Pilot allocation scheme based on coalition game for TDD massive MIMO systems

Hui Zhi[1*] and Xiaoguang Ding[2]

Abstract

Pilot contamination is a major factor that restricts the system performance of time division duplex (TDD) massive multiple-input multiple-output (MIMO) systems. Reasonable pilot allocation is an effective method to reduce pilot contamination. However, most of the existing pilot allocation methods focus on cells of the same size, a regular shape, an equal number of users per cell, and/or independent fading channels. When they are generalized or modified to actual scenes (in which cells have irregular shapes, users are randomly distributed, and channels are correlated fading channels), they are inflexible and cannot guarantee the target performance requirement. Considering both flexibility of pilot allocation and system performance, this paper introduces the idea of a coalition game into pilot allocation and proposes a pilot allocation scheme based on the coalition game. Different utility functions are analyzed under correlated fading channels; the adjustment principle of coalition structure and the coalition formation algorithm are given. Simulation results show that the proposed scheme can be used in any actual scene as well as guarantee the target performance requirement.

Keywords: Massive MIMO, Pilot allocation scheme, Coalition game

1 Introduction
1.1 The research background of this paper

With the appearance of smartphones and tablets, wireless communication requires increasingly higher communication rates. Therefore, many technologies have been proposed to improve the communication rate of 5G, and massive multiple-input multiple-output (MIMO) is one of these technologies. Compared with traditional wireless networks, massive MIMO significantly enhances the spectral and power efficiency [1].

However, the excellent performance of massive MIMO depends on the accuracy of channel estimation. For conventional channel estimation in time division duplex (TDD) massive MIMO systems, users transmit pilot signals in the uplink, and base stations (BSs) receive the pilot signal and estimate channel state information (CSI). Then, BSs can use the CSI for uplinking received signal detection and downlinking transmission precoding. Therefore, the accuracy

of channel estimation is one of the key factors that affect the system performance of uplink and downlink. In pilot-based channel estimation, the number of pilots is finite and limited by the channel coherence time, so mobile users in different cells might have to reuse the same pilot. Pilot reuse seriously affects the channel estimation accuracy at BSs and further influences the system performance of uplink and downlink, where the influence is known as pilot contamination [2]. Recent studies have shown that pilot contamination is one of the key factors that affect the performance of TDD massive MIMO systems [1–4]. Therefore, many methods have been proposed to reduce pilot contamination, such as pilot contamination precoding [4], time-shifted pilots [5, 6], spatial domain-based pilot allocation [7], design and transmit power control of pilot sequences [8, 9], and pilot reuse [10, 11].

For instance, reference [4] shows that the performance of TDD massive MIMO system is limited by pilot contamination. Furthermore, reference [4] considers the uplink precoding and downlink precoding (called pilot contamination precoding) based on slow fading coefficients to reduce pilot contamination. Papers [5, 6]

* Correspondence: zhihui_0902@163.com
[1]Key Lab of Intelligent Computing and Signal Processing (Ministry of Education), Anhui University, Hefei 230601, China
Full list of author information is available at the end of the article

discuss the time-shifted pilot method in hexagonal cells, which means cells are divided into different groups, and cells in different groups transmit pilots in different time slots; analyses show that this method can effectively reduce pilot contamination. However, this method assumes that the number of users in each cell is fixed and equal. In addition, it requires strict system synchronization, making implementation complex and difficult. Paper [7] proposes a spatial domain method with the main idea that users with high channel orthogonality can use the same pilot. This method also assumes that cells have a regular shape and that the number of users in each cell is fixed and equal. Furthermore, this method is highly complex because BSs need to know the CSIs between all users and all BSs to carry out the search for highly orthogonal channels. Paper [8] gives a design for non-orthogonal pilot signals, and paper [9] proposes a pilot power allocation method based on cell grouping to alleviate pilot contamination. The methods in [8, 9] both assume that cells have a regular shape and the number of users per cell is fixed. Paper [10] presents a pilot reuse method based on cell grouping and analyzes the system performance of different pilot reuse factors. In paper [11], a fractional pilot reuse method is proposed in which users are classified into cell-center users and cell-edge users, where the cell-center users reuse the same pilot subset and the cell-edge users use the remaining pilots according to cell grouping.

Studies [12, 13] try to introduce game theory into pilot allocation to reduce pilot contamination. However, the pilot allocation methods in [12, 13] both consider the games between cells. Simply because of the games between cells, the suitable application scene of [12, 13] is still fixed and has an equal number of users in each cell, although reference [13] attempts to discuss irregularly shaped cells.

Reviewing the studies in [4–13], reasonable pilot allocation can effectively reduce pilot contamination and improve system performance. However, in order to facilitate theoretical analysis and simulation, most of these pilot allocation methods assume that cells have a regular shape (most are hexagons), the number of users per cell is equal, and/or channels are independent. However, in reality, the cell shapes are affected by various factors, such as landforms and buildings, and do not always have a regular shape. Moreover, mobile users are randomly distributed in reality, so it is impossible for the number of users in each cell to be equal or fixed. When the number of base station antennas is very large, it is difficult to ensure the distance between the antennas, and it is difficult to ensure that all channels are independent.

Most of these pilot allocation methods (in references [4–13]) assume that the number of users per cell is equal

or fixed for two main reasons. First, when the number of users in each cell is equal or fixed, it is easy to allocate pilots. For example, in references [4–13], in order to avoid the use of the same pilot in adjacent cells (pilot contamination is large in this case), these pilot allocation methods [4–13] always reserve a certain number of pilots for each cell. If the number of users in each cell is equal or fixed, it is convenient to estimate the number of pilots per cell, so that when pilots are allocated, the pilots of adjacent cells can be differentiated. Second, when the number of users in each cell is equal, the dimension of the channel matrix in each cell is equal, which is convenient for merging in mathematical operation in theoretical analysis.

If we modify these methods [4–13] to the actual scenes in which all users are randomly distributed and the number of users per cell is unequal, the system performance will not be guaranteed. For example, if the number of users is large in a cell, using these methods [4–13], pre-allocated pilots are insufficient, and pilot reuse will cause large pilot contamination. On the other hand, if the number of users in a cell is small, using these methods [4–13], the pilot resource cannot be fully utilized, and this will cause the waste of pilot resources and then affect overall performance. Obviously, this point reflects that these methods [4–13] lack flexibility.

The assumption of a regular cell shape is also beneficial to pilot allocation. When the cell shape is regular, cells can be grouped first. For example, as shown in Fig. 1a, taking pilot reuse factor 3 as an example, the cells in the picture have three types of background pattern. The cells with the same background pattern are grouped together and reuse the same pilot subset (the total pilot set has been divided into three orthogonal pilot subsets). Cells with different background patterns use different orthogonal pilot subsets. References [9–11] use similar cell grouping pilot allocation methods. However, the actual situation is that the cell shape is irregular. The pilot multiplexing in 11 cells shown in Fig. 1b is based on the pilot allocation method in Fig. 1a. It can be seen that in Fig. 1b, because of the irregular cell shape, the same pilot subset is used in adjacent cells, which will cause serious pilot contamination; thus, the system performance cannot be guaranteed.

So we can conclude that if we modify these methods [4–13] to the actual scenes (in which cells have irregular shapes, users are randomly distributed), the system performance will not be guaranteed.

Papers [14, 15] propose different greedy algorithms to allocate pilots in the scene of correlated fading channels. When compared with the methods in [4–13], greedy algorithms in papers [14, 15] are flexible and can be easily
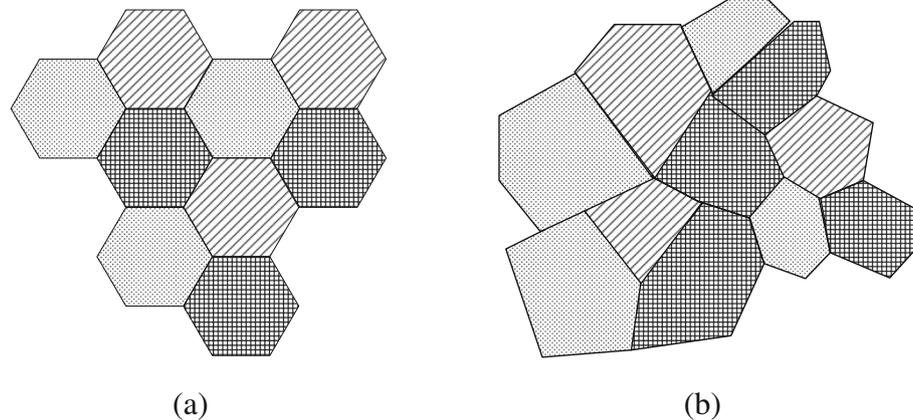
Fig. 1 Pilot allocation based on cell grouping

generalized or modified to the actual scenes, but they still cannot guarantee the target performance requirement. Paper [14] presents a statistical greedy pilot scheduling algorithm, which focuses on users in a single cell and searches users with small channel correlations to use the same pilot. This algorithm does not consider the application scenario of multi-cells and the correlation of the channels between different users and different BSs, so it cannot reduce the interference among different cells and thus cannot guarantee system performance in multi-cell scenarios. That is to say, although this algorithm can be generalized to multi-cell scenarios of arbitrary BS locations and an unequal number of users per cell, it cannot guarantee target performance satisfaction. Thus, it is essentially different from our work; the proposed pilot allocation scheme in our work can be used in arbitrary scenarios as well as guarantee system performance.

Paper [15] proposes a greedy approach; its main idea is selecting one user from each cell (there are P users per cell) to form a user set during each search, and users in the same set use the same pilot. The principle of user selection is to minimize the sum mean square error of channel estimation for the users using the same pilot. The greedy approach assumes the number of users per cell is equal; it cannot be used in the case of different numbers of users per cell, so it is different from our work and we have to make modifications in simulations. In addition, it only considers the sum-mean-square error of users with the same pilot and does not consider the total performance of all users; furthermore, it is a one-time search for all users and cannot achieve global optimization, whereas the proposed pilot allocation scheme in our work considers the total performance of all users and performs a cyclic search for all users. Therefore, in comparison, the greedy approach in [15] cannot meet the needs of total performance, whereas

our scheme can meet overall performance requirements through cyclic search.

1.2 The main contents of this paper

In order to find the flexible pilot allocation method that can be used in any actual scenes and provide performance guarantee, we introduce the idea of a coalition game into pilot allocation and propose a pilot allocation scheme based on the coalition game in this paper. We consider the game between users and divide users into different sub-coalitions. Because this coalition game is among users, the proposed pilot allocation scheme is more flexible and can be used in any actual scene. Because the purpose of each coalition adjustment is to improve the target performance (average utility function), the target performance can be satisfied through the cyclic search of coalition adjustment. The contributions of this paper are as follows.

1) To be applied in arbitrary actual scene (with irregular shape cells, unequal number of users per cell, and correlated fading channels) and provide target performance guarantee, a novel pilot allocation scheme based on the coalition game is proposed.
   For the game between users, the coalition structure is defined, different utility functions are analyzed, and the coalition formation algorithm is given.
2) The implementation method of the proposed pilot allocation scheme is presented.
3) Simulations are developed to show that the proposed scheme can be used in any actual scene, effectively reduce pilot contamination, and provide a certain target performance guarantee. Moreover, the comparisons of computational complexity between different schemes are discussed.

The remainder of this paper is organized as follows. The system model for the massive MIMO system is described in Section 2. Different utility functions are analyzed under correlated fading channels in Section 3, including mean square error of channel estimation (MSE-CE), mean square error of signal detection (MSE-SD), received signal-to-interference-plus-noise ratio (SINR), and spectrum efficiency (SE). In Section 4, the definition of the coalition game is given, and then the coalition structure, the adjustment principle of coalition structure, and the condition of final stability are defined. After that, the coalition formation algorithm is given, a pilot allocation scheme based on the coalition game is proposed, and the implementation method of the proposed scheme is discussed. In Section 5, simulations are developed with respect to different utility functions, and the proposed pilot allocation scheme is compared with the greedy approach in reference [15] and the random coalition scheme. Section 6 concludes this paper. The proof of theorem is given in the Appendices 1, 2, and 3.

2 System model

Consider the massive MIMO system as shown in Fig. 2. There are L cells and N users in a fixed area, where we focus on the uplink. Each cell is assigned an index in the set $\mathscr{L} = \{1, 2, \cdots, L\}$. The kth user in the jth cell is denoted by user (j, k). Assume cell j has K_j active users, thus $\sum_{j=1}^{L} K_j = N$. All users form the set $\mathcal{N}$ with $\mathcal{N} = \{(1, 1), \cdots (1, K_1), \cdots (j, 1), \cdots (j, K_j), \cdots (L, 1), \cdots (L, K_L)\}$. Each cell has one BS, and one cell and its BS use the same index number, i.e., BS j is in the jth cell. Each BS is equipped with an array of M antennas, the shape of antennas at BSs is uniformly spaced linear array (ULA) [15], and each user is equipped

with a single antenna. Assume that BSs are subject to a random distribution and that the users are subject to another random distribution, the distributions of BSs and users are independent. Each user selects and accesses its nearest BS according to the strength of the received BS signals, so the number of users in each cell ($K_j, j \in \{1, 2, \cdots, L\}$) is not equal.

Assume $z_{jk}^u \in \mathbb{R}^2$ and $z_l^b \in \mathbb{R}^2$ are the locations of user (j, k) and BS l respectively, and z_{jk} is known at all BSs. $d_l(z_{jk})$ is the distance from user (j, k) to BS l. The channel between user (j, k) and BS l is defined as $\boldsymbol{h}_{ljk} \in \mathbb{C}^{M \times 1}$, and we consider correlated Rayleigh fading channels [15–17]. (For other types of correlated channel models and independent channel models, our pilot allocation scheme is also applicable.) The channel $\boldsymbol{h}_{ljk}$ is the superposition of F arriving paths,

$$\boldsymbol{h}_{ljk} = \frac{1}{\sqrt{F}} \sum_{f=1}^{F} a\left(\theta_{ljk}^{(f)}\right) \alpha_{ljk}^{(f)} \tag{1}$$

where $a(\theta_{ljk}^{(f)}) \in \mathbb{C}^{M \times 1}$ is the antenna steering vector corresponding to angle-of-arrivals (AoAs) $\theta_{ljk} \in [0, 2\pi)$, f is the path index, and $\alpha_{ljk}^{(f)} \overset{i.i.d.}{\sim} \mathcal{CN}(0, \beta_{ljk})$ is the channel coefficient of the fth path, β_{ljk} denotes the large-scale fading coefficient and $\beta_{ljk} = \delta \|z_{jk}^u - z_l^b\|_2^{-\nu}$, where ν is the path loss exponent, δ is a constant of channel coefficient, and we let δ be a constant for arbitrary path f and arbitrary channel $\boldsymbol{h}_{ljk}$ in our simulations.

The expression of $a(\theta)$ is

$$a(\theta) = \begin{bmatrix} 1 \\ e^{-j2\pi \frac{D}{\lambda} \cos(\theta)} \\ \vdots \\ e^{-j2\pi(M-1)\frac{D}{\lambda} \cos(\theta)} \end{bmatrix} \tag{2}$$
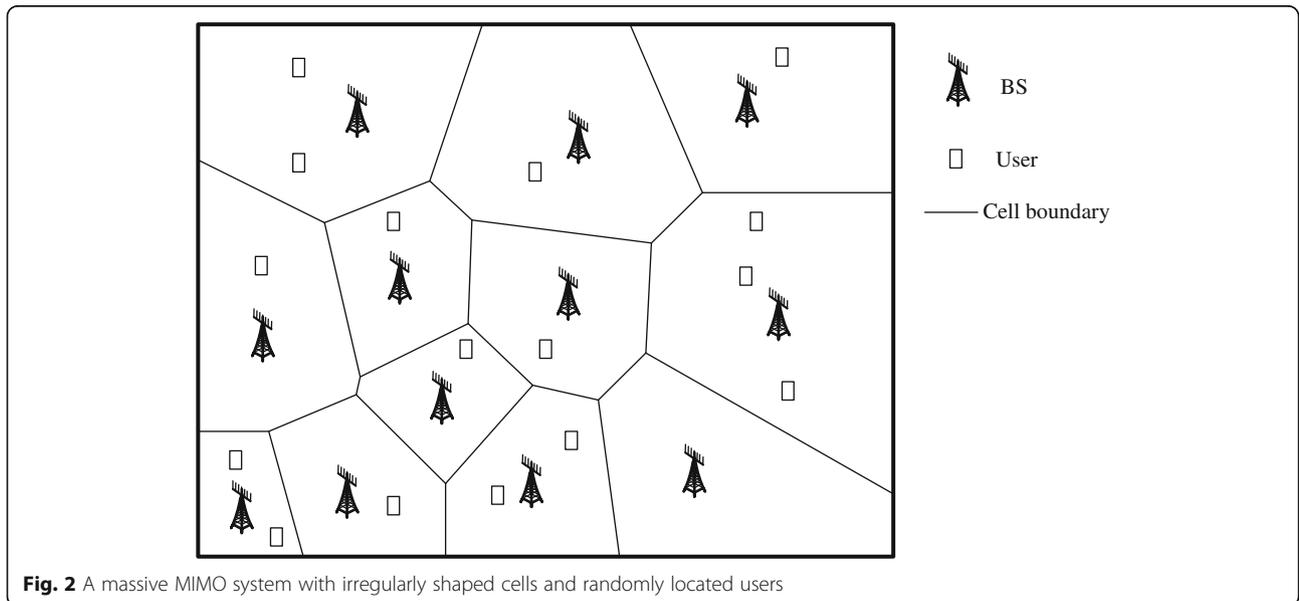
Fig. 2 A massive MIMO system with irregularly shaped cells and randomly located users

where D is the interval of antenna spacing at the BS, and λ is the signal wavelength.

As an approximation, we consider a ring of radius r_s comprising many scatters around the user. In this case [16], θ_{ljk} obeys a uniform distribution on $(\theta_{ljk}^{\min}, \theta_{ljk}^{\max})$, where $\theta_{ljk}^{\min} < \theta_{ljk}^{\max} \in [0, \pi)$, and

$$\theta_{ljk}^{\min} = \arctan\left(\frac{[z_{jk}^u]_2 - [z_l^b]_2}{[z_{jk}^u]_1 - [z_l^b]_1}\right) - \arcsin\left(\frac{r_s}{\left\|z_{jk}^u - z_l^b\right\|_2}\right) \tag{3}$$

$$\theta_{ljk}^{\max} = \arctan\left(\frac{[z_{jk}^u]_2 - [z_l^b]_2}{[z_{jk}^u]_1 - [z_l^b]_1}\right) + \arcsin\left(\frac{r_s}{\left\|z_{jk}^u - z_l^b\right\|_2}\right) \tag{4}$$

The channel covariance matrix is $\mathbf{R}_{ljk} = \mathbb{E}\{\boldsymbol{h}_{ljk}\boldsymbol{h}_{ljk}^H\}$ $= \frac{1}{F}\beta_{ljk}\sum_{f=1}^{F}\int_{\theta_{ljk}^{\min}}^{\theta_{ljk}^{\max}} a(\theta_{ljk}^{(f)})[a(\theta_{ljk}^{(f)})]^H d\theta_{ljk}^{(f)}$, where $\mathbb{E}\{\cdot\}$ denotes the expectation.

3 Analysis of different utility functions

The uplink transmission time-frequency resources are divided into frames consisting of T_c seconds and B_c Hz; thus, each frame contains $Q = T_cB_c$ transmission symbols (P pilot symbols and $Q-P$ data symbols, pilot symbols and data symbols are transmitted in different time slots). Assume channels between BSs and users have constant channel response within a frame but vary between frames. Therefore, the number of available orthogonal pilot sequences is P.

Assume that the coalition game divides all users into P sub-coalitions ($N > P$). P pilots are allocated to the P sub-coalitions respectively. That is, users in the same sub-coalition use the same pilot, users in different sub-coalitions use different pilots, and P pilots are mutually orthogonal. The analysis of utility function must be carried out under a certain coalition structure, so we first give the coalition structure definition in definition 1.

Definition 1 (Coalition structure). User set $\mathcal{N}$ is divided into P disjoint sub-coalitions $\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_P$, and $\cup_{n=1}^P \mathcal{S}_n = \mathcal{N}$, $\cap_{n=1}^P S_n = \varnothing$. Coalition structure $\mathcal{S}$ is the coalitional division state, i.e., $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_P\}$.

3.1 Analysis of MSE-CE

MSE-CE means the mean square error of channel estimation. During the uplink pilot transmission, all users transmit pilots at the same time, thus the received pilot signal at the BS l is

$$\boldsymbol{y}_l^{\text{pilot}} = \sum_{j=1}^{L}\sum_{k=1}^{K_j} \sqrt{t_{jk}}\; \boldsymbol{h}_{ljk}\boldsymbol{w}_{jk}^H + \boldsymbol{N}_l^{\text{pilot}} \tag{5}$$

where $\boldsymbol{w}_{jk} \in \mathbb{C}^{P \times 1}$ is the pilot sequence used by user (j, k), and $\mathbb{E}\{\boldsymbol{w}_{jk}^H\boldsymbol{w}_{jk}\} = 1$. t_{jk} is the pilot transmit power of user (j, k). In order to avoid the near-far issue in the uplink, here we consider the transmit power control and let $t_{jk} = T/\beta_{jjk}$, where T is a design parameter of transmit power. $\boldsymbol{N}_l^{\text{pilot}} \in \mathbb{C}^{M \times P}$ is the equivalent additive white Gaussian noise (AWGN) at the receiver with independent and identically distributed (i.i.d.) elements of $\mathcal{CN}(0, \sigma^2)$, where σ^2 is the variance of each element in noise .

After decorrelation and power normalization of $\boldsymbol{y}_l^{\text{pilot}}$, BS l can obtain the channel observation of $\boldsymbol{h}_{ljk}$, which is

$$\bar{\bar{\boldsymbol{h}}}_{ljk} = \frac{\boldsymbol{y}_l^{\text{pilot}}}{\sqrt{t_{jk}}}\boldsymbol{w}_{jk} = \boldsymbol{h}_{ljk} + \sum_{\substack{(m, n) \in \Lambda_{\mathcal{S}}(j, k) \\ (m, n) \neq (j, k)}} \frac{\sqrt{t_{mn}}}{\sqrt{t_{jk}}}\boldsymbol{h}_{lmn} + \frac{1}{\sqrt{t_{jk}}}\boldsymbol{N}_l^{\text{pilot}}\boldsymbol{w}_{jk} \tag{6}$$

where $\Lambda_{\mathcal{S}}(j, k)$ is one of the sub-coalitions in coalition structure $\mathcal{S}$, and user (j, k) belongs to the sub-coalition $\Lambda_{\mathcal{S}}(j, k)$.

According to the definition of minimum mean square error (MMSE) estimation and expressions (14), (15) in reference [14], the MMSE estimation of channel $\boldsymbol{h}_{ljk}$ at BS l, which is based on the channel observation $\bar{\bar{\boldsymbol{h}}}_{ljk}$, is given by

$$\hat{\boldsymbol{h}}_{ljk} = \mathbf{R}_{ljk}\mathbf{C}_{ljk}^{-1}\bar{\bar{\boldsymbol{h}}}_{ljk} \tag{7}$$

where $\mathbf{C}_{ljk} = \sum_{(m,n) \in \Lambda_{\mathcal{S}}(j, k)} \frac{t_{mn}}{t_{jk}}\mathbf{R}_{lmn} + \frac{\sigma^2}{t_{jk}}\mathbf{I}_M$ is the covariance matrix of $\bar{\bar{\boldsymbol{h}}}_{ljk}$.

Similar to expression (7), we can obtain the channel estimation of $\boldsymbol{h}_{jjk}$ (the channel between user (j, k) and its own BS j), which is

$$\begin{aligned}\hat{\boldsymbol{h}}_{jjk} &= \mathbf{R}_{jjk}\mathbf{C}_{jjk}^{-1}\bar{\bar{\boldsymbol{h}}}_{jjk} \\ &= \mathbf{R}_{jjk}\mathbf{C}_{jjk}^{-1}\left(\boldsymbol{h}_{jjk} + \sum_{\substack{(m, n) \in \Lambda_{\mathcal{S}}(j, k) \\ (m, n) \neq (j, k)}} \frac{\sqrt{t_{mn}}}{\sqrt{t_{jk}}}\boldsymbol{h}_{jmn} + \frac{1}{\sqrt{t_{jk}}}\boldsymbol{N}_j^{\text{pilot}}\boldsymbol{w}_{jk}\right)\end{aligned} \tag{8}$$

with covariance $\mathbf{R}_{\hat{\boldsymbol{h}}_{jjk}} = \mathbf{R}_{jjk}\mathbf{C}_{jjk}^{-1}\mathbf{R}_{jjk}$. Thus, we can determine that the channel estimation error is $\tilde{\boldsymbol{h}}_{jjk} = \boldsymbol{h}_{jjk} - \hat{\boldsymbol{h}}_{jjk}$. From the orthogonality principle of MMSE

estimation [14, 18], channel estimation error $\tilde{\boldsymbol{h}}_{jjk}$ is independent of $\hat{\boldsymbol{h}}_{jjk}$ and the covariance of $\tilde{\boldsymbol{h}}_{jjk}$ is $\mathbf{R}_{\tilde{h}_{jjk}}$ $=\mathbf{R}_{jjk}-\mathbf{R}_{\hat{h}_{jjk}}=\mathbf{R}_{jjk}-\mathbf{R}_{jjk}\mathbf{C}_{jjk}^{-1}\mathbf{R}_{jjk}$. Note that $\hat{\boldsymbol{h}}_{jjk}$ and $\mathbf{R}_{\tilde{h}_{jjk}}$ are also mean and covariance of $\boldsymbol{h}_{jjk}$ conditioned on $\overline{\overline{\boldsymbol{h}}}_{jjk}$, respectively [18].

The mean square error of channel estimation (MSE-CE) $\hat{\boldsymbol{h}}_{jjk}$ means the sum of squares estimation errors for each element of $\boldsymbol{h}_{jjk}$ (you can refer to expression (18) in [14]); so the MSE-CE of $\hat{\boldsymbol{h}}_{jjk}$ is   MSE-CE$_{jk}$

$$= \varepsilon_{\tilde{h}_{jjk}}^{\mathrm{MSE\text{-}CE}} = \mathbb{E}\{(\tilde{\boldsymbol{h}}_{jjk})^H \tilde{\boldsymbol{h}}_{jjk}\} = \mathbb{E}\{\mathrm{tr}\{\tilde{\boldsymbol{h}}_{jjk}\tilde{\boldsymbol{h}}_{jjk}^H\}\} = \mathrm{tr}\{\mathbf{R}_{\tilde{h}_{jjk}}\}\ .$$

We define the sum MSE-CE of all users as the sum of the estimation mean square error (MSE) of the channels between each user and its own BS. Thus, the average MSE-CE per user is

$$\mathrm{MSE\text{-}CE}_{\mathrm{ave}} = \frac{1}{N}\sum_{j=1}^{L}\sum_{k=1}^{K_j} \mathrm{tr}\left\{\mathbf{R}_{\tilde{h}_{jjk}}\right\} \tag{9}$$

3.2 Analysis of MSE-SD and received SINR

MSE-SD means the mean square error of signal detection. During the uplink data transmission, all users send their data to all BSs at the same time, and the received signal at BS j is

$$\begin{aligned}
\boldsymbol{y}_j^{\mathrm{data}} &= \sum_{m=1}^{L}\sum_{n=1}^{K_m} \sqrt{t_{mn}^d}\ \boldsymbol{h}_{jmn}x_{mn} + \boldsymbol{n}_j^{\mathrm{data}} \\
&= \sum_{m=1}^{L}\sum_{n=1}^{K_m} \boldsymbol{g}_{jmn}\ x_{mn} + \boldsymbol{n}_j^{\mathrm{data}}
\end{aligned} \tag{10}$$

where $x_{mn}\in\mathbb{C}^{1\times 1}$ represents the symbol transmitted from user (m,n), $\mathbb{E}\{|x_{mn}|^2\}=1$, $\boldsymbol{y}_j^{\mathrm{data}}=[y_{j1},\cdots,y_{jM}]^T\in \mathbb{C}^{M\times 1}$, $\boldsymbol{n}_j^{\mathrm{data}}$ is AWGN, and $\boldsymbol{n}_j^{\mathrm{data}}\sim\mathcal{CN}(0,\sigma^2\mathbf{I}_M)$, $\boldsymbol{g}_{jmn}$ $=\sqrt{t_{mn}^d}\boldsymbol{h}_{jmn}$.

Let $\boldsymbol{H}_{jm}=[\boldsymbol{h}_{jm1},\boldsymbol{h}_{jm2},\cdots,\boldsymbol{h}_{jmK_m}]\in\mathbb{C}^{M\times K_m}$, and $\boldsymbol{T}_m^d$ $= \mathrm{diag}(\sqrt{t_{m1}^d},\cdots,\sqrt{t_{mK_m}^d})$, $\boldsymbol{x}_m=[x_{m1},\cdots,x_{mK_m}]^T$, $\boldsymbol{G}_{jm}=$ $\boldsymbol{H}_{jm}\boldsymbol{T}_m^d=[\boldsymbol{g}_{jm1},\cdots,\boldsymbol{g}_{jmK_m}]$; thus, the received signal at BS j can be rewritten as $\boldsymbol{y}_j^{\mathrm{data}}=\sum_{m=1}^L \boldsymbol{H}_{jm}\ \boldsymbol{T}_m^d\boldsymbol{x}_m + \boldsymbol{n}_j^{\mathrm{data}}$ $=\sum_{m=1}^L \boldsymbol{G}_{jm}\boldsymbol{x}_m + \boldsymbol{n}_j^{\mathrm{data}}$.

To distinguish the symbol of each user, the BS j selects the combination matrix $\boldsymbol{W}_{jj}^{\mathrm{scheme}}\in\mathbb{C}^{M\times K_j}$, which is directly multiplied with the received signal, and the

processed data of all K_j users in the jth cell can be written as

$$\hat{\boldsymbol{x}}_j = \left(\boldsymbol{T}_j^d\right)^{-1}\left(\boldsymbol{W}_{jj}^{\mathrm{scheme}}\right)^H \boldsymbol{y}_j^{\mathrm{data}} \tag{11}$$

Considering the maximum ratio combining (MRC), zero-forcing combining (ZFC), and MMSE schemes of the received signal $\boldsymbol{y}_j^{\mathrm{data}}$, thus

$$\boldsymbol{W}_{jj}^{\mathrm{scheme}} = \begin{cases} \hat{\boldsymbol{H}}_{jj}\boldsymbol{D}_{jj}\ , & \text{for MRC} \\ \hat{\boldsymbol{H}}_{jj}\left(\hat{\boldsymbol{H}}_{jj}^H\hat{\boldsymbol{H}}_{jj}\right)^{-1}, & \text{for ZRC} \\ \hat{\boldsymbol{H}}_{jj}\left(\hat{\boldsymbol{H}}_{jj}^H\hat{\boldsymbol{H}}_{jj}+\sigma^2\left(\boldsymbol{T}_j^d\right)^{-2}\right)^{-1}, & \text{for MMSE} \end{cases} \tag{12}$$

where $\hat{\boldsymbol{H}}_{jj}=[\hat{\boldsymbol{h}}_{jj1},\hat{\boldsymbol{h}}_{jj2},\cdots,\hat{\boldsymbol{h}}_{jjK_j}]\in\mathbb{C}^{M\times K_j}$, and $\boldsymbol{T}_j^d= \mathrm{diag}(\sqrt{t_{j1}^d},\cdots,\sqrt{t_{jK_j}^d})$, $\boldsymbol{D}_{jj}= \mathrm{diag}(\frac{1}{\delta_{jj1}},\cdots,\frac{1}{\delta_{jjK_j}})$ is the diagonal matrix parameter that is set for $\mathbb{E}\{(\boldsymbol{w}_{jjk}^{MRC})^H\boldsymbol{h}_{jjk}\}$ $=1$, and $\delta_{jjk}=\mathrm{tr}\{\mathbf{R}_{\hat{h}_{jjk}}\}$.

Theorem 1 *For a given coalition structure $\mathcal{S}$, the mean square error of signal detection (MSE-SD) for user (j,k) at BS j for MRC, ZFC, and MMSE schemes are given by the following expressions* (13), (14) *and* (15) *respectively,*

$$\begin{aligned}
\mathrm{MSE\text{-}SD}_{jk}^{\mathrm{mrc}} &= \frac{1}{\delta_{jjk}^2}\mathrm{sum}\left\{\mathbf{R}_{\tilde{h}_{jjk}}\circ\mathbf{R}_{\hat{h}_{jjk}}^T\right\} + \sum_{(l,m)\notin\Lambda(j,k)}\frac{t_{lm}^d}{t_{jk}^d}\frac{1}{\delta_{jjk}^2}\mathrm{sum}\left\{\mathbf{R}_{jlm}\circ\mathbf{R}_{\hat{h}_{jjk}}^T\right\} \\
&+ \sum_{(l,m)\in\Lambda(j,k)}\frac{t_{lm}^d}{t_{jk}^d}\frac{1}{\delta_{jjk}^2}\{\frac{t_{lm}^d}{t_{jk}^d}\left|\mathrm{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\mathbf{R}_{\hat{h}_{jjk}}^T\right\}\right|^2 \\
&+\mathrm{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\mathbf{R}_{\hat{h}_{jjk}}^T\right\}\} + \frac{\sigma^2}{t_{jk}^d\delta_{jjk}}-1
\end{aligned} \tag{13}$$

$$\begin{aligned}
\mathrm{MSE\text{-}SD}_{jk}^{\mathrm{zfc}} &= \sum_{(l,m)\notin\Lambda(j,k)}\frac{t_{lm}^d}{t_{jk}^d}\mathrm{sum}\left\{\mathbf{R}_{jlm}\circ\boldsymbol{Q}_1^T\right\} \\
&+ \sum_{(l,m)\in\Lambda(j,k)}\frac{t_{lm}^d}{t_{jk}^d}\{\frac{t_{lm}^d}{t_{jk}^d}\left|\mathrm{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\boldsymbol{Q}_2^T\right\}\right|^2 \\
&+\mathrm{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\boldsymbol{Q}_1^T\right\}\} + \frac{\sigma^2}{t_{jk}^d}\cdot[\boldsymbol{Q}_3]_{k,k}-1
\end{aligned} \tag{14}$$

$$\begin{aligned}
\mathrm{MSE\text{-}SD}_{jk}^{\mathrm{mmse}} &= \sum_{(l,m)\notin\Lambda(j,k)}\frac{t_{lm}^d}{t_{jk}^d}\mathrm{sum}\left\{\mathbf{R}_{jlm}\circ\boldsymbol{Q}_5^T\right\} \\
&+ \sum_{(l,m)\in\Lambda(j,k)}\frac{t_{lm}^d}{t_{jk}^d}\{\frac{t_{lm}^d}{t_{jk}^d}\left|\mathrm{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\boldsymbol{Q}_6^T\right\}\right|^2 \\
&+\mathrm{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\boldsymbol{Q}_5^T\right\}\} + \frac{\sigma^2}{t_{jk}^d}\cdot[\boldsymbol{Q}_7]_{k,k}-2[\boldsymbol{Q}_4]_{k,k}-1
\end{aligned} \tag{15}$$

where $\boldsymbol{Q}_1 = \mathbb{E}_{\boldsymbol{h}}\{\boldsymbol{w}_{jjk}^{zfc}(\boldsymbol{w}_{jjk}^{zfc})^H\}$, $\boldsymbol{Q}_2 = \mathbb{E}_{\boldsymbol{h}}\{\hat{\boldsymbol{h}}_{jjk}(\boldsymbol{w}_{jjk}^{zfc})^H\}$, $\boldsymbol{Q}_3$ $= \mathbb{E}_{\boldsymbol{h}}\{(\hat{\boldsymbol{H}}_{jj}^H\hat{\boldsymbol{H}}_{jj})^{-1}\}$, $\boldsymbol{Q}_4 = \mathbb{E}_{\boldsymbol{h}}\{\sigma^2(\hat{\boldsymbol{H}}_{jj}^H\hat{\boldsymbol{H}}_{jj}+\sigma^2(\boldsymbol{T}_j^d)^{-2})^{-1}\}$,

$Q_5 = \mathbb{E}_h\{w_{jjk}^{\text{mmse}}(w_{jjk}^{\text{mmse}})^H\}$, $Q_6 = \mathbb{E}_h\{\hat{h}_{jjk}(w_{jjk}^{\text{mmse}})^H\}$, $Q_7 = \mathbb{E}_h\{(W_{jj}^{\text{mmse}})^H W_{jj}^{\text{mmse}}\}$, w_{jjk}^{zfc} is the kth column of the matrix $W_{jj}^{\text{zfc}} = \hat{H}_{jj}(\hat{H}_{jj}^H \hat{H}_{jj})^{-1}$, and w_{jjk}^{mmse} is the kth column of the matrix $W_{jj}^{\text{mmse}} = \hat{H}_{jj}(\hat{H}_{jj}^H \hat{H}_{jj} + \sigma^2(T_j^d)^{-2})^{-1}$.

For a given coalition structure $\mathcal{S}$, the uplink received SINR for user (j, k) at BS j for MRC, ZFC, and MMSE schemes are given by the expressions (17), (18) and (19) respectively.

Proof The proofs of Theorem 1 for MRC, ZFC, and MMSE schemes are given in Appendices 1, 2, and 3, respectively.

3.3 Analysis of spectrum efficiency

Based on previous theoretical analyses, according to Shannon capacity and referring to expression (11) in [19], the uplink spectrum efficiency (SE) of user (j,k) is

$$\text{SE}_{jk}^{\text{scheme}} = \left(1 - \frac{P}{Q}\right) \cdot \log_2\left(1 + \text{SINR}_{jk}^{\text{scheme}}\right) \qquad (16)$$

where the $\text{SINR}_{jk}^{\text{scheme}}$ for MRC, ZFC, and MMSE schemes are given by expressions (17), (18), and (19), respectively.

4 Pilot allocation scheme based on a coalition game

As mentioned in Section 3, we use the idea of partition form [20] in the coalition game, and all users will be divided into P sub-coalitions ($N > P$). P pilots are allocated to the P sub-coalitions respectively. That is, users are players in the coalition game. Users in the same sub-coalition use the same pilot and interact with each other in channel estimation. Users in different sub-coalitions use different pilots, and they do not affect each other in channel estimation. Therefore, there are mutually influencing and mutually constraining relationships between users, which coincides with the idea of the coalition game. For this reason, we introduce the idea of the coalition game into pilot allocation. The definition of the coalition game is given as follows.

Definition 2 (Coalition game). The coalition game is $\mathcal{G} = \langle \mathcal{N}, \{u_{j,k}\}_{(j,k)\in\mathcal{N}} \rangle$, where $\mathcal{N}$ is the set of game players; that is to say, users are players in the coalition game. $u_{j,k}$ is the utility function of player (j, k) (user (j, k)), and it is also called the payoff function in game theory.

In this coalition game, each player must belong to a sub-coalition, and the definition of coalition structure is given in definition 1.

The purpose of the coalition game is to optimize utility function, so different performance metrics can be used as the utility function in the coalition game. In this paper, we use the MSE-CE ($u_{j,k}(\mathcal{S}) = \text{MSE-CE}_{jk}(\mathcal{S})$), MSE-SD ($u_{j,k}(\mathcal{S}) = \text{MSE-SD}_{jk}^{\text{scheme}}(\mathcal{S})$), received SINR ($u_{j,k}(\mathcal{S}) = \text{SINR}_{jk}^{\text{scheme}}(\mathcal{S})$), and SE ($u_{j,k}(\mathcal{S}) = \text{SE}_{jk}^{\text{scheme}}(\mathcal{S})$) as the utility functions in the coalition game when we want to satisfy the corresponding system performance. The optimization of a certain utility function is also the minimization of pilot contamination corresponding to this performance.

According to the partition form definition in the coalition game theory [20], we need to describe three elements: (1) the coalition structure, (2) the adjustment principle of the coalition structure, and (3) the final stability of the coalition structure.

During each adjustment of the coalition structure, one user leaves its sub-coalition and joins another sub-coalition. This adjustment reduces interference of the original sub-coalition of the user and increases the interference of the newly joined sub-coalition.

The definition of coalition structure is given in definition 1. To benefit each user, the coalition structure needs to continuously adjust until final stability. For a certain adjustment, assume player (j, k) leaves its sub-coalition $\Lambda_{\mathcal{S}}(j, k)$ and joins another sub-coalition $\mathcal{S}_i$, where $\mathcal{S}_i \in \{\{\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_P\} \setminus \Lambda_{\mathcal{S}}(j, k)\}$. That is, the entire coalition structure changes from the original $\mathcal{S}$ into a new structure $\mathcal{S}^\Phi$. We describe this adjustment as $\mathcal{S} \xrightarrow{(j,k)} \mathcal{S}^\Phi$. According to the individual stability concept in the coalition game [21], the adjustment of a user (j, k) leaving its sub-coalition $\Lambda_{\mathcal{S}}(j, k)$ and joining another sub-coalition $\mathcal{S}_i$ is permissible if this adjustment can

$$\text{SINR}_{jk}^{\text{mrc}} = \cfrac{1}{\sum\limits_{(l,m)\notin\Lambda(j,k)}^{L} \frac{t_{lm}^d}{t_{jk}^d}\frac{1}{\delta_{jjk}^2}\text{sum}\left\{\mathbf{R}_{jlm}\circ\mathbf{R}_{\hat{h}_{jjk}}^T\right\} + \sum\limits_{(l,m)\in\Lambda(j,k)}^{L} \frac{t_{lm}^d}{t_{jk}^d}\frac{1}{\delta_{jjk}^2}\left\{\frac{t_{lm}^d}{t_{jk}^d}\left|\text{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\mathbf{R}_{\hat{h}_{jjk}}^T\right\}\right|^2 + \text{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\mathbf{R}_{\hat{h}_{jjk}}^T\right\}\right\} + \frac{\sigma^2}{t_{jk}^d\delta_{jjk}} - 1}$$

$$(17)$$

$$\text{SINR}_{jk}^{\text{zfc}} = \frac{1}{\sum\limits_{(l,m)\notin\Lambda(j,k)}^{L} \frac{t_{lm}^d}{t_{jk}^d}\text{sum}\{\mathbf{R}_{jlm}\circ\boldsymbol{Q}_1^T\} + \sum\limits_{(l,m)\in\Lambda(j,k)}^{L} \frac{t_{lm}^d}{t_{jk}^d}\left\{\frac{t_{lm}^d}{t_{jk}^d}\left|\text{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\boldsymbol{Q}_2^T\right\}\right|^2 + \text{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\boldsymbol{Q}_1^T\right\}\right\} + \frac{\sigma^2}{t_{jk}^d}\cdot[\boldsymbol{Q}_3]_{k,k}-1}$$

$$(18)$$

$$\text{SINR}_{jk}^{\text{mmse}} = \frac{1-[\boldsymbol{Q}_4]_{k,k}}{\sum\limits_{(l,m)\notin\Lambda(j,k)}^{L} \frac{t_{lm}^d}{t_{jk}^d}\text{sum}\{\mathbf{R}_{jlm}\circ\boldsymbol{Q}_5^T\} + \sum\limits_{(l,m)\in\Lambda(j,k)}^{L} \frac{t_{lm}^d}{t_{jk}^d}\left\{\frac{t_{lm}^d}{t_{jk}^d}\left|\text{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\boldsymbol{Q}_6^T\right\}\right|^2 + \text{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\boldsymbol{Q}_5^T\right\}\right\} + \frac{\sigma^2}{t_{jk}^d}\cdot[\boldsymbol{Q}_7]_{k,k}-1+[\boldsymbol{Q}_4]_{k,k}}$$

$$(19)$$

strictly improve its utility function and does not reduce the average utility function of all users. Therefore, we have definition 3.

Definition 3 (Principle of adjustment). *The adjustment $\mathcal{S} \xrightarrow{(j,k)} \mathcal{S}^\Phi$ is permissible if $u_{j,k}(\mathcal{S}^\Phi)$ is better than $u_{j,k}(\mathcal{S})$ and $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S}^\Phi)$ is not worse than $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S})$.*

In definition 3, that $u_{j,k}(\mathcal{S}^\Phi)$ is better than $u_{j,k}(\mathcal{S})$ means $u_{j,k}(\mathcal{S}^\Phi) < u_{j,k}(\mathcal{S})$ for MSE-CE and MSE-SD, and $u_{j,k}(\mathcal{S}^\Phi) > u_{j,k}(\mathcal{S})$ for SINR and SE. That $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S}^\Phi)$ is not worse than $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S})$ means $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S}^\Phi) \le \frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S})$ for MSE-CE and MSE-SD, and $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S}^\Phi) \ge \frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S})$ for SINR and SE.

From definition 3, we can see that any user that wants to join another sub-coalition should be beneficial to itself without damaging the average interest of all users. Thus, the average utility function will be optimal at the final stability.

This adjustment comes to an end as defined in definition 4, which is similar to the definition of final stability in [21].

Definition 4 (Final stability). *The whole coalition structure $\mathcal{S}$ is stable if there is no permission of adjustment $\mathcal{S} \xrightarrow{(j,k)} \mathcal{S}^\Phi$ for all users $(j,k)\in\mathcal{N}$ and all sub-coalitions, or $\alpha \ge \tau$, or the average utility function ($\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S})$) is not worse than U for a given performance index U (where $U \ge U_{\min}$ for MSE-CE and MSE-SD, $U \le U_{\min}$ for SINR and SE).*

In definition 4, assume the average utility function is equal to $U_{\min}$ when there is no permission of adjustment $\mathcal{S} \xrightarrow{(j,k)} \mathcal{S}^\Phi$ for all users $(j,k)\in\mathcal{N}$ and all sub-coalitions. That is to say, $U_{\min}$ is the global optimal value of average utility function. In addition, for the given performance index of the average utility function, $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S})$ is not worse than U means $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S}^\Phi) \le U$ for MSE-CE and MSE-SD, and $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S}^\Phi) \ge U$ for SINR and SE.

According to the discussions above, the coalition formation algorithm is given in algorithm 1. In definition 4 and algorithm 1, α is the current number of searches and τ is the limitation of total number of searches, which is used to prevent the coalition formation algorithm from falling into an endless loop.

We can see from algorithm 1 that each adjustment of the coalition structure further optimizes the average utility function. Taking the utility function MSE-CE as an example, the purpose of every searching in algorithm 1 is to reduce the average MSE-CE, and every adjustment of the coalition structure can reduce the average MSE-CE. According to the conditions of final stability in definition 4, algorithm 1 is convergent. Therefore, when the condition of final stability is that $\frac{1}{N}\sum\limits_{(l,m)\in\mathcal{N}} u_{l.m}(\mathcal{S})$ is not worse than U for a given performance index U (where $U \ge U_{\min}$ for MSE-CE), this algorithm can search continuously until the performance U is satisfied, it is a state of finite convergence. Therefore, this pilot allocation scheme can provide performance guarantee.

Algorithm 1 Coalition formation algorithm

Initialize: random coalition structure $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_P\}$, and $\alpha = 0$

   Repeat

      1. For all users $(j, k) \in \mathcal{N}$, do cyclic search

      2. For all $\mathcal{S}_i \in \{\{\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_P\} \setminus \Lambda_{\mathcal{S}}(j, k)\}$ ($i = 1, 2, \cdots, P$), do search

        Ask for the permission of adjustment $\mathcal{S} \xrightarrow{(j,k)} \mathcal{S}^{\Phi}$, i.e., judge $u_{j,k}(\mathcal{S}^{\Phi})$ is better than $u_{j,k}(\mathcal{S})$ and $\frac{1}{N} \sum_{(l,m) \in \mathcal{N}} u_{l,m}(\mathcal{S}^{\Phi})$ is not worse than $\frac{1}{N} \sum_{(l,m) \in \mathcal{N}} u_{l,m}(\mathcal{S})$.

        If adjustment $\mathcal{S} \xrightarrow{(j,k)} \mathcal{S}^{\Phi}$ is permissible, (j, k) leaves its sub-coalition and joins sub-coalition $\mathcal{S}_i$, update the coalition structure ($\mathcal{S} = \mathcal{S}^{\Phi}$).

        Else if adjustment $\mathcal{S} \xrightarrow{(j,k)} \mathcal{S}^{\Phi}$ is not permissible, coalition structure remains unchanged ($\mathcal{S} = \mathcal{S}$), end if.

      End (Corresponding to 2)

     Increase the number of searches, i.e., $\alpha = \alpha + 1$.

   End (Corresponding to 1)

   Return to **repeat**, until there is no permission of adjustment $\mathcal{S} \xrightarrow{(j,k)} \mathcal{S}^{\Phi}$, or $\alpha \geq \tau$, or $\frac{1}{N} \sum_{(l,m) \in \mathcal{N}} u_{l,m}(\mathcal{S})$ is not worse than U.

We go on to discuss the implementation method of the proposed pilot allocation scheme. The utility function calculations require the distance of each user to each BS. If algorithm 1 operates at BSs, all BSs should cooperate and exchange many messages. Therefore, both the calculations of utility functions and the processes of algorithm 1 can be set at the upper layer of BSs, i.e., completed by the base station controller (BSC). Each BS sends its own location and user locations to BSC. BSC calculates distances, utility functions, and the coalition formation algorithm, and then BSC sends the results of the coalition formation algorithm (i.e., the results of pilot allocation) to each user through BS forwarding. This can greatly reduce the complexity of the whole system. Although the computational complexity of this scheme is still somewhat high, with the general improvement of computing power, processing delays of BSC can be reduced. In addition, BSC needs to send the results of pilot allocation to each user; this will increase the signaling delays of the system. However, this is similar to the case that the receiver transmits the channel estimation back to the sender, only the transmission delay between BSs and BSC is added; however, the transmission delay of this kind of optical fiber line is very small, so we think this signal transmission can be hopefully achievable.

Stated thus, the proposed scheme can be hopefully applied in actual situations.

5 Simulation results and discussions

In this section, simulation results are provided to illustrate the system performance of the proposed pilot allocation scheme based on the coalition game. All experiments are Monte Carlo simulations (based on the results of theoretical analysis of different utility functions) under correlated Rayleigh fading channels. Then, we compare the proposed pilot allocation scheme with the pilot allocation scheme based on the greedy approach in [15] and the random coalition scheme (i.e., randomized pilot allocation) in terms of different performance and computational complexity.

The greedy approach in [15] assumes that there are P users per cell and there are P orthogonal pilots. The greedy approach selects one user from each cell to form a user set during each search, all users are grouped into P sets by P searches. Users in the same set use the same pilot, and users in the different sets use different pilots. The principle of user selection is to minimize the sum mean square error of channel estimation for the users using the same pilot.

The greedy approach in [15] assumes the number of users per cell is equal, and it cannot be used in the case of different number of users per cell, so we have to make modifications to facilitate the comparison. When simulating this greedy approach, $\eta = \lceil \frac{N}{P} \rceil$ (η represents the average number of users in each set) users are selected to form a set during each search, regardless of whether the selected users are in different cells. In addition, in all the following simulations, this greedy approach uses the MMSE combining scheme. Taking the MSE-CE as an example, the performance comparison of the greedy approach before and after the modification is shown in Fig. 3 (when using parameters in Table 1). We can see that the values of MSE-CE are very close before and after modification. As to other performance (like MSE-SD, SINR, SE), the results are similar according to our simulations; thus, we omit these results to save space. So, we can compare our scheme with the modified greedy approach, and the performance improvement can be seen in Figs. 4, 5, 6, 7, 8, 9, 10, 11, and 12. The greedy approach in Figs. 4, 5, 6, 7, 8, 9, 10, 11, and 12 is referring to the greedy approach after modification.

Simulation parameters are given in Table 1. We use a fixed area of 2000 m × 2000 m; BSs are subject to a uniform distribution, and the users are subject to a uniform distribution; and the distributions of BSs and users are independent. Each user selects and accesses its nearest BS. About the simulation parameters in Table 1, we refer to the simulation parameters in references [7, 16, 17]. For the setting of the ring radius of scatters, we consider a non-line-of-sight (non-LOS) scenario with 50 scattering paths as in references [7, 16, 17].

Fig. 3 Comparison of the greedy approach before and after the modification
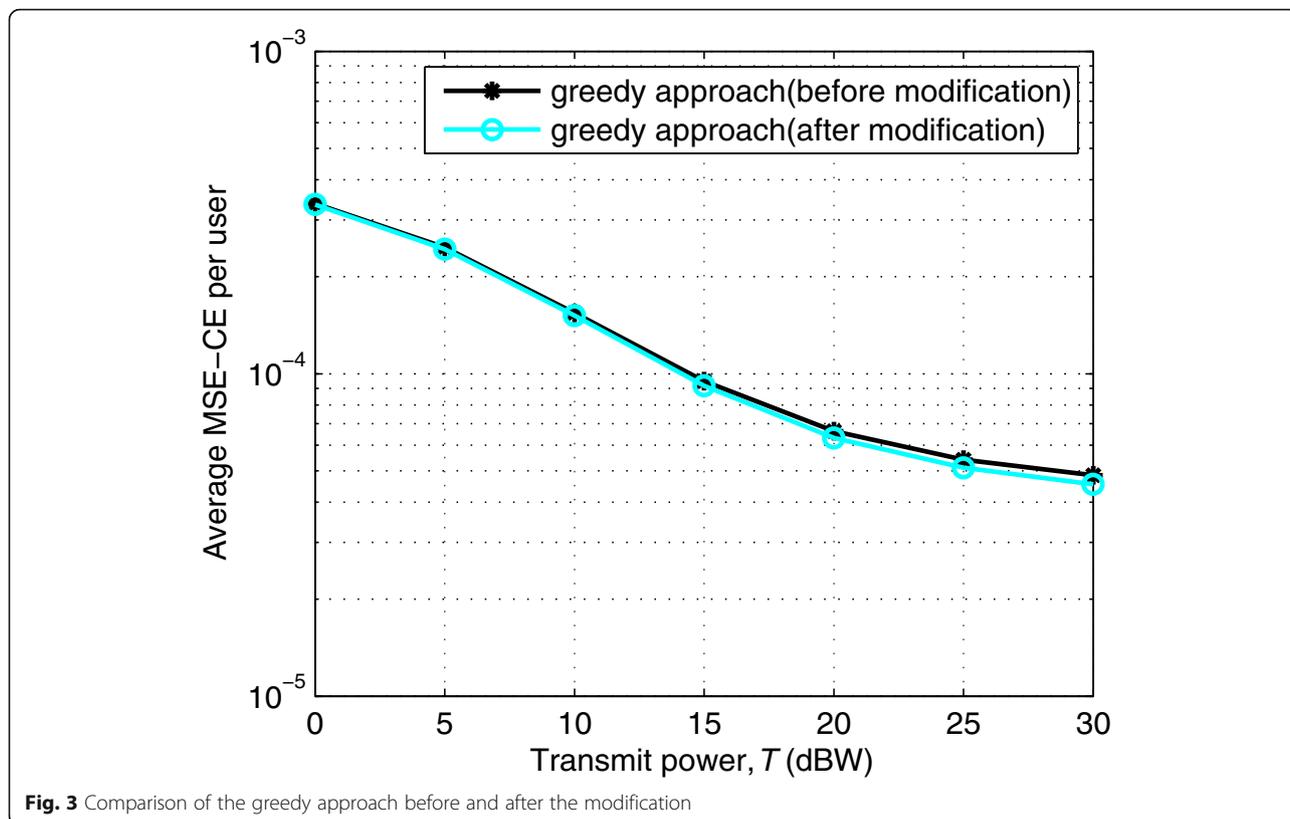
In the following simulations, for all coalition game pilot allocation schemes, we show the optimal results of the coalition formation algorithm, i.e., the condition of the final stability is the no permission of adjustment $\mathcal{S} \stackrel{(j,k)}{\rightarrow} \mathcal{S}^{\Phi}$. However, we know that the proposed pilot allocation scheme based on the coalition game can provide a certain performance guarantee through a user cyclic

Table 1 Simulation parameters

Parameter	Symbol	Value
Number of cells	L	4
Number of users	N	12
Path-loss exponent	v	2.5
Number of antennas	M	100
Number of paths	F	50
Number of symbols in each frame	Q	15
Ring radius of scatters around users	r_s	100 m
Antenna spacing at BS	D	$\lambda/2$
Channel coefficient	δ	10
Standard deviation of noise power	σ^2	10 W
Wave length	λ	0.1 m

search. For a given performance (average utility function) index U, the proposed scheme can search continuously until the performance index U is satisfied. Taking the utility function of MSE-CE as an example, Fig. 4 shows the relationship between the average MSE-CE and the number of searches (α) for a certain location of BSs and users. We can see that, with the increase of α, the average MSE decreases, and when $\alpha \geq 22$, the average MSE reaches its minimum value. To satisfy the average MSE-CE performance index of 10^{-4}, we should set the total number of search limitation τ to greater than or equal to 14.

Figure 5 shows the average MSE-CE per user versus the transmit power T with $P = 4$. The unit of transmission power T is the watt in decibels (i.e., dBW). The "coalition game, MSE-CE" means the pilot allocation scheme based on the coalition game when utility function is MSE-CE. The "greedy approach" means the greedy approach in [15]. We can see that the coalition game for MSE-CE can achieve lower average MSE than the greedy approach and random coalition. Although the object function of the greedy approach in [15] is also MSE-CE, the greedy approach is worse than the coalition game for MSE-CE, because the greedy approach in [15] is a one-time search for users, but the coalition
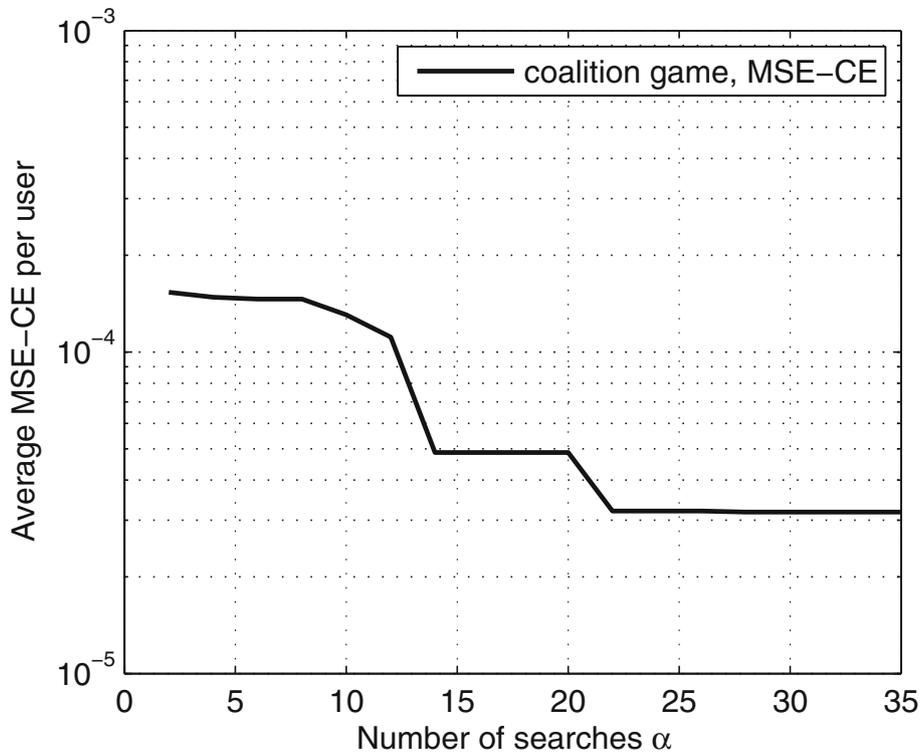
Fig. 4 Average MSE-CE versus the number of searches (α) with $P = 4$ and $T = 20$ dBW
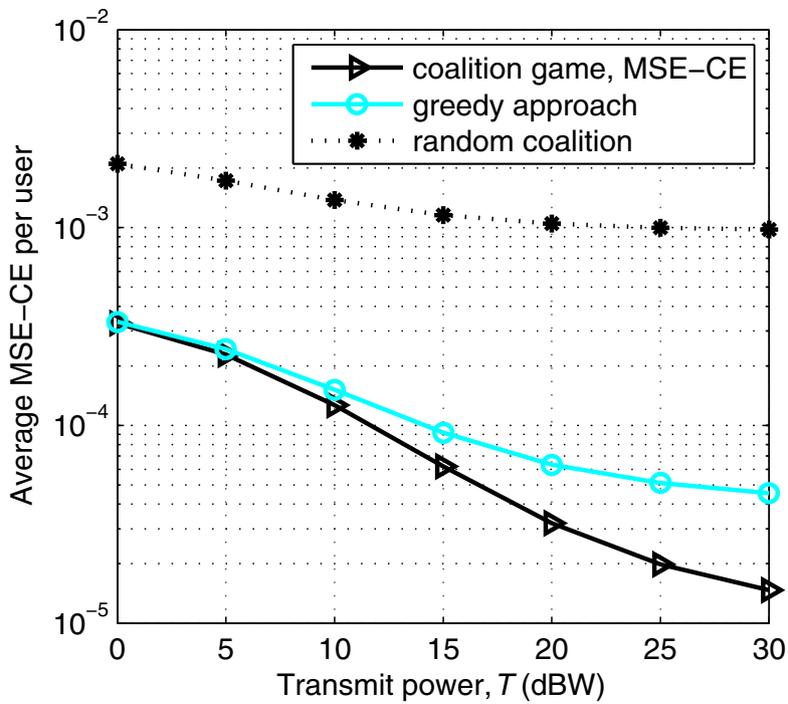
Fig. 5 Average MSE-CE versus transmit power T with $P = 4$
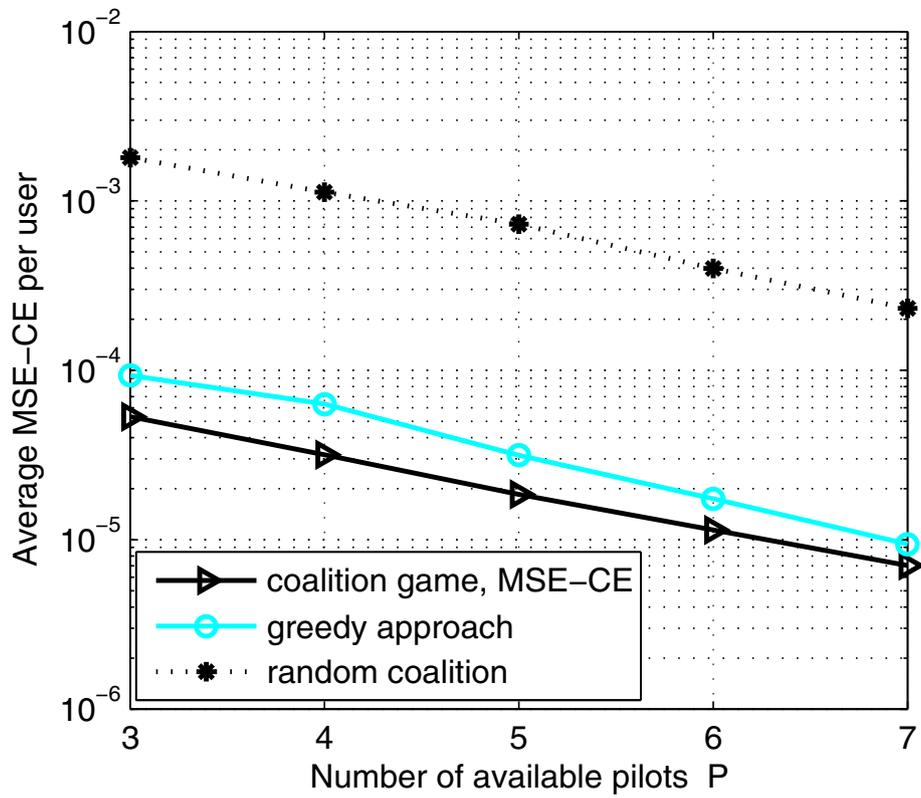
Fig. 6 Average MSE-CE versus the number of available pilots *P* with *T* = 20 dBW
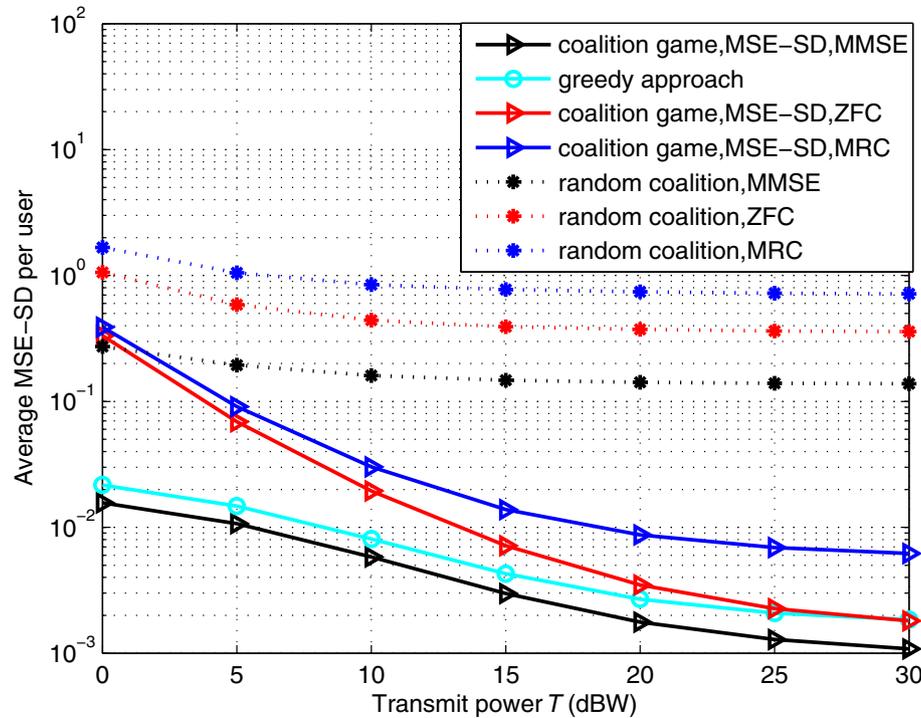
Fig. 7 Average MSE-SD versus transmit power *T* with *P* = 4
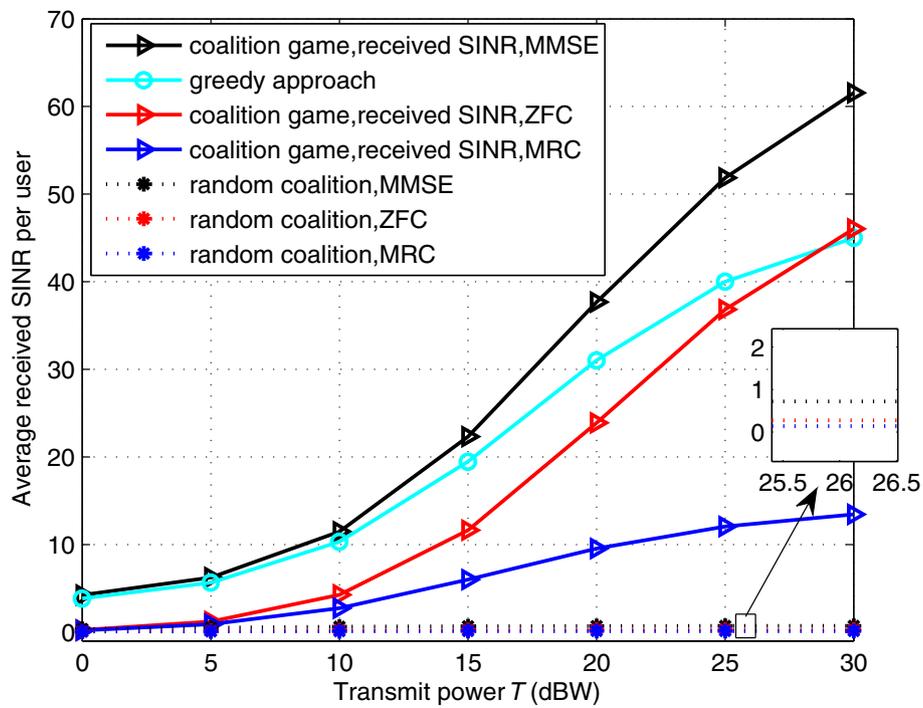
Fig. 8 The average received SINR versus transmit power T with $P = 4$
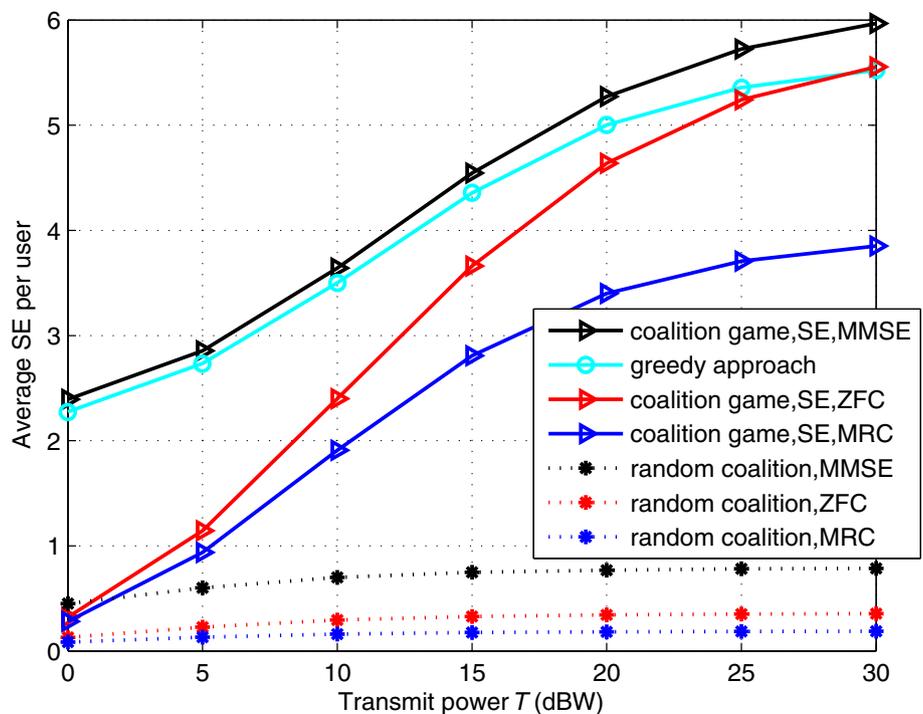
Fig. 9 Average spectrum efficiency (SE) versus transmit power T with $P = 4$
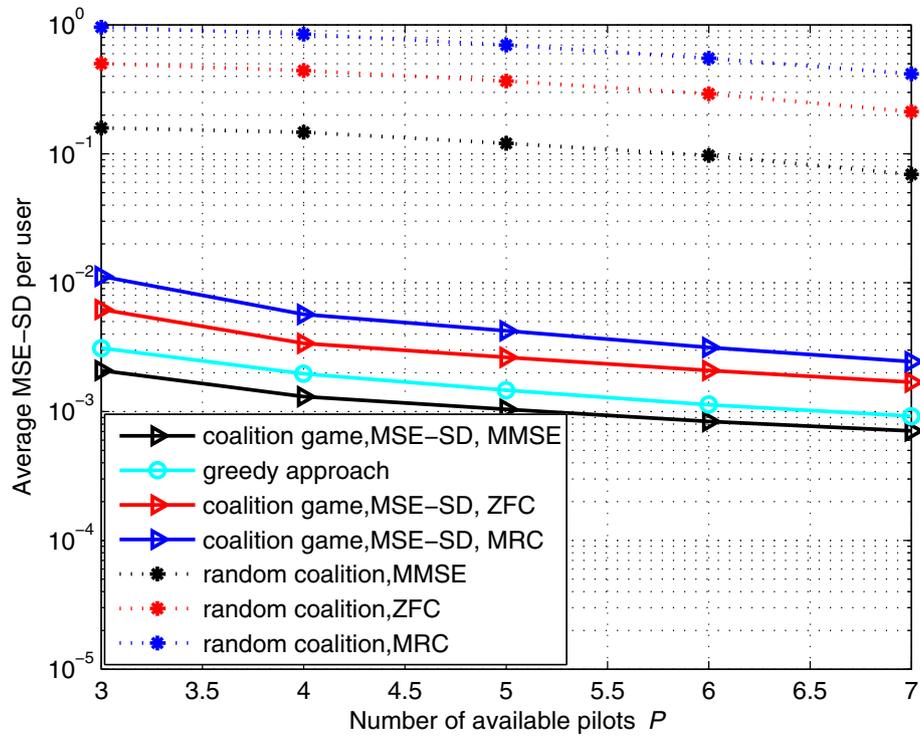
Fig. 10 Average MSE-SD versus the number of pilots with $T = 20$ dBW
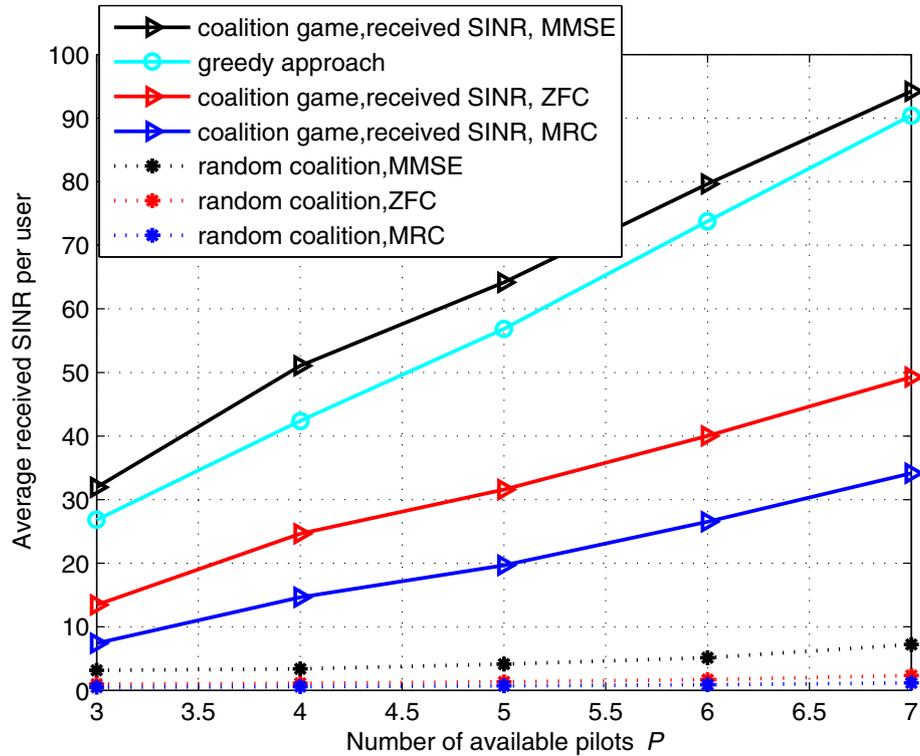
Fig. 11 Average received SINR versus the number of pilots with $T = 20$ dBW

Fig. 12 Average SE versus the number of pilots with $T = 20$ dBW
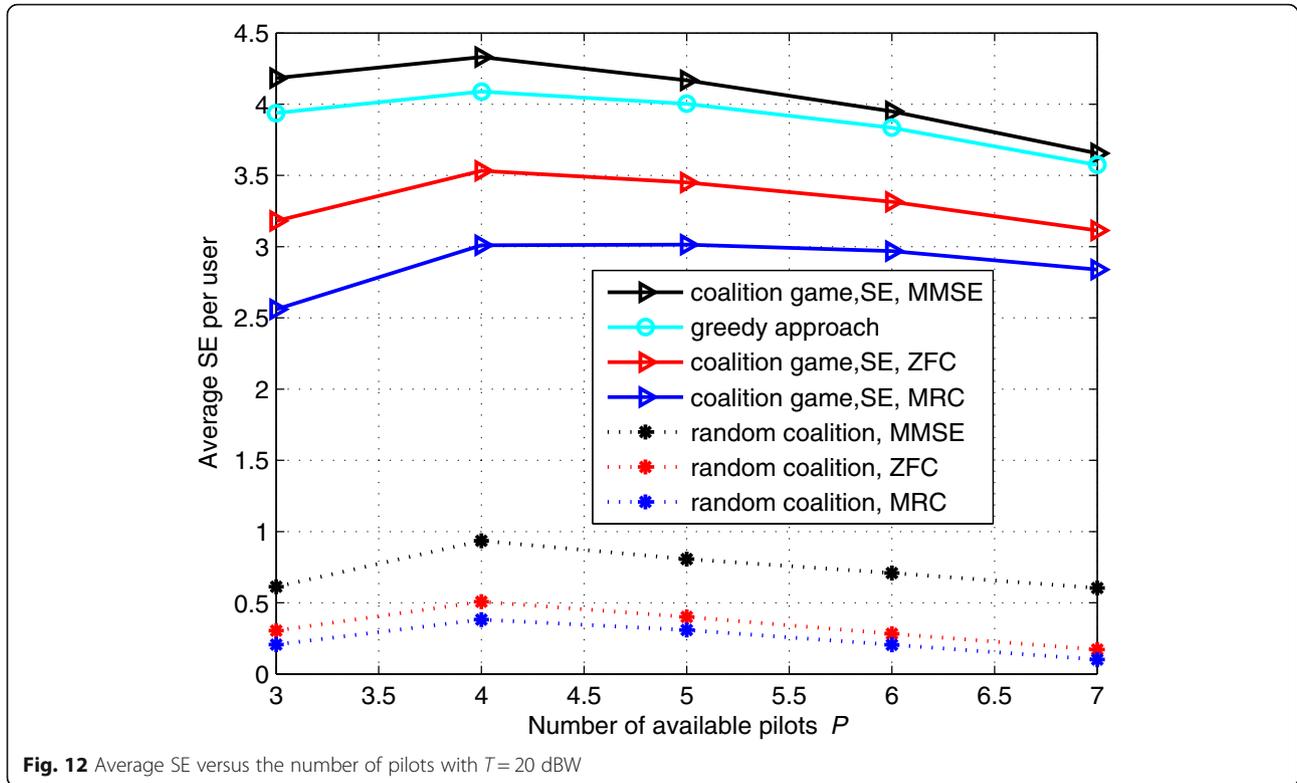
game for MSE-CE is a cyclic search for users. In addition, the greater the transmit power, the more obvious the performance improvements.

Figure 6 shows the average MSE-CE per user versus the number of available pilots P with $T = 20$ dBW. We can see that the coalition game for MSE-CE can achieve lower average MSE-CE than the greedy approach. The performance of random alliance is the worst because of random allocation of pilots. When the number of pilots becomes larger, the average MSE-CE of all schemes becomes smaller. This is because the larger the number of pilots, the larger the number of sub-coalitions and the smaller the interference between users.

Figure 7 shows the average MSE-SD per user versus the transmit power T with $P = 4$. The "coalition game, MSE-SD, MMSE," "coalition game, MSE-SD, ZFC," and "coalition game, MSE-SD, MRC" means the pilot allocation scheme based on the coalition game when utility function is MSE-SD with the MMSE scheme, when utility function is MSE-SD with the ZFC scheme, and when utility function is MSE-SD with the MRC scheme, respectively. It can be seen that the coalition game for MSE-SD with MMSE is better than the coalition game for MSE-SD with ZFC, and the coalition game for MSE-SD with ZFC is better than the coalition game for MSE-SD with MRC. That is because the MRC scheme cannot reduce interference and noise, and the ZFC

scheme can reduce interference, but cannot reduce noise, whereas the MMSE scheme can reduce both interference and noise. When the transmit power is small, the performance of the greedy approach is slightly worse than that of the coalition game for MSE-SD with MMSE and better than that of the coalition game for MSE-SD with ZFC and the coalition game for MSE-SD with MRC. Because the purpose of the greedy approach in [15] is to minimize the sum of normalized MSE-CE of all users, its performance is better than the coalition game for MSE-SD with ZFC and the coalition game for MSE-SD with MRC. The greedy approach is worse than the coalition game for MSE-SD with MMSE because the greedy approach is a one-time search for users, but the coalition game for MSE-SD with MMSE is a cyclic search for users. With the increase of transmit power T, the advantage of the coalition game for MSE-SD with MMSE relative to the greedy approach is increasingly obvious. In addition, for each combining scheme (MMSE, ZFC, or MRC), the coalition game is better than random coalition.

Figure 8 shows the average received SINR per user versus the transmit power T with $P = 4$. The utility functions of all coalition game schemes in Fig. 8 are received SINR. When transmit power T becomes larger, the average received SINR of the coalition game schemes and greedy approach become larger. But random coalition

schemes change little with the increase of T, that is because the interference between all users is increasing when transmit power T becomes larger. From Fig. 8, we can see that the coalition game for received SINR with MMSE is better than the greedy approach, and other results are similar to Fig. 7.

Figure 9 shows the average SE per user versus transmit power T with $P = 4$. The utility functions of all coalition game schemes in Fig. 9 are SE. We can see that when transmit power T becomes larger, the average SE of the coalition game schemes and greedy approach become larger. When comparing these different schemes, the results are similar to Figs. 7 and 8.

Figure 10 shows the average MSE-SD per user versus the number of available pilots P with $T = 20$ dBW. The utility functions of all coalition game schemes in Fig. 10 are MSE-SD. It can be seen that the coalition game for MSE-SD with MMSE is better than the coalition game for MSE-SD with ZFC, and the coalition game for MSE-SD with ZFC is better than the coalition game for MSE-SD with MRC. All these results are the same as Fig. 7. In addition, the larger the number of available pilots, the smaller the average MSE-SD of all schemes. This is because when the number of pilots becomes larger, fewer users use the same pilot; thus, the interference between users can be reduced, and the system performance can be improved.

Figure 11 shows the average received SINR per user versus the number of available pilots with $T = 20$ dBW. The utility functions of all coalition game schemes in Fig. 11 are received SINR. We can get the similar results as in Fig. 8.

Figure 12 shows the average SE per user versus the number of available pilots with $T = 20$ dBW. The utility functions of all coalition game schemes in Fig. 12 are SE. The results in Fig. 12 are different from those

in Fig. 9. When the number of pilots becomes larger, the average SE of all schemes first becomes larger and then becomes smaller. That is because when the number of pilots is very small, the number of sub-coalitions is small and the interference between users is large, so the SE is small. On the other hand, when the number of pilots becomes very large, the number of data symbols in each frame becomes small and the coefficient in front of the log in expression (16) becomes small, so the SE is also small. That is to say, there exists the optimal number of pilots to maximize SE. We can see in Fig. 12 that the optimal number of pilots is four for all schemes.

Table 2 shows the complexity (in terms of the multiplication times) during each search for different schemes. It is assumed that calculating the inversion of an M-dimensional matrix needs about M^3 multiplications, $\eta = \lceil \frac{N}{P} \rceil$ represents the average number of users in each sub-coalition, and $\xi = \lceil \frac{N}{L} \rceil$ represents the average number of users in each cell, where $\lceil * \rceil$ means the round up number of *. The expressions in Table 2 are obtained according to the multiplication times of Eqs. (9),(13)~(19), and we have assumed that $\mathbf{R}_{jlm}$ is known according to large-scale channel fading and the locations of users and BSs and that $\mathbf{R}_{\hat{h}_{jlm}}$, $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$, $\mathbf{Q}_4$, $\mathbf{Q}_5$, $\mathbf{Q}_6$, and $\mathbf{Q}_7$ are obtained through statistical average of channel estimation and calculated once during each search. $o(NM^2)$ represents higher order infinitesimal of NM^2.

Table 3 shows the average number of searches (i.e., $\overline{\alpha}$) for each scheme, and also means the calculation times of average utility function in coalition formation algorithm. The values in Table 3 are obtained through Monte Carlo simulations, and they are the average number of searches for 10,000 user and BS locations. For all

Table 2 The multiplication times during each search for different schemes

	Different combining schemes		
	MMSE	ZFC	MRC
Coalition game for different utility function			
MSE-SD	$(3N + 2\eta)NM^3 +$ $[(N + \eta)\xi^2 + 3\xi^2 + 2(N + \eta)]NM^2 + o(NM^2)$	$(3N + 2\eta)NM^3 +$ $[(N + \eta)\xi^2 + \xi^2 + 2(N + \eta)]NM^2 + o(NM^2)$	$(3N + 2\eta)NM^3 +$ $[N + \eta + 1]NM^2 + o(NM^2)$
Received SINR	$(3N + 2\eta)NM^3 +$ $[(N + \eta)\xi^2 + 3\xi^2 + 2(N + \eta)]NM^2 + o(NM^2)$	$(3N + 2\eta)NM^3 +$ $[(N + \eta)\xi^2 + \xi^2 + 2(N + \eta)]NM^2 + o(NM^2)$	$(3N + 2\eta)NM^3 +$ $[N + \eta + 1]NM^2 + o(NM^2)$
SE	$(3N + 2\eta)NM^3 +$ $[(N + \eta)\xi^2 + 3\xi^2 + 2(N + \eta)]NM^2 + o(NM^2)$	$(3N + 2\eta)NM^3 +$ $[(N + \eta)\xi^2 + \xi^2 + 2(N + \eta)]NM^2 + o(NM^2)$	$(3N + 2\eta)NM^3 +$ $[N + \eta + 1]NM^2 + o(NM^2)$
MSE-CE	$3NM^3 + o(NM^2)$ (without combining schemes)		
Random coalition	0 (for all combining schemes)		
Greedy approach in [15]			
MSE-CE	$3NM^3 + o(NM^2)$ (for MMSE combining scheme)		

Table 3 The average number of searches ($\overline{a}$) for different schemes

	Different combining schemes		
	MMSE	ZFC	MRC
Coalition game for different utility function			
MSE-SD	36.1	35.2	34.2
Received SINR	36.5	35.8	34.2
SE	36.5	35.8	34.2
MSE-CE	24 (without combining schemes)		
Random coalition	0 (for all combining schemes)		
Greedy approach in [15]			
MSE-CE	12 (for MMSE combining scheme)		

coalition game pilot allocation schemes in Table 3, the condition of the final stability is the no permission of adjustment $\mathcal{S} \overset{(j,k)}{\rightarrow} \mathcal{S}^{\Phi}$.

In Table 3, the average number of search of random coalition is 0 because random coalition means randomized pilot allocation and there is no adjustment of the coalition structure. The average number of search of the greedy approach in [15] is 12 because the greedy approach is a one-time search for all users and the number of users is 12. The average number of search of the coalition game for different utility functions (MSE-SD, received SINR, or SE) is very close, and they are all close to 36, which means each user is searched about 3 times on average. The average number of searches of the coalition game for MSE-CE is 24, which means each user is searched about 2 times on average.

The total complexity of each scheme is the result of the multiplication of the corresponding items in Tables 2 and 3. Combining Tables 2 and 3, when comparing the total multiplication times, we can see that random coalition has the lowest complexity, the greedy approach is higher than random coalition, the coalition game for MSE-CE is higher than the greedy approach, and the coalition game for MSE-SD, received SINR, and SE are very close and has the highest complexity. The total multiplication times of the coalition game for different utility functions (MSE-SD, received SINR, or SE) are also very close, and they are higher than the coalition game for MSE-CE and the greedy approach, although their average search time does not differ much. That is because the multiplication times during each search of the coalition game for different utility functions (MSE-SD, received SINR, or SE) are larger than the coalition game for MSE-CE and the greedy approach. In a word, the total multiplication times of the coalition game for MSE-CE and the greedy approach is much higher than that of random coalition,

but lower than that of the coalition game for MSE-SD, received SINR, and SE.

6 Conclusions

A pilot allocation scheme based on the coalition game is proposed for TDD massive MIMO systems. This scheme can be suitable for any actual scene in which cells are irregularly shaped (BSs distribute randomly), numbers of users per cell are not equal (users distribute randomly), and channels are correlated fading channels, and this scheme can provide performance guarantee. The definition of the coalition structure is given, and different utility functions are analyzed. To benefit each user, the adjustment principle of the coalition structure is set, and the coalition formation algorithm is provided. According to the coalition formation algorithm, the pilot allocation scheme based on the coalition game (for different utility functions) is simulated and compared with the greedy approach in [15] and random coalition. And then, the computational complexity in terms of multiplication times for different pilot allocation schemes is compared. Simulation results show that when considering the performance comparison of different pilot allocation schemes, the results for different performance metrics (different utility functions) are similar. The coalition game for each utility function with MMSE achieves the best performance with high computational complexity. For each utility function and for each combining scheme (MMSE, ZFC, or MRC), the coalition game is better than random coalition. When comparing the performance of different combining schemes, with regard to the coalition game for each utility function, MMSE is better than ZFC and ZFC is better than MRC. The performance of the greedy approach in [15] is between the coalition game for each utility function with MMSE and the coalition game for each utility function with ZFC. In addition, we obtain the optimal number of pilots to maximize SE through simulations.

7 Appendix 1

According to expression (11), the mean square error of signal detection (MSE-SD) for user (j, k) is

$$
\begin{aligned}
\text{MSE-SD}_{jk}^{\text{scheme}} = \mathbb{E}\Bigg\{ & \Bigg| \frac{1}{\sqrt{t_{jk}^d}} \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \left(\sum_{m=1}^{L}\sum_{n=1}^{K_m} \sqrt{t_{mn}^d}\, \boldsymbol{h}_{jmn} x_{mn} + \boldsymbol{n}_j^{\text{data}}\right) \\
& -x_{jk}\Big|^2 \Bigg\} = \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \hat{\boldsymbol{h}}_{jjk} - 1 \right|^2 \right\} + \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \tilde{\boldsymbol{h}}_{jjk} \right|^2 \right\} \\
& + \sum_{l=1}^{L}\sum_{m=1}^{K_l} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \boldsymbol{h}_{jlm} \right|^2 \right\} - \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \boldsymbol{h}_{jjk} \right|^2 \right\} \\
& + \frac{1}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \boldsymbol{n}_j^{\text{data}} \right|^2 \right\}
\end{aligned}
$$

$$(20)$$

Then, according to reference ([19], Lemma 2), the received SINR for user (j, k) at BS j is

$$\text{SINR}_{jk}^{\text{scheme}} = \frac{\left| \mathbb{E}_{\boldsymbol{h}}\left\{ \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \boldsymbol{h}_{jjk} \right\} \right|^2}{\sum_{l=1}^{L}\sum_{m=1}^{K_l} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \boldsymbol{h}_{jlm} \right|^2 \right\} - \left| \mathbb{E}_{\boldsymbol{h}}\left\{ \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \boldsymbol{h}_{jjk} \right\} \right|^2 + \frac{1}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \boldsymbol{n}_j^{\text{data}} \right|^2 \right\}} \quad (21)$$

For the MRC scheme, we can obtain $\mathbb{E}_{\boldsymbol{h}}\{|(\boldsymbol{w}_{jjk}^{\text{mrc}})^H \boldsymbol{h}_{jjk} - 1|^2\}$ $= \mathbb{E}_{\boldsymbol{h}}\{|(\boldsymbol{w}_{jjk}^{\text{mrc}})^H \hat{\boldsymbol{h}}_{jjk} - 1|^2\} + \mathbb{E}_{\boldsymbol{h}}\{|(\boldsymbol{w}_{jjk}^{\text{mrc}})^H \tilde{\boldsymbol{h}}_{jjk}|^2\}$. That is the first and the second items of expression (20).

According to the definition of the MRC scheme, we obtain

$$\mathbb{E}_{\boldsymbol{h}}\left\{ \left(\boldsymbol{w}_{jjk}^{\text{mrc}}\right)^H \boldsymbol{h}_{jjk} \right\} = \mathbb{E}_{\boldsymbol{h}}\left\{ \left(\boldsymbol{w}_{jjk}^{\text{mrc}}\right)^H \hat{\boldsymbol{h}}_{jjk} \right\}$$
$$= \frac{1}{\delta_{jjk}} \mathbb{E}_{\boldsymbol{h}}\left\{ \hat{\boldsymbol{h}}_{jjk}^H \hat{\boldsymbol{h}}_{jjk} \right\} = 1 \quad (22)$$

Therefore, the first item of expression (20) is

$$\mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{mrc}}\right)^H \hat{\boldsymbol{h}}_{jjk} - 1 \right|^2 \right\} = 0. \quad (23)$$

When $M \to \infty$, $\mathbb{E}_{\boldsymbol{h}}\{|(\boldsymbol{w}_{jjk}^{\text{mrc}})^H \boldsymbol{h}_{jjk}|^2\} \approx |\mathbb{E}_{\boldsymbol{h}}\{(\boldsymbol{w}_{jjk}^{\text{mrc}})^H \boldsymbol{h}_{jjk}\}|^2$ $= 1$. This is the result of the fourth item of expression (20).

Because $\tilde{\boldsymbol{h}}_{jjk}$ is independent of $\boldsymbol{w}_{jjk}^{\text{wrc}}$, the second item of expression (20) is

$$\mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{mrc}}\right)^H \tilde{\boldsymbol{h}}_{jjk} \right|^2 \right\} = \frac{1}{\delta_{jjk}^2} \mathbb{E}_{\boldsymbol{h}}\left\{ \hat{\boldsymbol{h}}_{jjk}^H \tilde{\boldsymbol{h}}_{jjk} \tilde{\boldsymbol{h}}_{jjk}^H \hat{\boldsymbol{h}}_{jjk} \right\}$$
$$= \frac{1}{\delta_{jjk}^2} \text{sum}\left\{ \mathbf{R}_{\tilde{h}_{jjk}} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\} \quad (24)$$

where $\boldsymbol{A} \circ \boldsymbol{B}$ is the Hadamard product of $\boldsymbol{A}$ and $\boldsymbol{B}$.

For the third item of expression (20) and the first item of the denominator in expression (21), two cases (i.e., $(l, m) \notin \Lambda_{\mathcal{S}}(j, k)$ and $(l, m) \in \Lambda_{\mathcal{S}}(j, k)$) are discussed.

When $(l, m) \notin \Lambda_{\mathcal{S}}(j, k)$,

$$\sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{mrc}}\right)^H \boldsymbol{h}_{jlm} \right|^2 \right\}$$
$$= \sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \mathbb{E}_{\boldsymbol{h}}\left\{ \hat{\boldsymbol{h}}_{jjk}^H \boldsymbol{h}_{jlm} \boldsymbol{h}_{jlm}^H \hat{\boldsymbol{h}}_{jjk} \right\}$$
$$= \sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \text{sum}\left\{ \mathbf{R}_{jlm} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\}$$

$$(25)$$

where the second equation exists because of the independence of $\hat{\boldsymbol{h}}_{jjk}$ and $\boldsymbol{h}_{jlm}$ for $(l, m) \notin \Lambda_{\mathcal{S}}(j, k)$.

When $(l, m) \in \Lambda_{\mathcal{S}}(j, k)$,

$$\sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{mrc}}\right)^H \boldsymbol{h}_{jlm} \right|^2 \right\}$$
$$= \sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \left(\mathbb{E}_{\boldsymbol{h}}\left\{ \hat{\boldsymbol{h}}_{jjk}^H \hat{\boldsymbol{h}}_{jlm} \hat{\boldsymbol{h}}_{jlm}^H \hat{\boldsymbol{h}}_{jjk} \right\} + \mathbb{E}_{\boldsymbol{h}}\left\{ \hat{\boldsymbol{h}}_{jjk}^H \tilde{\boldsymbol{h}}_{jlm} \tilde{\boldsymbol{h}}_{jlm}^H \hat{\boldsymbol{h}}_{jjk} \right\} \right)$$

$$(26)$$

According to expression (7) and the analysis above, when $(l, m) \in \Lambda_{\mathcal{S}}(j, k)$, $\hat{\boldsymbol{h}}_{jlm} = \sqrt{\frac{t_{lm}^d}{t_{jk}^d}} \mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \hat{\boldsymbol{h}}_{jjk}$. Substituting it into expression (26), we can get (A-8) for $(l, m) \in \Lambda_{\mathcal{S}}(j, k)$

$$\mathbb{E}_{\boldsymbol{h}}\left\{ \hat{\boldsymbol{h}}_{jjk}^H \hat{\boldsymbol{h}}_{jlm} \hat{\boldsymbol{h}}_{jlm}^H \hat{\boldsymbol{h}}_{jjk} \right\} = \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \hat{\boldsymbol{h}}_{jjk}^H \mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \hat{\boldsymbol{h}}_{jjk} \hat{\boldsymbol{h}}_{jjk}^H \left(\mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \right)^H \hat{\boldsymbol{h}}_{jjk} \right\}$$
$$= \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \hat{\boldsymbol{h}}_{jjk}^H \mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \hat{\boldsymbol{h}}_{jjk} \right|^2 \right\}$$
$$\approx^{(a)} \frac{t_{lm}^d}{t_{jk}^d} \left| \mathbb{E}_{\boldsymbol{h}}\left\{ \hat{\boldsymbol{h}}_{jjk}^H \mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \hat{\boldsymbol{h}}_{jjk} \right\} \right|^2$$
$$= \frac{t_{lm}^d}{t_{jk}^d} \left| \text{sum}\left\{ \left(\mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \right) \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\} \right|^2$$

$$(27)$$

where equation (a) exists for $\mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1}$ is a Hermitian matrix, and $\hat{\boldsymbol{h}}_{jjk}^H \mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \hat{\boldsymbol{h}}_{jjk} = \text{sum}\{(\mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1}) \circ (\hat{\boldsymbol{h}}_{jjk} \hat{\boldsymbol{h}}_{jjk}^H)^T\}$ is a real number. When $M \to \infty$, $\mathbb{E}_{\boldsymbol{h}}\{|\hat{\boldsymbol{h}}_{jjk}^H \mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \hat{\boldsymbol{h}}_{jjk}|^2\}$ $\approx |\mathbb{E}_{\boldsymbol{h}}\{\hat{\boldsymbol{h}}_{jjk}^H \mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \hat{\boldsymbol{h}}_{jjk}\}|^2$. In addition, due to the mutual independence of $\hat{\boldsymbol{h}}_{jjk}$ and $\boldsymbol{h}_{jlm}$, $\mathbb{E}_{\boldsymbol{h}}\{\hat{\boldsymbol{h}}_{jjk}^H \tilde{\boldsymbol{h}}_{jlm} \tilde{\boldsymbol{h}}_{jlm}^H \hat{\boldsymbol{h}}_{jjk}\}$ $= \text{sum}\{\mathbf{R}_{\tilde{h}_{jlm}} \circ \mathbf{R}_{\hat{h}_{jjk}}^T\}$. Therefore, the third item of expression (20) (the first item of the denominator in expression (21)) is

$$\sum_{l=1}^{L}\sum_{m=1}^{K_l} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{scheme}}\right)^H \boldsymbol{h}_{jlm} \right|^2 \right\}$$
$$= \sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{mrc}}\right)^H \boldsymbol{h}_{jlm} \right|^2 \right\}$$
$$+ \sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}}\left\{ \left| \left(\boldsymbol{w}_{jjk}^{\text{mrc}}\right)^H \boldsymbol{h}_{jlm} \right|^2 \right\}$$
$$= \sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \text{sum}\left\{ \mathbf{R}_{jlm} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\}$$
$$+ \sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \left\{ \frac{t_{lm}^d}{t_{jk}^d} \left| \text{sum}\left\{ \left(\mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \right) \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\} \right|^2 \right.$$
$$\left. + \text{sum}\left\{ \mathbf{R}_{\tilde{h}_{jlm}} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\} \right\}$$

$$(28)$$

For the last item of expression (20) and the last item of the denominator in expression (21),

$$\frac{1}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}} \left\{ \left| \left(\boldsymbol{w}_{jjk}^{\mathrm{mrc}} \right)^H \boldsymbol{n}_j^{\mathrm{data}} \right|^2 \right\} = \frac{1}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \mathbb{E}_{\boldsymbol{h}} \left\{ \hat{\boldsymbol{h}}_{jjk}^H \boldsymbol{n}_j^{\mathrm{data}} \left(\boldsymbol{n}_j^{\mathrm{data}} \right)^H \hat{\boldsymbol{h}}_{jjk} \right\}$$

$$= \frac{\sigma^2}{t_{jk}^d \delta_{jjk}^2} \mathbb{E}_{\boldsymbol{h}} \left\{ \hat{\boldsymbol{h}}_{jjk}^H \hat{\boldsymbol{h}}_{jjk} \right\} = \frac{\sigma^2}{t_{jk}^d \delta_{jjk}} \tag{29}$$

Combining expressions (20), (23), (24), (28), and (29), the MSE-SD for user (j, k) at BS j is

$$\mathrm{MSE\text{-}SD}_{jk}^{\mathrm{mrc}} = \frac{1}{\delta_{jjk}^2} \mathrm{sum} \left\{ \mathbf{R}_{\tilde{h}_{jjk}} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\}$$

$$+ \sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \mathrm{sum} \left\{ \mathbf{R}_{jlm} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\}$$

$$+ \sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \left\{ \frac{t_{lm}^d}{t_{jk}^d} \left| \mathrm{sum} \left\{ \left(\mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \right) \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\} \right|^2 \right.$$

$$\left. + \mathrm{sum} \left\{ \mathbf{R}_{\tilde{h}_{jlm}} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\} \right\} + \frac{\sigma^2}{t_{jk}^d \delta_{jjk}} - 1 \tag{30}$$

Combining expressions (21), (22), (28), and (29), the received SINR for user (j, k) at BS j is

$$\mathrm{SINR}_{jk}^{\mathrm{mrc}} = \frac{1}{\displaystyle\sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \mathrm{sum} \left\{ \mathbf{R}_{jlm} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\}}$$

$$+ \sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \frac{1}{\delta_{jjk}^2} \left\{ \frac{t_{lm}^d}{t_{jk}^d} \left| \mathrm{sum} \left\{ \left(\mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \right) \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\} \right|^2 \right.$$

$$\left. + \mathrm{sum} \left\{ \mathbf{R}_{\tilde{h}_{jlm}} \circ \mathbf{R}_{\hat{h}_{jjk}}^T \right\} \right\} + \frac{\sigma^2}{t_{jk}^d \delta_{jjk}} - 1 \tag{31}$$

8 Appendix 2

We consider the ZFC scheme. For ZFC, the $\boldsymbol{w}_{jjk}^{\mathrm{zfc}}$ is the kth column of the matrix $\boldsymbol{W}_{jj}^{\mathrm{zfc}} = \hat{\boldsymbol{H}}_{jj} (\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj})^{-1}$, and thus, $(\boldsymbol{W}_{jj}^{\mathrm{zfc}})^H \hat{\boldsymbol{H}}_{jj} = \mathbf{I}_{K_j}$ and $(\boldsymbol{W}_{jj}^{\mathrm{zfc}})^H \boldsymbol{W}_{jj}^{\mathrm{zfc}} = (\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj})^{-1}$. Therefore

$$\mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \boldsymbol{h}_{jjk} \right\} = \mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{zfc} \right)^H \left(\hat{\boldsymbol{h}}_{jjk} + \tilde{\boldsymbol{h}}_{jjk} \right) \right\}$$

$$= \mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \hat{\boldsymbol{h}}_{jjk} \right\} = 1 \tag{32}$$

Therefore, the first item of expression (20) is

$$\mathbb{E}_{\boldsymbol{h}} \left\{ \left| \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \hat{\boldsymbol{h}}_{jjk} - 1 \right|^2 \right\} = 0 \tag{33}$$

Because $\tilde{\boldsymbol{h}}_{jjk}$ is independent of $\boldsymbol{w}_{jjk}^{\mathrm{zfc}}$, the second item of expression (20) is

$$\mathbb{E}_{\boldsymbol{h}} \left\{ \left| \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \tilde{\boldsymbol{h}}_{jjk} \right|^2 \right\} = \mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \tilde{\boldsymbol{h}}_{jjk} \tilde{\boldsymbol{h}}_{jjk}^H \boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right\}$$

$$= \mathrm{sum} \left\{ \mathbf{R}_{\tilde{h}_{jjk}} \circ \boldsymbol{Q}_1^T \right\} \tag{34}$$

Here, we assume that $\boldsymbol{Q}_1 = \mathbb{E}_{\boldsymbol{h}} \{ \boldsymbol{w}_{jjk}^{\mathrm{zfc}} (\boldsymbol{w}_{jjk}^{\mathrm{zfc}})^H \}$. The fourth item of expression (20) is

$$\mathbb{E}_{\boldsymbol{h}} \left\{ \left| \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \boldsymbol{h}_{jjk} \right|^2 \right\}$$

$$= \mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \hat{\boldsymbol{h}}_{jjk} \hat{\boldsymbol{h}}_{jjk}^H \boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right\} + \mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \tilde{\boldsymbol{h}}_{jjk} \tilde{\boldsymbol{h}}_{jjk}^H \boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right\}$$

$$= 1 + \mathrm{sum} \left\{ \mathbf{R}_{\tilde{h}_{jjk}} \circ \boldsymbol{Q}_1^T \right\} \tag{35}$$

Using a similar analysis method as MRC in Appendix 1, we can divide the third item of expression (20) (the first item of the denominator in expression (21)) into two cases.

When $(l, m) \notin \Lambda_{\mathcal{S}}(j, k)$,

$$\sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}} \left\{ \left| \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \boldsymbol{h}_{jlm} \right|^2 \right\}$$

$$= \sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \boldsymbol{h}_{jlm} \boldsymbol{h}_{jlm}^H \boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right\} \tag{36}$$

$$= \sum_{(l,m) \notin \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \mathrm{sum} \left\{ \mathbf{R}_{jlm} \circ \boldsymbol{Q}_1^T \right\}$$

When $(l, m) \in \Lambda_{\mathcal{S}}(j, k)$,

$$\sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_{\boldsymbol{h}} \left\{ \left| \left(\boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right)^H \boldsymbol{h}_{jlm} \right|^2 \right\}$$

$$= \sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \left(\mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{zfc} \right)^H \hat{\boldsymbol{h}}_{jlm} \hat{\boldsymbol{h}}_{jlm}^H \boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right\} \right.$$

$$\left. + \mathbb{E}_{\boldsymbol{h}} \left\{ \left(\boldsymbol{w}_{jjk}^{zfc} \right)^H \tilde{\boldsymbol{h}}_{jlm} \tilde{\boldsymbol{h}}_{jlm}^H \boldsymbol{w}_{jjk}^{\mathrm{zfc}} \right\} \right)$$

$$= \sum_{(l,m) \in \Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \left\{ \frac{t_{lm}^d}{t_{jk}^d} \left| \mathrm{sum} \left\{ \left(\mathbf{R}_{jlm} \mathbf{R}_{jjk}^{-1} \right) \circ \boldsymbol{Q}_2^T \right\} \right|^2 \right.$$

$$\left. + \mathrm{sum} \left\{ \mathbf{R}_{\tilde{h}_{jlm}} \circ \boldsymbol{Q}_1^T \right\} \right\} \tag{37}$$

where $\boldsymbol{Q}_2 = \mathbb{E}_{\boldsymbol{h}} \{ \hat{\boldsymbol{h}}_{jjk} (\boldsymbol{w}_{jjk}^{\mathrm{zfc}})^H \}$.

In addition, the last item of expression (20) (the last item of the denominator in expression (21)) is

$$\frac{1}{t_{jk}^d} \mathbb{E}_h\left\{\left|\left(\boldsymbol{w}_{jjk}^{zfc}\right)^H \boldsymbol{n}_j^{data}\right|^2\right\} = \frac{1}{t_{jk}^d} \mathbb{E}_h\left\{\left(\boldsymbol{w}_{jjk}^{zfc}\right)^H \boldsymbol{n}_j^{data}\left(\boldsymbol{n}_j^{data}\right)^H \boldsymbol{w}_{jjk}^{zfc}\right\}$$

$$= \frac{\sigma^2}{t_{jk}^d} \mathbb{E}_h\left\{\left(\boldsymbol{w}_{jjk}^{zfc}\right)^H \boldsymbol{w}_{jjk}^{zfc}\right\} = \frac{\sigma^2}{t_{jk}^d} \cdot [\boldsymbol{Q}_3]_{k,k} \tag{38}$$

where $\boldsymbol{Q}_3 = \mathbb{E}_h\{(\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj})^{-1}\}$ and $[\boldsymbol{Q}_3]_{k, k}$ means the kth row and kth column element of matrix $\boldsymbol{Q}_3$.

Combining expressions (20), (33), (34), (35), (36), (37), and (38), the MSE-SD for user (j, k) at BS j is

$$\text{MSE-SD}_{jk}^{zfc} = \sum_{(l,m)\notin\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \text{ sum}\left\{\mathbf{R}_{jlm}\circ\boldsymbol{Q}_1^T\right\}$$

$$+ \sum_{(l,m)\in\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d}\left\{\frac{t_{lm}^d}{t_{jk}^d}\left|\text{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\boldsymbol{Q}_2^T\right\}\right|^2\right.$$

$$\left.+\text{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\boldsymbol{Q}_1^T\right\}\right\} + \frac{\sigma^2}{t_{jk}^d} \cdot [\boldsymbol{Q}_3]_{k,k} - 1 \tag{39}$$

Combining expressions (21), (32), (36), (37) and (38), the received SINR for user (j, k) at BS j is

$$\text{SINR}_{jk}^{zfc} = \cfrac{1}{\begin{aligned}&\sum_{(l,m)\notin\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \text{sum}\left\{\mathbf{R}_{jlm}\circ\boldsymbol{Q}_1^T\right\}\\ &+ \sum_{(l,m)\in\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d}\left\{\frac{t_{lm}^d}{t_{jk}^d}\left|\text{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\boldsymbol{Q}_2^T\right\}\right|^2\right.\\ &\left.+\text{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\boldsymbol{Q}_1^T\right\}\right\} + \frac{\sigma^2}{t_{jk}^d} \cdot [\boldsymbol{Q}_3]_{k,k} - 1\end{aligned}} \tag{40}$$

9 Appendix 3

We consider the MMSE scheme, the $\boldsymbol{w}_{jjk}^{mmse}$ is the kth column of the matrix $\boldsymbol{W}_{jj}^{mmse} = \hat{\boldsymbol{H}}_{jj}(\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj} + \sigma^2(\boldsymbol{T}_j^d)^{-2})^{-1}$. Because $(\boldsymbol{W}_{jj}^{zfc})^H \hat{\boldsymbol{H}}_{jj} + \sigma^2(\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj} + \sigma^2(\boldsymbol{T}_j^d)^{-2})^{-1} = \mathbf{I}_{K_j}$, $(\boldsymbol{w}_{jjk}^{mmse})^H \hat{\boldsymbol{h}}_{jjk} = [(\boldsymbol{W}_{jj}^{zfc})^H \hat{\boldsymbol{H}}_{jj}]_{k,k} = 1 - [\sigma^2(\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj} + \sigma^2(\boldsymbol{T}_j^d)^{-2})^{-1}]_{k,k}$, thus

$$\mathbb{E}_h\left\{\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \hat{\boldsymbol{h}}_{jjk}\right\} = 1 - \mathbb{E}_h\left\{\left[\sigma^2\left(\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj} + \sigma^2\left(\boldsymbol{T}_j^d\right)^{-2}\right)^{-1}\right]_{k,k}\right\} = 1 - [\boldsymbol{Q}_4]_{k,k} \tag{41}$$

Therefore, the first item of expression (20) is

$$\mathbb{E}_h\left\{\left|\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \hat{\boldsymbol{h}}_{jjk} - 1\right|^2\right\} = \left|[\boldsymbol{Q}_4]_{k,k}\right|^2 \tag{42}$$

where $\boldsymbol{Q}_4 = \mathbb{E}_h\{\sigma^2(\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj} + \sigma^2(\boldsymbol{T}_j^d)^{-2})^{-1}\}$. Because $\tilde{\boldsymbol{h}}_{jjk}$ is independent of $\boldsymbol{w}_{jjk}^{zfc}$, the second item of expression (20) is

$$\mathbb{E}_h\left\{\left|\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \tilde{\boldsymbol{h}}_{jjk}\right|^2\right\} = \mathbb{E}_h\left\{\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \tilde{\boldsymbol{h}}_{jjk}\tilde{\boldsymbol{h}}_{jjk}^H \boldsymbol{w}_{jjk}^{mmse}\right\}$$

$$= \text{sum}\left\{\mathbf{R}_{\tilde{h}_{jjk}}\circ\boldsymbol{Q}_5^T\right\} \tag{43}$$

where $\boldsymbol{Q}_5 = \mathbb{E}_h\{\boldsymbol{w}_{jjk}^{mmse}(\boldsymbol{w}_{jjk}^{mmse})^H\}$. The fourth item of expression (20) is

$$\mathbb{E}_h\left\{\left|\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \boldsymbol{h}_{jjk}\right|^2\right\} = \mathbb{E}_h\left\{\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \hat{\boldsymbol{h}}_{jjk}\hat{\boldsymbol{h}}_{jjk}^H \boldsymbol{w}_{jjk}^{mmse}\right\}$$

$$+ \mathbb{E}_h\left\{\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \tilde{\boldsymbol{h}}_{jjk}\tilde{\boldsymbol{h}}_{jjk}^H \boldsymbol{w}_{jjk}^{mmse}\right\} \tag{44}$$

$$\approx \left|1 - [\boldsymbol{Q}_4]_{k,k}\right|^2 + \text{sum}\left\{\mathbf{R}_{\tilde{h}_{jjk}}\circ\boldsymbol{Q}_5^T\right\}$$

where the last approximate equality exists when $M \to \infty$.

Using a similar analysis method as MRC in Appendix 1, we can divide the third item of expression (20) (the first item of the denominator in expression (21)) into two cases. The results is

$$\sum_{l=1}^{L}\sum_{m=1}^{K_l} \frac{t_{lm}^d}{t_{jk}^d} \mathbb{E}_h\left\{\left|\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \boldsymbol{h}_{jlm}\right|^2\right\}$$

$$= \sum_{(l,m)\notin\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \text{ sum}\left\{\mathbf{R}_{jlm}\circ\boldsymbol{Q}_5^T\right\}$$

$$+ \sum_{(l,m)\in\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d}\left\{\frac{t_{lm}^d}{t_{jk}^d}\left|\text{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\boldsymbol{Q}_6^T\right\}\right|^2\right.$$

$$\left.+\text{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\boldsymbol{Q}_5^T\right\}\right\} \tag{45}$$

where $\boldsymbol{Q}_6 = \mathbb{E}_h\{\hat{\boldsymbol{h}}_{jjk}(\boldsymbol{w}_{jjk}^{mmse})^H\}$.

In addition, the last item of expression (20) (the last item of the denominator in expression (21)) is5

$$\frac{1}{t_{jk}^d} \mathbb{E}_h\left\{\left|\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \boldsymbol{n}_j^{data}\right|^2\right\}$$

$$= \frac{1}{t_{jk}^d} \mathbb{E}_h\left\{\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \boldsymbol{n}_j^{data}\left(\boldsymbol{n}_j^{data}\right)^H \boldsymbol{w}_{jjk}^{mmse}\right\} \tag{46}$$

$$= \frac{\sigma^2}{t_{jk}^d} \mathbb{E}_h\left\{\left(\boldsymbol{w}_{jjk}^{mmse}\right)^H \boldsymbol{w}_{jjk}^{mmse}\right\} = \frac{\sigma^2}{t_{jk}^d} \cdot [\boldsymbol{Q}_7]_{k,k}$$

where $\boldsymbol{Q}_7 = \mathbb{E}_h\{(\boldsymbol{W}_{jj}^{mmse})^H \boldsymbol{W}_{jj}^{mmse}\} = \mathbb{E}_h\{(\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj} + \sigma^2(\boldsymbol{T}_j^d)^{-2})^{-1} \hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj}(\hat{\boldsymbol{H}}_{jj}^H \hat{\boldsymbol{H}}_{jj} + \sigma^2(\boldsymbol{T}_j^d)^{-2})^{-1}\}$.

Combining expressions (20), (42), (43), (44), (45), and (46), the MSE-SD for user (j, k) at BS j is

$$\text{MSE-SD}_{jk}^{\text{mmse}} = \sum_{(l,m)\notin\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \text{sum}\{\mathbf{R}_{jlm}\circ\mathbf{Q}_5^T\}$$
$$+ \sum_{(l,m)\in\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d}\left\{\frac{t_{lm}^d}{t_{jk}^d}\left|\text{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\mathbf{Q}_6^T\right\}\right|^2\right.$$
$$\left.+\text{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\mathbf{Q}_5^T\right\}\right\} + \frac{\sigma^2}{t_{jk}^d}\cdot[\mathbf{Q}_7]_{k,k}-2[\mathbf{Q}_4]_{k,k}-1 \tag{47}$$

Combining expressions (21), (41), (45), and (46), the received SINR for user (j,k) at BS j is

$$\text{SINR}_{jk}^{\text{mmse}} = \frac{1-[\mathbf{Q}_4]_{k,k}}{I_{jk}^{\text{mmse}}} \tag{48}$$

where $I_{jk}^{\text{mmse}} = \sum_{(l,m)\notin\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d} \text{sum}\{\mathbf{R}_{jlm}\circ\mathbf{Q}_5^T\}$

$$+ \sum_{(l,m)\in\Lambda(j,k)} \frac{t_{lm}^d}{t_{jk}^d}\left\{\frac{t_{lm}^d}{t_{jk}^d}\left|\text{sum}\left\{\left(\mathbf{R}_{jlm}\mathbf{R}_{jjk}^{-1}\right)\circ\mathbf{Q}_6^T\right\}\right|^2\right.$$
$$\left.+\text{sum}\left\{\mathbf{R}_{\tilde{h}_{jlm}}\circ\mathbf{Q}_5^T\right\}\right\} \begin{array}{l} +\frac{\sigma^2}{t_{jk}^d}\cdot[\mathbf{Q}_7]_{k,k}-1+[\mathbf{Q}_4]_{k,k} \end{array}.$$

Abbreviations
AoA: Angle-of-arrival; AWGN: Additive white Gaussian noise; BS: Base station; BSC: Base station controller; CSI: Channel state information; MIMO: Multiple-input multiple-output; MMSE: Minimum mean square error; MRC: Maximum ratio combining; MSE: Mean square error; MSE-CE: Mean square error of channel estimation; MSE-SD: Mean square error of signal detection; SE: Spectrum efficiency; SINR: Signal-to-interference-plus-noise ratio; TDD: Time division duplex; ULA: Uniformly spaced linear array; ZFC: Zero-forcing combining

Availability of data and materials
Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Authors' contributions
HZ has contributed 70% of the work, and the rest is contributed by XD. Both authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Key Lab of Intelligent Computing and Signal Processing (Ministry of Education), Anhui University, Hefei 230601, China. [2]QuantumCTek Co., Ltd., Hefei 230088, China.

References
1. L. Lu, G.Y. Li, A.L. Swindlehurst, et al., An overview of massive MIMO: benefits and challenges. IEEE J. Sel. Top. Sign. Proces. **8**(5), 742–758 (2014).
2. Z. Gong, C. Li, F. Jiang, Pilot contamination mitigation strategies in massive MIMO systems. IET Commun. **11**(16), 2403–2409 (2017).
3. T.L. Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans. Wirel. Commun. **9**(11), 3590–3600 (2010).
4. A. Ashikhmin, T. Marzetta, *Pilot Contamination Precoding in Multi-Cell Large Scale Antenna Systems. Proceedings International Symposium on Information Theory (ISIT)* (2012), pp. 1137–1141.
5. M.H.C. Garcia, J. Luo, in *Proceedings of 2018 IEEE Wireless Communications and Networking Conference (WCNC)*. Time-shifted pilots multiplexed with uplink data and unequal power allocation (IEEE, Barcelona, 2018), pp. 1–6.
6. W. Chang, Y.K. Hua, S.F. Liao, Partial overlapped time-shifted pilots for massive MIMO systems. IEEE Commun. Lett. **21**(11), 2480–2483 (2017).
7. H. Echigo, T. Ohtsuki, W. Jiang, Y. Takatori, in *Proceedings of 2017 IEEE Global Communications Conference (GLOBECOM)*. Fair pilot assignment based on AOA and pathloss with location information in massive MIMO (IEEE, Singapore, 2017), pp. 1–6.
8. H. Wang, W. Zhang, Y. Liu, Q. Xu, P. Pan, On design of non-orthogonal pilot signals for a multi-cell massive MIMO system. IEEE Wireless Commun. Lett. **4**(2), 129–132 (2015).
9. P. Liu, S. Jin, T. Jiang, Q. Zhang, M. Matthaiou, Pilot power allocation through user grouping in multi-cell massive MIMO systems. IEEE Trans. Commun. **65**(4), 1561–1574 (2017).
10. H. Zhi, Q. Yuan, J. Zhu, Y. Hu, *A pilot allocation scheme for TDD massive MIMO system enabled HetNet*, 9th IEEE International Conference on Communication Software and Networks (ICCSN, GuangZhou, 2017), pp. 443–447.
11. J. Fan, W. Li, Y. Zhang, J. Deng, in *Proceedings of IEEE 85th Vehicular Technology Conference (VTC Spring)*. Fractional pilot reuse with vertical sectorization in massive MIMO systems (IEEE, Sydney, 2017), pp. 1–5.
12. H. Ahmadi, A. Farhang, N. Marchetti, A. MacKenzie, A game theoretic approach for pilot contamination avoidance in massive MIMO. IEEE Wireless Commun. Lett. **5**(1), 12–15 (2016).
13. R. Mochaourab, E. Björnson, M. Bengtsson, Adaptive pilot clustering in heterogeneous massive MIMO networks. IEEE Trans. Wirel. Commun. **15**(8), 5555–5568 (2016).
14. L. You, X. Gao, X. Xia, N. Ma, Y. Peng, Pilot reuse for massive MIMO transmission over spatially correlated rayleigh fading channels. IEEE Trans. Wirel. Commun. **14**(6), 3352–3366 (2015).
15. H. Yin, D. Gesbert, M. Filippou, Y. Liu, A coordinated approach to channel estimation in large-scale multiple-antenna systems. IEEE J. Sel. Areas Commun. **31**(2), 264–273 (2013).
16. L.S. Muppirisetty, H. Wymeersch, J. Karout, G. Fodor, in *Proceedings of 2015 IEEE Global Communications Conference (GLOBECOM)*. Location-aided pilot contamination elimination for massive MIMO systems (IEEE, San Diego, 2015), pp. 1–5.
17. L.S. Muppirisetty, T. Charalambous, J. Karout, G. Fodor, H. Wymeersch, Location-aided pilot contamination avoidance for massive MIMO systems. IEEE Trans. Wirel. Commun. **17**(4), 2662–2674 (2018).
18. T. Kailath, A.H. Sayed, B. Hassibi, *Linear Estimation* (Prentice-Hall, Upper Saddle River, 2000).
19. E. Björnson, E.G. Larsson, M. Debbah, Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated? IEEE Trans. Wirel. Commun. **15**(2), 1293–1308 (2016).
20. M.J. Osborne, A. Rubinstein, *A Course in Game Theory* (MIT Press, Cambridge, 1994).
21. A. Bogomolnaia, M.O. Jackson, The stability of hedonic coalition structures. Games Econ. Behavior. **38**(2), 201–230 (2002).