Pilot Allocation and Sum-rate Analysis in Distributed Massive MIMO Systems

Ramiz Raoof, Huiling Zhu, and Jiangzhou Wang
School of Engineering and Digital Arts, University of Kent, Canterbury, CT2 7NT, United Kingdom
Email: {rar34, h.zhu, and j.z.wang}@kent.ac.uk

***Abstract*—In distributed massive multi-input multi-output (DM-MIMO) systems, orthogonal pilot sequences are generally utilized to acquire the channel state information (CSI). However, this highly restricts the number of users simultaneously served. In this paper, a pilot reuse within a single cell DM-MIMO system is proposed to serve more users than the available pilot sequences. The reuse in this strategy is applied so that maximum achievable sum-rate is satisfied with the constraint of predefined pilot resource. On this basis, two users in different subcells separated by a large distance and satisfying a specific signal to interference plus noise ratio (SINR) level can share the same pilot sequence. An expression for SINR is derived for any pair of users who use the same pilot. Based on this expression, an algorithm is proposed to choose which pairs of users are able to use the same pilot with the constraint of satisfying the minimum SINR required for these users. The simulation results demonstrate that the uplink achievable sum-rate for the proposed strategy is higher than both cases when no pilot reuse or random pilot reuse are considered.**

***Index Terms*—Distributed antenna system, distributed massive MIMO system, pilot allocation, pilot contamination, pilot reuse.**

I. Introduction

Wireless communication systems have continuously been evolved to provide explosive spectral efficiency (SE) for meeting the vast growth in wireless traffic. Distributed massive multi-input multi-output (DM-MIMO) technology is an attractive direction to provide high data rates for future wireless communication systems [1]–[3]. Distributed antennas system (DAS) with spatially separated antennas provides power saving and capacity merits over co-located antenna system (CAS) [4], [5]. Combined with large antenna arrays of massive MIMO to specially multiplex many users on the same channel resources [6], [7], DM-MIMO was introduced to further enhance the data rate performance. In DM-MIMO, multiple remote radiohead (RRHs) are distributed over a cell and connected to a central unit (CU) via high speed links, such as optical fiber. Each RRH consists of multiple remote antenna units (RAUs). All the signal precess operations are performed at the CU so that the multiple RRHs can cooperatively communicate with the users in order to improve system performance [1]–[3].

Accurate channel state information (CSI) for all the channels between RRHs and users play an essential role in DM-MIMO transmission. In order to obtain accurate CSIs in the CU, orthogonal pilot sequences are generally employed for different users [17]. In time-division duplexing (TDD) massive MIMO transmission, CSIs can be acquired from the users' uplink pilots and uplink/downlink radio-channel reciprocity. This enables allocating only a few uplink pilot transmission as the same number of single-antenna users simultaneously served, to permit training a large antenna array [8]–[10]. However, when the number of users grows excessively, large pilot overhead inevitably affects the system sum-rate. One of the possible solutions for this problem is to reuse pilot radio resources among sets of users with tolerable interference, which was first discussed in [6].

Pilot reuse can cause pilot contamination, which limits the quality of the channel estimation and degrades system performance. In the literature, several approaches have been introduced to mitigate pilot contamination when pilots are reused among cells for different architectures of MIMO systems, including co-located massive MIMO [11]–[13], Network MIMO [14] and small cells scenarios [15], [16]. Among these researchers, the angle of arrival (AOA)-based methods have assumed pilot reuse within the same cell in co-located massive MIMO to reduce the pilot overhead [9], [10]. Nevertheless, pilot reuse and pilot contamination analysis within a single cell DM-MIMO system still need to be investigated. Where DM-MIMO has a different structure from the above mentioned systems, as mentioned earlier, and users in DM-MIMO can be served by multiple RRHs. We also remark, unlike above mentioned works that reuse pilot in a fixed reuse pattern, this paper considers a dynamic reuse strategy to allocate the same pilot for a pair of users within the single cell after guarantee they satisfy a specific SINR level. It is required in this strategy to adaptively change the number of times of pilot reuse according to the number of users or their locations.

This work suggests the pilot reuse within a single cell DM-MIMO system in the uplink transmission. By applying pilot reuse, maximum achievable sum-rate is satisfied with the constraints of fixed pilot resources and meeting the minimum SINR requirements of users who share their pilots. The dynamic pilot reuse strategy is proposed to allow the CU to estimate the channel for additional users. This strategy is implemented when the number of users is larger than the number of available orthogonal pilot sequences. A cell is divided into a number of subcells and a pair of users, each located in a different subcell, are allowed to share the same pilot under this scenario. To restrict the effect of intra-cell pilot contamination produced by pilot reuse, it is assumed that a pilot is shared by two users, and pilot reuse is only employed for a pair of users who satisfy specific SINR requirements which means they need to be located in sufficiently distant subcells and suffering only low pilot contamination. An algorithm for pilot allocation is presented to find the pairs of users that cope with these requirements. To enable the algorithm to know the potential SINR, a formula for the SINR is derived for any pair of users who share the same pilot by supposing both minimum mean square error (MMSE) channel estimator and detector. Simulation results verify the effectiveness of the proposed strategy in terms of the uplink achievable sum-rate compared with both no reuse case and random pilot reuse case in the uplink DM-MIMO system.

$Notations$: In this paper, bold symbols refer to matrices or vectors. $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote transpose, conjugate and Hermitian transpose respectively. $\mathbb{E}\{\cdot\}$ means the expectation operation, $\|\cdot\|$ represents the second-order norm and $\hat{a}$ is the estimation value of a. $\mathbf{I}_N$ denotes the $N \times N$ dimensional identity matrix, and 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

The DM-MIMO architecture in a single cell with N hexagonal subcells is shown in Fig. 1. One RRH is located in the center of each subcell. M antennas are equipped with each RRH, and the total number of antennas in the cell is NM. The set of antennas of the j^{th} RRH is supposed as $\mathcal{M}_j$. Furthermore, RRHs are physically connected to CU through high-speed links, where the CU performs the joint detection of the received signals from multiple RRHs. It is further assumed that the j^{th} subcell has a set of active single-antenna users, denoted by $\mathcal{K}_j$ and the total number of active users within the cell is $\sum_{j=1}^{N}|\mathcal{K}_j| = K$. It is also considered that $M \gg |\mathcal{K}_j|$ for all RRHs, and active users are randomly distributed within the cell.

TDD mode is assumed 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 employed in the CU with the received pilot signals, and the downlink channels are acquired by exploiting channel reciprocity. The time-frequency resource units (RUs) are allocated within frames with a length of T_c symbols and W_c subcarriers. One time-frequency RU is a unit of one subcarrier over one symbol period. The number of the RUs per frame is $S = T_cW_c$. Within these S RUs, B RUs are reserved for pilot sequences in each frame. This leaves room for $S-B$ RUs, which are split between uplink and downlink data transmission. The set of pilot sequences $\boldsymbol{\mathcal{V}}$ is defined as:

$$\boldsymbol{\mathcal{V}} = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k, ..., \mathbf{v}_B\}, \tag{1}$$

where each $\mathbf{v}_k \in \mathbb{C}^{B\times 1}$ in set $\boldsymbol{\mathcal{V}}$ is mutually orthogonal with each other, that is $\boldsymbol{\mathcal{V}}^H\boldsymbol{\mathcal{V}} = B\mathbf{I}_B$ [2], [17]. By transmitting these pilot signals over B RUs, only B users are able to send their pilots without interfering with each other [17].

We use the notation $\langle j,k\rangle$ to describe the k^{th} user in the j^{th} subcell. By assuming both a small and a large scale fading channel, the uplink channel response for user $\langle j,k\rangle$ to the m^{th} antenna in the i^{th} RRH is given by:

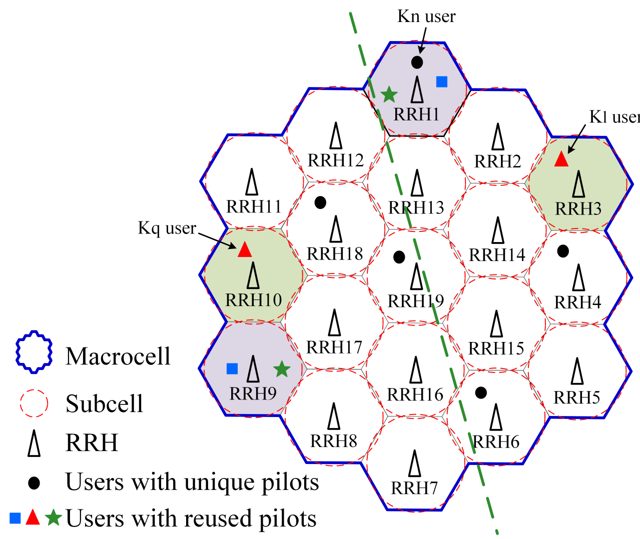

Fig. 1. Structure of a single cell DM-MIMO with 19 subcells and illustration of the Pilot reuse strategy by supposing $B = 8$ and $K = 11$.

$$g_{\langle j,k\rangle,i,m} = \beta^{1/2}_{\langle j,k\rangle,i} h_{\langle j,k\rangle,i,m}, \tag{2}$$

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

$$\mathbf{g}_{\langle j,k\rangle,i} = \beta^{1/2}_{\langle j,k\rangle,i}[h_{\langle j,k\rangle,i,1}, h_{\langle j,k\rangle,i,2}, ..., h_{\langle j,k\rangle,i,M}]^T, \tag{3}$$

By assuming a synchronous pilot transmission from all the active users, the received signal matrix of pilots at the i^{th} RRH is expressed as [12], [17], [18]

$$\mathbf{y}_{p,i} = \sqrt{\rho_p}\sum_{j=1}^{N}\sum_{k\in\mathcal{K}_j}\mathbf{g}_{\langle j,k\rangle,i}\mathbf{v}_k^T + \mathbf{z}_{p,i} \in \mathbb{C}^{M\times B}, \tag{4}$$

where ρ_p is the transmit power for pilots. $\mathbf{v}_k \in \mathbb{C}^{B\times 1}$ is the pilot signal transmitted by the k^{th} user. In addition, $\mathbf{z}_{p,i} \in \mathbb{C}^{M\times B}$ is complex additive white Gaussian noise with zero mean and unit covariance variables.

The CU has the following observation of the channel from user $\langle j,k\rangle$ to the i^{th} RRH

$$\tilde{\mathbf{y}}_{p,\langle j,k\rangle,i} = \frac{1}{\sqrt{\rho_p}B}\mathbf{y}_{p,i}\mathbf{v}_k^* \in \mathbb{C}^{M\times B}. \tag{5}$$

III. Sum-rate Analysis

In some cases, there is a probability that K is larger than the number of available pilot sequences (B). As a result, it is proposed in this paper to allow pilot reuse among some users within the same cell in order to serve more users when K is larger than B. Therefore, a dynamic pilot reuse strategy is suggested to permit a pair of users to share one pilot, meanwhile different pilots are assigned for different pairs of these users. An algorithm will be presented and explained later to determine the selection of user pairs for pilot reuse. For the rest of the users, no pilot reuse is considered, and unique pilot sequences will be allocated to serve them. This is shown in Fig. 1, in which it is considered that $B = 8$ and $K = 11$. This means that $K-B = 3$, reusing three pilots is required to serve all the users in Fig. 1. The users with unique pilot sequences are illustrated with the black circles, and the pairs of users with reused pilot sequences are represented by the same figures (square, triangle and star), where each pair is marked with the same color (blue, red and green). 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 of these users, the result of (5) for any user $\langle j,k\rangle$ should be calculated [12], [18]

$$\tilde{\mathbf{y}}_{p,\langle j,k\rangle,i} = \sum_{k\in\mathcal{I}} \mathbf{g}_{\langle j,k\rangle,i} + \frac{\mathbf{z}_{p,i}\mathbf{v}_k^*}{\sqrt{p_p}B} \;\; \forall j,k,i, \tag{6}$$

where $\mathcal{I}$ is the set that includes user $\langle j,k\rangle$ and the second user sharing the same pilot with user $\langle j,k\rangle$, or it includes only user $\langle j,k\rangle$ in the case of no reuse.

Based on [19], CU computes the minimum mean squared error (MMSE) estimation of user $\langle j,k\rangle$

$$\hat{\mathbf{g}}_{\langle j,k\rangle,i} = \beta_{\langle j,k\rangle,i}(\sum_{k\in\mathcal{I}} \beta_{\langle j,k\rangle,i} + \frac{1}{\sqrt{\rho_p}B})^{-1}\tilde{\mathbf{y}}_{p,\langle j,k\rangle,i} \;\; \forall j,k,i, \tag{7}$$

where $\hat{\mathbf{g}}_{\langle j,k\rangle,i} \sim \mathcal{CN}(0,\Phi_{\langle j,k\rangle,i}\mathbf{I}_M)$, and $\Phi\langle j,k\rangle,i$ is the same between user $\langle j,k\rangle$ and all the antennas of any RRH, and it can be defined as [9], [19]

$$\Phi_{\langle j,k\rangle,i} = \beta^2_{\langle j,k\rangle,i}(\sum_{k\in\mathcal{I}} \beta_{\langle j,k\rangle,i} + \frac{1}{\sqrt{\rho_p B}})^{-1} \;\; \forall j,k,i. \tag{8}$$

The channel $\mathbf{g}_{\langle j,k\rangle,i}$ can be decomposed as $\mathbf{g}_{\langle j,k\rangle,i} = \hat{\mathbf{g}}_{\langle j,k\rangle,i} + \tilde{\mathbf{g}}_{\langle j,k\rangle,i}$, where $\tilde{\mathbf{g}}_{\langle j,k\rangle,i}$ is the estimated error and it is statistically independent of $\hat{\mathbf{g}}_{\langle j,k\rangle,i}$ because of the orthogonality property of MMSE estimation and the joint Gaussianity of both vectors. The covariance of $\tilde{\mathbf{g}}_{\langle j,k\rangle,i}$ is $((\beta_{\langle j,k\rangle,i} - \Phi_{\langle j,k\rangle,i})\mathbf{I}_M)$.

During the uplink data transmission, the received data signal vector for the i^{th} RRH can be written as

$$\mathbf{y}_{d,i} = \sqrt{\rho_d}\sum_{j=1}^{N}\sum_{k\in\mathcal{K}_j} \mathbf{g}_{\langle j,k\rangle,i}x_k + \mathbf{z}_{d,i} \in \mathbb{C}^{M\times 1}, \tag{9}$$

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, and $\mathbf{x} = [x_1, x_2, ..., x_K] \in \mathbb{C}^{K\times 1}$ is the transmitted data signal vector of all the K users.

It is assumed that the MMSE detection is applied in CU for recovering user data, where the detector for user $\langle j,k\rangle$ at i^{th} RRH is given as [9], [19].

$$\hat{\mathbf{a}}_{\langle j,k\rangle,i} = (\hat{\mathbf{G}}\hat{\mathbf{G}}^H + \mathbf{Z} + \frac{1}{\rho_d}\mathbf{I}_N)^{-1}\hat{\mathbf{g}}_{\langle i,k\rangle,i}, \tag{10}$$

where the k^{th} column of $\hat{\mathbf{G}}$ is $\hat{\mathbf{g}}_{\langle j,k\rangle,i}$ and $\mathbf{Z} = \sum_{j=1}^{N}\sum_{k\in\mathcal{K}_j}\mathbb{E}(\tilde{\mathbf{g}}_{\langle j,k\rangle,i}\tilde{\mathbf{g}}^H_{\langle j,k\rangle,i})$ is the covariance matrix of the channel estimation errors.

It is assumed that users are only served by their nearest RRHs. For a pair of users that shares the same pilot, the sets of RRHs serving these two users should not overlap. This is essential to enable the CU to recognize the estimated channel for the users with pilot reuse, also the received signals from the nearest RRHs are more effective compared with distant RRHs. On this basis, only the CSIs from any user to the nearest RRHs are needed in order to detect the signal of that user. However, by referring to (10), it is still required to know the CSIs from user $\langle j,k\rangle$ to all other RRHs in the cell to detect the signals of these users by using MMSE. Therefore, the CSIs will be measured from any user without pilot reuse to all the RRHs, but only the nearest RRHs will be taken into account for data transmission. On the other side, for users with pilot reuse, two sets of RRHs will be considered. The first set, which is denoted by G_k, represents the group of RRHs serving user $\langle j,k\rangle$. This set is only considered for the purpose of detecting the signal of user $\langle j,k\rangle$. The second set, which is denoted by Y_k, represents the largest possible set of nearest RRHs serving user $\langle j,k\rangle$ so they do not overlap with the corresponding set of RRHs of the second user whose pilot is shared with user $\langle j,k\rangle$. As an example, to specify Y_k for the pair $\langle l,k_l\rangle$ and $\langle q,k_q\rangle$ of users shown in Fig. 1, both Y_{k_l} and Y_{k_q} are represented by the RRHs that are located on the right and the left sides of the green dashed line respectively. The aim of having Y_{k_l} is to provide $\mathbf{g}_{\langle l,k_l\rangle,i}, i \in Y_{k_l}$ for the detection purposes of the signals of all the users within the subcells of Y_{k_l}. However, during the detection of data signals of the users who are located within any subcell of Y_{k_l}, the interference from the users with pilot reuse who are located within the subcells of Y_{k_q} will be ignored as $\mathbf{g}_{\langle q,k_q\rangle,i}, i \in Y_{k_q}$ will not be available. This assumption can be acceptable as these users, within the subcells of Y_{k_q}, are located in far away subcells from the users of the subcells of Y_{k_l}.

Afterwards, the MMSE detection is applied in CU. The estimated data symbol for user $\langle j,k\rangle$ is represented by [2]

$$\hat{\mathbf{x}}_k = \sum_{i\in G_k} \hat{\mathbf{a}}^H_{\langle j,k\rangle,i}\mathbf{y}_{d,i}, \tag{11}$$

Then, the ergodic achievable sum-rate for DM-MIMO system can be obtained from [9], [10]

$$\mathrm{R}_T = (1-\frac{B}{S})\sum_{j=1}^{N}\sum_{k\in\mathcal{K}_j}\log_2(1+\mathrm{SINR}_{\langle j,k\rangle}), \quad (12)$$

where the SINR for user $\langle n,k_n\rangle$, which represents a user who has unique pilot sequence, is described in (13). Additionally, the SINR for user $\langle l,k_l\rangle$, whose pilot is reused with user $\langle q,k_q\rangle$, is presented in (14).

In order to simplify the SINR in both (13) and (14), the deterministic equivalent analysis from [19] is employed here. The results in [19] are asymptotic as the SINR formulas are derived by considering that the number of antennas grows considerably. According to [19, Theorem 5], the SINR for MMSE beamformer is represented as shown in [19, equation (25)]. This formula is modified to match the single cell DM-MIMO scenario of this paper as shown in (15) and (16), for a user without pilot reuse and a user with pilot reuse respectively. Where the elements of the channel vector from any user to the antennas of the base station in [19] are replaced by the channel vector from the same user to all the antennas of all the RRHs, except for users with pilot reuse as mentioned before. In the denominator of both (15) and (16), the first term represents the power of noise and the second term is the sum power of other user interference and the third or the last term in (16) is the pilot contamination. By utilizing [19, Theorem 1 and Theorem 2] to calculate the parameters $\delta_{\langle j,k\rangle}, \mu_{\langle j,k\rangle}, \theta_{\langle j,k\rangle}$and $\mathbf{T}^{'}$. These parameters depend on $\mathbf{Z} = \sum_{j=1}^{N}\sum_{k=1}^{K}\mathbf{R}_{\langle\mathbf{j},\mathbf{k}\rangle} - \mathbf{\Phi}_{\langle\mathbf{j},\mathbf{k}\rangle}$. Where $\mathbf{R}_{\langle\mathbf{j},\mathbf{k}\rangle}$ is a diagonal matrix with the diagonal entries of $\beta_{\langle j,k\rangle,i,m}$, where $\beta_{\langle j,k\rangle,i,m} = \beta_{\langle j,k\rangle,i}$ for $m \in \mathcal{M}_i$, $m = 1,...,M$, and $m = 1,...,N$. The same for $\mathbf{\Phi}_{\langle\mathbf{j},\mathbf{k}\rangle}$, which is a diagonal matrix with the diagonal entries of $\phi_{\langle j,k\rangle,i,m}$, where $\phi_{\langle j,k\rangle,i,m} = \phi_{\langle j,k\rangle,i}$ for $m \in \mathcal{M}_i$, $m = 1,...,M$, and $m = 1,...,N$. By employing [19, Theorem 1 and Theorem 2] to calculate the parameters $\delta_{\langle j,k\rangle}, \mathbf{T}^{'}, \mu_{\langle j,k\rangle}$ and $\theta_{\langle j,k\rangle}$, which are all functions of $\mathbf{R}_{\langle\mathbf{j},\mathbf{k}\rangle}$ and $\mathbf{\Phi}_{\langle\mathbf{j},\mathbf{k}\rangle}$. The expressions of these parameters are not included here due to space limitation.

In this paper, the dynamic reuse strategy aims to allocate the pilots for users ($\mathcal{P}\langle i,k\rangle$) with reusing pilots. This problem can be formulated as

$$\max_{\mathcal{P}\langle i,\,k\rangle} \mathbb{E}\{\mathrm{R}_T\} \quad (17a)$$
$$s.t.\ \mathrm{SINR}_{\langle l,k_l\rangle\langle q,k_q\rangle} \geq \gamma_{th} \quad (17b)$$
$$\mathrm{SINR}_{\langle q,k_q\rangle\langle l,k_l\rangle} \geq \gamma_{th} \quad (17c)$$
$$|\mathcal{V}| = B \quad (17d)$$
$$G_{k_l} \cap G_{k_q} = \emptyset \quad (17e)$$
$$Y_{k_l} \cap Y_{k_q} = \emptyset \quad (17f)$$

where $p_{\langle j,k\rangle}$ is the pilot used by user $\langle j,k\rangle$.

Maximum achievable sum-rate is satisfied (17a) by applying the dynamic pilot reuse strategy. The reuse of the pilots between the two users k_l and k_q located in different subcells, l and q respectively, is considered. The constraints (17b) and (17c) mean that each user that shares its pilot should satisfy the minimum SINR requirement (γ_{th}). The threshold γ_{th} is chosen according to the minimum SINR requirements of the users. Although each user could have a different SINR requirement, one threshold value is considered for all users for simplicity purposes. In addition, constraint (17d) represents that only B pilot sequences are available, so K should be larger than B to apply the reuse. Furthermore, the number of times of reuse is s, where $1 < s < \eta$, η is $\min(N_p, B)$, and N_p equals $\frac{K}{2}$ if K is even or $(\frac{K}{2}-1)$ if K is odd. For the rest of $(K-2s)$ users, no reuse is considered. Finally, both (17e) and (17f) indicate that the two sets G_{k_l} and G_{k_q}, and the sets Y_{k_l} and Y_{k_q} must not be overlapped, as explained earlier.

IV. Algorithm for dynamic pilot reuse

The long distance gap that can be found among some subcells will be exploited to reuse a pilot between a pair of users each located in a different subcell. An algorithm is presented in this section to describe the procedures over which the selection of these pairs of users is achieved. The pseudocode of this algorithm is clarified in **Algorithm 1**. In this algorithm, a pilot can be reused once with another user located in a different subcell, and no pilot can be reused in the same subcell. The purpose of reusing pilots once is to restrict the intra-cell pilot contamination to one source for each user, especially as the reuse happens within the same cell. The two users in each pair should meet γ_{th}. No reuse will apply if no pair satisfies the required γ_{th}, and only B users will be served in this case. **Algorithm 1** should update the pilot allocation for users every time K changes or the location of the users changes.

The decision to select the pairs of users that able to share the same pilot can be summarized in two stages: In the first stage, a pair of users separated by maximum possible distance is first chosen. Then, the potential SINRs are calculated for the two users in this pair by supposing that the pilot for the first user is reused by the second user. If the results of the potential SINRs are not less than γ_{th}, then it is possible to allocate one pilot for these two users.

The full explanation of **Algorithm 1** can be described as follows:

1) *Initialization:* In line 1, an initialization of the pilot allocation is performed. The number of pairs of users that use the same pilots (v) is also initialized.
2) *User Pairs Classification:* A search is achieved in line 2 for all the possibilities of user pairs configurations in the cell, where each user in each pair should be located in different subcells. In addition, no user can be located in two different pairs. Afterwards, these possibilities of pairs are ordered in $[\theta_1, \theta_2, ..., \theta_f, ..., \theta_F]$ according to the separation distance between them, where θ_1 represents the pair with largest separation distance.
3) *Users Selection:* In line 4, the pair θ_1 is initially chosen, and the two users of this pair are considered as ($\langle l,k_l\rangle$ and $\langle q,k_q\rangle$).
4) *Pilot Reuse Condition:* For the pair of users ($\langle l,k_l\rangle$ and $\langle q,k_q\rangle$) selected in line 4, the potential values of $\mathrm{SINR}_{\langle l,k_l\rangle\langle q,k_q\rangle}$ and $\mathrm{SINR}_{\langle q,k_q\rangle\langle l,k_l\rangle}$ are calculated based on equation (16), which in turn depends on the large scale

$$\mathrm{SINR}_{\langle n,k_n\rangle}=\frac{\rho_d(\sum_{i\in G_{k_n}}\mathbf{a}^H_{\langle n,k_n\rangle,i}\hat{\mathbf{g}}_{\langle n,k_n\rangle,i})^2}{\rho_d(\sum_{i\in G_{k_n}}\mathbf{a}^H_{\langle n,k_n\rangle,i}\tilde{\mathbf{g}}_{\langle n,k_n\rangle,i})^2+\rho_d\sum_{j=1}^{N}\sum_{\substack{k\in\mathcal{K}_j\\ k\neq k_n}}(\sum_{i=1}^{N}\mathbf{a}^H_{\langle n,k_n\rangle,i}\mathbf{g}_{\langle j,k\rangle,i})^2+(\sum_{i=1}^{N}\mathbf{a}^H_{\langle n,k_n\rangle,i}\mathbf{z}_{d,i})^2} \tag{13}$$

$$\mathrm{SINR}_{\langle l,k_l\rangle\langle q,k_q\rangle}=\frac{\rho_d(\sum_{i\in G_{k_l}}\mathbf{a}^H_{\langle l,k_l\rangle,i}\hat{\mathbf{g}}_{\langle l,k_l\rangle,i})^2}{\rho_d(\sum_{i\in G_{k_l}}\mathbf{a}^H_{\langle l,k_l\rangle,i}\tilde{\mathbf{g}}_{\langle l,k_l\rangle,i})^2+\rho_d(\sum_{i\in G_{k_l}}\mathbf{a}^H_{\langle l,k_l\rangle,i}\mathbf{g}_{\langle q,k_q\rangle,i})^2+\rho_d\sum_{j=1}^{N}\sum_{\substack{k\in\mathcal{K}_j\\ k\neq k_l\\ k\neq k_q}}(\sum_{i=1}^{N}\mathbf{a}^H_{\langle l,k_l\rangle,i}\mathbf{g}_{\langle j,k\rangle,i})^2+(\sum_{i=1}^{N}\mathbf{a}^H_{\langle l,k_l\rangle,i}\mathbf{z}_{d,i})^2} \tag{14}$$

$$\mathrm{SINR}_{\langle n,k_n\rangle}=\frac{\delta^2_{\langle n,k_n\rangle}}{\frac{1}{\rho_d(NM)^2}\ \mathrm{tr}\ \mathbf{\Phi}_{\langle \mathbf{n},\mathbf{k_n}\rangle}\mathbf{T}'+\frac{1}{(\mathbf{NM})^2}\sum_{j=1}^{\mathbf{N}}\sum_{\substack{k=1\\ k\neq k_n}}\mu_{\langle \mathbf{n},\mathbf{k_n}\rangle}} \tag{15}$$

$$\mathrm{SINR}_{\langle l,k_l\rangle\langle q,k_q\rangle}=\frac{\delta^2_{\langle l,k_l\rangle}}{\frac{1}{\rho_d(NM)^2}\ \mathrm{tr}\ \mathbf{\Phi}_{\langle \mathbf{l},\mathbf{k_l}\rangle}\mathbf{T}'+\frac{1}{(\mathbf{NM})^2}\sum_{j=1}^{\mathbf{N}}\sum_{\substack{k=1\\ k\neq k_l}}\mu_{\langle \mathbf{l},\mathbf{k_l}\rangle}+|\theta_{\langle \mathbf{q},\mathbf{k_q}\rangle}|^{\mathbf{2}}} \tag{16}$$

Algorithm 1 Algorithm for dynamic pilot reuse

INPUT: System parameters: K, M, B, N, η and threshold value γ_{th};
Large scale fading coefficients $\beta_{\langle j,k\rangle,i}$.
OUTPUT: Pilot allocation: $\{\mathcal{P}\langle i,k\rangle\}$.

1: $\{\mathcal{P}\langle i,k\rangle\}=0, v=0, f=1.$
2: In descending order, sort all the possible configurations of users pairs into $[\theta_1,\theta_2,...,\theta_f,...,\theta_F]$ according to the separation distance.
3: **while** $v<\eta$ OR $f<F$ **do**
4:   **for** a pair of users ($\langle l,k_l\rangle$ and $\langle q,k_q\rangle$) with order θ_f **do**
5:     **if** $\mathrm{SINR}_{\langle l,k_l\rangle\langle q,k_q\rangle}\geq\gamma_{th}$ AND $\mathrm{SINR}_{\langle q,k_q\rangle\langle l,k_l\rangle}\geq\gamma_{th}$ **then**
6:       $p\langle l,k_l\rangle=p\langle q,k_q\rangle$.
7:       Increment number of reuse $v=v+1$.
8:       Increment pair order $f=f+1$.
9:       Go to line 4.
10:     **else**
11:       Increment pair order $f=f+1$.
12:       Go to line 4.
13:     **end if**
14:   **end for**
15: **end while**
16: **if** $v=\eta$ OR $f=F$ **then**
17:   Randomly allocate the rest of pilots.
18: **end if**

fading coefficients ($\beta_{\langle j,k\rangle,i}$). Also, (16) includes the effect of the estimation error on the SINR of users because of pilot reuse. These values of SINRs are compared with γ_{th}. If both of them are larger than or equal γ_{th}, one pilot will be allocated for the two users $p\langle l,k_l\rangle=p\langle q,k_q\rangle$, as shown in line 6. Later, a new pair of users is selected and tested (Steps 8-9 or 11-12).

5) *Pilot Reuse Termination:* As shown in line 3, the above procedures to reuse pilots will be repeated a number of times equal to η or when all the pairs of users are tested (f=F). The parameter v in line 7 counts how many pilots are reused. When v equals η or when all pairs of users are tested, reusing pilots will be terminated. It is necessary needed here to allocate the remaining pilots, if there is, randomly as shown in lines 16 and 17. For users with unique pilots, data rates can be calculated based on the SINR values computed from equation (15).

The worst case for **Algorithm 1** from the complexity aspect is to search over all the combinations of user pairs to find the pairs that can meet the pilot reuse condition. The number of pair combinations for K users is $\sum_{i=1}^{K-1} i$. This is less than the complexity of the pilot allocation algorithms of other works, such as [15]. Some examples are shown in Table 1 to compare the complexity of **Algorithm 1** with that of the algorithm of [15] by supposing minimum pilot reuse factor ($\tau=2$), which is the simplest case from the complexity side in [15].

TABLE I
COMPLEXITY COMPARISON

	$K=12$	$K=16$
Algorithm 1	66	120
Algorithm of [15]	156	272

V. SIMULATION RESULTS

The performance of the dynamic pilot reuse strategy in the uplink of single cell TDD DM-MIMO system is evaluated in this section. Some of the simulation parameters are listed in Table 2. The users are randomly distributed in the cell where each subcell contains a different number of users. The large scale fading coefficient $\beta_{\langle j,k\rangle,i}$ can be calculated as $\beta_{\langle j,k\rangle,i}=z_{\langle j,k\rangle,i}/(d_{\langle j,k\rangle,i}/r_o)^{\alpha}$, where $z_{\langle j,k\rangle,i}$ represents the shadow fading and it obeys a log-normal distribution, which is represented by, $10\log_{10}(z_{\langle j,k\rangle,i})\sim\mathcal{CN}(0,\sigma_{shad})$, and $d_{\langle j,k\rangle,i}$ is the distance between user $\langle j,k\rangle,i$ to the i_{th} RRH [6].

In Fig. 2, the uplink achievable sum-rate is plotted for the proposed reuse strategy with $B=8$, $S=32$, $M=20$, $|G_k|=2$ and K changes from 6 to 16. When $K=6$ or 8, no reuse is applied as $K\ngtr B$. While when $K=16$, all the 8 pairs of users could reuse their pilots depending if they meet the reuse condition. For comparison purposes, the no reuse case

TABLE II
SIMULATION PARAMETERS

Parameter	Value
Number of subcells or RRHs N	19
Cell radius R	2000m
Normalize radius r_o	100m
Decay exponent α	3
Shadow fading standard deviation σ_{shad}	8 dB
Transmit power for pilots ρ_p	40 dBm
Transmit power for data ρ_d	40 dBm

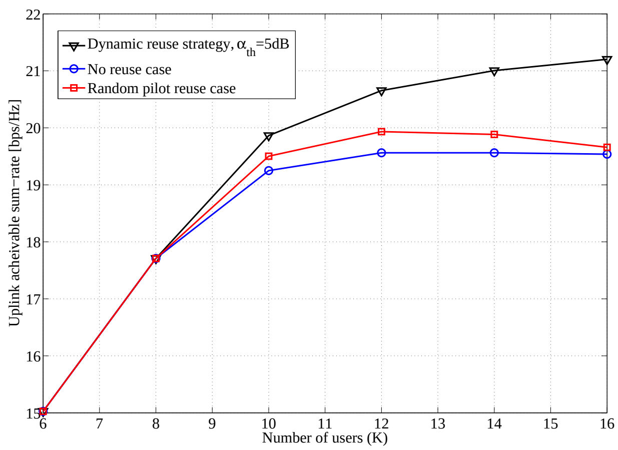

Fig. 2. Achievable uplink sum-rate versus the number of users.

and the random pilot reuse case are also shown in Fig. 2. In the no reuse case, unique pilot is allocated for each user, this means $B = K$ is assumed in all values of K. Furthermore, pilot reuse is applied randomly in the case of random pilot reuse, where B is also constant as in the proposed strategy. The same value of the number of times of reuse of the dynamic strategy is considered for random pilot reuse case as well. However, random pilot allocation is performed for random pilot reuse without taking into account the separation distance between the two users in each pair, and without applying the SINR constraints. As shown, the achievable sum-rate of the dynamic reuse strategy outperforms the no reuse case and random pilot reuse case. This is because of the effectiveness of the proposed reuse strategy and also because of the prelog factor $(1 - \frac{B}{S})$ in (12), which is constant for the dynamic pilot reuse case for any value of K since B is constant as well. On the other hand, the prelog factor $(1 - \frac{B}{S})$ is reduced as K increases in the no reuse case, where a greater part of the transmission frame is employed for the pilot signalling and a smaller part is allocated for data transmission. Although B is also constant for the random pilot reuse case, the random reuse for pilots can lead to severe intra-cell pilot contamination. This is particularly occurred when a pair of users with low separation distance shares the same pilot, and reduces the achievable sum-rate performance.

Fig. 3 shows the uplink achievable sum-rate for the dynamic reuse strategy versus the number of users with $S = 32$, $M = 20$ and for various values of SINR thresholds γ_{th} = 1dB, 5dB

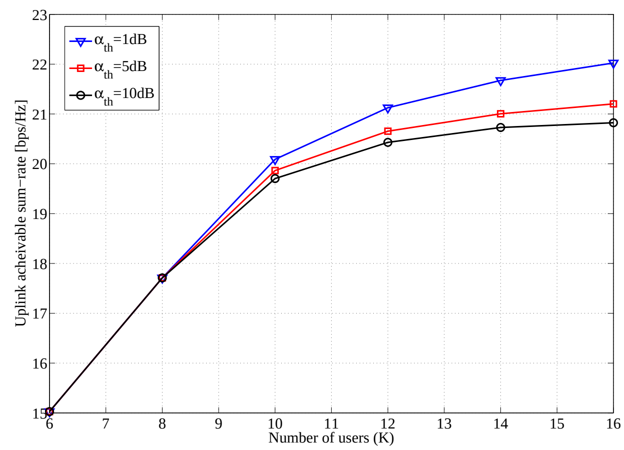

Fig. 3. Uplink achievable sum-rate versus the number of users at different values of threshold γ_{th}.

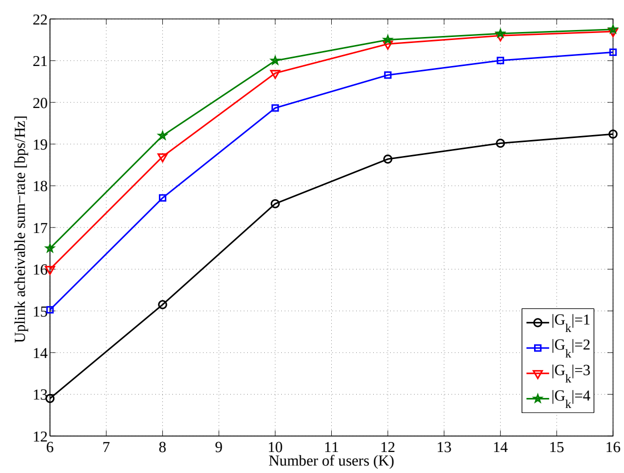

Fig. 4. Uplink achievable sum-rate versus the number of users at different various lengths of G_k.

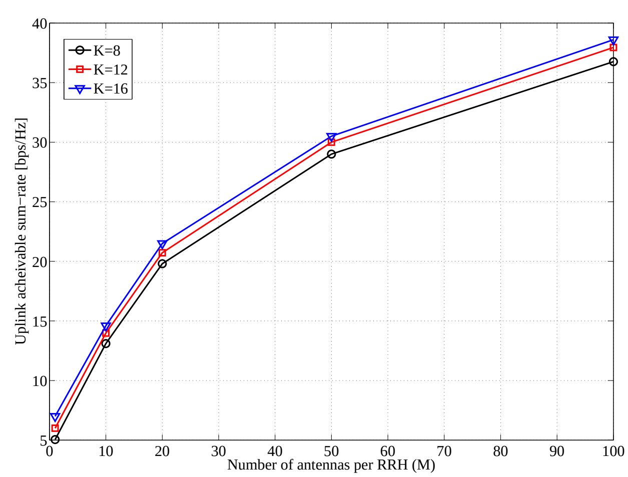

Fig. 5. Uplink achievable sum-rate versus the number of antennas per RRH (M) .

and 10dB. The achievable sum-rate of the dynamic pilot reuse strategy is better with less values of γ_{th} as more pairs can easily satisfy the small values of γ_{th}. In the opposite, few pairs of users can meet the high values of γ_{th}. This indicates that through γ_{th}, a trade-off can be performed between the number of users that can be served and pilot interference.

Fig. 4 plots the uplink achievable sum-rate for the dynamic reuse strategy versus the number of users with $S = 32, M = 20$ and for various lengths of $G_k = 1, 2, 3$ and 4. The rate performance of dynamic reuse strategy is generally improved as $|G_k|$ increases, where more of the nearest RRHs will serve the users. Thus, more uplink signals for these users will be coherently detected by the CU. However, when $|G_k|$ is highly increased, as when $|G_k| = 3$ or 4, the two sets of G_k of a pair of users that shares a single pilot will be closer to each other. Hence, more pilot contamination will be received from close RRHs, and this will negatively affect the sum-rate performance especially when high pilot reuse is needed when K is high.

Finally, the uplink achievable sum-rate for the dynamic reuse strategy is shown in Fig. 5, by supposing B = 8 and S = 32 with $M = 1, 10, 20, 50$ and 100, and for three cases ($K = 8, 12$ and 16). It is obvious that the sum-rate performance is continuously improved by increasing the number of antennas. This indicates that it is possible to reuse more pilots if larger number of antennas is utilized.

VI. Conclusion

In this paper, a dynamic pilot reuse strategy in a single cell TDD DM-MIMO system was proposed. Firstly, an achievable uplink sum-rate expression was derived for any pair of users who share the same pilots. After that, an algorithm was suggested to decide which pairs of users have acceptable potential data rate levels in order to allocate the same orthogonal pilot for them. The proposed strategy was shown to have better uplink achievable sum-rate performance compared with the no reuse case and random pilot reuse case.

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] D. Qiao, Y. Wu, and Y. Chen, "Massive MIMO architecture for 5G networks: Co-located, or distributed?" in *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, Aug. 2014, pp. 192–197.

[4] H. Zhu and J. Wang, "Radio resource allocation in multiuser distributed antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 10, pp. 2058–2066, Oct. 2013.

[5] D. Wang, J. Wang, X. You, Y. Wang, M. Chen, and X. Hou, "Spectral efficiency of distributed MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 10, pp. 2112–2127, Oct. 2013.

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

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

[8] V. Saxena, "Pilot contamination and mitigation techniques in massive MIMO systems," *Ph.D. dissertation, Department of Electrical and Information Technology, Lund University, Lund, Sweden,*, Feb. 2014.

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

[10] Q. He, L. Xiao, X. Zhong, and S. Zhou, "Increasing the sum-throughput of cells with a sectorization method for massive MIMO," *IEEE Communications Letters*, vol. 18, no. 10, pp. 1827–1830, Oct. 2014.

[11] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[12] X. Yan, H. Yin, M. Xia, and G. Wei, "Pilot sequences allocation in TDD massive MIMO systems," in *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2015, pp. 1488–1493.

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

[14] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3226–3239, 2012.

[15] Q. Sun, J. Wang, S. Jin, C. Xu, X. Gao, and K. K. Wong, "Rate analysis and pilot reuse design for dense small cell networks," in *2015 IEEE International Conference on Communications (ICC)*, Jun. 2015, pp. 2845–2850.

[16] 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 (ICC)*, May 2016, pp. 1–7.

[17] E. Bjrnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293–1308, Feb. 2016.

[18] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO : Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[19] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.