairman of Graduate Studies with the Department of Electrical and Computer Engineering from 2004 to 2007. He is currently a Full Professor with the Department of Electrical and Computer Engineering, McGill University. His research focuses on the study of advanced algorithms for the processing of communication signals by digital means. His interests span many areas of statistical signal processing, including detection and estimation, sensor array processing, adaptive filtering, and applications thereof to broadband communications and audio processing, where he has co-authored nearly 250 referred publications. His research has been funded by the Natural Sciences and Engineering Research Council of Canada, the Fonds de Recherche sur la Nature et les Technologies from the Government of Quebec, and some major industrial sponsors, including Nortel Networks, Bell Canada, InterDigital, and Microsemi.

Dr. Champagne was an Associate Editor of the *EURASIP Journal on Applied Signal Processing* from 2005 to 2007, the IEEE SIGNAL PROCESSING LETTERS from 2006 to 2008, and the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2010 to 2012, and a Guest Editor for two special issues of the *EURASIP Journal on Applied Signal Processing* published in 2007 and 2014. He has also served on the Technical Committees of several international conferences in the fields of communications and signal processing. In particular, he was the Registration Chair of the IEEE ICASSP 2004, the Co-Chair of the Antenna and Propagation Track for the IEEE VTC–Fall 2004, the Co-Chair of the Wide Area Cellular Communications Track for the IEEE PIMRC 2011, the Co-Chair of the Workshop on D2D Communications for the IEEE ICC 2015, and the Publicity Chair of the IEEE VTC–Fall 2016.

Geoffrey Ye Li (S'93–M'95–SM'97–F'06) received the B.S.E. and M.S.E. degrees from the Department of Wireless Engineering, Nanjing Institute of Technology, Nanjing, China, in 1983 and 1986, respectively, and the Ph.D. degree from the Department of Electrical Engineering, Auburn University, AL, in 1994.

He was a Teaching Assistant and then a Lecturer with Southeast University, Nanjing, from 1986 to 1991, a Research Assistant and a Teaching Assistant with Auburn University, from 1991 to 1994, and a Post-Doctoral Research Associate with the University of Maryland at College Park, Maryland, from 1994 to 1996. He was with AT&T Labs-Research, Red Bank, NJ, as a Senior Technical Staff Member and then a Principal Technical Staff Member from 1996 to 2000. Since 2000, he has been with the School of Electrical and Computer Engineering, Georgia Institute of Technology, as an Associate Professor and then a Full Professor. He has also been holding a Cheung Kong Scholar title at the University of Electronic Science and Technology of China since 2006.

His general research interests include statistical signal processing and communications, with emphasis on cross-layer optimization for spectral- and energy-efficient networks, cognitive radios and opportunistic spectrum access, and practical issues in LTE systems. In these areas, he has authored around 200 journal papers in addition to around 40 granted patents and numerous conference papers. He was an IEEE Fellow for his contributions to signal processing for wireless communications in 2005. His publications have been cited around 28 000 times and he has been recognized as the World's Most Influential Scientific Mind, also known as an Highly-Cited Researcher, by Thomson Reuters. He received the 2010 Stephen O. Rice Prize Paper Award, the 2013 WTC Wireless Recognition Award, the 2017 Award for Advances in Communication from the IEEE Communications Society, the 2013 James Evans Avant Garde Award, and the 2014 Jack Neubauer Memorial Award from the IEEE Vehicular Technology Society. He also received the 2015 Distinguished Faculty Achievement Award from the School of Electrical and Computer Engineering, Georgia Tech. He has been involved in editorial activities for over 20 technical journals for the IEEE, including the founding Editor-in-Chief of the IEEE 5G Tech Focus. He has organized and chaired many international conferences, including the Technical Program Vice-Chair of the IEEE ICC 2003, the Technical Program Co-Chair of the IEEE SPAWC 2011, the General Chair of the IEEE GlobalSIP 2014, and the Technical Program Co-Chair of the IEEE VTC 2016 (Spring).