SPECIAL SECTION ON COMMUNICATIONS IN HARSH ENVIRONMENTS

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 27, 2020, accepted March 17, 2020, date of publication March 25, 2020, date of current version April 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2983215

Joint Pilot Allocation and Pilot Sequence Optimization in Massive MIMO Systems

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

Corresponding author: Feng Zhao (zhaofeng@guet.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61871466, in part by the Key Science and Technology Project of Guangxi under Grant AB19110044, and in part by the Guangxi Natural Science Foundation Innovation Research Team Project under Grant 2016GXNSFGA380002.

ABSTRACT Pilot contamination (PC) is recognized as a bottleneck for the achievable throughput of multi-cell massive multiple-input multiple-output (MIMO) systems. In this paper, we propose a joint pilot allocation and pilot sequences optimization (JPA-PSO) scheme to mitigate the effects of PC and maximize the system spectral efficiency (SE). Specifically, we first construct a range of pilot sequence in line with the coherent interval. Then, we establish an allocation basis by defining a similarity to distinguish PC among users. An angle of arrival (AOA) positioning method is designed to obtain the user's location information. Under the location information, we can monitor and estimate the distance between users in different cells to denote the potential interference of transmit power among users. Then, we execute pilot allocation strategy in line with the similarity and distance between users in different cells, that is, assign orthogonal pilots to the users with greater similarity and shorter distance. Finally, we provide a pilot sequence optimization algorithm to maximize the system SE. Simulation results verify that combining pilot allocation and pilot sequence optimization can significantly improve the system SE.

INDEX TERMS Massive multiple-input multiple-output, pilot contamination, pilot allocation, pilot sequence optimization, spectral efficiency.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) has achieved great success in throughput, and has been as one of the most critical technologies for the fifth generation (5G) and beyond 5G (B5G) wireless communication [1]–[3]. To explore the benefits of massive MIMO, many researchers bend over backwards to design revolutionary communication technologies. By deploying large-scale antennas at the base station (BS), a massive MIMO system has substantial potential in suppressing multiuser interference and providing disruptive improvements in spectral efficiency (SE) and energy efficiency (EE) [4]–[6]. In line with resource saving and highly scalable features, the time division duplex (TDD) model is used to achieve communication between BS and users. Since the channel state information (CSI) is uncertainty knowledge and needs to be obtained by the BSs that operate in line with TDD [7], [8]. In practice, the pilot sequences are utilized to estimate the CSIs. Due to the coherence time cannot be extended infinitely, the pilot length is controlled within a narrow range and the same pilot sequences is used for all cells, which leads to the problem of pilot contamination (PC). Due to the existence of PC, the system capacity of the network is limited even with innumerable BS antennas. In other words, PC is the main factor leading to system performance degradation. Therefore, PC is recognized as a bottleneck for the achievable system capacity of multi-cell massive MIMO systems [9]–[11].

In order to alleviate the impact of PC, many pilot allocation schemes have been designed to restrict the interference among users. A basic approach of pilot reuse was studied in [12], which focuses on designing pilot reuse patterns to prove the relationship between PC and pilot consumption, and each pilot reuse pattern can provide a certain performance gain by consuming different number of pilots. The pilot allocation scheme that ignores correlations between users is proposed in [13], which to reduce PC by maximizing the performance of users with poor channel conditions in

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

each cell. However, the system performance is undesired since this scheme sacrifices the users with superior conditions in exchange for the availability of users with poor conditions. The authors of [14] have designed a supervised learning scheme of pilot allocation, which completes pilot allocation through four steps: acquiring data, preprocessing data, system modeling and post processing. This solution achieves high system capacity with short time complexity. In [15], smart antennas are used for pilot allocation, which can mitigate PC and improve system performance by separating users who send the same pilot in physical space. Nevertheless, the interference intensity between users is not distinguished, which will lead to poor performance of edge users. The graph algorithm is described in [16], which adopts iterative search to obtain interference threshold, and uses the channel characteristics and interference threshold to construct an interference graph, then assigns pilots to user based on the interference graph. In [17], a weighted pilot allocation scheme is proposed and it designs a PC weighted graph for pilot allocation. Due to its allocation algorithm is greedy, the system performance cannot be well guaranteed. In [18], a clustering pilot allocation scheme is proposed by clustering users who have significant mutual interference. Moreover, the authors of [19] design a dictionary of pilot and pilot transmit power, and based on this dictionary, an asymptotic technique is adopted to design the algorithm of maximizing the minimum SE. In addition, a heuristic algorithm which takes into account the pilot length is proposed in [20] and it is mainly to reduce the channel estimation error. However, this algorithm uses traversal search for pilot length optimization, which will lead to high algorithm complexity with the coherence interval extend.

In this paper, we studied the advantages and challenges of massive MIMO uplink system, based on which a joint pilot allocation and pilot sequence optimization (JPA-PSO) scheme is proposed to restrain PC and improve the system SE. First, we formulate two pilot optimization problems by analyzing the relationship between SE and PC and the relationship between SE and pilot sequence: pilot allocation lies in maximizing the system SE by suppressing the PC; pilot sequence optimization focus on maximizing the system SE by optimizing the pilot sequence length. To solve these problems, we construct a range of pilot sequence based on the coherent interval. Then, we establish an allocation basis by defining a similarity to distinguish PC among users. An angle of arrival (AOA) positioning method is designed to obtain the user's location information. Under the location information, we can monitor and estimate the distance between users in different cells to denote the potential interference of transmit power among users. Then, we execute pilot allocation strategy according to the similarity and distance between users in different cells, that is, assign orthogonal pilots to the users with greater similarity and shorter distance. Finally, we use extrapolation and interpolation methods to solve the problem of pilot sequence optimization for maximizing the system SE. Simulation results verify that combining pilot allocation and pilot sequence optimization can significantly restrain PC and improve system SE. In addition, the effectiveness of our proposed JPA-PSO scheme is superior to traditional pilot allocation schemes, and almost catches up to the exhaustive search with low computational complexity.

II. SYSTEM MODEL

In this paper, a multi-cell massive MIMO uplink system with L hexagonal cells that operate based on a TDD protocol is considered, and its model is demonstrated in Fig. 1. Each cell is composed of a BS and K single-antenna users, BS is deployed with M antennas to suppress interference and communicates with K users simultaneously [4]. In this model, we adopt TDD to achieve communication between users and BS, and exploit the block fading model, thus the channels is smooth in one coherent time block, and change independently in the next coherent time block [11]. For convenience, we use the symbol $\langle j, k\rangle$ to represent the kth user in the jth cell, and BS i is represent the BS of the ith cell. In the channel setup, $\mathbf{h}_{j,k}^{i} = \mathbf{Z}_{j,k}^{i}\sqrt{\beta_{j,k}^{i}}$ denotes channel vector from the user $\langle j, k\rangle$ to the BS i, where $\mathbf{Z}_{j,k}^{i}$ represents the small-scale fading vector, and obeys independent and identically distribution, that is, $\mathbf{Z}_{j,k}^{i} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. And $\beta_{j,k}^{i} = s_{j,k}^{i}/\left(r_{j,k}^{i}/R\right)^{\alpha}$ is the large-scale fading coefficient, where $s_{j,k}^{i}$ represents the shadow fading from the user $\langle j, k\rangle$ to the BS i, $r_{j,k}^{i}$ denotes the distance between the user $\langle j, k\rangle$ and the BS i, α denotes the path loss exponent of links, and R is the radius of cell [9].

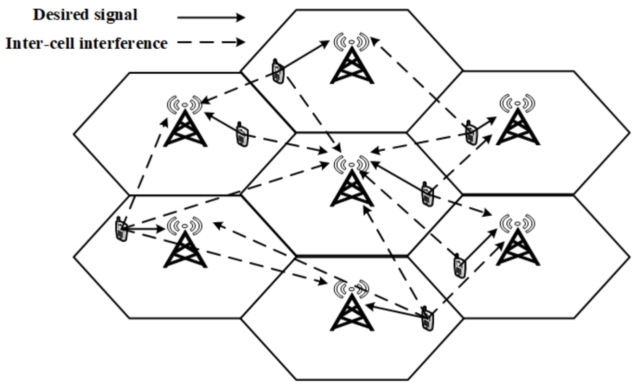


FIGURE 1. Massive MIMO uplink system model.

In practice, the CSI of user is uncertainty knowledge and needs to be obtained through the BS with the assistance of pilots. Hence, we assume that there are τ pilots for channel estimation, and a pilot sequence $\mathbf{\Phi}$ is constructed, which can be expressed as $\mathbf{\Phi} = [\phi_1, \phi_2, \cdots, \phi_\tau]^T \in \mathcal{C}^{\tau\times 1}$, where ϕ is pilot symbol. The length of $\mathbf{\Phi}$ is τ and $\mathbf{\Phi}$ is reused in each cell.

To detect the transmission signal for the user $\langle i, k\rangle$, supposing that BS i is equipped with a matching filter receiver [11]. After data detection, the user's signal-to-interference plus-noise ratio (SINR) can be obtained based on the detection results. And the uplink SINR of user $\langle i, k\rangle$

is [13], [19]

$$\begin{aligned}\mathrm{SINR}_{i,k}^{UL} &= \frac{\rho_{i,k}\left|(\tilde{\mathbf{h}}_{i,k}^{i})^{H}\mathbf{h}_{i,k}^{i}\right|^{2}}{\sum_{j=1,j\neq i}^{L}\rho_{j,k}\left|(\tilde{\mathbf{h}}_{j,k}^{i})^{H}\mathbf{h}_{j,k}^{i}\right|^{2}+|\boldsymbol{u}_{i,k}|^{2}} \\ &\approx \frac{\rho_{i,k}\hat{\rho}_{i,k}(\beta_{i,k}^{i})^{2}}{\sum_{j=1,j\neq i}^{L}\rho_{j,k}\hat{\rho}_{j,k}(\beta_{j,k}^{i})^{2}}, M\rightarrow\infty\,, \qquad (1)\end{aligned}$$

where $\rho_{i,k}$ and $\hat{\rho}_{i,k}$ represent the pilot and data transmission power of user $\langle i,k\rangle$, respectively. And $\tilde{\mathbf{h}}_{i,k}^{i}$ represents the channel estimate of $\mathbf{h}_{i,k}^{i}$. This formula indicates that if the BS antenna scale is enormous, the channels will be orthogonal to each other and the approximate equality above holds. Since the channels are orthogonal, the uncorrelated interference $\boldsymbol{u}_{i,k}$ disappears gradually, and the survivor in the denominator is PC, which is not interfered by the BS antennas. Then the expression of total system SE applied to pilot allocation and pilot sequence length optimization is

$$U_{\mathrm{SE}}^{(\tau)} = \left(1-\frac{\tau}{T}\right)\sum_{i=1}^{L}\sum_{k=1}^{K}E\left\{\log_2\left(1+\mathrm{SINR}_{i,k}^{UL}\right)\right\}\,, \qquad (2)$$

where T denotes the coherent interval length (total transmission symbols of the pilot and the data within a coherent interval).

III. PROPOSED JOINT PILOT ALLOCATION AND PILOT LENGTH OPTIMIZATION

A. PROBLEM FORMULATION

During the uplink transmission phase, pilot and data signals are transmitted to the destination through the coherent interval and each coherent interval is split into two areas with different load functions, pilots occupying one functional area and data occupying another functional area. Since the length of the coherence interval cannot be infinitely extended, when numerous users access the system, the length of pilot sequence is restricted. As a result, a pilot sequence set is reuse in each cell and it is a major factor of PC. To protect users from PC and suppress the system interference as much as possible, we constructed a range of pilot sequence length in line with the coherence interval, that is $K \leq \tau \leq T$.

From equation (1), we can observe that the PC does not disappear no matter how magnificent the BS antenna scale. Hence, to avoid the users with strong mutual interference sharing the same pilot symbol, our goal is to reduce the PC and maximize the total system SE, which can be formulated as

$$\max_{\{SINR\}}\left(1-\frac{\tau}{T}\right)\sum_{i=1}^{L}\sum_{k=1}^{K}E\left\{\log_2\left(1+\mathrm{SINR}_{i,k}^{UL}\right)\right\} \qquad (3)$$

where the constraint condition is SINR, and it is given in (1). Meanwhile, for a given coherence interval T, τ is the range occupied by pilots, then $T-\tau$ is the range occupied by data symbols. Decreasing τ can increase the $T-\tau$ and improve the system SE, but will enlarge the PC. Therefore, to balance the tradeoff between pilot sequence length and data transmission symbols, we propose a pilot sequence optimization problem and its optimization variable is the pilot sequence length. The purpose of pilot sequence optimization is suppress interference and maximize the system SE, which can be formulated as

$$\begin{aligned}&\max_{\{\tau\}}\left(1-\frac{\tau}{T}\right)\sum_{i=1}^{L}\sum_{k=1}^{K}E\left\{\log_2\left(1+\mathrm{SINR}_{i,k}^{UL}\right)\right\},\\ &\text{s.t.}\quad \tau\geq K,\\ &\qquad\ \ \tau\leq T \qquad (4)\end{aligned}$$

and the constraint condition is pilot sequence length.

In order to solve these problems, a JPA-PSO scheme is proposed in the following subsection.

B. ANALYSIS OF SIMILARITY AND DISTANCE

Equation (1) shows that if the transmit power is given, reduce the PC, that is, reduce the $\sum_{j=1,j\neq i}^{L}(\beta_{j,k}^{i})^{2}$ can increase the SINR of user, is equivalent to improve the system SE. Therefore, PC and system performance are closely connected with the large-scale fading coefficient. To identify PCs between users, we establish a pilot allocation basis by defining a similarity, which represents the first priority. We use the symbol $f_{i,k}^{j,k}$ to represent the similarity between two users, and i and j represent two different cells. The similarity value, $f_{i,k}^{j,k}$ is

$$f_{i,k}^{j,k} = (\beta_{i,k}^{j})^{2}/(\beta_{j,k}^{j})^{2}+(\beta_{j,k}^{i})^{2}/(\beta_{i,k}^{i})^{2} \qquad (5)$$

Clearly, stronger $f_{i,j}^{j,k}$ indicates that the similarity between users is strong, and they need to use orthogonal pilots, otherwise it will lead to more serious PC. Note that, the $\beta_{j,k}^{i}$ is closely bound up with the distance $r_{j,k}^{i}$ between the user $\langle j,k\rangle$ and BS i. Due to the randomness of user location, there may be several users with the same distance to the BS, and this means that their large-scale fading coefficient values are equal, which will result in multiple users with the same similarity. For example, consider three users A, B and C, A and B are in jth cell with the same distance from the BS j. C is the user of ith cell who is closer to A. When the distances of A and B to the BS i are also the same, the similarities f_{A}^{C} between A and C and f_{B}^{C} between B and C may be the same. Moreover, since the distance between A and C is shorter than the distance between B and C, the interference between A and C is greater than the interference between B and C when same pilot is allocated to them. Therefore, simply using similarity to distinguish PC among users is not accurate enough. To make up for this shortcoming, we introduce the distance between users to establish another pilot allocation basis.

Considering the users who are close to each other in the edge of different cells and bear the interference of pilot sharing, when they transmit signals to the BS synchronously, they suffer great interference of transmit power from each other, which affect the channel estimation and data detection of BS. Therefore, the distance between users in different cells is far away, the system performance can be better. Moreover, considering that, there may be multiple users with the same similarity and their interference intensity can be distinguished by distance, so the distance between users in different cells is taken as the second basis of pilot allocation.

Assuming that the BS is deployed uniform linear antenna array, an AOA locating method is designed to obtain the user's location information. The BS senses the arrival direction of signal through the antenna array and according to the signal arrival direction calculates the relative azimuth or angle between the BS and users, then the user's location can be calculated by triangulation with the user's azimuth angle obtained above. Due to the limitation of the angular resolution of the antenna equipment, the positioning accuracy of the AOA is limited. The greater the distance between the BS and the mobile equipment, the lower the measurement accuracy of the AOA. Therefore, to improve the measurement accuracy of AOA, we design each BS only obtain the location information of users in its cell, and adopted the measurement method of AOA positioning and signal wave path difference to obtain the azimuth angle and elevation angle between users and BS. The user's location information is shown in Fig. 2. Each BS is taken as the origin of the space coordinate of the corresponding cell, and the central BS is taken as the origin of the space coordinate of the whole system, denoted as $(0, 0, 0)$. The radius of the cell is R. Let us take two antennas near the center of the BS as an example. If the horizontal spacing d_h and vertical spacing d_v between the antenna array elements are known, and α_1 and α_2 represent the horizontal angle at which the signal reaches the adjacent antennas, then the phase difference between two adjacent antenna received signals is computed as $\Delta\alpha = |\alpha_1 - \alpha_2|$. When the signal wavelength is λ, the wave path difference of adjacent signal can be calculated by the phase difference $\Delta\alpha$, i.e., $d_h * \cos\theta = \lambda * \frac{\Delta\alpha}{2\pi}$. Thus, the azimuth angle between the user and BS is computed as $\theta = \arg\cos\left(\frac{\lambda\Delta\alpha}{2\pi d_h}\right)$. Similarly, β_1 and β_2 represent the elevation angle at which the signal reaches the adjacent antennas, then the phase difference between two adjacent antenna received signals is computed as $\Delta\beta = |\beta_1 - \beta_2|$. When the signal wavelength is λ, the wave path difference of adjacent signal can be calculated by the phase difference $\Delta\beta$, i.e., $d_v * \sin\omega = \lambda * \frac{\Delta\beta}{2\pi}$. Thus, the elevation angle between the user and the BS is computed as $\omega = \arg\sin\left(\frac{\lambda\Delta\beta}{2\pi d_v}\right)$. Then, the space location information of user can be described by the tuple (x, y, z) completely, and the elements

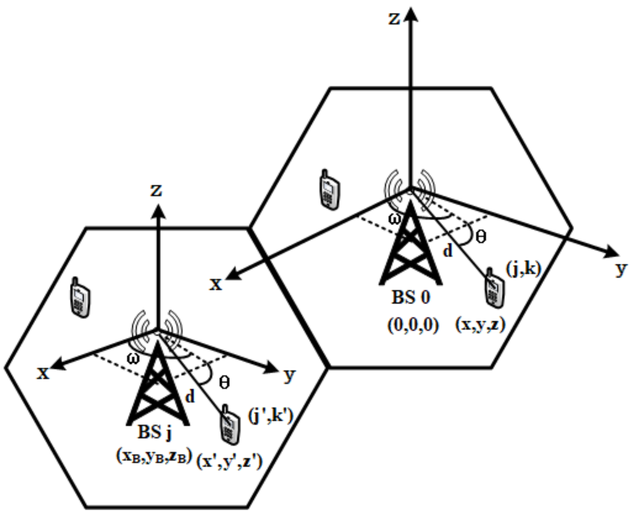


FIGURE 2. User location information.

$$\begin{cases} x &= d\cos\theta\cos\omega + x_B \\ y &= d\cos\theta\sin\omega + y_B \\ z &= d\sin\theta + z_B \end{cases} \tag{6}$$

where $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $\omega \in [0, 2\pi)$, and d represents the distance from user to BS, and it is the known prior information obtained by BS estimation. The distance between two users can be calculated as

$$r = \sqrt{|x - x'|^2 + |y - y'|^2 + |z - z'|^2} \tag{7}$$

In the same way, (x', y', z') denotes the space location information of another user. Small r indicates that the potential interference of transmit power between users is strong. In order to suppress the interference of the transmission power as much as possible, users who are closer should use orthogonal pilots, otherwise the system SE will be degrade.

C. EXTRAPOLATION AND INTERPOLATION

Equation (1) and (2) show that both short and long pilot sequence lengths can affect the system performance, shorter pilot sequence length can improve SE, but will introduce serious PC and degrade the SINR. Therefore, within the range $K \le \tau \le T$, we can find an optimal pilot sequence length τ, which makes the system performance the best. The pilot sequence length optimization problem is described in subsection A and we can observe that it is a nonlinear programming with a single peak. We can do a one-dimensional search using the traversal method to get the peak. However, the computational complexity of the traversal search method will be high when the coherence interval is large. Therefore, we adopt extrapolation and interpolation method to solve this problem and its computational complexity is much lower than the traversal search method. The core idea of the extrapolation and interpolation method is the algorithm starts from a certain initial point, and then advances in the direction of increasing function value with appropriate stride, until the function value decreases. At this time, we can obtain three points, and the function value of these three points meets low on both sides and high in the middle and the interval where the maximum value is located can be determined according to the three points. Then, interpolate a point between any two points in the interval obtained above and compare the magnitude of the function values about these points. Then the left and right neighbors of the maximum function value will be used as a new interval, which can shorten the interval where the maximum value is located again, and finally determine the position of the maximum value.

D. PROPOSED JPA-PSO SCHEME

In this paper, inspired by the weight algorithm [17] and pilot length optimization [20], a JPA-PSO scheme was proposed. By analyzing the incomplete reliability of simply using similarity to distinguish PC, we consider the interference of transmit power and propose the distance between users in different cells to make up for this shortcoming, and analyzing the relationship between pilot sequences and SE, a conclusion that the pilot sequence optimization is a nonlinear programming is summarized. Then we design the JPA-PSO scheme. In particular, we construct a range of pilot sequence length for pilot allocation and pilot sequence optimization. Then, the similarity and distance are defined as the first priority and the second priority for pilot allocation, respectively. If the user satisfies the first priority, the user will be assigned the orthogonal pilot first. If the first priority cannot be selected, the second priority will be used as the decision condition, and the users who meet the first and second priorities are assigned the orthogonal pilot. If all the orthogonal pilots have been allocated, the remaining users will reuse the same pilots, repeat the steps above, until all users are assigned pilots. Finally, we perform a nonlinear search over the range $K \leq \tau \leq T$ by extrapolation and interpolation to obtain the maximum SE. The detailed steps of the proposed JPA-PSO scheme is designed in Algorithm 1 and described in the following seven parts.

1) INTRODUCTION

In this letter, the range $K \leq \tau \leq T$ is constructed for pilot allocation and pilot sequence optimization, in which we perform a nonlinear search to obtain the maximum SE. During the search, the corresponding pilot sequence length will be used for pilot allocation and the pilot allocation is performed in each search.

2) PILOT ALLOCATION INITIALIZATION

First, we define D as an empty set for storing users who are assigned pilot, then calculate the similarity between users in different cells and compare similarity values across all users, finally assign orthogonal pilots to users with the greatest similarity. At the same time, users with the greatest similarity are saved in D, while the remaining users are candidates.

3) USER SELECTION

For user selection, we introduce two priority matrices $\mathcal{F}$ and θ with L rows and K columns, $\mathcal{F}$ for storing the accumulation of the similarity between candidate and the users in D, which is defined as the first priority of user selection. And θ for storing the accumulation of the distance between candidate and the users in D, which is defined as the second priority of user selection. Then, the candidate who satisfies the maximum $\mathcal{F}(\mathrm{j},\mathrm{k})$ will be selected as preferred candidate of pilot allocation. If the number of candidates who satisfy the maximum $\mathcal{F}(\mathrm{j},\mathrm{k})$ have multiple, we will select the candidate who satisfies both the maximum $\mathcal{F}(\mathrm{j},\mathrm{k})$ and minimum $\theta(\mathrm{j},\mathrm{k})$ as the preferred candidate of pilot allocation.

Algorithm 1 Proposed JPA-PSO Algorithm

```
Require: $K, L$
Ensure: $\{p_{j,k}\}\ \forall j, k, \quad \tau^*_{\mathrm{SE}}$
$\tau = h = K, K \leq \tau \leq h < T, \Delta > 0$
while $U^{(h)}_{SE} \geq U^{(\tau)}_{SE}$ do.
    $\{p_{j,k}\} = 0$, define $n$ initial users.
    $\{\langle j_n, k_n \rangle\} = \arg\max_{\forall i,j} f^{j,k}_{i,k}$.
    $p_{j_n,k_n} = n, D = \{\langle j_n, k_n \rangle\}, C = \{KL \backslash D\}$.
    while $C \neq \emptyset$ do
        $\mathcal{F}(\mathrm{j},\mathrm{k}) = \sum_{i,k \in D, j,k \in C} f^{j,k}_{i,k}$.
        $\theta(\mathrm{j},\mathrm{k}) = \sum_{i,k \in D, j,k \in C} r^{j,k}_{i,k}$.
        $\langle j_x, k_x \rangle = \arg\max \{\mathcal{F}(\mathrm{j},\mathrm{k})\}$.
        if $\langle j_x, k_x \rangle \geq 2$ then
            $\langle j_x, k_x \rangle = \arg\min \{\theta(\mathrm{j},\mathrm{k})\}$.
        end if
        $\psi = \{\mathcal{P} : \tau \backslash p_{j_x,k}, \forall \mathrm{k}\}$.
        $\mathcal{P} = \arg\min \sum_{i,k \in D, p_{j_x,k_x} \in \psi} f^{j_x,k_x}_{i,k}$.
        $p_{j_x,k_x} = \mathcal{P}, D = D \cup \{\langle j_x, k_x \rangle\}, C = C \backslash D$.
    end while
    $\tau = h, h = h + 2\Delta$.
    Calculate $U^{(\tau)}_{SE}$ and $U^{(h)}_{SE}$ by (2).
end while
Determine the interval $[\tau - 2\Delta, \tau + 2\Delta]$ where the maximum SE (EE) is located.
Calculate $U^{(\tau)}_{SE}$ within the interval $[\tau - 2\Delta, \tau + 2\Delta]$ by (2).
$\tau^*_{\mathrm{SE}} = \max_{\tau \in [\tau - 2\Delta, \tau + 2\Delta]} U^{(\tau)}_{SE}$.
```

4) PILOT ALLOCATION

If the preferred candidate of pilot allocation is chosen, we first construct a pilot set ψ, which represents the orthogonal pilots available in the cell where the preferred candidate is located. Then, the available pilots in the pilot set ψ are assigned to the preferred candidate, and the accumulations of similarity between the users in D and the preferred candidate are calculated in order, then the preferred candidate makes a comparison and decides to choose the pilot with the least accumulation of similarity. Finally, save the preferred candidate into D. Repeat steps 2) and 3), until all candidates are assigned pilot.

5) PILOT LENGTH INITIALIZATION

We set initial interval of the pilot sequence length as $[K, T]$, initial point as $\tau = h = K$ and give the initial stride $\Delta > 0$.

6) EXTRAPOLATION

To determine the interval where the maximum SE is, we use extrapolation to perform a nonlinear search over the interval $[K, T]$. First, the pilot sequence length takes a step forward from the initial point $\tau = K$ with a stride 2Δ and get point

$\tau+2\Delta$, then calculate $U_{SE}^{(\tau)}$ and $U_{SE}^{(\tau+2\Delta)}$ by (2) and compare these two function values. If $U_{SE}^{(\tau+2\Delta)} \geq U_{SE}^{(\tau)}$, continue to increase the pilot length with a stride 2Δ, repeat the steps described above until $U_{SE}^{(\tau+2\Delta)} \leq U_{SE}^{(\tau)}$. Then, we can get three points $\tau-2\Delta$, τ and $\tau+2\Delta$, and the function values of these three points satisfy low on both sides and high on middle, i.e., $U_{SE}^{(\tau-2\Delta)} \leq U_{SE}^{(\tau)} \geq U_{EE}^{(\tau+2\Delta)}$. Finally, we can determine that the point of the maximum SE is within the interval $[\tau-2\Delta, \tau+2\Delta]$.

7) INTERPOLATION

After determining the interval $[\tau-2\Delta, \tau+2\Delta]$ in which the maximum SE is located, we calculate the SE and compare the size of SE with τ within the interval $[\tau-2\Delta, \tau+2\Delta]$, then select the pilot sequence length that maximizes SE as the optimal pilot length. If the interval $[\tau-2\Delta, \tau+2\Delta]$ is large, we can interpolate a point $\tau+\Delta$ between τ and $\tau+2\Delta$. Then, compare the magnitude of $U_{SE}^{(\tau)}$ at points $\tau-2\Delta$, τ, $\tau+\Delta$ and $\tau+2\Delta$, and the left and right neighbors of the maximum function $U_{SE}^{(\tau)}$ are used as a new interval, which can short the interval where the maximum SE is located again. Repeat the steps described above and finally determine the location of the maximum SE.

The overall algorithm complexity of pilot allocation and pilot sequence optimization is $\mathcal{O}\left(\frac{\tau-K}{2\Delta}\tau(KL)^3\right)$, which comprises the computational complexity $\frac{\tau-K}{2\Delta}$ for optimizing the pilot sequence length by extrapolation and interpolation, and the computational complexity $\tau(KL)^3$ required to perform pilot allocation. In addition, it is more reliable and feasible compared to the computational complexity of exhaustive solution for combining pilot allocation and pilot sequence optimization, which is $\mathcal{O}(T(C_{\tau}^{K})^{L-1})$.

IV. SIMULATION RESULTS

In this section, we utilize Monte Carlo to perform system simulation, and several numerical results are provided to evaluate the system performance of the proposed JPA-PSO scheme. A massive MIMO system comprise L cells, L BSs, and KL users is designed to simulate actual uplink communication. In the model setup, the radius of the cell is $500m$, users are randomly and evenly scattered in the safe and available areas, and BS deployed M antennas is constructed in the most advantageous position of the cell and communicates with K users simultaneously [4]. We choose the coherent interval length $T = 196$, and the detailed simulation parameters are summarized in table 1. To prove and discuss the reliability and availability of our design scheme, the system simulation of traditional pilot allocation is provided for comparison. For clarification, the previously designed schemes of pilot allocation are collectively referred to as traditional pilot allocation.

TABLE 1. Simulation parameters.

Simulation parameters	symbol	Values
Number of cells	L	$L = 19$
Number of BS antennas	M	$10 \leq M \leq 1000$
Number of users in each cell	K	10
Number of orthogonal pilots	τ	$k \leq \tau \leq T$
Cell radius	R	$500m$
Minimum cell radius	r	$50m$
Path loss exponent	α	3
Bandwidth	B	$20MHZ$
Log normal shadowing fading	σ_{shadow}	$8dB$
Data transmit power	ρ	$15dB$
Pilot transmit power	$\hat{\rho}$	$10dB$
Circuit power per user equipment	ρ_{cir}	$1dB$
Coherence interval	T	196

Fig. 3 compares the SINR of the traditional pilot allocation, exhaustive search, exhaustive search with pilot sequence optimization (Exhaustive search & PSO), the proposed pilot assignment and the proposed JPA-PSO scheme, where the traditional pilot allocations include random pilot assignment and smart pilot assignment. Random pilot assignment in which users reuse pilots without any regular pattern in each cell, due to its randomness, the PC of users in different cells is serious, resulting in a low SINR. Smart pilot assignment aim at maximizing the minimum SINR in each cell, since it sacrifices the largest SINR of the overall system in exchange for the availability of users with poor channel conditions, the SINR has not improved much. Exhaustive search obtains the largest SINR by finding users with the least PC and assigning them the same pilot. The proposed JPA-PLO scheme considers PC and power interference, uses the similarity and the distance between users to indicate mutual interference between users, and assigns orthogonal pilots to users with greater similarity and shorter distances. As can be seen from Fig. 3, the SINR of the proposed JPA-PSO scheme surpasses traditional pilot allocations, the proposed pilot assignment and exhaustive search, almost catches up to exhaustive search with pilot sequence optimization.

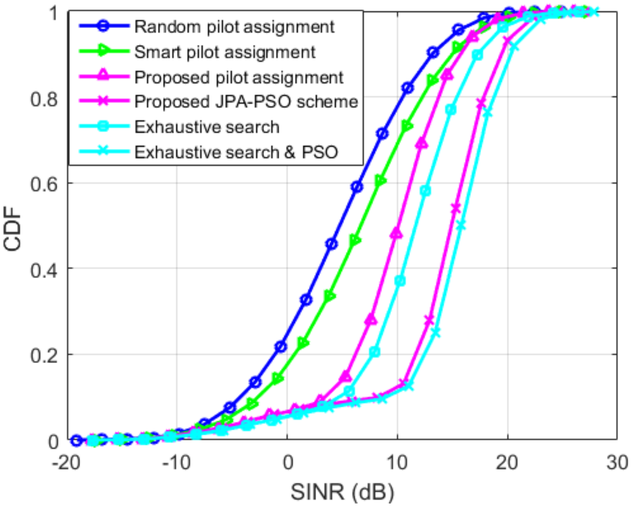


FIGURE 3. Cumulative distribution function (CDF) of SINR of different schemes with $L = 19$, $\tau = K = 10$, $\rho = 15dB$, $\hat{\rho} = 10dB$ and $M = 128$.

Fig. 4 indicates the increasing trend of average system rate under the BS antenna scale. The results in Fig. 4 show that the more BS antennas, the greater improvement the

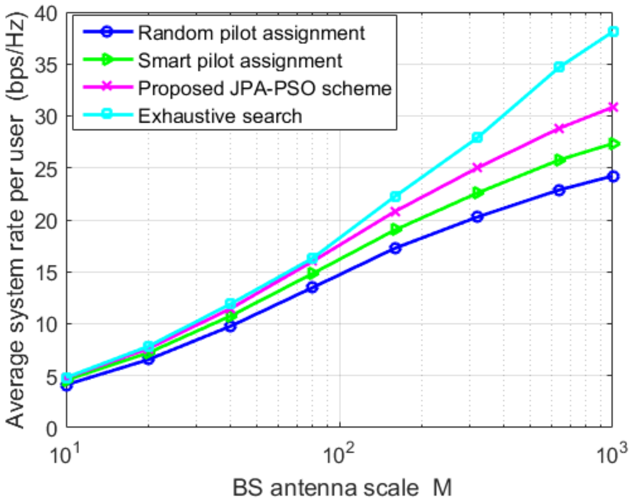


FIGURE 4. **Average system rate under the BS antenna scale M with $L = 19$, $\tau = K = 10$, $\rho = 15dB$ and $\hat{\rho} = 10dB$.**

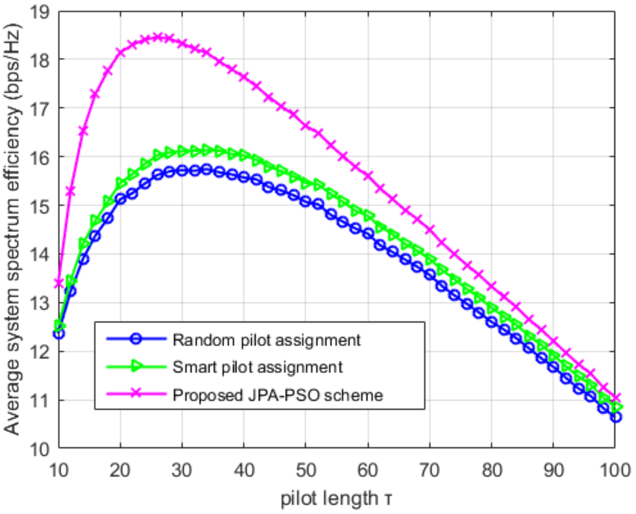


FIGURE 5. **Average system SE against the pilot sequence length τ with $L = 19$, $K = 10$, $\rho = 15dB$, $\hat{\rho} = 10dB$ and $M = 128$.**

average system rate of traditional pilot allocations, exhaustive search and the proposed JPA-PSO scheme. The reason for this phenomenon is that tremendous BS antenna scale can realize channel orthogonality, which can reduce mutual interference between users, thus improve the system rate. Compared with the traditional pilot allocations, the proposed JPA-PSO scheme can achieve higher rate, and this is because the proposed JPA-PSO scheme aims to Reduce PC and power interference, which can largely reduce the channel estimation error, thus effectively improve the system rate.

Fig. 5 indicates the relationship between the average system SE and the pilot sequence length τ. From the Fig. 5, we can find that all the curves first rise and then fall, it can be determined that SE is a single-peak function, which has a maximum function value. When $\tau = 34$, the system SE of tradition pilot allocations reach a maximum, when $\tau = 26$, the system SE of the proposed JPA-PSO scheme reaches a maximum. Within the pilot sequence length $K \leq \tau \leq T$, the SE of proposed JPA-PSO scheme is superior to the SE of tradition pilot allocations. When $K \leq \tau \leq 26 (K \leq \tau \leq 34)$, the system SE improves with the pilot sequence length extends. This phenomenon can be explained as increasing pilot sequence length can weaken the PC, thus improving the SE of the whole network. When $26 < \tau < T (34 < \tau < T)$, increasing the pilot sequence length, SE drops rapidly, and this is because increasing pilot sequence length τ can compress the number of data transmission symbols and reduces the system SE.

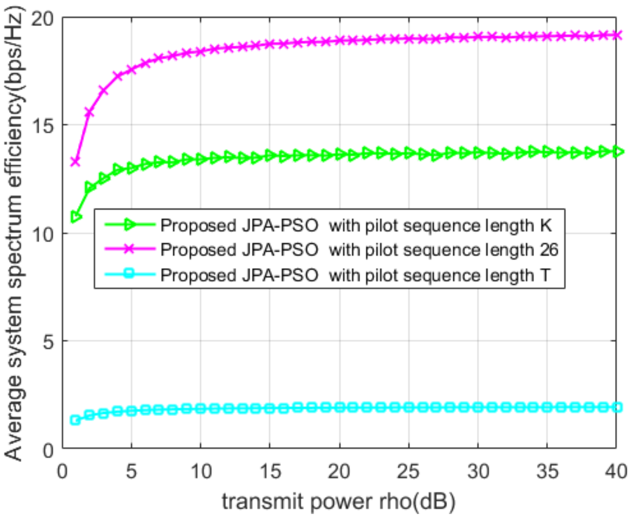


FIGURE 6. **Average system SE against the transmit power with $L = 19$, $K = 10$ and $M = 128$.**

Fig. 6 shows the relationship between average system SE and the transmit power. For accuracy, we take the proposed JPA-PSO scheme based on different pilot sequence lengths as optimization objective, and $\tau = K$, Optimal pilot sequence length $\tau = 26$ and $\tau = T$ are considered, then the average system SE is obtained by 500 times simulation. We can observe that when the transmission power is increased to a certain degree, these three schemes converge to a stable state. When the transmit power is about $25dB$, $20dB$ and $15dB$, the average system SE of the proposed JPA-PSO scheme with $\tau = 26$, $\tau = K$ and $\tau = T$ converge, respectively. Moreover, with the increase of transmit power, average system SE of these three options also improve. Besides, the proposed JPA-PSO scheme with $\tau = 26$ can greatly improve the system SE. However, the system SE does not change greatly under the transmit power gradually advances in the rising direction, and the reason is that larger transmit power affect the channel estimation and data detection of users in different cells.

V. CONCLUSION

This paper promoted the optimization of SE in massive MIMO system by proposing a joint pilot allocation and pilot sequence optimization scheme to restrain PC and maximize the system SE. A PC problem of pilot allocation and a nonlinear programming problem of pilot sequence optimization were formulated. To solve these problems, we first construct a range of pilot sequence based on the coherent interval. Then, we establish an allocation basis by defining a similarity to distinguish PC among users. An angle of arrival (AOA) positioning method is designed to obtain the user's location information. Under the location information, we can monitor and estimate the distance between users in different cells to denote the potential interference of transmit power among users. Then, we perform pilot allocation strategy in line with the similarity and distance, that is, assign orthogonal pilots to the users with greater similarity and shorter distance. Finally, we adopt extrapolation and interpolation methods to obtain the optimal pilot sequence. Simulation results have proven that combining pilot allocation and pilot sequence optimization can significantly restrain PC and improve system SE. Besides, the effectiveness of our proposed JPA-PSO scheme is superior to traditional pilot allocation schemes, and almost catches up to the exhaustive search with low computational complexity. Moreover, combined with pilot allocation, pilot sequence and power optimization can maximize system SE.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] F. Boccardi, R. W. Heath, Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] E. Björnson, J. Hoydis, and L. Sanguinetti, *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency*, vol. 11. New York, NY, USA: Now Publishers, 2017, pp. 3–4.

[5] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[6] D. Tes and P. Viswanath, *Fundamentals Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[7] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A comprehensive survey of pilot contamination in massive MIMO—5G system," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 905–923, 2nd Quart., 2016.

[8] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[9] F. Rusek, D. Persson, B. Kiong Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[10] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser mimo systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[11] T. V. Chien and E. Björnson, *Massive MIMO Communications*. Linköping, Sweden: Linköping Univ. Electronic Press, 2017.

[12] L. You, X. Gao, X.-G. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, Jun. 2015.

[13] X. Zhu, Z. Wang, L. Dai, and C. Qian, "Smart pilot assignment for massive MIMO," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1644–1647, Sep. 2015.

[14] K. Kim, J. Lee, and J. Choi, "Deep learning based pilot allocation scheme (DL-PAS) for 5G massive MIMO system," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 828–831, Apr. 2018.

[15] Z. Zhao, Z. Chen, and Y. Liu, "Cell sectorization-based pilot assignment scheme in massive MIMO systems," in *Proc. Wireless Telecommun. Symp. (WTS)*, New York, NY, USA, Apr. 2015, pp. 1–5.

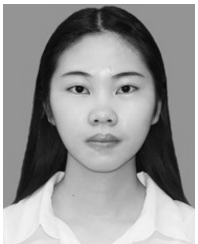
[16] X. Zhu, L. Dai, and Z. Wang, "Graph coloring based pilot allocation to mitigate pilot contamination for multi-cell massive MIMO systems," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1842–1845, Oct. 2015.

[17] X. Zhu, L. Dai, Z. Wang, and X. Wang, "Weighted-graph-coloring-based pilot decontamination for multicell massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2829–2834, Mar. 2017.

[18] H. Gao, T. Zhang, C. Feng, and Y. Wang, "Clustering based pilot allocation algorithm for mitigating pilot contamination in massive MIMO systems," in *Proc. Asia–Pacific Microw. Conf. (APMC)*, Kyoto, Japan, Nov. 2018, pp. 878–880.

[19] T. Van Chien, E. Bjornson, and E. G. Larsson, "Joint pilot design and uplink power allocation in multi-cell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2000–2015, Mar. 2018.

[20] S. Xu, H. Zhang, J. Tian, D. Wu, and D. Yuan, "Pilot length optimization for spectral and energy efficient D2D communications underlay massive MIMO networks," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Maui, HI, USA, Mar. 2018, pp. 855–860.



XINHUA NIE received the B.Eng. degree in electronic information engineering from the Changchun Institute of Technology, China, in 2018. She is currently pursuing the M.E. degree in communication engineering with the Guilin University of Electronic Technology, China, in 2018. Her research focuses on pilot allocation in multi-cell massive MIMO systems.

FENG ZHAO received the Ph.D. degree in communications and information systems from Shandong University, China, in 2007. He is currently a Professor with the Key Laboratory of Complex System Optimization and Big Data Processing, Guangxi Colleges and Universities, Yulin Normal University, Yulin, China. His research interests include wireless communications, signal processing, and information security.

• • •