Journal of Electrical Engineering, Vol. 77, No. 2, 2026, pp. 105-112
https://reference-global.com/journal/JEE

Pilot-based clustering method in multi-ARS cell-free MIMO systems

Hai-Nam Le, Be Thi Bich Hoai, Pham Thanh Hiep, Nguyen Thu Phuong

Cell Free (CF) model, with the advantages of improved performance, high load capacity, and uniform communication quality, is a potential solution for 5G and B5G networks. Furthermore, the combination of the CF and advanced Unmanned Aerial Vehicle (UAV) technology further enhances the overall service quality and flexibility of the system. In this work, we examine the Multi-ARS CF model, where a large number of UAVs acting as Aerial Relay Stations (ARSs) are coordinated by the Ground Base Station (GBS) and cooperate to serve a large number of Mobile Stations (MSs) within the same frequency and time resources. The communication protocol utilizes the time division duplex (TDD) mechanism, with the implementation of the Minimum Mean Square Error (MMSE) method for estimating the uplink channel. A closed-form expression for MS throughput is derived. Moreover, we propose a pilot assignment algorithm at the master ARSs focusing on minimizing "pilot contamination" in the MS-master ARS communication link. Additionally, a clustering algorithm based on pilots is implemented. The system's performance is assessed through the Cumulative Distribution Function (CDF) of the MS throughput and compared with other pilot assignment algorithms, such as random and the Greedy algorithm. The results reveal that the pilot assignment and clustering algorithm successfully address the pilot contamination issue and enhance the overall system performance.

Keywords: cell-free, multi-ARS, clustering, pilot assignment, greedy algorithm, minimum mean square error

1 Introduction

The rapid expansion of the Internet of Things (IoT) and Artificial Intelligence (AI) necessitates wireless networks with significant capacity and ubiquitous connectivity. While conventional cellular networks deliver high speeds at cell centers, they often suffer from performance degradation at the cell edge due to path loss, shadowing, and inter-cell interference [1]. To address speed inconsistency, distributed network architectures like the Cell-Free (CF) model have emerged as viable solutions [2]. Introduced in [3] and inspired by distributed systems [4], the CF model eliminates cell boundaries and associated limitations. In CF systems, a large number of distributed Access Points (APs) jointly serve Mobile Stations (MSs), ensuring uniform service quality and mitigating edge-user issues [5]. Integrating terrestrial and non-terrestrial networks further enhances this potential [6].

Concurrently, UAVs acting as Aerial Relay Stations (ARSs) offer flexible deployment and superior Line-of-Sight (LoS) conditions, particularly in remote or post-disaster scenarios where ground infrastructure is compromised [7, 8]. Compared to terrestrial networks, ARSs provide cost-effectiveness and efficient coverage [9, 10]. Operational aspects of UAVs, including deployment and channel modeling, have been extensively reviewed in [11]. While CF-mMIMO [3] and UAV communications [10] have been studied separately, their integration remains underexplored, particularly in the context of practical clustering and pilot assignment [11]. However, the realization of UAV-aided cell-free networks faces two primary challenges. First, "pilot contamination" arises when the number of users exceeds the available orthogonal pilot sequences, leading to severe inter-user interference during channel estimation. Second, the standard cell-free approach, where every ARS serves every user, results in an excessive computational and backhaul data load. Unlike prior works that address these issues in isolation, this paper proposes a unified framework tackling both.

To address these challenges, we propose a novel Multi-ARS CF model utilizing multiple UAVs as ARSs operating as a unified network. Our main contributions include: (i) designing a channel model following ITU/3GPP guidelines [12] and utilizing MMSE for channel estimation; (ii) proposing a pilot assignment algorithm at master ARSs and a clustering method to minimize interference; and (iii) evaluating performance against random and Greedy algorithms [3], demonstrating superior throughput and reliability.

The paper is organized as follows: Section 2 describes the system model. Section 3 details the proposed algorithms. Section 4 presents numerical results, and Section 5 concludes. Table 1 lists key notations.

Le Quy Don Technical University, 236 Hoang Quoc Viet, Hanoi, Vietnam
namlh@lqdtu.edu.vn, bichhoaibt@gmail.com, phamthanhhiep@gmail.com, phuong.nt@mta.edu.vn

https://doi.org/10.2478/jee-2026-0011, Print (till 2015) ISSN 1335-3632, On-line ISSN 1339-309X
© This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Table 1. Mathematical symbols in this paper

Notation	Description
$(.)^*$	Conjugate
$(.)^H$	Conjugate-transpose
$\mathbb{E}\{\cdot\}$	Expectation operator
$\|.\|$	Euclidean norm
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbb{C}^{m\times n}$	Vector or matrix with size of $m \times n$
ρ^p, ρ^u, ρ^d	Normalized pilot, uplink, and downlink transmit powers
$\eta_k^{\mathrm{p}}, \eta_k^{\mathrm{u}}, \eta_k^{\mathrm{d}}$	Pilot, uplink data and downlink data power control coefficient of kth MS
K	Number of MSs
M	Number of antennas at ARS
A	Number of ARSs

2 System model

We consider a multi-ARS CF system with A ARSs serving K terrestrial MSs ($A > K$) in a large area (Fig. 1). ARSs use Decode-and-Forward (DF) and are coordinated by a GBS via ideal backhaul. The system operates in TDD mode with coherence interval $\tau^c = T^c \times B^c$ divided into uplink training, uplink data, and downlink data phases (Fig. 2). It is noted that the MSs are assumed to be static or slow-moving, and the ARSs hover during operation. Although UAVs may experience vibrations, the resulting Doppler spread is negligible within the short coherence interval. However, for high-mobility scenarios or moving ARSs, the Doppler effect becomes significant and should be included in future studies.

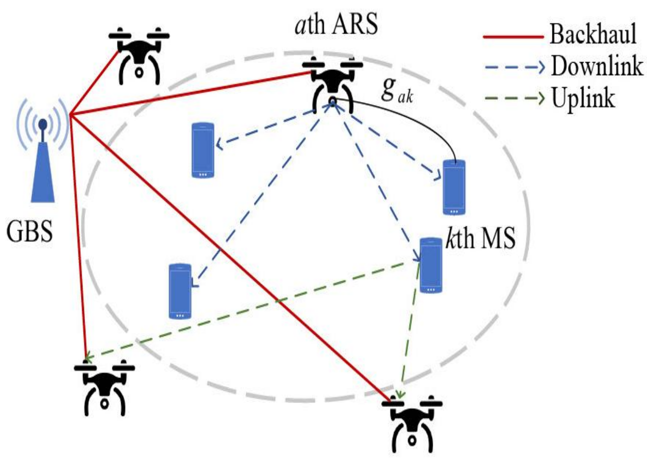

Fig. 1. The system model of multi-ARS SC

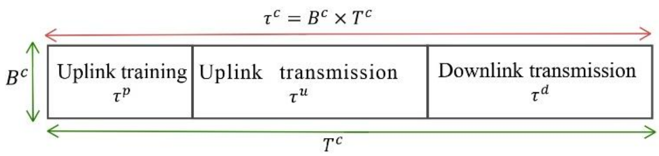

Fig. 2. Time-division duplex (TDD) protocol

The channel vector $\mathbf{g}_{ak} \in \mathbb{C}^{M\times 1}$ includes large-scale fading β_{ak} and small-scale fading $\mathbf{h}_{ak}$ [13] as

$$\mathbf{g}_{ak} = \sqrt{\beta_{lk}}\mathbf{h}_{ak}, \quad \beta_{ak} = 10^{\frac{\mathrm{PL}_{ak}}{10}}, \tag{1}$$

where PL_{ak} follows the 3GPP UMi scenario for aerial vehicles [12], and is described in Eqn. (2).

$$\mathrm{PL}_{ak} = \begin{cases} \mathrm{PL}_{\mathrm{LoS}} = \max\{\mathrm{PL}', 30.9 + (22.25 - 0.5\log_{10}h_{\mathrm{ARS}})\log_{10}d_{3\mathrm{D}} + 20\log_{10}f_c\}, \\ \mathrm{PL}_{\mathrm{NLoS}} = \max\{\mathrm{PL}_{\mathrm{LoS}}, 32.4 + (43.2 - 7.6\log_{10h_{\mathrm{ARS}}})\log_{10}d_{3\mathrm{D}} + 20\log_{10}f_c\} \end{cases} \tag{2}$$

2.1 Uplink training and data transmission

In the training phase, ARSs estimate channels using MMSE based on received pilots $\mathbf{y}_a^p$ by

$$\mathbf{y}_a^p = \sqrt{\tau^p\rho^p}\sum_{k=1}^K \mathbf{g}_{ak}\sqrt{\eta_k^p}\boldsymbol{\phi}_k^H + \mathbf{w}_a^p, \tag{3}$$

where ρ^p and η_k^{p} are the normalized pilot transmit power and pilot power control coefficient of kth MS, respectively.

The estimated channel $\hat{\mathbf{g}}_{ak}$ and its mean squared value γ_{ak} are

$$\hat{\mathbf{g}}_{ak} = c_{ak}\breve{\mathbf{y}}_{ak}^p, \quad \gamma_{ak} = \sqrt{\tau^p\rho^p\eta_k}\beta_{ak}c_{ak}, \tag{4}$$

where c_{ak} depends on pilot contamination.

During uplink data transmission, ARSs detect signals using matched filtering. The GBS combines signals to obtain $\mathbf{r}_k^{\mathrm{u}}$. The achievable rates are derived using the 'Use-and-then-Forget' (UatF) bound. It is important to note that in the conventional full-coordinate CF model described by Eqns. (5) and (6), the summation is performed over all ARSs ($a = 1, \ldots, A$). The achievable uplink rate is derived as in Eqn. (5), where ρ^u and η_k^{u} respectively represent the normalized uplink transmit power and uplink data power control coefficient of kth MS [7, 11].

2.2 Downlink transmission

Using conjugate beamforming, ARSs transmit

$$\mathbf{x}_a = \sqrt{\rho^d}\sum_{k=1}^{K}\sqrt{\eta_{ak}^{\mathrm{d}}}\,\hat{\mathbf{g}}_{ak}^{*}q_k.$$

The downlink rate is represented in Eqn. (6), where ρ^d and η_k^{d} denote the normalized downlink transmit power and downlink data power control coefficient of kth MS, respectively [8].

$$R_k^{\mathrm{u}} = \log_2\left(1+\frac{M^2\rho^{\mathrm{u}}\eta_k^{\mathrm{u}}\left(\sum_{a=1}^{A}\gamma_{ak}\right)^2}{M^2\rho^{\mathrm{u}}\sum_{k'\neq k}^{K}\eta_{k'}^{\mathrm{u}}\|\boldsymbol{\Phi}_k^H\boldsymbol{\Phi}_{k'}\|^2\left(\sum_{a=1}^{A}\gamma_{ak}\sqrt{\frac{\eta_{k'}\beta_{ak'}}{\eta_k\,\beta_{ak}}}\right)^2+M\rho^{\mathrm{u}}\sum_{k'=1}^{K}\eta_{k'}^{\mathrm{u}}\sum_{a=1}^{A}\gamma_{ak}\beta_{ak'}+M\sum_{a=1}^{A}\gamma_{ak}}\right) \tag{5}$$

$$R_k^{\mathrm{d}} = \log_2\left(1+\frac{M^2\rho^{\mathrm{d}}\left(\sum_{a=1}^{A}\sqrt{\eta_{ak}^{\mathrm{d}}}\gamma_{ak}\right)^2}{M^2\rho^{\mathrm{d}}\sum_{k'\neq k}^{K}\|\phi_{k'}^H\phi_k\|^2\left(\sum_{a=1}^{A}\sqrt{\eta_{ak'}^{\mathrm{d}}}\gamma_{ak'}\frac{\sqrt{\eta_k\,\beta_{ak}}}{\sqrt{\eta_{k'}\beta_{ak'}}}\right)^2+M\rho^{\mathrm{d}}\sum_{k'=1}^{K}\sum_{a=1}^{A}\eta_{ak'}^{\mathrm{d}}\gamma_{ak'}\beta_{ak}+1}\right) \tag{6}$$

3 Pilot assignment and clustering algorithm

In the initial CF model, the fact that all ARSs serve the MS contributes to increasing the throughput. However, the mechanism also makes transmission time increase due to the ARSs located far from the MSs. Furthermore, certain ARSs contribute insignificantly to the signal power while generating a heavy backhaul load. This situation can increase the risk of backhaul overload, especially under limited backhaul channel conditions. Limiting the number of ARSs serving an MS may slightly degrade the achievable throughput; however, an appropriate subset selection strategy can render this degradation negligible.

3.1 Pilot assignment

Assume that the kth MS is only served by the ARS in the subset $\mathcal{D}_k \subset \{1,\ldots,A\}$, $k=1,\ldots,K$. In the TDD protocol, the pilots with a length of τ^{p} samples are transmitted for channel estimation. Each MS should use a unique orthogonal pilot to avoid interference during the uplink training phase, known as "pilot contamination" [14]. However, since the pilots are τ^{p}-dimensional vectors, we can only find a set of maximum τ^{p} mutually orthogonal sequences $\boldsymbol{\Phi}_{t_1},\ldots,\boldsymbol{\Phi}_{\tau^{\mathrm{p}}} \in \mathbb{C}^{\tau^{\mathrm{p}}}$. Therefore, in practical systems with $K > \tau^{\mathrm{c}}$, multiple MSs must be allocated with identical pilots.

We denote the index of the pilot assigned to MS k as $t_k \in \{1,\ldots,\tau^{\mathrm{p}}\}$ and $\mathcal{P}_k$ as the set of the MSs assigned the same pilot as MS k as follows:

$$\mathcal{P}_k = \{i: t_i = t_k, i = 1,\ldots K\} \subset \{1,\ldots,K\} \tag{7}$$

The product, when multiplying the received pilot signal by the known pilot in the uplink training phase, is

$$\begin{aligned}\breve{\mathbf{y}}_{ak}^{p} &= \boldsymbol{\Phi}_{t_k}^{H}\mathbf{y}_a^{p}\\ &= \sqrt{\tau^p\rho^p\eta_k^p}\mathbf{g}_{ak} + \sqrt{\tau^p\rho^p}\sum_{k'=1}^{K}\sqrt{\eta_{k'}^{p}}\mathbf{g}_{ak'}\boldsymbol{\Phi}_{t_k}^{H}\boldsymbol{\Phi}_{t_{k'}} + \mathbf{w}_a^p\boldsymbol{\Phi}_{t_{k'}}^{H}\\ &= \sqrt{\tau^p\rho^p\eta_k^p}\mathbf{g}_{ak} + \sqrt{\tau^p\rho^p}\sum_{\substack{k'\in\mathcal{P}_k\\ k'\neq k}}^{K}\sqrt{\eta_{k'}^{p}}\mathbf{g}_{ak'} + \mathbf{w}_a^p\boldsymbol{\Phi}_{t_k}^{H}\end{aligned} \tag{8}$$

Removing the power control coefficients, $\eta_k^p = 1$ $\forall k$, we obtain

$$\breve{\mathbf{y}}_{ak}^{p} = \sqrt{\tau^p\rho^p}\mathbf{g}_{ak} + \sqrt{\tau^p\rho^p}\sum_{\substack{k'\in\mathcal{P}_k\\ k'\neq k}}^{K}\mathbf{g}_{ak'} + \mathbf{w}_a^p\boldsymbol{\Phi}_{t_k}^{H} \tag{9}$$

The second component in Eqn. (9) introduces interference due to pilot reusing and leads to errors in the channel estimations.

There are $(\tau^{\mathrm{p}})^K$ possible assignments in the system with K MSs and τ^p pilots. Finding an optimal assignment scheme for a large number of MSs requires significant computational effort. Therefore, allocating pilots based on local sub-optimal strategies is more practical.

In this work, a pilot assignment algorithm executed at the master ARSs, ARSs with the largest β_{ak}, will be presented. With the design perspective that the master ARSs contribute the most to the signal power of the MSs, the pilot assignment method allows the master ARSs to choose the pilot with the least interference for the MS. Specifically, the master ARS selects the pilot sequence by calculating the sum of β from itself to the MSs that have been assigned pilots previously. Then, the pilot with the smallest value is chosen to be assigned to the MS.

It is assumed that due to the short duration of the coherence interval ($T_c \approx 1$ ms), the physical displacement of UAVs is negligible; thus, the large-scale fading coefficients β_{ak} are considered constant during this period, ensuring that the Master ARS selection remains stable within each coherence interval. In the rare event of identical path losses, the ARS with the lowest index is selected.

The algorithm is carried out through two basic steps as follows.

- ***Step 1:*** Initially, the first τ^p MSs are sequentially assigned the τ^{p} pilot sequences from the orthogonal pilots set.

- ***Step 2:*** Assign pilots to $(K-\tau^{\mathrm{p}})$ remaining MSs. Specifically, from the $(\tau^{\mathrm{p}}+1)$th MS to the Kth MS, each MS sequentially reuses the pilots that minimize the occurrence of "pilot contamination".

- The MS determines the master ARS as the ARS with the best channel conditions.

$$a_k = \underset{a\in\{1,\dots,A\}}{\arg\max}\ \beta_{ak}. \tag{10}$$

- Master ARS determine best pilot according to

$$\tau = \underset{t\in\{1,\dots,\tau^{\mathrm{p}}\}}{\arg\min} \sum_{\substack{i=1\\ t_i=t}}^{k-1} \beta_{a_k i}\,. \tag{11}$$

The algorithm creates the most favorable conditions for communication between the MS and its master ARS, which contributes the most to the useful signal power transmitted and received by the MS.

3.2 Clustering Algorithm

The clustering algorithm relies on the assumption that each ARS should serve only one MS per pilot. The algorithm organizes the MSs into groups, each group containing only the MSs assigned to different pilots. The strategy helps to minimize the pilot interference and improve the performance of the wireless communication system. Additionally, limiting the number of MSs that an ARS has to serve to τ^p reduces the computational complexity of channel estimation and signal processing. Thus, the communication process becomes simpler and more efficient.

For each pilot, the ARS selects the MS with the highest channel gain to serve.

$$i \leftarrow \underset{k\in\{1,\dots,K\}:t_k=t}{\arg\max}\ \beta_{ak}. \tag{12}$$

The algorithm is executed locally at each ARS, making the system more flexible and simpler. The pilot assignment and pilot-based clustering algorithm are demonstrated as follows:

Initialization: Set up the initial clusters. $\mathcal{D}_1,\dots,\mathcal{D}_K = \varnothing$

Assign pilot for τ^p first MSs: Assign τ^p pilots in orthogonal pilots set to τ^p MSs:

- $t_k \leftarrow k$ with $k = 1,\dots,\tau^{\mathrm{p}}$

Assign pilots to the remaining MSs: Reuse τ^p pilot for the remaining $(K-\tau^p)$ MSs, where $k=\tau^{\mathrm{p}}+1,\dots,K$:

- Determine the master ARS of MS k:
$$a_k = \underset{a\in\{1,\dots,A\}}{\arg\max}\ \beta_{ak}$$
- Find the pilot that causes the least pilot interference to the master ARS:
$$\tau = \underset{t\in\{1,\dots,\tau^{\mathrm{p}}\}}{\arg\min} \sum_{\substack{i=1\\ t_i=t}}^{k-1} \beta_{a_k i}$$
- Assign the determined pilot to the MS k: $\tau_k \leftarrow \tau$.

Cluster based on the assigned pilots: With each ARS a , $a=1,\dots,A$ and each pilot t , $t=1,\dots,\tau^p$:

- Determine the best MS assigned the pilot t to be served by ARS a:
$$i \leftarrow \underset{k\in\{1,\dots,K\}:t_k=t}{\arg\max}\ \beta_{ak}$$
- Add ARS a into the cluster of ARSs serving the MS i: $\mathcal{D}_i \leftarrow \mathcal{D}_i \cup \{a\}$.

End: Finish the algorithm and return the result of pilot $t_1,\dots,t_K$ and cluster $\mathcal{D}_1,\dots,\mathcal{D}_K$.

Disregarding power optimization, the formula for calculating the uplink transmission rate is determined as follows.

For the proposed clustering method, the k-th MS is served exclusively by the subset of ARSs $\mathcal{D}_k$ rather than the entire set of A ARSs. Therefore, by replacing the global summation $\sum_{a=1}^{A}$ with the cluster-specific summation $\sum_{a\in\mathcal{D}_k}$ we obtain the specific achievable uplink and downlink rates for the clustered system as shown in Eqns. (13) and (14).

$$R_k^{\mathrm{u}} = \log_2\left(1+\frac{M^2\rho^{\mathrm{u}}\left(\sum_{a\in\mathcal{D}_k}\gamma_{ak}\right)^2}{M^2\rho^{\mathrm{u}}\sum_{\substack{k'\in P_k\\ k'\neq k}}\left(\sum_{a\in\mathcal{D}_k}\gamma_{ak}\frac{\beta_{ak'}}{\beta_{ak}}\right)^2+M\rho^{\mathrm{u}}\sum_{k'=1}^{K}\sum_{a\in\mathcal{D}_k}\gamma_{ak}\beta_{ak'}+M\sum_{a\in\mathcal{D}_k}\gamma_{ak}}\right) \tag{13}$$

$$R_k^{\mathrm{d}} = \log_2\left(1 + \frac{M^2\rho^{\mathrm{d}}\left(\sum_{a\in\mathcal{D}_k}\gamma_{ak}\right)^2}{M^2\rho^{\mathrm{d}}\sum_{k'\neq k}^{K}\left\|\phi_{k'}^H\phi_k\right\|^2\left(\sum_{a\in\mathcal{D}_k}\gamma_{ak'}\sqrt{\frac{\beta_{ak}}{\beta_{ak'}}}\right)^2 + M\rho^{\mathrm{d}}\sum_{k'=1}^{K}\sum_{a\in\mathcal{D}_k}\gamma_{ak'}\beta_{ak}+1}\right) \tag{14}$$

The clustering algorithm limits the number of MSs that each ARS needs to serve to a fixed number, thus ensuring that there is no overload on the backhaul link, even when the total number of MSs becomes large. However, global optimality is not guaranteed in the decentralized pilot assignment and clustering at each ARS. Centralized strategies for pilot assignment and clustering can be implemented to achieve a higher level of global optimality.

A key advantage of the proposed algorithm is its low computational complexity. The Greedy algorithm requires iterative rate calculations with a complexity of $\mathcal{O}(I\cdot K^2\cdot A)$, where I is the number of iterations. In contrast, the proposed method relies on local sorting of channel gains, resulting in a significantly lower complexity of $\mathcal{O}(K\cdot A\cdot \log A)$, making it highly suitable for real-time applications involving mobile UAVs.

4 Numerical results and discussion

We evaluate the performance of the multi-ARS CF system by comparing the performance of the system using the proposed algorithm with other pilot assignment algorithm such as random and Greedy algorithms [3]. In the Greedy algorithm, the MSs are initially assigned pilots randomly. Then, the MS with the lowest transmission rate will be reassigned a pilot. The new pilot is the one that causes the "pilot contamination" and is determined by the following formula:

$$\sum_{a=1}^{A}\sum_{k'\neq k*}^{K}\beta_{ak'}\left|\phi_{k*}^H\phi_{k'}\right|^2, \tag{15}$$

where k^* is the worst MS. Pilot reassignment is done through several iterations, with the number of iterations determined by system requirements.

4.1 Parameters and setup

The simulation parameters used in the system are summarized in Tab. 2, which are consistently adopted in previously published studies such as [7, 8 and 11]. The ARSs are randomly distributed according to a continuous uniform distribution in a 3D space within a 1×1 km^2 area, and the height range is $22.5\ \mathrm{m}\leq h_{ARS}\leq 300$ m. The distance has been clearly expressed and calculated in the Path Loss equation of the channel model (LoS-NLoS). In our experiments, we have computed the CDF for the uplink throughput of each MS.

Table 2. Parameters of the system used for simulation

Parameter	Value
Carrier frequency	1.9 GHz
Bandwidth (B)	20 MHz
Coherence bandwidth (B_c)	200 kHz
Coherence time (T_c)	1 ms
ARS height range	$22.5\ \mathrm{m}\leq h_{ARS}$ ≤ 300 m
$\rho^{\mathrm{p}} = \rho^{\mathrm{u}} = \rho^{\mathrm{d}}$	100 mW
Coherence interval τ^c	200 samples
The number of antennas per ARS (M)	1 antenna
Area distribution of ARS and K users	1 km^2

4.2 Results and discussion

Figure 3 illustrates the system performance of a multi ARS CF system with $A = 30, K = 20$ and $\tau^p = 10$ using the proposed algorithm, random pilot assignment algorithm, and Greedy algorithm with 5 iterations. The pilot reuse ratio is relatively high, $K/\tau^{\mathrm{p}} = 2$. The solid line represents the minimum rate of the MSs, while the dashed line represents the average rate. In both assessments, the quality of the random pilot selection method is the worst. With $K/\tau^{\mathrm{p}} = 2$, on average, there are 2 MSs using the same pilot in the ideal case. Due to the randomness, the ratio may be higher, resulting in a significant level of "pilot contamination". The Greedy algorithm demonstrates superior performance in terms of minimum rate, whereas the pilot assignment at master ARS and the pilot-based clustering algorithm exhibit superiority when considering the average value. The reason is that the Greedy algorithm reassigns pilots by enhancing the rate of the worst MS. The approach involves identifying the MS with the lowest rate, then reallocating a new pilot to minimize "pilot contamination". Therefore, the Greedy algorithm greatly improves the minimum rate. However, when considering the entire network, the pilot assignment algorithm at the master ARS is more comprehensive by eliminating the "pilot contamination" level at the master ARS of all MSs in the network. Thanks to that, all MSs' rates are guaranteed to be relatively high, increasing the overall performance.

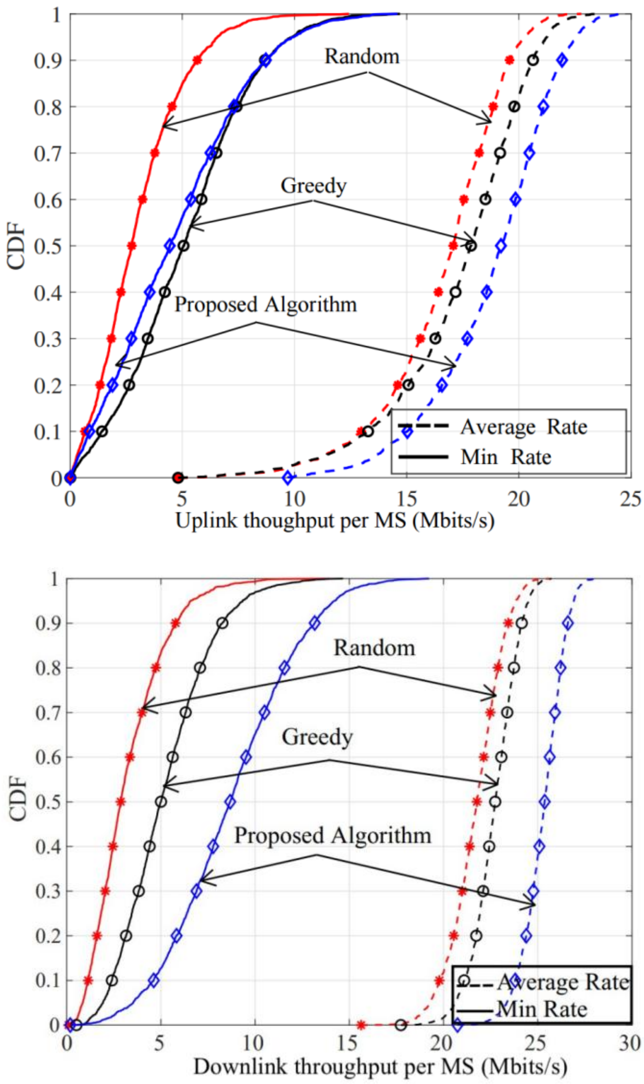

Fig. 3. CDFs of the minimum throughput and average throughput of the multi-ARS CF model using the proposed algorithm, random and Greedy algorithm, with $A = 30$, $K = 20$ and $\tau^{\mathrm{p}} = 10$

To further emphasize the impact of the pilot reuse ratio on the performance, we compare the CDF of the per-MS throughput of the systems using the proposed algorithm, random and Greedy algorithm with a higher reuse ratio, $K/\tau^{\mathrm{p}} = 4$, specifically $A = 30$, $K = 20$ and $\tau^{\mathrm{p}} = 5$ in Fig. 4. The enhancement in the system performance of the proposed algorithm is more prominently demonstrated in this scenario. The reason is that the proposed algorithm reduces the "pilot contamination" in the primary communication path of all MSs, ensuring that contamination remains within an acceptable range even if the pilot reuse ratio is high. In contrast, in the remaining two algorithms, the "contamination" level increases with the pilot reuse ratio. Although the Greedy algorithm can improve the min rates, it is not enough to significantly improve the network's overall performance.

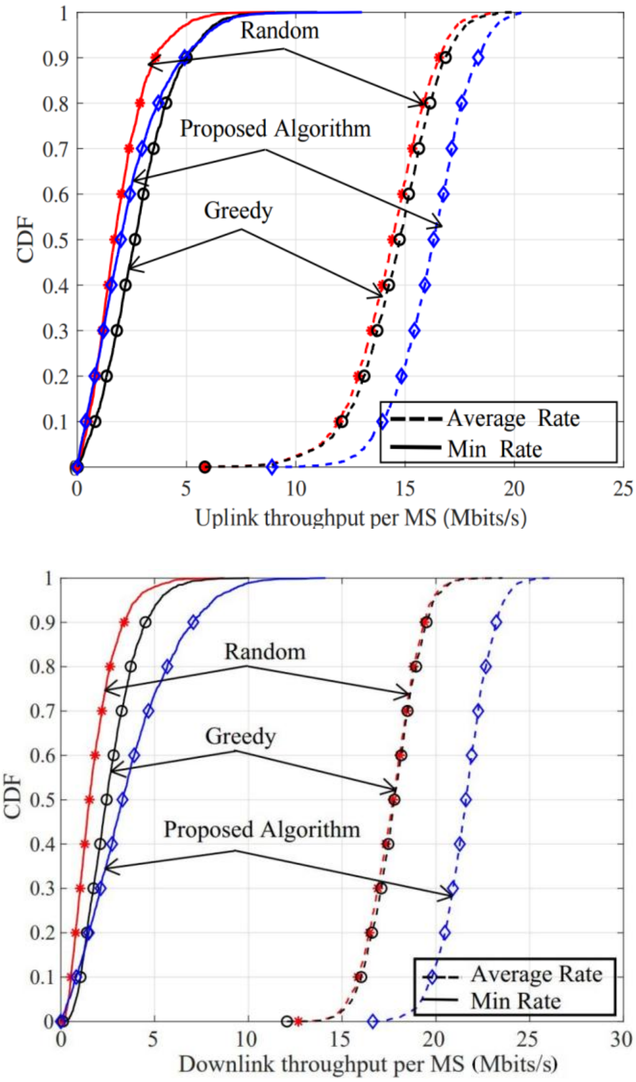

Fig. 4. CDFs of the minimum throughput and average throughput of the multi-ARS CF model using the proposed algorithm, random and Greedy algorithm, with $A = 30$, $K = 20$ and $\tau^{\mathrm{p}} = 5$

On the contrary, when the ratio $K/\tau^{\mathrm{p}} < 1$, i.e., the number of orthogonal pilots is enough to provide all MSs, the performance of the three pilot assignment algorithms becomes similar. In Fig. 5, the system performance of a Multi ARS CF system with $A = 30, K = 20$ and $\tau^{p} = 50$ is described. The performance of the random pilot assignment algorithm remains the worst due to its random nature. Although there are enough orthogonal pilots to allocate to all MSs, there is still a possibility of multiple MSs being assigned the same pilot, resulting in pilot interference. However, the probability is relatively low, so the difference in quality is not significant. The Greedy algorithm follows a similar approach by beginning with a random assignment, but Greedy demonstrates notable enhancement compared to the random algorithm. Although not too significant, the proposed algorithm shows enhanced quality in comparison to the other two methods.

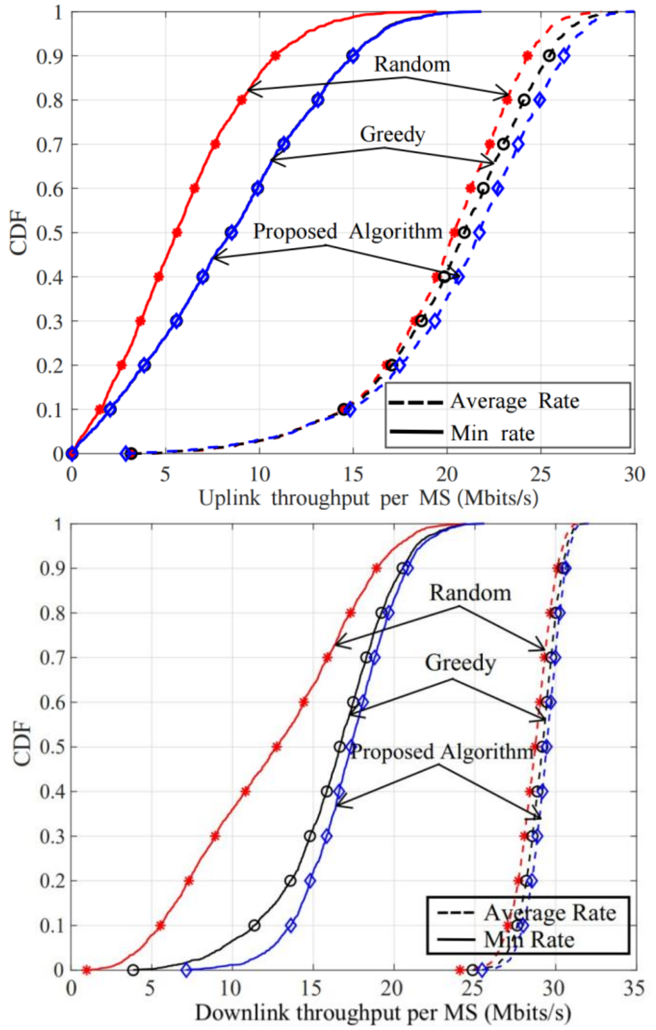

Fig. 5. CDFs of the minimum throughput and average throughput of the multi-ARS CF model using the proposed, random and Greedy algorithms, with $A = 30$, $K = 20$, and $\tau^{\mathrm{p}} = 50$

The reason is that Step 1 of the algorithm assigns entirely orthogonal pilots to all MSs without any random element involved. As a result, the system avoids any "pilot contamination", leading to significantly improved throughput.

The proposed algorithm also offers notable enhancements to the SC model. In Fig. 6, we illustrate the performance of different pilot assignment algorithms in the SC model with $A = 30$, $K = 20$, and $\tau^{\mathrm{p}} = 10$. The results obtained are similar to those of the typical CF model. The performance of the proposed algorithm remains efficient in the SC model, outperforming the other two methods.

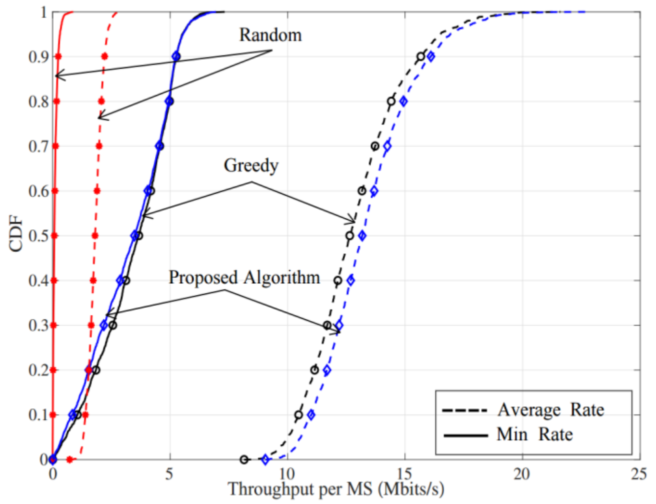

Fig. 6. CDFs of the minimum throughput and average throughput of the multi-ARS SC model using the proposed, random, and Greedy algorithms, with $A = 30$, $K = 20$, and $\tau^{\mathrm{p}} = 10$

5 Conclusion

In this paper, we have proposed a novel pilot-based clustering and assignment strategy for Multi-ARS Cell-Free systems to address pilot contamination and backhaul load challenges. By locally optimizing pilot reuse at the master ARS, the proposed method significantly mitigates interference while maintaining low computational complexity. Simulation results demonstrate that our approach consistently outperforms both Random and Greedy algorithms in terms of user throughput, especially under high pilot reuse conditions. Future work will also compare the proposed pilot-based clustering with other user-centric approaches, such as large-scale fading threshold-based clustering, to further evaluate trade-offs between interference mitigation and macro-diversity.

References

[1] H. Tataria, K. Haneda, A. F. Molisch, M. Shafi, and F. Tufvesson, "Standardization of propagation models for terrestrial cellular systems: A historical perspective," International Journal of Wireless Information Networks, vol. 28, pp. 20–44, 2021.

[2] A. Davis and H. Chang, "A survey of wireless sensor network architectures," International Journal of Computer Science and Engineering Survey, vol. 3, no. 6, p. 1, 2012.

[3] 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, 2017.

[4] S. Zhou, M. Zhao, X. Xu, J. Wang, and Y. Yao, "Distributed wireless communication system: a new architecture for future public wireless access," IEEE Communications Magazine, vol. 41, no. 3, pp. 108–113, 2003.

[5] J. Zheng, J. Zhang, H. Du, D. Niyato, B. Ai, M. Debbah, and K. B. Letaief, "Mobile cell-free massive mimo: Challenges, solutions, and future directions," IEEE Wireless Communications, 2024.

[6] J. Kassam, D. Castanheira, A. Silva, R. Dinis, and A. Gameiro, "A review on cell-free massive mimo systems," Electronics, vol. 12, no. 4, p. 1001, 2023.

[7] V. S. Nguyen, B. A. Duc, T. M. Hoang, X. N. Tran, P. T. Hiep, and N. T. Phuong, "Spatial pilot reassignment algorithm for channel estimation stage of cell-free multi-ars communication systems," Digital Signal Processing, vol. 156, p. 104830, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S105120042400455X

[8] B. A. Duc, T. M. Hoang, N. T. Phuong, X. N. Tran, and P. T. Hiep, "Power optimization and arss selection strategies in downlink cell-free multi-arss communication systems," Arabian Journal for Science and Engineering, vol. 50, no. 8, pp. 5647–5664, 2025. [Online]. Available: https://doi.org/10.1007/s13369-024-09441-4

[9] B. Alzahrani, O. S. Oubbati, A. Barnawi, M. Atiquzzaman, and D. Alghazzawi, "UAV assistance paradigm: State-of-the-art in applications and challenges," Journal of Network and Computer Applications, vol. 166, p. 102706, 2020.

[10] X. Li, Q. Wang, Y. Liu, T. A. Tsiftsis, Z. Ding, and A. Nallanathan, "UAV-aided multi-way NOMA networks with residual hardware impairments," IEEE Wireless Communications Letters, vol. 9, no. 9, pp. 1538–1542, 2020.

[11] B. D. Anh, T. M. Hoang, N. T. Phuong, X. N. Tran, and P. T. Hiep, "Pilot and data power optimization problems in multiple aerial relay stations cell-free communication systems," International Journal of Communication Systems, vol. 37, 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:269426328

[12] 3GPP, "Technical specification group radio access network; study on enhanced LTE support for aerial vehicles," TR 36.777, Dec 2017.

[13] S. Buzzi and C. D'Andrea, "Cell-free massive mimo: User-centric approach," IEEE Wireless Commun. Lett., vol. 6, no. 6, pp. 706–709, Aug. 2017.

[14] T. L. Marzetta, "How much training is required for multiuser mimo?" in 2006 fortieth asilomar conference on signals, systems and computers. IEEE, 2006, pp. 359–363.

Received 15 January 2026