SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

September 2023, Vol. 66 192302:1–192302:13
https://doi.org/10.1007/s11432-022-3653-y

Intelligent reflecting surface assisted untrusted NOMA transmissions: a secrecy perspective

Dawei WANG[1,2], Xuanrui LI[1,2], Yixin HE[1,2], Fuhui ZHOU[3,4*] & Qihui WU[3,4]

[1]*School of Electronics and Information, Northwestern Polytechnical University, Xi'an* 710072*, China;*
[2]*Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen* 518057*, China;*
[3]*College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing* 211106*, China;*
[4]*Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, Nanjing University of Aeronautics and Astronautics, Nanjing* 211106*, China*

Received 7 September 2022/Revised 26 November 2022/Accepted 13 December 2022/Published online 28 August 2023

Abstract In this paper, we investigate the intelligent reflecting surface (IRS) assisted secure transmissions for the untrusted non-orthogonal multiple access (NOMA) network threatened by an external eavesdropper, where the near user and two untrusted far users receive the privacy information from the transmitter assisted by the IRS's passive reflection. To protect the privacy information against internal and external eavesdropping, we first provide a comprehensive analysis of the secrecy rates, which can offer suitable methods for separately handling the internal and external secrecy performance. Following the above results, we formulate a secrecy rate maximization problem by optimizing the transmission power allocation and phase shifts. To achieve this goal, we decouple the non-convex optimization problem into the external security problem and the internal security problem. Then, we propose a heuristic simulated annealing algorithm to iteratively solve the above problems, where the semi-definite relaxation (SDR) technique, the singular value decomposition (SVD) technique, and the Lagrange multiplier method are utilized to relax the problems and optimize the power allocation and phase shifts. Finally, numerical results show that the proposed scheme can improve the secrecy rate compared with the existing IRS-aided secure NOMA transmission schemes.

Keywords intelligent reflecting surface, non-orthogonal multiple access, security, untrusted users

Citation Wang D W, Li X R, He Y X, et al. Intelligent reflecting surface assisted untrusted NOMA transmissions: a secrecy perspective. Sci China Inf Sci, 2023, 66(9): 192302, https://doi.org/10.1007/s11432-022-3653-y

1 Introduction

Fifth-generation (5G) mobile communication technology has been widely commercialized owing to its high data rate, low latency, energy savings, and large-scale equipment connections [1]. In 5G networks, the non-orthogonal multiple access (NOMA) technology has been proven to have great application potential in alleviating the spectrum shortage problem [2]. Compared to the orthogonal multiple access (OMA) schemes, the NOMA technique removes the limitation of orthogonality and improves spectrum efficiency by allowing multiple users to share the same time-frequency resources with different power levels [3, 4]. To cancel the multi-user interference, successive interference cancellation (SIC) is utilized [5], which demodulates the signals of all users in order to degrade the users' intra-interference [6, 7]. Besides, the NOMA technique is widely investigated in the current researches. In [8], the authors adopted this technique to improve the achievable sum rate in multiple-input multiple-output (MIMO)-based multi-user visible light communication (VLC) systems, and gained a significant transmission rate improvement. The authors in [9, 10] showed that the NOMA technology could effectively utilize multipath fading in wireless channels and improved the performance of cooperative communication systems. In addition, NOMA could utilize limited spectrum resources in the dense network, effectively alleviating severe data congestion and avoiding low access efficiency [11]. The authors in [12] showed that NOMA could improve the system performance and energy efficiency in multiple application areas [13, 14].

* Corresponding author (email: zhoufuhui@ieee.org)

The intelligent reflecting surface (IRS) technology is also a promising technique to assist NOMA communications [15]. IRS is comprised of low-cost passive elements that reflect the incident signal with the different phase shifts. By adaptively changing the phase shifts, IRS can create a reconfigurable radio environment with high energy and spectrum efficiency for wireless communications [16]. In addition, IRS can be regarded as a multi-antenna relay, which is different from the conventional relay. Specifically, IRS works as a reconfigurable scatter without any energy supply for information transmissions [17]. Existing researches about IRS mainly focus on the enhancement of signal coverage and energy efficiency [18], where the power consumption is reduced by optimizing the beamforming vector and passive reflection phase shift [19]. Moreover, IRS is also an effective method to improve the quality of service (QoS) for cell-edge users [20]. The authors in [21] studied the joint beamforming and phase shift optimization for unmanned aerial vehicles (UAV) downlink networks. The authors in [22, 23] investigated the IRS-aided downlink multi-cluster NOMA transmission, where each cluster was serviced by one IRS. In the uplink NOMA network, IRS also improved the performance of remote users, and boosted the spectrum and energy efficiency in multi-point transmissions [24].

Since the wireless electromagnetic wave is transmitted in the open wireless media, the wireless privacy information is vulnerable to being eavesdropped [25], especially in a NOMA network with multiple users utilizing the same spectrum. Currently, multiple works have studied the information security of these networks. The authors in [26] utilized the instantaneous channel state information (CSI) to enhance the secrecy performance of a full-duplex (FD) NOMA system in vehicle-to-vehicle (V2V) communications. The authors in [26] considered a practical eavesdropping scenario with imperfect CSI. In [27], the authors studied the energy-harvesting scenario with the external eavesdropper. The authors in [28] investigated the millimeter-wave and Terahertz secure communication to maximize the system secrecy rate. The authors in [29] proposed two schemes to protect the information security of the NOMA network with untrusted users. Moreover, the authors in [30] proposed an internal distrust model for the near users, but the far users' information security was ignored. Furthermore, the authors in [31] investigated a two-way secure communication network, and optimized the IRS phase shift to protect the privacy information by destructively adding the IRS-reflected and non-IRS-reflected signals at the untrusted users. In the above studies, the information security from the internal or external threats was separately studied for IRS-aided NOMA networks. However, for the multiple untrusted users network, the information security threats from both internal and external users. To the best of the authors' knowledge, there is no work yet that has investigated the information security problems by considering both the internal and external eavesdropping in IRS-aided NOMA networks.

Motivated by the above discussions, we will study the secure transmissions for the IRS-aided NOMA networks with both the internal and external eavesdropping. In this network, we first provide a comprehensive analysis of the secrecy rates. Then, by optimizing the transmission power allocation and phase shifts, we formulate a secrecy rate maximization problem. To achieve this goal, we decouple the non-convex problem into the external security problem and the internal security problem. In addition, a heuristic algorithm is proposed to iteratively solve the above problems. The contributions of this work can be summarized as follows.

- In the proposed system, one near user and two untrusted far users will securely receive their privacy information against the external and internal threats. In addition, due to the long transmission range for the far users, an IRS is deployed to improve the secrecy performance. To the best of the authors' knowledge, this is the first work that studies the information security against the external and internal eavesdropping in the IRS-aided NOMA networks.
- For the proposed system, in order to protect the privacy information against untrusted users as well as external eavesdroppers, we propose a novel priority-based secure NOMA transmission scheme for the IRS-aided networks, where the privacy information is transmitted through the NOMA technique and assisted by the IRS. According to the priorities of untrusted users, we formulate a secrecy rate maximization problem, where the NOMA transmission power and IRS's phase shifts are optimized. To solve the above non-convex problem, we decouple the original problem into the external security problem and the internal security problem. Moreover, we propose an improved simulated annealing iteration algorithm, where the semi-definite relaxation (SDR), singular value decomposition (SVD), and Lagrange multiplier methods are utilized to obtain the optimal power allocation and phase shifts.
- In the simulation section, we investigate the performance of our proposed scheme and compare it with other benchmark schemes. Numerical results show that the proposed scheme can improve the secrecy rate of all users compared with the existing IRS-aided secure transmission scheme. Even though

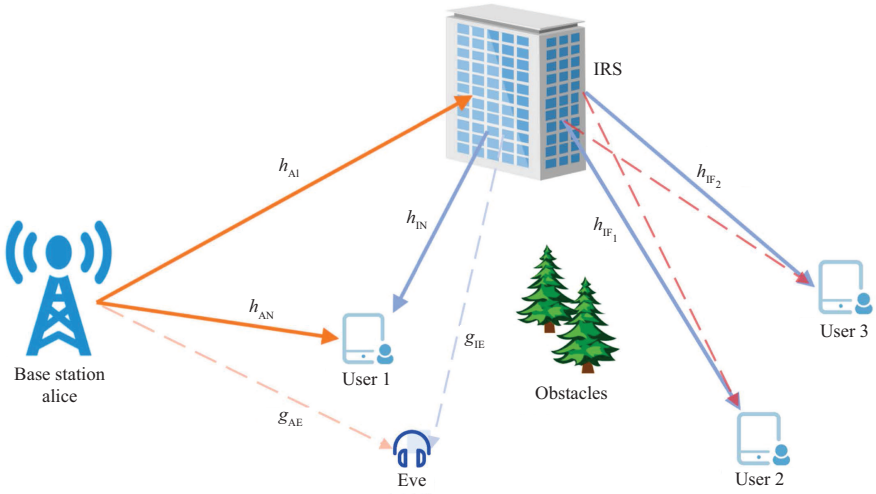

Figure 1 (Color online) IRS-assisted NOMA secure transmission system with untrusted users and external eavesdropper.

the multiple users' security priority and external eavesdropping make the problem complicated to allocate the transmission power and design the phase shift, we propose a heuristic algorithm to solve the secrecy rate maximization problem.

The rest of this paper is organized as follows. In Section 2, we establish the basic IRS-assisted NOMA system model and define the relevant formulas. In Section 3, we formulate the optimization problem to optimize the phase shift and power allocation. Section 4 elaborates the proposed iterative optimization scheme based on the heuristic simulated annealing algorithm and analyzes its complexity. In Section 5, we simulate the proposed scheme and study the system performance through numerical simulations. In the end, conclusion is drawn in Section 6.

2 System model

We consider an IRS-aided NOMA network as shown in Figure 1, which consists of a base station (BS), denoted as Alice, an IRS, a near user, denoted as U_1, two far users, denoted as U_2 and U_3, respectively, and an eavesdropper Eve denoted as E, which is located around U_1. The IRS consists of M ($M \geqslant 1$) passive reflecting elements, and includes a smart controller that can adjust the phase shifts to assist the NOMA secure transmission intelligently and improve the information reception quality [32–34]. All users and BS are equipped with only one antenna. Due to the obstacles' blocking and/or the several path losses owing to the long distance, we assume that there is a weak downlink path between Alice and the far users. U_1 and E are located close to Alice and have direct links to Alice. E is deployed as the equipment with the ability to work as a normal user, and also has the potential to access for harvesting energy from radio frequency signals from other users. Moreover, it has to harvest enough energy to ensure its normal working.

In this network, we assume that U_1 is trusted with U_2 and U_3. Meanwhile, U_2 and U_3 are untrusted with each other, and they have the corresponding security levels. The low-level user threatens the high-level user's information security by decoding its signal, which means that the high-level user tends to distrust the low-level user, and treats the low-level legal user as the potential internal eavesdropper. We suppose that the security level is expressed as $U_2 > U_3$. In addition, all users' information is threatened by Eve. This system model can be applied in the communication scenario where the far users are mobile users that just access the network and untrusted with each other[1)]. In each time slot, Alice utilizes NOMA to simultaneously transmit all users' signals. Since there are three users in the system, the base-band signal transmitted by Alice can be expressed as

$$x = \sum_{k=1}^{3} \sqrt{a_k P} x_k, \tag{1}$$

1) The model is easy to be expanded to multiple users scenarios where there are multiple near and far users.

where α_k $(k=1,2,3)$ represents the weight of power allocation for the k-th data stream, $a_k \in [0,1]$; P is the total power of Alice; x_1, x_2, and x_3 represent the data flows, which will be sent to U_1, U_2, and U_3, respectively; $x_1, x_2, x_3 \sim \mathcal{CN}(0,1)$.

In the proposed scheme, U_1 can receive the information directly from Alice and be assisted by IRS. U_2 and U_3 only receive the signal assisted by IRS. Therefore, the received signals at U_1, U_2, and U_3 are denoted as

$$\begin{cases} y_1 = (\boldsymbol{h}_{\mathrm{IN}}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} + h_{\mathrm{AN}})x + n_1, \\ y_2 = \boldsymbol{h}_{\mathrm{IF}_1}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} x + n_2, \\ y_3 = \boldsymbol{h}_{\mathrm{IF}_2}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} x + n_3, \end{cases} \tag{2}$$

where $\boldsymbol{\Theta} = \mathrm{diag}\{\mathrm{e}^{\mathrm{j}\Theta_1}, \ldots, \mathrm{e}^{\mathrm{j}\Theta_m}, \ldots, \mathrm{e}^{\mathrm{j}\Theta_M}\}$ denotes the IRS's phase-shift matrix, among which $\Theta_m \in [0, 2\pi)$ denotes the phase shift of the m-th reflecting element, $m = 1, \ldots, M$; $\boldsymbol{h}_{\mathrm{IN}}$, $\boldsymbol{h}_{\mathrm{IF}_1}$, and $\boldsymbol{h}_{\mathrm{IF}_2} \in \mathbb{C}^{M\times 1}$ denote the channel vectors between IRS and U_1, between IRS and U_2, between IRS and U_3, respectively; $\boldsymbol{h}_{\mathrm{AI}} \in \mathbb{C}^{M\times 1}$ and $h_{\mathrm{AN}} \in \mathbb{C}^{1\times 1}$ denote the vectors between Alice and IRS, and between Alice and U_1, respectively [35, 36]; n_1, n_2, and n_3 denote the additive white Gaussian noise (AWGN) at U_1, U_2, and U_3, respectively. For simplicity, we assume that all AWGNs in this model have the same variance σ^2, and $n_k \sim \mathcal{CN}(0, \sigma^2)$, $k = 1, 2, 3$. For the eavesdropper E, the received signal can be expressed as

$$y_e = (\boldsymbol{g}_{\mathrm{IE}}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} + g_{\mathrm{AE}})x + n_e, \tag{3}$$

where $\boldsymbol{g}_{\mathrm{IE}} \in \mathbb{C}^{M\times 1}$ and $g_{\mathrm{AE}} \in \mathbb{C}^{1\times 1}$ denote the channel vectors between IRS and Eve, and between Alice and Eve, respectively; $n_e \sim \mathcal{CN}(0, \sigma^2)$ is the received noise at E. In addition, E will harvest the energy from the radio-frequency signal to power itself. The minimum harvested energy at E should be great than Q_E. The harvested energy is derived as

$$P_E = \left| \boldsymbol{g}_{\mathrm{IE}}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} + g_{\mathrm{AE}} \right|^2 P\eta, \tag{4}$$

where η is the energy conversion efficiency factor, $\eta \in (0,1)$, and we can neglect the noise power since the noise power is much less than the signal.

For the NOMA transmission, we adopt the SIC technology to decode the desired signal. In the downlink of the NOMA system, the weak user decodes the data directly which means that the strong user signal is regarded as interference. The strong user decodes the weak user's data first and then eliminates the interference to decode itself data. We set $R_{i\to j}$ as the rate for U_i decoding U_j's data, where $i = 1, 2, 3$ and $j = 1, 2, 3$. According to the SIC technology and the security priority levels, the decoding rates at U_1 can be derived as

$$\begin{cases} R_{1\to 3} = \log_2 \left(1 + \dfrac{a_3 P \left| \boldsymbol{h}_{\mathrm{IN}}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} + h_{\mathrm{AN}} \right|^2}{(a_1 + a_2) P \left| \boldsymbol{h}_{\mathrm{IN}}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} + h_{\mathrm{AN}} \right|^2 + \sigma^2} \right), \\ R_{1\to 2} = \log_2 \left(1 + \dfrac{a_2 P \left| \boldsymbol{h}_{\mathrm{IN}}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} + h_{\mathrm{AN}} \right|^2}{a_1 P \left| \boldsymbol{h}_{\mathrm{IN}}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} + h_{\mathrm{AN}} \right|^2 + \sigma^2} \right), \\ R_{1\to 1} = \log_2 \left(1 + \dfrac{a_1 P \left| \boldsymbol{h}_{\mathrm{IN}}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} + h_{\mathrm{AN}} \right|^2}{\sigma^2} \right), \end{cases} \tag{5}$$

where $R_{1\to 1}$ denotes the transmission rate of U_1. The decoding rates at U_2 can be written as

$$\begin{cases} R_{2\to 3} = \log_2 \left(1 + \dfrac{a_3 P \left| \boldsymbol{h}_{\mathrm{IF}_1}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} \right|^2}{(a_1 + a_2) P \left| \boldsymbol{h}_{\mathrm{IF}_1}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} \right|^2 + \sigma^2} \right), \\ R_{2\to 2} = \log_2 \left(1 + \dfrac{a_2 P \left| \boldsymbol{h}_{\mathrm{IF}_1}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} \right|^2}{a_1 P \left| \boldsymbol{h}_{\mathrm{IF}_1}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} \right|^2 + \sigma^2} \right), \end{cases} \tag{6}$$

where $R_{2\to 2}$ denotes the transmission rate of U_2. The user U_3 has the lowest security level. The transmission rate of U_3 is expressed as

$$R_{3\to 3} = \log_2 \left(1 + \frac{a_3 P \left| \boldsymbol{h}_{\mathrm{IF}_2}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} \right|^2}{(a_1 + a_2) P \left| \boldsymbol{h}_{\mathrm{IF}_2}^{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\mathrm{AI}} \right|^2 + \sigma^2} \right). \tag{7}$$

According to the security levels of all users, the user with a higher security level has the possibility and ability to decode the information of the user with a lower security level. Thus, the eavesdropping rate for U_2 threatened by U_3 is derived as

$$R_{3\to 2} = \log_2\left(1 + \frac{a_2 P\left|\boldsymbol{h}_{\mathrm{IF}_2}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{AI}}\right|^2}{a_1 P\left|\boldsymbol{h}_{\mathrm{IF}_2}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{AI}}\right|^2 + \sigma^2}\right). \tag{8}$$

Then, the internal secrecy rate is defined as the decoding rate $R_{i\to i}$ (the rate of U_i decoding its own information) minus the eavesdropping rate $R_{j,i}$ (the rate of potential eavesdroppers decoding this message), where U_j represents internal eavesdroppers [28]. In order to protect the information against the internal threat, we need to guarantee that $R_{i\to i}$ is larger than $R_{j\to i}$. Therefore, for the security requirement of U_2, the maximum transmission rate of U_2 should not be smaller than the potential eavesdropping rate at U_3, and the secrecy rate is defined as

$$R_{\mathrm{in}} = \left(R_{2\to 2} - R_{3\to 2}\right)^+, \tag{9}$$

where $(r)^+ \triangleq \max(r, 0)$.

For the external eavesdropping, the wiretap rate is derived as [30]

$$R_{e\to i} = \log_2\left(1 + \mathrm{SINR}_{e,i} = \frac{\alpha_i P\left|\boldsymbol{g}_{\mathrm{IE}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{AI}} + g_{\mathrm{AE}}\right|^2}{\sum_{j=1}^{i-1}\alpha_j P\left|\boldsymbol{g}_{\mathrm{IE}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{AI}} + g_{\mathrm{AE}}\right|^2 + \sigma^2}\right), \quad i = 1, 2, 3. \tag{10}$$

Therefore, the external secrecy rate is derived as

$$R_{\mathrm{ex}} = \min_{i=1,2,3}\left(R_{i\to i} - R_{e\to i}\right)^+. \tag{11}$$

3 Secrecy rate maximization

The goal of this work is to maximize the internal and external secrecy rates by optimizing the power allocation $\boldsymbol{\alpha} = [a_1, a_2, a_3]^{\mathrm{T}}$ and phase shift $\boldsymbol{\Theta}$. The optimal problem is formulated as

$$\begin{aligned}
(\mathrm{P1}): \max_{\boldsymbol{\alpha},\boldsymbol{\Theta}} \quad & \min\{R_{\mathrm{ex}}, R_{\mathrm{in}}\} && \text{(12a)}\\
\text{s.t.} \quad & R_{\mathrm{ex}} \geqslant Q_1, && \text{(12b)}\\
& R_{\mathrm{in}} \geqslant Q_2, && \text{(12c)}\\
& R_{i\to i} \leqslant R_{j\to i}, \quad 1 \leqslant j \leqslant i \leqslant 3, && \text{(12d)}\\
& P_E \geqslant Q_E, && \text{(12e)}\\
& \sum_{i=1}^{3} a_i \leqslant 1, && \text{(12f)}\\
& a_i \geqslant 0, \ \forall i, && \text{(12g)}\\
& 0 \leqslant \Theta_m \leqslant 2\pi, \ \forall m, m = 1, 2, \ldots, M, && \text{(12h)}
\end{aligned}$$

where Q_1 and Q_2 denote the internal and external secrecy rate constraints, respectively. In (12a), our goal is to maximize the minimum value of R_{in} and R_{ex}, such that we can optimize the overall secrecy performance in the system. Eqs. (12b) and (12c) are external and internal security constraints, respectively. Eq. (12d) is the constraint for the SIC decoding principle. To make sure the successful SIC decoding process, we should restrict the rates when weak users decode itself less than the rates when they are decoded by strong users. Eq. (12e) means the overall harvested energy by E that should be large than the required energy Q_E to satisfy the eavesdropper's energy requirement. This is because E can be seen as one of the untrusted near users that need enough energy to support normal communication activities like U_1. Eqs. (12f) and (12g) are power allocation constraints restricted by SIC decoding requirements. Eq. (12h) is unit modulus constraints on the diagonals of Θ, which should be limited between 0 and 2π. Since this problem is non-convex and it is difficult to optimize the secrecy rate with the traditional convex algorithms, we will propose an iterative algorithm to solve the optimization problem.

4 Optimal power allocation and phase shift design

To deal with the formulated optimization problem, we decouple P1 into the external and internal cases and generate two subproblems, P2 and P3, respectively. Then, a heuristic simulated annealing algorithm is proposed to process the above two optimization problems iteratively.

The optimization problem P1 can be decoupled as

$$\begin{aligned}
(\mathrm{P2}):\max_{\boldsymbol{\alpha},\boldsymbol{\Theta}}\quad & R_{\mathrm{ex}} && (13a)\\
\text{s.t.}\quad & R_{\mathrm{ex}}\geqslant Q_1, && (13b)\\
& P_E\geqslant Q_E, && (13c)\\
& \sum_{i=1}^{3}a_i\leqslant 1, && (13d)\\
& a_i\geqslant 0,\ \forall i, && (13e)\\
& 0\leqslant\Theta_m\leqslant 2\pi,\ \forall m, m=1,2,\ldots,M. && (13f)
\end{aligned}$$

and

$$\begin{aligned}
(\mathrm{P3}):\max_{\boldsymbol{\alpha},\boldsymbol{\Theta}}\quad & R_{\mathrm{in}} && (14a)\\
\text{s.t.}\quad & R_{\mathrm{in}}\geqslant Q_2, && (14b)\\
& R_{i\to i}\leqslant R_{j\to i}, 1\leqslant j\leqslant i\leqslant 3, && (14c)\\
& \sum_{i=1}^{3}a_i\leqslant 1, && (14d)\\
& a_i\geqslant 0,\ \forall i, && (14e)\\
& 0\leqslant\Theta_m\leqslant 2\pi,\ \forall m, m=1,2,\ldots,M, && (14f)
\end{aligned}$$

where (P2) and (P3) are the external and internal security problems, respectively.

4.1 External security problem

For P2, since the size of $\boldsymbol{\Theta}$ is difficult to be calculated accurately in the iterative process, we use the Gaussian random process and SVD method to find the estimated variable. We first simplify R_{ex} by setting $i=1$ condition as an example to show the process. As $\boldsymbol{\Theta}=\mathrm{diag}\{\mathrm{e}^{\mathrm{j}\Theta_1},\ldots,\mathrm{e}^{\mathrm{j}\Theta_m},\ldots,\mathrm{e}^{\mathrm{j}\Theta_M}\}$, $e_m=\mathrm{e}^{\mathrm{j}\Theta m}$ can be derived as $\boldsymbol{l}=[e_1,\ldots,e_M,1]^{\mathrm{H}}$. For $|\boldsymbol{h}_{\mathrm{IN}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{AI}}+h_{\mathrm{AN}}|^2$ in $R_{1\to 1}$, we use $\boldsymbol{q}=\mathrm{diag}(\boldsymbol{h}_{\mathrm{IN}}^{\mathrm{H}})\boldsymbol{h}_{\mathrm{AI}}$, and simplify $\sigma=\sigma'\sqrt{P}$. Thus, we have

$$\boldsymbol{H}_N=\begin{bmatrix}\boldsymbol{q}\boldsymbol{q}^{\mathrm{H}} & \boldsymbol{q}h_{\mathrm{AN}}\\ \boldsymbol{q}^{\mathrm{H}}h_{\mathrm{AN}} & 0\end{bmatrix},\quad \boldsymbol{L}=\boldsymbol{l}\boldsymbol{l}^{\mathrm{H}},$$

where $\boldsymbol{L}$ needs to satisfy $\boldsymbol{L}\geqslant 0$ and $\mathrm{rank}(\boldsymbol{L})=1$. We rewrite it as $\mathrm{Tr}(\boldsymbol{H}_N\boldsymbol{L})+|h_{\mathrm{AN}}|^2$. For $a_1P|\boldsymbol{g}_{\mathrm{IE}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{AI}}+g_{\mathrm{AE}}|^2$ in $\mathrm{SINR}_{e,1}$, we use $\boldsymbol{p}=\mathrm{diag}(\boldsymbol{g}_{\mathrm{IE}}^{\mathrm{H}})\boldsymbol{h}_{\mathrm{AI}}$. We define

$$\boldsymbol{G}_N=\begin{bmatrix}\boldsymbol{p}\boldsymbol{p}^{\mathrm{H}} & \boldsymbol{p}g_{\mathrm{AE}}\\ \boldsymbol{p}^{\mathrm{H}}g_{\mathrm{AE}} & 0\end{bmatrix}.$$

We rewrite it as $\mathrm{Tr}(\boldsymbol{G}_N\boldsymbol{L})+|g_{\mathrm{AE}}|^2$. Since the rank-one constraint is non-convex, we can use the SDR technique to relax (P2) as

$$\begin{aligned}
(\mathrm{P2.1}):\max_{\boldsymbol{\alpha},\boldsymbol{L},Q_1}\quad & Q_1 && (15a)\\
\text{s.t.}\quad & \frac{\sigma'^2+a_1\mathrm{Tr}(\boldsymbol{H}_N\boldsymbol{L})+|h_{\mathrm{AN}}|^2}{\sigma'^2+a_1\mathrm{Tr}(\boldsymbol{G}_N\boldsymbol{L})+|g_{\mathrm{AE}}|^2}\geqslant 2^{Q_1}, && (15b)\\
& \mathrm{Tr}(\boldsymbol{G}_N\boldsymbol{L})\geqslant Q_E, && (15c)
\end{aligned}$$

$$\boldsymbol{L} \geqslant 0, \boldsymbol{L}_{m,m} == 1, \tag{15d}$$

$$\sum_{i=1}^{3} a_i \leqslant 1, \tag{15e}$$

$$a_i \geqslant 0, \ \forall i, \tag{15f}$$

$$0 \leqslant \Theta_m \leqslant 2\pi, \ \forall m, \tag{15g}$$

where (15b) is equivalent to (13b). $\boldsymbol{L}$ is still difficult to be optimized even though we have relaxed this problem. To deal with this problem, we utilize the Charnes-Cooper technique to reformulate it as a semi-definite programming (SDP) problem. Define

$$t = \frac{1}{\sigma'^2 + a_1 \mathrm{Tr}(\boldsymbol{G}_N \boldsymbol{L})}, \quad [\boldsymbol{T}_i]_{j,k} = \begin{cases} 1, & j = k = i, \\ 0, & \text{otherwise.} \end{cases}$$

Then, let $\boldsymbol{X} = t\boldsymbol{L}$. Since the high-rank optimization problem is still non-convex, we use SDR to relax it again by ignoring the rank-one constraint. We can transform the original problem as

$$(\text{P2.2}): \max_{\boldsymbol{\alpha}, t, \boldsymbol{T}} \quad Q_1 \tag{16a}$$

$$\text{s.t.} \quad t\sigma'^2 + \mathrm{Tr}(\boldsymbol{X}\boldsymbol{H}_N) + |h_{\mathrm{AN}}|^2 \geqslant 2^{Q_1}, \tag{16b}$$

$$t \geqslant \frac{1}{\sigma'^2 + Q_E}, \tag{16c}$$

$$\mathrm{Tr}(\boldsymbol{T}_i \boldsymbol{X}) = t, \quad \forall i, i = 1, 2, \ldots, N_t + 1, \tag{16d}$$

$$\sum_{i=1}^{3} a_i \leqslant 1, \tag{16e}$$

$$a_i \geqslant 0, \ \forall i, \tag{16f}$$

$$\boldsymbol{X}, \ t \geqslant 0. \tag{16g}$$

This convex problem can be solved by the CVX tool [37], and the Gaussian randomization method can be used to obtain an approximate solution to cope with this condition.

Since the phase shift $\boldsymbol{L}$ is too complex in the solution process, the SVD technology can also be utilized. Therefore, we can decompose $\boldsymbol{L}$, and then use fewer computation resources to achieve better results. Specifically, $\boldsymbol{L}$ can be written as $\boldsymbol{L} = \boldsymbol{V}\mathrm{diag}(\lambda)\boldsymbol{V}^{-1}$ and $\widetilde{\boldsymbol{L}} = \boldsymbol{U}\sqrt{\mathrm{diag}}(\lambda)\boldsymbol{V}^{\mathrm{T}}$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrixes. We can estimate $\boldsymbol{L}$ by

$$\widehat{\boldsymbol{L}} = \widehat{\boldsymbol{l}}\,\widehat{\boldsymbol{l}}^{\mathrm{H}}, \quad \widehat{\boldsymbol{l}} = \mathrm{e}^{\mathrm{j}\widehat{\Theta}}, \quad \widehat{\Theta} = \left\{ \frac{\widehat{\boldsymbol{l}}}{\widehat{\boldsymbol{l}}_{M+1}} \right\}_{(1:M)}, \tag{17}$$

where $\{\boldsymbol{x}\}_{(1:M)}$ means that the vector $\boldsymbol{x}$ has M elements. Based on the above discussions, we only need to process each $\widehat{\boldsymbol{L}}$ when it is optimized at every iteration. It is much easier than solving the original problem (P2.2).

4.2 Internal security problem

In internal secrecy rate optimization P3, we notice that (14b) and (14c) are non-convex constraints because of two coupled variables $\boldsymbol{\alpha}$ and $\boldsymbol{\Theta}$. To process it approximately, we define $\boldsymbol{\alpha}^n$ and $\boldsymbol{\Theta}^n$ as their values after the n-th algorithmic iteration. To decouple $\boldsymbol{\alpha}$ and $\boldsymbol{\Theta}$, we divide the original problem P3 into two subproblems: phase shift problem (P3.1) and power allocation problem (P3.2). In the internal phase shift problem, after relaxing the problem, we simplify it by using the SDR technology. To find the optimal value in each iteration, we exploit the bisection search technique. Then, we introduce the Lagrange function to optimize $\boldsymbol{\alpha}$.

4.2.1 *Internal phase shift optimization*

In each iteration $n = 1, 2, \ldots, \infty$, we treat $\boldsymbol{\alpha}$ as a given power allocation coefficient. Therefore, $\boldsymbol{\Theta}$ is optimized as

$$\begin{aligned}(\text{P3.1}): \max_{\boldsymbol{\Theta}, Q_2} \quad & Q_2 && \text{(18a)}\\ \text{s.t.} \quad & R_{\text{in}} \geqslant Q_2, && \text{(18b)}\\ & R_{i\to i} \leqslant R_{j\to i}, 1 \leqslant j \leqslant i \leqslant 3, && \text{(18c)}\\ & 0 \leqslant \Theta_m \leqslant 2\pi, \ \forall m, m = 1, 2, \ldots, M, && \text{(18d)}\end{aligned}$$

where (18b) can be expressed as

$$R_{\text{in}} = \log_2 \left(\frac{\sigma'^2 + a_1 \left| \boldsymbol{h}_{\text{IF}_2}^{\text{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\text{AI}} \right|^2}{\sigma'^2 + a_1 \left| \boldsymbol{h}_{\text{IF}_1}^{\text{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\text{AI}} \right|^2} \cdot \frac{\sigma'^2 + (a_1 + a_2) \left| \boldsymbol{h}_{\text{IF}_1}^{\text{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\text{AI}} \right|^2}{\sigma'^2 + (a_1 + a_2) \left| \boldsymbol{h}_{\text{IF}_2}^{\text{H}} \boldsymbol{\Theta} \boldsymbol{h}_{\text{AI}} \right|^2} \right) \geqslant Q_2. \tag{19}$$

We first introduce matrix $\boldsymbol{L}$ that is same as in Subsection 4.1, and exploit the SDR technique to process this non-convex constraint. We define

$$\boldsymbol{H}_{F_1} = \begin{bmatrix} \boldsymbol{q}_1 \boldsymbol{q}_1^{\text{H}} & \boldsymbol{q}_1 \\ \boldsymbol{q}_1^{\text{H}} & 0 \end{bmatrix}, \quad \boldsymbol{H}_{F_2} = \begin{bmatrix} \boldsymbol{q}_2 \boldsymbol{q}_2^{\text{H}} & \boldsymbol{q}_2 \\ \boldsymbol{q}_2^{\text{H}} & 0 \end{bmatrix},$$

where $\boldsymbol{q}_1 = \text{diag}(\boldsymbol{h}_{\text{IF}_1}^{\text{H}}) \boldsymbol{h}_{\text{AI}}$ and $\boldsymbol{q}_2 = \text{diag}(\boldsymbol{h}_{\text{IF}_2}^{\text{H}}) \boldsymbol{h}_{\text{AI}}$. Through a mathematical transformation, Eq. (19) can be relaxed as

$$R_{\text{in}} = \log_2 \left(\frac{1 + \frac{a_2}{a_1} - \frac{a_2 \sigma'^2}{a_1^2 \text{Tr}(\boldsymbol{H}_{F_1} \boldsymbol{L})}}{1 + \frac{a_2}{a_1} - \frac{a_2 \sigma'^2}{a_1^2 \text{Tr}(\boldsymbol{H}_{F_2} \boldsymbol{L})}} \right) \geqslant Q_2, \tag{20}$$

which can be further simplified as

$$\frac{2^{Q_2}}{\text{Tr}(\boldsymbol{H}_{F_2} \boldsymbol{L})} - \frac{1}{\text{Tr}(\boldsymbol{H}_{F_1} \boldsymbol{L})} \geqslant \eta_a (2^{Q_2} - 1), \tag{21}$$

where $\eta_a = (a_1 + a_2) a_1 / a_2$. The problem will be rewritten as

$$\begin{aligned}(\text{P3.1A}): \max_{\boldsymbol{L}, Q_2} \quad & Q_2 && \text{(22a)}\\ \text{s.t.} \quad & \frac{2^{Q_2}}{\text{Tr}(\boldsymbol{h}_{F_2} \boldsymbol{L})} - \frac{1}{\text{Tr}(\boldsymbol{h}_{F_1} \boldsymbol{L})} \geqslant \eta_a (2^{Q_2} - 1), && \text{(22b)}\\ & R_{i\to i} \leqslant R_{j\to i}, 1 \leqslant j \leqslant i \leqslant 3, && \text{(22c)}\\ & \boldsymbol{L} \geqslant 0, \ \text{rank}(\boldsymbol{L}) = 1. && \text{(22d)}\end{aligned}$$

We exploit the bisection search technique to find the final $\boldsymbol{\Theta}$ and Q_2 during the n-th iteration. First, we artificially determine the values of $Q_{2\max}$ and $Q_{2\min}$. Using $Q_2 = \frac{Q_{2\max} + Q_{2\min}}{2}$ to drive Q_2 to solve (P3.2). The updates of $Q_{2\max}$ and $Q_{2\min}$ depend on whether (P3.1A) can be solved or a feasible $\boldsymbol{L}$ can be found based on the proposed algorithm. Thus, through the n-th iteration, Q^n and $\boldsymbol{L}^n$ can be derived. Besides, $\boldsymbol{L}$ will be calculated as a feasible solution by the Gaussian randomization method in Subsection 4.1, and then we can obtain $\boldsymbol{\Theta}$ by $\boldsymbol{L}$.

4.2.2 *Internal power allocation optimization*

Then, we focus on the power allocation problem and the optimization of $\boldsymbol{\alpha} = \{a_1, a_2, a_3\}$. In this problem, we should find an optimal solution based on two variables. Therefore, we utilize the Lagrange multiplier method to optimize $\boldsymbol{\alpha}$ with given $\boldsymbol{\Theta}$ in each iteration. First, we rewrite the problem as

$$\begin{aligned}(\text{P3.2}): \max_{\boldsymbol{\alpha}} \quad & R_{\text{in}} && \text{(23a)}\\ \text{s.t.} \quad & R_{\text{in}} \geqslant Q_2, && \text{(23b)}\end{aligned}$$

$$R_t \geqslant R_{i\to i}, \forall i, \tag{23c}$$

$$\sum_{i=1}^{3} a_i \leqslant 1, \tag{23d}$$

$$a_i \geqslant 0, \ \forall i, \tag{23e}$$

where $R_t = \min\{R_{1\to 3}, R_{1\to 2}, R_{2\to 3}\}$. To find the local optimization solution of R_{in}, we consider α_1, α_2, and α_3 as three unknown numbers x, y, and z. Thus, Eqs. (23b)–(23d) can be transformed to three Lagrangian constraint equations.

$$\begin{cases} f(x,y,z) = R_{\text{in}}(x,y,z) - Q_2, \\ \psi(x,y,z) = R_t(x,y,z) - R_{i\to i}(x,y,z), \\ \phi(x,y,z) = x + y + z - 1. \end{cases} \tag{24}$$

We introduce Lagrange multipliers λ and μ and derive a Lagrange function as

$$F(x,y,z) = f(x,y,z) + \mu\psi(x,y,z) + \lambda\phi(x,y,z). \tag{25}$$

We can solve the above equations by letting the following derivatives equal to zero as

$$\begin{cases} F_x = f_x(x,y,z) + \mu\psi_x(x,y,z) + \lambda\phi_x(x,y,z) = 0, \\ F_y = f_y(x,y,z) + \mu\psi_y(x,y,z) + \lambda\phi_y(x,y,z) = 0, \\ F_z = f_z(x,y,z) + \mu\psi_z(x,y,z) + \lambda\phi_z(x,y,z) = 0, \\ F_\mu = \psi(x,y,z) = 0, \\ F_\lambda = \phi(x,y,z) = 0. \end{cases} \tag{26}$$

Then, we obtain x_0, y_0, z_0, μ and λ, and find the potential extreme three-dimensional point (x_0, y_0, z_0). If we determine it as the extreme point, we can calculate the extreme value of $f(x,y,z)$. Afterwards, Q_2 and optimized $\boldsymbol{\alpha}$ will be clearly presented. Otherwise, we go through the next iteration and repeat the procedures above.

4.3 Overall algorithm

Based on the simulated annealing algorithm, the overall algorithm is summarized in Algorithm 1. First, the appropriate initialization parameters are selected artificially. In our algorithm, n, k and i represent the iterations' numbers of computation, temperature, and the expected Q, respectively. $f(x)$ is defined as a disturbance function with an artificial error factor ϵ set. We also define a stability function as $T_f = g(Q_1, Q_2)$, which presents an undulant relationship between Q_1 and Q_2. The greater the fluctuation of Q_1 and Q_2 over some time, the higher the value of the stability function will be. Specifically, this function can be replaced by the classical inverse Markov function. When it is stable enough as required, the algorithm exits the operation. Next, Q^i is determined according to the dichotomy. Through adding disturbances by random number generators, Q^i after each iteration can moderately approach the optimal value. The Metropolis criteria indicates that the new results are accepted in terms of the energy and probability. The returned $\boldsymbol{L}^{i+1}$ is the estimated value of $\boldsymbol{\Theta}^{i+1}$ after SVD decomposition is applied. Then, we determine which part to optimize by comparing the internal and external secrecy rates. According to the process of simulated annealing, the computation will gradually cool down and be solved circularly. When the stability of two rates is the best and the current temperature is not higher than that set as T_k, we stop solving and get one of the local optimal solutions.

Compared with the traditional simulated annealing algorithm, our iterative algorithm introduces two user-defined functions related to optimization, connects the fluctuation relationship of iterative value with annealing temperature, and looks for the optimal solution globally as much as possible on the basis of fully optimizing the known model. The introduction of such heuristic algorithms and new solutions is helpful for the optimization of traditional IRS-NOMA. In the subsequent simulation, we also investigate the feasibility and efficiency of this algorithm.

Algorithm 1 Simulated annealing iteration algorithm

Initialize $\boldsymbol{\alpha}^i$, $\boldsymbol{\Theta}^i$, $Q_{1\max}$, $Q_{1\min}$, $Q_{2\max}$, $Q_{2\min}$, $n=0$, $k=0$, $i=0$.
Define $f(x)=x^{i*}-x^i-\epsilon$, initialize temperature $T_k=T_0$. Define stability function $T_f=g(Q_1,Q_2)$.
repeat
  Adjust Q^i_{xmax}, Q^i_{xmin} and $Q^i=\frac{Q^i_{\mathrm{xmax}}+Q^i_{\mathrm{xmin}}}{2}$;
  repeat
    Add disturbances Q^{i*} by random number generator and obtain new result Q_1^{i*}, Q_2^{i*};
    Compute $f(Q_1)$ and $f(Q_2)$ and solve the feasibility problem in (22);
    if (22) can be solved **then**
      if $f(Q_2)\geqslant 0$ **then**
        Accept the new result $Q^i_{2\min}=Q^{i*}$;
      else
        Accept new result according to Metropolis criteria $Q_{2\min}=Q^{0*}$;
      end if
    else
      $Q^i_{2\max}=Q^{i*}$;
    end if
    return $\boldsymbol{L}^{i+1}$;
    repeat
      Apply SVD technique to $\boldsymbol{L}^{i+1}$, obtain $\boldsymbol{\Theta}^{i+1}=\mathrm{diag}(\widehat{\Theta})$ according to Subsection 4.1;
    until $\boldsymbol{\Theta}^i$ is feasible for (22);
    return $\boldsymbol{\Theta}^{i+1}$;
    Solve (23) for given $\boldsymbol{\Theta}^{i+1}$ and obtain $\boldsymbol{\alpha}^{i+1}$;
    Compare Q_1^i and Q_2^i;
    if $Q_1^i\geqslant Q_2^i$ **then**
      Only optimize internal rate, solve (14), obtain Q_2^{i+1};
    else
      Only optimize external rate, solve (13), obtain Q_1^{i+1};
    end if
    $i=i+1$;
    if $\boldsymbol{\Theta}^{i+1}$, $\boldsymbol{\alpha}^{i+1}$, Q_1^{i+1}, Q_2^{i+1} are local optimal **then**
      $n=n+1$;
    end if
  until $n>n(T_k)$;
  $k=k+1$, lower the temperature T_k;
until $T_k\leqslant T_f$;
Finally obtain the global optimal solutions;
return $\boldsymbol{\Theta}=\boldsymbol{\Theta}^{i+1}$, $\boldsymbol{\alpha}=\boldsymbol{\alpha}^{i+1}$, $Q_1=Q_1^{i+1}$, and $Q_2=Q_2^{i+1}$.

Complexity analysis. The complexity of Algorithm 1 comes from two parts, which are from solving (13) and (14). Let L_1 and L_2 represent the iteration numbers of the two processes before successfully obtaining the approximate local optimal solutions, respectively. In (13), there are $M+3$ variables, 1 linear matrix inequality (LMI) constraint of size 1, $M+4$ linear constraints, and 1 logarithm inequality constraint. The computation complexity of the SDP problem is $\mathcal{O}((M+4)4.5)$, where $\mathcal{O}(\cdot)$ is the big-$\mathcal{O}$ notation. The complexity of solving (13) is

$$\mathcal{O}(n_1L_1\sqrt{M+n_1\log(n_1)+5}[n_1+2n_1(N+1)+2n_1^2\log(n_1)+M+(M+4)4.5+n_1^2+5]), \tag{27}$$

where $n_1=\mathcal{O}(M+3)$. In order to solve (14), there are $M+3$ variables, 4 logarithm inequality constraints and $M+4$ linear constraints. Besides, the SDR and the bisection search techniques are used with complexity of $\mathcal{O}((M+4)4.5)$ and $\mathcal{O}(\log(n_2))$, respectively. According to the same analysis above, the complexity is

$$\mathcal{O}(n_2L_2\sqrt{M+4\cdot n_2\log(n_2)+4}[8n_2^2N\log(n_2)+M+\mathcal{O}((M+4)4.5)+\log(n_2)+n_2^2+4]), \tag{28}$$

where $n_2=\mathcal{O}(M+3)$.

5 Simulation results

In this section, the performance of the IRS-assisted NOMA system is evaluated through numerical simulation with MATLAB. We assume that the BS-to-user channel is Rayleigh fading, and the path loss is determined by d^{-3}, where d is the distance. Both the BS-to-IRS and the IRS-to-user channels are seen as Rician fading, and the path loss is denoted as d^{-2}. Besides, the noise power is $\sigma^2=-74$ dBm. We assume that BS, U_1, U_2, U_3, and IRS are located at (0, 0, 0), (0, 20 m, 0), (0, 200 m, 0), (0, 250 m,

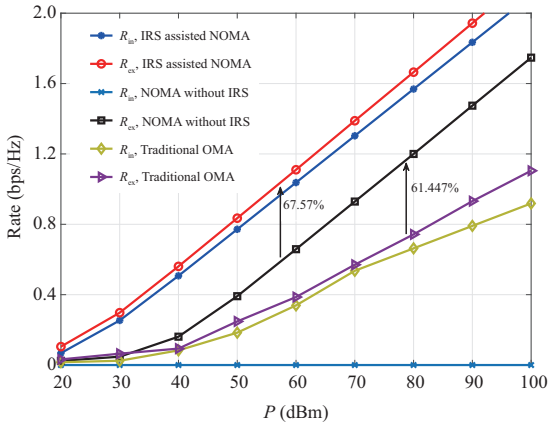

Figure 2 (Color online) The secrecy rate comparison with different benchmarks.

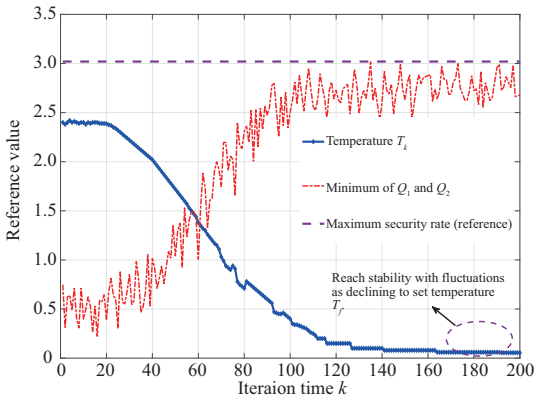

Figure 3 (Color online) Simulated annealing algorithm simulation results.

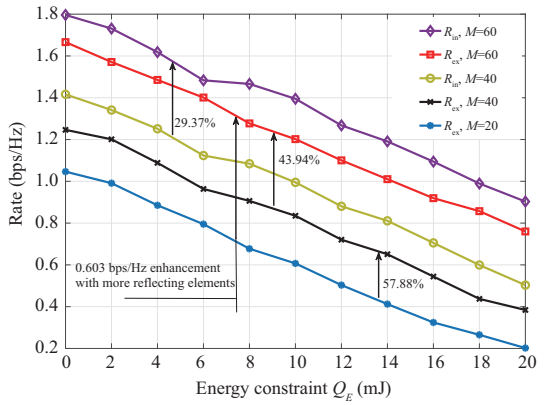

Figure 4 (Color online) Secrecy rate versus energy constraint at Eve.

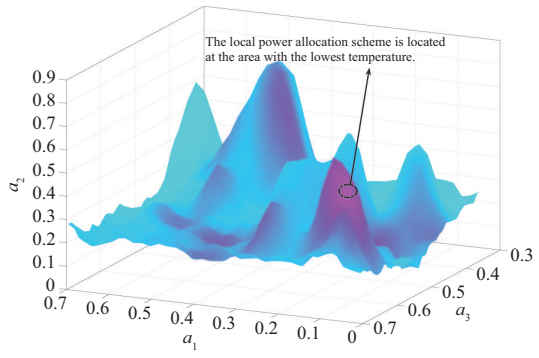

Figure 5 (Color online) The power allocation optimization result by the proposed algorithm.

0), and (0, 100 m, 50 m) in a 3D area, respectively. E is located around U_1, taking 10 m as the radius. Obstacles are filled in the area between (0, 20 m, 0) and (0, 200 m,0). In the iteration section, $M = 50$ and $T_0 = 150$°C. To study the auxiliary role of NOMA technology and IRS technology, we considered two benchmarks, i.e., traditional OMA [38], NOMA without IRS [39], and IRS-assisted NOMA [40]. Then, we investigate the transmission power and other constraints as the influence factors on the secrecy rate. We mainly focus on two kinds of secrecy rates value Q_1, Q_2, and system performance when energy constraint Q_E changes. The convergence and stability of the proposed algorithm are researched in the end.

Figure 2 plots the total secrecy rate performance of per-user versus the BS transmission power P. We focus on the impact of the existence of IRS and the difference between OMA and NOMA technology on overall security performance. First, we observe that the IRS-assisted NOMA achieves the best rate gain compared to the others. It is observed that the utilization of NOMA can improve around 61% secrecy rate because it gets rid of the limitation of orthogonality and each user can be allocated more resources compared to the traditional OMA. It is worth mentioning that if we are devoted to improving system security performance, we should optimize the phase shifts and IRS gains in IRS-far user links as much as possible. The combination of IRS and NOMA technology can significantly improve system performance. When IRS does not exist, there is no direct link from the BS to U_2 and U_3. Thus matrix $\boldsymbol{\Theta}$ is $\mathbf{0}$ in (6), (8) and (9), and R_{in} will be constantly 0. Besides, for the case of $P = 58$ dBm, the proposed IRS-assisted NOMA improves the system security performance by over 67% compared to the traditional NOMA without IRS and is higher than OMA.

In Figure 3, we plot some results of global optimization for the proposed simulated annealing algorithm. Due to the different units of temperature and rate, we use the idea of normalization to set their reference

values, so as to roughly reflect the changing trend and quantitative processing. Our goal is to maximize the minimum value of R_{ex} and R_{in} as much as possible, and get the optimal solution of the theory when they are stable enough. It can be seen from Figure 3 that with the progress of iteration, the secrecy rate experiences a rapid rise period before $k = 100$, then basically reaches stability despite fluctuations, and the temperature gradually reaches the set value of the stability function T_f. Therefore, this algorithm can achieve the optimal secrecy rate. Compared with the traditional IRS-NOMA optimization, although we partially simplify $\boldsymbol{\alpha}$ and $\boldsymbol{\Theta}$, we optimize them through the proposed algorithm. Therefore, $\boldsymbol{\alpha}$ and $\boldsymbol{\Theta}$ can be updated in the iterations which make the proposed algorithm have high flexibility and unity.

In Figure 4, we study the relationship between the energy constraint Q_E at Eve and the secrecy rate. In addition to the energy harvested by eavesdropper through physical devices, the number of reflecting elements M also affects the secrecy rate. With the increase of energy constraints, the secrecy rate decreases significantly. At the same time, under the same energy constraint, the number of elements of IRS is positively correlated with the secrecy rate. It is observed that this enhancement is higher especially within the low elements conditions. When M is adjusted from 20 to 40, there is roughly 58% increase on the external secrecy rate. This increase becomes lower with more reflecting elements equipped, and about 30% enhancement is achieved when M is changed from 40 to 60. It is worth noting that when the energy constraint of the eavesdropper is high enough, the system may even obtain a very small secrecy rate. At this time, although we optimize the phase shifts of the IRS and the power allocation of the BS, it is difficult to make the system endowed with sufficient external security. Therefore, when designing the communication network, we need to control the energy constraint Q_E within 20 mJ.

We present the optimization results of internal power allocation using our proposed algorithm in several iterations in Figure 5. In the three-dimensional space composed of $(\alpha_1, \alpha_2, \alpha_3)$, the color at each point represents the final temperature T_k value in Algorithm 1. The lower temperature is shown as the deeper blue point. We have identified a local power allocation scheme that is located at the area with the lowest temperature at the end of the iterative algorithm. According to our proposed algorithm, the point with lower temperature has higher accuracy and more optimized results. It can be seen that the area mentioned in the graph contains the locally optimized solutions. At the same time, we can also find the other optimal solutions area that are replaced by purple regions. It allows us to find an effective alternative solution when the best advantage is not appropriate, which increases the fitness and flexibility of Algorithm 1. Moreover, the method can also be extended to the case of multi-user scenario. However, when the number of users is large enough, the difficulty of system operation and processing increases remarkably, and the optimization performance becomes worse gradually.

6 Conclusion

In this paper, we proposed a novel priority-based secure transmission scheme for the IRS-aided NOMA networks against both the internal and external eavesdropping. First, according to the priority of all users, we formulated a secrecy rate maximization problem, where the NOMA transmission power and IRS's phase shifts were optimized. Second, to solve the above non-convex problem, we decoupled the original problem into the external security problem and the internal security problem. Then, to solve decoupled non-convex problems, we proposed an iterative algorithm, where the SDR technique, SVD technique, and Lagrange multiplier method were adopted to derive the power allocation and phase shifts. Finally, numerical results showed that the proposed scheme could improve the secrecy rate of all users compared with the existing IRS-aided secure NOMA schemes. Even though the multiple users' security priority levels and the external eavesdropping made it complex to allocate the transmission power and designed the phase shifts, the proposed scheme still could achieve a high secrecy rate.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 62271399, 62222107, 62071223, 62031012), Key Research and Development Program of Shaanxi Province (Grant No. 2022KW-07), Young Elite Scientist Sponsorship Program by China Association for Science and Technology, National Key Research and Development Program of China (Grant No. 2020YFB1807003), and Foundation of the Science, Technology, and Innovation Commission of Shenzhen Municipality (Grant No. JCYJ20190806160218174).

References

1 Ji X S, Huang K Z, Jin L, et al. Overview of 5G security technology. Sci China Inf Sci, 2018, 61: 081301
2 Saily M, Estevan C B, Gimenez J J, et al. 5G radio access network architecture for terrestrial broadcast services. IEEE Trans Broadcast, 2020, 66: 404–415

3 Lv L, Li Z, Ding H Y, et al. Secure NOMA and OMA coordinated transmission schemes in untrusted relay networks. Sci China Inf Sci, 2021, 64: 209302
4 Abbasi O, Ebrahimi A, Mokari N. NOMA inspired cooperative relaying system using an AF relay. IEEE Wireless Commun Lett, 2019, 8: 261–264
5 He Y, Wang D, Huang F, et al. Downlink and uplink sum rate maximization for HAP-LAP cooperated networks. IEEE Trans Veh Technol, 2022, 71: 9516–9531
6 Ding Z, Liu Y, Choi J, et al. Application of non-orthogonal multiple access in LTE and 5G networks. IEEE Commun Mag, 2017, 55: 185–191
7 Yadav A, Quan C, Varshney P K, et al. On performance comparison of multi-antenna HD-NOMA, SCMA, and PD-NOMA schemes. IEEE Wireless Commun Lett, 2021, 10: 715–719
8 Chen C, Zhong W D, Yang H L, et al. On the performance of MIMO-NOMA-based visible light communication systems. IEEE Photon Technol Lett, 2018, 30: 307–310
9 Huang R L, Wan D H, Ji F, et al. Performance analysis of NOMA-based cooperative networks with relay selection. China Commun, 2020, 17: 111–119
10 Li Y, Baduge G A A. Underlay spectrum-sharing massive MIMO NOMA. IEEE Commun Lett, 2019, 23: 116–119
11 Liu G, Wang Z Q, Hu J W, et al. Cooperative NOMA broadcasting/multicasting for low-latency and high-reliability 5G cellular V_2X communications. IEEE Int Things J, 2019, 6: 7828–7838
12 Kusaladharma S, Zhu W P, Ajib W, et al. Rate and energy efficiency improvements of massive MIMO-based stochastic cellular networks with NOMA. IEEE Trans Green Commun Netw, 2021, 5: 1467–1481
13 Wang D W, Zhou F H, Lin W S, et al. Cooperative hybrid nonorthogonal multiple access-based mobile-edge computing in cognitive radio networks. IEEE Trans Cogn Commun Netw, 2022, 8: 1104–1117
14 He C F, Hu Y, Chen Y, et al. Joint power allocation and channel assignment for NOMA with deep reinforcement learning. IEEE J Sel Areas Commun, 2019, 37: 2200–2210
15 You C S, Zheng B X, Zhang R. Fast beam training for IRS-assisted multiuser communications. IEEE Wireless Commun Lett, 2020, 9: 1845–1849
16 Wu Q Q, Zhang R. Towards smart and reconfigurable environment: intelligent reflecting surface aided wireless network. IEEE Commun Mag, 2020, 58: 106–112
17 Zhu J Y, Huang Y M, Wang J H, et al. Power efficient IRS-assisted NOMA. IEEE Trans Commun, 2021, 69: 900–913
18 Fang F, Xu Y Q, Pham Q V, et al. Energy-efficient design of IRS-NOMA networks. IEEE Trans Veh Technol, 2020, 69: 14088–14092
19 Jiao S Y, Fang F, Zhou X T, et al. Joint beamforming and phase shift design in downlink UAV networks with IRS-assisted NOMA. J Commun Inf Netw, 2020, 5: 138–149
20 Wang D W, He T M, Zhou F H, et al. Outage-driven link selection for secure buffer-aided networks. Sci China Inf Sci, 2022, 65: 182303
21 Wang D W, Wu M H, Wei Z X, et al. Uplink secrecy performance of RIS-based RF/FSO three-dimension heterogeneous networks. IEEE Trans Wireless Commun, 2023. doi: 10.1109/TWC.2023.3292073
22 Xie X M, Fang F, Ding Z G. Joint optimization of beamforming, phase-shifting and power allocation in a multi-cluster IRS-NOMA network. IEEE Trans Veh Technol, 2021, 70: 7705–7717
23 Wang H, Liu C, Shi Z, et al. On power minimization for IRS-aided downlink NOMA systems. IEEE Wireless Commun Lett, 2020, 9: 1808–1811
24 Lin C, Chang Q, Li X X. Uplink NOMA signal transmission with convolutional neural networks approach. J Syst Eng Electron, 2020, 31: 890–898
25 Li Y B, Zhang H J, Long K P. Joint resource, trajectory, and artificial noise optimization in secure driven 3-D UAVs with NOMA and imperfect CSI. IEEE J Sel Areas Commun, 2021, 39: 3363–3377
26 Wang D W, He Y X, Yu K P, et al. Delay-sensitive secure NOMA transmission for hierarchical HAP-LAP medical-care IoT networks. IEEE Trans Ind Inf, 2022, 18: 5561–5572
27 Wang D W, Wu M H, He Y X, et al. An HAP and UAVs collaboration framework for uplink secure rate maximization in NOMA-enabled IoT networks. Remote Sens, 2022, 14: 4501
28 Li N, Li M, Liu Y W, et al. Intelligent reflecting surface assisted NOMA with heterogeneous internal secrecy requirements. IEEE Wireless Commun Lett, 2021, 10: 1103–1107
29 Cao Y, Zhao N, Chen Y F, et al. Secure transmission via beamforming optimization for NOMA networks. IEEE Wireless Commun, 2020, 27: 193–199
30 Qiao J P, Alouini M S. Secure transmission for intelligent reflecting surface-assisted mmWave and terahertz systems. IEEE Wireless Commun Lett, 2020, 9: 1743–1747
31 Wijewardena M, Samarasinghe T, Hemachandra K T, et al. Physical layer security for intelligent reflecting surface assisted two-way communications. IEEE Commun Lett, 2021, 25: 2156–2160
32 Cao X L, Yang B, Huang C W, et al. Massive access of static and mobile users via reconfigurable intelligent surfaces: protocol design and performance analysis. IEEE J Sel Areas Commun, 2022, 40: 1253–1269
33 Cao X L, Yang B, Huang C W, et al. Reconfigurable intelligent surface-assisted aerial-terrestrial communications via multi-task learning. IEEE J Sel Areas Commun, 2021, 39: 3035–3050
34 Huang C, Zappone A, Alexandropoulos G C, et al. Reconfigurable intelligent surfaces for energy efficiency in wireless communication. IEEE Trans Wireless Commun, 2019, 18: 4157–4170
35 Wei L, Huang C, Alexandropoulos G C, et al. Channel estimation for RIS-empowered multi-user MISO wireless communications. IEEE Trans Commun, 2021, 69: 4144–4157
36 Wei L, Huang C, Alexandropoulos G C, et al. Multi-user holographic MIMO surfaces: channel modeling and spectral efficiency analysis. IEEE J Sel Top Signal Process, 2022, 16: 1112–1124
37 de Maio A, Huang Y, Palomar D P, et al. Fractional QCQP with applications in ML steering direction estimation for radar detection. IEEE Trans Signal Process, 2021, 59: 172–185
38 Hashemi R, Beyranvand H, Mili M R, et al. Energy efficiency maximization in the uplink delta-OMA networks. IEEE Trans Veh Technol, 2021, 70: 9566–9571
39 Kim B, Park Y, Hong D. Partial Non-orthogonal multiple access (P-NOMA). IEEE Wireless Commun Lett, 2019, 8: 1377–1380
40 Gong C H, Yue X W, Wang X Y, et al. Intelligent reflecting surface aided secure communications for NOMA networks. IEEE Trans Veh Technol, 2022, 71: 2761–2773