

A pilot allocation method for multi-cell multi-user massive MIMO system

LI Yiming¹, DU Liping^{1,2,*}, and CHEN Yueyun¹

1. School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China;

2. Wireless and Space Communication Laboratory, Shunde Graduate School of University of Science and Technology Beijing, Foshan 528000, China

Abstract: Pilot contamination can spoil the accuracy of channel estimation and then has become one of the key problems influencing the performance of massive multiple input multiple output (MIMO) systems. This paper proposes a method based on cell classification and users grouping to mitigate the pilot contamination in multi-cell massive MIMO systems and improve the spectral efficiency. The pilots of the terminals are allocated one-bit orthogonal identifier to diminish the cell categories by the operation of exclusive OR (XOR). At the same time, the users are divided into edge user groups and central user groups according to the large-scale fading coefficients by the clustering algorithm, and different pilot sequences are assigned to different groups. The simulation results show that the proposed method can effectively improve the spectral efficiency of multi-cell massive MIMO systems.

Keywords: massive multiple input multiple output (MIMO), pilot allocation, cell classification, users grouping (UG), spectral efficiency.

DOI: [10.23919/JSEE.2021.000033](https://doi.org/10.23919/JSEE.2021.000033)

1. Introduction

The feature of massive multiple input multiple output (MIMO) is that the base station (BS) has many antennas (tens or even hundreds) [1,2]. As the key technology of the 5th generation (5G) wireless communication system [3–5], massive MIMO has many advantages such as high transmission rate, energy efficiency and spectral efficiency compared with the traditional MIMO technology [6–9]. Its advantages rely mostly on the accuracy of channel state information (CSI). At present, the massive MIMO system generally uses pilots for channel estimation. A pilot sequence is inserted in the transmitting data, and the receiver performs channel estimation by using the re-

ceived pilot symbols. Ideally, the massive MIMO system assigns each user a mutually orthogonal pilot sequence. When the coherence time is short, the length of the pilot sequence and the size of the pilot set are often limited. If the same pilot sequence is assigned to different users, it induces the estimation error of CSI, which is the problem of pilot contamination [10,11]. The influence of pilot contamination on system performance in multi-cell massive MIMO systems was studied in [12,13]. The research shows that the pilot contamination not only seriously affects the accuracy of the CSI, but also has a great influence on the downlink precoding of the system. Therefore, different pilot allocation strategies have been proposed to come up with the pilot contamination without increasing the transmit power and without changing the pilot structure. It is also an effective way to improve the spectral efficiency of the massive MIMO system. At present, pilot scheduling schemes include time-shifted pilots, superimposed pilot (SP), power control and users grouping.

The method of time-shifting pilots is to rearrange the transmission order of the uplink pilots of different cells in the frame structure [14,15]. The zero-forcing (ZF) precoding algorithm and the time-shifted pilots method were used to mitigate pilot contamination in [16]. Only one cell transmits pilots while the rest cells transmit data. However, when the users within different cells are very close, both data and pilots transmitted at the same time may cause interference and time synchronization is required. In [17], a partial-time-shifted pilot scheme was proposed. The scheme finds a better tradeoff between the users accommodation and orthogonality of pilots. However, it has stronger interference.

In the SP method, the pilots are periodically superimposed and transmitted on the data information in different time slots [18–20]. In [21], a time-multiplexed pilot sequence is combined with SP to mitigate pilot contamination. The interference caused by transmitting the pilot to-

Manuscript received February 13, 2020.

*Corresponding author.

This work was supported by the Scientific and Technological Innovation Foundation of Shunde Graduate School, USTB (BK19CF002).

gether with the data will be greatly reduced. In [22], using the theory of structured compressive sensing (CS), the structured subspace pursuit (SSP) algorithm was proposed with low pilot overhead.

The power control method is that different user groups transmit pilots with controlled power in their time slots [23,24]. A single-cell pilot power allocation method was proposed in [25]. This method allocates the pilot power of each user to maximize the sum-rate of downlink under the condition of total pilot power constraints. In [26], the problem of pilot and data power allocation in multi-cell massive MIMO systems was studied, and a joint allocation scheme of pilot and data power was proposed. The total transmit power of the system is reduced based on the signal-to-interference-plus-noise-ratio (SINR) requirements and power constraints of each user. The method proposed in [27] defined the relative channel estimation error (RCEE) metric and used the expected value of RCEE to improve the minimum uplink achievable rate. In [28, 29], the cells were separated into two groups while the pilot training phase was divided into two segments. Each group occupies one segment of pilot time slot, but the pilot contamination from all neighboring cells cannot be completely eliminated. The length of the time slot occupied by the pilot is twice as much as the conventional method. In [30], the cells were classified into three categories. The cells of each category are assigned orthogonal cell identification, which are added to the beginning of the pilot sequence. This method eliminates pilot contamination at the expense of the triple length of pilot time slot in the conventional method.

The method of users grouping (UG) is to allocate different pilots among user groups classified with different criteria to lower the pilot contamination [31–33]. The traditional UG method uses the distance between users and the BS. Fang et al. [34] used the interference strength between BS in the target cell and the users of the interfering cells to classify the interfering cell into two categories. The first type performs optimal allocation of pilots, and the second type randomly allocates pilots. This method effectively improves the downlink sum-rate of the system, but the pilot contamination generated by the second type of cell is still relatively large. In [35], users were divided into edge user groups and central user groups by using large-scale fading coefficients. The central users multiplex the same pilots, while the edge users with poor channel quality use orthogonal pilots. This method reduces pilot contamination of edge users and improves the sum-rate of the system with longer length of pilot time slot.

This paper proposes a method based on orthogonal cell identification and UG to reduce pilot contamination in the massive MIMO system. In the hexagonal cellular net-

work structure, the cells are divided into two categories. Users in the cells are divided into two groups: central group and edge group. One-bit orthogonal cell identification is added to the beginning of the pilot sequence. The central group multiplexes the same pilots and the edge group is assigned orthogonal pilots. The contribution of the paper lies in the following:

(i) One-bit orthogonal identifier ahead the pilot sequence marks the cells into two categories. Combining one-bit orthogonal identifier with the UG can reduce the length of the pilot and then improve the spectral efficiency of the system.

(ii) We propose a new UG method based on the Euclidean distances of the large-scale fading coefficient which considers the space information in each category, the users are grouped by using the cluster algorithm.

The numerical simulation results show that the proposed method can effectively reduce inter-cell pilot contamination in multi-cell systems and improve the spectral efficiency.

The rest of the paper is organized as follows. In Section 2, we review the multi-cell massive MIMO system model and the influence of pilot contamination on the system. The proposed method, namely the pilot allocation method based on one-bit orthogonal identifier and UG, is discussed in Section 3. The simulation results are shown in Section 4. Section 5 concludes the paper.

2. System model

Fig. 1 shows a multi-cell massive MIMO system model. There are L hexagonal cells in the cellular network, and each cell includes one BS equipped with M antennas, serving K ($M \gg K$) single antenna users. In order to facilitate the analysis of the problem, the target cell is surrounded by six neighbor cells. We assume that the system uses the time division duplex (TDD) mode. Based on the channel reciprocity criterion, the channel information can be estimated according to the uplink pilot transmitted by the user. The ZF precoding is used at the BS.

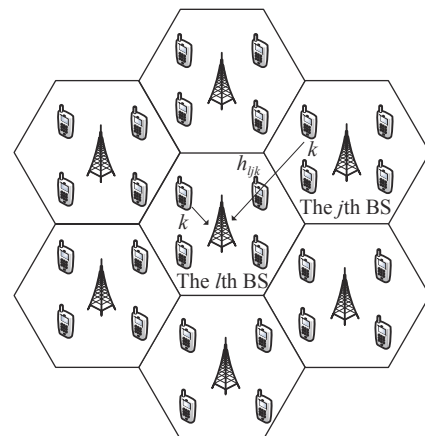


Fig. 1 Multi-cell massive MIMO system model

The channel quality is determined by the small-scale fading coefficient and the large-scale fading coefficient. $\mathbf{H}_{lj} \in \mathbf{C}^{M \times K}$ is the channel matrix between all users in the j th ($1 < j < L$) cell and the BS of the l th cell:

$$\mathbf{H}_{lj} = \mathbf{G}_{lj} \sqrt{\mathbf{D}_{lj}}, \quad (1)$$

$$\mathbf{G}_{lj} = \begin{bmatrix} g_{l1j1} & \cdots & g_{l1jK} \\ \vdots & \ddots & \vdots \\ g_{lMj1} & \cdots & g_{lMjK} \end{bmatrix}. \quad (2)$$

All elements of $\mathbf{G}_{lj} \in \mathbf{C}^{M \times K}$ are independent identical distribution (i.i.d). g_{lmjk} is the small-scale fading coefficient between the k th user in the j th cell and the m th antenna of the l th BS. We assume that the large-scale fading coefficients of a user to all antennas of the BS are the same because the distance between the antenna array is negligible compared with the distance between the users and the BS [36]. $\mathbf{D}_{lj} \in \mathbf{C}^{K \times K}$ is a diagonal matrix. The elements of $\mathbf{D}_{lj} = [\beta_{lj1}, \beta_{lj2}, \dots, \beta_{ljK}]$ are the large-scale fading coefficient between the k th user in the j th cell and the l th BS:

$$\beta_{ljk} = \frac{z_{ljk}}{(r_{ljk}/R)^\kappa} \quad (3)$$

where κ ($\kappa \geq 2$) is the path loss index, R is the radius of the cell, r_{ljk} is the distance between the k th user in the j th cell and the l th BS, and z_{ljk} represents the shadow fading and possesses a lognormal distribution $10\lg(z_{ljk}) \sim CN(0, \sigma_{\text{shadow}})$.

2.1 Channel estimation

For the unlink of the massive MIMO system, it is assumed that the pilot sequence transmitted by all users in the j th cell is $\boldsymbol{\varphi}_j = [\boldsymbol{\theta}_{j1}, \boldsymbol{\theta}_{j2}, \dots, \boldsymbol{\theta}_{jK}]^T$, while the pilot length $\tau \geq K$ and $\boldsymbol{\varphi}_j \boldsymbol{\varphi}_j^H = \mathbf{I}_K$. The training information received by the BS of the l th cell is

$$\mathbf{Y}_l = \sqrt{p_p \tau} \sum_{j=1}^L \mathbf{H}_{lj} \boldsymbol{\varphi}_j + \mathbf{N}_l \quad (4)$$

where p_p is the pilot transmit power, and \mathbf{N}_l is the additive white Gaussian noise in the l th cell. With the least square (LS) channel estimation, the estimation of CSI of the BS in the l th cell to the users in the j th cell is shown as

$$\begin{aligned} \hat{\mathbf{H}}_{lj} &= \frac{1}{\sqrt{p_p \tau}} \mathbf{Y}_l \boldsymbol{\varphi}_j^H = \\ &= \frac{1}{\sqrt{p_p \tau}} \left(\sum_{j=1}^L \mathbf{H}_{lj} \sqrt{p_p \tau} \boldsymbol{\varphi}_j + \mathbf{N}_l \right) \boldsymbol{\varphi}_j^H = \\ &= \mathbf{H}_{lj} + \left(\sum_{j=1, j \neq l}^L \mathbf{H}_{lj} \boldsymbol{\varphi}_j \right) \boldsymbol{\varphi}_j^H + \frac{1}{\sqrt{p_p \tau}} \mathbf{N}_l \boldsymbol{\varphi}_j^H. \end{aligned} \quad (5)$$

It can be known from the second term of (5) that the received signal is influenced by the same pilot of other cells. $\hat{\mathbf{h}}_{lk}$ is the k th column of $\hat{\mathbf{H}}_{lj} \in \mathbf{C}^{M \times K}$, which follows the distribution of $CN \sim (0_M, \zeta_{lk} \mathbf{I}_M)$, where

$$\zeta_{lk} = \sum_{j=1}^L \beta_{ljk} \boldsymbol{\theta}_{jk} \boldsymbol{\theta}_{jk}^H + \frac{1}{p_p \tau}. \quad (6)$$

The mean square error (MSE) per antenna of the LS estimation is

$$\eta_{lk}^{\text{LS}} = \frac{1}{M} \mathbf{E} \left[(\mathbf{h}_{lk} - \hat{\mathbf{h}}_{lk}) (\mathbf{h}_{lk} - \hat{\mathbf{h}}_{lk})^H \right] = \zeta_{lk} - \beta_{lk} \quad (7)$$

where \mathbf{h}_{lk} is the k th column of \mathbf{H}_{lj} . We define the error matrix as $\Delta \mathbf{H}_{lj} = \hat{\mathbf{H}}_{lj} - \mathbf{H}_{lj}$, $\Delta \mathbf{h}_{lk}$ is the k th column of $\Delta \mathbf{H}_{lj} \in \mathbf{C}^{M \times K}$.

2.2 Downlink data transmission

For the downlink of the multi-cell massive MIMO system in the TDD mode, the received signal y_{lk} of the k th user in the l th cell can be expressed as

$$y_{lk} = \sqrt{p_s} \sum_{j=1}^L \sum_{n=1}^K \mathbf{h}_{jlk}^T \mathbf{w}_{jn} x_{jn} + n_{lk} \quad (8)$$

where p_s is the BS transmit power, and x_{jn} is the data sent by the j th BS to the n th user in the cell. \mathbf{W}_j is precoding matrix of the j th cell, and \mathbf{w}_{jn} is the n th column of $\mathbf{W}_j \in \mathbf{C}^{M \times K}$.

In the downlink of the multi-cell massive MIMO system, we use ZF precoding to improve the anti-interference ability. The precoding matrix of the l th cell is $\mathbf{W}_l = \hat{\mathbf{H}}_{ll}^* (\hat{\mathbf{H}}_{ll}^T \hat{\mathbf{H}}_{ll}^*)^{-1}$, and $\text{tr}(\mathbf{W}_l^H \mathbf{W}_l) = K$. \mathbf{w}_{ln} is the n th column of \mathbf{W}_l . The received signal y_{lk} of the k th user in the l th cell can be expressed as

$$\begin{aligned} y_{lk} &= \sqrt{p_s} \sum_{j=1}^L \sum_{n=1}^K \mathbf{h}_{jlk}^T \mathbf{w}_{jn} x_{jn} + n_{lk} = \\ &= \sqrt{p_s} \sum_{n=1}^K \mathbf{h}_{llk}^T \mathbf{w}_{ln} x_{ln} + \sqrt{p_s} \sum_{j \neq l}^L \sum_{n=1}^K \mathbf{h}_{jlk}^T \mathbf{w}_{jn} x_{jn} + n_{lk} = \\ &= \sqrt{p_s} \sum_{n=1}^K (\hat{\mathbf{h}}_{llk}^T - \Delta \mathbf{h}_{llk}^T) \mathbf{w}_{ln} x_{ln} + \\ &= \sqrt{p_s} \sum_{j=1, j \neq l}^L \sum_{n=1}^K \mathbf{h}_{jlk}^T \mathbf{w}_{jn} x_{jn} + n_{lk} = \\ &= \sqrt{p_s} \hat{\mathbf{h}}_{llk}^T \mathbf{w}_{lk} x_{lk} + \sqrt{p_s} \sum_{n \neq k}^K \hat{\mathbf{h}}_{llk}^T \mathbf{w}_{ln} x_{ln} - \\ &= \sqrt{p_s} \sum_{n=1}^K \Delta \mathbf{h}_{llk}^T \mathbf{w}_{ln} x_{ln} + \sqrt{p_s} \sum_{j \neq l}^L \sum_{n=1}^K \mathbf{h}_{jlk}^T \mathbf{w}_{jn} x_{jn} + n_{lk} \end{aligned} \quad (9)$$

where the first item is the expected signal, and the other items represent inter-cell interference, channel estimation error, intra-cell interference and Gaussian white noise. x_{lk} is the data of the l th BS sent to the k th user. p_s is the BS transmission power. \mathbf{w}_{lk} is the k th column of \mathbf{W}_l . \mathbf{w}_{jn} is the n th column of \mathbf{W}_j . We assume $\hat{\mathbf{h}}_{ll}^T \mathbf{w}_{lk} = \delta_{ki}$. Then $\delta_{ki} = 0$ when $k \neq i$. Therefore, the SINR of the k th user in the l th cell is

$$\text{SINR}_{lk} = \frac{|\hat{\mathbf{h}}_{ll}^T \mathbf{w}_{lk}|^2}{\frac{|n_l|^2}{p_s} + \sum_{n=1}^K |\Delta \mathbf{h}_{ll}^T \mathbf{w}_{ln}|^2 + \sum_{j \neq l}^L \sum_{n=1}^K |\mathbf{h}_{jl}^T \mathbf{w}_{jn}|^2}. \quad (10)$$

When $M \rightarrow \infty$, (10) can be expressed as

$$\text{SINR}_{lk} \rightarrow \frac{|\hat{\mathbf{h}}_{ll}^T \mathbf{w}_{lk}|^2}{\sum_{n=1}^K |\Delta \mathbf{h}_{ll}^T \mathbf{w}_{ln}|^2 + \sum_{j \neq l}^L \sum_{n=1}^K |\mathbf{h}_{jl}^T \mathbf{w}_{jn}|^2}. \quad (11)$$

The spectral efficiency of the k th user in the l th cell is

$$R_{lk} = \frac{T - \tau}{T} \mathbb{E} \{ \log_2(1 + \text{SINR}_{lk}) \}. \quad (12)$$

Since $\log_2(1 + \text{SINR}_{lk})$ is a convex function, according to Jensen's inequality, the lower bound of the spectral efficiency of the k th user in the l th cell is

$$R_{lk} \geq \frac{T - \tau}{T} \log_2 \{ 1 + \mathbb{E}(\text{SINR}_{lk}) \}. \quad (13)$$

We substitute (11) into (13), the spectral efficiency of the l th cell can be expressed as

$$R_l = \frac{T - \tau}{T} \left\{ \sum_{k=1}^K \log_2 \left\{ 1 + \mathbb{E} \left[\frac{|\hat{\mathbf{h}}_{ll}^T \mathbf{w}_{lk}|^2}{\sum_{n=1}^K |\Delta \mathbf{h}_{ll}^T \mathbf{w}_{ln}|^2 + \sum_{j \neq l}^L \sum_{n=1}^K |\mathbf{h}_{jl}^T \mathbf{w}_{jn}|^2} \right] \right\} \right\}. \quad (14)$$

It can be seen from (12)–(14) that when the number of antennas approaches infinity, the spectral efficiency is related to the precoding matrix and the channel estimate. Since the precoding matrix is determined by the channel estimation, we can improve the accuracy of precoding by reducing the channel estimation error. From (5)–(7), the result of channel estimation is related to the large-scale fading coefficient and the pilots when the LS method is used. Therefore, when the large-scale fading coefficient is determined, the spectral efficiency of the system can be improved by designing the pilot allocation method.

3. Pilot allocation method based on UG and one-bit orthogonal identifier

3.1 UG based on clustering algorithm

To reduce the pilot contamination, we group users of each cell into center users and edge users using the clustering algorithm which is based on large-scale fading coefficients of users. Here the Euclidean distances of large-scale fading coefficients are used to measure the space information between users and BS. The center users and the edge users are expressed as K_{jc} and K_{je} , respectively. The channel quality between the k th user and the j th BS as the square of the large-scale fading coefficient is

$$S_j = [s_{j1}, s_{j2}, \dots, s_{jK}] = [\beta_{jj1}^2, \beta_{jj2}^2, \dots, \beta_{jjK}^2]. \quad (15)$$

The Euclidean distances can be expressed as

$$d = \sqrt{\sum_{m \in K_{jc}} (s_{jm} - s_c)^2} + \sqrt{\sum_{n \in K_{je}} (s_{jn} - s_e)^2}. \quad (16)$$

The UG algorithm based on clustering can be described as follows:

Step 1 Preprocess the channel quality. If some channel quality s_{jk} is greater than the predetermined threshold μ , this user is directly added to the central group K_{jc} , and its channel quality is removed from S_j .

Step 2 Randomly select two values s_c and s_e from S_j as the center of each group.

Step 3 Compute the Euclidean distances between each elements of S_j and the two centers. Divide each user in S_j into two groups according to whether it is close to s_c or s_e .

Step 4 Update the center of two new groups s_c and s_e according to the mean of each group.

Step 5 Repeat Step 3 and Step 4 until the users in all groups no longer change, or (16) does not change significantly.

Step 6 Repeat Step 1 to Step 5 until all cells complete the users grouping.

Algorithm 1 UG based on the clustering algorithm

1. **Input:** β_{jkk} ($j = 1, 2, \dots, L; k = 1, 2, \dots, K$)

2. **Output:** K_{jc}, K_{je} ($j = 1, 2, \dots, L$)

3. **for** $j=1:L$

4. **Calculate** $S_j = [s_{j1}, s_{j2}, \dots, s_{jK}]$ by (15)

5. **for** $k=1:K$

6. **if** $s_{jk} \geq \mu$

7. $k \in K_{jc}$

8. $S_j = S_j / s_{jk}$

9. **end if**

10. **end for**

11. Randomly select two elements in S_j , defined as

s_c and s_e

12. **Calculate** the Euclidean distance between the elements in S_j and two centers. Divide users into K_{jc} and K_{je} .

13. **Recalculate** the center of each category s_c and s_e .

14. Repeat Step 11 and Step 12 until K_{jc} and K_{je} no longer change or (16) does not change significantly.

15. **Return** K_{jc}, K_{je}

16. **end for**

Because of the division of users in each cell, the pilot sequences are divided into two groups P_1 and P_2 as well. The edge users of cells use the mutually orthogonal pilot sequence of group P_2 , and the central users of cells share the pilot sequences P_1 to reduce the extra pilot overhead.

After the users are grouped, the pilot length τ' is

$$\tau' = \max\{a_1, a_2, \dots, a_L\} + (b_1 + b_2 + \dots + b_L) \quad (17)$$

where a_1, a_2, \dots, a_L is the number of users in the center group of each cell, and b_1, b_2, \dots, b_L is the number of users in the edge group.

3.2 Cell classification method based on one-bit orthogonal identifier

We assume that there are seven cells in the massive MIMO system, and each cell has K users. The middle cell is the target cell. To further reduce the pilot contamination, all cells are classified into two categories, L_1 (yellow cells) and L_2 (blue cells). The pilot length of each user is τ , and one-bit orthogonal cell identifications $A_i = [0, 1]$ ($i = 1, \dots, L$) are allocated to each cell. As shown in Fig. 2, L_1 is assigned 1, and L_2 is 0.

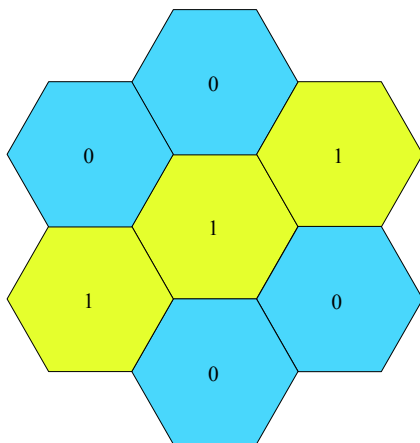


Fig. 2 One-bit orthogonal cell identifications allocated to two cell classifications

The orthogonal identification is added to the beginning of the pilot sequence. In Fig. 3, the one-bit orthogonal identifier is added to the head of the pilot sequence in the same cell category.

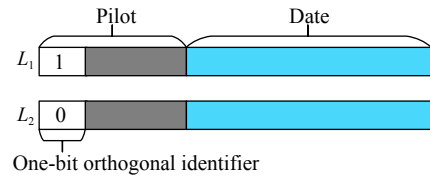


Fig. 3 Assignment of pilot in a time slot based on one-bit orthogonal identifier

When the BS receives the uplink pilot of users, the one bit of the head of each pilot sequence is taken out to be B . Then perform an exclusive OR (XOR) operation between B and the one-bit identification A_i , which can be shown as

$$B \oplus A_i = \begin{cases} 0 \\ 1 \end{cases}. \quad (18)$$

If the result of (18) is 0, it indicates that the pilot received by the BS is from the same category of cell. Then the pilot is sent to the BS processor that performs downlink channel estimation and data precoding according to the received pilot. If the result is 1, it will consider the corresponding pilot from the other category, so the corresponding pilot signal is filtered out. In this way, the pilot contamination of the blue cells is eliminated, so the target cell is only contaminated by pilots of two neighboring cells.

Therefore, our proposed method makes UG based on the clustering algorithm and the cell classification method based on the one-bit orthogonal identifier. It can be described as follows:

Step 1 Use the method in Subsection 3.2 to classify cells and assign the one-bit orthogonal identifier to users in each cell.

Step 2 Use the method in Subsection 3.1 to group users in each cell to obtain the center user groups K_{jc} and edge user groups K_{je} .

Step 3 Assign pilot sequences P_1 and P_2 to K_{jc} and K_{je} , respectively.

Step 4 Add the pre-assigned one-bit orthogonal identifier to the beginning of the pilot sequence of users for each cell.

In our proposed method, the pilots of the target cell are only contaminated by the pilots of users in the edge user group of two yellow neighboring cells, and the length of the pilot sequence τ can be expressed as

$$\tau = \max_{i \in L_1} \{a_i\} + \sum_{i \in L_1} b_i + 1 \quad (19)$$

where a_i is the number of central users in the i th cell, and b_i is the number of edge users in the i th cell.

The spectral efficiency of the l th cell of the proposed method can be expressed as

$$R_i = \frac{T - \left(\max_{i \in L_1} \{a_i\} + \sum_{i \in L_1} b_i + 1 \right)}{T} \sum_{k=1}^K \log_2 \{1 + E(\text{SINR}_{ik})\}. \quad (20)$$

The MSE per antenna of the LS estimation can be expressed as

$$\eta_{lk}^{\text{LS}} = \begin{cases} \sum_{j \in L_1} \beta_{ljk} \theta_{jk} \theta_{lk}^{\text{H}} + \frac{1}{p_p \tau} - \beta_{llk}, & k \in K_{lc} \\ \frac{1}{p_p \tau}, & k \in K_{le} \end{cases} \quad (21)$$

where k_j is the k th user in the j th cell. When $k_j \in K_{je}$, $\theta_{jk} \theta_{lk}^{\text{H}} = 0$.

3.3 Computational complexity analysis

In our proposed methods, the computational complexity is mainly made by three stages. The first stage is to directly list the users whose large-scale fading coefficient is larger than the preset threshold in the center group, and the complexity is $O(K)$. The second stage is to classify the remaining users by using the clustering algorithm. The complexity is $O(nK_i t_i)$, where n is the number of categories ($n = 2$), K_i is the number of remaining users after the first stage for the i th cell, t_i is the number of iterations. The complexity of the second stage can also be expressed as $O(K_i t_i)$. The third stage is that the base station takes out one bit of the head of each pilot sequence for XOR operation and the complexity is $O(K)$.

As for the UG method in [35], it needs two stages: the first stage is to calculate the mean of the maximum and the minimum of the square of the large-scale fading coefficient from the users to the base station in each cell, and its complexity is $O(K)$; the second stage is to group users according to the mean, and the complexity is $O(K)$.

It can be seen from the analysis that the second stage makes the computational complexity of the proposed method slightly higher than the UG method in [35]. The additional complexity is in proportion to the number of users in each cell and the iteration number of the clustering algorithm. From our experiment observation, the iteration number of the clustering algorithm is between 2 to 10, but it can dynamically classify users according to large-scale fading coefficients and obtain better results of UG.

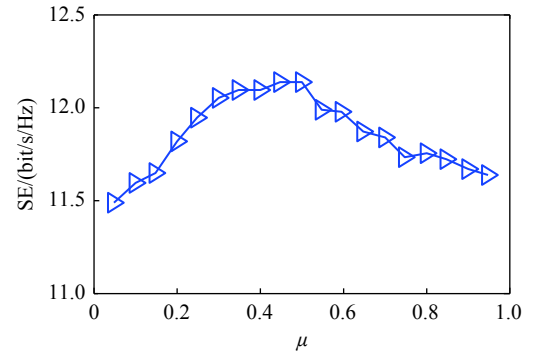
4. Simulation results

In this section, we will study the performance of the proposed scheme through Monte Carlo simulation. The system parameters are given in Table 1.

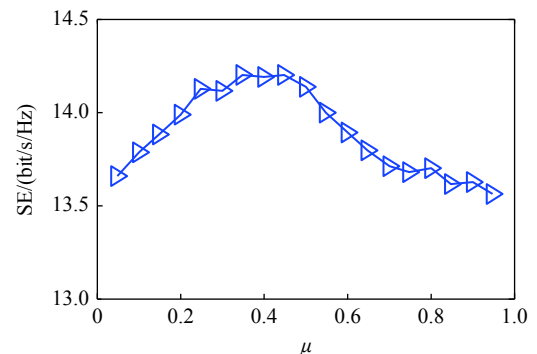
Table 1 Basic parameters of system simulation

| Parameter | Value |
|--|--------|
| Number of cells L | 7 |
| Cell radius R/m | 500 |
| Number of BS antennas M | 16–256 |
| Number of users in each cell K | 10–40 |
| Path loss κ | 3.8 |
| Signal-to-noise ratio (SNR)/dB | 0–20 |
| Log normal shadowing fading $\sigma_{\text{shadow}}/\text{dB}$ | 8 |

Fig. 4 shows the spectral efficiency of the target cell which changes with the predetermined threshold μ under different simulation parameters. The users are grouped based on the clustering algorithm. It is shown that the spectral efficiency (SE) first increases with μ and then decreases after it reaches the peak, and finally approaches a fixed value. It is known that the number of pre-selected central users decreases with μ . If μ is too small, more users are included into the pre-selected central users and some of them may be users with poor channel quality. When μ is large, some users with high channel quality may be not selected. Both cases can affect the clustering result. The SE reaches the maximum with μ ranging from 0.35 to 0.5 and trends to constant after $\mu = 0.7$.



(a) SNR=5 dB, $K=10$, $M=128$



(b) SNR=10 dB, $K=10$, $M=64$

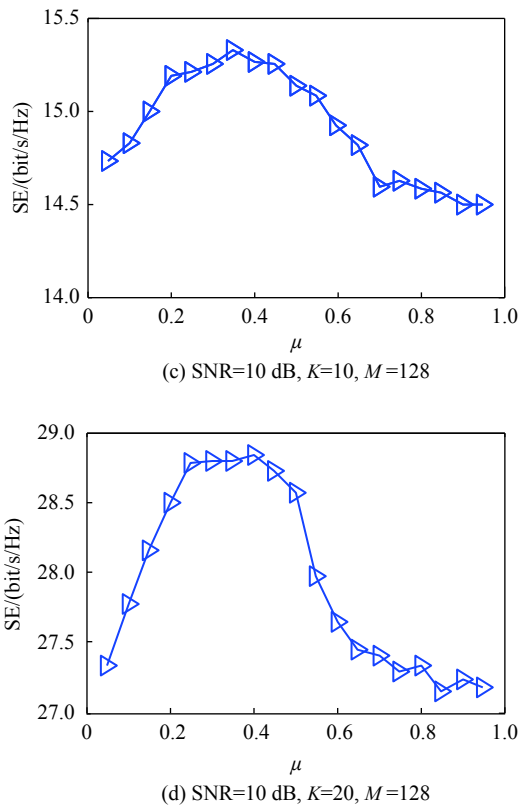


Fig. 4 Predetermined threshold μ on the spectral efficiency of the target cell

Fig. 5 shows the achievable sum-rate of the target cell of our proposed method (cell classification and UG based on the clustering algorithm), the UG method in [35], the method without pilot contamination (the users in all cells use orthogonal pilot sequences) and the traditional method (each cell multiplexes the same pilot sequences) when the numbers of users K increases. The number of antennas is 128. The SNR is 10 dB. The predetermined threshold μ is 0.35. As the number of users increases, the achievable sum-rate of the four methods tends to increase. The proposed method and the UG method are close to the method without pilot contamination. The proposed method has a little better performance than the UG method, which indicates that these sum-rates of the two methods are similar.

Fig. 6 shows the SE of the target cell of our proposed method, the UG method in [35], the method without pilot contamination and the traditional method when the number of users K increases. The number of antennas is 128. The SNR is 10 dB. The predetermined threshold μ is 0.35. It is shown that the SE of the four methods increases with the number of users. The SE of these four methods is similar when the number of users is small. As the number of users increases, the advantages of the proposed method will be more obvious. It is known that the

pilot length will increase with the number of users. The time slot of the UG method is $\max\{a_1^{\text{UG}}, a_2^{\text{UG}}, \dots, a_L^{\text{UG}}\} + (b_1^{\text{UG}} + b_2^{\text{UG}} + \dots + b_L^{\text{UG}})$, where $a_1^{\text{UG}}, a_2^{\text{UG}}, \dots, a_L^{\text{UG}}$ is the number of users in each center group, and $b_1^{\text{UG}}, b_2^{\text{UG}}, \dots, b_L^{\text{UG}}$ is the number of users in the each edge group by this method. The time slot of the method without pilot contamination is $7K$ and the SE of this method is the least. From (18), the proposed method reduces the time slot occupied by transmitting pilots, then improves the SE.

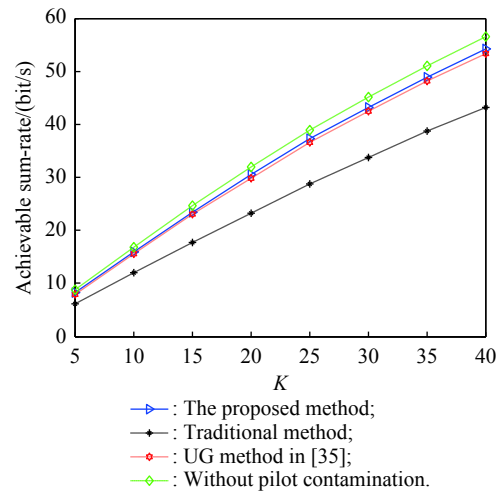


Fig. 5 Influence of the number of users K on the achievable sum-rate of the target cell

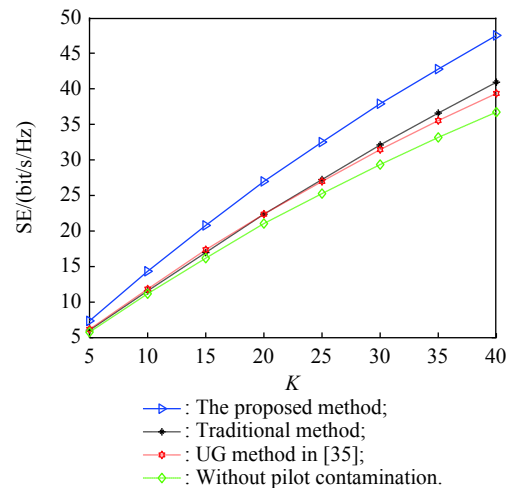


Fig. 6 Influence of the number of users K on SE of the target cell

Fig. 7 shows the spectral efficiency of the target cell of our proposed method, the UG method in [35], the method without pilot contamination and the traditional method when the SNR increases. The number of antennas is 128. The number of users in each cell is 15. The predetermined threshold μ is 0.35. It can be seen in Fig. 7, the SE of the four methods increases with the SNR. The SE of

our proposed method is better than that of the other three methods. The SE of the UG method is similar to the method without pilot contamination. The traditional method has the lowest SE with a high SNR.

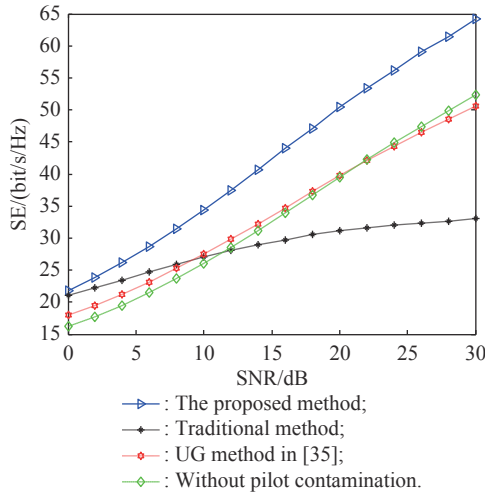


Fig. 7 Influence of SNR on SE of the target cell

Fig. 8 shows the SE of the target cell of our proposed method, the UG method in [35], the method without pilot contamination and the traditional method when the antennas number increases. The SNR is 5 dB. The number of users in each cell is 15. The predetermined threshold μ is 0.35. The SE of four methods increases with the number of antennas from 16 to 256. The proposed method is better than the other three methods. Compared with the UG method, the SE of the proposed method is improved by 4 bit/s/Hz when $M=256$.

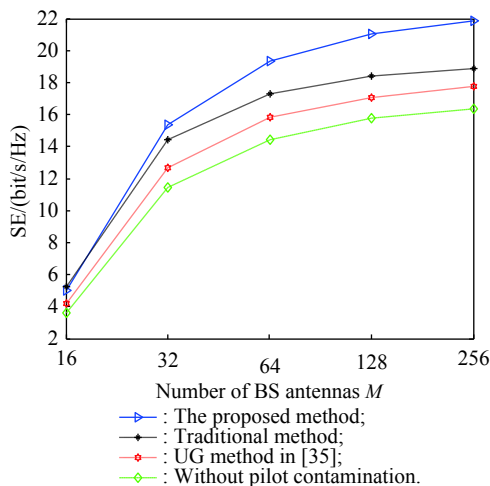


Fig. 8 Influence of the number of antennas on SE of the target cell

From the results of sum-rate and SE, the advantages of the proposed method are more obvious.

5. Conclusions

This paper proposes a method based on the one-bit orthogonal identifier and UG to mitigate pilot contamination in multi-cell massive MIMO systems while improving the SE. The pilots between different categories of cells classified by the one-bit orthogonal identifier are filtered out. At the same time, users of each cell are grouped according to the large-scale fading coefficient to reduce pilot contamination between cells of the same category. The simulation results show that the proposed method can effectively improve the SE of multi-cell massive MIMO systems.

References

- [1] ASAAD S, RABIEI A M, MULLER R R. Massive MIMO with antenna selection: fundamental limits and applications. *IEEE Trans. on Wireless Communications*, 2018, 17(12): 8502–8516.
- [2] MARZETTA T L. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans. on Wireless Communications*, 2010, 9(11): 3590–3600.
- [3] BJORNSON E, WYMEERSCH H, SANGUINETTI L, et al. Massive MIMO is a reality? What is next? Five promising research directions for antenna arrays. *Digital Signal Processing*, 2019, 94(1): 3–20.
- [4] CHIH-LIN I, HAN S, XU Z, et al. 5G: rethink mobile communications for 2020+. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2016, 374(2062): 20140432.
- [5] LARSSON E G, EDFORS O, TUFVESSON F, et al. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 2013, 52(2): 186–195.
- [6] TAN W J, FENG X A, LIU G Y, et al. Spectral efficiency of massive MIMO systems with multiple sub-arrays antenna. *IEEE Access*, 2018, 6: 31213–31223.
- [7] KAMGA G N, XIA M H, AISSA S, et al. Spectral-efficiency analysis of massive MIMO systems in centralized and distributed schemes. *IEEE Trans. on Communications*, 2016, 64(5): 1930–1941.
- [8] NGO H Q, LARSSON E G, MARZETTA T L. Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Trans. on Communications*, 2013, 61(4): 1436–1449.
- [9] NEUMANN D, WIESE T, JOHAM M, et al. A bilinear equalizer for massive MIMO systems. *IEEE Trans. on Signal Processing*, 2018, 66(14): 3740–3751.
- [10] BJORNSON E, LARSSON E G, DEBBAH M. Optimizing multi-cell massive MIMO for spectral efficiency: how many users should be scheduled? *Proc. of the IEEE Global Conference on Signal and Information*, 2014: 612–616.
- [11] HOYDIS J, TEN BRINK S, DEBBAH M. Massive MIMO in the UL/DL of cellular networks: how many antennas do we need. *IEEE Journal on Selected Areas in Communications*, 2013, 31(2): 160–171.
- [12] ZHANG X Q, WU S B, YU S S, et al. Spectral efficiency of massive MIMO networks with pilot contamination and channel covariance matrix estimation. *IET Communications*, 2019, 13(1): 59–65.
- [13] LU L, LI G Y, SWINDLEHURST A L, et al. An overview of massive MIMO: benefits and challenges. *IEEE Journal of Se-*

- lected Topics in Signal Processing, 2014, 8(5): 742–758.
- [14] MORALES J L L, ROY S. Channel estimation using time-shifted pilot sequences in non-cooperative cellular TDD networks with large antenna arrays. Proc. of the Asilomar Conference on Signals, Systems and Computers, 2013: 1258–1262.
- [15] WEI T, FENG W, SHI R, et al. Coordinated multi-cell multi-user pilot allocation in massive MIMO systems with time-shifted pilots. Proc. of the International Conference on Wireless Communications & Signal Processing, 2015. DOI: 10.1109/WCSP.2015.7341265.
- [16] JIN S, WANG X T, LI Z, et al. On massive MIMO zero-forcing transceiver using time-shifted pilots. *IEEE Trans. on Vehicular Technology*, 2016, 65(1): 59–74.
- [17] CHANG W, HUA Y K, LIAO S F. Partial overlapped time-shifted pilots for massive MIMO systems. *IEEE Communications Letters*, 2017, 21(11): 2480–2483.
- [18] YANG X, ZHANG S M, GAO B, et al. A low complexity joint user grouping and resource allocation algorithm in massive MIMO systems. Proc. of the 19th IEEE International Conference on Communication Technology, 2019: 914–919.
- [19] WANG Z Y, ZHOU T, XU T H, et al. Joint user grouping and power control of virtual MIMO in fog computing. Proc. of the IEEE Globecom Workshops, 2019. DOI: 10.1109/GCWshps45667.2019.9024447.
- [20] KONG Q, FENG W J. Pilot allocation algorithm based on user grouping in massive MIMO systems. Proc. of the 5th IEEE International Conference on Computer and Communications, 2019: 1120–1125.
- [21] LAGO L A, ZHANG Y, HE Z, et al. Pilot decontamination based on superimposed pilots in massive MIMO systems. Proc. of the 88th IEEE Vehicular Technology Conference, 2018. DOI: 10.1109/VTCFall.2018.8690763.
- [22] WANG Z C, GAO Z, DAI L L. Structured compressive sensing based superimposed pilot design in downlink large-scale MIMO systems. *Electronics Letters*, 2014, 50(12): 896–898.
- [23] GHAZANFARI A, CHENG H V, BJORNSON E, et al. Enhanced fairness and scalability of power control schemes in multi-cell massive MIMO. *IEEE Trans. on Communications*, 2020, 68(5): 2878–2890.
- [24] WANG J T. Joint MMSE equalization and power control for MIMO system under multi-user interference. *IEEE Communications Letters*, 2012, 16(1): 54–56.
- [25] ZHENG X R, ZHANG H, XU W, et al. Optimized pilot power allocation for heterogeneous users in massive MIMO downlinks. Proc. of the International Conference on Information and Communication Technology Convergence, 2014: 246–250.
- [26] ZHANG Y, ZHU W P, OUYANG J. Energy efficient pilot and data power allocation in multi-cell multi-user massive MIMO communication systems. Proc. of the 84th IEEE Vehicular Technology Conference, 2016. DOI: 10.1109/VTCFall.2016.7880987.
- [27] LIU P, JIN S, JIANG T, et al. Pilot power allocation through user grouping in multi-cell massive MIMO systems. *IEEE Trans. on Communications*, 2017, 65(4): 1561–1574.
- [28] WANG H R. Pilot contamination reduction in very large MIMO multi-cell TDD systems. *Journal of Signal Processing*, 2013, 29(2): 171–180. (in Chinese)
- [29] WANG H R, WANG Y H, HANG Y M, et al. Pilot contamination reduction in very large MIMO cellular network. *Journal on Communications*, 2014, 35(1): 24–33. (in Chinese)
- [30] XU Y H, WU Y. An approach of pilot contamination reduction based on power control and orthogonal identification. Proc. of the 3rd IEEE International Conference on Cloud Computing and Big Data Analysis, 2018: 476–480.
- [31] SHAFIN R, LIU L J. Superimposed pilot for multi-cell multi-user massive FD-MIMO systems. *IEEE Trans. on Wireless Communications*, 2020, 19(5): 3591–3606.
- [32] VERENZUELA D, BERGSTROM A, BJORNSON E. Optimal power control for superimposed pilots in uplink massive MIMO systems. Proc. of the 52nd Asilomar Conference on Signals, Systems, and Computers, 2018: 499–503.
- [33] BAO D, QIN G D, DONG Y Y. A superimposed pilot-based integrated radar and communication system. *IEEE Access*, 2020, 8: 11520–11533.
- [34] FANG X, ZHANG J F, CAO H Y, et al. Dynamic pilot assignment in large scale MIMO systems. *Journal of Electronics & Information Technology*, 2016, 38(8): 1901–1907. (in Chinese)
- [35] WU Y C, LIU T, CAO M, et al. Pilot contamination reduction in massive MIMO systems based on pilot scheduling. *EURASIP Journal on Wireless Communications and Networking*, 2018, 1: 21. DOI: 10.1186/s13638-018-1029-1.
- [36] ALEXEI A, LIANGBIN L, MARZETTA T L. Interference reduction in multi-cell massive MIMO systems with large-scale fading precoding. *IEEE Trans. on Information Theory*, 2018, 64(9): 6340–6361.

Biographies



LI Yiming was born in 1995. He is currently pursuing his master's degree in the information and communication engineering from the School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB). His research interest is wireless communication. E-mail: liyiming95111@126.com



DU Liping was born in 1975. She is a professor in the School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB) and Shunde Graduate School of USTB. Her research interests are wireless communication, signal processing and pattern recognition. E-mail: Dlp2001@ies.ustb.edu.cn



CHEN Yueyun was born in 1966. She is a professor in the School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB). She is an IEEE member. Her research interests are wireless mobile communication system theory, broadband wireless communication technology and wireless resource management. E-mail: chenyy@ustb.edu.cn