Achievable Rate Optimization for Aerial Intelligent Reflecting Surface-Aided Cell-Free Massive MIMO System

TAO ZHOU, KUI XU (Member, IEEE), XIAOCHEM XIA, WEI XIE, AND JIANHUI XU
College of Communications Engineering, PLA Army Engineering University, Nanjing 210007, China
Corresponding author: Kui Xu (lgdxukui@sina.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 62071485, Grant 61901519, Grant 61771486, and Grant 61671472, in part by the Basic Research Project of Jiangsu Province under Grant BK20192002, and in part by the Natural Science Foundation of Jiangsu Province under Grant BK20201334 and Grant BK20181335.

ABSTRACT The intelligent reflecting surface (IRS) is considered a core technology of next-generation mobile communication. It has significant advantages in enhancing network coverage, spectrum efficiency, energy efficiency, and deployment cost. Compared with the conventional massive multiple-input-multiple-output (MIMO) system, cell-free massive MIMO overcomes the limitation imposed by inter-cell interference in traditional cellular mobile networks and realizes coherent transmission centered on users. In this article, we consider a new application scenario for the IRS—an aerial IRS (AIRS)-aided cell-free massive MIMO system where multiple APs serve several users through an AIRS. The users are in a “shadow area” where we cannot provide good quality of service (QoS) due to the remote location and the shelter of tall buildings. Our goal is to optimize the power allocation and beamforming of each AP, the placement and reflection phase shift parameters of the AIRS to maximize the user’s achievable rate. Firstly, we consider the optimization for fixed placement of the AIRS, where we propose a joint optimization strategy to maximize the achievable rate of the user. Then, we propose a fast optimal location search algorithm base on the path loss model to determine the optimal location of the AIRS and decrease the computational complexity. The simulation results show that the proposed methods can improve the performance for the achievable rate of the system. To the best of our knowledge, this article is the first to study the application scenario for the combination of IRS technology and a cell-free massive MIMO system.

INDEX TERMS Aerial intelligent reflecting surface (AIRS), cell-free massive multiple-input-multiple-output (MIMO), optimization.

I. INTRODUCTION

Over the past decade, the mobile communication industry has developed rapidly. Many key technologies have been studied in an effort to achieve the goal of a 1000-fold rate increase. In past studies, the industry generally believed that wireless channels are uncontrollable due to the uncertainty of the propagation environment. Most of the research on improving communication quality still focuses on optimizing the receiver and transmitter to adapt to the wireless channel. In fact, using an intelligent reflecting surface (IRS), we can realize intelligent and reconfigurable wireless communication based on the propagation environment.

The associate editor coordinating the review of this manuscript and approving it for publication was Prakash Periasamy.

IRS is also called a reconfigurable intelligent surface (RIS) [1], [2] or software-controlled metasurface [3]–[5]. It is a new technology that solves the problem of area coverage, energy spectrum efficiency, and spectrum efficiency and is considered as one of the core technologies supporting next-generation mobile communication. Unlike traditional parabolic antennas or phased array antennas, IRS is a linear or planar structure that is controlled by a smart controller connected to the base station (BS) or the access point (AP). An IRS contains many low-cost passive array elements, each of which has a sub-wavelength structure. Each element can independently change the phase or amplitude of the incident signal using the smart controller. By jointly designing the reflection phase of each element, we can realize the coherent superposition of reflected signals at specific positions to
achieve beamforming [6]. This is equivalent to “changing” the direction of the reflected signal so that it points to a specific location. When the direct path between the transmitter and the receiver is blocked, the channel of the direct path is weak. An IRS can be deployed at an appropriate location to establish a high-quality reflection channel to reflect the signal from the transmitter to the receiver. By adjusting the reflection coefficient of each element in the IRS, the signals reflected by the IRS can be superposed coherently at the receiver to improve the quality of communication. In fact, adjusting the reflection coefficient of the IRS is equivalent to changing the reflecting channel, which is why the IRS can realize real-time reconfigurable communication based on the propagation environment.

It is worth noting that although an IRS looks similar to a traditional multiple-input–multiple-output (MIMO) relay, there are obvious differences. First, relay operations such as decoding and forwarding (DF) and amplifying and forwarding (AF) need to demodulate the received signals before forwarding them; consequently, the relays must possess signal processing capabilities. In contrast, an IRS usually does not possess signal processing capability but instead simply reflects signals passively. Therefore, there is no need to equip the IRS with a transmission radio frequency (RF) chain [7]. Second, compared with traditional relays, IRS consumes little energy, and no additional thermal noise is introduced during the reflection. IRS deployment costs are also low; they can even be deployed on urban building walls. One way to improve the spectrum efficiency of the communication system is to deploy a large number of antennas at the AP or BS. However, due to the limitations of antenna size and deployment cost, the number of antennas that can be equipped at the receiver and transmitter is limited. Considering the cost and energy consumption, we want the number of antennas for APs and users to be as small as possible. Using the IRS, we can achieve a performance gain that is similar to the performance of massive MIMO arrays. Based on the above differences, the IRS has attracted increasing attention and has broad application prospects.

The existing studies on IRS mainly focus on the performance analysis and optimization of an IRS-assisted cellular network. In [8], the performance of an IRS-aided single-input–single-output (SISO) communication system is studied, in which the authors compare an IRS with a DF relay. The results show that the deployment of an IRS with a large number of reflection elements can achieve better performance under the premise of minimizing the total transmit power and maximizing energy efficiency when the system requires a high data rate. In [9], an IRS is deployed in a multiple-input–single-output (MISO) system to assist the communication between a multiple-antenna AP and several single-antenna users. By jointly optimizing the transmit beamforming at the AP and the phase shift matrix of the IRS, the problem of minimizing transmit power under the constraint of the user signal-to-noise ratio (SNR) is solved. In [10], the capacity performance of an IRS aided-MIMO communication system is studied, and an IRS is used to assist the communication between a multi-antenna BS and a multi-antenna user. An alternative iterative strategy is proposed to jointly optimize the IRS phase shift matrix and transmit the covariance matrix to improve the system capacity performance. With the deepening research on IRS, academia has proposed a new IRS architecture—the aerial intelligent reflecting surface (AIRS). Unlike conventional IRSs, AIRS can be deployed on aerial platforms such as unmanned aerial vehicles (UAVs) thus can provide stronger line-of-sight (LoS) links [11]. Moreover, the AIRS can adapt to different propagation environments by changing its position. With this mobility, AIRS can realize the panoramic reflection of incident signals [12].

Cell-free massive MIMO has been considered a core architecture of the next-generation mobile communication network. It overcomes the limitation of inter-cell interference in traditional cellular mobile networks. In a cell-free massive MIMO system, multiple APs serve a small number of users in the coverage area simultaneously and realize user-centered coherent transmission under the control of a central processing unit (CPU) [13]–[16].

Research on cell-free massive MIMO is being increasingly perfected. However, some problems exist that need to be considered. In the actual deployment of a cell-free massive MIMO system, the layout of urban buildings is uneven, and it is difficult to achieve uniform coverage in the deployment of APs. For areas with a remote location or the shelter of tall buildings, the communication between users and APs is blocked; we call these “shadow areas”. In a cell-free massive MIMO system, we can choose the appropriate APs to serve a user based on their location, namely, AP clustering. Specifically, we can serve users by clustering APs for these shadow areas, but the service quality may not be significantly improved because these areas are far from APs. One simple way to improve the users’ quality of service is to build more APs in these areas to achieve high-density coverage. However, the deployment cost and energy consumption increase as more APs are deployed. Can we serve these shadow area users in another way? Indeed, we can deploy an IRS to support the communication between users in a shadow area and the APs.

The addition of an IRS for a cell-free massive MIMO system solves the above problems. For users in these shadow areas, the direct link is usually weak, so direct link transmission has poor performance. In this case, an AIRS can be used to support communication. Specifically, we can use an AIRS that is deployed on a UAV to support the communication between the APs and the user. In this system, the smart controller of the AIRS is connected to the APs with a CPU. All transmission parameters are adjusted by the CPU; and with channel estimation and some location technologies, the locations of APs, users and all the channel state information (CSI) are known by the CPU. We call this information the “user fingerprint”. The “shadow area” users can be identified if the CPU obtains this “user fingerprint information”. Then, an AIRS is deployed by the CPU to assist...
with communication. The advantages of adding AIRS to a cell-free massive MIMO system are listed as follows.

- The cost of deploying an AIRS is lower than the cost of deploying more APs.
- The power consumption of an AIRS is also lower than that of APs.
- The AIRS is deployed on a flyer such as a UAV or a balloon, so it can dynamically adjust its own location to serve users in different propagation environments.
- The addition of AIRS can improve the area coverage in a cell-free massive MIMO system.

The main contributions of this article are summarized as follows. In this article, we extend the application of the IRS from traditional cellular networks to a cell-free Massive MIMO system. We study an AIRS-aided cell-free massive MIMO downlink system where numbers of APs serve a single-antenna user in the “shadow area” through an AIRS. We propose an interactive iterative optimization algorithm to maximize the achievable rate of the user at a fixed position. Then we consider the effect of AIRS’s location on achievable rate performance and propose a fast optimal location search algorithm to find the optimal position of the AIRS.

The organization of this article is as follows. In Section II, we will introduce the system model of the AIRS-aided cell-free massive MIMO system and formulate the problem to be solved in this article. In Section III, we begin by considering the case of AIRS with a fixed position; we use a joint iteration method to jointly optimize the phase shift of AIRS, the transmit power allocation and the precoding vectors of all APs. Then, we consider the case of AIRS with a movable position, and we use a fast optimal location search algorithm to obtain the optimal placement of the AIRS. In Section IV, we will perform a simulation of a real scene to verify the performance of the proposed algorithm. Finally, we will conclude this article in Section V.

The notations used in this article are listed below. We use $j = \sqrt{-1}$ to express an imaginary unit. For readability, we use bold capital letters to denote matrices. A vector is denoted by an italic and bold lowercase letter. Here, we denote the space of $M \times N$ as a complex-valued matrix by $\mathbb{C}^{M \times N}$. For a vector $g$, we use $|g|$ to denote the $\ell_1$-norm and use $\|g\|$ to denote the $\ell_2$-norm. Additionally, we use $\text{arg}(g)$ to denote a vector, each element of which is a phase of the corresponding element in $g$. We use diag($g$) to denote a diagonal matrix, where each diagonal element is the corresponding element in $x$. For a matrix $F$, $F^T$ and $F^H$ denote its transpose and conjugate transpose, respectively. For a plural $C$, we use $|C|$ to express the modulus of $C$, and we use $\arg(C)$ to denote the phase of $C$.

II. SYSTEM AND CHANNEL MODEL

A. SYSTEM MODEL

As shown in Fig. 1, some APs serve several users in this system under the control of a CPU. The CPU knows which user is in the shadow area by “user fingerprint information”, and it deploys an AIRS to assist with the communication. The AIRS is deployed on a UAV, so it can move aerially. In practice, the AIRS cannot be placed at any point in the air because it needs to be placed away from the shelters, and the flight altitude of the AIRS is also limited. The placeable area is expressed as $A$.

We assume there are $N$ APs and $K$ shadow-area users. Each AP is deployed with $M$ antennas, and the AIRS consists of $R$ reflection elements. For the sake of discussion, we assume there is only one user in the shadow area, thus, $K = 1$. We will research multi-user cases in our future work. To simplify the description, we set up a coordinate system, where the AP$_n$ and the user is in the x-o-z plane. $A$ is in the x-o-z plane. So, the coordinates of the AIRS are $l = (x, 0, h)$. In this case, we will optimize the horizontal coordinate $x$ and height $h$ of the AIRS.

All the direct links are blocked by tall buildings, so we assume they do not exist. The transmit power of each AP is controlled by the CPU. All APs transmit at the full allocated power. We define the power allocation coefficient of $n$-th AP as $\lambda_n$. The power allocation vector is defined as $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N)$; then, we have $\sum_{n=1}^{N} \lambda_n = 1$. $P$ represents the total transmit power. All APs are serving the user at the same time. The signal received at the user can be expressed as:

$$y = (h_1(l)\Theta G_1(l)\sqrt{P\lambda_1w_1} + h_2(l)\Theta G_2(l)\sqrt{P\lambda_2w_2} + \ldots + h_N(l)\Theta G_N(l)\sqrt{P\lambda_Nw_N})x + \delta$$

where $h_i(l) \in \mathbb{C}^{1 \times R}$ represents the channel from AIRS to the user. $G_n(l) \in \mathbb{C}^{R \times M}$ denotes the channel from the $n$-th AP to the AIRS. $l$ denotes the coordinates of the AIRS. In this article, we only consider the optimization for the height $h$ and horizontal position $x$ of the AIRS, so $l = (x, h)$. As we can see, the channels $h_i(l)$ and $G_n(l)$ change with $l$. $\Theta \in \mathbb{C}^{R \times R}$ is the phase shift matrix of the AIRS. $\Theta = \text{diag}(\eta_1e^{j\phi_1}, \eta_2e^{j\phi_2}, \ldots, \eta_R e^{j\phi_R})$, where $\eta_i \in [0, 1]$, $i = 1, 2, \ldots, R$, represents the amplitude reflection coefficient of the corresponding element at the AIRS. In practice, we usually set the value of all amplitude reflection coefficients to 1 to achieve the whole reflection. $\phi_i$ represents the phase shift coefficient of the $i$-th element at the AIRS. Here, we assume the AIRS has a continuous phase shift,
so $\theta_i \in [0, 2\pi), i = 1, 2, \ldots, R$. $w_n$ is the beamforming vector at the $n$-th AP, which satisfies $|w_n|^2 \leq 1$. $\delta$ is the additive white Gaussian noise (AWGN) at the user.

### B. CHANNEL MODEL
Suppose that both APs and the AIRS are sub-wavelength uniform linear arrays (ULAs), the spacing between each antenna of the AP and each element of the AIRS is $d = 0.5\lambda$, and $\lambda$ is the wavelength. The antennas at each AP are placed in position $l$, and the AIRS is placed parallel to the ground. So the baseband equivalent channels $G_n$ and $h_r$ is expressed as

$$G_n(l) = \sum_{i=1}^{L_1} \sqrt{\beta_G(i)} a_r^H(\epsilon G_n(i))$$

$$h_r(l) = \sum_{i=1}^{L_2} \sqrt{\beta_h(i)} a_t^H(\epsilon h_r(i))$$

where $L_1$ and $L_2$ denote the total number of paths for the channel $G_n$ and $h_r$, respectively. For each channel, we set $i = 1$ to denote the LoS path and $i > 1$ to denote the non-line-of-sight (NLoS) paths. $\phi_{G_n(i)}$ is the angle of arrival (AoA) of the $i$-th path from the $n$-th AP to the AIRS, $\phi_{h_r(i)}$ is the angle of departure (AoD) of the $i$-th path from the $n$-th AP to the AIRS. $a_r(\epsilon G_n(i))$ and $a_t(\epsilon h_r(i))$ are the vectors of antenna arrays. We can express them as follows

$$a_r(\epsilon G_n(i)) = [1, e^{-j2\pi \frac{d}{\lambda} \cos(\phi_{G_n(i)}), e^{-j2\pi \frac{d}{\lambda} \cos(\phi_{G_n(i)}, i > 1} \ldots, e^{-j(R-1)2\pi \frac{d}{\lambda} \cos(\phi_{G_n(i)}, i > 1} \ldots, e^{-j\pi \frac{d}{\lambda} \cos(\phi_{G_n(i)}), i > 1} \ldots, e^{-j\pi \frac{d}{\lambda} \cos(\phi_{G_n(i)}, i > 1} \ldots, e^{-j(R-1)\pi \frac{d}{\lambda} \cos(\phi_{G_n(i)}, i > 1} \ldots, e^{-jM\pi \frac{d}{\lambda} \cos(\phi_{G_n(i)}, i > 1} ]^T$$

$$a_t(\epsilon h_r(i)) = [1, e^{-j2\pi \frac{d}{\lambda} \sin(\phi_{h_r(i)}), e^{-j2\pi \frac{d}{\lambda} \sin(\phi_{h_r(i)}, i > 1} \ldots, e^{-j\pi \frac{d}{\lambda} \sin(\phi_{h_r(i)}), i > 1} \ldots, e^{-j\pi \frac{d}{\lambda} \sin(\phi_{h_r(i)}}, i > 1} \ldots, e^{-j\pi \frac{d}{\lambda} \sin(\phi_{h_r(i)}, i > 1} \ldots, e^{-j(R-1)\pi \frac{d}{\lambda} \sin(\phi_{h_r(i)}, i > 1} \ldots, e^{-jM\pi \frac{d}{\lambda} \sin(\phi_{h_r(i)}, i > 1} ]^T$$

$\beta_n$ is the path loss of the corresponding channel, which can be expressed as

$$\beta_n(d(I)) = \beta_0 \left(\frac{d(I)}{d_0}\right)^{-\alpha}.$$  

$\beta_0$ denotes the path loss at $d_0 = 1$, and $\alpha$ represents the path loss factor.

Since the AIRS is deployed in the air, for each channel, the LoS path is much stronger than the NLoS path, we can assume that all the channels forming APs to the user through the AIRS only have LoS path. Then the channels $G_n$ and $h_r$ can be simplified as

$$G_n(l) = \sqrt{\beta_G(i)} a_r(\epsilon G_n(i))$$

$$h_r(l) = \sqrt{\beta_h(i)} a_t^H(\epsilon h_r(i))$$

### C. PROBLEM FORMULATION
In this article, we mainly consider the user’s achievable rate optimization for the AIRS-aided cell-free massive MIMO system. Therefore, the purpose is to maximize the achievable rate of the user by jointly designing the transmit power allocation vector $\lambda$, the precoding vector of each AP $w_n$, $n = 1, \ldots, N$, the phase shift matrix $\Theta$, and the placement of the AIRS $I, I \in A$. Here we assume all the CSI are known by the CPU. However the CSI is difficult to obtain, channel estimation for IRS-aided cell-free Massive MIMO system will be studied in our future work. When the AIRS is placed in position $l$, the received SNR at the user can be expressed as

$$\gamma(I) = \frac{\left|\sum_{n=1}^{N} h_r(I)\Theta G_n(I)\sqrt{\lambda_n w_n}\right|^2}{\sigma^2}.$$  

The achievable rate of the user can be expressed as

$$Ra(I) = \log_2\left(1 + \frac{\left|\sum_{n=1}^{N} h_r(I)\Theta G_n(I)\sqrt{\lambda_n w_n}\right|^2}{\sigma^2}\right).$$

Thus, the objective function is formulated as

$$\text{P1 :}$$

$$\max_{\lambda, w_n, \Theta} Ra(I) = \log_2\left(1 + \frac{\left|\sum_{n=1}^{N} h_r(I)\Theta G_n(I)\sqrt{\lambda_n w_n}\right|^2}{\sigma^2}\right)$$

s.t. $\sum_{n} \lambda_n = 1$;

$\|w_n\|^2 \leq 1, n = 1, \ldots, N$;

$\theta_r \in [0, 2\pi)$, $r = 1, \ldots, R$.

### III. OPTIMIZATION STRATEGY
As we can see from (P1), it is very difficult to optimize all system parameters at the same time, so we will use a gradual optimization strategy. First, we consider the optimization strategy of AIRS with a fixed position. Then, we expand to the case to an AIRS with a variable position.

#### A. OPTIMIZATION FOR AIRS WITH FIXED POSITION
In this case, we do not consider the impact of AIRS location on system performance and assume that the AIRS is placed in a fixed position. So, the objective function (P1) can be simplified to

$$\text{P2 :}$$

$$\max_{\lambda, w_n, \Theta} Ra = \log_2\left(1 + \frac{\left|\sum_{n=1}^{N} h_r\Theta G_n\sqrt{\lambda_n w_n}\right|^2}{\sigma^2}\right)$$

s.t. $\sum_{n} \lambda_n = 1$;

$\|w_n\|^2 \leq 1, n = 1, \ldots, N$;

$\theta_r \in [0, 2\pi), r = 1, \ldots, R$. 

VOLUME 9, 2021
We will use an interactive iterative optimization method to solve (P2). Since there is only one user in the system and all APs serve the user simultaneously, there is no interference between different users. When the power distribution coefficient of each AP $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_N]$ and the phase shift matrix $\Theta$ are determined, the optimal precoding vector of each AP is the maximum ratio transmission (MRT) precoding [17], which can be written as

$$w_n^{opt} = \frac{(h_n \Theta G_n)^H}{\|h_n \Theta G_n\|_2}, \quad n = 1, \ldots, N. \tag{14}$$

As we can see from (14), the $w_n$ mainly depend on the channel $G_n$, thus we can simplified $w_n^{opt}$ as

$$w_n^{opt} = \frac{a_r(\phi_{G_n})}{\|a_r(\phi_{G_n})\|_2} \tag{15}$$

This method is studied in [12], which means we just beamforming the signal at each AP to the AIRS and do not consider the channel $h$. We will compare the two different beamforming methods later.

When $w_n$ and $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_N]$ are fixed, the channel $h$ is determined, which can be expressed as

$$h = \sum_{n=1}^{N} G_n \sqrt{\lambda_n} w_n. \tag{16}$$

Then, the problem (P2) becomes

$$P3: \max_{\Theta} \log_2(1 + P \frac{|h_r \Theta h|}{\sigma^2}) \quad \text{s.t.} \quad \theta_r \in [0, 2\pi]. \tag{17}$$

To solve (P2), we need to set all the values of the diagonal matrix $\Theta$ so that the reflected signals are coherently superimposed at the user. Therefore, the value of $\theta_n$ should satisfy the following equation:

$$\theta_n^{opt} = \Omega - \text{arg}([h_r]_n [h]_n), \quad n = 1, \ldots, R. \tag{18}$$

$\Omega$ is a reference alignment phase; here, we can set $\Omega = \pi$. $[h]_n$ and $[h_r]_n$ are the $n$-th elements of $h$ and $h_r$. When $\theta_n$, $n = 1, \ldots, R$ is determined, we can use the current gain of $q_n = h \Theta G_n$ to reallocate the transmit power of each AP.

$$\lambda_n^{opt} = \frac{\|q_n\|_2^2}{\sum_{n=1}^{N} \|q_n\|_2^2}, \quad n = 1, \ldots, N. \tag{19}$$

To obtain better achievable rate performance, we need to repeat the above steps until the achievable rate of the user converges. The specific algorithmic implementation is in Algorithm 1.

**B. OPTIMIZATION OF VARIABLE POSITION**

Now we consider the optimization of the AIRS with a variable position. If the AIRS can move freely in $\mathcal{A}$, we need to find the optimal placement to maximize system performance for the user’s achievable rate. So, the objective function is again (P1).

**Algorithm 1** Interactive Iterative Optimization Algorithm (A1)

1. Initialize $\lambda$ and $\Theta$. Obtain the initial $w_n, n = 1, \ldots, N$ by (14) or (15).
2. Obtain $h$ by (16)
3. for $r = 1 \rightarrow R$
4.   Optimize $\Theta$ by (18).
5. end
6. Obtain $q_n, n = 1, \ldots, N$, $q_n = h \Theta G_n$. Reallocate the transmit power according to (19). Obtain a new $\lambda$. Obtain the optimized $w_n, n = 1, \ldots, N$ by (14) or (15).
7. Calculate the achievable rate $Ra$ by (11). Check convergence, if yes, stop; if no, go to Step 2.

To solve this problem, we need to consider the impact of AIRS placement. If the AIRS can be placed in any position, the problem will be easy to solve. In this case, the optimal height $h$ is the lowest in $\mathcal{A}$ that has the shortest distance to both the AIRS and the user. However, the AIRS cannot be placed anywhere due to the flight height limit, and the placement requires it avoids building shelter. $\mathcal{A}$ usually does not include all areas of the air and tends to be irregular. It is difficult to optimize all the parameters at the same time. We find that if the placement of the AIRS is determined, the optimization problem returns to the fixed placement case, and we can use Algorithm 1 to optimize the other parameters. In other words, in each location in $\mathcal{A}$, we can obtain an optimized achievable rate for the user. One simple method to solve this problem is the exhaustive search. Specifically, we can calculate the achievable rates of all locations in $\mathcal{A}$ and then choose the point with the highest achievable rate as the optimal location of the AIRS. The algorithmic implementation is in Algorithm 2.

**Algorithm 2** Exhaustive Search Algorithm (A2)

1. List all legal location points for the AIRS $(x_c, 0, h_c) \in \mathcal{A}, c = 1, \ldots, C$.
2. For each point $(x_c, 0, h_c), c = 1, \ldots, C$, use Algorithm 1 to calculate the corresponding achievable rate $Ra_c, c = 1, \ldots, C$.
3. Choose the point with the maximum achievable rate $(x^*, 0, h^*)$ as the optimal location of the AIRS.

The exhaustive search algorithm seems easy to implement. However, it often requires substantial computing resources and time because the area $\mathcal{A}$ is very large. To obtain the optimal placement of the AIRS, we need to pass through all of the locations in $\mathcal{A}$, which in most cases is unrealistic. Therefore, we propose a fast optimal location search algorithm to optimize the placement of AIRS.

It can be seen from (15) that for each location $l$ in $\mathcal{A}$, the corresponding optimal precoding vector of the
n-th AP is
\[ w_n(t) = \frac{\mathbf{a}_t(\phi_{G_n(t)})}{\|\mathbf{a}_t(\phi_{G_n(t)})\|_2} \] (20)

The path loss of channels \( \mathbf{G}_n \) and \( \mathbf{h}_t \) can be expressed as follows:
\[ \beta_{G_n(t)} = \frac{\beta_0}{\| I - I_n \|_2}, \quad n = 1, \ldots, N, \] (21)
\[ \beta_{h_t} = \frac{\beta_0}{\| I - I_0 \|_2}. \] (22)

Taking (20)-(22) into (11), we can obtain the optimal achievable rate \( R_a \) of the user when the AIRS is placed at \( I \), which can be expressed as
\[
R_a = \log_2(1 + \frac{MP}{\sigma^2} \frac{\beta_0}{\| I - I_0 \|_2} \| \mathbf{a}_t(\phi_{G_1(t)}) \|_2^2) \times \Theta(\sqrt{\lambda_1 \| I - I_1 \|_2^2} \mathbf{a}_t(\phi_{G_1(t)})
+ \sqrt{\lambda_2 \| I - I_2 \|_2^2} \mathbf{a}_t(\phi_{G_2(t)}) + \ldots + \sqrt{\lambda_N \| I - I_N \|_2^2} \mathbf{a}_t(\phi_{G_N(t)})^2)
= \log_2(1 + \frac{MP \beta_0}{\sigma^2} |B_1 A_1 + B_2 A_2 + \ldots + B_N A_N|^2) \] (23)

where
\[ A_n = \sqrt{\lambda_n} \mathbf{a}_t(\phi_{G_n(t)}) \Theta_{\mathbf{a}_t}(\phi_{G_n(t)}), \quad n = 1, \ldots, N, \] (24)
\[ B_n = \frac{1}{\| I - I_0 \|_2 - I_n \|_2}, \quad n = 1, \ldots, N. \] (25)

It is difficult to find the maximum value of such a function directly. Therefore, we will use an approximate optimal solution to find the optimal AIRS placement. From the expression of \( A_n \), we find that if \( N = 1 \), then there is only one AP serving the user by AIRS. When we find the optimal phase shift matrix of the AIRS, \( A_1 = R \) [12]. In this case, \( A_n \) depends on the number of AIRS elements \( R \). Therefore, in any location of \( A_n \), when the optimal \( \Theta \) is determined, the achievable rate depends on \( B_1 \). \( N > 1 \) means that there is more than one APs serving the user. Although both \( A_n \) and \( B_n \) change with location \( I \), compared with \( B_n \), \( A_n \) is much less sensitive to location because when \( I = I_n, n = 1, \ldots, N \), the denominator of \( B_n \) is close to 0. Therefore, we only need to consider the following equation
\[ D(I) = B_1 + B_2 + \ldots + B_N \]

\[ \text{FIGURE 2. The impact of } h \text{ on } B_1. \]

Let \( I = (x, 0), I_0 = (w_0, 0, 0), I_n = (w_n, 0, 0), n \) \( = 1, \ldots, N \); then, for each \( I_n = (w_n, 0, 0), n \) \( = 1, \ldots, N \), we have the following equation:
\[
B_n = \frac{1}{\sqrt{\| I - I_0 \|_2 - I_n \|_2}}
= \frac{1}{\sqrt{(x-h^2)^2 + (x-w_n^2)^2}}
= \left( \frac{1}{\sqrt{h^2 + t^2}} \right)^2 (x - w_n^2)^2 \] (27)

where \( t = x - w_0, \Delta = w_n - w_0 \). After polynomial expansion, for a given \( h \), using the Cardano formula in [12], the optimal \( x^* \) is
\[
\left( \frac{1}{2} \pm \frac{1}{4} - \frac{h}{\Delta} \right) \Delta + w_0 \quad \text{where} \quad \sigma_h \leq \frac{1}{2} \] (28)

where \( \sigma_h = \frac{h}{\Delta} \). So the optimal horizontal position is depend on \( \sigma_h \).

We draw a function graph of \( B_1 \) to verify the conclusion in (24), we set \( I_1 = (100, 0, 0), I_0 = (400, 0, 0) \) and \( I = (x, 0, h), x \epsilon [0, 500], h \epsilon [50, 100, 150, 200] \). As shown in Fig. 2, with the increase in \( h \), the maximum-value points of \( B_1 \) change from \( \left( \frac{1}{2} \pm \sqrt{\frac{1}{4} - \left( \frac{h}{\Delta} \right)^2} \right) \) \( \Delta + w_0 \) to the midpoint
\[ \frac{1}{2} \Delta + w_0 \]. If \( \frac{h}{\Delta} \leq 0.5 \), the optimal \( x^* \) is close to \( AP_n \) or the user. If \( \frac{h}{\Delta} \geq 0.5 \), the optimal \( x^* \) is the midpoint of \( AP_n \) and the user. \( D(I) \) is the sum of \( B_n, n = 1, \ldots, N \). We use \( B_m \) to denote the largest one among \( B_n, n = 1, \ldots, N \). In fact, \( B_m \) is the nearest AP to the AIRS,
so $B_m$ determines the maximum-value point of $D(I)$. Usually, to maximize reflection performance, the AIRS should not be placed in the AP area; the optimal location of the AIRS should between $AP_m$ and the user. This is because other APs are far away from the user, the path loss is higher if we place the AIRS between these APs and the user. Thus, we can simplify and use $B_m$ to optimize $I$ instead of using $B_n$, $n = 1, \ldots, N$.

For each AIRS placement height $h$, we can find the optimal $x^*$ according to (28). Then, we calculate the optimal locations corresponding to all heights and the achievable rate of users when the AIRS is placed at these locations. We can choose the point with the largest optimized achievable rate as the optimal placement of the AIRS.

Although calculating all the heights of $A$ seems complicated, the range for the AIRS altitude $h$ is smaller than that for the AIRS horizontal position $x$ due to the limited flight altitude of AIRS. Thus, for a given $h$, we only need to calculate several horizontal positions of $A$ and not all the horizontal positions. The computing resources and time are greatly reduced using this method. The specific algorithm is in Algorithm 3.

**Algorithm 3 Fast Optimal Location Search Algorithm (A3)**

**Require:** Range of $A$; horizontal coordinate of each AP $x_n$, $n = 1, \ldots, N$; horizontal coordinate of user $x_0$.

**Ensure:** The optimal height $h^{opt}$ and the optimal horizontal coordinate $x^{opt}$ of the AIRS

1. For each height $h \in A$, find the horizontal range of $A$ which can be expressed as $[x_{min}^h, x_{max}^h]$. Find the coordinate of closest AP to the AIRS $x_m$, $\Delta = x_m - x_0$;
2. For $h = H_{min} \rightarrow H_{max}$ in $A$, obtain $\sigma_h = |\frac{h}{\Delta}|$
   if $\sigma_h > 0.5$, the theoretical optimal horizontal coordinate of the AIRS at height $h$ is $x_h^* = \frac{1}{2}\Delta + x_0$.
   else the optimal $x$ of height $H$ is $x_h^* = \left(1 - \sqrt{\frac{1}{4} - \left(\frac{h}{\Delta}\right)^2}\right)\Delta + x_0$.

   Obtain $x_h^*, h \in [H_{min}, H_{max}]$ in $A$.

3. For $h = H_{min} \rightarrow H_{max}$ in $A$.
   Check the relationship between $x_h^*$ and $[x_{min}^h, x_{max}^h]$.
   if $x_h^* \in [x_{min}^h, x_{max}^h]$, the optimal $x_h^* = x_h^*$. If there are two optimal $x_h^*$, choose the $x_h^*$ with a higher achievable rate using Algorithm 1.
   else compare the achievable rate at $x_{min}^h, x_{max}^h$, and $x_h^*$ and choose the one with the highest achievable rate as $x_h^*$.

4. Find all $x_h^*, h \in [H_{min}, H_{max}]$. Calculate the optimized achievable rate of the corresponding position $(x_h^*, h)$ by Algorithm 1, choose the $x_h^*, h^* \in [H_{min}, H_{max}]$ with the largest achievable rate as the optimal placement of the AIRS. $h^{opt} = h^*, x^{opt} = x_h^*$.

**IV. SIMULATION RESULTS**

In this section, we will conduct a simulation in an actual scenario to verify the performance of the proposed optimization strategy. As shown in Fig. 3, there are 4 APs and a single-antenna user in this system. Due to the shelter of buildings and the uneven distribution of APs, the user is in a “shadow area”, and we use an AIRS to assist the communication between the APs and the user. The AIRS is deployed on a UAV, so it can move in the air. We set up a three-dimensional rectangular coordinate system based on the line from the AP closest to the user to the user as the $x$-axis. All APs are placed on the $x$-$o$-$y$ plane. The AIRS is placed on the $x$-$o$-$z$ plane.

**FIGURE 3. Simulation scene.**

**A. FIXED PLACEMENT**

First, we will consider the AIRS with a fixed position. We set the AIRS to be at a fixed location to verify the performance of Algorithm 1. The parameter settings are summarized in detail in Table 1.

| Parameters | Value |
|------------|-------|
| Number of APs $N$ | 4 |
| Number of antennas at each AP $M$ | 10 |
| Number of elements at the AIRS $K$ | 100 |
| Bandwidth $B$ | 1 MHz |
| Power spectral density of the noise | $-170$ dBm/Hz |
| Path loss exponent $\alpha$ | 2.2 |
| Reference distance $d_0$ | 30 dB |
| The Rician factor $c$ | 10 |
| The coordinates of $AP_1$ (m) | (200.00, 100.00) |
| The coordinates of $AP_2$ (m) | (300.00, 0.00) |
| The coordinates of $AP_3$ (m) | (600.00, 0.00) |
| The coordinates of $AP_4$ (m) | (400.100.00) |
| The coordinates of the AIRS (m) | (900.0, 150.0) |
| The coordinates of the user (m) | (1200.0, 0.0) |

In Fig. 4, we analyze the performance of Algorithm 1. We compared four different settings:

1) **Without any optimization:** we do not optimize the phase shift matrix $\Theta$ of the AIRS or the power allocation vector $\lambda$ of the APs. All the APs are allocated equal power, and each AP performs MRT.
2) **Only power allocation optimized**: In this case, we optimize only the power allocation vector \( \lambda \). The phase shift matrix \( \Theta \) of the AIRS is randomly set, and each AP performs MRT.

3) **Only phase shift optimized**: In this case, we allocate equal power to each AP and optimize the phase shift matrix \( \Theta \) of the AIRS. Each AP conducts MRT.

4) **Proposed algorithm with whole channel beamforming**: In this case, we optimize the power allocation vector \( \lambda \) the beamforming vectors \( w_n \) and the phase shift matrix \( \Theta \) of the AIRS using the proposed Algorithm 1 in section III. We use (14) to get the optimal \( w_n \).

5) **Proposed algorithm with local channel beamforming**: we optimize the achievable rate using the proposed Algorithm 1, the difference is that we use (15) to get the optimal \( w_n \).

As shown in Fig. 5, compared with no optimization, we obtain a significant performance gain with the proposed Algorithm 1. Compared with optimizing the power allocation vector \( \lambda \) alone, we obtain a higher achievable rate after optimizing the phase shift matrix \( \Theta \). This is because if we do not optimize \( \Theta \), the signals cannot be coherently superimposed on the user, and the use of the AIRS increases the transmission distance, which leads to a larger path loss. Using the proposed Algorithm 1 can realize a significant performance increase in the user’s achievable rate. As we can see from Fig. 5, the local beamforming method has better achievable rate performance than the whole beamforming method.

In Fig. 5, we study the convergence of Algorithm 1. We set the total transmit power to 20 dbm. We find that Algorithm 1 has good astringency; after several iterations, we obtain a good achievable rate performance.

### B. UNFIXED PLACEMENT

Next, we turn to the optimization of AIRS with variable positions. As shown in Fig. 3, due to the flight height of AIRS itself and the limitation of building occlusion, the AIRS’s mobile range is limited to \( A' \). In this case, we consider the optimization of the horizontal coordinate (x-axis) and the height of the AIRS (z-axis). We will use Algorithm 3 to optimize AIRS placement. The specific distance parameters are listed in Table 2.

#### TABLE 2. The parameter settings.

| Parameters                        | Value       |
|-----------------------------------|-------------|
| Height of shelter \( h_0 \) (m)   | 70          |
| \( h_1 \) (m)                     | 77          |
| Maximum flight altitude of AIRS (m) | 200        |
| The reference coordinate of the AIRS (m) | (900,0,150) |
| The coordinate of the user (m)    | (12000,0,0) |
| Total transmit power \( P^* \)    | 20 dbm      |
| \( d_1 \) (m)                     | 350         |
| \( d_2 \) (m)                     | 200         |
| \( d_r \) (m)                     | 50          |

In Fig. 6, we compare the accuracy of the proposed fast optimal location search algorithm A3 with that of the exhaustive search algorithm A2. For each height of \( A' \), we can use A2 to accurately find an optimal placement point. However, exhaustive search requires substantial computing time and computing resource overhead because we need to calculate the optimized achievable rate for each point in \( A' \). First, for each height of \( A \), we use A2 to find the point with the highest achievable rate and record the maximum achievable rate. Then, we use A3 to find the optimal points and obtain the maximum achievable rate at each height. Finally, we compare the difference between the maximum achievable rates found by the two algorithms at each altitude. As we can see from Fig. 6, the user’s achievable rate at the optimal horizontal point at each height is in agreement with the result of the exhaustive search using Algorithm 2 approximately. This means that using the proposed A3, we can obtain a performance gain approximate to that of A2 without the exhaustive search of regions so the amount of calculation has been reduced.
In Fig. 7, we compare the user’s achievable rate performance given position optimization and non-optimization. In the case of non-optimization, we randomly set the reference coordinate to \( I = (900, 0, 150) \). Then, we use Algorithm 3 to obtain the optimal coordinate \( I^{opt} = (1160, 0, 112) \). From Fig. 7, we find that after optimizing the placement of the AIRS, we can obtain further performance improvement in terms of the achievable rate under the same total transmit power constraint. In other words, we can optimize the placement to reduce the total transmit power.

In Fig. 8, we study the influence of different element numbers for the AIRS \( R \) on the system performance. At the same total transmit power, we can increase \( R \) to improve the user’s achievable rate rather than increase the total transmit power. Furthermore, if we optimize the placement of the AIRS, the performance can be further improved. In practice, we can increase \( R \) and optimize the AIRS’s placement to obtain higher service quality for the user and save transmit power. However, compared with increasing the number of elements of the AIRS, we can obtain significant performance gain by optimizing the placement of the AIRS. In other words, we can have fewer elements by optimizing their placement. Thus, it is necessary to optimize the placement of the AIRS.

V. CONCLUSION

In this article, we initially consider the combination of an AIRS and a cell-free massive MIMO and study an AIRS-assisted cell-free massive MIMO system. The mobility of the AIRS means that we can use it to dynamically adjust and control the service quality of the system to meet the needs of users in a cell-free massive MIMO system. With the addition of the AIRS, the coverage, energy consumption and spectrum efficiency of future wireless networks will be greatly improved. The purpose of this article is to determine how to maximize the user’s achievable rate by optimizing the power allocation for each AP, the beamforming vector at each AP, the placement of the AIRS and the reflection phase shift of the AIRS. First, we assume that the AIRS is placed at a fixed position, and we use an interactive iterative optimization algorithm to jointly design the power allocation vector, the precoding vector at each AP and the phase shift matrix of the AIRS. Then, we consider the AIRS with a variable position, and we propose a fast optimal location search algorithm to obtain the optimal placement of the AIRS. The simulation results show that we can obtain a significant performance improvement in the achievable rate by using the proposed methods.

This article first applies the AIRS to a cell-free massive MIMO system. However, the research of this article needs to be further improved. In this article, we only consider the single-user case, and the multi-user case needs to be studied. We also assume that the AIRS is placed in the x-o-z plane and only optimize the height and horizontal coordinates of the AIRS. The optimization needs to be extended to the whole three-dimensional coordinate system. The above problems will be our future work.
REFERENCES

[1] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” IEEE Trans. Wireless Commun., vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[2] E. Basar, M. D. Renzo, J. D. Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, “Wireless communications through reconfigurable intelligent surfaces,” IEEE Access, vol. 7, pp. 116753–116773, 2019.

[3] C. Liaskos, S. Nie, A. Tsiliaridou, A. Pitsillides, S. Ioannidis, and I. Akylidiz, “A new wireless communication paradigm through software-controlled metasurfaces,” IEEE Commun. Mag., vol. 56, no. 9, pp. 162–169, Sep. 2018.

[4] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, “Massive MIMO is a reality—What is next?: Five promising research directions for antenna arrays,” Digit. Signal Process., vol. 94, pp. 3–20, Nov. 2019.

[5] M. D. Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin, J. de Rosny, A. Bounceu, G. Lerosey, and M. Fink, “Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come,” EURASIP J. Wireless Commun. Netw., vol. 2019, no. 1, pp. 1–20, 2019.

[6] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” IEEE Commun. Mag., vol. 58, no. 1, pp. 106–112, Jan. 2020.

[7] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming,” IEEE Trans. Wireless Commun., vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[8] E. Björnson, O. Ozdogan, and E. G. Larsson, “Intelligent reflecting surface versus decode- and-forward: How large surfaces are needed to beat relaying?” IEEE Wireless Commun. Lett., vol. 9, no. 2, pp. 244–248, Feb. 2020.

[9] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1–6.

[10] S. Zhang and R. Zhang, “Capacity characterization for intelligent reflecting surface aided MIMO communication,” IEEE J. Sel. Areas Commun., vol. 38, no. 8, pp. 1823–1838, Aug. 2020.

[11] W. Wang, J. Tang, N. Zhao, X. Liu, X. Y. Zhang, Y. Chen, and Y. Qian, “Joint precoding optimization for secure SWIPT in UAV-aided NOMA networks,” IEEE Trans. Commun., vol. 68, no. 8, pp. 5026–5040, Aug. 2020.

[12] H. Lu, Y. Zeng, S. Jin, and R. Zhang, “Enabling panoramic full-angle reflection via aerial intelligent reflecting surface,” in Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops), Dublin, Ireland, Jun. 2020, pp. 1–6.

[13] Z. Chen and E. Björnson, “Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry,” IEEE Trans. Commun., vol. 66, no. 11, pp. 5205–5219, Nov. 2018.

[14] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, “Precoding and power optimization in cell-free massive MIMO systems,” IEEE Trans. Wireless Commun., vol. 16, no. 7, pp. 4445–4459, Jul. 2017.

[15] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, “Cell-free massive MIMO: Uniformly great service for everyone,” in Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC), Jun. 2015, pp. 201–205.

[16] E. Björnson and L. Sanguinetti, “Scalable cell-free massive MIMO systems,” IEEE Trans. Commun., vol. 68, no. 7, pp. 4247–4261, Jul. 2020.

[17] D. Tse and P. Viswanath, Fundamentals of Wireless Communication. Cambridge, U.K.: Cambridge Univ. Press, 2005.

**TAO ZHOU** was born in 1997. He received the B.E. degree from Sichuan University in 2019, where he is currently pursuing the M.S. degree with the Institution of Communications Engineering, PLA Army Engineering University. His research interests include MIMO techniques, cell-free massive MIMO, and intelligent reflecting surface.

**KUI XU** (Member, IEEE) was born in 1982. He received the B.E. degree in wireless communications and the Ph.D. degree in software-defined radio from the PLA University of Science and Technology, Nanjing, China, in 2004 and 2009, respectively. He is currently an Associate Professor with the College of Communications Engineering, PLA Army Engineering University. His research interests include broadband wireless communications, signal processing for communications, network coding, and wireless communication networks. He has authored about 100 articles in refereed journals and conference proceedings. He holds five patents in China.

Prof. Xu served on the Technical Program Committee of the IEEE WCSP2014, WCSP2015, ICPADM2015, and SIRS2015 TPC. He received the 2010 Ten Excellent Doctor Degree Dissertation Award of PLAUST and the URSI Young Scientists Award in 2014. He serves as a Reviewer for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE COMMUNICATIONS LETTERS, and the IEEE SIGNAL PROCESSING LETTERS.

**XIAOCHEN XIA** was born in China in 1987. He received the B.E. degree in electronic science and technology from Tianjin University in 2010 and the Ph.D. degree in communication and information system from the PLA University of Science and Technology in 2017. His research interests include relaying networks, full-duplex communication, network coding, MIMO techniques. He received the 2013 Excellent Master Degree Dissertation Award of Jiangsu Province, China.

**WEI XIE** was born in 1977. He received the B.E. degree in microwave communication from the Institute of Communications Engineering, Nanjing, China, in 1999, and the M.S. degree in communications and information system from the PLA University of Science and Technology, Nanjing, in 2002. He is currently an Associate Professor with the College of Communications Engineering, PLA Army Engineering University. His research interests include broadband wireless communications, signal processing for communications, network coding, and wireless communication networks.

**JIANHUI XU** was born in 1990. She received the master’s degree in communication engineering from the PLA University of Science and Technology, Nanjing, China, in 2016. She is currently a Lecturer with PLA Army Engineering University. Her research interests include wireless communications and network coding.