Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 4044783, 10 pages
https://doi.org/10.1155/2022/4044783

*Research Article*

# Pilot Allocation and Data Power Optimization Based on Access Point Selection in Cell-Free Massive MIMO

**Zhiwen Duan [1] and Feng Zhao [2]**

[1]The Key Laboratory of Cognitive Radio and Information Processing, Guilin University of Electronic Technology, Guilin 541004, China
[2]The Guangxi Colleges and Universities Key Laboratory of Complex System Optimization and Big Data Processing, Yulin Normal University, Yulin 537000, China

Correspondence should be addressed to Feng Zhao; zhaofeng@guet.edu.cn

Received 8 February 2022; Accepted 18 March 2022; Published 7 April 2022

Academic Editor: Yingjie Wang



Focusing on the pilot contamination problem of cell-free massive multi-input multi-output (MIMO), a pilot allocation algorithm based on access point (AP) selection is proposed. To improve the system performance further, data power optimization is carried out. First, the AP that serves each user is selected through the user-centered idea and the large-scale fading matrix. Use the same service AP numbers between users to measure the pilot contamination intensity and then assign orthogonal pilots to strong potential pilot contamination users. A spectrum efficiency (SE) scheme to maximize the minimum user is proposed for the fairness problem in data power optimization. It is proved that the objective function of data power optimization accords with linear programming, and a dichotomy is given to solve the problem. Simulation results show that the pilot allocation algorithm based on AP selection can significantly improve the total SE of the system and reduce the computational complexity of the system. At the same time, the data power optimization algorithm improves minimum SE for users in the system.

## 1. Introduction

With the progress and development of the times, wireless communication has facilitated people's lives, and the amount of wireless data has also increased significantly. For example, wireless sensors collect a large amount of data in the physical world; new intelligent cyberphysical systems (CPS) collect data in different dimensions: the widespread use of the Internet of Things in smart cities and industrial 4.0 [1–6]. In the future, a large amount of wireless data will put forward broader connection requirements for 6G and even higher versions of wireless communication technology. It also put forward higher performance requirements. To meet the requirements of high energy efficiency, high SE, and ubiquitous network connection requirements for the next-generation wireless communication, especially the communication demand of high-density user (UE) scenario, some scholars put forward the concept of cell-free massive multi-input multi-output (CF massive MIMO), and this technology has become one of the critical technologies of 6G [7]. Through the distributed deployment of APs, CF massive MIMO reduces the distance between UE and AP, providing consistent and good service for UE. At the same time, because there is no boundary of cellular network, it avoids the problems that UEs have to switch cell services frequently, and the boundary service quality is poor. It also has the characteristics of channel hardening, strong macrodiversity, and the ability to resist multiuser interference [8–10]. However, when AP acquires channel state information (CSI) through the pilot information sent by UEs, due to many UEs and the limitation of the coherence time, the pilot will be multiplexed among multiple UEs, which leads to the problem of pilot contamination. Pilot contamination is the obstacle and bottleneck to improve the performance of CF massive MIMO [11].

To solve the pilot contamination problem of CF massive MIMO and improve SE, researchers mainly design pilot allocation and data power optimization schemes. References

[7, 12, 13] use prior information, such as user location and large-scale fading matrix, to design pilot allocation methods, while references [14–16] design a power optimization scheme from the aspects of reducing power consumption and improving SE. A greedy pilot allocation algorithm is proposed in reference [7]. It can improve the SE of the worst users. However, the initial random pilot allocation ignores the potential pilot contamination between users, which may not substantially improve the system's performance in subsequent iterations. Reference [12] proposes a pilot allocation algorithm based on tabu search. An iterative algorithm is constructed to avoid optimal local results by defining the domain, searching the objective function, and introducing a taboo list. Finally, it is proved that the pilot allocation method outperforms the random and greedy pilot allocation methods. Reference [13] proposed a pilot allocation scheme based on the Hungarian algorithm. First, several users are selected by the size of the large-scale fading matrix, the number of users chosen is equal to the number of orthogonal pilots, and then these pilots are assigned to users. Then, an optimization problem is constructed, and finally, the Hungarian algorithm is used to solve the problem. In reference [14], an optimization problem of minimizing the total transmission power under the condition of satisfying the user's quality of service is proposed. It is proved that the problem is a linear programming problem, and the optimal global solution can be found in polynomial time. Reference [15] proposed a fractional power control method, which determines the power control coefficient by calculating the ratio of the large-scale fading value of a single user to the sum of the large-scale fading values of all users. This method can effectively suppress the interference between users. In the case that both the user and AP have multiple antennas, the reference [16] adopts the power distribution strategy of sum-rate maximization and minimum rate maximization on the uplink and downlinks, respectively. By solving the power optimization problem on the uplink and downlinks, the total data rate of the uplink and the fairness of the user performance of the downlink are improved, respectively. From the reference [14–19], it can be seen that data power optimization can improve the system's performance, such as improving the fairness among users and improving the system's overall performance. For the pilot allocation problem in reference [7, 12, 13], it is assumed that all AP services to every user will increase the computational complexity of the system. It is necessary to eliminate pilot contamination in CF massive MIMO and consider the practicability of dense users. To solve this problem, with the idea of user-centered, the literature [17, 18] can effectively reduce the computational complexity of the system by selecting part of AP to serve users. Still, it does not consider the pilot contamination problem among users after choosing AP. Therefore, In the CF massive MIMO, it is necessary to perform AP selection, allocate the pilot, and optimize the data power simultaneously.

In this paper, we study the advantages and existing challenges of the CF massive MIMO uplink system and propose a pilot allocation algorithm based on AP selection to suppress the influence of pilot contamination and improve the system's SE. Lastly, optimizing data power improves the minimum SE of users in the system. First of all, by analyzing the large-scale fading matrix between users and AP, the AP serving user is selected based on the large-scale fading matrix and user-centered idea. The user AP service matrix is constructed. Then, based on the AP service matrix, the number of the same service AP between the user is obtained, that is, the AP coincidence degree. Users of the same AP service are divided into the same group for pilot allocation. However, groups with more users will reuse more pilots. Therefore, groups with more users are preferentially selected for pilot allocation so as to use as many mutually orthogonal pilots as possible to reduce interference. The AP coincidence degree measures the potential pilot contamination intensity between UEs for the same group of UEs. The UEs with high pilot contamination intensity prioritize assigning orthogonal pilots. For the problem of fairness, a data power optimization scheme to max-min user's SE is proposed. The data power optimization scheme is proved to belong to the linear programming problem. And the dichotomy is used to solve the problem. Simulation results show that the pilot allocation and data power optimization method based on AP selection can effectively suppress the pilot contamination problem and improve the system's total SE. It also can reduce the computational complexity of the system and improve the fairness between users.

## 2. System Model

This paper studies the uplink system of CF massive MIMO based on time division duplexing (TDD) mode and the system model (see Figure 1) [19].

This system is mainly composed of $M$ APs, $K$ UEs, and several CPUs; each UE and AP are assumed to be a single antenna. AP is randomly assigned to the covered area, assuming that the UE is at low speed or static state. All APs are connected to the CPU through the backhaul link and perform signal processing in the CPU. In the case of partial AP serving users, the CPU uses the selection information of AP to select the signals collected by partial AP for user channel estimation and signal detection.

Suppose that the total length of each coherent block is $T = \tau_p + \tau_u + \tau_d$, where $\tau_p$ is the pilot length, $\tau_u$ is the length of the uplink data transmission, and $\tau_d$ is the length of data transmission in the downlink. The channel matrix between the $k$-th user and the $m$-th AP is expressed by $g_{m,k}$. It is assumed that the channel is a correlated Rayleigh fading channel, then $g_{m,k}$ is the large-scale fading matrix and $g_{m,k} \sim N_{\mathbb{C}}(0, R_{m,k})$.

In the system setting, it is assumed that AP and UE do not have a priori CSI at the beginning of the coherent interval; so, channel estimation is needed in each coherent interval. Therefore, the communication in the uplink includes two stages: uplink pilot transmission and uplink data transmission. In the pilot transmission stage, each UE is assigned a pilot. The received pilot information is used for channel estimation. The estimated channel is used to detect the received data, thereby calculating the SE of each UE.

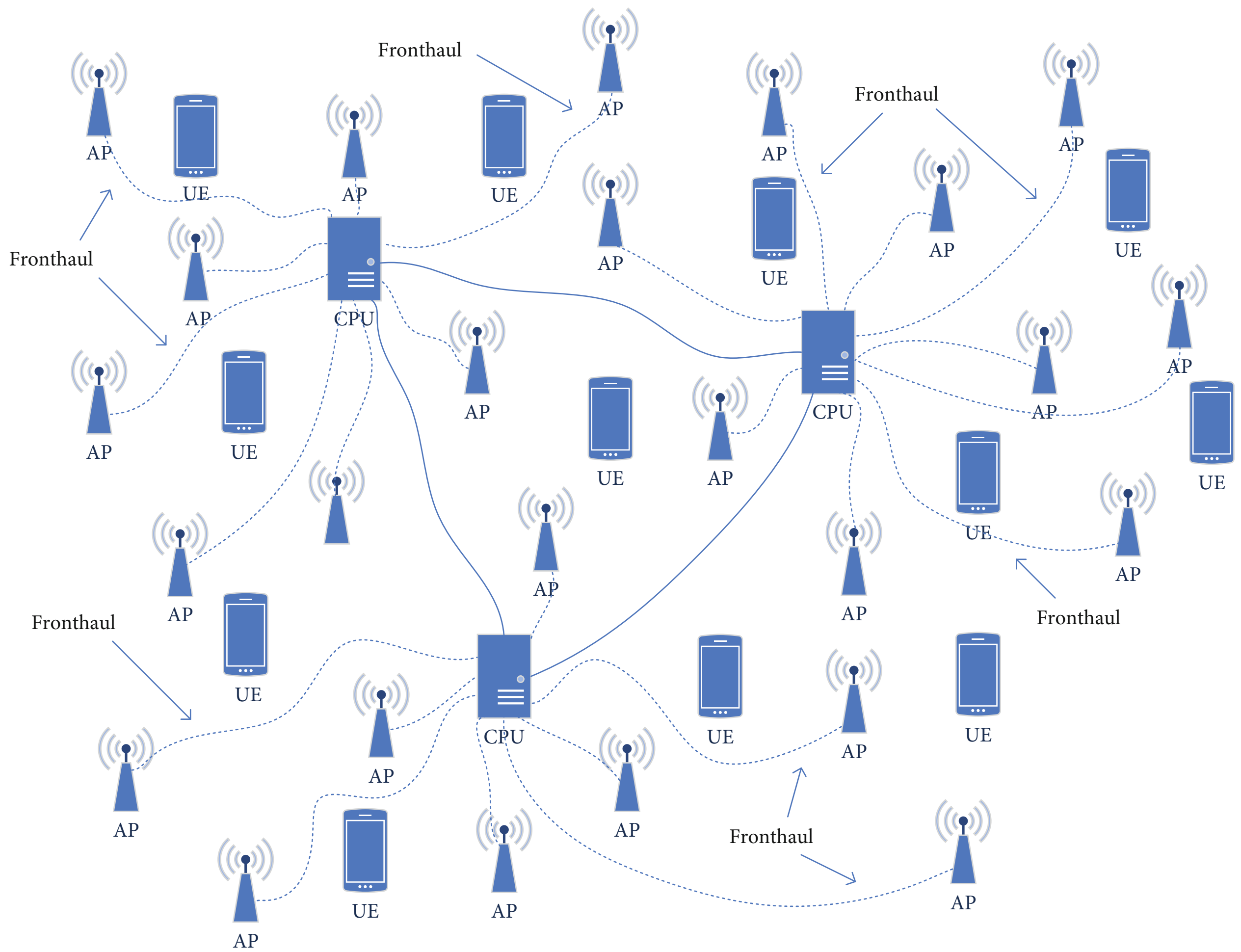


Figure 1: CF massive MIMO network system.

*2.1. Uplink Pilot Transmission and Uplink Data Transmission.* The set of $\varphi^u = \{\varphi_1, \varphi_2 \cdots \varphi_{\tau_p}\}$ represents all orthogonal pilots, and $\|\varphi_t\|^2 = \tau_p$. In the uplink pilot transmission phase, the pilot signal received by the $m$-th AP can be expressed as $y_{p,m}$:

$$y_{p,m} = \sum_{j=1}^{K} \sqrt{p_j^p} g_{m,j} \varphi_{t_j}^{\mathrm{T}} + \mathbf{N}_m, \tag{1}$$

where $p_j^p$ is the pilot transmission power of user $j$, and $\mathbf{N}_m$ represents the noise received by the $m$-th AP. Using minimum mean squared error (MMSE) to estimate the channel, then the estimated channel between the $k$-th user and the $m$-th AP $\widehat{g}_{m,k}$ [20] is

$$\widehat{g}_{m,k} = \sqrt{p_k^p \tau_p} \mathbf{R}_{m,k} \Psi_{t_k,m}^{-1} y_{p,m\,t_k}, \tag{2}$$

where

$$\Psi_{t_k,m} = \sum_{j \in B_t} \tau_p p_j^p \mathbf{R}_{m,j} + \sigma^2 \mathbf{I}, \tag{3}$$

where $B_t$ represents the set of users who use pilot $t$ in the pilot transmission phase. From Equation (3), it can be seen that there will be pilot contamination among users who reuse the same pilot.

In the uplink data transmission phase, similar to the pilot transmission phase, then the data signal received by the $m$-th AP can be expressed as $y_{u,m}$:

$$y_{u,m} = \sum_{j=1}^{K} \sqrt{p_j^u} g_{m,j} x_j + \mathrm{N}_m, \tag{4}$$

where $x_j$ is the data signal sent by user $j$, and $p_j^u$ is the data transmission power of user $j$. Before the AP selection is made, it is necessary to combine all the signals received by the AP to detect the data. Channel estimation, detection matrix, and data detection all adopt MMSE. The achievable SE is [19, 20]

$$\mathrm{SE}_k^{(ul,1)} = \frac{\tau_p}{T} \mathbb{E}\left\{\log_2\left(1 + \mathrm{SINR}_k^{(ul,1)}\right)\right\}, \tag{5}$$

where

$$\mathrm{SINR}_k^{(ul,1)} = p_k^u \widehat{g}_k^{\mathrm{H}} \left( \sum_{j=1, j\neq k}^{K} p_j^u \widehat{g}_j \widehat{g}_j^{\mathrm{H}} + \sum_{j=1}^{K} p_j^u \mathbf{C}_j + \sigma^2 \mathbf{I}_M \right)^{-1} \widehat{g}_k. \tag{6}$$

## 3. Pilot Assignment

*3.1. AP Selection.* For users, the AP located near the user contributes the maximum SE. In contrast, the AP, which is far from the UE, has less gain in macrodiversity. The potential pilot contamination among users who assign the same pilot is mainly related to its large-scale fading matrix. Therefore, AP is selected with the help of a user-centered idea and each user's large-scale fading matrix [21, 22]. To sum up, the AP selection formula for the $k$-th user is

$$\sum_{m=k^{(1)}}^{k^{(Q)}} \frac{\bar{R}_{m,k}}{\sum_{m'=1}^{M} R_{m',k}} \geq \beta\%, \tag{7}$$

where $\{k^{(1)}, k^{(2)} \cdots, k^{(Q)}\}$ represents the $Q$ APs selected by user $k\,(Q \leq K)$, the descending set of the large-scale fading matrix between the UE and the AP is $\{\bar{R}_{k^{(1)},k}, \bar{R}_{k^{(2)},k}, \cdots, \bar{R}_{k^{(Q)},k}\}$, and $\beta$ represents a set constant. Assuming that the set of $Q$ APs selected by the $k$-th user is represented by $A_k$, the corresponding AP service matrix $D_{k,m}$ is defined as

$$D_{k,m} = \begin{cases} 1 & \text{if m} \in A_k, \\ 0 & \text{if m} \notin A_k. \end{cases} \tag{8}$$

The position where the service matrix $D_{k,m} = 1$ represents that the $m$-th AP serves the $k$-th UE. Through the service matrix, the AP coincidence matrix $B \in \mathbb{C}^{K\times K}$ between two users and the matrix $d \in \mathbb{C}^{1\times M}$ of the number of AP service users can be obtained, which is defined as follows:

AP coincidence matrix is as follows:

$$B_{k,k'} = \sum_{m=1}^{M} b_{k,k'}^{m}, \tag{9}$$

where

$$b_{k,k'}^{m} = \begin{cases} 1 & \text{if } D_{k,m} = D_{k',m} = 1 \\ 0 & \text{else} \end{cases}. \tag{10}$$

The number of users served by the m-th AP can be expressed as

$$d(m) = \sum_{k=1}^{K} D_{k,m}, \tag{11}$$

which is an example of the user's AP selection (see Figure 2). The AP of several service users is selected according to the user's large-scale fading matrix. As mentioned earlier, the CPU selects signals received by a part of AP according to the AP selection information for channel estimation and subsequent signal detection. If there is the same AP between users, the channel estimation error will be generated when the same pilot is assigned to the two users. As shown the UE1, UE2, and UE3 in Figure 2, the service matrix $D_{1,1} = 1, D_{1,2} = 1, D_{1,3} = 1, D_{1,4} = 1, D_{2,3} = 1, D_{2,4} = 1, D_{2,5} = 1, D_{3,6} = 1, D_{3,7} = 1, D_{3,8} = 1$, UE1, and UE2 share AP3 and AP4, for coincidence matrix$B_{1,2} = 2$, but there is no case of sharing the same AP between UE2 and UE3, for coincidence matrix $B_{2,3} = 0$. Therefore, the scheme of assigning the same pilot between UE1 and UE2 will cause more pilot contamination and degrade the system's performance than that of UE2 and UE3. The pilot allocation algorithm in this paper is also based on this idea, and the detailed pilot allocation algorithm is described later.

To sum up, the service matrix is a manifestation of the user's choice of service AP. In contrast, the size of the coincidence matrix represents the number of coincident AP between two users and the potential pilot contamination intensity between two users and then designs the pilot allocation algorithm through this value. In the case of users at low speed or even static, the service matrix and coincidence matrix can be considered to be constant for a period of time; so, with the help of AP selection theory, the detection matrix uses part of MMSE (P-MMSE). For the SINR in Equation (6), there are [19]

$$\mathrm{SINR}_k^{(\mathrm{ul},2)} = p_k^u \widehat{g}_k^{\mathrm{H}} \mathbf{D}_k \left( \sum_{j\in O_k} p_j^u \mathbf{D}_k \widehat{g}_j \widehat{g}_j^{\mathrm{H}} \mathbf{D}_k + \mathbf{Z}_k' \right)^{\dagger} \mathbf{D}_k \widehat{g}_k, \tag{12}$$

where

$$\mathbf{Z}_k' = \mathbf{D}_k \left( \sum_{j\in O_k} p_j^u \mathbf{C}_j + \sigma^2 \mathbf{I}_M \right) \mathbf{D}_k, \tag{13}$$

where $O_k$ represents users who have part of the same AP service as user $k$. The proof of Equation (12) can refer to the related content in reference [23].

*3.2. Pilot Allocation.* Inspired by the user-centered idea and the scalable problem of CF massive MIMO [19, 22], a pilot allocation scheme based on AP selection is proposed in this paper. To solve the problem of a large amount of calculation and pilot contamination in the case of total AP service for every user, the AP that provides services for each user is selected by the user's large-scale fading matrix. The degree of AP coincidence between users is proposed to quantify the possible pilot contamination intensity when users after selecting AP. For the pilot allocation of users, the AP that serves the most users is first selected, and the pilot allocation of users is carried out under the AP. For users with high repetition, orthogonal pilots are assigned first. To reduce channel estimation error, users under the same AP assign orthogonal pilots as much as possible [24]. Then, select other users under the AP to assign pilots until all users are


FIGURE 2: User selects AP sample.

assigned pilots. If all orthogonal pilots are assigned, other users will reuse the same pilots. The detailed steps of the proposed scheme are described in Algorithm 1 and described in the following five parts.

*3.2.1. Initialization of Pilot Allocation.* The length of the pilot is defined as $\tau_p$, and $\varphi^u = \{\varphi_1, \varphi_2 \cdots \varphi_{\tau_p}\}$ is the set of orthogonal pilots. $U$ is defined as the set of unassigned users, the set of AP service users is defined as $U_1$, and the set of unassigned users under this AP is $U_2 = U \cap U_1$.

*3.2.2. Confirmation of AP Priority.* For users under the same AP service, the subsequent channel estimation and signal detection will lead to greater errors and degrade the system performance if the same pilot is assigned. Therefore, when assigning pilots, the users of an AP service are taken as a group to assign pilots. For the AP to serve more users, to reduce pollution, the users under the AP are given priority to assign orthogonal pilots. Hence, the order of the selected AP is according to the number of AP service users. Select the AP through the matrix d in the Equation (11). If the AP in the selection is $m_1$, then

$$m_1 = \arg\max (d). \tag{14}$$

*3.2.3. Confirmation of User Priority.* For users' choice, the potential pilot contamination intensity is defined according to the AP coincidence degree between users. For users with a high coincidence degree, their potential pilot contamination is strong; so, orthogonal pilots are given priority. Through the coincident matrix $B$, select the user with a high coincidence degree of AP and mark it as $u$, and then

$$u = \arg\max_{o \in U_2} \left( \sum_{k=1}^{K} B_o \right). \tag{15}$$

*3.2.4. Pilot Selection.* The cumulative pilot contamination between the user and the user who has been assigned pilot is calculated by Equation (3), to obtain the pilot with minimum cumulative pilot contamination, which is recorded as

**Input:** $K, M, D, R, \tau_p, F, d$
**Output:** $P$
initialization: $m = 1$, pilot allocation matrix $P$ and initial user set $U$.
**while** $m \leq M$ and $U \neq 0$ do
  select the AP $m_1$ by Equation (14), then set $d(m_1) = -1$ and get a set of $S$ users $U_1$.
  **if** $m = 1$ **then**
    $U_2 = \{u_1, u_2, \cdots, u_s\}$ by sorting $U_1$ in descending order by Equation (15), and $U = U \setminus U_2$.
    **if** $S \leq \tau_p$ **then**
      assign orthogonal pilots to the user set $U_2$ in turn according to the pilot set $\varphi_1^u = \{\varphi_1, \varphi_2 \cdots \varphi_{\tau_p}\}$.
      add 1 for reuse times of $\varphi_2^u = \{\varphi_1, \varphi_2 \cdots \varphi_S\}$.
    **else**
      first assign all orthogonal pilots, and unassigned user $U_3 = \{u_{\tau_p+1}, \cdots, u_s\}$.
      **while** $U_3 \neq 0$ **do**
        select multiplexed pilots $\varphi_i$ through Equation (16), user $u_i$ in $U_3$, then assign.
        add 1 for reuse times of $\varphi_i$, and $U_3 = U_3 \setminus u_i$.
      **end while**
    **end if**
  **else**
    unassigned users $U_2 = U_1 \cap U$ (assuming $I$ users), and $U = U \setminus U_2$.
    **if** $U_2 \neq 0$ **then**
      $U_3 = \{u_1, u_2, \cdots, u_I\}$ in descending order by Equation (15) for $U_2$.
      **while** $U_3 \neq 0$ **do**
        select multiplexed pilots $\varphi_i$ through Equation (16), user $u_i$ in $U_3$ and the number of pilot multiplexing is less than $F$.
        add 1 for reuse times of $\varphi_i$, and $U_3 = U_3 \setminus u_i$.
      **end while**
    **end if**
  **end if**
  $m = m + 1$.
**end while**

Algorithm 1: Pilot assignment based on AP selection.

$\varphi$, and then

$$\varphi = \arg \min_{t_k} tr\left(\Psi_{t_k,m}\right). \tag{16}$$

*3.2.5. Other.* To avoid the extreme situation that the multiplexing times of the same pilot is too much and the multiplexing times of other pilots are too few, which leads to more severe pilot contamination, the maximum multiplexing times of each pilot is set as $F$. When the number of users in an AP is less than or equal to the pilot length, all orthogonal pilots can be assigned to minimize pilot contamination among the same group of users.

## 4. Data Power Optimization

For users, the smaller data transmission power will affect the communication quality of users, and the larger data transmission power will cause power waste and increase the interference between users. Therefore, data power optimization can effectively reduce power waste and interference while ensuring a certain communication quality. Based on pilot allocation, the optimization of data power can further improve the performance of the system. Here, the max-min fairness problem to increase user fairness is proposed, which is described as follows:

$$\begin{aligned} &\max_{p_k^u} \min_{k=1,2,\cdots K} \mathrm{SINR}_k, \\ &\text{subject to}\; p_k^u \leq p_{\max}^u, \forall k \in [1, 2, \cdots K], \\ &p_k^u \geq 0, \forall k \in [1, 2, \cdots K], \end{aligned} \tag{17}$$

where $p_{\max}^u$ is the maximum power value of the transmitted data, and the equivalent form of the formula (17) is

$$\begin{aligned} &\max_{p_k^u, \kappa} \kappa \\ &\text{subject to}\; \mathrm{SINR}_k \geq \kappa, \forall k \in [1, 2, \cdots K], \\ &p_k^u \leq p_{\max}^u, \forall k \in [1, 2, \cdots K], \\ &p_k^u \geq 0, \forall k \in [1, 2, \cdots K]. \end{aligned} \tag{18}$$

Regarding $\kappa$ as a variable, we can see that the optimization problem of (18) is a linear programming problem [25], which can be solved by CVX [26]. Here, the dichotomy is used to solve the convex optimization problem, and the algorithm for solving the optimization problem (18) is shown in Algorithm 2:

**Input:** $\kappa_{\min}$, $\kappa_{\max}$, $\xi$.
**Output:** $p_k^u$.
initialize: $\kappa_{\min}$, $\kappa_{\max}$, and threshold $\xi$.
while $\|\kappa_{\max} - \kappa_{\min}\| > \xi$ do.
  let $\kappa = (\kappa_{\min} + \kappa_{\max})/2$, and solve the optimization problem (18) with CVX.
  **if** problem solved **then**
  $\kappa_{\min} = \kappa$, and $\kappa_{\max}$ unchanged.
  else
    $\kappa_{\min}$ unchanged, and $\kappa_{\max} = \kappa$.
   **end if**
**end while**.

Algorithm 2: Data power optimization algorithm.

## 5. Simulation Result

In this section, Monte Carlo is used to simulate the above algorithm in MATLAB, and the performance of pilot allocation algorithm and data power optimization algorithm based on AP selection is obtained.

*5.1. Parameter Setting.* Here, we consider a scenario with $M$ APs and $K$ UEs, where both AP and UEs are single antennas. UE and AP are randomly distributed in the area of $1 \times 1\text{km}^2$. In order to avoid the boundary effect, the encircling method is used to deal with it [7]. When the system runs in TDD mode, the maximum transmission power of $p_{\max} = 100\text{mW}$ for data and pilot and the pilot length is $\tau_p = 10$; the coherence time is $T = 200$. The detailed parameters are as follows (see Table 1).

## 6. Results and Discussion

According to the above parameters, MATLAB is used to simulate the proposed pilot allocation method. The proposed pilot allocation method is compared with the greedy allocation method in reference [7], the pilot allocation method based on dissimilarity clustering (DCPA) in reference [27], the random pilot allocation method, and the ideal state without pilot contamination. In the figure, the proposed method is represented by proposed, greedy represents the greedy method, DCPA represents the dissimilarity cluster based pilot assignment method, and random represents the random pilot allocation. NoPC represents the one without pilot contamination.

The relationship between the sum SE and the cumulative distribution function (CDF) of the above methods in CF massive MIMO systems is compared (see Figure 3). P-MMSE is used for data detection, and scalable represents a method based on AP selection. As shown in Figure 3, the proposed method is better than the previously mentioned methods in terms of total SE. For the greedy pilot allocation method, because the potential contamination between users is ignored in the initial allocation, and then the greedy method is used to improve the performance, even if the number of iterations is increased, the subsequent pilot allocation may fall into a loop, resulting in no great improvement in performance, but the computational complexity of this method is relatively simple. When initially assigning pilots, the method proposed in this paper considers the potential pilot contamination between users based on the coincidence degree of AP, then by assigning orthogonal pilots to users with high coincidence degree, the pilot contamination problem between users is effectively suppressed, and the performance of the system is improved. At the same time, the performance of the random pilot allocation scheme is the worst. Still, in the case of no pilot contamination, the obtained channel is the actual channel in the uplink channel estimation and detection, which has no interference between users; its performance is also the best. Still, this method is ideal and cannot be realized in practice.

The impact of the number of AP on system performance (see Figure 4): as can be seen from Figure 4, with the increase in the number of AP, the system's performance continues to improve. To increase the number of AP, the degree of channel hardening between users and AP continues to improve. Meanwhile, the channel interference between users continues to reduce; so, the performance continues to improve. As mentioned in reference [28], the number of antennas in an area has a certain influence on the hardening ratio of the channel. With the increase of the number of antennas, the hardening ratio of the channel increases, the interference between channels decreases, and the system's performance is improved. However, as can be seen from Figure 4, when the number of AP is small, the performance is significantly enhanced by increasing the number of AP. Later, with the increase in the number of APs, performance improvement tends to be smooth, and the amount of data received and processed continues to increase; so, the trade-off between the impact of AP number on performance and complexity is also worth studying. At the same time, what is considered here is the case that the AP is a single antenna. In future research, each AP has multiple antennas in B5G and even 6G, which is also a trend and a place worth studying [29].

The relationship between AP selection ratio and average SE (see Figure 5): as the AP selection ratio increases, the system performance continues to increase. However, the more APs select, the greater the number of calculations. At the same time, the pilot contamination between users is also growing. It can be seen that the performance gap between the curve of NoPC and the curve of the other two methods is getting larger and larger. For the case of the full selection

Table 1: Simulation parameters.

| Simulation parameters | Symbol | Numerical value |
|---|---|---|
| AP number | $M$ | 400 |
| Number of users | $K$ | 40 |
| Pilot length | $\tau_p$ | 10 |
| Maximum pilot multiplexing times | $F$ | 6 |
| Transmission power | $p_{\max}$ | 100 mW |
| Coherent time | $T$ | 200 |
| Noise figure | $\delta$ | 7 dB |
| Path loss | $\phi$ | 3.76 |
| Shadow fading | $\sigma$ | 10 dB |

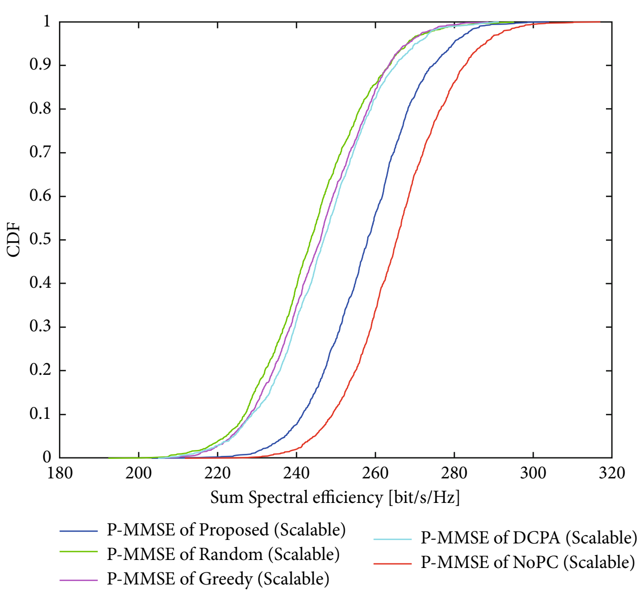


Figure 3: Sum spectral efficiency and CDF.

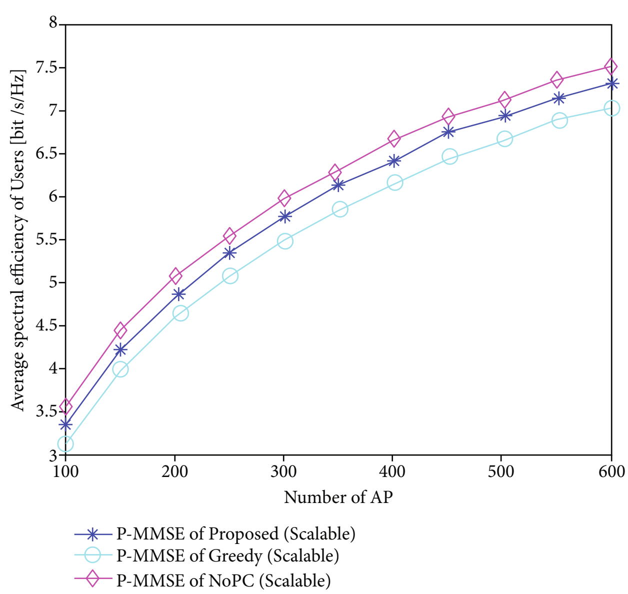


Figure 4: Average spectral efficiency and AP number.

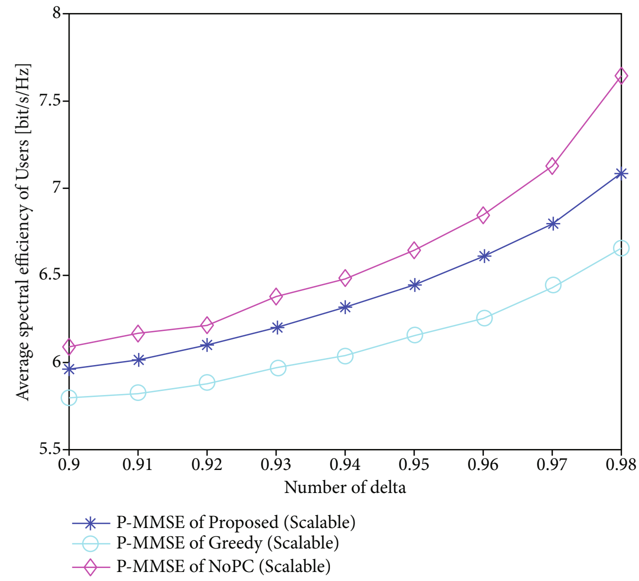


Figure 5: Average spectral efficiency and AP selection ratio.

of AP, the load of the backhaul link from AP to CPU increases, especially in the places where the number of users is large or dense, and the selection based on AP can effectively reduce the backhaul link's load. At the same time, the method proposed in this paper can also effectively suppress pilot contamination, which is a suitable compromise method.

The relationship between AP selection ratio and average SE (see Figure 5): as the AP selection ratio increases, the system performance continues to increase. However, the more APs select, the greater the number of calculations. At the same time, the pilot contamination between users is also growing. It can be seen that the performance gap between the curve of NoPC and the curve of the other two methods is getting larger and larger. For the case of the full selection of AP, the computational complexity of the system is large, especially in places where the number of users is large or dense, and the selection based on AP can effectively reduce the computational complexity of the system. At the same time, the method proposed in this paper can also effectively suppress pilot contamination, which is a suitable compromise method.

The relationship between the user's average number of AP selected and the total number of AP is under different region sizes (see Figure 6). As shown in Figure 6, the average number of AP chosen by users in various areas is much less than the total number of AP. According to the previous results, it is only necessary to select a small amount of APs for each user, rather than all the AP to serve a particular user. It reduces the computational complexity of the system and makes the actual implementation of CF massive MIMO possible.

The CDF diagram of the minimum SE of the system with data power optimization is shown (see Figure 7). It can be seen that the minimum user SE with data power optimization is significantly better than the minimum user SE without data power optimization. It shows that data power

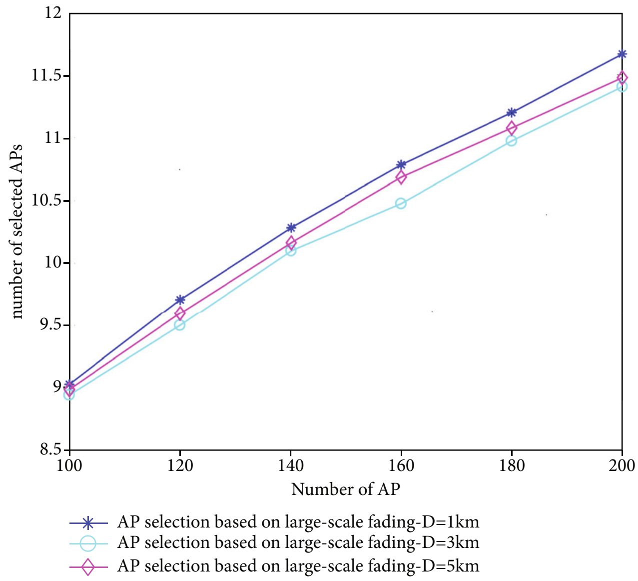


FIGURE 6: The average number of AP selected by the user and the total number of AP.

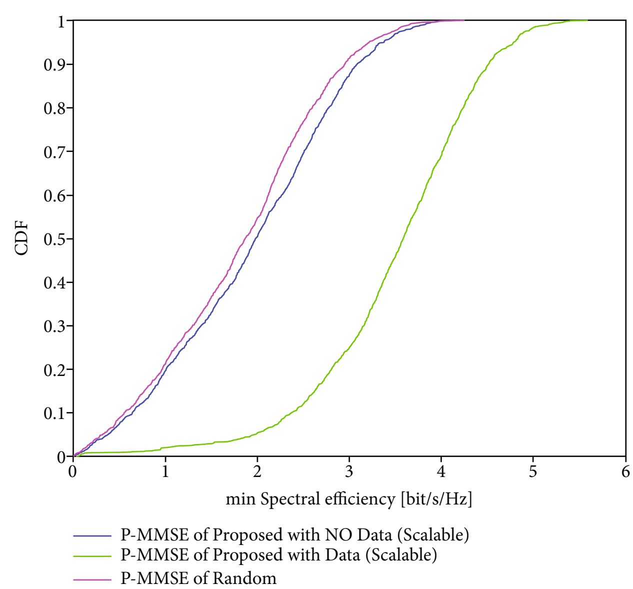


FIGURE 7: Minimum spectral efficiency and CDF.

optimization can improve system performance and fairness among users, mainly because users' interference intensity will increase when they are in full power. The interference to the users with poor performance will be more significant. Hence, the data power optimization fully considers this point, thus improving the minimum SE of users, while power optimization can also save part of the power consumption.

## 7. Conclusions

In this paper, aiming at the pilot contamination problem of CF massive MIMO uplink, a pilot allocation method based on AP selection is proposed. And the data power is optimized, which effectively reduces the computational complexity of the system also improves the fairness between users. Firstly, with the help of the user-centered idea, the AP service is selected for each user through the a priori large-scale fading matrix. The potential pilot contamination among users is quantified by the AP coincidence degree of each user. The users with large potential pilot contamination prioritize assigning orthogonal pilots then propose the data power optimization. Through the transformation of the formula, the optimization problem is turned into a convex optimization problem, and the dichotomy is used to solve it. Simulation results show that the proposed algorithm can effectively suppress pilot contamination while reducing the computational complexity of the system and improving the system's total SE and fairness between users. In future research, for the pilot allocation problem and data power optimization problem of CF massive MIMO multi-antennas, the issue of channel aging when users move at a certain speed and the actual hardware loss is worthy of in-depth study.

## Data Availability

The simulation code data used to support the findings of this study have not been made available. Because the code supporting this paper is a laboratory project, the source program cannot be made public for the time being.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61871466 and in part by the Key Science and Technology Project of Guangxi under Grant AB19110044.

## References

[1] S. Cheng, Z. Cai, J. Li, and H. Gao, "Extracting kernel dataset from big sensory data in wireless sensor networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 4, pp. 813–827, 2017.

[2] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 2, pp. 766–775, 2020.

[3] J. Q. Li, F. R. Yu, G. Deng, C. Luo, Z. Ming, and Q. Yan, "Industrial internet: a survey on the enabling technologies, applications, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1504–1526, 2017.

[4] B. Holfeld, D. Wieruch, T. Wirth et al., "Wireless communication for factory automation: an opportunity for LTE and 5G systems," *IEEE Communications Magazine*, vol. 54, no. 6, pp. 36–43, 2016.

[5] Z. Cai and Z. He, "Trading private range counting over big IoT data," in *2019 IEEE 39th International Conference on*

*Distributed Computing Systems (ICDCS)*, pp. 144–153, Dallas, TX, USA, 2019.

[6] X. Zheng and Z. Cai, "Privacy-preserved data sharing towards multiple parties in industrial IoTs," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 968–979, 2020.

[7] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[8] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: a new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99878–99888, 2019.

[9] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "Energy efficiency of the cell-free Massive MIMO uplink with optimal uniform quantization," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 4, pp. 971–987, 2019.

[10] T. C. Mai, H. Q. Ngo, M. Egan, and T. Q. Duong, "Pilot power control for cell-free massive MIMO," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 11264–11268, 2018.

[11] J. Koh, Y. Lim, C. Chae, and J. Kang, "On the feasibility of full-duplex large-scale MIMO cellular systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6231–6250, 2018.

[12] H. Liu, J. Zhang, X. Zhang, A. Kurniawan, T. Juhana, and B. Ai, "Tabu-search-based pilot assignment for cell-free massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2286–2290, 2020.

[13] S. Buzzi, C. D'Andrea, M. Fresia, Y. Zhang, and S. Feng, "Pilot assignment in cell-free massive MIMO based on the Hungarian algorithm," *IEEE Wireless Communications Letters*, vol. 10, no. 1, pp. 34–37, 2021.

[14] T. H. Nguyen, T. K. Nguyen, H. D. Han, and V. D. Nguyen, "Optimal power control and load balancing for uplink cell-free multi-user massive MIMO," *IEEE Access*, vol. 6, pp. 14462–14473, 2018.

[15] J. Zheng, J. Zhang, E. Björnson, and B. Ai, "Cell-free massive MIMO with channel aging and pilot contamination," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, pp. 1–6, Taipei, Taiwan, 2020.

[16] S. Buzzi, C. D'Andrea, A. Zappone, and C. D'Elia, "User-centric 5G cellular networks: resource allocation and comparison with the cell-free Massive MIMO approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1250–1264, 2020.

[17] E. Björnson and L. Sanguinetti, "A new look at cell-free massive MIMO: making it practical with dynamic cooperation," in *2019 IEEE 30th annual international symposium on personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–6, Istanbul, Turkey, 2019.

[18] G. Femenias, N. Lassoued, and F. Riera-Palou, "Access point switch ON/OFF strategies for green cell-free mssive MIMO networking," *IEEE Access*, vol. 8, pp. 21788–21803, 2020.

[19] E. Björnson and L. Sanguinetti, "Scalable cell-free Massive MIMO systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.

[20] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, 2020.

[21] H. Liu, J. Zhang, S. Jin, and B. Ai, "Graph coloring based pilot assignment for cell-free massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9180–9184, 2020.

[22] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: user-centric approach," *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 706–709, 2017.

[23] B. Emil, H. Jakob, and S. Luca, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[24] C. Pan, H. Mehrpouyan, Y. Liu, M. Elkashlan, and N. Arumugam, "Joint pilot allocation and robust transmission design for ultra-dense user-centric TDD C-RAN with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2038–2053, 2018.

[25] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[26] M. Grant and S. Boyd, *CVX: Matlab Software for Convex Programming, Version 2.1*, 2014, http://cvxr.com/cvx.

[27] G. Femenias and F. Riera-Palou, "Cell-free millimeter-wave massive MIMO systems with limited fronthaul capacity," *IEEE Access*, vol. 7, pp. 44596–44612, 2019.

[28] A. Á. Polegre, F. Riera-Palou, G. Femenias, and A. G. Armada, "Channel hardening in cell-free and user-centric massive MIMO networks with spatially correlated Ricean fading," *IEEE Access*, vol. 8, pp. 139827–139845, 2020.

[29] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1637–1660, 2020.