Power Minimization for Secure Multi-User MISO NOMA System With Energy Harvesting

Jiasi Zhou, Yanjing Sun, *Member, IEEE*, Qi Cao, Song Li, *Member, IEEE*, Zhi Sun, *Member, IEEE*, and Xiaolin Wang

***Abstract*—Non-orthogonal multiple access (NOMA) has been deemed a promising technology to achieve high spectral efficiency. This paper investigates transmit power optimization for a multi-user multiple-input single-output (MISO) NOMA downlink system endowed with energy harvesting and physical layer security. More specifically, the system is designed to satisfy the data-link security requirements of its legitimate users while charging energy harvesting devices. To achieve that goal, we investigate a joint design on beamforming vectors and artificial noise (AN) covariance matrix. Then, an iterative difference-of-two-convex-function (D.C.) algorithm is proposed by applying the successive convex approximation (SCA) approach. To further reduce the computational complexity, a sub-optimal alternating D.C. (ADC) algorithm is devised by decoupling beamforming and mapping the AN to the null space of legitimate users. In the ADC algorithm, optimal null space and beamforming vectors are alternately solved by convex programming. Furthermore, the convergence of the proposed D.C. and ADC algorithms is derived. Extensive simulation results show that the two proposed AN-aided algorithms significantly reduce the transmit power cost over OMA and NOMA scheme without AN, especially when there are more eavesdroppers.**

***Index Terms*—Physical layer security, energy harvesting, non-orthogonal multiple access, beamforming.**

I. Introduction

THE rapid growth in personal mobile terminals such as smartphones and tablets has imposed more demanding requirements on the base station (BS) for higher spectral and energy efficiency. To cope with these challenges, non-orthogonal multiple access (NOMA) has recently received considerable attention from both industry and academia [1], [2]. Unlike the conventional orthogonal multiple access (OMA) scheme, NOMA multiplexes multiple users' signals over the same radio resources e.g. time/frequency/code while carefully designing the power levels based on users' channel gain [3]–[5]. As a result, mutual interference stronger than the desired signal can be eliminated via successive interference cancellation (SIC) at the receiver.

Meanwhile, the advances in Internet of Things (IoT) technology have enabled effective information collection from massive low-power devices. However, since these devices are mostly battery-powered and deployed in complex environments, changing batteries is time-consuming and of high inconvenience. Limited lifetime is the key bottleneck of signal collection. To tackle this challenge, harvesting energy from the radio frequency (RF) signals has been proposed to prolong the lifetime of these power-constrained devices [6]–[9]. In addition to simply harvesting energy from RF signals, it has been shown that simultaneous wireless information and power transfer (SWIPT) can be achieved by using different optimization criteria. The three most investigated criteria include minimizing transmission power subject to quality of service (QoS) requirements [10]–[12], optimizing energy efficiency or data rate subject to the transmit power cost [13], [14] and finally, maximizing the worst user's QoS under the transmission power constraint [15].

Despite its many advantages, energy harvesting has incurred information security concerns due to the broadcast nature of the wireless medium. More specifically, in addition to harvesting energy from the RF signals, neighboring devices may be able to eavesdrop legitimate users' information by decoding the signal. Therefore, it is imperative for energy harvesting to enhance its information transmission safety. In contrast to the conventional cryptographic security, physical layer security approach requires fewer overheads and computational complexity [16], [17]. We consider employing the physical layer security approach to guarantee information transmission safety in this work. Specifically, BS equipped with multi-antenna can steer legitimate users' beamforming vectors and inject artificial-noise (AN) to weaken the decoding capacity of eavesdroppers. AN injection is motivated by the fact that it is insufficient to prevent eavesdropping by using beamforming vectors alone if the eavesdroppers and legitimate users have very similar channel vectors. However, given unlimited transmit power, an accurate maximum achievable secrecy rate without AN injection is hard

Manuscript received August 12, 2019; revised November 29, 2019 and May 6, 2020; accepted June 1, 2020. Date of publication June 25, 2020; date of current version October 13, 2020. This work was supported in part by the Future Scientists Program of China University of Mining and Technology under Grant 2020WLKXJ028 and in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant KYCX20_1992. The review of this article was coordinated by Dr. H. Zhu. *(Corresponding author: Yanjing Sun.)*

Jiasi Zhou and Song Li are with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China (e-mail: jiasi_zhou@cumt.edu.cn; lisong@cumt.edu.cn).

Yanjing Sun is with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China, and also with the School of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: yjsun@cumt.edu.cn).

Qi Cao is with The Chinese University of Hong Kong (Shenzhen), Shenzhen 518172, China (e-mail: caoqi@cuhk.edu.cn).

Zhi Sun is with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China, and also with the Department of Electrical Engineering, the State University of New York at Buffalo, Buffalo, NY 14260 USA (e-mail: zhisun@buffalo.edu).

Xiaolin Wang is with the School of Mines, China University of Mining and Technology, Xuzhou 221116, China (e-mail: graciouswxl@126.com).

Digital Object Identifier 10.1109/TVT.2020.3005143

to calculate or simulate. In this paper, we verify that the secrecy rate is saturated by designing a reasonable metric.

A. Related Work and Motivation

Issues about jointly designing power splitting ratio and beamforming vectors for energy harvesting systems are addressed in [18]–[23]. In [18], the energy efficiency is maximized for a multiuser multiple-input single-output (MISO) SWIPT system. The authors of [19] then achieved the same goal with nearly closed-form solution by adopting Lagrangian relaxation under the zero-forcing beamforming scheme. In [20], optimal solution and suboptimal solutions are derived for minimizing the total transmission cost in a multiuser MISO SWIPT system. In [21] and [22], the stronger user's rate is maximized in a cooperative two-user MISO NOMA system subject to the QoS of weaker user. A similar system model is considered in [23], but its focus is to maximize energy efficiency.

With the goal of maximizing secrecy rate, some secure communication schemes have been studied [24]–[30]. In [24], an optimal power allocation strategy with closed-form expression is derived for a single-input single-output (SISO) system. In [25], optimal power allocation among useful signals and AN is determined for a two-user MISO NOMA system. In [26], the transmit scheme based on AN is designed for MISO systems in the presence of one legitimate user and multiple eavesdroppers. In [27] and [28], the beamforming and AN are designed for cognitive radio MISO systems, assuming that AN is mapped to the null space of legitimate receivers. In [29], the asymptotic-optimal transmit beamforming is derived in a SWIPT multi-user multiple-input multiple-output (MIMO) wiretap channel under the assumption of statistical channel state information. In [30], the secrecy beamforming vector is designed for a cooperative NOMA system with multiple eavesdroppers.

However, maximizing the secrecy sum rate alone will only guarantee the legitimate users' information security while ignoring their data rates as well as the system's energy efficiency. Thus, a system should be designed to satisfy the data-link security requirements of its legitimate users as well as to improve the energy efficiency or reduce transmit power cost. The authors in [31] and [32] extended the transmit power minimization problem above to the MIMO SWIPT system. In [33], the transmit power minimization scheme for a multi-user MIMO SWIPT secrecy energy harvesting communication system is investigated. Besides the works mentioned above, the transmission strategies for NOMA with energy harvesting have also been considered [34]–[37]. In [34], the authors consider a MISO multicarrier AN-aided NOMA system, where a full-duplex BS serves multiple half-duplex uplink and downlink users. The optimal resource allocation algorithm is designed to maximize system throughput. Furthermore, the secrecy outage probability in closed-form is derived in [35], assuming only one eavesdropper. In [36], AN-aided beamforming is designed to enhance the security of a MISO SWIPT system with one eavesdropper. In [37], the power minimization problem is investigated for secure MISO NOMA cognitive radio systems. However, regarding the power harvesting devices as the potential eavesdroppers is missing in the research of physical layer secure transmit schemes. Besides, assuming only one eavesdropper is an overly optimistic [24], [25], [29], [33], [35], [36], which drives us to investigate a more general scenario. In some complex scenarios, such as IoT or battery-powered systems, the device with single-antenna still has its benefits, e.g. low power cost. We thus study transmit power minimization problem for AN-aided MISO NOMA systems in the presence of multiple energy harvesting eavesdroppers.

B. Main Contributions

Specifically, the required minimum signal-to-interference-noise ratio (SINR) of the legitimate users, the minimum harvested energy level and the minimum secrecy rate are satisfied first. This goal is achieved by optimizing the design of the beamforming and AN covariance matrices. However, to ensure the physical layer security in multiple eavesdroppers scenarios, the maximum and minimum operators are concluded in the formulated optimization problem. It results in that the secrecy rate constraints are hard to transform into convex sets. Besides, due to the coupling between beamforming and AN covariance matrices, such power minimization problem is a non-convex problem. To cope with this challenge, two algorithms are devised, namely the difference-of-two-convex-functions/sets (D.C.) algorithm and the sub-optimal alternating D.C. (ADC) algorithm. The main contributions of this work are summarized as follows:

- To circumvent the coupling between the beamforming and its conjugate transposition, the semidefinite programming (SDP) is employed to decouple the beamforming. Furthermore, the successive convex approximation (SCA) is adopted to recast the primary problem into a convex form. The transformed optimization problem can be solved by the proposed local optimal D.C. iterative algorithm.
- To reduce the computation cycles of D.C. programming, an alternating optimization algorithm is developed to decouple the beamforming and the AN covariance matrices. In the ADC algorithm, the covariance matrix of AN is first mapped to the null space of legitimate users. Then, two problems are solved by convex optimization framework.
- The convergence of the proposed D.C. and ADC algorithms is proved, which can be applied to the optimization problems with D.C. constraints or D.C. objective function. The rank of AN covariance matrix in ADC is shown to be one by using the Karush-Kuhn-Tucker (KKT) conditions, which ensures that it can be decomposed into an one-dimensional vector.
- Simulation results show that the performance of the ADC algorithm with low computing time is close to that of the D.C. algorithm. It is also shown that the proposed AN-aided algorithms significantly reduce the transmit power cost over OMA and NOMA schemes without AN, especially when there are more eavesdroppers.

C. Organization and Notation

The remainder of the paper is organized as follows: In Section II, we elaborate the system model and problem formulation. After that, a joint beamforming and AN covariance

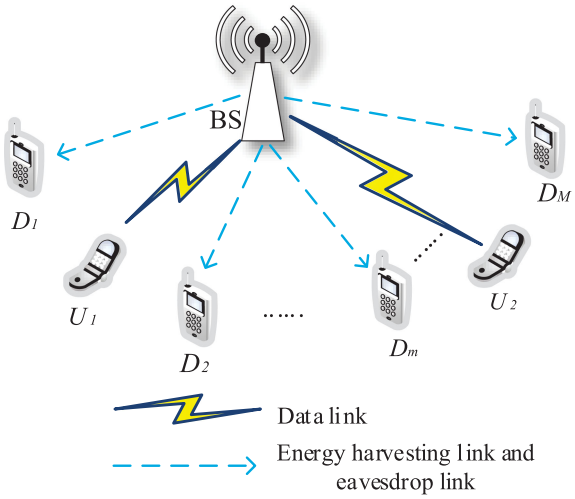


Fig. 1. Multi-user MISO NOMA system with energy harvesting devices.

matrices optimization algorithm is developed and analyzed in Section III. ADC algorithm is proposed in Section IV to solve the optimization problem with low computation cycles. Furthermore, the rank of AN covariance matrix and the convergence of the ADC algorithm are established. In Section V, simulation results are shown to validate the effectiveness of the proposed algorithms. Finally, a summary is provided in Section VI.

Notations: We use the uppercase and lowercase boldface letters to denote matrices and vectors, respectively. $(\mathbf{A})^H$, $\lambda_{\max}(\mathbf{A})$, $\mathrm{Tr}(\mathbf{A})$ and $\mathrm{Rank}(\mathbf{A})$ are the Hermitian transpose, maximal eigenvalue, trace and rank of matrix $\mathbf{A}$, respectively. Furthermore, $\mathbf{A} \succeq \mathbf{0}$ denotes that $\mathbf{A}$ is positive semidefinite matrix. In addition, $\mathbb{E}\{b\}$ and b^* stand for the expected value and the optimum of variable b, respectively. Finally, for $\mathbf{a} \in \mathbb{C}^{N\times 1}$, $\max(\mathbf{a})$ and $|a|$ denote the maximal elements and the norm of vector $\mathbf{a}$, respectively.

II. System Model and Problem Formulation

In this section, we consider a multi-user MISO NOMA system as shown in Fig. 1 in which one N-antenna BS transmits to two single-antenna legitimate users in the presence of M energy harvesting single-antenna devices who may potentially be eavesdroppers. For presentational simplicity, we call the legitimate users simply as *users* while energy harvesting devices as *devices* in the sequel. The channel vectors between the BS and the k-th user and the m-th device are denoted respectively, as $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$, $\forall k \in \{1,2\}$ and $\mathbf{g}_m \in \mathbb{C}^{N\times 1}$, $\forall m \in \mathcal{M} = \{1,2,\ldots,M\}$. The transmitted signal from the BS is given by

$$\mathbf{x} = \mathbf{w}_1 x_1 + \mathbf{w}_2 x_2 + \mathbf{v} x_{AN}, \tag{1}$$

where $x_k \in \mathbb{C}$ and $\mathbf{w}_k \in \mathbb{C}^{N\times 1}$ are the desired signal and the beamformer of the k-th user, respectively. Furthermore, $\mathbf{v}$ and x_{AN} stand for the beamformer and AN to prevent information leakage, respectively. The transmitted signal has been normalized, i.e. $\mathbb{E}|x_1|^2 = \mathbb{E}|x_2|^2 = \mathbb{E}|x_{AN}|^2 = 1$. Thus, the received signal of the k-th user can be expressed as

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_0, \tag{2}$$

where $n_0 \sim \mathcal{CN}(0,1)$ is additive white Gaussian noise (AWGN) with zero-mean and unit variance.

Without loss of generality, the first user is identified as the stronger user and the second user is the weaker user, i.e. $|\mathbf{h}_1|^2 > |\mathbf{h}_2|^2$. Following the assumption that the stronger user can perfectly eliminate the interference from the weaker user via SIC in [38], the resulting SINR of the stronger user can be written as

$$\gamma_1 = \frac{\left|\mathbf{h}_1^H \mathbf{w}_1\right|^2}{\left|\mathbf{h}_1^H \mathbf{v}\right|^2 + 1}. \tag{3}$$

In contrast, the weaker user decodes its desired signal by treating the signal designated to the stronger user as interference. As a result, the SINR of the weaker user can be expressed as

$$\tilde{\gamma}_2 = \frac{\left|\mathbf{h}_2^H \mathbf{w}_2\right|^2}{\left|\mathbf{h}_2^H \mathbf{w}_1\right|^2 + \left|\mathbf{h}_2^H \mathbf{v}\right|^2 + 1}. \tag{4}$$

The SINR that the stronger user decodes the weaker user's signal is given by

$$\gamma_{1\to 2} = \frac{\left|\mathbf{h}_1^H \mathbf{w}_2\right|^2}{\left|\mathbf{h}_1^H \mathbf{w}_1\right|^2 + \left|\mathbf{h}_1^H \mathbf{v}\right|^2 + 1}. \tag{5}$$

According to [39], the stronger user in NOMA can completely eliminate the interference from the weaker user if $\gamma_{1\to 2}$ is greater than or equal to $\tilde{\gamma}_2$. Thus, the minimum required SINR for the weaker user is given by

$$\gamma_2 = \min(\tilde{\gamma}_2, \gamma_{1\to 2}). \tag{6}$$

Next, we model the signal received by the energy harvesting devices. The received signal of the m-th passive eavesdropper can be written as

$$y_{e,m} = \mathbf{g}_m^H \mathbf{x} + n_e, \tag{7}$$

where $n_e \sim \mathcal{CN}(0,1)$.

Similar to (5), the m-th device regards the stronger user's signal as interference when trying to decode the signal of the weaker user. Thus, its SINR can be written as

$$\gamma_{e,m\to 2} = \frac{\left|\mathbf{g}_m^H \mathbf{w}_2\right|^2}{\left|\mathbf{g}_m^H \mathbf{w}_1\right|^2 + \left|\mathbf{g}_m^H \mathbf{v}\right|^2 + 1}. \tag{8}$$

In this work, users and all devices are treated indiscriminately, so devices can employ SIC. If the device cannot eliminate the weaker users' signal before eavesdropping the stronger user, the approach of transforming (8) can be adopted. However, similar to [40], [41], we make a pessimistic assumption that the device has successfully decoded the weaker user's signal and wiped out it before attempting to detect the stronger user's signal. This means that the eavesdropper's capability is overestimated. If the enlarged eavesdropping capacity is still smaller than the required threshold, information safety can be ensured. We now examine the SINR of the m-th device attempting to detect the stronger user

$$\gamma_{e,m\to 1} = \frac{\left|\mathbf{g}_m^H \mathbf{w}_1\right|^2}{\left|\mathbf{g}_m^H \mathbf{v}\right|^2 + 1}. \tag{9}$$

We aim to prevent information leakage, so the power splitting scheme[1] is not employed at the devices. The power received by the m-th device can be represented as

$$\hat{P}_{e,m} = \left|\mathbf{g}_m^H \mathbf{w}_1\right|^2 + \left|\mathbf{g}_m^H \mathbf{w}_2\right|^2 + \left|\mathbf{g}_m^H \mathbf{v}\right|^2. \tag{10}$$

Note that the energy harvested from the noise environment is ignored as the energy collected from the AWGN is negligible as compared to that from the wireless signals [43]. In this paper, a practical non-linear energy harvesting model is adopted. According to [44], the harvesting power of devices, denoted by $P_{e,m}$, can be modeled as:

$$P_{e,m} = \frac{\psi_{e,m} - P_{e,m}^{\max}\Psi_{e,m}}{1 - \Psi_{e,m}}, \tag{11}$$

where $\psi_{e,m} = \frac{P_{e,m}^{\max}}{1+\exp[-a_{e,m}(\hat{P}_{e,m}-b_{e,m})]}$ and $\Psi_{e,m} = \frac{1}{1+\exp(a_{e,m}b_{e,m})}$. $a_{e,m}$ and $b_{e,m}$ denote parameters that reflect the circuit specifications, such as the resistance, the capacitance, and diode turn-on voltage. $P_{e,m}^{\max}$ is the maximum harvested power when the energy harvesting circuit is saturated.

Finally, we formulate the transmit power minimization problem while satisfying the QoS requirements of the users, energy harvesting constraints and the minimum secrecy rate constraints. The optimization problem can be formulated as

$$Q_1: \quad \min_{\mathbf{w}_1,\mathbf{w}_2,\mathbf{v}} |\mathbf{w}_1|^2 + |\mathbf{w}_2|^2 + |\mathbf{v}|^2,$$

subject to

$$\begin{aligned}
&C1: \ \gamma_k \geq \Gamma_k, \quad k \in \{1,2\},\\
&C2: \ P_{e,m} \geq \theta_m, \quad m \in \mathcal{M},\\
&C3: \ \log(1+\gamma_k) - \max_{m\in\mathcal{M}} \log(1+\gamma_{e,m\to k}) \geq \rho_k,\\
&\qquad\qquad k \in \{1,2\},
\end{aligned} \tag{12}$$

where $\Gamma_k = 2^{R_k} - 1$ and R_k is the k-th user's rate requirement for $\forall k \in \{1,2\}$, $C1$ denotes the users' rate constraints while $C2$ imposes the minimum energy harvested by the devices. Furthermore, $C3$ sets the minimum secrecy rate constraints for the users. Finally, the constants R_1, R_2, θ_m, ρ_1 and ρ_2 in constraints $C1 - C3$ denote the threshold of the users' rate, energy harvested by the m-th device and the minimum secrecy rate of the users, respectively.

III. Joint Beamforming and AN Optimization Algorithm

Inspection of the optimization problem Q_1 shown in (12) reveals that it is non-convex due to the mutual coupling of beamforming and quadratic constraints. In this section, we propose an efficient D.C. algorithm by employing SDP with rank relaxation to jointly optimize the beamforming and AN covariance matrices. We utilize a penalty function to guarantee that rank-one constraint is satisfied. Fortunately, simulation results indicate that beamforming vectors have been already rank-one, we thus omit the constructed penalty function in simulation setup. Furthermore, we prove the convergence of the proposed D.C. algorithm.

[1]Power splitting scheme splits the received signal with suitable power ratio for energy harvesting and information decoding separately [42].

Through algebraic simplification, the constraint $C2$ can be equivalently recast as

$$\begin{aligned}
\hat{P}_{e,m} &\geq b_{e,m}\\
&\quad - \frac{1}{a_{e,m}} \ln\left\{\frac{P_{e,m}^{\max}}{\theta_m(1-\Psi_{e,m}) + P_{e,m}^{\max}\Psi_{e,m}} - 1\right\} \triangleq \phi_m.
\end{aligned} \tag{13}$$

To proceed, we first introduce three auxiliary variables, namely $\mathbf{W}_k = \mathbf{w}_k\mathbf{w}_k^H$ and $\mathbf{V} = \mathbf{v}\mathbf{v}^H$ for $\forall k \in \{1,2\}$. Defining $\mathbf{H}_k = \mathbf{h}_k\mathbf{h}_k^H$ and $\mathbf{G}_m = \mathbf{g}_m\mathbf{g}_m^H$ for $\forall m \in \mathcal{M}$, we can rewrite Q_1 as

$$Q_2: \quad \min_{\mathbf{W}_1,\mathbf{W}_2,\mathbf{V}} \mathrm{Tr}(\mathbf{W}_1 + \mathbf{W}_2 + \mathbf{V}),$$

subject to

$$\begin{aligned}
&C1: \ \frac{\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)}{\mathrm{Tr}(\mathbf{H}_1\mathbf{V})+1} \geq \Gamma_1,\\
&C2: \ \frac{\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_2)}{\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1)+\mathrm{Tr}(\mathbf{H}_k\mathbf{V})+1} \geq \Gamma_2, k \in \{1,2\},\\
&C3: \ \mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1) + \mathrm{Tr}(\mathbf{G}_m\mathbf{W}_2) + \mathrm{Tr}(\mathbf{G}_m\mathbf{V})\\
&\qquad\qquad \geq \phi_m, m \in \mathcal{M},\\
&C4: \ \log\left(1 + \frac{\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)}{\mathrm{Tr}(\mathbf{H}_1\mathbf{V})+1}\right)\\
&\qquad - \max_{m\in\mathcal{M}} \log\left(1 + \frac{\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)}{\mathrm{Tr}(\mathbf{G}_m\mathbf{V})+1}\right) \geq \rho_1,\\
&C5: \ \log\left(1 + \min_{k\in\{1,2\}} \frac{\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_2)}{\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1)+\mathrm{Tr}(\mathbf{H}_k\mathbf{V})+1}\right)\\
&- \max_{m\in\mathcal{M}} \log\left(1 + \frac{\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_2)}{\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)+\mathrm{Tr}(\mathbf{G}_m\mathbf{V})+1}\right) \geq \rho_2,\\
&C6: \ \mathrm{Rank}(\mathbf{W}_1) = \mathrm{Rank}(\mathbf{W}_2) = \mathrm{Rank}(\mathbf{V}) = 1,
\end{aligned} \tag{14}$$

where $C6$ is imposed to satisfy

$$\mathrm{Rank}\left(\mathbf{W}_k\right) \leq \min\left(\mathrm{Rank}(\mathbf{w}_k), \mathrm{Rank}(\mathbf{w}_k^H)\right). \tag{15}$$

Since $C4$ and $C5$ involve non-convex sets, it is non-trivial to solve Q_2 directly by using any existing solvers. To cope with this problem, we begin with rewriting $C1$ and $C2$ in terms of equivalent convex sets for $k \in \{1,2\}$.

$$\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1) - \Gamma_1\mathrm{Tr}(\mathbf{H}_1\mathbf{V}) - \Gamma_1 \geq 0, \tag{16}$$

$$\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_2) - \Gamma_2\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1) - \Gamma_2\mathrm{Tr}(\mathbf{H}_k\mathbf{V}) - \Gamma_2 \geq 0. \tag{17}$$

We then review in Appendix A the main knowledge about the convex functions/sets that will be used. Next, we remove the deleterious effects of maximum operator included in $C4$. We introduce a new auxiliary variable t while setting that SINR of all eavesdroppers who attempt to decode the stronger user's signal is smaller than or equal to $(t-1)$. Thus, its SINR can be

rewritten as

$$\frac{\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)}{\mathrm{Tr}(\mathbf{G}_m\mathbf{V})+1} \le t-1, \tag{18}$$

for $\forall m \in \mathcal{M}$.

Adopting *Case 1* in Appendix A and introducing auxiliary variable δ_m^2 for $\forall m \in \mathcal{M}$, (18) is decomposed into two constraints

$$\begin{cases} \dfrac{\delta_m^2}{t} \le \mathrm{Tr}(\mathbf{G}_m\mathbf{V}), \\ \mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)+\mathrm{Tr}(\mathbf{G}_m\mathbf{V})-t+1 \le \delta_m^{(n)}\left(2\delta_m-\delta_m^{(n)}\right), \end{cases} \tag{19}$$

for $\forall m \in \mathcal{M}$.

After introducing the auxiliary variable t, $C4$ can be rewritten as:

$$\log\left(\frac{\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1+\mathbf{H}_1\mathbf{V})+1}{t\mathrm{Tr}(\mathbf{H}_1\mathbf{V})+t}\right) \ge \rho_1. \tag{20}$$

The form is the same as *Case 2* in Appendix A. By introducing an auxiliary variable μ^2, we can derive

$$\begin{cases} \mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1+\mathbf{H}_1\mathbf{V})+1-2^{\rho_1}t \ge \mu^2, \\ \dfrac{2\mu^{(n)}\mu}{t^{(n)}} - \dfrac{{\mu^{(n)}}^2 t}{{t^{(n)}}^2} \ge 2^{\rho_1}\mathrm{Tr}(\mathbf{H}_1\mathbf{V}). \end{cases} \tag{21}$$

We first remove the minimum operator in $C5$ and have

$$\begin{aligned} &\log\left[1+\frac{\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_2)}{\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1)+\mathrm{Tr}(\mathbf{H}_k\mathbf{V})+1}\right] \\ &-\log\left[1+\max_{m\in\mathcal{M}}\frac{\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_2)}{\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)+\mathrm{Tr}(\mathbf{G}_m\mathbf{V})+1}\right] \ge \rho_2, \end{aligned} \tag{22}$$

for $\forall k \in \{1,2\}$.

After that, the auxiliary variables λ, ψ_m, and β_k are utilized for $\forall m \in \mathcal{M}$ and $\forall k \in \{1,2\}$. Adopting a similar approach in handling $C4$, the constraint $C5$ can be divided into the following four constraints,

$$\begin{cases} \dfrac{\psi_m^2}{(\lambda-1)} \le \mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)+\mathrm{Tr}(\mathbf{G}_m\mathbf{V})+1, \\ \mathrm{Tr}(\mathbf{G}_m\mathbf{W}_2) \le \psi_m^{(n)}\left(2\psi_m-\psi_m^{(n)}\right), \\ \beta_k^2 \le \mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1+\mathbf{H}_k\mathbf{W}_2+\mathbf{H}_k\mathbf{V})+1-2^{\rho_2}\lambda, \\ \dfrac{2\beta_k^{(n)}\beta_k}{\lambda^{(n)}} - \dfrac{{\beta_k^{(n)}}^2\lambda}{{\lambda^{(n)}}^2} \ge 2^{\rho_2}\mathrm{Tr}\left(\mathbf{H}_k\left(\mathbf{W}_1+\mathbf{V}\right)\right), \end{cases} \tag{23}$$

for $\forall m \in \mathcal{M}$ and $\forall k \in \{1,2\}$.

So far, in addition to the rank-one constraint of beamforming matrices and AN covariance matrix in Q_2, the all constraints have been transformed into convex sets. Based on the above simplification and relaxing the rank-one constraints, Q_2 can be reformulated as

$$\begin{aligned} Q_3\text{:} \quad & \min_{\mathbf{W}_1,\mathbf{W}_2,\mathbf{V},\mathbf{t},\boldsymbol{\mu},\boldsymbol{\delta},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\psi}} \mathrm{Tr}(\mathbf{W}_1+\mathbf{W}_2+\mathbf{V}) \\ \text{subject to :} \quad & (16),(17),(19),(21),(23),C3, \\ & \mathbf{W}_1 \succeq \mathbf{0}, \mathbf{W}_2 \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}, \end{aligned} \tag{24}$$

Algorithm 1: D.C. Algorithm.

1: **Initialize:** Set up variables initial value $t^{(1)}$, $\mu^{(1)}$, $\delta_m^{(1)}$, $\lambda^{(1)}$, $\beta_k^{(1)}$, $\psi_m^{(1)}$, $m \in \mathcal{M}$, $k \in \{1,2\}$, make the iteration index $n=1$, and set the maximum tolerance ξ.
2: **while** not convergence **do**
3: Solve Q_3 by employing CVX in MATLAB.
4: Update $t^{(n+1)}=t^{(*)}$, $\mu^{(n+1)}=\mu^{(*)}$, $\delta_m^{(n+1)}=\delta_m^{(*)}$, $\lambda^{(n+1)}=\lambda^{(*)}$, $\beta_k^{(n+1)}=\beta_k^{(*)}$, $\psi_m^{(n+1)}=\psi_m^{(*)}$.
5: Update iteration index $n=n+1$.
6: **end while**

where the optimization variables $\boldsymbol{\delta}=(\delta_1,\delta_2,\ldots,\delta_M)$, $\boldsymbol{\beta}=(\beta_1,\beta_2)$, and $\boldsymbol{\psi}=(\psi_1,\psi_2,\ldots,\psi_M)$.

It is worth noting that problem Q_3 is convex and can be solved directly by using any standard solvers, such as CVX [45]. The proposed iterative D.C. algorithm for transmit power minimization is summarized in Algorithm 1.

However, it is analytically intractable to show that problem Q_2 and problem Q_3 are equivalent because the rank-one constraints imposed on the beamforming and AN have been relaxed. However, even if the rank of optimal beamforming matrices are not one, we can employ penalty function to find rank-one optimal beamforming matrices as follows: If $\mathbf{A}$ is a semidefinite matrix, we have $\mathrm{Tr}(\mathbf{A}) \ge \lambda_{\max}(\mathbf{A})$, so we can deduce that $\mathbf{A}$ is rank-one if we can guarantee $\mathrm{Tr}(\mathbf{A}) \le \lambda_{\max}(\mathbf{A})$. In other words, if $\mathrm{Rank}(\mathbf{W}_k)=\mathrm{Rank}(\mathbf{V})=1$, $\mathbf{W}_k$ and $\mathbf{V}$ have only one non-zero eigenvalue, respectively. Thus, the constraint of rank-one can be transformed as

$$\sum_{k=1}^{2}\left[\mathrm{Tr}(\mathbf{W}_k)-\lambda_{\max}(\mathbf{W}_k)\right]+\mathrm{Tr}(\mathbf{V})-\lambda_{\max}(\mathbf{V}) \le 0, \tag{25}$$

for $\forall k \in \{1,2\}$. Note that the function $\lambda_{\max}(\mathbf{W}_k)$ is convex on the set of Hermitian matrices. Therefore, we aim to minimize the following formula

$$\begin{aligned} &\mathrm{Tr}(\mathbf{W}_1+\mathbf{W}_2+\mathbf{V}) \\ &\quad+\kappa\left(\sum_{k=1}^{2}\left[\mathrm{Tr}(\mathbf{W}_k)-\lambda_{\max}(\mathbf{W}_k)\right]\mathrm{Tr}(\mathbf{V})-\lambda_{\max}(\mathbf{V})\right), \end{aligned} \tag{26}$$

where $\kappa>0$ is a constant. If the penalty factor κ is chosen to be large enough, then the difference between $\mathrm{Tr}(\mathbf{W}_1+\mathbf{W}_2+\mathbf{V})$ and $\lambda_{\max}(\mathbf{W}_1+\mathbf{W}_2+\mathbf{V})$ can be minimized. Otherwise, even $\sum_{k=1}^{2}[\mathrm{Tr}(\mathbf{W}_k)-\lambda_{\max}(\mathbf{W}_k)]+\mathrm{Tr}(\mathbf{V})-\lambda_{\max}(\mathbf{V})$ is small, the system power consumption is over-high because the penalty factor is large enough. However, simulation results indicate rank-one optimal beamforming matrices always exist, so penalty function in simulation is omitted.

At present, the proof of convergence of the D.C. algorithm is suitable only if the objective function is a D.C. function [46], [47]. However, in this paper, the constraints is D.C. functions, which can not be solved by existing work about the D.C. algorithm convergence. In the following theorem, we develop the general proof about the convergence of D.C. algorithm. This

proof approach can be used to the optimization problems with D.C. constraints or D.C. objective function.

Theorem 1: The D.C. algorithm can converge to a local optimal value after finite iterations.

Proof: See Appendix B. ■

Complexity: D.C. algorithm in problem Q_3 involves three linear matrix inequalities (LMIs) constraints of sizes N and $5M+9$ linear constraints. The number of variables is $(3N^2+2M+5)$. Thus, the complexity order is $\mathcal{O}(n\alpha_1\sqrt{3N+5M+9}(3N^3+n(3N^2+5M+9)+n^2)$, where $n=\mathcal{O}(3N^2+2M+5)$ and α_1 denotes iteration number.

IV. A Sub-Optimal Alternating D.C. Algorithm

In this section, in order to reduce the computation cycles of D.C. algorithm, we propose a sub-optimal alternating D.C. algorithm. To ensure the information security and not excessively damage the users' transmission capacity simultaneously, the AN is mapped to the null space of the users' channel vectors. We then decouple the AN covariance matrix and the beamforming matrices of users by alternating optimization.

The AN vector is based on the users' channel information denoted by $\mathbf{H}_{N\times 2}=[\mathbf{h}_1,\mathbf{h}_2]$. Let $\tilde{\mathbf{V}}\in\mathbb{C}^{N\times(N-2)}$ denote the null space of $\mathbf{H}$, i.e. $\mathbf{H}^H\tilde{\mathbf{V}}=\mathbf{0}$ and $|\tilde{\mathbf{V}}|^2=1$. Owing to the AN vector has been mapped to the null space of the users' channel information, the AN vector can be expressed as

$$\mathbf{v}=\tilde{\mathbf{V}}\mathbf{u} \tag{27}$$

where $\mathbf{u}\in\mathbb{C}^{(N-2)\times 1}$ is the vector for searching the optimal null space of $\mathbf{H}$, and $\mathrm{Tr}(\tilde{\mathbf{V}}\mathbf{u}\mathbf{u}^H\tilde{\mathbf{V}}^H)$ denotes the AN power consumption. In order to simplify expression, we set $\tilde{\mathbf{g}}_m^H=\mathbf{g}_m^H\tilde{\mathbf{V}}$, $\tilde{\mathbf{G}}_m=\tilde{\mathbf{g}}_m\tilde{\mathbf{g}}_m^H$ and $\mathbf{U}=\mathbf{u}\mathbf{u}^H$ for $\forall m\in\mathcal{M}$. After using $\mathrm{Tr}(\mathbf{H}_k\mathbf{V})=0$ and giving the fixed beamforming matrices of the users, the constraint $C3$ in Q_2 can be transformed into

$$\mathrm{Tr}(\tilde{\mathbf{G}}_m\mathbf{U})\geq\frac{\phi_m}{\eta_m}-\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)-\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_2)\triangleq\tau_m \tag{28}$$

for $\forall m\in\mathcal{M}$.

Similar to (28), the constraints $C4$ and $C5$ in problem Q_2 can respectively be expressed as

$$\begin{cases}\mathrm{Tr}(\tilde{\mathbf{G}}_m\mathbf{U})\geq\dfrac{\mathrm{Tr}(G_m\mathbf{W}_1)}{2^{-\rho_1}\left(1+\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)\right)-1}-1\triangleq\eta_m,\\ \mathrm{Tr}(\tilde{\mathbf{G}}_m\mathbf{U})\geq\dfrac{\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_2)}{2^{-\rho_2}\left(1+\frac{\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_2)}{\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1)+1}\right)-1}-1\\ \qquad\qquad-\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)\triangleq\varphi_{m,k},\end{cases} \tag{29}$$

for $\forall m\in\mathcal{M}$ and $\forall k\in\{1,2\}$.

Summarizing (28) and (29), we have

$$\mathrm{Tr}(\tilde{\mathbf{G}}_m\mathbf{U})\geq\max(\tau_m,\eta_m,\varphi_{m,1},\varphi_{m,2})\triangleq\omega_m, \tag{30}$$

for $\forall m\in\mathcal{M}$.

Given $\mathbf{W}_1$ and $\mathbf{W}_2$ that can satisfy constraints $C1$ and $C2$, the system power consumption is determined by the AN. We divide the power consumption of AN into two categories according to the value of $\boldsymbol{\omega}$, where $\boldsymbol{\omega}$ is a vector composed of ω_m. If $\max(\boldsymbol{\omega})\leq 0$, users' information is well-protected without being eavesdropped. As a result, no AN is required, i.e. $\mathbf{U}=\mathbf{0}$. However, if $\max(\boldsymbol{\omega})>0$, the given precoding matrices cannot guarantee that information of the users is being protected.

Driven by reducing the power consumption brought by AN covariance matrix, we thus aim to search the users' optimal null space. This problem can be efficiently solved by maximizing the worst-device supporting proportion[2] at the unit supplied power. The optimization problem can be formulated as

$$\begin{aligned}Q_4:\quad&\max_{\mathbf{U}}\min_{i\in\mathcal{I}}\frac{\mathrm{Tr}(\tilde{\mathbf{G}}_i\mathbf{U})}{\varpi_i}\\ \text{subject to}:\quad&C1:\ \mathrm{Tr}(\tilde{\mathbf{V}}\mathbf{U}\tilde{\mathbf{V}}^H)\leq 1,\\ &C2:\ \mathbf{U}\succeq\mathbf{0},\\ &C3:\ \mathrm{Rank}(\mathbf{U})=1,\end{aligned} \tag{31}$$

where $\mathcal{I}$ is the device index set with $\omega_m>0$ and $\boldsymbol{\varpi}$ is a vector composed of positive elements in $\boldsymbol{\omega}$ with ϖ_i for $i\in\mathcal{I}$ being the elements in $\boldsymbol{\varpi}$. $C1$ denotes the unit power constraint while $\tilde{\mathbf{G}}_i$ is a new channel matrix corresponding to $\boldsymbol{\varpi}$, $C3$ is imposed to guarantee that $\mathbf{U}$ can be decomposed into an one-dimensional vector.

Lemma 1: The optimal value of Q_4 is achieved when $\mathrm{Tr}(\tilde{\mathbf{V}}\mathbf{U}\tilde{\mathbf{V}}^H)=1$.

Proof: We prove Lemma 1 by contradiction. Given $C2$ and $C3$, we assume that the beamforming reaches the optimal value at $\tilde{\mathbf{U}}$ with $\mathrm{Tr}(\tilde{\mathbf{V}}\tilde{\mathbf{U}}\tilde{\mathbf{V}}^H)<1$. Then, we can find a value denoted by $\delta>0$ to satisfy $\frac{(1+\delta)\mathrm{Tr}(\tilde{\mathbf{G}}_i\tilde{\mathbf{U}})}{\varpi_i}>\frac{\mathrm{Tr}(\tilde{\mathbf{G}}_i\tilde{\mathbf{U}})}{\varpi_i}$ and other constraints, which contradicts with the assumption that Q_4 reaches its maximum at the $\tilde{\mathbf{U}}$. ■

In order to decouple the max-min operator, we equivalently transform Q_4 into the following problem

$$\begin{aligned}Q_5:\quad&\max_{\mathbf{U},\Gamma}\Gamma\\ \text{subject to}:\quad&C1-C3,\\ &C4:\ \frac{\mathrm{Tr}(\tilde{\mathbf{G}}_i\mathbf{U})}{\varpi_i}\leq\Gamma,\quad\forall i\in\mathcal{I}.\end{aligned} \tag{32}$$

The constraints $C1-C3$ are the same as constraints in problem Q_4. Note that Q_5 is convex except $C3$. Unfortunately, $C3$ is essential to guarantee the equivalence between Q_5 and Q_2. Next, we prove the existence of the optimal solution with rank-one by using the following theorem.

Theorem 2: Q_5 always has an optimal solution with $\mathrm{Rank}(\mathbf{U})=1$.

Proof: See Appendix C. ■

As mentioned, we can get the users' optimal null space with the given beamforming matrices. Then, we calculate the optimal beamforming matrices based on the current AN covariance

[2]The supporting proportion is the ratio between the provided trace and the required trace. The inverse of the proportions supported at the unit power is the power required by AN covariance matrix. However, the power is decided by the worst device, so we investigate the maximum of worst-device.

matrix. The optimization problem and the constraints are the same as the problem Q_2 except $\mathbf{V}$ is a constant matrix.

Similar to (16) and (17), the constraints $C1$ and $C2$ in Q_2 can be recast as

$$\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1) \geq \Gamma_1, \tag{33}$$

$$\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_2) - \Gamma_2\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1) - \Gamma_2 \geq 0. \tag{34}$$

The inequalities (33) and (34) are established because $\mathrm{Tr}(\mathbf{H}_k\mathbf{V}) = 0$ for $\forall k \in \{1, 2\}$.

Similar to $C4$ in D.C. programming, we introduce the auxiliary variable t, and define that

$$\frac{\mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1)}{\mathrm{Tr}(\tilde{\mathbf{G}}_m\mathbf{U}) + 1} \leq t - 1, \quad \forall m \in \mathcal{M}, \tag{35}$$

which can define a convex set. After introducing the auxiliary variable t, the remaining constraint of $C4$ can be rewritten as

$$\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1) + 1 \geq 2^{\rho_1}t. \tag{36}$$

Both sides of the inequality (36) are affine function that can define a convex set.

Adopting a similar approach in handling $C5$ in problem Q_2, we also introduce the auxiliary variable λ, ψ_m and β_k for $\forall m \in \mathcal{M}$ and $\forall k \in \{1, 2\}$. After using $\mathrm{Tr}(\mathbf{H}_k\mathbf{V}) = 0$, $C5$ can be decomposed into

$$\begin{cases} \dfrac{\psi_m^2}{(\lambda - 1)} \leq \mathrm{Tr}(\mathbf{G}_m\mathbf{W}_1) + \mathrm{Tr}(\tilde{\mathbf{G}}_m\mathbf{U}) + 1, \\ \mathrm{Tr}(\mathbf{G}_m\mathbf{W}_2) \leq \psi_m^{(n)}\left(2\psi_m - \psi_m^{(n)}\right), \\ \beta_k^2 \leq \mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1 + \mathbf{H}_k\mathbf{W}_2) + 1 - 2^{\rho_2}\lambda, \\ \dfrac{2\beta_k^{(n)}\beta_k}{\lambda^{(n)}} - \dfrac{{\beta_k^{(n)}}^2\lambda}{{\lambda^{(n)}}^2} \geq 2^{\rho_2}\mathrm{Tr}(\mathbf{H}_k\mathbf{W}_1). \end{cases} \tag{37}$$

Through the above simplification, all constraints have been transformed into convex sets. As a result, the original problem, neglecting rank-one constraint, can be recast as

$$\begin{aligned} Q_6\text{:} \min_{\mathbf{W}_1, \mathbf{W}_2, \mathbf{t}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\psi}} \ & \mathrm{Tr}(\mathbf{W}_1 + \mathbf{W}_2 + \tilde{\mathbf{V}}\mathbf{U}) \\ \text{subject to :} \ & C3, (33), (34), (35), (36), (37), \\ & \mathbf{W}_1 \succeq \mathbf{0}, \mathbf{W}_2 \succeq \mathbf{0}. \end{aligned} \tag{38}$$

So far, the problem Q_6 can be solved directly by CVX. In order to make the description about the sub-optimal alternating D.C. algorithm more readable, the specific process of sub-optimal ADC algorithm is relegated to Algorithm 2. Problem Q_6 cannot guarantee the equivalence with the problem Q_2, but we can deduce the existence of the optimal solution with rank-one from simulation results.

Since we use alternating optimization to find the sub-optimal beamforming, the convergence proof mentioned above is not suitable to this algorithm. We therefore develop the next theorem to verify the convergence.

Theorem 3: The sub-optimal ADC algorithm can converge to a sub-optimal value after finite iterations.

Algorithm 2: Sub-Optimal Alternating D.C. Algorithm (ADC).

1: **Initialize:** Set initial AN covariance matrix $\mathbf{U}^{(1)}$, set the iteration index $q = 1$, $n = 1$ and the maximum tolerance ξ and τ.
2: Set variables initial value $\lambda^{(1)}$, $\beta_k^{(1)}$, $\psi_m^{(1)}$, for $\forall m \in \mathcal{M}$ and $\forall k \in \{1, 2\}$.
3: **while** not convergence **do**
4: **while** not convergence **do**
5: Solve Q_6 by employing CVX without considering $\mathrm{Rank}(\mathbf{U}) = 1$.
6: Update $\psi_m^{(q+1)} = \psi_m^{(*)}$, $\lambda^{(q+1)} = \lambda^{(*)}$, $\beta_k^{(q+1)} = \beta_k^{(*)}$.
7: Update iteration index $q = q + 1$.
8: Repeat 5–7 until convergence.
9: Output ${\mathbf{W}_1}^{(*)}$, ${\mathbf{W}_2}^{(*)}$.
10: **end while**
11: Update ${\mathbf{W}_1}^{(n)} = {\mathbf{W}_1}^{(*)}$, ${\mathbf{W}_2}^{(n)} = {\mathbf{W}_2}^{(*)}$.
12: **if** $\max(\boldsymbol{\omega}) \leq 0$ **then**
13: break;
14: **else**
15: Solve Q_5 by employing CVX.
16: **end if**
17: Update iteration index $n = n + 1$.
18: Update iteration index $\mathbf{U}^{(n)} = \mathbf{U}^{(*)}$.
19: Update $\lambda^{(n)}$, $\beta_k^{(n)}$, $\psi_m^{(n)}$, for $\forall m \in \mathcal{M}$ and $\forall k \in \{1, 2\}$ according to (58)
20: **end while**

Proof: See Appendix D. ■

Complexity: In ADC, the original non-convex optimization problem is decomposed into two convex optimization subproblems. The outer problem Q_5 consists of one LMI constraint of size $N - 2$ and $I + 1$ linear constraints. The number of variables is $((N-2)^2 + 1)$. Therefore, the complexity order is $\mathcal{O}(n\alpha_2\sqrt{N + I - 1}((N-2)^3 + n((N-2)^2 + 5I + 1) + n^2))$, where $n = \mathcal{O}((N-2)^2 + 1)$ and α_2 denotes iteration index. The inter problem Q_6 contains two LMI constraints of size N and $4M + 8$ linear constraints. The number of variables is $(2N^2 + M + 4)$. Therefore, the complexity order is $\mathcal{O}(n\alpha_2\alpha_3\sqrt{2N + 4M + 8}(2N^3 + n(2N^2 + 4M + 8) + n^2))$, where $n = \mathcal{O}(2N^2 + M + 4)$ and α_3 denotes iteration number.

V. Simulation Results

In this section, the performance of the two proposed algorithms, namely D.C. and sub-optimal ADC algorithm, are evaluated by computer simulations. In the following simulation, the results are obtained by averaging over 300 Rayleigh-distributed channel realizations. The circuit parameters are referenced in [36]. Without loss of generality, all of eavesdroppers have the same energy requirements. In this paper, $\phi_m = 7$dBm, which means that the harvested energy is saturated. In ADC, we

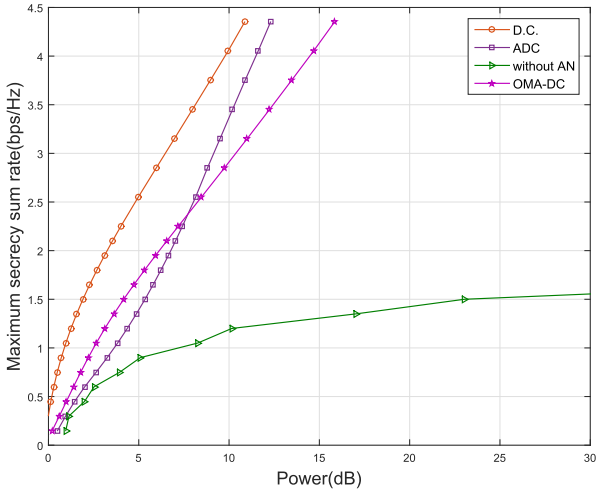


Fig. 2. Achievable secrecy rate comparison of different algorithms when $M \geq N$.

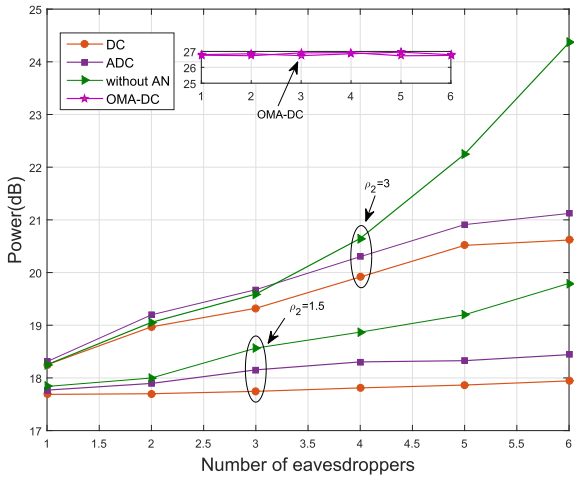


Fig. 3. Minimum transmit power versus the number of devices with different algorithms.

set the initial power of AN covariance matrix at $10M$ mW.[3] In order to test the effectiveness of our proposed algorithms, they are compared with the conventional NOMA *without* AN and the OMA based on D.C. (OMA-DC) algorithm *with* AN covariance matrix.

- *conventional NOMA scheme without AN:* In this algorithm, all processes except AN covariance matrix are the same as our proposed algorithms. We aim to accurately determine the effect of AN injection.
- *OMA-DC scheme:* In OMA-DC scheme with AN covariance matrix, the users are allocated to the equal slot. Devices can only eavesdrop the information at corresponding halve slot and harvest energy at the whole time slot.

In simulation results, we design a reasonable metric, which can simulate the maximum supported security rate when AN is not injected and given power is unlimited. It can also indicate the performance gap between injecting AN or not, as shown in Fig. 2.

At present, an intuitive idea is that beamforming design without AN-aided cannot ensure absolute information secrecy. However, given unlimited transmit power, the maximum supported security rate when AN is not injected is hard to calculate or simulate. To accurately determine the role of AN injection, we design a reasonable metric. The goal of the metric is to guarantee that system power cost can be expressed even if it is infinite, which is

$$f(P) = \upsilon P_r + (1-\upsilon)\frac{\Upsilon}{\upsilon^a + \delta}, \tag{39}$$

[3]To prevent information leakage and the power required by AN covariance matrix at (n)-th iteration is less than or equal to the last iteration, so the initial power given to AN should be large enough. However, because the ADC algorithm only find the sub-optimal value, in order to prevent the algorithm has converged at the large value, the power should be set as small as possible. Considering the trade-off between the information leakage and power consumption, we adopt the linear relationship between initial power of AN and the number of eavesdroppers.

where υ is the probability that conventional NOMA can satisfy the minimum secrecy rate with finite power cost, P_r is the real power consumption when the conventional NOMA scheme can meet the QoS. Υ, a and δ are constants. $\upsilon^a + \delta$ and Υ denote the penalty factor and the power consumption that the QoS of secrecy rate can not be satisfied. In this paper, $\Upsilon = 90$dB, $a = 6$ and $\delta = 0.01$.

By adopting (39), we simulate transmit power versus maximum secrecy sum rate between four algorithms for $M = 8$, $N = 5$, and $R_1 = R_2 = 0$, as shown in Fig. 2. In Fig. 2, the most important observation is that the optimum secrecy sum rate converges to a precise value if AN is not injected. The proposed algorithms can achieve any secrecy rate if the system is provided sufficient power. We can observe that OMA-DC algorithm can achieve higher secrecy sum rate than ADC algorithm under low transmit power. This may be caused by the following factors. Compared to OMA, NOMA can significantly boost the transmission rate at high SNR, but the advantage of NOMA is not particularly obvious at low SNR. In ADC algorithm, AN is mapped to null space of legitimate users, leading to the loss of solution space. However, with the increase of transmit power, the advantage of NOMA transmission scheme makes up for this drawback.

In Fig. 3, we present minimum transmit power comparison for different values of ρ_2 with $N = 8$, $R_1 = 6$, $R_2 = 3$, and $\rho_1 = 3$. We can observe that the performance of proposed sub-optimal ADC algorithm is very comparable to that of D.C. algorithm. AN is almost unnecessary under a small number of eavesdroppers, but the initial power is relatively high. This leads to that performance of ADC is exceeded by the conventional NOMA scheme because ADC is a sub-optimum algorithm. With the increase of devices, we can also find that the performance gap between our proposed algorithms and the conventional NOMA scheme become larger. The power cost of OMA scheme is almost unchanged, probably because QoS of the stronger user is the main constraint while the secrecy rate constraint has been relaxed. Simulation result shows that our proposed two

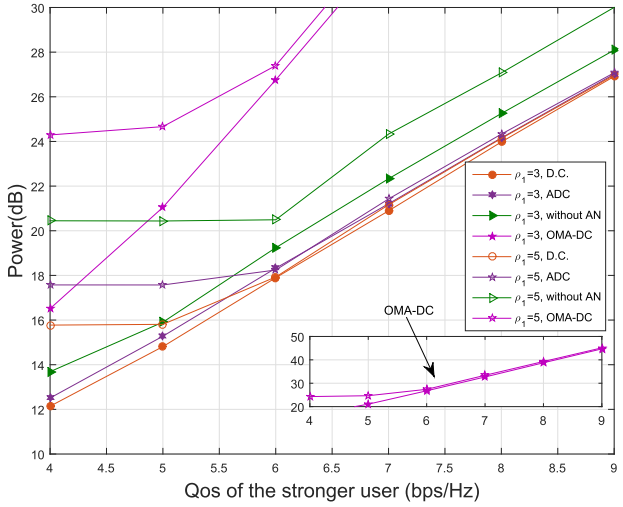


Fig. 4. Minimum transmit power versus QoS of the stronger user with different algorithms.

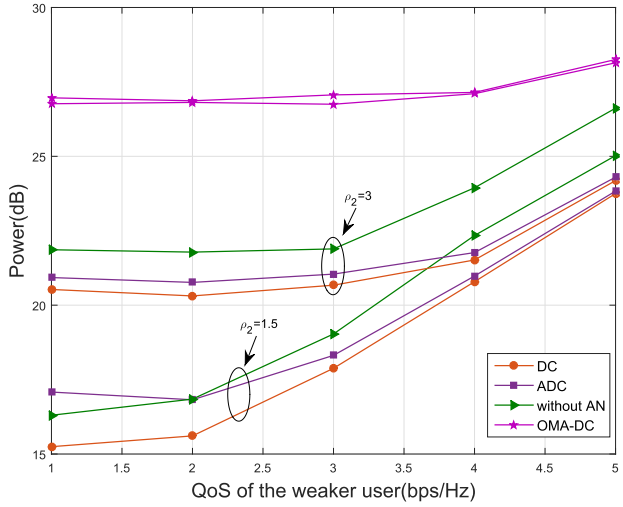


Fig. 5. Minimum transmit power versus QoS of the weaker user with different algorithms.

algorithms outperform the NOMA scheme *without* AN and OMA-DC scheme, especially when there exist multi-devices.

Fig. 4 depicts minimum transmit power versus QoS of the stronger user with different algorithms with $N = 8$, $M = 5$, $R_2 = 3$, and $\rho_2 = 1.5$. It can be seen that the power consumption remains unchanging when the range of the stronger user's QoS from 4 bps/Hz to 5 bps/Hz at $\rho_1 = 5$. Because the secrecy rate constraint is bigger than QoS of the stronger user, the constraint of the stronger user's rate has been relaxed. With the increase of minimum rate requirement of the stronger user, we observe that the performance differences among the three algorithms (except OMA-DC scheme) are decreasing. Meanwhile, the system power cost gap between the same algorithm under different secrecy rates is reducing. The reason is that constraint of secrecy rate is gradually relaxed, which means that the effect of injecting AN is gradually fading. However, our proposed two algorithms are always superior to the two baseline schemes because D.C. and ADC algorithms can search at least the optimum of NOMA without AN.

In Fig. 5, we simulate the minimum transmit power versus the QoS of the weaker user in terms of the different ρ_2 with $N = 8$, $M = 5$, $R_1 = 6$, and $\rho_1 = 3$. Similar to Fig. 4, the system transmit power consumption also remains a static value when minimum secrecy rate is greater than the weaker user's QoS. Because the optimum null space can be found in sub-optimal ADC algorithm and D.C. is also a local-optimum solution, its performance can be closely related to that of D.C. algorithm. Compared to D.C. algorithm and the conventional NOMA scheme, the transmit power consumption of sub-optimal ADC algorithm is higher when $\rho_2 = 1.5$ and the weaker user's QoS from 1 to 2. This is because the initial power of AN is relatively high, which leads to the algorithm has converged at the higher power cost.

Fig. 6 depicts the transmit power minimization with respect to different number of BS antennas and minimum secrecy rate of the weaker user with $M = 5$, $R_1 = 6$, $R_2 = 3$, and $\rho_1 = 3$.

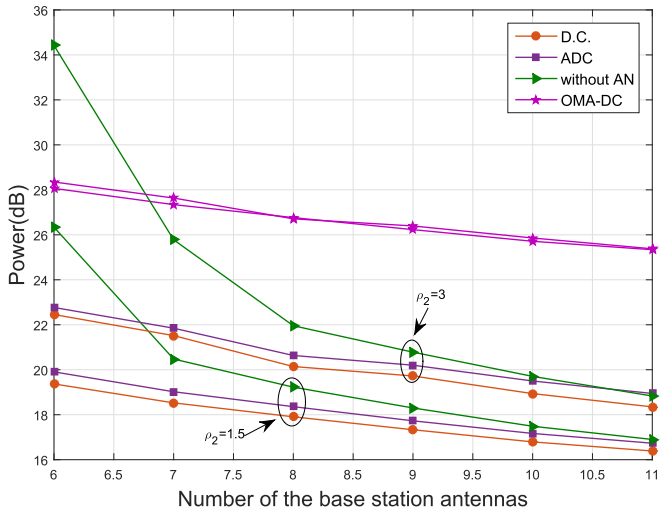


Fig. 6. Minimum transmit power versus the number of BS antennas with different algorithms.

From Fig. 6, we can see that the transmit power consumption of all algorithms will decreases with the increase of number of base station antenna. This is because the BS with more antennas can provide the higher multipling gain, which is very useful to improve transmission rate under the same transmit power. Besides, we can find that the sub-optimal ADC algorithm and the conventional NOMA without AN can approximate the performance of D.C. algorithm. The reason for this phenomenon is twofold. Firstly, because information security transmission can be satisfied nearly by designing beamforming matrices with the increase of users' beamforming dimension. Secondly, because the users' null space dimension increases, the linear combinations of null space vector can approximately express all beamforming vectors.

In Table I, we compare the computation cycles between four different algorithms with $N = 8$, $R_1 = 6$, $R_2 = 3$, and

TABLE I
COMPUTATION CYCLES COMPARISON BETWEEN FOUR DIFFERENT ALGORITHMS

(a) $\rho_2 = 1.5$

# of devices	1	2	3	4	5	6
D.C.	3.82	4.97	6.33	7.76	9.23	10.88
ADC	3.62	4.00	4.40	4.77	5.16	5.35
without AN	2.31	2.77	3.43	4.23	5.39	6.77
OMA-DC	1.68	2.10	2.65	3.26	3.82	4.49

(b) $\rho_2 = 3$

# of devices	1	2	3	4	5	6
D.C.	6.79	9.35	11.48	13.39	15.22	17.12
ADC	5.43	6.33	7.11	7.84	8.55	9.21
without AN	4.07	4.96	6.16	7.60	9.12	10.86
OMA-DC	3.77	4.67	5.53	6.59	7.38	8.35

$\rho_1 = 3$. We can observe that ADC always outperforms the D.C. algorithm because ADC contains fewer variables and the optimization of AN is a convex optimization problem. Beside, ADC is better than conventional NOMA scheme in computation cycles when there are multiple eavesdroppers in the system. The OMA-DC algorithm has the shortest computation cycle while its performance is the worst.

VI. CONCLUSION

In this paper, with the goal of power minimization, a multi-user MISO NOMA system involving energy harvesting and physical layer security is investigated. To establish secure data-link and meanwhile charge energy harvesting devices, the system is subject to the users' QoS, energy harvesting requirement and minimum secure rate. By jointly designing the beamforming and AN, we propose D.C. algorithm to solve the non-convex optimization problem by employing SCP and rank-one relaxation. Secondly, to reduce the computation cycles of D.C. algorithm, the AN covariance matrix is mapped to the null space of the users. Then, the sub-optimal ADC algorithm is proposed. Furthermore, the convergence of the proposed D.C. and ADC algorithms is derived. The proof method is suit to all optimization problems with D.C. constraints or D.C. objective function. From our simulation results, it indicates that, given unlimited transmit power, there exists a secrecy rate saturation for the scheme without AN. On the other hand, how high the required secrecy rate is, a transmission scheme with AN can be found to satisfy the requirement. Besides, the efficacy and superiority of our proposed algorithms are verified.

APPENDIX A
REVIEW OF CONVEX FUNCTIONS/SETS

Let $\mathbf{W}$ and $\mathbf{V} \in \mathbb{C}^{N\times N}$ denote the semidefinite variable matrices and t be a nonnegative variable. We consider the following two cases.

Case 1: A domain is defined by a fractional inequality, as follows:

$$\frac{\mathrm{Tr}(\mathbf{W})}{\mathrm{Tr}(\mathbf{V})+1} \le t-1. \tag{40}$$

Similar to $C1$ and $C2$ in Q_2, (40) can be also rewritten as

$$\mathrm{Tr}(\mathbf{W}) + \mathrm{Tr}(\mathbf{V}) - t + 1 \le t\mathrm{Tr}(\mathbf{V}). \tag{41}$$

However, (41) remains to be non-convex due to the coupling between t and $\mathrm{Tr}(\mathbf{V})$. To circumvent this problem, an auxiliary variable δ^2 is introduced to decompose (41) into two constraints,

$$\frac{\delta^2}{t} \le \mathrm{Tr}(\mathbf{V}), \tag{42}$$

and

$$\mathrm{Tr}(\mathbf{W}) + \mathrm{Tr}(\mathbf{V}) - t + 1 \le \delta^2. \tag{43}$$

Note that the functions in both sides of (43) are convex function, which drives us to develop an approximate function by using first-order Taylor expansion. We propose to replace δ^2 with its first-order Taylor expansion around a fixed point $\delta^{(n)}$. Then, (43) becomes

$$\mathrm{Tr}(\mathbf{W}) + \mathrm{Tr}(\mathbf{V}) - t + 1 \le \delta^{(n)}\left(2\delta - \delta^{(n)}\right). \tag{44}$$

The transformation from (43) to (44) is effective because the first-order Taylor expansion of δ^2 narrows the domain defined by (43). This means that the transmit power minimization we are concerned with is solved at the sub-region of original domain, so the solution can satisfy the system requirement.

Case 2: The constraint can be written as follows,

$$\log\left(\frac{\mathrm{Tr}(\mathbf{W})+1}{t\mathrm{Tr}(\mathbf{V})+t}\right) \ge \rho_1, \tag{45}$$

where ρ_1 is a constant.

Then, (45) can be further transformed as

$$\mathrm{Tr}(\mathbf{W}) + 1 - 2^{\rho_1}t \ge 2^{\rho_1}t\mathrm{Tr}(\mathbf{V}). \tag{46}$$

To decouple t and $\mathrm{Tr}(\mathbf{V})$, an auxiliary variable μ^2 is introduced to decompose (46) as

$$\mathrm{Tr}(\mathbf{W}) + 1 - 2^{\rho_1}t \ge \mu^2, \tag{47}$$

and

$$\frac{\mu^2}{t} \ge 2^{\rho_1}\mathrm{Tr}(\mathbf{V}). \tag{48}$$

(48) remains a non-convex set. However, the functions in its both sides are convex function [48]. We therefore rewrite (48) by replacing $\frac{\mu^2}{t}$ by its first-order Taylor expansion at the fixed point $(\mu^{(n)}, t^{(n)})$,

$$\frac{2\mu^{(n)}\mu}{t^{(n)}} - \frac{{\mu^{(n)}}^2 t}{{t^{(n)}}^2} \ge 2^{\rho_1}\mathrm{Tr}(\mathbf{V}), \tag{49}$$

which is a convex set.

APPENDIX B
PROOF OF THEOREM 1

The domain of Q_3 is determined by the intersection between convex sets and the sets determined by difference of two convex functions, which known as D.C. constraint. Let $f(Q)$ and $g(Q)$ denote the corresponding two convex functions respectively, where Q contains the beamforming and auxiliary variables. The D.C. constraints in Q_3 are equivalent to $f(Q) - g(Q) \le 0$. The

initial point of (n)-th iteration is denoted by $Q^{(n-1)}$. $\tilde{g}(Q)\big|_{Q^{(n-1)}}$ denotes the first-order Taylor expansion of $g(Q)$ at $Q = Q^{(n-1)}$. As we have shown above, since the function is convex, the first-order Taylor expansion will be below the original function $g(Q)$. The local optimum can be guaranteed because, compared with the original domain, the current domain is shrinked. Let $\mathbb{D}(Q^{(n-1)})$ denote the domain determined by the $(n-1)$-th iteration initial point. After employing D.C. programming, the output value is $\mathrm{Tr}(Q^{(n)})$, so the $\mathrm{Tr}(Q^{(n)})$ is the local optimal value in the $\mathbb{D}(Q^{(n-1)})$. However, $\mathrm{Tr}(Q^{(n)}) \leq \mathrm{Tr}(Q^{(n-1)})$ can not be established because we can not guarantee that $Q^{(n-1)}$ is inside of the $\mathbb{D}(Q^{(n-1)})$. Substituting $Q^{(n-1)}$ into the first-order Taylor expansion of $f(Q) - g(Q)$, we can obtain

$$f\left(Q^{(n-1)}\right) - \tilde{g}\left(Q^{(n-1)}\right)\Big|_{Q^{(n-1)}} \overset{(a)}{=} f\left(Q^{(n-1)}\right) - g\left(Q^{(n-1)}\right) \overset{(b)}{\leq} 0. \tag{50}$$

The equality (a) holds because the first-order Taylor expansion of function $g(Q^{(n-1)})$ is equal to $g(Q^{(n-1)})$ at the point $Q^{(n-1)}$. The inequality (b) is established because $Q^{(n-1)}$ is the feasible point at the last convex optimization process. Above expression indicates that $Q^{(n-1)}$ belongs to the domain defined by $Q^{(n-1)}$. Thus, $Q^{(n-1)}$ and $Q^{(n)}$ belong to the domain defined by $Q^{(n-1)}$ at the same time. We can deduce that $\mathrm{Tr}(Q^{(n)}) \leq \mathrm{Tr}(Q^{(n-1)})$ because $Q^{(n)}$ is no worse than $Q^{(n-1)}$ in the $\mathbb{D}(Q^{(n-1)})$. Since the energy consumption cannot be reduced indefinitely, we can deduce that the D.C. algorithm converges to a local optimal value. ■

APPENDIX C
PROOF OF THEOREM 2

We define the Lagrange dual function about Q_5 as

$$L(\mathbf{U}, \boldsymbol{\Delta}, \Gamma, \lambda, \beta) = \Gamma + \sum_{i=1}^{I} \lambda_i \left(\frac{\mathrm{Tr}(\tilde{\mathbf{G}}_i \mathbf{U})}{\varpi_i} - \Gamma \right) + \beta \left(\mathrm{Tr}(\tilde{\mathbf{V}} \mathbf{U} \tilde{\mathbf{V}}^H) - 1 \right) - \mathrm{Tr}(\mathbf{U}\boldsymbol{\Delta}), \tag{51}$$

where $\lambda_i \geq 0$, $\beta \geq 0$ and $\boldsymbol{\Delta} \succeq \mathbf{0}$ are the Lagrange multipliers associated with the problem constraints. It can be verified the partial derivative of $L(\mathbf{U}, \boldsymbol{\Delta}, \Gamma, \lambda, \beta)$ with respect to $\mathbf{U}$ is equal to 0 at optimal point, so we can obtain

$$\frac{\partial L}{\partial \mathbf{U}} = \sum_{i=1}^{I} \lambda_i \frac{\tilde{\mathbf{G}}_i}{\varpi_i} + \beta \tilde{\mathbf{V}}^H \tilde{\mathbf{V}} - \boldsymbol{\Delta} = \mathbf{0}, \tag{52}$$

$$\boldsymbol{\Delta}\mathbf{U} = \mathbf{0},$$

$$\mathbf{U} \succeq \mathbf{0}. \tag{53}$$

For both sides of (52), we can get the following expression by post-multiplying $\mathbf{U}$.

$$\lambda_j \frac{\tilde{\mathbf{G}}_j}{\varpi_j}\mathbf{U} + \sum_{i=1, i\neq j}^{I} \lambda_i \frac{\tilde{\mathbf{G}}_i}{\varpi_i}\mathbf{U} + \beta \tilde{\mathbf{V}}^H \tilde{\mathbf{V}}\mathbf{U} - \boldsymbol{\Delta}\mathbf{U} = \mathbf{0}. \tag{54}$$

There exists at least one $\lambda_j > 0$. Otherwise, we can improve Γ. Combining (53) and (54), we can obtain

$$\lambda_j \frac{\tilde{\mathbf{G}}_j}{\varpi_j}\mathbf{U} + \sum_{i=1, i\neq j}^{I} \lambda_i \frac{\tilde{\mathbf{G}}_i}{\varpi_i}\mathbf{U} + \beta \tilde{\mathbf{V}}^H \tilde{\mathbf{V}}\mathbf{U} = \mathbf{0}, \tag{55}$$

which indicates that

$$\mathrm{Rank}\left(-\lambda_j \frac{\tilde{\mathbf{G}}_j}{\varpi_j}\mathbf{U}\right) = \mathrm{Rank}\left(\left(\sum_{i=1, i\neq j}^{I} \lambda_i \frac{\tilde{\mathbf{G}}_i}{\varpi_i} + \beta \tilde{\mathbf{V}}^H \tilde{\mathbf{V}}\right)\mathbf{U}\right). \tag{56}$$

Because $\tilde{\mathbf{V}}$ is the null space of users, the rank of $\tilde{\mathbf{V}}^H\tilde{\mathbf{V}}$ is $N-2$. In Lemma 1, we have demonstrated that the problem Q_5 can be maximized at the power constraint boundary, so Lagrange multiplier, β, is bigger than 0. Otherwise, the power constraint is not tight. Meanwhile, $\lambda_i \geq 0$ indicates that $\lambda_i \frac{\tilde{\mathbf{G}}_i}{\varpi_i}$ is a positive semidefinite matrix. As above description, we can obtain $(\sum_{i=1, i\neq j}^{I} \lambda_i \frac{\tilde{\mathbf{G}}_i}{\varpi_i} + \beta \tilde{\mathbf{V}}^H \tilde{\mathbf{V}})$ is a full rank matrix.

$$\mathrm{Rank}(\mathbf{U}) = \mathrm{Rank}\left(-\lambda_j \frac{\tilde{\mathbf{G}}_j}{\varpi_j}\mathbf{U}\right) \leq \mathrm{Rank}\left(\lambda_j \frac{\tilde{\mathbf{G}}_j}{\varpi_j}\right) \overset{(a)}{=} 1. \tag{57}$$

The equality (a) holds because the rank of the $\tilde{\mathbf{G}}_j$ is 1. ■

APPENDIX D
PROOF OF THEOREM 3

On the above, we have proved the convergence of the D.C. algorithm, so the convergence about D.C. programming in ADC algorithm is omitted. Now, we concern about the convergence of alternating optimization algorithm. Firstly, We set that the users' beamforming, AN covariance matrix and power consumption are denoted by $\mathbf{W}^{(n-1)}$, $\mathbf{U}^{(n-1)}$ and $P(\mathbf{W}^{(n-1)}, \mathbf{U}^{(n-1)})$ at the $(n-1)$-th iteration, respectively, where $\mathbf{W}^{(n-1)}$ contains two users' beamformer. Given the initial point of AN at the (n)-th iteration, $\mathbf{U}^{(n-1)}$, we set the initial point of $\lambda^{(n)}$, $\beta^{(n)}$ and $\psi^{(n)}$ as follows

$$\lambda^{(n)} = \max\left(1 + \frac{\mathrm{Tr}\left(\mathbf{G}_m \mathbf{W}_2^{(n-1)}\right)}{\mathrm{Tr}\left(\mathbf{G}_m \mathbf{W}_1^{(n-1)}\right) + \mathrm{Tr}\left(\tilde{\mathbf{G}}_m \mathbf{U}^{(n-1)}\right) + 1}\right),$$

$$\beta_k^{(n)} = \sqrt{\mathrm{Tr}\left(\mathbf{H}_k \mathbf{W}_1^{(n-1)} + \mathbf{H}_k \mathbf{W}_2^{(n-1)}\right) + 1 - 2^{\rho_2}\lambda^{(n-1)}},$$

$$\psi_m^{(n)} = \sqrt{\mathrm{Tr}\left(\mathbf{G}_m \mathbf{W}_2^{(n-1)}\right)}, \tag{58}$$

for $\forall m \in \mathcal{M}$ and $\forall k \in \{1, 2\}$. The initial point is $(\mathbf{U}^{(n-1)}, \lambda^{(n)}, \beta^{(n)}, \psi^{(n)})$ at the (n)-th iteration in the first layer. Combining $(\lambda^{(n)}, \beta^{(n)}, \psi^{(n)})$ in (58) and the last AN covariance matrix $\mathbf{U}^{(n-1)}$, we can guarantee that the beamforming matrices $(\mathbf{W}_1^{(n-1)}, \mathbf{W}_2^{(n-1)})$ can satisfy the problem Q_6. We obtain the optimal users' beamforming, $\mathbf{W}^{(n)}$, by using CVX to solve the Q_6, and we can derive

$$P\left(\mathbf{W}^{(n-1)}, \mathbf{U}^{(n-1)}\right) \overset{(a)}{\geq} P\left(\mathbf{W}^{(n)}, \mathbf{U}^{(n-1)}\right). \tag{59}$$

The inequality (a) holds because $\mathbf{W}^{(n)}$ is local optimum by CVX at $\mathbf{U} = \mathbf{U}^{(n-1)}$ and $(\mathbf{W}^{(n-1)}, \mathbf{U}^{(n-1)})$ is also included in the domain of optimization. After obtaining the optimal beamformer $(\mathbf{W}^{(n)}, \mathbf{U}^{(n-1)})$, we regard $(\mathbf{W}^{(n)}, \mathbf{U}^{(n-1)})$ as the initial point to solve the problem Q_5, and $(\mathbf{W}^{(n)}, \mathbf{U}^{(n)})$ is outputted as the optimal point, so the energy relation can be expressed as

$$P\left(\mathbf{W}^{(n)}, \mathbf{U}^{(n-1)}\right) \overset{(b)}{\geq} P\left(\mathbf{W}^{(n)}, \mathbf{U}^{(n)}\right). \tag{60}$$

The inequality (b) holds because $\mathbf{U}^{(n)}$ is the optimal AN covariance matrix for the optimization problem Q_5 when $\mathbf{W} = \mathbf{W}^{(n)}$ and $(\mathbf{W}^{(n)}, \mathbf{U}^{(n-1)})$ can be found even if it has reached the optimum. Combining (59) and (60), we can obtain

$$P\left(\mathbf{W}^{(n-1)}, \mathbf{U}^{(n-1)}\right) \geq P\left(\mathbf{W}^{(n)}, \mathbf{U}^{(n)}\right). \tag{61}$$

which indicates that the resulting value of algorithm 2 is non-increasing. However, considering the power consumption can not be infinitely small, we deduce that algorithm 2 converges after finite iterations. ■

ACKNOWLEDGMENT

The authors would like to sincerely thank Prof. Man-on Pun for his helpful and insightful comments.

REFERENCES

[1] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[2] H. Haci, H. Zhu, and J. Wang, "Performance of non-orthogonal multiple access with a novel asynchronous interference cancellation technique," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1319–1335, Mar. 2017.

[3] G. Gui, H. Sari, and E. Biglieri, "A new definition of fairness for non-orthogonal multiple access," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1267–1271, Jul. 2019.

[4] E. C. Cejudo, H. Zhu, and O. Alluhaibi, "On the power allocation and constellation selection in downlink NOMA," in *Proc. IEEE 86th Veh. Technol. Conf.*, Sep. 2017, pp. 1–5.

[5] Y. Yin, Y. Peng, M. Liu, J. Yang, and G. Gui, "Dynamic user grouping-based NOMA over Rayleigh fading channels," *IEEE Access*, vol. 7, pp. 110 964–110 971, 2019.

[6] X. Lu, P. Wang, D. Niyato, I. K. Dong, and Z. Han, "Wireless networks with Rf energy harvesting: A contemporary survey," *IEEE Commun. Surv. Tut.*, vol. 17, no. 2, pp. 757–789, Nov. 2015.

[7] J. Zhang, C. Yuen, C. Wen, S. Jin, K. Wong, and H. Zhu, "Large system secrecy rate analysis for SWIPT MIMO wiretap channels," *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 1, pp. 74–85, Jan. 2017.

[8] P. Kamalinejad, C. Mahapatra, Z. Sheng, S. Mirabbasi, V. C. M. Leung, and Y. L. Guan, "Wireless energy harvesting for the Internet of Things," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 102–108, Jun. 2015.

[9] Y. Xu, G. Li, Y. Yang, M. Liu, and G. Gui, "Robust resource allocation and power splitting in SWIPT enabled heterogeneous networks: A robust min-imax approach," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10799–10811, Dec. 2019.

[10] S. Timotheou, I. Krikidis, and B. Ottersten, "MISO interference channel with QoS and RF energy harvesting constraints," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2013, pp. 4191–4196.

[11] S. Timotheou, I. Krikidis, G. Zheng, and B. Ottersten, "Beamforming for MISO interference channels with QoS and RF energy transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2646–2658, May 2014.

[12] Q. Shi, W. Xu, T. Chang, Y. Wang, and E. Song, "Joint beamforming and power splitting for MISO interference channel with SWIPT: An socp relaxation and decentralized algorithm," *IEEE Trans. Signal Process.*, vol. 62, no. 23, pp. 6194–6208, Dec. 2014.

[13] M. Basharat, W. Ejaz, M. Naeem, A. M. Khattak, A. Anpalagan, and O. Alfandi, "Energy efficient resource allocation for NOMA in cellular IoT with energy harvesting," in *Proc. IEEE 13th ICET*, Dec. 2017, pp. 1–6.

[14] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[15] S. Zhang, H. Zhang, and L. Yang, "Max-min fair robust beamforming design for multi-user MISO SWIPT systems," in *Proc. IEEE Int. Conf. Commun. Syst.*, Dec. 2016, pp. 1–5.

[16] M. Jiang, Y. Li, Q. Zhang, Q. Li, and J. Qin, "Secure beamforming in downlink MIMO nonorthogonal multiple access networks," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1852–1856, Dec. 2017.

[17] F. Cheng, G. Gui, N. Zhao, Y. Chen, J. Tang, and H. Sari, "UAV-relaying-assisted secure transmission with caching," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3140–3153, May 2019.

[18] Q. Vu, L. Tran, R. Farrell, and E. Hong, "An efficiency maximization design for SWIPT," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2189–2193, Dec. 2015.

[19] Q. Shi, C. Peng, W. Xu, M. Hong, and Y. Cai, "Energy efficiency optimization for MISO SWIPT systems with zero-forcing beamforming," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 842–854, Feb. 2016.

[20] Q. Shi, L. Liu, W. Xu, and R. Zhang, "Joint transmit beamforming and receive power splitting for MISO SWIPT systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3269–3280, Jun. 2014.

[21] Y. Xu *et al.*, "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4874–4886, Sep. 2017.

[22] B. Su, Q. Ni, and W. Yu, "Robust transmit beamforming for SWIPT-enabled cooperative NOMA with channel uncertainties," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4381–4392, Jun. 2019.

[23] Y. Yuan, Y. Xu, Z. Yang, P. Xu, and Z. Ding, "Energy efficiency optimization in full-duplex user-aided cooperative SWIPT NOMA systems," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5753–5767, Aug. 2019.

[24] Y. Zhang, H. Wang, Q. Yang, and Z. Ding, "Secrecy sum rate maximization in non-orthogonal multiple access," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 930–933, May 2016.

[25] Y. Feng, S. Yan, Z. Yang, N. Yang, and J. Yuan, "Beamforming design and power allocation for secure transmission with NOMA," *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 2639–2651, May 2019.

[26] Q. Li and W. Ma, "Spatially selective artificial-noise aided transmit optimization for MISO multi-eves secrecy rate maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2704–2717, May 2013.

[27] B. Fang, Z. Qian, W. Shao, and W. Zhong, "Precoding and artificial noise design for cognitive MIMOME wiretap channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6753–6758, Aug. 2016.

[28] V. Nguyen, T. Q. Duong, O. A. Dobre, and O. Shin, "Secrecy rate maximization in a cognitive radio network with artificial noise aided for MISO multi-eves," in *Proc. IEEE Int. Conf. Commun.*, May 2016, pp. 1–6.

[29] J. Zhang, C. Yuen, C. Wen, S. Jin, K. Wong, and H. Zhu, "Large system secrecy rate analysis for SWIPT MIMO wiretap channels," *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 1, pp. 74–85, Jan. 2016.

[30] Y. Hao and T. Lv, "SWIPT-aided secure beamforming design for downlink cooperative NOMA systems," in *Proc. Global Wireless Summit*, Nov. 2018, pp. 364–369.

[31] Z. Zhu, Z. Chu, N. Wang, S. Huang, Z. Wang, and I. Lee, "Beamforming and power splitting designs for an-aided secure multi-user MIMO SWIPT systems," *IEEE Trans. Inf. Forensics Secur.*, vol. 12, no. 12, pp. 2861–2874, Dec. 2017.

[32] H. Zhang, C. Li, Y. Huang, and L. Yang, "Secure beamforming for SWIPT in multiuser MISO broadcast channel with confidential messages," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1347–1350, Aug. 2015.

[33] Z. Zhu, Z. Chu, N. Wang, S. Huang, Z. Wang, and I. Lee, "Beamforming and power splitting designs for an-aided secure multi-user MIMO SWIPT systems," *IEEE Trans. Inf. Forensics Secur.*, vol. 12, no. 12, pp. 2861–2874, Dec. 2017.

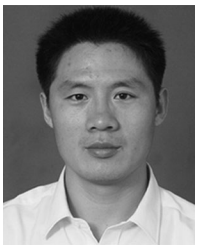
[34] Y. Sun, D. W. K. Ng, J. Zhu, and R. Schober, "Robust and secure resource allocation for full-duplex MISO multicarrier NOMA systems," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 4119–4137, Sep. 2018.

[35] L. Lv, Z. Ding, Q. Ni, and J. Chen, "Secure MISO-NOMA transmission with artificial noise," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6700–6705, Jul. 2018.

[36] F. Zhou, Z. Chu, Y. Wu, N. Al-Dhahir, and P. Xiao, "Enhancing phy security of MISO NOMA SWIPT systems with a practical non-linear EH model," in *Proc. IEEE Int. Conf. Commun. Workshops*, May 2018, pp. 1–6.

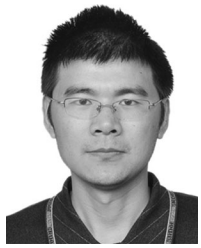
[37] F. Zhou, Z. Chu, H. Sun, R. Q. Hu, and L. Hanzo, "Artificial noise aided secure cognitive beamforming for cooperative MISO-NOMA using SWIPT," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 918–931, Apr. 2018.
[38] E. C. Cejudo, H. Zhu, J. Wang, and O. Alluhaibi, "A fast algorithm for resource allocation in downlink multicarrier NOMA," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2019, pp. 1–5.
[39] Q. Sun, S. Han, C. I, and Z. Pan, "Energy efficiency optimization for fading MIMO non-orthogonal multiple access systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 2668–2673.
[40] M. Tian, Q. Zhang, S. Zhao, Q. Li, and J. Qin, "Secrecy sum rate optimization for downlink MIMO nonorthogonal multiple access systems," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1113–1117, Aug. 2017.
[41] Y. Zhang, H. Wang, Q. Yang, and Z. Ding, "Secrecy sum rate maximization in non-orthogonal multiple access," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 930–933, May 2016.
[42] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: Architecture design and rate-energy tradeoff," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754–4767, Nov. 2013.
[43] H. Bao, C. Zhang, L. Wu, and M. Li, "Design of physical layer secure transmission scheme based on SWIPT NOMA systems," in *Proc. IEEE 17th Int. Conf. Commun. Technol.*, Oct. 2017, pp. 6–9.
[44] F. Zhou, Z. Chu, H. Sun, and V. C. M. Leung, "Resource allocation for secure MISO-NOMA cognitive radios relying on SWIPT," in *Proc. IEEE Int. Conf. Commun.*, May 2018, pp. 1–6.
[45] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming, version 2.1," Mar. 2014. [Online]. Available: http://cvxr.com/cvx
[46] Q. Cao, Y. Sun, W. Shi, Q. Ni, and B. Wang, "Max-min weighted achievable rate for full-duplex MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 37–40, Feb. 2019.
[47] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local D.C. programming," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.
[48] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Jiasi Zhou received the B.S. degree in electronic and communication engineering from Pingdingshan University, in 2016. He is currently working toward the Ph.D. degree with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China. His current research interests include cache-enabled C-RAN, rate splitting multiple access, non-orthogonal multiple access, and physical layer security.



Yanjing Sun (Member, IEEE) received the Ph.D. degree in information and communication engineering from the China University of Mining and Technology, Xuzhou, China, in 2008. He is currently a Professor with the School of Information and Control Engineering and the Director of Network Information Center, China University of Mining and Technology. His current research interests include IBFD communications, embedded real-time systems, wireless sensor networks, and cyber-physical systems.

Qi Cao received the B.S. degree in electronic communication engineering from the University of Liverpool, in 2013, the M.S. degree in communications signal processing from Imperial College, in 2014, and the Ph.D. degree from the School of Information and Control Engineering of China University of Mining and Technology, in 2018. He is now a Postdoc with the Chinese University of Hong Kong, Shenzhen. His research interests include MIMO wireless communications, multi-user transmission and full-duplex and cooperative communication, and his current research focuses on AI driven radio resource management.



Song Li (Member, IEEE) received the B.Sc. and M.Sc. degrees in communication and information system from Tianjin University, Tianjin, China, in 2007 and 2009, respectively, and the Ph.D. degree in signal and information processing from the Beijing University of Posts and Telecommunications, Beijing, China, in 2012. He is currently an Associate Professor with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China. His current research interests include fifth-generation (5G) wireless networks, cyber-physical system, and industrial Internet of Things.

Zhi Sun (Member, IEEE) received the B.S. degree in telecommunication engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2004, the M.S. degree in electronic engineering from Tsinghua University, Beijing, in 2007, and the Ph.D. degree in electrical and computer engineering from Georgia Institute of Technology, Atlanta, USA, in 2011. He was a Postdoctoral Fellow with the Georgia Institute of Technology, Atlanta, USA, from 2011 to 2012. He joined the Department of Electrical Engineering at University at Buffalo, State University of New York, Buffalo, NY, USA, as an Assistant Professor, in 2012. Currently, he is an Associate Professor with the Department of Electrical Engineering at University at Buffalo. He currently serves as an Editor of Elsevier Computer Networks Journal (Elsevier). His research interests include wireless communication and networking in extreme environments, metamaterial enhanced communication and security, physical-layer security, wireless intra-body networks, wireless underground networks, wireless underwater networks, and cyber physical systems. Prof. Sun was the recipient of the National Science Foundation (NSF) Faculty Early Career Development (CAREER) Award in 2017, the University at Buffalo (UB) Exceptional Scholar – Young Investigator Award in 2017, the Best Demo Award in IEEE Infocom 2017, the Best Paper Award in IEEE Global Communications Conference (GLOBECOM) in 2010, the Broadband Wireless Networking (BWN) Researcher of the Year Award at Georgia Institute of Technology in 2009, and the Outstanding Graduate Award at Tsinghua University in 2007. He currently serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Xiaolin Wang received the Ph.D. degree from the China University of Mining and Technology, in 2012. She is currently an Associate Professor with the School of Mines, China University of Mining and Technology, Xuzhou, China. Her current research interests include mining system engineering, mining communications and so on.