

Pilot Allocation and Sum-rate Analysis in Cell-free Massive MIMO Systems

Ramiz Sabbagh, Cunhua Pan, and Jiangzhou Wang

School of Engineering and Digital Arts, University of Kent, Canterbury, CT2 7NT, United Kingdom

Email: {rar34, c.pan, and j.z.wang}@kent.ac.uk

Abstract—This paper deals with the challenging issue of the unaffordable channel training overhead in the dense cell-free massive multi-input multi-output (MIMO) system when a high number of users are being simultaneously served. By adopting the user-centric cluster method, a dynamic pilot reuse (DPR) scheme is proposed to allow a pair of users to share a single pilot sequence. Specifically, the proposed reuse scheme is achieved with the objective of maximizing the uplink achievable sum-rate subject to users' signal to interference plus noise ratio (SINR) requirements and pilot resources constraints. On this basis, the SINR expression is derived for any user sharing its pilot with another by utilizing both minimum mean squared error (MMSE) detection and channel estimation. A low complexity pilot reuse algorithm is then developed based on the separation distance between users. The iterative grid search (IGS) method is employed to find the threshold that can be utilized in the proposed algorithm to maximize the sum-rate. Finally, simulation results are presented to show the effectiveness of the DPR scheme with the optimized threshold in terms of the uplink achievable sum-rate.

I. INTRODUCTION

Distributed massive multi-input multi-output (DM-MIMO) is an attractive technology to meet the vast growth in the requirements for huge data rates in future wireless communication systems [1], [2]. A distributed antenna system (DAS) with spatially separated antennas provides a better coverage, power saving and capacity merits than a co-located antenna system (CAS) at the cost of high fronthaul requirements [3]. Since DM-MIMO combines both DAS and massive MIMO concepts, it will bring all the benefits from these two systems [4]. In DM-MIMO, multiple remote radio heads (RRHs) are spread out over a large area and connected to a central processing unit (CPU) via high speed links, such as optical fibers. All the signal processing operations are performed at the CPU so that the multiple RRHs can cooperatively communicate with the users in order to improve system performance [1], [2].

By adopting the time-division duplexing (TDD) mode, DM-MIMO exploits the channel reciprocity property. If the RRHs obtain the channel state information (CSI) through pre-defined orthogonal pilot sequences transmitted by users in the uplink, the CSI will be valid during downlink transmission as well due to the channel reciprocity. In a dense DM-MIMO network, tremendous amount of CSIs are to be measured to perform the centralized signal processing operations [5]. If all users are assigned orthogonal pilots, it will cost unaffordable pilot resources which increase linearly with the number of users.

Thus, some users should be allowed to use the same pilot to reduce the training overhead [6]. A smart pilot reuse scheme should be developed to reduce the channel estimation error.

In this work, the DM-MIMO is considered without cell boundary and so this system can be called a cell-free massive MIMO [4], [7]. The user-centric cluster method is further adopted here to reduce the computational complexity at the CPU [8], [9]. In this method, each user selects its own serving cluster from the nearby RRHs since the distant RRHs contribute little to the user signal quality. This method allows users' clusters to overlap with each other. In the literature, some papers [4], [7], [10] have been produced on the pilot allocation design or overhead reduction in the cell-free massive MIMO. In [10], non-orthogonal pilot codes are proposed with applying sectorization in which RRHs can be turned on or off to prevent pilot collisions which may reduce system performance. It should be also noticed that the user-centric cluster method has not been included in this work. A downlink pilot scheme was proposed in [7] based on that less channel hardening can be expected in the cell-free massive MIMO compared with that of a co-located massive MIMO. However, this paper did not take into account the effect of pilot reuse. In [4], both random and greedy pilot assignment were investigated along with the power control. Nonetheless, single antenna RRHs were considered in this work.

Pilot reuse design has been extensively studied in co-located and distributed massive MIMO [11]–[15]. In the proposed reuse schemes, the same pilots are generally assigned to the users with different angles or least channel correlations based on the special characteristics of massive MIMO. Some of these works [11]–[13] adopted cell-based approaches which are not applicable in the cell-free scenario where the notion of cell is no longer valid. Other approaches included the data-aided channel estimation with superimposed training design as in [14], and the blind channel estimation based on singular value decomposition (SVD) as in [15]. However, such approaches suffered high computational complexity when applied to a dense cell-free massive MIMO with a large network size.

In this paper, a dynamic pilot reuse (DPR) scheme is proposed in the cell-free massive MIMO system to deal with the case when the number of users is larger than the number of available pilot sequences. The proposed reuse scheme is developed to allow the same pilot to be allocated to a pair of

users with some practical constraints. Specifically, our main objective is to maximize the achievable sum-rate with a fixed number of pilots. Another constraint is imposed that the signal to interference plus noise ratio (SINR) for each user is no less than a specific threshold where the impacts of pilot contamination are considered. For any pair of users sharing the same pilot, it is a requirement that no common RRHs simultaneously serve the two users. To limit the effect of pilot contamination, it is assumed that a pilot is shared by a pair of users, and the pilot reuse is only employed for the pair of users that are sufficiently separated so they can meet the SINR constraint. The achievable sum-rate for all users is first analysed. An algorithm for pilot allocation is then proposed to find the pairs of users that meet the SINR constraint. The SINR threshold is optimized to maximize the sum-rate. Simulation results verify the effectiveness of the proposed reuse scheme in terms of the uplink achievable sum-rate.

Notations : In this paper, boldface lower and upper case symbols represent vectors and matrices, respectively. $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote transpose, conjugate and Hermitian transpose, respectively. $\mathbb{E}\{\cdot\}$ means the expectation operation and \hat{a} is the estimation value of a . \mathbf{I}_N denotes the $N \times N$ dimensional identity matrix, and $\text{tr}(A)$ is the trace of A . Finally, $\mathbb{C}^{M \times N}$ is the set of complex matrices with M rows and N columns.

II. SYSTEM MODEL

A dense network with N RRHs and \bar{K} users is considered in this paper as shown in Fig. 1. Each RRH and each user are equipped with M antennas and a single antenna, respectively. Denote the set of RRHs and users as $\mathcal{N} = \{1, \dots, N\}$ and $\bar{\mathcal{K}} = \{1, \dots, \bar{K}\}$, respectively. It is supposed that K users are admitted in this network and the set of these users is denoted by $\mathcal{K} \subseteq \bar{\mathcal{K}}$. In addition, the set of antennas of the i^{th} RRH is denoted as $\mathcal{M}_i = \{1, \dots, M\}$. It is also assumed that $NM \gg K$, and the users are randomly distributed within the coverage area. Furthermore, RRHs are physically connected to the CPU through high-speed links, where the CPU performs the joint detection of the received signals from multiple RRHs.

TDD mode is adopted in this scenario, which indicates that the CSI is the same for both uplink and downlink transmissions. The channel estimation of the uplink channels is executed in the CPU with the received pilot signals. It is assumed that S time-frequency resource units (RUs) are allocated within one transmission frame. Within these S RUs, B RUs are reserved for pilot sequences in each frame. The remaining $(S - B)$ RUs are dedicated for uplink and downlink data transmission.

Denote the available pilot set as $\mathcal{V} = \{1, \dots, B\}$, and the corresponding orthogonal pilot sequences as $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k, \dots, \mathbf{v}_B\} \in \mathbb{C}^{B \times B}$, where each $\mathbf{v}_k \in \mathbb{C}^{B \times 1}$ in set \mathbf{V} is mutually orthogonal with each other, that is $\mathbf{V}^H \mathbf{V} = B \mathbf{I}_B$ [2], [16]. By using the orthogonal pilot assignment scheme, only B users are able to send their pilots without interfering with each other [16]. In order to serve more users, a pilot reuse scheme is required. We denote an arbitrary pilot reuse scheme as $\mathcal{P}(\mathcal{K}, \mathcal{V}) = (k, \pi_k) : k \in \mathcal{K}, \pi_k \in \mathcal{V}$, where (k, π_k) denotes that the k^{th} user is allocated with the pilot sequence \mathbf{v}_{π_k} .

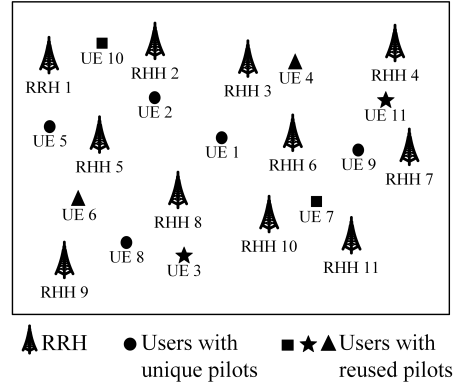


Fig. 1. Structure of a cell-free massive MIMO and an illustration of the proposed DPR scheme by supposing $B = 8$ and $K = 11$. Unique pilots are allocated for users with circle shapes, and a single pilot is allocated for a pair of users with any one of the following shapes: square, triangle and star.

The uplink channel response for the k^{th} user to the m^{th} antenna in the i^{th} RRH is given by [12]:

$$g_{k,i,m} = \beta_{k,i}^{1/2} h_{k,i,m}, \quad (1)$$

where $\beta_{k,i}$ represents the large scale fading coefficient between the k^{th} user to the i^{th} RRH, which consists of both the path loss and the shadow fading. $\beta_{k,i}$ is assumed to be the same between the k^{th} user and all M antennas of any RRH. $h_{k,i,m}$ is the small scale fading coefficient between the k^{th} user to the m^{th} antenna in the i^{th} RRH, and it is assumed that each element of $h_{k,i,m}$ is an i.i.d. complex Gaussian variable with zero mean and unit variance. The channel response $M \times 1$ vector from the k^{th} user to the i^{th} RRH is given by [1]

$$\mathbf{g}_{k,i} = \beta_{k,i}^{1/2} [h_{k,i,1}, h_{k,i,2}, \dots, h_{k,i,M}]^T. \quad (2)$$

By assuming a synchronous pilot transmission from all users, the received signal matrix of pilots at the i^{th} RRH is

$$\mathbf{Y}_{p,i} = \sqrt{\rho_p} \sum_{k \in \mathcal{K}} \mathbf{g}_{k,i} \mathbf{v}_{\pi_k}^T + \mathbf{Z}_{p,i} \in \mathbb{C}^{M \times B}, \quad (3)$$

where ρ_p is the transmit power for pilots. \mathbf{v}_{π_k} is the pilot transmitted by the k^{th} user, where $\mathbf{v}_{\pi_k} \in \mathbf{V}$. Also, $\mathbf{Z}_{p,i} \in \mathbb{C}^{M \times B}$ is complex additive white Gaussian noise with zero mean and unit covariance variables. The CPU has the following observation of the channel from the k^{th} user to the i^{th} RRH [2], [5]

$$\tilde{\mathbf{y}}_{p,k,i} = \frac{1}{\sqrt{\rho_p B}} \mathbf{Y}_{p,i} \mathbf{v}_{\pi_k}^* \in \mathbb{C}^{M \times 1}. \quad (4)$$

III. SUM-RATE ANALYSIS

In this paper, the scenario that K is larger than the number of available pilot sequences (B) is discussed. Therefore, the DPR scheme is proposed to measure the CSIs for more users by permitting a pair of users to share one pilot and different pilots are assigned to different pairs. In this section, given the pairs of users using the same pilot according to the proposed DPR scheme, the achievable sum-rate for data transmission

is analysed. Based on the analytical results and the target of maximizing the sum-rate, an algorithm for the dynamic pilot reuse is proposed in the next section.

After reusing the pilots for some user pairs, no pilot reuse is considered for the rest of the users, and unique pilot sequences are allocated to them. As illustrated in Fig. 1, ($B = 8$) and ($K = 11$) are considered. This means that three pilots ($K - B$) are required to be reused in order to serve all the users in Fig. 1. The users with unique pilot sequences are represented with the black circles, and the pair of users using the same pilot sequence is represented by one of the following shapes: square, triangle and star. As a result, two groups of users will be created: users with unique pilot sequences, and users with reused pilot sequences. In order to obtain the channel estimation for these users, the equation (4) for any user should be expanded as [2], [5]

$$\tilde{\mathbf{y}}_{p,k,i} = \sum_{k \in \mathcal{I}} \mathbf{g}_{k,i} + \frac{\mathbf{z}_{p,i} \mathbf{v}_{\pi_k}^*}{\sqrt{\rho_p B}} \quad \forall k, i, \quad (5)$$

where \mathcal{I} is the set that includes the pair of users $\{k, k'\}$ in the case of pilot reuse where the two users share the same pilot, or \mathcal{I} includes only the k^{th} user in the case of no reuse. Based on [17], CPU computes the MMSE estimation of the k^{th} user

$$\hat{\mathbf{g}}_{k,i} = \beta_{k,i} \left(\sum_{k \in \mathcal{I}} \beta_{k,i} + \frac{1}{\sqrt{\rho_p B}} \right)^{-1} \tilde{\mathbf{y}}_{p,k,i} \quad \forall k, i, \quad (6)$$

where $\hat{\mathbf{g}}_{k,i} \sim \mathcal{CN}(0, \Phi_{k,i} \mathbf{I}_M)$ and $\Phi_{k,i}$ is given by

$$\Phi_{k,i} = \beta_{k,i}^2 \left(\sum_{k \in \mathcal{I}} \beta_{k,i} + \frac{1}{\sqrt{\rho_p B}} \right)^{-1} \quad \forall k, i. \quad (7)$$

The channel $\mathbf{g}_{k,i}$ can be decomposed as $\mathbf{g}_{k,i} = \hat{\mathbf{g}}_{k,i} + \tilde{\mathbf{g}}_{k,i}$, where $\tilde{\mathbf{g}}_{k,i}$ is the channel estimation error and it is statistically independent of $\hat{\mathbf{g}}_{k,i}$ due to the orthogonal property of MMSE estimation and the joint Gaussianity of both vectors, the covariance of $\tilde{\mathbf{g}}_{k,i}$ is $((\beta_{k,i} - \Phi_{k,i}) \mathbf{I}_M)$. During the uplink data transmission, the received data signal vector for the i^{th} RRH is

$$\mathbf{y}_{d,i} = \sqrt{\rho_d} \sum_{k \in \mathcal{K}} \mathbf{g}_{k,i} x_k + \mathbf{z}_{d,i} \in \mathbb{C}^{M \times 1}, \quad (8)$$

where ρ_d is the transmit power for data and all users have the same transmit power, $\mathbf{z}_{d,i} \in \mathbb{C}^{M \times 1}$ is complex additive white Gaussian noise with zero mean and unit covariance variables. x_k is the transmitted data signal from the k^{th} user, $k \in \mathcal{K}$.

It is assumed that the MMSE detection is applied in the CPU for recovering user data. The detector $M \times 1$ vector for the k^{th} user at i^{th} RRH is given as [17]

$$\hat{\mathbf{a}}_{k,i} = (\hat{\mathbf{G}}_i \hat{\mathbf{G}}_i^H + \mathbf{Z}_i + \frac{1}{\rho_d} \mathbf{I}_M)^{-1} \hat{\mathbf{g}}_{k,i}, \quad (9)$$

where $\hat{\mathbf{G}}_i$ is the estimated channel response $M \times K$ matrix from all the users to the i^{th} RRH, and the k^{th} column of $\hat{\mathbf{G}}_i$ is $\hat{\mathbf{g}}_{k,i}$. In addition, $\mathbf{Z}_i = \sum_{k \in \mathcal{K}} \mathbb{E}(\tilde{\mathbf{g}}_{k,i} \tilde{\mathbf{g}}_{k,i}^H)$ is the covariance matrix of the channel estimation errors for the i^{th} RRH.

The user-centric cluster method is adopted here to reduce the computational cost for dense cell-free massive MIMO. This means that users are only served by a cluster of the nearest RRHs since distant RRHs contribute little to the user's signal quality due to the severe path loss. The notation \mathcal{G}_k is used to represent the group of RRHs serving the k^{th} user. For a pair of users sharing the same pilot, the clusters of RRHs serving these two users should not overlap. In other words, the users served by at least one common RRH should be allocated with difference pilots. This is essential to enable the CPU to differentiate the CSIs for the users sharing a single pilot. Based on (9), the CSIs from all users to the i^{th} RRH are needed. However, for users with pilot reuse, the instantaneous CSIs that can be provided are only from a user to its serving cluster of RRHs. For the CSI from the user to the RRHs outside its cluster, only the large-scale fading (path loss and shadowing) can be tracked. The assumption of the availability of the large-scale fading is valid because of the slow rate at which this parameter varies compared with the small-scale fading, thus it can be obtained at a much lower cost.

The MMSE detection is then applied in the CPU. The estimated data symbol for the k^{th} user is represented by

$$\hat{x}_k = \sum_{i \in \mathcal{G}_k} \hat{\mathbf{a}}_{k,i}^H \mathbf{y}_{d,i}. \quad (10)$$

The ergodic achievable sum-rate can be written as

$$R_T = \left(1 - \frac{B}{S}\right) \sum_{k \in \mathcal{K}} \log_2(1 + \text{SINR}_k). \quad (11)$$

The SINR for the k^{th} user, whose pilot is reused with the k'^{th} user, is given in (12). For a user with unique pilot, the SINR can still be obtained from (12) but the second term of the pilot contamination in the denominator of (12) should be neglected.

In order to simplify the SINR in (12), the deterministic equivalent analysis from [17] is employed here. The results in [17] are asymptotic since the SINR formulas were derived by considering the number of antennas increases considerably. However, simulation results in [17] demonstrate that even for a small number of antennas the fit between simulation and approximation is acceptable. According to [17, Theorem 5], the SINR for MMSE detector is given in [17, equation (25)]. This formula is modified to match the cell-free massive MIMO scenario of this paper as shown in (13) and (14), for a user without pilot reuse and a user with pilot reuse, respectively. The elements of the channel vector from the k^{th} user to the antennas of the BS in [17] are replaced by the channel vector from the same user to all the antennas of the serving set of RRHs (\mathcal{G}_k). In the denominator of both (13) and (14), the first term represents the power of noise, the second term is the sum interference power from other users and the third or the last term in (14) is the pilot contamination. The parameters of $\delta_k, \mu_{\langle q, k \rangle}, \vartheta_k$ and \mathbf{T} can be calculated by using [17, Theorem 1 and Theorem 2]. These parameters depend on

$$\text{SINR}_{\langle k, k' \rangle} = \frac{\rho_d (\sum_{i \in \mathcal{G}_k} \hat{\mathbf{a}}_{k,i}^H \hat{\mathbf{g}}_{k,i})^2}{\rho_d [(\sum_{i \in \mathcal{G}_k} \hat{\mathbf{a}}_{k,i}^H \hat{\mathbf{g}}_{k,i})^2 + (\sum_{i \in \mathcal{G}_k} \hat{\mathbf{a}}_{k,i}^H \mathbf{g}_{k',i})^2 + \sum_{\substack{q \in \mathcal{K} \\ q \neq k \\ q \neq k'}} (\sum_{i \in \mathcal{G}_k} \hat{\mathbf{a}}_{k,i}^H \mathbf{g}_{q,i})^2] + (\sum_{i \in \mathcal{G}_k} \hat{\mathbf{a}}_{k,i}^H \mathbf{z}_{d,i})^2} \quad (12)$$

\mathbf{Z}_i , which is defined in (9), and can also be calculated from $\mathbf{Z}_i = \sum_{k \in \mathcal{K}} (\mathbf{R}_{k,i} - \Phi_{k,i}) \forall i$ [17]. $\mathbf{R}_{k,i}$ is a diagonal matrix with the diagonal entries of $\beta_{k,i,m}$, where $\beta_{k,i,m} = \beta_{k,i}$ for $m \in \mathcal{M}_i$, $m = 1, \dots, M$, and $i = 1, \dots, N$. Similarly, $\Phi_{k,i}$ is a diagonal matrix with the diagonal entries of $\phi_{k,i,m}$, where $\phi_{k,i,m} = \phi_{k,i}$ for $m \in \mathcal{M}_i$, $m = 1, \dots, M$, and $i = 1, \dots, N$. Then, employ [17, Theorem 1 and Theorem 2] to calculate the parameters δ_k , \mathbf{T}' , $\mu_{\langle q,k \rangle}$ and ϑ_k , which are all functions of $\mathbf{R}_{k,i}$ and $\Phi_{k,i}$. The expressions of these parameters are not included here due to space limitation.

$$\text{SINR}_k = \frac{\delta_k^2}{\frac{1}{\rho_d (NM)^2} \text{tr } \Phi_k \mathbf{T}' + \frac{1}{(NM)^2} \sum_{\substack{q \in \mathcal{K} \\ q \neq k}} \mu_{\langle q,k \rangle}} \quad (13)$$

$$\text{SINR}_{\langle k, k' \rangle} = \frac{\delta_k^2}{\frac{1}{\rho_d (NM)^2} \text{tr } \Phi_k \mathbf{T}' + \frac{1}{(NM)^2} \sum_{\substack{q \in \mathcal{K} \\ q \neq k}} \mu_{\langle q,k \rangle} + |\vartheta_{k'}|^2} \quad (14)$$

In this paper, the DPR scheme is developed to allocate the pilots with the objective of maximizing the achievable uplink sum-rate in (11) under some practical constraints and by employing the user-centric cluster method. Hence, this problem can be formulated as

$$\max_{\mathcal{P}(\mathcal{K}, \mathcal{V})} \mathbb{E}\{\mathbf{R}_T\}, \quad (15a)$$

$$\text{s.t. } \text{SINR}_{\langle k, k' \rangle} \geq \gamma_{th}, \quad (15b)$$

$$|\mathcal{V}| = B, \quad (15c)$$

$$\mathcal{G}_k \cap \mathcal{G}_{k'} = \emptyset. \quad (15d)$$

It can be seen from (15) that the reuse of the pilots between pairs of users is considered. For any pair of users $\{k, k'\}$, the following constraints are provided: First, the constraint (15b) means that each user in a pair with the same pilot should satisfy a SINR threshold (γ_{th}). This threshold will be optimized as shown later to satisfy the objective in (15a). No SINR constraint is considered for the users with unique pilots. Secondly, constraint (15c) represents the fact that only B pilots are available. Furthermore, the number of times of pilot reuse is denoted as ω , where ω is upper bounded by $\min(N_p, B)$, and

$$N_p = \begin{cases} \frac{K}{2}, & \text{if } K \text{ is even} \\ \frac{K}{2} - 1, & \text{if } K \text{ is odd.} \end{cases} \quad (16)$$

Finally, (15d) indicates that the two RRHs clusters \mathcal{G}_k and $\mathcal{G}_{k'}$ which are based on the user-centric cluster method must not be overlapped, as explained earlier.

IV. ALGORITHM FOR DYNAMIC PILOT REUSE

The optimal solution of the problem in (15a) can be obtained by the exhaustive search method, in which it is required to test

all possible pilot allocation schemes and choose the one that achieves the aim of maximizing the sum-rate. However, the exhaustive search method may be at the cost of high complexity, especially in the case of a high number of users and RRHs. Therefore, a low complexity algorithm for pilot allocation is presented in this section. This algorithm will be referred by the dynamic pilot reuse (DPR) algorithm in the rest of this paper. In the DPR algorithm, two users separated by a large distance can share the same pilot to ensure a marginal channel estimation error. The reuse is only allowed when the potential SINRs for the two users sharing the same pilot are still higher than the SINR threshold γ_{th} . The SINRs of these users can be calculated based on (14), which includes the effect of pilot contamination. Additionally, (14) depends on the users' large scale fading coefficients $\beta_{k,i}$. This means that it is required in the DPR algorithm to have the large-scale channel gains, which change very slowly and can be accurately obtained. This enables the DPR algorithm to know the potential SINRs for users irrespective of the actual channel realization.

The explanation of the DPR algorithm can be described as follows: A search is effected for the possible configurations of user pairs with non-overlapping serving set of RRHs as in (15d). The set of user pairs are ordered descendingly in $\mathcal{Q} = [q_1, \dots, q_f, \dots, q_F]$ according to the separation distance, where F is the number of pairs. The pair q_1 is initially chosen, and the two users of this pair are considered as $\{k, k'\}$. For this pair, the potential values of $\text{SINR}_{\langle k, k' \rangle}$ and $\text{SINR}_{\langle k', k \rangle}$ are calculated based on equation (14) and compared with γ_{th} . If both of them are larger than or equal to γ_{th} , one pilot will be allocated for the two users. After that, the other user pairs in \mathcal{Q} that include any of the two users k or k' will be omitted from \mathcal{Q} where this will reduce the search field. Later, a new pair in of users in \mathcal{Q} is selected and tested. The above procedures to reuse pilots will be repeated a number of times equal to ω or when all the pairs of users are tested ($f = F$). When the number of reused pilots equals ω or when all pairs of users are tested, reusing pilots will be terminated. It is necessary here to allocate the remaining pilots, if there is, randomly to the rest of the users. If all pairs satisfy the SINR constraint and in the end some pilots are still available, the pairs with lowest average SINR will be separated and a unique pilot will be allocated to each user in these pairs.

The value of γ_{th} utilized in the DPR algorithm is calculated by considering the iterative grid search (IGS) method [12] to find the threshold that provides the objective of maximum sum-rate in (15a). In IGS method, it is first required to measure the minimum and maximum values of γ_{th} , i.e. γ_{min} and γ_{max} , respectively. In the extreme case when $\gamma_{th} \leq \gamma_{min}$, any pair of users can meet the SINR constraint in (15b) whatever

the separation distance between them. In this case, maximum number of users can be served, but high pilot contamination is expected. On the other hand, when $\gamma_{th} \geq \gamma_{max}$, no pair is able to meet the SINR constraint. This case indicates that no reuse can be applied and the number of users that can be served is the same as the number of pilots (B) with no pilot contamination. The near-optimal threshold should be located between $[\gamma_{min}$ and $\gamma_{max}]$ and can be found by sampling the interval $[\gamma_{min}$ and $\gamma_{max}]$ by \mathbb{N} points with uniform subinterval $\Delta = (\gamma_{max} - \gamma_{min})/(\mathbb{N} - 1)$. At each point, the sum-rate is calculated, and the threshold that provides the maximum sum-rate value will be considered as $\gamma_{max}^{(1)}$. These procedures will be repeated by supposing that the new threshold interval is, $[\gamma_{max}^{(1)} - \Delta^{(1)}/2, \gamma_{max}^{(1)} + \Delta^{(1)}/2]$. These steps will be applied \mathbb{T} times, and the final near-optimal threshold will be $\gamma_{th} = \gamma_{max}^{(T)}$. The number of times of reuse at γ_{th} is considered as ω .

The computational complexity of the DPR algorithm is investigated in this part. The maximum number of the possible combinations of user pairs that satisfy (15d) is $F = \sum_{f=1}^{K-1} f$. Therefore, the highest complexity for the DPR algorithm is achieved when it is required to search over all the combinations of user pairs to find the pairs that can meet the pilot reuse condition based on a specific threshold γ_{th} . By including the pairs classification and the search over user pairs, the complexity of the DPR algorithm is $\mathcal{O}(2F)$ in the case of using a fixed SINR threshold (γ_{th}). Using the IGS method to find γ_{th} that provides the maximum sum-rate requires additional iterations equal to $(\mathbb{N}\mathbb{T})$. As a result, the upper bound of the total complexity when the optimized threshold is used will be $\mathcal{O}((\mathbb{N}\mathbb{T} + 2)F)$.

V. SIMULATION RESULTS

The uplink performance of the proposed pilot reuse scheme is evaluated in this section. The dense cell-free massive MIMO system is assumed to be deployed in a square area of 4 Km \times 4 Km. Both users and RRHs are randomly and uniformly placed in the this area. The large scale fading coefficient $\beta_{k,i}$ can be calculated as $\beta_{k,i} = z_{k,i}/(d_{k,i}/r_o)^\alpha$, where $z_{k,i}$ is the shadow fading and it obeys a log-normal distribution, which is represented by, $10 \log_{10}(z_{k,i}) \sim \mathcal{CN}(0, \sigma_{shad})$, and $d_{k,i}$ is the distance between the k^{th} user to the i^{th} RRH [6]. Unless otherwise specified, each user is served by its nearest two RRHs based on the user-centric cluster method, which indicates that $|\mathcal{G}_k| = 2, \forall k$. The other simulation parameters are set as follows: $N = 19, S = 64, M = 20$, normalized radius $r_o = 100\text{m}$, decay exponent $\alpha = 3, \sigma_{shad} = 8$ dB, transmit power for pilots and data $\rho_p = \rho_d = 40$ dBm and the IGS parameters are $\mathbb{N} = 10$ and $\mathbb{T} = 2$. The performance of the DPR scheme with the optimized threshold is compared with the following schemes:

- 1) Orthogonal pilot allocation: A unique orthogonal pilot is allocated for each user in this scheme. Thus, the number of admitted users is equal to the number of pilots (B).
- 2) Random pilot reuse: In this scheme, user pairs are selected randomly to allocate the same pilot for them without taking into account any SINR constraint.

- 3) DPR scheme with no threshold: This scheme adopts the DPR algorithm as well but allocating the pilots to user pairs is performed based only on the separation distance between the two users in each pair. Therefore, the SINR constraint will not be taken into account in this case too.
- 4) DPR scheme with a fixed threshold: The DPR algorithm is utilized here with a fixed SINR threshold which means that the IGS method is not adopted in this case.

To guarantee the fairness, the number of times of pilot reuse ω in the DPR scheme with both no threshold and a fixed threshold and the random pilot reuse scheme is supposed to be the same as ω in the DPR scheme with the optimized threshold.

In Fig. 2, the uplink achievable sum-rate is plotted versus the number of available pilots B with $\bar{K} = 32$. As expected, the orthogonal pilot allocation scheme has the worst performance since no users are allowed to reuse the pilots. On the other hand, the performance of the DPR scheme with the optimized threshold outperforms the random pilot reuse scheme, the DPR scheme with fixed SINR thresholds (5 dB and 10 dB) and no threshold cases. The random selection of user pairs to reuse their pilots in the random pilot reuse scheme may lead to a severe pilot contamination in pairs with low separation distances. Furthermore, a fixed threshold is generally less effective than the optimized one in the DPR scheme although all users guarantee their SINR requirements through the fixed threshold. The sum-rate performance at the threshold of 5 dB is better than the 10 dB as more users can be admitted in the case of 5 dB compared with the 10 dB as shown in Fig. 3. Using the IGS method to continuously find the threshold that provides the maximum sum-rate gives the superiority to the DPR scheme when this threshold is adopted. However, additional computational complexity is needed due to the use of the IGS method compared with the fixed threshold.

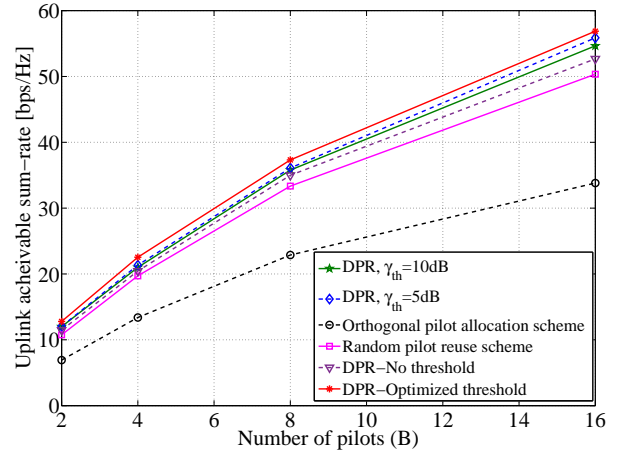


Fig. 2. Uplink achievable sum-rate versus the number of pilots (B).

Fig. 3 shows the number of admitted users K versus the number of pilots B for the orthogonal pilot allocation and the DPR scheme with both the optimized threshold and fixed thresholds (5 dB and 10 dB) when $\bar{K} = 32$. The numbers

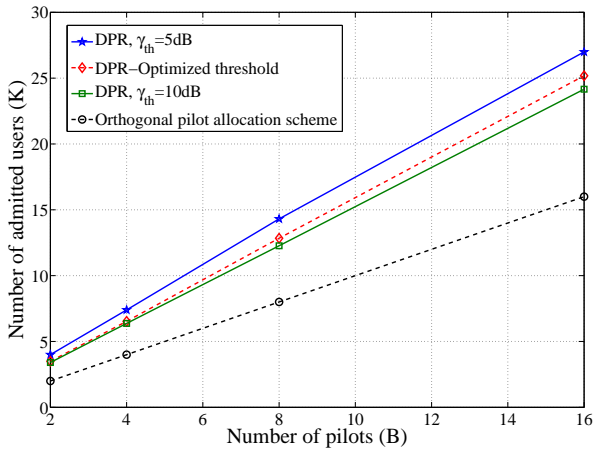


Fig. 3. Number of admitted users versus the available pilots (B).

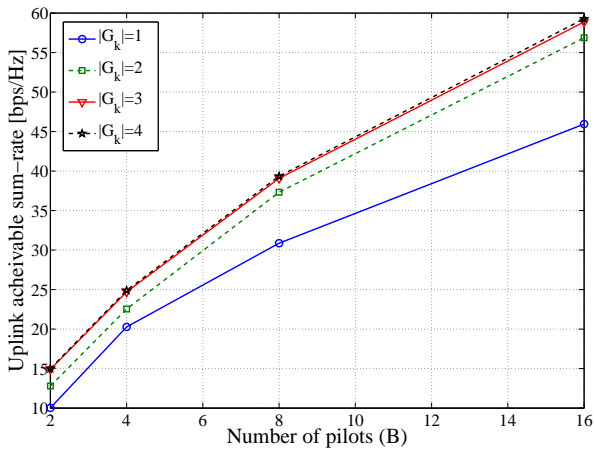


Fig. 4. The impact of the size of \mathcal{G}_k on the uplink achievable sum-rate.

of admitted users for all cases increases with increasing the number of available pilots. However, the orthogonal pilot allocation scheme can serve the lowest number of users since no users are allowed to reuse the pilots. For the fixed threshold DPR scheme, the number of admitted users increases when the threshold decreases as more users can satisfy the SINR constraint. For the optimized threshold DPR scheme, the number of admitted users depends on the threshold itself, where the aim here is to obtain the maximum sum-rate which does not guarantee a maximum number of admitted users.

Finally, Fig. 4 plots the uplink sum-rate for the DPR scheme with the optimized threshold versus B with $\bar{K} = 32$, and for various sizes of serving cluster of the user-centric method ($|\mathcal{G}_k| = 1, 2, 3$ and 4). The rate performance generally improves as $|\mathcal{G}_k|$ increases, where more of the nearest RRHs serve the users. Thus, more uplink signals for these users will be coherently detected by the CPU. However, when $|\mathcal{G}_k|$ is highly increased, such as when $|\mathcal{G}_k| = 3$ or 4 , the two sets of \mathcal{G}_k of a pair of users that share a single pilot will be closer to each other. Hence, higher pilot contamination could be received by RRHs located close to the second user using the same pilot. This will

negatively affect the rate performance.

VI. CONCLUSION

In this paper, the DPR scheme in a TDD cell-free massive MIMO system was proposed with the objective of maximizing the sum-rate subject to some practical constraints. Firstly, an expression for SINR was derived for any pair of users sharing the same pilots. An algorithm with low complexity was then proposed based on the separation distance between users. Simulation results showed that the proposed scheme with the optimized threshold can achieve a better sum-rate performance compared with other pilot reuse schemes.

VII. ACKNOWLEDGEMENT

The first author acknowledges the Ministry of Higher Education and Scientific Research-IRAQ / Ninevah University for their financial support of his PhD research program.

REFERENCES

- [1] K. T. Truong and R. W. Heath, "The viability of distributed antennas for massive MIMO systems," in *2013 Asilomar Conference on Signals, Systems and Computers*, Nov. 2013, pp. 1318–1323.
- [2] M. Sadeghi, C. Yuen, and Y. H. Chew, "Sum rate maximization for uplink distributed massive MIMO systems with limited backhaul capacity," in *2014 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 308–313.
- [3] W. Ali, J. Wang, H. Zhu, and J. Wang, "Distributed antenna system based frequency switch scheme evaluation for high-speed railways," in *2017 IEEE International Conference on Communications*, May 2017, pp. 1–6.
- [4] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, March 2017.
- [5] Q. He, L. Xiao, X. Zhong, and S. Zhou, "Increasing the sum-throughput of cells with a sectorization method for massive MIMO," *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1827–1830, Oct. 2014.
- [6] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov., 2010.
- [7] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "How much do downlink pilots improve cell-free massive MIMO?" in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–7.
- [8] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint precoding and RRH selection for user-centric green MIMO C-RAN," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2891–2906, May 2017.
- [9] C. Pan, H. Mehrpouyan, Y. LIU, M. Elkashlan, and A. Nallanathan, "Joint pilot allocation and robust transmission design for ultra-dense user-centric TDD C-RAN with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2018.
- [10] O. Y. Bursalioglu, C. Wang, H. Papadopoulos, and G. Caire, "RRH based massive MIMO with "on the fly" pilot contamination control," in *2016 IEEE International Conference on Communications*, May 2016, pp. 1–7.
- [11] H. Yang and T. L. Marzetta, "Performance of pilot reuse in multi-cell massive MIMO," in *2015 IEEE International Black Sea Conference on Communications and Networking*, May 2015, pp. 157–161.
- [12] X. Zhu, L. Dai, and Z. Wang, "Graph coloring based pilot allocation to mitigate pilot contamination for multi-cell massive MIMO systems," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1842–1845, Oct. 2015.
- [13] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, February 2013.
- [14] K. Upadhyya, S. A. Vorobyov, and M. Vehkaperä, "Superimposed pilots are superior for mitigating pilot contamination in massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2917–2932, June 2017.
- [15] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3111–3124, June 2014.
- [16] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1293–1308, Feb. 2016.
- [17] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.