Joint Time Allocation, Beamforming, and Antenna Location for Pinching Antenna-assisted Downlink Multiuser Systems

Vu Phong Pham*, Quang Nhat Le*, and Trung Q. Duong*‡

*Memorial University, Canada, e-mail:{vppham, qnle, tduong}@mun.ca
‡Queen's University Belfast, UK, e-mail: trung.q.duong@qub.ac.uk

***Abstract*—In this paper, we jointly optimize the time allocation, beamforming, and antenna's locations to maximize the total throughput of all users in pinching antenna (PA)-assisted downlink multiuser systems. The problem is constrained by the limited power budget at the base station, the minimal distances between the PAs, and the minimum throughput of each user. The considered problem is inherently nonconvex due to the complex channel structure and the couplings of variables in the user throughput expressions. To address this, we propose an alternating optimization method to separate the problem into three subproblems that utterly depend on the fractional-time allocation, beamforming, and PA locations, respectively. These subproblems are then tackled by direct solutions, successive convex algorithm, and Armijo backtracking in augmented Lagrangian expression. Simulation results validate the superior performance of our proposed algorithm compared to the baseline methods and highlight its potential for practical implementation.**

I. Introduction

Pinching antenna systems (PASS) are emerging as a promising technology for reconfiguring the parameters of wireless channels [1]. The flexibility of pinching antennas (PAs) along the dielectric waveguides has brought many key advantages compared to existing fixed deployment and flexible antennas. In particular, the PA system is considerably more cost-effective and easier to deploy than the other conventional antennas. In addition, PAs are capable of repositioning their positions to provide strong, reliable, and adjustable line-of-sight (LoS) links for users, which leads to substantial improvements in overall system performance. PASS can be implemented simultaneously with other transmission approaches, such as non-orthogonal multiple access (NOMA), multiple-input multiple-output (MIMO) and integrated sensing and communication (ISAC) to further improve the performance of these network models [2]–[5].

In specific, a PA-assisted downlink communication system was investigated in [2], where multiple PAs are mounted on a single waveguide to optimize PA locations. Another study [3] investigated an equivalent system model, with the intention to optimize the locations of PAs and power allocation of the base station (BS) to maximize the total throughput of all terrestrial users. Meanwhile, [4] addressed an ISAC system assisted by PASS to maximize the total data throughput of all users. The problem is formulated by jointly optimizing the positions of PAs and the transmit power of mobile users. In [5], the authors illustrated the system model in which the access point is equipped with multiple waveguides attached to one PA. The studied problem is to maximize the weighted sum-rate of the users by optimizing the power budget and positions of PAs, which can be solved via fractional programming and block coordinate descent algorithm.

These aforementioned studies have successfully verified the effectiveness of PASS integrated with other promising approaches. However, to the best of our knowledge, none of these existing works have investigated the integration of the fractional-time method. This approach not only improves the achievable throughput of all users but also enables the system to support a greater number of users than the available antennas at the base station [6], [7]. Hence, in this paper, we apply the PA-assisted downlink multiuser communication combined with the fractional-time allocation approach to maximize the total rate of all users. The problem is formed by jointly optimizing fractional-time allocation, beamforming, and positioning of PAs under maximum power, minimum rate, and position constraints of PA. Since the investigated problem is inherently nonconvex and challenging to solve, we propose an alternating optimization (AO) method to split into different subproblems, which are simplified and can be solved via various inner convex approximations for beamforming, combined with a closed-form analytical solution for fractional-time allocation and a backtracking line search method to guarantee a stationary point of the optimization problem.

II. System Model and Problem Formulation

In this study, we consider a downlink (DL) communication system assisted by PAs, as shown in Fig. 1. The terrestrial network is designed in a $2D \times 2D$ square and without loss of generosity, we assume that the origin point lies in the center of the area. A terrestrial BS connected to N waveguides at height H is used to transmit the signal to $2K$ terrestrial users. Let us denote $\mathcal{N} = \{1, \dots, N\}$ and $\mathcal{K} = \{1, \dots, K\}$. The waveguides have the same length $2D$ and are parallel to the x-axis. The feed point of the waveguide n is set to $\overline{\mathbf{\Psi}}_n(-D, y_n, H)$, where $y_n = D - \frac{2nD}{N}, \forall n \in \mathcal{N}$, to ensure that each waveguide is equally spaced with each other.

Taking into account the terrestrial area, there are K users located in a central zone, named Zone 1 or the inner zone, and denoted as $u_{1k}(x_{1k}, y_{1k}, 0), \forall k \in \mathcal{K}$. Similarly, K users,

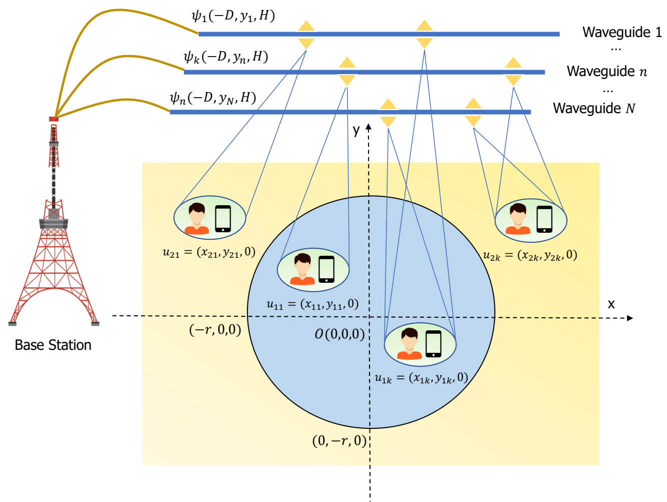

Fig. 1. An illustration of the joint fractional-time allocation and pinching antenna location for the downlink communication network.

denoted as $u_{2k}(x_{2k}, y_{2k}, 0), \forall k \in \mathcal{K}$ are located in the other zone, called Zone 2 or the outer zone. Each waveguide has two separate PAs that can freely move to serve the users in different zones. In particular, the inner zone can be served by one PA in each waveguide in duration τ_1 and the service in the outer zone can be provided by the other PAs in τ_2. The coordinates of the two PAs attached in the waveguides can be set as x_{1n} and $x_{2n}, \forall n \in \mathcal{N}$, respectively. Accordingly, the location of the PAs that operate in the inner zone and the outer zone can be set as $\mathbf{\Psi}_{i,n} = (x_{i,n}, y_n, H)$, where $x_{i,n}$ is the x-coordinate of PAs and $i \in \{1, 2\}$. The signal transmitted between BS and user u_{ik} in each zone can be expressed as

$$y_{ik} = \mathbf{h}_{ik}^{\mathbf{H}} \mathbf{w}_{ik} s_{ik} + \mathbf{h}_{ik}^{\mathbf{H}} \sum_{j \neq k} \mathbf{w}_{ij} s_{ij} + \sigma_{ik}, \quad \forall i \in \{1, 2\}, k \in \mathcal{K}. \tag{1}$$

Herein, s_{ik} is the information desired to transfer to user u_{ik}, $\mathbf{w}_{ik} \in \mathbb{C}^{N \times 1}$ is the beamforming vectors associated with s_{ik}. In addition, σ_{ik} is the additive white Gaussian noise (AWGN) and $\mathbf{h}_{ik} = [h_{ik,1}, \ldots, h_{ik,N}]^T$ denotes the channel vector of N antennas to the user u_{ik}. Specifically, the channel coefficient between the ground user u_{ik} and the PA $\mathbf{\Psi}_{i,n}$ can be given by

$$h_{ik,n} = \frac{\eta^{\frac{1}{2}} e^{-\alpha \left\| \overline{\mathbf{\Psi}}_n - \mathbf{\Psi}_{i,n} \right\|}}{\left\| \mathbf{\Psi}_{i,n} - u_{ik} \right\|} \times e^{-j\left(\frac{2\pi}{\lambda} \left\| \mathbf{\Psi}_{i,n} - u_{ik} \right\| + \frac{2\pi}{\lambda_g} \left\| \overline{\mathbf{\Psi}}_n - \mathbf{\Psi}_{i,n} \right\| \right)}, \tag{2}$$

where c is the speed of light, f_c denotes the carrier frequency in the free space, $\lambda = c/f_c$ denotes the wavelength in the free space environment and $\eta = c^2/(16\pi^2 f_c^2)$. We also denote α as the attenuation coefficient and $\lambda_g = \lambda/\eta_{\text{neff}}$ is the guided wavelength with η_{neff} denoting the effective refractive index of the waveguide. The achievable data rate for the pinching-antenna system at user u_{ik} can now be formulated as

$$R_{ik}(\mathbf{\Psi}, \tau_i, \mathbf{w}) = \tau_i \log_2 \left(1 + \frac{\left\| \mathbf{h}_{ik}^{\mathbf{H}} \mathbf{w}_{ik} \right\|^2}{\sum_{j \neq k} \left\| \mathbf{h}_{ik}^{\mathbf{H}} \mathbf{w}_{ij} \right\|^2 + \sigma_{ik}^2} \right). \tag{3}$$

Here, $\mathbf{\Psi} \triangleq \{\mathbf{\Psi}_{Xn}\}, \forall n \in \mathcal{N}, X \in \{1, 2\}$, $\boldsymbol{\tau} \triangleq \{\tau_1, \tau_2\}$, and $\mathbf{w} \triangleq \{\mathbf{w}_1, \mathbf{w}_2\}, \mathbf{w}_1 \triangleq \{\mathbf{w}_{1k}\}_{k \in \mathcal{K}}$ and $\mathbf{w}_2 \triangleq \{\mathbf{w}_{2k}\}_{k \in \mathcal{K}}$ are the beamforming vectors for Zone 1 and Zone 2, respectively. Note that the positions of the PAs in each waveguide must be synchronized so that the power allocated to each PA can be suitably calculated. The maximum power the BS can provide to these waveguides can be indicated as $P_{\max}$. Following [8], [9], the maximum power used by PAs can be approximately given by

$$P_a = P_{\max} F \sin^2(\kappa L), \tag{4}$$

where F is the coupling efficiency between the waveguide and the dielectric PA, κ is the coupling efficiency and L is the coupling length. From (4), we can control the power used by each PA by adjusting the coupling length L. For simplicity, we assume that the coupling efficiencies between the waveguides and the PAs attached to them are identical. We also assume that $x_{1,n} < x_{2,n}$, indicating that within each waveguide, PA 1 first consumes power to serve users located in the inner zone, while PA 2 utilizes the remaining power to support users in the outer zone. The maximum available power for the first PA is written as

$$\left\| \mathbf{w}_1 \right\|^2 \leq F P_{\max}. \tag{5}$$

Then, the maximum available power for the second PA is given by

$$\left\| \mathbf{w}_2 \right\|^2 \leq F(P_{\max} - \left\| \mathbf{w}_1 \right\|^2). \tag{6}$$

The goal of this paper is to maximize the sum data rate of the DL network by jointly optimizing the locations of the PAs, fractional-time allocation, and beamforming vectors. The optimization problem can be expressed as follows:

$$\max_{\mathbf{\Psi}, \boldsymbol{\tau}, \mathbf{w}} \; R_{\text{total}}(\mathbf{\Psi}, \boldsymbol{\tau}, \mathbf{w}) \triangleq \sum_{(i,k) \in \mathcal{L}} R_{ik}(\mathbf{\Psi}, \tau_i, \mathbf{w}) \tag{7a}$$

$$\text{s.t.} \; R_{ik}(\mathbf{\Psi}, \tau_i, \mathbf{w}) \geq \overline{R}_{\min}, \tag{7b}$$

$$-D \leq x_{i,n} \leq D, \forall n \in \mathcal{N}, i \in \{1, 2\}, \tag{7c}$$

$$x_{2n} - x_{1n} \geq \frac{\lambda}{2}, \forall n \in \mathcal{N}, \tag{7d}$$

$$0 < \tau_1, \tau_2 < 1, \tag{7e}$$

$$\tau_1 + \tau_2 \leq 1, \tag{7f}$$

$$\left\| \mathbf{w}_1 \right\|^2 \leq F P_{\max}, \tag{7g}$$

$$\left\| \mathbf{w}_2 \right\|^2 \leq F(P_{\max} - \left\| \mathbf{w}_1 \right\|^2), \tag{7h}$$

where $\mathcal{L} = \{(i, k) | i \in \{1, 2\}, k \in \mathcal{K}\}$. The sum rate maximization objective is formed in (7a). The first constraint (7b) represents the minimum throughput requirement for each user of the terrestrial network. Constraints (7c) and (7d) characterize the location of the PA and require the minimum distance of two PAs on the same waveguide to avoid mutual coupling. Notably, the wavelength spacing is required to be sufficiently large to prevent coupling between PAs mounted on different wavelengths. Constraints (7e) and (7f) denote the time allocation constraints. The last two constraints (7g) and (7h) set the limitation of the power used in each PA.

III. Proposed Solutions

It can be seen that problem (7) is a nonconvex problem due to the nonconcavity of the objective (7a) and constraint (7b). In addition, antenna positioning $\mathbf{\Psi}$ plays a crucial role in phase variations and channel coefficients due to its complex structure within channel components. Therefore, the proposed problem (7) is very challenging to tackle. As such, we adopt an AO method along with inner convex approximation algorithm to effectively solve this problem. The problem will be divided into three subproblems in beamforming, fractional-time allocation, and positioning of PAs to solve iteratively while the other variables are fixed.

A. Beamforming subproblem

The beamforming subproblem is constructed by keeping other variables fixed, that is, $\mathbf{\Psi} = \mathbf{\Psi}^{(t)}$ and $\boldsymbol{\tau} = \boldsymbol{\tau}^{(t)}$. By applying these modifications, the original problem (7) can be reformulated as

$$\max_{\mathbf{w}} \sum_{(i,k)\in\mathcal{L}} R_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w}) \quad (8a)$$

$$\text{s.t. } R_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w}) \geq \overline{R}_{\min}, \quad (8b)$$

$$\|\mathbf{w}_1\|^2 \leq FP_{\max}, \quad (8c)$$

$$\|\mathbf{w}_2\|^2 \leq F(P_{\max} - \|\mathbf{w}_1\|^2). \quad (8d)$$

It can be seen that the constraints (8c) and (8d) are already convex, but the problem remains nonconvex due to the nonconcavity of $R_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w})$. Therefore, to address this issue, we first reformulate the user throughput according to [10] as

$$R_{ik}(\mathbf{\Psi}^{(t)},\tau_i^{(t)}, \mathbf{w}) = \log_2\left(1 + \frac{\Re\big\{(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{w}_{ik}\big\}^2}{\sum_{j\neq k}\big\|(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{w}_{ij}\big\|^2 + \sigma_{ik}^2}\right). \quad (9)$$

The aforementioned reformulation can only be held if the trusted region for $\Re\big\{(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{w}_{ik}\big\}$ is satisfied. This constraint can be specifically written as

$$\Re\Big\{(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{w}_{ik}\Big\} \geq 0. \quad (10)$$

Following these developments, we approximate the problem concerned using a lower convex approximation for $R_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w})$ as [11]

$$\ln\left(1 + \frac{x^2}{y}\right) \geq \ln\left(1 + \frac{\bar{x}^2}{\bar{y}}\right) - \frac{\bar{x}^2}{\bar{y}} + \frac{2x\bar{x}}{\bar{y}} - \frac{\bar{x}^2(x^2+y)}{\bar{y}(\bar{y}+\bar{x}^2)}, \quad (11)$$

where $\bar{x}, \bar{y}$ are real positive constants and x, y are real positive variables. By employing (11) into the transmission rate of each user, we derive the lower bound approximation as

$$R_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w}) \geq \mathcal{R}_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w}), \quad (12)$$

where $\mathcal{R}_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w})$ is constructed as

$$\mathcal{R}_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)},\mathbf{w}) = a_{ik}^{(t)} + b_{ik}^{(t)}\Re\Big\{(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{w}_{ik}\Big\} - c_{ik}^{(t)}\Big(\sum_{j\in\mathcal{K}}\Big\|(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{w}_{ij}\Big\|^2 + \sigma_{ik}^2\Big). \quad (13)$$

Herein, we denote $x_{ik}^{(t)} = \Re\{(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{w}_{ik}^{(t)}\}, y_{ik}^{(t)} = \sum_{j\neq k}||(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{w}_{ij}^{(t)}||^2 + \sigma_{ik}^2, a_{ik}^{(t)} = \log_2(1 + (x_{ik}^{(t)})^2/y_{ik}^{(t)}) \times \tau_i^{(t)} - \tau_i^{(t)}(x_{ik}^{(t)})^2/(y_{ik}^{(t)}\ln 2)$, $b_{ik}^{(t)} = 2\tau_i^{(t)}x_{ik}^{(t)}/(y_{ik}^{(t)}\ln 2)$ and $c_{ik}^{(t)} = \tau_i^{(t)}(x_{ik}^{(t)})^2/[y_{ik}^{(t)}((x_{ik}^{(t)})^2 + y_{ik}^{(t)})\ln 2]$. Since the lower bound approximation of the transmission rate is found, we can rewrite the beamforming subproblem as follows:

$$\max_{\mathbf{w}} \sum_{(i,k)\in\mathcal{L}} \mathcal{R}_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w}) \quad (14a)$$

$$\text{s.t. } \mathcal{R}_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w}) \geq \overline{R}_{\min}, \quad (14b)$$

$$\|\mathbf{w}_1\|^2 \leq FP_{\max}, \quad (14c)$$

$$\|\mathbf{w}_2\|^2 \leq F(P_{\max} - \|\mathbf{w}_1\|^2). \quad (14d)$$

This reformulated problem is now a convex problem due to the concavity of $\mathcal{R}_{ik}(\mathbf{\Psi}^{(t)}, \tau_i^{(t)}, \mathbf{w})$. Hence, it can be addressed efficiently by applying a traditional convex program as CVXPY. The procedure is outlined in Algorithm 1.

Algorithm 1 Proposed Algorithm for Optimizing $\mathbf{w}^{(t+1)}$

```
Initialization: Set ε_beam as a small error tolerance and M_max
    as the maximum number of iterations. Set m = 0 and create a
    feasible point w^(t_0) = w^(t)
while | (R_total(Ψ^(t), τ^(t), w^(t_m)) − R_total(Ψ^(t), τ^(t), w^(t_{m−1}))) / R_total(Ψ^(t), τ^(t), w^(t_{m−1})) | ≥ ε_beam
    and m < M_max do
    Update m = m + 1;
    Find the optimal point w* by solving (14).
    Update the parameters: w^(t_m) = w*;
end while
Return: Optimal beamforming vector w^(t+1) = w^(t_m)
```

B. Antenna positioning subproblem

The subproblem of maximizing the total throughput of all users by arranging PAs positions can be expressed as follows:

$$\max_{\mathbf{\Psi}} \sum_{(i,k)\in\mathcal{L}} R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)}) \quad (15a)$$

$$\text{s.t. } R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)}) \geq \overline{R}_{\min}, \quad (15b)$$

$$-D \leq x_{i,n} \leq D, \forall n \in \mathcal{N}, i \in \{1,2\}, \quad (15c)$$

$$x_{2n} - x_{1n} \geq \frac{\lambda}{2}, \forall n \in \mathcal{N}. \quad (15d)$$

Due to the nonconcavity of $R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)})$, (15) remains nonconvex and challenging to solve. However, for each PA, the coordinate $\Psi_{i,n}$ is based on $x_{i,n}$; therefore, we will use a projected gradient descent method to maximize the following objective. Since the problem has constraint (15b), which is computationally intractable, we have to create a penalty term before calculating the gradient values with respect to the

PA positions $x_{i,n}$. The subproblem can be reformulated as $\min_{\mathbf{\Psi}} -\sum_{(i,k)\in\mathcal{L}} R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)})$ such that (15b), (15c) and (15d) are satisfied. Then, the augmented Lagrangian objective can be constructed as

$$\begin{aligned}\tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi}) = &-\sum_{(i,k)\in\mathcal{L}} R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)}) \\ &+ \sum_{(i,k)\in\mathcal{L}} \lambda_{ik} \left(\overline{R}_{\min} - R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)})\right) \\ &+ \sum_{(i,k)\in\mathcal{L}} \frac{\rho_{ik}}{2} \left(\overline{R}_{\min} - R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)})\right)^2,\end{aligned} \tag{16}$$

where λ_{ik} and ρ_{ik} are the Lagrange multipliers associated with the constraints (15b) and the penalty parameters for the quadratic terms. The problem now is to minimize $\tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi})$ with convex constraints (15c) and (15d) due to real variables. To calculate the gradient values of $\tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi})$, we first rewrite this objective as

$$R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)}) = \tau_i^{(t)} \log_2\left(1 + \varphi^{(t)}(\mathbf{\Psi})\right), \tag{17}$$

where $\mathbf{\Phi}_{ik}^{(t)} = \sum_{j\neq k} \mathbf{w}_{ij}^{(t)}(\mathbf{w}_{ij}^{(t)})^{\mathbf{H}}$, $S_{ik}^{(t)}(\mathbf{\Psi}) = ||\mathbf{h}_{ik}^{\mathbf{H}}\mathbf{w}_{ik}^{(t)}||^2$, $T_{ik}^{(t)}(\mathbf{\Psi}) = (\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{\Phi}_{ik}^{(t)}\mathbf{h}_{ik}^{(t)}$ and $\varphi^{(t)}(\mathbf{\Psi}) = S_{ik}^{(t)}(\mathbf{\Psi})/(T_{ik}^{(t)}(\mathbf{\Psi}) + \sigma_{ik}^2)$. Moreover, to further simplify the intractable form of $\mathbf{h}_{ik}$, we can rephrase the channel coefficient from PA $\mathbf{\Psi}_{i,n}$ to user u_{ik} in (2) as

$$\begin{aligned}h_{ik,n} = &\frac{\eta^{\frac{1}{2}} e^{-\alpha(x_{i,n}+D)}}{\sqrt{(x_{i,n}-x_{ik})^2 + \mathcal{A}_{i,n}}} \\ &\times e^{-j\left(\frac{2\pi}{\lambda}\sqrt{(x_{i,n}-x_{ik})^2+\mathcal{A}_{i,n}} + \frac{2\pi}{\lambda_g}(x_{i,n}+D)\right)},\end{aligned} \tag{18}$$

where $\mathcal{A}_{i,n} \triangleq (y_{i,n} - y_{ik})^2 + H^2$. Then, we can calculate the gradient values of $R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)})$ with respect to $x_{i,n}$ by the chain rule as follows:

$$\begin{aligned}\frac{\partial \tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi})}{\partial x_{i,n}} &= \sum_{(i,k)\in\mathcal{L}} M_{ik}^{(t)}(\mathbf{\Psi}) \frac{\partial R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)})}{\partial x_{i,n}}, \\ \frac{\partial R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)})}{\partial x_{i,n}} &= \frac{\tau_i^{(t)}}{(1+\varphi^{(t)}(\mathbf{\Psi}))\ln 2} \frac{\partial \varphi^{(t)}(\mathbf{\Psi})}{\partial x_{i,n}}, \\ \frac{\partial \varphi^{(t)}(\mathbf{\Psi})}{\partial x_{i,n}} &= \frac{1}{T_{ik}^{(t)}(\mathbf{\Psi}) + \sigma_{ik}^2} \frac{\partial S_{ik}^{(t)}(\mathbf{\Psi})}{\partial x_{i,n}} \\ &\quad - \frac{S_{ik}^{(t)}(\mathbf{\Psi})}{(T_{ik}^{(t)}(\mathbf{\Psi}) + \sigma_{ik}^2)^2} \frac{\partial T_{ik}^{(t)}(\mathbf{\Psi})}{\partial x_{i,n}}, \\ \frac{\partial S_{ik}^{(t)}(\mathbf{\Psi})}{\partial x_{i,n}} &= 2\Re\left\{(\mathbf{w}_{ik}^{(t)})^{\mathbf{H}} \frac{\partial \mathbf{h}_{ik}}{\partial x_{i,n}} \mathbf{h}_{ik}^{\mathbf{H}}\mathbf{w}_{ik}^{(t)}\right\}, \\ \frac{\partial T_{ik}^{(t)}(\mathbf{\Psi})}{\partial x_{i,n}} &= 2\Re\left\{(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}\mathbf{\Phi}_{ik}^{(t)} \frac{\partial \mathbf{h}_{ik}}{\partial x_{i,n}}\right\}, \\ \frac{\partial \mathbf{h}_{ik}}{\partial x_{i,n}} &= [0, \ldots, \mathcal{C}(x_{i,n})h_{ik,n}, \ldots, 0],\end{aligned} \tag{19}$$

where $M_{ik}^{(t)}(\mathbf{\Psi}) = -(\rho_{ik}(\overline{R}_{\min} - R_{ik}(\mathbf{\Psi}, \tau_i^{(t)}, \mathbf{w}^{(t)})) + 1 + \lambda_{ik})$ and $\mathcal{C}(x_{i,n}) = -\alpha - (x_{i,n} - x_{ik})/(\mathcal{A}_{i,n} + (x_{i,n} - x_{ik})^2) - j(2\pi(x_{i,n} - x_{ik})/(\lambda\sqrt{\mathcal{A}_{i,n} + (x_{i,n} - x_{ik})^2})) + 2\pi/\lambda_g)$. Then, the PA position variables $x_{i,n}$ are updated through the backtracking line search method combined with the Armijo condition. The updated results are then projected into the convex set defined by constraints (15c) and (15d) to ensure feasibility. Let $\mathbf{x} \triangleq \{x_{i,n}\}_{i\in\{1,2\},n\in\mathcal{N}}$ and $\nabla_{\mathbf{x}}\tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi}_p^{(t)}) = [\partial\tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi}_p^{(t)})/\partial x_{i,n}]^T_{i\in\{1,2\},n\in\mathcal{N}}$. The proposed two-stage projected gradient descent (PGD)-based method is demonstrated in Algorithm 2, which guarantees convergence to a stationary point.

Algorithm 2 Two-stage PGD-based Method and Backtracking for Optimizing $\mathbf{\Psi}^{(t+1)}$

Initialization: Set $p = 0$, maximum update iteration $\mathcal{P}_{\max}$, maximum backtracking iteration $\mathcal{Q}_{\max}$; generate the initial feasible points $\mathbf{\Psi}_0^{(t)} = \mathbf{\Psi}^{(t)}, \boldsymbol{\lambda} \triangleq \{\lambda_{ik}\}_{(i,k)\in\mathcal{L}}, \boldsymbol{\rho} \triangleq \{\rho_{ik}\}_{(i,k)\in\mathcal{L}}$, Armijo parameter μ, initial step length α_0 and small convergence parameter ε_{pos}
while $\left|\frac{R_{\text{total}}(\mathbf{\Psi}_p^{(t)}, \boldsymbol{\tau}^{(t)}, \mathbf{w}^{(t)}) - R_{\text{total}}(\mathbf{\Psi}_{p-1}^{(t)}, \boldsymbol{\tau}^{(t)}, \mathbf{w}^{(t)})}{R_{\text{total}}(\mathbf{\Psi}_{p-1}^{(t)}, \boldsymbol{\tau}^{(t)}, \mathbf{w}^{(t)})}\right| \geq \varepsilon_{\text{pos}}$ **and** $p < \mathcal{P}_{\max}$ **do**
  Calculate gradient: $\mathbf{g}_p = \nabla_{\mathbf{x}}\tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi}_p^{(t)})$;
  Set $\alpha = \alpha_0$ and $q = 0$;
  while $q < \mathcal{Q}_{\max}$ **do**
    Compute projected point: $\mathbf{\Psi}_p^{\text{Proj}} = \text{Proj}[\mathbf{\Psi}_p^{(t)} - \alpha\mathbf{g}_p]$;
    if $\tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi}_p^{\text{Proj}}) \leq \tilde{R}^{\mathbf{\Psi}}_{\text{total}}(\mathbf{\Psi}_p^{(t)}) - \mu\alpha\|\mathbf{g}_p\|_2^2$ **then**
      $\mathbf{\Psi}_{p+1}^{(t)} = \mathbf{\Psi}_p^{\text{Proj}}$;
      $\alpha_{p+1} = \alpha$;
      break
    else
      $\alpha = \alpha/2$;
      $q = q + 1$;
    end if
  end while
  Update $\lambda_{ik} \leftarrow \max\left\{0, \lambda_{ik} + \rho_{ik}\left(\overline{R}_{\min} - R_{ik}(\mathbf{\Psi}_{p+1}^{(t)})\right)\right\}$, $\forall(i,k)\in\mathcal{L}$;
  Update $p = p + 1$;
end while
Output: Optimal antenna positions $\mathbf{\Psi}^{(t+1)} = \mathbf{\Psi}_p^{(t)}$

C. Time allocation subproblem

The next subproblem is to maximize the total transmission rate from the BS to the ground users by optimizing the time allocation at each PA. The following problem can be solved by setting $\mathbf{\Psi} = \mathbf{\Psi}^{(t)}$ and $\mathbf{w} = \mathbf{w}^{(t)}$ as

$$\max_{\boldsymbol{\tau}} \sum_{(i,k)\in\mathcal{L}} R_{ik}(\mathbf{\Psi}^{(t)}, \tau_i, \mathbf{w}^{(t)}) \tag{20a}$$

$$\text{s.t. } R_{ik}(\mathbf{\Psi}^{(t)}, \tau_i, \mathbf{w}^{(t)}) \geq \overline{R}_{\min}, \tag{20b}$$

$$0 < \tau_1, \tau_2 < 1, \tag{20c}$$

$$\tau_1 + \tau_2 \leq 1. \tag{20d}$$

Since (20) is a closed-form problem with two variables, we can directly attain the solution of this subproblem. It can be clearly seen that constraint (20d) must hold to the equality due to the maximizing condition. In addition, constraint (20c) can be tightened using constraint (20b). As other variables remain unchanged, the minimum value of τ_i can be derived as

$$\tau_i \geq \tau_i^{\min} \triangleq \max_{k\in\mathcal{K}} \left\{ \overline{R}_{\min} \left[C_{ik}^{(t)}(\mathbf{\Psi}^{(t)}, \mathbf{w}^{(t)}) \right]^{-1} \right\}. \tag{21}$$

Here, $C_{ik}^{(t)}(\mathbf{\Psi}^{(t)}, \mathbf{w}^{(t)})$ is the maximum transmission capacity from BS to user u_{ik}. The formulation of $C_{ik}^{(t)}(\mathbf{\Psi}^{(t)}, \mathbf{w}^{(t)})$ can be expressed as

$$C_{ik}^{(t)}(\mathbf{\Psi}^{(t)}, \mathbf{w}^{(t)}) = \log_2 \left(1 + \frac{\left\| (\mathbf{h}_{ik}^{\mathbf{H}})^{(t)} \mathbf{w}_{ik}^{(t)} \right\|^2}{\sum_{j\neq k} \left\| (\mathbf{h}_{ik}^{\mathbf{H}})^{(t)} \mathbf{w}_{ij}^{(t)} \right\|^2 + \sigma_{ik}^2} \right), \tag{22}$$

where $(\mathbf{h}_{ik}^{\mathbf{H}})^{(t)}$ is constructed by the fixed antenna position $\mathbf{\Psi}^{(t)}$. The objective (20a) of the interested subproblem can be rewritten as

$$R_{\text{total}}^{\boldsymbol{\tau}} = \tau_1 C_{\text{s1}}^{(t)} + \tau_2 C_{\text{s2}}^{(t)}, \tag{23}$$

where $C_{\text{s1}}^{(t)} \triangleq \sum_{k\in\mathcal{K}} C_{1k}^{(t)}(\mathbf{\Psi}^{(t)}, \mathbf{w}^{(t)})$ and $C_{\text{s2}}^{(t)} \triangleq \sum_{k\in\mathcal{K}} C_{2k}^{(t)}(\mathbf{\Psi}^{(t)}, \mathbf{w}^{(t)})$. Since the maximization of (23) with respect to (21) admits a closed-form solution, the optimal time allocation is directly given as

$$\boldsymbol{\tau}^{(t+1)} = \begin{cases} \{1-\tau_2^{\min}, \tau_2^{\min}\}, & \text{if } C_{\text{s1}}^{(t)} \geq C_{\text{s2}}^{(t)}, \\ \{\tau_1^{\min}, 1-\tau_1^{\min}\}, & \text{if } C_{\text{s1}}^{(t)} < C_{\text{s2}}^{(t)}. \end{cases} \tag{24}$$

D. Proposed AO Algorithm

Denote $\boldsymbol{\delta} = \{\mathbf{w}, \mathbf{\Psi}, \boldsymbol{\tau}\}$, $\boldsymbol{\delta}^{(t)} = \{\mathbf{w}^{(t)}, \mathbf{\Psi}^{(t)}, \boldsymbol{\tau}^{(t)}\}$, $\boldsymbol{\xi}_{(t)} = \{\mathbf{w}, \mathbf{\Psi}, \boldsymbol{\tau}^{(t)}\}$ and $\boldsymbol{\xi}_{(t)}^{(i)} = \{\mathbf{w}^{(t_i)}, \mathbf{\Psi}^{(t_i)}, \boldsymbol{\tau}^{(t)}\}$. Throughout the subproblems developed above, the algorithm proposed to tackle (7) is given in Algorithm 3. This algorithm uses two different AO loops for optimization, where the first AO loop is used to derive the optimal solution to the problem (7) and the second AO loop is used to optimize the beamforming $\mathbf{w}^{(t+1)}$ and PA positions $\mathbf{\Psi}^{(t+1)}$ in iteration t of the first AO loop. Specifically, in the second AO loop, the optimal beamforming $\mathbf{w}^{(t+1)}$ and PA positions $\mathbf{\Psi}^{(t+1)}$ are first conducted in iteration t. Then, using the direct solution in (24), the optimal time allocation $\boldsymbol{\tau}^{(t+1)}$ is derived using $\mathbf{w}^{(t+1)}$ and $\mathbf{\Psi}^{(t+1)}$. The counting index is updated in both AO loops and continues to operate until convergence is reached.

Algorithm 3 Proposed AO Algorithm for solving (7)

Initialization: Set $t = 0$, $i = 0$, maximum number of iteration $N_{\max}^1, N_{\max}^2$, convergence thresholds $\varepsilon_1, \varepsilon_2$ and generate the initial feasible points $(\mathbf{w}^{(0)}, \mathbf{\Psi}^{(0)}, \boldsymbol{\tau}^{(0)})$
while $\frac{R_{\text{total}}(\boldsymbol{\delta}^{(t)}) - R_{\text{total}}(\boldsymbol{\delta}^{(t-1)})}{R_{\text{total}}(\boldsymbol{\delta}^{(t-1)})} \leq \varepsilon_1$ or $t > N_{\max}^1$ **do**
  while $\frac{R_{\text{total}}(\boldsymbol{\xi}_{(t)}^{(i)}) - R_{\text{total}}(\boldsymbol{\xi}_{(t)}^{(i-1)})}{R_{\text{total}}(\boldsymbol{\xi}_{(t)}^{(m-1)})} \leq \varepsilon_2$ or $t > N_{\max}^2$ **do**
    Implement Algorithm 1 to obtain the optimal beamforming $\mathbf{w}^{(t_{i+1})}$;
    Implement Algorithm 2 to achieve the optimal PA positions $\mathbf{\Psi}^{(t_{i+1})}$;
    Update $i = i + 1$;
  end while
  Result: Optimal beamforming $\mathbf{w}^{(t+1)} = \mathbf{w}^{(t_{i+1})}$ and PA positions $\mathbf{\Psi}^{(t+1)} = \mathbf{\Psi}^{(t_{i+1})}$
  Using (24) and $\mathbf{w}^{(t+1)}, \mathbf{\Psi}^{(t+1)}$ to attain the optimal time allocation $\boldsymbol{\tau}^{(t+1)}$
end while
Output: Optimal solution $\boldsymbol{\delta}^* = \{\mathbf{w}^{(t)}, \mathbf{\Psi}^{(t)}, \boldsymbol{\tau}^{(t)}\}$

IV. Simulation Results and Discussions

A. Simulation Settings

In this section, the performance of the proposed algorithm and other benchmarks is evaluated via numerical simulations. In the network model, the number of waveguides used is $N = 4$ to serve $K = 3$ users in each zone and $D = 10$m. The height of the wavelengths is set to $H = 3$m. The attenuation coefficient is set to $\alpha = 0.08$ dB/m, the carrier frequency is set to $f_c = 28$GHz and the effective refractive index of the waveguide is set to $\eta_{\text{neff}} = 1.4$ [1]. The maximum power used by the BS is $P_{\max} = 40$dBm, the noise power at the receiver is $\sigma^2 = -90$dBm and the coupling efficiency is set to $F = 0.9$. The position of the users are randomly selected in the network area $2D \times 2D$. The simulations are implemented in a Python environment using `matplotlib` package, and the CVXPY package is used to solve the beamforming subproblem (8).

B. Numerical Results and Discussions

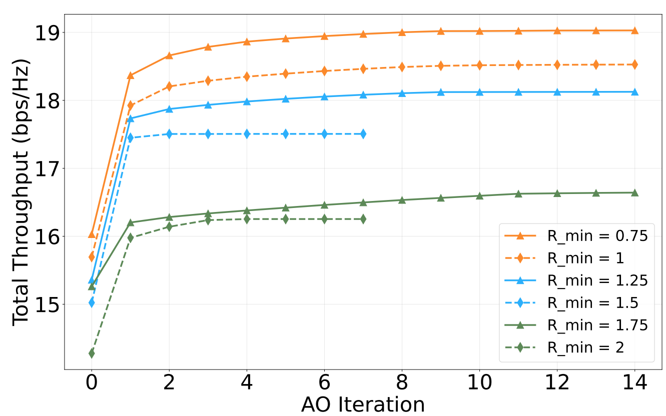

Fig. 2. Convergence behavior of the proposed AO algorithm demonstrating stable performance under varying minimum rate constraints $R_{\min}$ with $P_{\max} = 40$dBm.

Fig. 2 illustrates the convergence behavior of the proposed AO method under different minimum rate requirements for users. It can be observed that for $R_{\min} = 1.5$ and $R_{\min} = 2$, the algorithm reaches the optimal solution within only 7 iterations, while additional iterations are required for other cases to achieve full convergence. On the other hand, the proposed AO algorithm exhibits a significant decrease when the requirement threshold $R_{\min}$ increases, reflecting the trade-off between the achievable system throughput and the fairness between users. This behavior is expected, as the fractional-time

PA system allocates a significant portion of the BS transmit power to users in the outer zone to satisfy the minimum rate demands. Consequently, the available power for users under favorable channel conditions is reduced, leading to an overall decrease in the achievable system throughput. The results confirm that the proposed AO algorithm effectively balances this trade-off, maintaining stable convergence properties and consistent performance across different rate constraints.

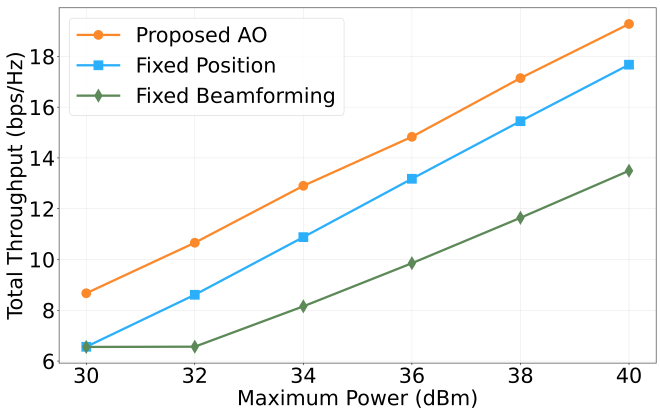

Fig. 3. Performance characteristics of the proposed AO algorithm evaluated under different maximum transmit power $P_{\max}$, with the minimum rate requirement set to $R_{\min} = 0.5$bps/Hz

Fig. 3 represents the performance behavior of the proposed AO method compared to two other different schemes. In the "Fixed Position" scheme, the PA positions are fixed at their initial values, and the AO algorithm is used to iteratively optimize the time allocation and beamforming variables. In contrast, the "Fixed Beamforming" approach maintains the initial beamforming configuration while optimizing the PA positions and time allocation variables using the AO approach. The comparison is carried out under different maximum transmit power $P_{\max}$, where the rate threshold of each ground user is set to $R_{\min} = 0.5$bps/Hz. One can see that the total throughput of all the considered schemes improves due to the expanded power budget available $P_{\max}$ for user transmission. However, the rate of improvement for the proposed AO remains higher, reflecting its ability to utilize the available resources. Meanwhile, other methods are constrained by their fixed parameters, which limit their adaptability to changing transmission power conditions. These results collectively illustrate the robustness and scalability of the proposed AO method, indicating that the integrated design of beamforming, PA positions and time allocation can deliver higher throughput and ensure stable convergence with adaptable performance in various system settings.

V. Conclusion

In this study, we address the integrated challenge of maximizing the total data rate of ground users through the joint optimization of fractional-time allocation, beamforming, and the positioning of PAs. To tackle this problem, we develop three distinct solution strategies that iteratively solve the three subproblems within an AO framework. In particular, an inner convex approximation is employed for the beamforming subproblem, a closed-form expression is derived for the fractional-time allocation, and the Armijo backtracking line search method is utilized to optimize the positions of the PAs. Numerical results have illustrated the superior performance of the proposed approach compared to other benchmark schemes. Moreover, the simulation results also highlight the effectiveness of the proposed AO method in maximizing the total data throughput under varying minimum rate requirements, as well as the efficiency of the inner convex approximation method in the beamforming subproblem. Future work may investigate the integration of advanced transmission techniques, such as MIMO and NOMA, to further improve the scalability and robustness of wireless network models for supporting a larger number of users.

Acknowledgments

This work was supported in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109, in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant Program RGPIN-2025-04941.

References

[1] Z. Ding, R. Schober, and H. V. Poor, "Flexible-antenna systems: A pinching-antenna perspective," *IEEE Trans. Commun.*, vol. 73, no. 10, pp. 9236–9253, Oct. 2025.
[2] X. Xie, F. Fang, Z. Ding, and X. Wang, "A low-complexity placement design of pinching-antenna systems," *IEEE Commun. Lett.*, vol. 29, no. 8, pp. 1784–1788, Aug. 2025.
[3] Y. Xu, Z. Ding, and G. K. Karagiannidis, "Rate maximization for downlink pinching-antenna systems," *IEEE Wireless Commun. Lett.*, vol. 14, no. 5, pp. 1215–1219, May 2025.
[4] Y. Qin, Y. Fu, and H. Zhang, "Joint antenna position and transmit power optimization for pinching antenna-assisted ISAC systems," *IEEE Wireless Commun. Lett.*, 2025.
[5] A. Bereyhi, S. Asaad, C. Ouyang, Z. Ding, and H. V. Poor, "Downlink beamforming with pinching-antenna assisted MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 24, no. 5, pp. 2567–2580, May 2025.
[6] A. A. Nasir, H. D. Tuan, and T. Q. Duong, "Fractional time exploitation for serving IoT users with guaranteed QoS by 5G spectrum," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 128–134, Oct. 2018.
[7] V.-D. Nguyen, H. D. Tuan, T. Q. Duong, O.-S. Shin, and H. V. Poor, "Joint fractional time allocation and beamforming for downlink multiuser MISO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6356–6369, Sep. 2017.
[8] Y. Liu, Z. Wang, X. Mu, C. Ouyang, X. Xu, and Z. Ding, "Pinching-Antenna Systems (PASS): Architecture Designs, Opportunities, and Outlook," *IEEE Commun. Mag.*, 2025.
[9] K. Okamoto, *Fundamentals of Optical Waveguides*, 2nd ed. Elsevier, 2006.
[10] Z. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed mimo receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
[11] Z. Sheng, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Beamforming optimization for physical layer security in miso wireless networks," *IEEE Trans. Signal Process.*, vol. 66, no. 14, pp. 3710–3723, May 2018.