TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Nghiên cứu tối ưu chùm tia trong hệ thống truyền không dây có sự trợ giúp của bề mặt phản xạ thông minh

Nguyễn Đỗ Dũng*, **Nguyễn Văn Hào, Đào Minh Hưng**

Khoa Kỹ thuật và Công nghệ, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: dd/mm/yyyy

TÓM TẮT

Bài báo này nghiên cứu bài toán tối ưu hóa thiết kế chùm tia trong hệ thống truyền thông không dây có sự hỗ trợ của bề mặt phản xạ thông minh (IRS), nhằm đáp ứng các yêu cầu nghiêm ngặt của giao tiếp độ trễ thấp cực kỳ đáng tin cậy (URLLC). Bằng cách thiết kế các vectơ chùm tia truyền tại trạm nhỏ (SC) với các giả thiết khác nhau về các góc phản xạ của IRS. Từ đó, bài báo đề xuất một thuật toán, trong đó đó chuyển đổi bài toán tối ưu phi lồi thành dạng bài toán lồi có thể giải được, thông qua các kỹ thuật xấp xỉ lồi kế tiếp (SCA) và thư giãn bán xác định (SDR). Kết quả mô phỏng cho thấy phương pháp đề xuất giúp cải thiện đáng kể thông lượng hệ thống và tốc độ dữ liệu của người dùng so với các thuật toán tham chiếu, đặc biệt là trong các cấu hình IRS và mức công suất truyền khác nhau. Nghiên cứu khẳng định hiệu quả và khả năng mở rộng của IRS trong việc nâng cao hiệu suất mạng trong các trường hợp URLLC.

Từ khóa: *Bề mặt phản chiếu thông minh (IRS), mạng không đồng nhất, thiết kế chùm tia, giao tiếp độ trễ thấp cực kỳ đáng tin cậy (URLLC).*

*Nguyễn Đỗ Dũng
Email: nguyendodung@qnu.edu.vn

QUY NHON UNIVERSITY
JOURNAL OF SCIENCE

Research on beamforming optimization in intelligent reflective surface-assisted wireless communication systems

Do Dung Nguyen*, Van Hao Nguyen, Minh Hung Dao

Faculty of Engineering and Technology, Quy Nhon University, Vietnam

Received: dd/mm/yyyy

ABSTRACT

This paper investigates the optimization of beamforming in intelligent reflecting surface (IRS)-assisted heterogeneous wireless communication systems, aiming to meet the stringent requirements of ultra-reliable low-latency communication (URLLC). By designing the transmit beamforming vectors at the small cell (SC) with different assumptions on IRS reflection angles, the proposed algorithm transforms a non-convex optimization problem into a solvable convex form using successive convex approximation (SCA) and semi-definite relaxation (SDR) techniques. Simulation results demonstrate that the proposed method significantly enhances system throughput and user data rates compared to benchmark algorithms, especially under varying IRS configurations and transmission power levels. The study confirms the effectiveness and scalability of IRS in improving network performance in URLLC scenarios.

Keyworks: *Intelligent Reflecting Surface(IRS), Heterogeneous network, Beamforming design, Ultra-Reliable Low-Latency Communication (URLLC)*

1. INTRODUCTION

Future wireless networks are anticipated to evolve toward intelligent and software-reconfigurable architectures, enabling ubiquitous communication among humans and mobile devices. These networks will possess the capability to sense, control, and optimize the wireless environment, thereby supporting low-power operation, high throughput, massive connectivity, and ultra-low-latency communication[1]. As communication scenarios in the 6th Generation (6G) mobile communication era become increasingly complex, large-scale MIMO and small-cell deployments are receiving growing attention as key network architectures[2,3]. This is particularly relevant for practical communication scenarios in densely populated areas with high user density. Ultra-Reliable and Low-Latency Communications (URLLC) represents a critical standard for 6G networks[4], employing short-packet transmission to satisfy stringent reliability and latency requirements. URLLC is especially targeted at mission-critical applications, including industrial automation, remote healthcare, and intelligent transportation, and other scenarios that demand ultra-low latency and ultra-high reliability[5].

Intelligent Reconfigurable Surface (IRS) is a technology capable of reconfiguring the wireless propagation environment, acting like a passive metal mirror or "wave collector." It can be programmed to modify the impinging electromagnetic field in a customizable man-

*Do Dung Nguyen
Email: nguyendodung@qnu.edu.vn

ner. IRS is considered a highly promising innovation due to its immense potential to achieve low power consumption, high energy efficiency, high-speed communication, massive connectivity, and low-latency wireless communication[1]. Simultaneously, accurate and low-overhead channel estimation is a critical challenge in IRS-based systems due to the large number of IRS elements and their unique hardware constraints. Reference focuses on an IRS-enabled multi-user Multi-Input Single-Output (MISO) uplink communication system and proposes a channel estimation framework based on parallel factor decomposition to obtain the cascaded channel model. It presents two iterative estimation algorithms for the channel between the base station and the IRS, and the channel between the IRS and the users. It demonstrates that the total rate achieved using the estimated channel always reaches the total rate of the perfect channel under various settings, thereby verifying the effectiveness and robustness of the proposed estimation algorithms[6]. Reference[7] investigates secure beamforming in a multi-IRS assisted millimeter-wave (mmWave) system. The study jointly optimizes the transmit beamforming and IRS control to maximize the secrecy rate, subject to total transmit power and unit modulus constraints. It proposes an alternating optimization algorithm based on Successive Convex Approximation (SCA) and manifold optimization. The results show that the proposed algorithm effectively improves the secrecy rate compared to traditional schemes. Moreover, Advances such as joint position and beamforming optimization for movable IRS (using product Riemannian manifold optimization) further extend the role of IRS security and flexibility in next-generation mmWave networks[8]. Currently, much research combines IRS with URLLC or heterogeneous network scenarios. For instance, the authors in[9] investigates the design of a resource allocation algorithm for an IRS-assisted MISO Orthogonal Frequency Division Multiple Access (OFDMA) multi-cell network, where IRS is deployed to create a virtual link to enhance the communication channel and improve reliability for URLLC users with unfavorable propagation conditions. In reference[10], the article performed maximizing the ratio of the total data rate to the total cross-layer interference for femtocell users in a two-tier OFDMA heterogeneous network with layer interference under imperfect channel state information. The optimization problem is studied by jointly optimizing the femtocell base station's transmit power and subcarrier allocation factors. The original problem is transformed into a convex optimization problem using quadratic transformation, variable relaxation, and Lagrange duality theory, and the results indicate an improved interference efficiency compared to traditional schemes.

However, references[6–10] either focus solely on URLLC without considering heterogeneous networks, or they only address the Shannon capacity scenario. In situations where network conditions become complex, it is important to consider cases where Shannon capacity alone may not meet user demands. To address this, and to satisfy the quality of service requirements for short-packet communication users in heterogeneous networks, this paper proposes an alternating iterative optimization algorithm that is more aligned with practical application needs. The main contributions are as follows:

(1) We propose a novel beamforming optimization algorithm for IRS-assisted heterogeneous wireless networks that designs the transmit beamforming vectors at the small cell with different assumptions on IRS reflection angles of the IRS to meet URLLC requirements.

(2) The original non-convex optimization problem is effectively transformed into a solvable convex form using successive convex approximation (SCA) and semi-definite relaxation (SDR) techniques.

(3) We conduct extensive simulations to evaluate the performance of the proposed algorithm under various system configurations, demonstrating its superiority over benchmark methods such as random phase shift and no-IRS schemes.

2. SYSTEM MODEL

2.1. Method of signal transmission to users

As shown in Fig. 1, the system considered in this paper is a downlink heterogeneous network, where the overall communication system consists of a base station (BS), small cell (SC),

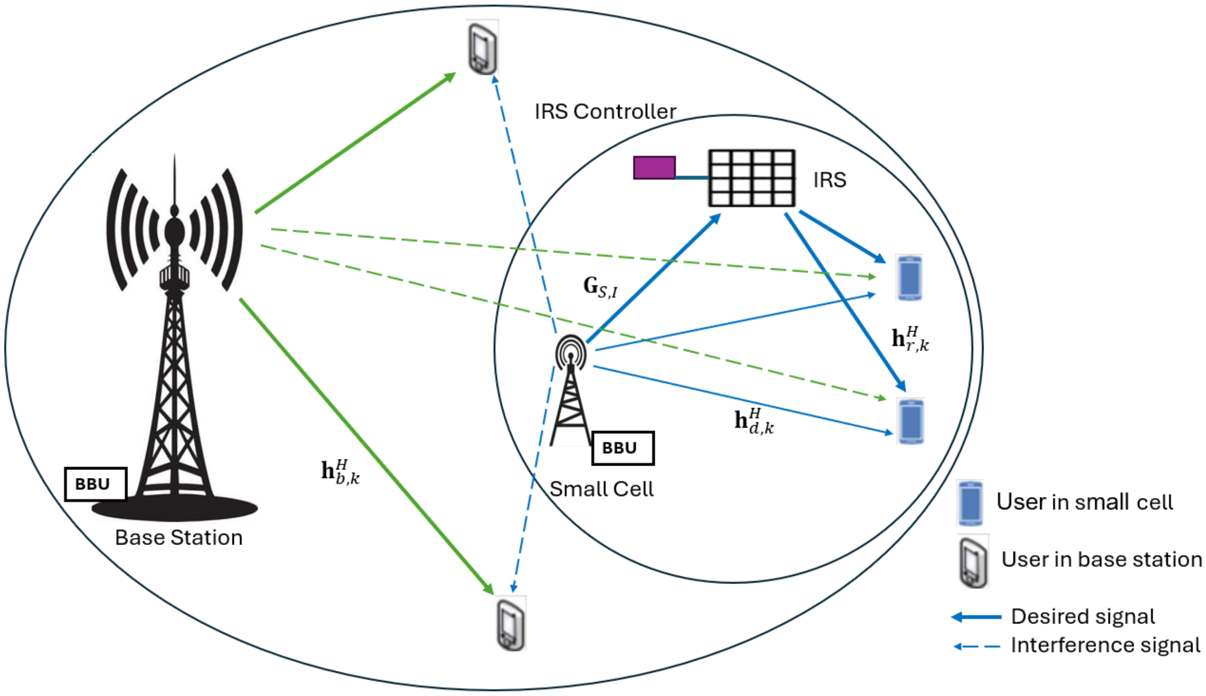

Figure 1. The heterogeneous network model with IRS support

users in BS, and users in SC assisted by intelligent reflecting surface (IRS) to meet the communication requirements of high reliability and low latency. Both the base station and small cell are equipped with multiple antennas $(M > 1)$. The total number of users in the communication area is $\mathcal{U}_B$, and the number of small cell users assisted by IRS is $\mathcal{U}_S = \{1, 2, \ldots, K\}$, with each user being a single-antenna user. The IRS is composed of $\mathcal{N} = \{1, 2, \ldots, N\}$ reflecting units, and the corresponding channel matrix is $\mathbf{\Phi} = \text{diag}(\mathbf{u}) \in \mathbb{C}^{N\times N}$, where $\mathbf{u} = [u_1, u_2, \ldots, u_N]^H = [e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_N}]^H \in \mathbb{C}^{N\times 1}$, with $\theta_i \in [0, 2\pi), \forall i \in \mathcal{N}$. The system operation involves the following coefficient channels: the transmission channel from SC station to IRS, $\mathbf{G}_{S,I} \in \mathbb{C}^{N\times M}$; reflective channel from IRS to users in SC k, $\mathbf{h}^h_{r,k} \in {}^{1\times N}$ and channel directly from SC to user in small cell, $\mathbf{h}_{d,k} \in \mathbb{C}^{1\times M}$ with $k \in \mathcal{U}_S$; the transmission channel from BS to user in base station k, $\mathbf{h}_{m,k} \in \mathbb{C}^{1\times M}$ with $k \in \mathcal{U}_B$. The IRS is equipped with an intelligent controller to coordinate its reception and reflection of signals during the transmission process, and it can obtain channel state information (CSI) from the SC to the IRS via a wireless connection[11]. Simultaneously, a basic baseband unit (BBU) is implemented between the BS and SC through a wireless backhaul, with both the SC and IRS controlled via wireless connection. It is assumed that the CSI of the entire system can be obtained at the BBU[12].

The signal transmitted from the BS station is given by

$$\mathbf{x}_B = \sum_{i\in\mathcal{U}_B} \mathbf{v}_i s_{b,i}, \tag{1}$$

where $\mathbf{v}_i \in \mathbb{C}^{M\times 1}$ is the beamforming vector for the i-th base station user transmitted from the BS, and $s_{b,i} \sim \mathcal{CN}(0, 1)$, $i \in \mathcal{U}_B$ is the data sent to the i-th base station user with $\mathbb{E}(|s_{m,i}|^2) = 1$. The transmission signal from the SC is expressed as:

$$\mathbf{x}_S = \sum_{k\in\mathcal{U}_S} \mathbf{w}_k s_{s,k}, \tag{2}$$

where $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$ is the beamforming vector for the k-th small cell user transmitted from the SC, and $s_{s,k} \sim \mathcal{CN}(0, 1)$, $k \in \mathcal{U}_S$, is the data sent to the k-th small cell user with $\mathbb{E}(|s_{s,k}|^2) = 1$. The maximum transmitted power at the SC, $P_{\max}$, is constrained by the following condition:

$$\sum_{k\in\mathcal{U}_S} \|\mathbf{w}_k\|^2 \le P_{\max}. \tag{3}$$

2.2. Maximum transmission rate for small cell users

To meet the high-reliability communication requirements and low-latency constraints, the received signal at the small cell users consists of three components: the desired signal from the SC, interference from the BS, and interference from other small cell users. The k-th small cell user is represented as

$$\begin{aligned} u_{s,k} = &(\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G}_{S,I}) \mathbf{w}_k s_{s,k} + \\ &\sum_{i \in \mathcal{U}_S, i \neq k} (\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G}_{S,I}) \mathbf{w}_i s_{s,i} + \\ &\mathbf{h}_{b,k}^H \mathbf{x}_B + n_k, \end{aligned} \tag{4}$$

where $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise (AWGN) at user k-th. Since this paper focuses on small cell users within the SC coverage area, the BS employs maximum ratio transmission (MRT) for precoding. The MRT beamforming vector[13] is given by $\mathbf{v}^{\text{MRT}} = \frac{\mathbf{h}_{b,k}}{\|\mathbf{h}_{b,k}\|}$. Thus, the signal to interference plus noise ratio (SINR) for a low-latency user k-th is:

$$\gamma_{s,k} = \frac{|(\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G}_{S,I}) \mathbf{w}_k|^2}{\sum_{i \in \mathcal{U}_S, i \neq k} |(\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G}_{S,I}) \mathbf{w}_i|^2 + \rho_k + \sigma_k^2}, \tag{5}$$

where $\rho_k = \sum_{i \in \mathcal{U}_B} |\mathbf{h}_{b,k}^H \mathbf{v}_i|^2$. Here $\mathbf{h}_{b,k}$ represents the channel from the BS to the k-th user in MS, and $\mathbf{v}_i$ denotes the beamforming vector associated with the i-th user.

The signal received by a base station user l-th also consists of three parts: the desired signal from the BS, the received multi-user interference signals from other users in BS, and the interference signals from the IRS. Thus, the received signal is denoted as:

$$\begin{aligned} u_{b,l} = &\mathbf{h}_{m,l}^H \mathbf{v}_l s_{m,l} + \sum_{i \in \mathcal{K}, i \neq l} \mathbf{h}_{m,l}^H \mathbf{v}_i s_{m,i} \\ &+ \mathbf{h}_{r,l}^H \mathbf{x}_s + n_l \end{aligned} \tag{6}$$

where $n_l \sim \mathcal{CN}(0, \sigma_l^2)$ is the additive white Gaussian noise for the l-th user. The SINR of the base station user l-th is given by

$$\gamma_{b,l} = \frac{|\mathbf{h}_{m,l}^H \mathbf{v}_l|^2}{\sum_{i \in (\mathcal{U}_B - \mathcal{U}_S), i \neq l} |\mathbf{h}_{m,l}^H \mathbf{v}_i|^2 + \sum_{j \in \mathcal{U}_S} |\mathbf{h}_{d,l}^H \mathbf{w}_j|^2 + \sigma_l^2} \tag{7}$$

The transmission rate of a small cell user (in bits per second per Hz) is approximated by equation[14]:

$$R(\gamma_{s,k}) = \log_2(1 + \gamma_k) - aQ^{-1}(\varepsilon) \sqrt{1 - (1 + \gamma_k)^{-2}} \tag{8}$$

This paper aims to maximize the system sum rate for $\mathcal{U}_S$ small cell users within the IRS-assisted region, while ensuring that their QoS[15] requirements are met. For base station users, outside the communication area, it is only necessary to satisfy the minimum SINR requirement. By jointly optimizing the transmit beamforming vectors $\{\mathbf{w}_k\}_{k \in \mathcal{U}_S}$ at the SC and the IRS reflection angles $\theta_i \in [0, 2\pi)$ for all $i \in \mathcal{N}$, the combined optimization problem can be formulated as:

$$\left.\begin{aligned} &\max_{\mathbf{w}_k, \mathbf{u}} \sum_{k \in \mathcal{U}_S} R(\gamma_{s,k}) \\ &\text{s.t.} \\ &\quad \text{C1: } R(\gamma_{s,k}) \geq L_k, \quad \forall k \in \mathcal{U}_S, \\ &\quad \text{C2: } \gamma_{b,l} \geq \text{SINR}_{BS}, \quad \forall l \in (\mathcal{U}_B - \mathcal{U}_S), \\ &\quad \text{C3: } \sum_{k \in \mathcal{U}_S} \|\mathbf{w}_k\|^2 \leq P_{\max}, \\ &\quad \text{C4: } |u_i| = 1, \quad \forall i \in \mathcal{N}. \end{aligned}\right\} \tag{9}$$

Constraint C1 in the optimization problem ensures the minimum QoS requirement for each small cell user; constraint C2 guarantees the minimum SINR requirement for base station users; and constraint C3 ensures that the transmission power at the SC does not exceed the maximum power requirement $P_{\max}$. Constraint C4 is the unit modulus constraint for the IRS elements. Problem (9) is a non-convex optimization problem, and its non-convexity mainly arises from the SINR in the objective function and the unit modulus constraint (C3). The transmission rate $R(\gamma_{s,k})$ is expressed in a non-linear form as shown in (8).

For most non-convex problems, there is currently no systematic solution. In the following sections of this paper, a series of approaches are proposed to tackle these problems, and an efficient alternative algorithm based on iterative optimization is presented.

3. DESIGNED ALGORITHM FOR OPTIMIZED BEAMFORMING

3.1. Optimization problem transformation

In this section, we solve optimization problems (9) with coupled variables to obtain optimal solution. This article focuses exclusively on the optimal beamforming at the SC. Therefore, we apply an optimization approach using a fixed-point iteration for the IRS phase shift vector $\mathbf{u}$. Finally, the optimization method is employed to solve the subproblems in (9).

For the processing $R(x)$ it can be expressed as $R(x) = f_1(x) - aQ^{-1}(\varepsilon)f_2(x)$, where $f_1(x)$ and $f_2(x)$ are two convex functions of x. $Q^{-1}(\varepsilon)$ is a positive constant, so the objective function $R(x)$ is the difference of two convex functions[16]. Using the first-order Taylor expansion of $f_2(x)$, we have:

$$\log_2(1 + x) - A_k x - B_k \tag{10}$$

where $A_k = aQ^{-1}(\varepsilon)f_2'(x^t)$, and $B_k = aQ^{-1}(\varepsilon)(f_2(x^t) - f_2'(x^t)x^t)$. Thus, the objective function for the optimal solution of this problem can be expressed in terms of the transmission rate in equation (10).

Next, we aim to optimize the beamforming vector $\mathbf{w}_k$ at the SC, assuming that the IRS reflection vector $\mathbf{u}$ is fixed. The following quantities are defined: $\mathbf{W}_k \triangleq \mathbf{w}_k\mathbf{w}_k^H; \mathbf{M}_k \triangleq \mathbf{m}_k\mathbf{m}_k^H; \mathbf{m}_k^H = \mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H\mathbf{\Phi}\mathbf{G}; \mathbf{m}_k = \mathbf{h}_{d,k} + \mathbf{G}^H\mathbf{\Phi}\mathbf{h}_{r,k}$. Therefore, the SINR in (5) can be rewritten as:

$$\gamma_{s,k} = \frac{\mathrm{Tr}(\mathbf{W}_k\mathbf{M}_k)}{\sum_{i\in\mathcal{U}_S, i\neq k}\mathrm{Tr}(\mathbf{W}_i\mathbf{M}_k) + \rho_k + \sigma_k^2}. \tag{11}$$

This formula expresses SINR as the ratio of the useful signal power to the sum of interference and noise powers for user k-th at the SC.

3.2. Optimization of the transmit beamforming vector at the SC

From (11), the optimization problem in expression (9) can be rewritten as follows:

$$\left.\begin{aligned}
&\max_{\mathbf{w}_k} \sum_{k\in\mathcal{U}_S} \log(1+\gamma_{s,k}) - A_k\gamma_{s,k} - B_x\\
&\text{s.t.}\\
&\quad \text{C1, C2, }\overline{\text{C3}}: \sum_{k\in\mathcal{U}_S}\mathrm{Tr}(\mathbf{W}_k) \le P_{\max},\\
&\quad \text{C5: } \mathbf{W}_k \ge 0, \quad \forall k\in\mathcal{U}_S,\\
&\quad \text{C6: } \mathrm{rank}(\mathbf{W}_k) \le 1, \quad \forall k\in\mathcal{U}_S.
\end{aligned}\right\} \tag{12}$$

Here, constraints C5 and C6 are introduced to ensure that $\mathbf{W}_k \triangleq \mathbf{w}_k\mathbf{w}_k^H$ remains valid after optimization. Next, an approximate treatment is made to the problem in expression (12). First, the objective function is expressed as the difference between two convex functions (DC programming)[17], $J_1 - J_2$, where

$$J_1 = \sum_{k\in\mathcal{U}_S}\log_2\left(\sum_{i\in\mathcal{U}_S}\mathrm{Tr}(\mathbf{W}_i\mathbf{M}_k) + \rho_k + \sigma_k^2\right) \tag{13}$$

and

$$J_2 = \sum_{k\in\mathcal{U}_S}\log_2\left(\sum_{i\in\mathcal{U}_S, i\neq k}\mathrm{Tr}(\mathbf{W}_i\mathbf{M}_k) + \rho_k + \sigma_k^2\right). \tag{14}$$

For any positive semi-definite matrix $\mathbf{W}_k, k\in\mathcal{U}_S$ where t represents the iteration index, we use the first-order Taylor expansion for J_2 at $\mathbf{W}_k^{(t_1)}$:

$$J_2(\mathbf{W}_k) \le J_2(\mathbf{W}_k^{(t_1)}) + \sum_{k\in\mathcal{U}_S}\mathrm{Tr}(\nabla_{\mathbf{W}_k}^H J_2(\mathbf{W}_k^{(t_1)})(\mathbf{W}_k - \mathbf{W}_k^{(t_1)})) \tag{15}$$

where $\nabla_{\mathbf{W}_k}^H J_2(\mathbf{W}_k^{(t_1)})$ is the gradient of J_2 with respect to $\mathbf{W}_k$. Therefore, the objective function can be rewritten as:

$$\sum_{k\in\mathcal{U}_S}\log_2(1+\gamma_k) \ge J_1 - J_2(\mathbf{W}_k^{(t_1)}) - \sum_{k\in\mathcal{U}_S}\mathrm{Tr}(\nabla_{\mathbf{W}_k}^H J_2(\mathbf{W}_k^{(t_1)})(\mathbf{W}_k - \mathbf{W}_k^{(t_1)})) \triangleq N_k \tag{16}$$

Since constraint C6 is also a non-convex constraint, to solve this type of problem, the semi-definite relaxation (SDR) technique[18] is employed. We define J_3 and J_4 to represent the following expressions:

$$J_3 = \mathrm{Tr}(\mathbf{W}_k \mathbf{M}_k) \tag{17}$$

and

$$J_4 = \sum_{j \in \mathcal{K}, j \neq k} \mathrm{Tr}(\mathbf{W}_j \mathbf{M}_k) + \rho_k + \sigma_k^2. \tag{18}$$

Both J_3 and J_4 are function of $\mathbf{W}_k$. According to reference[19], for $x > 0$, where $x \in \mathbb{R}$, x can be treated as a specific value after each iteration, so the following inequality (19) holds:

$$\frac{1}{x} \geq \frac{2}{\bar{x}} - \frac{x}{\bar{x}^2} \tag{19}$$

Substituting $\gamma_k = \frac{J_3}{J_4}$, we get:

$$\frac{J_3}{J_4} \geq J_3 \left(\frac{2}{\bar{J}_4^{(t_1)}} - \frac{J_4}{(\bar{J}_4^{(t_1)})^2} \right) = \frac{2J_3}{\bar{J}_4^{(t_1)}} - \frac{J_3 J_4}{(\bar{J}_4^{(t_1)})^2} \tag{11}$$

In particular, to approximate the product $J_3 J_4$, we write:

$$J_3 J_4 = \frac{(J_3 + J_4)^2}{4} - \frac{(J_3 - J_4)^2}{4}. \tag{20}$$

Similarly, a first-order Taylor expansion is applied to $(J_3 - J_4)^2$, after this treatment, the result is as follows:

$$\begin{aligned} \frac{J_3}{J_4} =& \frac{2J_3}{\bar{J}_4} - \frac{1}{\bar{J}_4^2} \left(\frac{(J_3 + J_4)^2}{4} - \frac{1}{4} (\bar{J}_4 - \bar{J}_3)^2 \right) \\ &+ 2(\bar{J}_4 - \bar{J}_3) \mathbf{M}_k (\mathbf{W}_k - \mathbf{W}_k^{(t_1)}), \end{aligned} \tag{21}$$

where $\bar{J}_4 = \sum_{i \in \mathcal{U}_S, i \neq k} \mathrm{Tr}(\mathbf{W}_i^{(t_1)} \mathbf{M}_k) + \rho_k + \sigma_k^2$, and $\bar{J}_3 = \mathrm{Tr}(\mathbf{W}_k^{(t_1)} \mathbf{M}_k)$. Furthermore, a Taylor approximation is applied to the convex function $(J_3 + J_4)^2$, and by rearranging inequality (11), the following result is obtained:

$$\begin{aligned} \frac{J_3}{J_4} \geq \frac{2J_3}{\bar{J}_4} - \frac{1}{\bar{J}_4^2} \Big(& \bar{J}_3 \bar{J}_4 + \\ 2\bar{J}_4 \mathrm{Tr} \left(\mathbf{M}_k \left(\mathbf{W}_k - \mathbf{W}_k^{(t_1)} \right) \right) & \Big) \triangleq J_{3,4} \end{aligned} \tag{22}$$

After this processing, the ratios of the two convex functions J_3/J_4. Thus, the objective function in expression (12) becomes:

$$\sum_{k \in \mathcal{U}_S} R(\gamma_{s,k}) = \sum_{k \in \mathcal{U}_S} N_k - A_k J_{3,4} - B_k. \tag{23}$$

Next, constraint C2 needs to be addressed. According to the minimum SINR requirement for macro-cell users, MRT precoding is applied for processing. Therefore, constraint C2 can be rewritten as

$$\begin{aligned} \overline{C2}: \sum_{k \in \mathcal{U}_S} \mathrm{Tr}(\mathbf{W}_k \mathbf{H}_{r,l}) + \rho_l + \sigma_l^2 + \leq \frac{1}{\mathrm{SINR_{BS}}}, \\ \forall l \in (\mathcal{U}_B - \mathcal{U}_S). \end{aligned} \tag{24}$$

Due to $\mathbf{W}_k \triangleq \mathbf{w}_k \mathbf{w}_k^{\mathrm{H}}$, it is included in constraint C6. To solve the problem, the semi-definite relaxation (SDR) technique is applied, and the original optimization problem (12) is reformulated as:

$$\left. \begin{aligned} &\max_{\mathbf{W}_k} \sum_{k \in \mathcal{U}_S} N_k - A_k J_{3,4} - B_k \\ &\text{s.t.} \\ &\quad \mathrm{C1}, \overline{\mathrm{C2}}, \overline{\mathrm{C3}}: \sum_{k \in \mathcal{U}_S} \mathrm{Tr}(\mathbf{W}_k) \leq P_{\max}, \ \forall k \in \mathcal{U}_S \\ &\quad \mathrm{C5}: \mathbf{W}_k \geq 0, \ \forall k \in \mathcal{U}_S \end{aligned} \right\} \tag{25}$$

Expression (25) is now a convex optimization problem that can be solved using standard convex optimization tools. To obtain the optimal solution, an iterative algorithm is employed to solve expression (25) for $\mathbf{W}_k$ across time slots t. The proposed algorithm is summarized in Table 1.

Table 1: Iterative beamforming vector design algorithm based on successive convex approximation (SCA)

Initialize the maximum number of iterations, $t^{\max}$, the iteration index t_1, and the variable set $\{\mathbf{W}_k^{(t_1)}\}$; set the optimization variables $\mathbf{u} = [u_1, u_2, \ldots, u_N]^H$ as a constant.

1. For i = 1, 2,..., do:

2. Under the given values of $\mathbf{W}_k^{(t_1)}$ and $\mathbf{u}$, solve problem (25) to obtain $\mathbf{W}_k^{(t_1+1)}$.

3. Update the iteration index: $t_1 = t_1 + 1$

4. Repeat steps 2–3 until convergence or until $t_1 = t_1^{\max}$

5. End for

4. NUMERICAL RESULTS

This section presents the simulation results for the proposed system and algorithm. The study considers all communication participants in an IRS-assisted heterogeneous network, which are assumed to be located on the same two-dimensional plane assumed as follows: i) Network Setup: Within the small cell coverage area, there are 2 IRS-assisted small cell users and 2 base station users. The base station region is centered at $(0, 0)$ meters, with the small cell base station located at $(20, 0)$ meters and the base station at $(-150, 0)$ meters. The IRS is positioned at $(70, 70)$ meters; ii) User distribution: Small cell users are randomly distributed within a circle of radius 5 meters centered at $(0, 80)$ meters. Base station users are randomly distributed within a circle of radius 5 meters centered at $(0, 25)$ meters; iii) System parameters: The noise power spectral density is set to -175 dBm/Hz, and the system bandwidth is 240 kHz. The channel is modeled with a Rician factor of 10 and a low-latency target decoding error probability of 10^{-5}. Each transmitted signal contains $L_k = 11$ bits. The system is equipped with 4 antennas at the base stations, and the IRS is configured with $N =$ 16, 36, or 64 reflecting elements. For base station users, the minimum SNR requirement is set to 15 dBm. In this study, the performance evaluation focuses on the average system throughput and the average data rate of small cell users, which are referred to as the system throughput and average user rate, respectively.

To evaluate and compare the performance of the proposed algorithm, two benchmark algorithms are considered:

(1) Random Phase Shift Algorithm: In this method, each IRS element is assigned a random phase shift. Specifically, for the optimization variable $\mathbf{u}$, a random number is selected uniformly from the interval $[0, 2\pi)$ and fixed at the beginning of the simulation. Subsequently, only the beamforming $\mathbf{W}_k$ is optimized.

(2) No-IRS Algorithm: In this approach, the channels associated with the IRS, namely $\mathbf{G}_{S,I}$ and $\mathbf{h}_{r,k}$, as well as the phase shift $\mathbf{\Phi}$, are removed. Optimization is then performed solely on the beamforming $\mathbf{W}_k$.

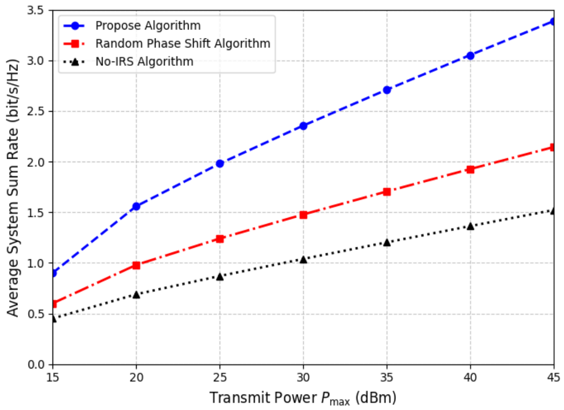

Figure 2. Relationship between P_{max} of the Small Cell (SC) and the Average Sum System Rate, with $K = 2$ and $N = 16$.

Fig. 2 illustrates the relationship between the maximum transmitted power (P_{max}) of small cell and the average system sum rate. As P_{max} increases, the average system sum rate shows a monotonically increasing trend. This behavior is influenced by constraint C3 in the optimization problem, where P_{max} directly impacts the feasible range of the beamforming optimization variable $\mathbf{W}_k$. At $P_{max} = 15$ dBm, the three algorithms exhibit different performance characteristics: the proposed optimization algorithm yields the highest average system sum rate, followed by the Random phase shift algorithm, while the "without IRS" case performs the worst. In the range from 15 to 40 dBm, there is a noticeable gap in the rate of increase between the proposed algorithm and the other two. However, the magnitude of this difference remains modest. The performance trend of the random phase shift algorithm closely follows that of the "without IRS" case. For P_{max} in the range of 40 to 45 dBm, the differences among the three algorithms become more pronounced. The proposed algorithm begins to exhibit a steeper growth rate compared to the other two. Similarly, the Random Phase Shift algorithm shows an increase in its growth rate relative to the "without IRS" case, which continues to show the smallest rate of increase.

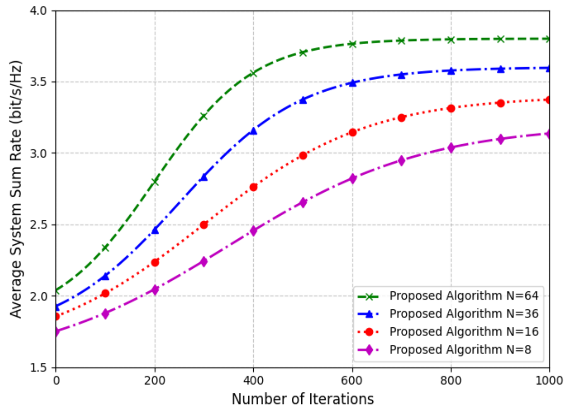

Figure 3. The relationship between number of iterations and the system's average sum rate with N = 8, 16, 36, 64, P_{max} = 30 dBm

Fig. 3 successfully validates the proposed algorithm's convergence and underscores the pivotal role of the IRS in enhancing network capacity. Specifically, Fig. 3 illustrates the relationship between the number of iterations and the average system sum rate (in bits/s/Hz) for a proposed algorithm under varying values of IRS reflecting elements N = 8, 16, 36, 64, with a fixed transmit power $P_{\max}$ = 30 dBm. The results clearly indicate that the proposed algorithm benefits significantly from increased IRS element counts. As N increases, the system sum rate improves across all iteration counts. For instance, at 1000 iterations, the algorithm with N = 64 achieves approximately 3.6 bits/s/Hz, while the configuration with N = 8 reaches only around 2.5 bits/s/Hz.In addition, the convergence of the algorithm is evident: all curves exhibit rapid initial growth in the system sum rate, particularly within the first 400 iterations. Beyond this point, the rate of improvement diminishes, indicating convergence toward a performance ceiling. The proposed algorithm demonstrates robust scalability and convergence properties. The consistent upward trend across all configurations confirms its effectiveness in iterative optimization settings.

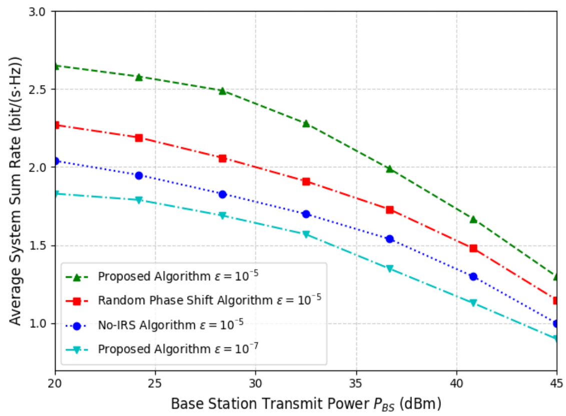

Figure 4. The relationship between the maximum transmitted power of base station (P_{BS}) and the system's average sum rate with K = 2, P_{max} = 30 dBm

From Fig. 4, it can be observed that as transmitted power at the base station P_{BS} increases, the average system sum rate decreases for all three algorithms. The influence of the BS on the average system sum rate for the SC users can be understood from expressions (9) and (12). In (9), the increase in BS transmission power results in a larger SC component in the denominator, thereby increasing interference for the SC users. This leads to a decrease in the SINR of the SC users, which in turn lowers the average system sum rate. Additionally, for the proposed algorithm, when the transmitted power at the BS is relatively low (e.g., between 20 and 32 dBm), the enhancement effect of the IRS remains significant, and as a result, the average system sum rate decreases more gradually. However, once the transmitted power at the BS exceeds the 32 dBm threshold, the excessive increase in BS power exacerbates its impact on SINR, causing a sharp acceleration in the rate of decrease of the average system sum rate. Moreover, it can be noted that the proposed algorithm does not always outperform the random phase shift and no-IRS cases. When the decoding error probability is set to $\varepsilon = 10^{-7}$, the average system sum rate is lower compared to all three algorithms when $\varepsilon = 10^{-5}$. This indicates that, as the transmitted power at the BS increases, the change in ε has a more substantial impact on the average system sum rate than in the random phase shift and "No-IRS" scenarios.

5. CONCLUSION

In this study, we proposed an efficient beamforming optimization algorithm for IRS-assisted heterogeneous wireless networks, targeting the enhancement of system throughput under URLLC constraints. By reformulating the original nonconvex problem using convex approximation techniques and iterative optimization, the algorithm achieves notable improvements in average system sum rate and convergence speed. Simulation results validate the superiority of the proposed method over traditional random phase shift and no-IRS schemes, particularly in scenarios with increased IRS elements and moderate transmission power. Furthermore, the analysis reveals the sensitivity of system performance to base station power and decoding error probability, highlighting the importance of careful parameter tuning. Overall, the proposed approach offers a promising solution for future high-performance wireless communication systems leveraging IRS technology.

Acknowledgements

This research is conducted within the framework of science and technology projects at instituutional level of Quy Nhon University under the project code T2025.896.16.

REFERENCES

1. C. Nwufo, O. Simpson, and Y. Sun. Reconfigurable Intelligent Surfaces (RIS) and Their Role in Next-Generation Wireless Networks: An Overview, *Transactions on Emerging Telecommunications Technologies*, **2025**, *36*(6), e70164.

2. S. Rajoria, A. Trivedi, and W. Wilfred Godfrey. A comprehensive survey: Small cell meets massive MIMO, *Physical communication* **2018** *26*, 40–49.

3. F. Qamar, S. H. A. Kazmi, K. A. Z. Ariffin, et al. Multi-antenna array-based massive mimo for b5g/6g: State of the art, challenges, and future research directions, *Information* **2024**, *15*, 442.

4. M. E. MHaque, F. Tariq, M. R. Khandaker, et al. A Comprehensive Survey of 5G URLLC and Challenges in the 6G Era, *preprint arXiv*, **2025**, arXiv:2508.20205.

5. D. Feng, L. Lai, J. Luo, et al. Ultra-reliable and low-latency communications: applications, opportunities and challenges, *Science China Information Sciences*, **2021** *64*(2), 120301.

6. W. Li, H. Chongwen, G. C. Alexandropoulos, et al. Channel estimation for RIS-empowered multi-user MISO wireless communications, *IEEE Transactions on Communications*, **2021** *69*(6), 4144-–4157.

7. Y. Xiu, J. Zhao, C. Yuen, et al. Secure beamformingfor multiple intelligent reflecting surfaces aided mmWave systems, *IEEE Communications Letters*, **2021**, *25*(2), 417-–421.

8. Y. Geng, T. H. Cheng, K. Zhong, K. C. Teh, et al. Movable IRS-Aided ISAC Systems: Joint Beamforming and Position Optimization, *arXiv preprint* arXiv:2509.04873.

9. W. R. Ghanem, V. Jamali, R. Schober. (2021, March). *Joint beamforming and phase shift optimization for multicell IRS-aided OFDMA-URLLC systems* In 2021 IEEE Wireless Communications and Networking Conference (WCNC), Nanjing, China, 2021.

10. Y. Xu, G. Gui, T. Ohtsuki, et al. Robust resource allocation for two-tier HetNets: An interference efficiency perspective, *IEEE Transactions on Green Communications and Networking*, **2021**, *5*(3), 1514-–1528.

11. B. Zheng, C. You and R. Zhang. Fast Channel Estimation for IRS-Assisted OFDM, *IEEE Wireless Communications Letters*, **2021**, *10*(3), 580–584.

12. Z. Zhang and L. Dai. A Joint Precoding Framework for Wideband Reconfigurable Intelligent Surface-Aided Cell-Free

Network, *IEEE Transactions on Signal Processing*, **2021**, *69*, 4085–4101.

13. R. Corvaja and A. G. Armada. Phase Noise Degradation in Massive MIMO Downlink With Zero-Forcing and Maximum Ratio Transmission Precoding, *IEEE Transactions on Vehicular Technology*, **2016**, *65*(10), 8052-8059.

14. T. Erseghe. Coding in the Finite-Blocklength Regime: Bounds Based on Laplace Integrals and Their Asymptotic Approximations, *IEEE Transactions on Information Theory*, **2016**, *62*(12), 6854–6883.

15. S. He, Z. An, J. Zhu, J. Zhang, Y. Huang and Y. Zhang. Beamforming Design for Multiuser uRLLC With Finite Blocklength Transmission, *IEEE Transactions on Wireless Communications*, **2021**, *20* (12) 8096–8109.

16. X. Yu, D. Xu, Y. Sun, et al. Robust and Secure Wireless Communications via Intelligent Reflecting Surfaces, *IEEE Journal on Selected Areas in Communications*, **2020**, *38*(11), 2637–2652.

17. S. Hu, Z. Wei, Y. Cai, et al. *Sum-Rate Maximization for Multiuser MISO Downlink Systems with Self-sustainable IRS*, GLOBECOM 2020-2020, IEEE Global Communications Conference, Taipei, Taiwan, 2020.

18. Q. Wu and R. Zhang. Weighted Sum Power Maximization for Intelligent Reflecting Surface Aided SWIPT, *IEEE Wireless Communications Letters*, **2020**, *9*(5), 586–590.

19. A. A. Nasir, H. D. Tuan, H. H. Nguyen, M. Debbah and H. V. Poor. Resource Allocation and Beamforming Design in the Short Blocklength Regime for URLLC, *IEEE Transactions on Wireless Communications*, **2021**, *20*(2), 1321–1335.