Pilot Allocation and Assignment Optimization in User-Centric Distributed Massive MIMO Networks

Thaissa T. R. Ueoka, Michael A. S. da Costa, Daynara D. Souza, Marx M. M. Freitas, André L. P. Fernandes, Gilvan S. Borges, André M. Cavalcante, and João C. W. A. Costa

***Abstract*— Distributed massive multiple-input multiple-output (MIMO) networks, also known as cell-free, are promising solution to increase efficiency in beyond 5G systems. Pilot-based uplink (UL) channel estimation directly influences transmission efficiency, as it is used to mitigate interference and noise from user equipment (UEs). In this context, this work uses genetic algorithm (GA) as a tool to optimize pilot allocation and assignment and maximize spectral efficiency (SE). First, we define the optimal amount of samples allocated to channel estimation that balances accuracy and overhead. Generally, this lead to fewer pilots than UEs. Therefore, the pilot assignment is also optimized to decrease interference between UEs reusing the same pilot. The results show that the optimal number of pilots presents a similar behavior when the number of UEs increases. The average SE is improved when GA is used to optimize pilot assignment compared with the baseline solution.**

***Keywords*— Cell-free massive MIMO, channel estimation, genetic algorithm, pilot allocation and assignment.**

I. Introduction

User-centric (UC) distributed massive multiple-input multiple-output (D-mMIMO) system, also known as cell-free (CF), is considered a key solution for increasing the transmission efficiency of fifth-generation (5G) and beyond wireless networks. The system consists of a large number of radio units (RUs) distributed in the coverage area to serve the user equipments (UEs). The system can overcome the main disadvantages of cellular systems by increasing macro-diversity and providing uniform spectral efficiency (SE) [1], [2].

The canonical cell-free (CCF) version assumes that all RUs serve all UEs. However, this strategy is not scalable as it requires enormous resource requirements from the network. In this sense, the scalable cell-free (SCF) approach emerged as an alternative to solve these disadvantages. The RU selection process limits the number of UE that each RU can serve to achieve a scalable system [1]–[3].

In D-mMIMO, pilots are used to estimate the communication channel between RUs and UEs only once due to the channel response being constant in the coherence block. The uplink (UL) channel estimation will serve to downlink (DL) by applying the principle of reciprocity. Then, the channel is estimated in the UL direction to perform data combination and precoding. The length of the pilot sequences affects the channel estimation quality and the overhead, influencing the system's SE performance [1], [4].

Thaissa T. R. Ueoka[1], Michael A. S. da Costa[1], Daynara D. Souza[1], Marx M. M. Freitas[1], André L. P. Fernandes[1], Gilvan S. Borges[2], André M. Cavalcante[3] and João C. W. A. Costa[1]. [1]Applied Electromagnetism Laboratory, Federal University of Pará - UFPA, Belém, Brazil, [2]Federal Institute of Pará - IFPA, Belém, Brazil, [3]Ericsson Research, Ericsson Telecomunicações Ltda., Indaiatuba, Brazil. E-mails: thaissa.ueoka@itec.ufpa.br; michael.costa@itec.ufpa.br; daynara@ufpa.br; marx@ufpa.br; andrelpf@ufpa.br; gilvan.borges@ifpa.edu.br; andre.mendes.cavalcante@ericsson.com; jweyl@ufpa.br. This work was supported by Ericsson Telecomunicações Ltda, CNPq and CAPES.

Due to the limited resource for pilots, there is a need to reuse them if the number of UEs is greater than the number of orthogonal pilot sequences. The UEs that share the same pilot suffer from the effect of "pilot interference" or pilot contamination. This interference reduces the estimation quality and causes the channel estimations of the UEs that share the pilot to be correlated [5]. This has a critical impact beyond channel estimation, as pilot contamination makes it difficult to mitigate interference between UEs that use the same pilot in the UL and DL directions [6], [7]. Therefore, properly determining which pilot sequences are assigned to UEs can improve the transmission capacity, reduce interference between UEs and increase performance.

In this work, genetic algorithm (GA) optimization is used as a tool to find the pilot allocation and assignment that presents the best SE performance. By optimizing these two parameters, we can obtain better channel estimates and interference suppression, in addition to balancing the impact of the estimation overhead on SE. The results obtained with GA optimization is compared with baseline solutions in [1].

II. System Model

A. Channel Model

We consider a D-mMIMO system consisting of L RUs, each equipped with N antennas, serving K UEs. The RUs are connected to the central processing units (CPUs) through fronthaul links. In this scenario, the channel vector $\mathbf{h}_{kl} \in \mathbb{C}^{N\times 1}$ between RU l and UE k is modeled as an independent Rician channel, being defined as [8]

$$\mathbf{h}_{kl} = \sqrt{\frac{\kappa_{kl}}{1+\kappa_{kl}}}\mathbf{h}_{kl}^{\mathrm{LOS}} + \sqrt{\frac{1}{1+\kappa_{kl}}}\mathbf{h}_{kl}^{\mathrm{NLOS}}, \tag{1}$$

where the first term corresponds to the deterministic component line-of-sight (LOS) and the second term is the random propagation component non-line-of-sight (NLOS). The Rician factor κ_{kl} represents the power ratio between the components LOS and NLOS, defined as $\kappa_{kl} = p_{\mathrm{LOS}}/(1-p_{\mathrm{LOS}})$, where p_{LOS} is the probability that the LOS component exists, but is zero for propagation links that are just NLOS [9]. The LOS component between the RU l and UE k can be written as

$$\mathbf{h}_{kl}^{\mathrm{LOS}} = \sqrt{\beta_{kl}}\left[1, e^{-\mathrm{j}\pi\sin(\varphi_{kl})}, \cdots, e^{-\mathrm{j}(N-1)\pi\sin(\varphi_{kl})}\right]^{\mathrm{T}} e^{j\theta_{kl}}, \tag{2}$$

where φ_{kl} is the angle-of-arrival (AoA) and β_{kl} is the large-scale fading gain, including path loss and shadowing. Besides, the term $\theta_{kl} \sim \mathcal{U}[0, 2\pi)$ denotes the random phase shifts that may occur due to UEs' mobility.

The multipath NLOS component, undergoes a correlated Rayleigh distribution, given by

$$\mathbf{h}_{kl}^{\mathrm{NLOS}} = \sqrt{\mathbf{R}_{kl}}\,\mathbf{g}_{kl}, \tag{3}$$

where $\mathbf{g}_{kl} \in \mathbb{C}^{N\times 1}$ is composed of elements that are complex independent and identically distributed (i.i.d.) Gaussian $\mathcal{N}_{\mathbb{C}}(0,1)$ random variables (RVs). The correlation matrix $\mathbf{R}_{kl}$ is calculated following the Gaussian spatial correlation model of local scattering presented in [1], with $\beta_{kl} = \mathrm{tr}\{\mathbf{R}_{kl}\}/N$ being the common large-scale gain.

B. UL Training

Knowledge of the channel from the UEs to the serving RUs is required to transmit the signals coherently. The channel estimation is performed only once for each coherence block, as the channels are considered constant throughout a coherence block and change independently from one block to another. During UL training, the UE transmits training signals, known as pilot training sequences, to the RUs. In each coherence block, τ_p samples are reserved for channel estimation, generating a set of τ_p pilot sequences. These training sequences are designed to be orthogonal to each other and have known properties, which allows channel estimation by the RUs, which satisfy

$$\boldsymbol{\phi}_{t_1}^{\mathrm{H}}\boldsymbol{\phi}_{t_2} = \begin{cases} \sqrt{\tau_p}, \text{ if } t_1 = t_2 \\ 0, \text{ if } t_1 \neq t_2. \end{cases} \tag{4}$$

If the number of UEs is greater than the number of pilots τ_p, the same sequence can be reused between them [7]. The pilot sequence of UE k is denoted by $\boldsymbol{\phi}_{t_k} \in \mathbb{C}^{\tau_p \times 1}$, $t_k \in \{1, \cdots, \tau_p\}$, and the signal received by RU l is given by

$$\mathbf{y}_l^{\mathrm{pilot}} = \sum_{k=1}^{K} \sqrt{\eta_k}\mathbf{h}_{kl}\boldsymbol{\phi}_{t_k} + \mathbf{n}_l, \tag{5}$$

where η_k is the pilot transmit power, $\mathbf{n}_l \sim N_{\mathbb{C}}(\mathbf{0}_{\tau_p}, \sigma_{ul}^2\mathbf{I}_{\tau_p})$ is an additive noise. RU l calculates an inner product between $\mathbf{y}_l^{\mathrm{pilot}}$ and $\boldsymbol{\phi}_{t_k}$ to get enough statistics for the estimate of $\mathbf{h}_{kl}$. Then, the specific pilot signal of UE k can be expressed as

$$\mathbf{y}_{t_k l}^{\mathrm{pilot}} = \boldsymbol{\phi}_{t_k}^{\mathrm{H}}\mathbf{y}_l^{\mathrm{pilot}} = \sum_{i\in\mathcal{P}_k} \sqrt{\tau_p\eta_i}\mathbf{h}_{il} + \mathbf{n}_{t_k l}, \tag{6}$$

where $\mathcal{P}_k \subset \{1, \ldots, K\}$ is the set of UEs that use the same pilot as UE k, and $\mathbf{n}_{t_k l} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0_N}, \sigma_{ul}^2\mathbf{I}_N\right)$ is the additive noise. Hence, we can derive the non-Bayesian least-square (LS) estimator that minimizes $\left|\left|\mathbf{y}_{t_k l}^{\mathrm{pilot}} - \sqrt{p_k\tau_p}\mathbf{h}_{kl}\right|\right|^2$, as

$$\widehat{\mathbf{h}}_{kl} = \frac{1}{\sqrt{\eta_k\tau_p}}\mathbf{y}_{t_k l}^{\mathrm{pilot}}. \tag{7}$$

The LS estimator is useful when statistics are unknown or unreliable, as it does not need statistical information [4].

C. DL Data Precoding

Precoding is the processing that RUs or CPUs use channel estimates to compensate for channel effects and mitigate interference to improve the quality and speed of DL data transmission. For comparison purposes, this work considers the following distributed and centralized scalable precoding schemes: maximum ratio (MR), local partial minimum mean square error (LP-MMSE), partial minimum mean square error (P-MMSE), and partial regularized zero-forcing (P-RZF).

One can define the distributed MR combining vector as

$$\mathbf{v}_{kl}^{\mathrm{MR}} = \hat{\mathbf{h}}_{kl}, \tag{8}$$

which has low complexity and maximizes the ratio $|\mathbf{v}_{kl}^{\mathrm{H}}\hat{\mathbf{h}_{kl}}|^2/||\mathbf{v}_{kl}||^2$ between the power of the desired signal and the squared norm of the combining vector. This approach ensures that all the received energy from the desired signal is coherently combined, as the combining vector is weighted according to the desired end-UE's channel response. Using the principle of duality, combining vectors in the UL are converted to precoding in the DL using

$$\mathbf{w}_{kl} = \sqrt{\rho_{kl}}\,\frac{\mathbf{v}_{kl}}{\sqrt{\mathbb{E}\left\{\mathbf{v}_{kl}^{\mathrm{H}}\mathbf{D}_{kl}\mathbf{v}_{kl}\right\}}}, \tag{9}$$

where ρ_{kl} is the transmission power of the DL and the connections between the UE k and RUs by defining $\mathbf{D}_{kl} = \mathbf{I}_N$ if RU l serves UE k, zero otherwise.

The LP-MMSE is a distributed method that minimizes the mean square error at the receiver. On the other hand, P-RZF is a centralized method that minimizes interference between the transmitted signals. In turn, P-MMSE is the centralized version of LP-MMSE. These methods are generally more robust than MR but require higher computational complexity [10].

D. DL Spectral Efficiency

SE quantifies the amount of information transmitted in a wireless communication system related to the used bandwidth. The typical unit for measuring SE is bits per second per hertz (bits/sec/Hz). The DL SE of UE k can be calculated as in [2]

$$\mathrm{SE}_k = \frac{\tau_c - \tau_p}{\tau_c}\log_2\left(1 + \mathrm{SINR}_k\right), \tag{10}$$

where $(\tau_c - \tau_p)/\tau_c$ is the pre-log factor, which is a fraction of samples per coherence block that is used to transmit the DL data and the term SINR_k is the signal-to-interference-plus-noise ratio (SINR), given by

$$\mathrm{SINR}_k = \frac{\left|\mathbb{E}\left\{\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{w}_k\right\}\right|^2}{\sum\limits_{i=1}^{K}\mathbb{E}\left\{|\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_i\mathbf{w}_i|^2\right\} - |\mathbb{E}\left\{\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{w}_k\right\}|^2 + \sigma_{\mathrm{dl}}^2}, \tag{11}$$

where $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$ and $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ are, respectively, the collective vectors of $\mathbf{w}_{kl}$ and $\mathbf{h}_{kl}$. For instance, $\mathbf{w}_k = \left[\mathbf{w}_{k1}^{\mathrm{T}}, ..., \mathbf{w}_{kL}^{\mathrm{T}}\right]^{\mathrm{T}}$ for $l \in \{1, \cdots, L\}$. Moreover, $\mathbf{D}_k = \mathrm{diag}\left(\mathbf{D}_{k1}, ..., \mathbf{D}_{kL}\right) \in \mathbb{N}^{M\times M}$ stands for the diagonal block matrix. Note that (10) represents the widely known hardening bound, which is a capacity lower bound valid for any choice of precoding vectors [1].

III. Optimal pilot allocation and attribution for maximization of spectral efficiency

This work uses GA to define the allocation and assignment of pilots to maximize SE. Pilot allocation refers to the number of coherence block samples for pilots (τ_p) that will be used for channel estimation. Pilot assignment defines which pilot will be used by each UE (t_k), which also implies determining which UEs should use the same pilot.

GA operates on a population of individuals representing potential solutions, where solutions are encoded as chromosomes. The process involves assessing the aptitude of each individual to select the fittest individuals, using resources such as crossing and mutation to generate offspring, and repeating these steps over several generations. The population is expected to evolve and approach optimal solutions. GA is a powerful and relevant tool for solving complex optimization problems, as with each new generation there is a tendency towards better results [11]. Thus, GA is a potential candidate to optimize the number of pilots and the pilot assignment in D-mMIMOs networks.

A. *Pilot Allocation Optimization*

In an ideal scenario, each UE would have its own pilot, but since the coherence block is limited, the division between pilots and information makes this approach unfeasible. Even though this scenario leads to less interference and better channel estimation, the estimation overhead reduces the pre-log factor. Thus, the main disadvantage of UL pilot-based estimation is still the need to use samples of additional resources in pilot training, requiring a small number of pilot sequences to keep the overhead low and, consequently, its reuse between the UE. Reusing pilots causes coherent interference and decreases the quality of the channel estimation. Reducing pilot reuse (increasing τ_p) does not improve the average SE, as the estimation overhead decreases the pre-log factor, even though there is less coherent interference and better channel estimation. It is essential to find a proper balance between the overhead and reliable performance. To this end, the optimization problem is formulated as

$$\underset{\tau_p}{\text{maximize}} \quad \sum_{k=1}^{K} \frac{1}{K} \mathrm{SE}_k(\tau_p) \tag{12a}$$

$$\text{subject to} \quad \tau_p \in \{1, 2, \cdots, \min\{f\tau_c, K\}\}, \tag{12b}$$

where the objective is to maximize the average SE by finding the optimal value for τ_p in (12a) [7]. Since the resources are limited, GA is used to find the optimum by applying restrictions to the equation (12b), which accounts for the fact that τ_p is at most equal to the number of UE, but it cannot occupy the entire coherence block, being limited to $f\%$ of τ_c. The optimization is needed since the best τ_p is not easily derived from the SE in (10) and is also conflicting. When the value of τ_p increases, the value inside the logarithm in (10) also increases. This happens because a higher τ_p improves the channel estimation in (7), consequently improving the precoding vectors computation and the SINR. On the other hand, an increase in τ_p decreases the SE pre-log factor in (10). Thus, the maximum pre-log value occurs when τ_p is reduced to zero. However, this is not feasible since at least one pilot is required to perform pilot-based channel estimation [1], [7].

B. *Pilot Assignment Optimization*

In D-mMIMO, there is an issue of pilot contamination when multiple UEs share the same pilot resource to estimate the channels. Pilot contamination can lead to channel estimation errors, signal quality degradation, and system capacity limitations. For accurate channel estimation and efficient data transmission, choosing which UEs will need to use the same pilot to achieve the best system performance is crucial [6]. The proper choice of pilot indices can significantly impact system performance, and finding a simple solution to the problem is challenging due to the enormous number of combinations. To demonstrate the number of possibilities in the pilot assignment, one can use the formula for permutation with repetition $P^R(n,r) = n^r$ for $n >= 0$ e $r >= 0$. Using it in this context, we conclude that the total number of possibilities equals τ_p^K. Then, the pilot assignment problem is formulated as

$$\underset{t_k}{\text{maximize}} \quad \sum_{k=1}^{K} \frac{1}{K} \mathrm{SE}_k(t_k) \tag{13a}$$

$$\text{subject to} \quad t_k \in \{1, 2, \cdots, \tau_p\}, \forall k \in \{1, ..., K\}, \tag{13b}$$

where the GA is used as a tool to determine which pilot index t_k is assigned to each UE, finding the optimal sequence of pilot indices that satisfies (13a). The constraint (13b) are applied to ensure that each pilot index is less than the total number of pilots τ_p and that all UEs receive a pilot. To reduce the complexity of the problem, we can consider that the firsts UEs will receive an orthogonal pilot. When the number of UEs exceeds the number of pilots, the GA will determine the optimal pilot indices for the remaining UEs. In this case, the number of possibilities is given by $\tau_p^{K-\tau_p}$ [7].

IV. Numerical Results

The propagation model adopted in this work follows 3GPP TR 38.901, considering urban micro (UMi) for external environments. The carrier frequency is 3.5 GHz and the bandwidth is 100 MHz. The noise figure is 8 dB, using coherence interval equal to 200 to simulate a high speed scenario [12], and other simulation parameters are given in Table I. The GA is performed using the MATLAB optimization toolbox. It is used 10 individuals and 10 generations for the pilot allocation optimization, while 40 individuals and 200 generations are used for pilot assignment optimization. For both scenarios, elitism is defined using 2 individuals, selection and crossover are based on the tournament and scattered methods, respectively.

TABLE I: Simulation parameters.

PARAMETER	VALUE
Numbers of antennas	2
Power of UL per UE	22 dBm
UE height, RU height	1.65 m, 11.65 m
RU total DL power	23 dBm

The channel between RUs and UEs is considered in the process of selecting RUs and determining which pilots will be assigned to which UE. The SCF method serves the UEs by a subset of RUs selected based on the best large-scale channel gains in each pilot. After that, the number of RUs connected to each UE is limited by $C_{\max} = 10$ [5], [12]. The baseline pilot assignment, the first τ_p UEs are assigned to mutually orthogonal pilots, and the remaining UEs are assigned to the pilot that causes the lowest pilot contamination [1]. Both UEs and RUs are uniformly distributed into the coverage area.

A. *Pilot Allocation*

In this section, it is considered $f = 0.6$ and $L = 200$ RUs, and a coverage area of $500\text{m} \times 500\text{m}$. Fig. 1 shows the system performance in terms of average SE with $K = 50$ UEs, are uniformly distributed, for different precoding schemes. The P-MMSE precoding performs better for τ_p lower than 12. After that, its performance decreases. For the P-RZF, despite initially reaching lower SE, it surpasses the P-MMSE after $\tau_p = 21$. It can be noticed that the centralized precoding schemes benefit more from using higher τ_p values than the distributed ones. This happens because centralized schemes can better mitigate interference, which needs good channel estimates to be achieved. The LP-MMSE and MR precoding have similar behavior, requiring a small number of pilots to maximize their SE. The LP-MMSE presents 20.5% higher SE than MR at the cost of greater computational complexity [1]. In all cases, the best SE when τ_p is in less than half of K UEs. Although the greater number of pilots the better the channel estimates and interference mitigation, the results show that this does not lead to the best SE due to the impact of the estimation overhead in the SE's pre-log factor.

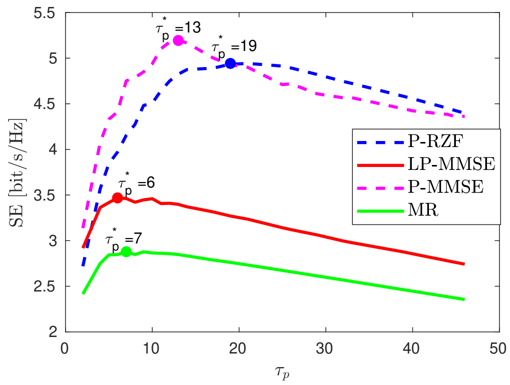


Fig. 1: Average DL SE vs. τ_p in SCF for different precoding schemes. Parameters setting: $L = 200$, $N = 2$ and $K = 50$ in a coverage area of $500\text{m} \times 500\text{m}$.

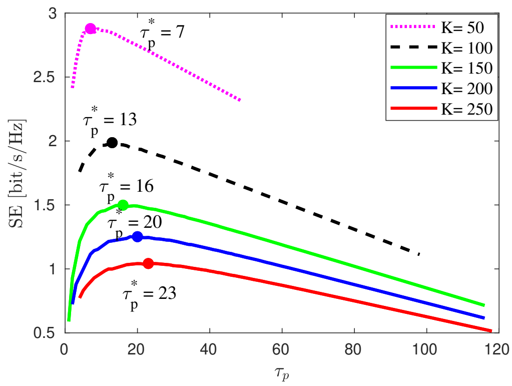


Fig. 2: Average DL SE vs. τ_p in the MR precoding varying K. Parameters setting: $L = 200$ and $N = 2$ in a coverage area of $500\text{m} \times 500\text{m}$.

Fig. 2 evaluates the average SE vs. the number of pilots τ_p for different numbers of UEs K, considering MR precoding. It can be noticed that the search space differs when the number of UEs changes. This happens because the maximum value that τ_p can assume corresponds to K for $K = 50$ and $K = 100$, while for larger values of K, τ_p is limited to $f\tau_c = 120$. It is observed that the smaller the number of UEs, the greater the SE due to less interference, and even when K is small, it is not advantageous to have a pilot for each UE due to the estimation overhead impact on the SE's pre-log factor. The optimal number of pilots changes by varying the number of UEs. It can be noticed that the optimal number of pilots is approximately 10% of the number of UEs the greater the number of UEs, to ensure the balance between accuracy and estimation overhead. Such behavior also occurs in other precoders, such as LP-MMSE and P-MMSE, but the results have been omitted to avoid redundancy.

B. *Pilot Assignment*

In this section, it is assumed $\tau_p = 4$, $L = 80$ RUs and $K = 8$ UEs in an area of $100\text{m} \times 100\text{m}$ to reduce the complexity of the optimization problem. Results are shown for MR and P-MMSE precoding schemes and for CCF and SCF RU selection. Fig. 3 compares the pilot assignment with and without GA, where each color represents a pilot. For CCF, the two precoders have the same pilot assignment distribution when using GA optimization. Like in the baseline method without GA, three UEs are assigned to the same pilot, increasing their interference. In both precoding schemes for SCF, each pilot is used by two UEs when GA optimization is performed, showing that the pilot assignment distribution is more balanced than the baseline method without GA, which potentially leads to less interference between UEs and improving channel estimation.

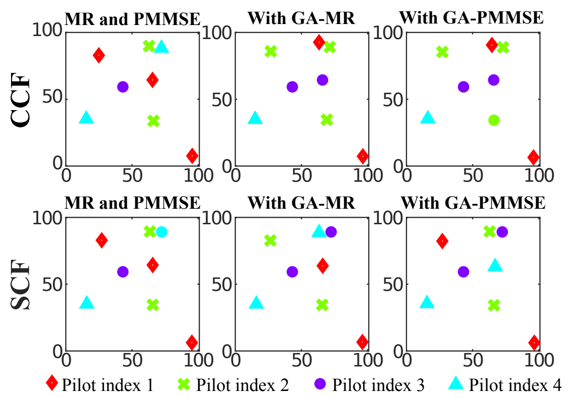
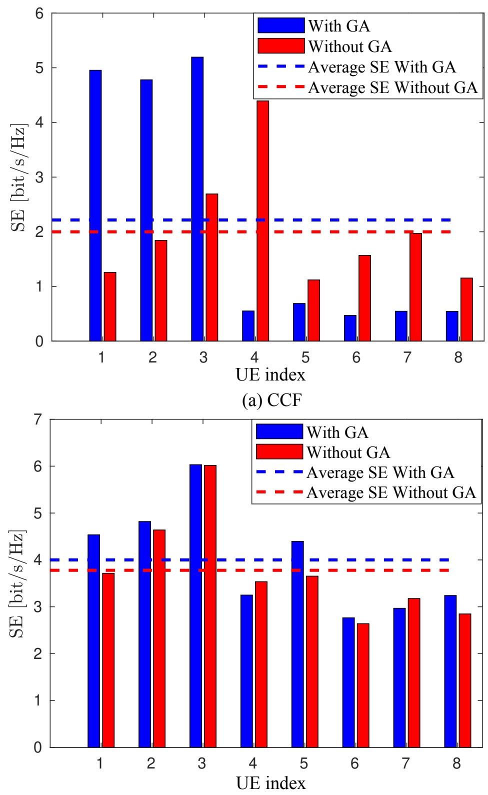


Fig. 3: Pilot assignment distribution in SCF and CCF, with and without GA. Parameters setting: $L = 80$, $N = 2$ and $K = 8$ in a coverage area of $100\text{m} \times 100\text{m}$.

Figs. 4 and 5 show the SE of each UE in the same scenario presented in Fig. 3, for MR and P-MMSE precoding schemes. In Fig. 4, the average SE for both SCF and CCF are higher using GA, even though the UEs specific SE is not always better with GA. It can also be noticed that the average SE for SCF is better than for CCF. For P-MMSE precoding using SCF in Fig. 5a, the SE of the UEs are very uniform, unlike for CCF in Fig. 5b. Despite that, GA pilot assignment optimization improves the average SE significantly for CCF. The results in

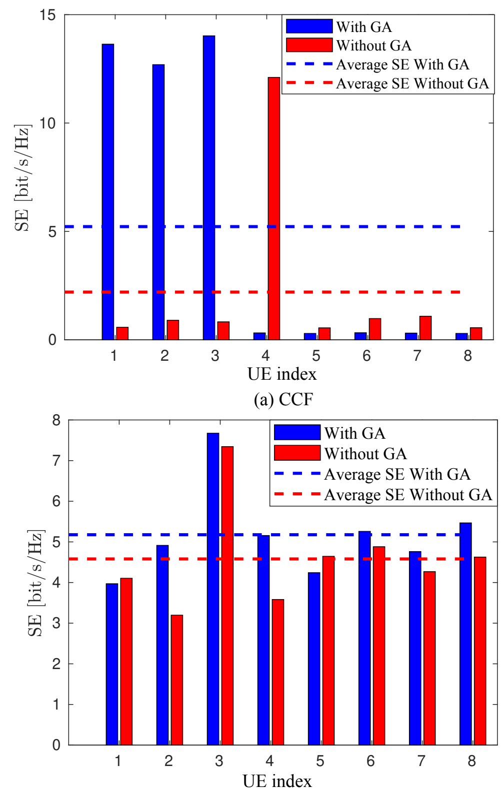
Figs. 4 and 5 demonstrate that the pilot assignment problem is highly influenced by the RU selection method.



(a) CCF

(b) SCF

Fig. 4: UEs' SE with SCF and CCF in MR precoding with and without GA. Parameters setting: $L = 80$, $N = 2$ and $K = 8$ in a coverage area of $100\text{m} \times 100\text{m}$.



(a) CCF

(b) SCF

Fig. 5: UEs' SE with SCF and CCF in P-MMSE precoding with and without GA. Parameters setting: $L = 80$, $N = 2$ and $K = 8$ in a coverage area of $100\text{m} \times 100\text{m}$.

V. Conclusions

This work has showed that the optimization of pilot allocation and assignment plays a crucial role in maximizing the SE of distributed massive MIMO networks. The results indicate that using GA, it was possible to find the optimal number of pilots that results in better system performance in terms of SE, being able to balance the accuracy and channel estimation overhead. Furthermore, the results show that utilizing distributed precoding requires fewer pilots to reach its maximum SE than centralized precoders. For the chosen simulation scenario, the results obtained by GA show that the optimal number of τ_p is approximately 10% of the total of UEs, when the number of UEs is high. However, other factors may influence this behavior, such as the size of the coverage area. Regarding the optimization of the pilot assignment, it is noted that the performance using GA increases in all cases, but is highly influenced by the RU selection scheme. As GA demands high computational complexity, it may be difficult to implement this solution in practical systems. Thus, the main findings of this work can be used to derive heuristic algorithms, since the baseline solutions are still not near the optimal. Future works also include the joint optimization of RU selection and pilot assignment, and the performance evaluation with linear minimum mean square error (MMSE) channel estimator instead of LS.

References

[1] Özlem Tugfe Demir, E. Björnson, and L. Sanguinetti, "Foundations of user-centric cell-free massive MIMO," *Foundations and Trends® in Signal Processing*, vol. 14, no. 3-4, pp. 162–472, 2021.

[2] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, 2020.

[3] M. Freitas *et al.*, "Matched-decision AP selection for user-centric cell-free massive MIMO networks," *IEEE Trans. Veh. Technol*, pp. 1–16, Jan. 2023.

[4] O. Özdogan, E. Björnson, and J. Zhang, "Performance of cell-free massive MIMO with Rician fading and phase shifts," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5299–5315, 2019.

[5] D. D. Souza *et al.*, "Effective channel DL pilot-based estimation in user-centric cell-free massive MIMO networks," in *Proc. IEEE Global Commun. Conf.*, Dec. 2022, pp. 705–710.

[6] L. Sanguinetti *et al.*, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 232–257, 2020.

[7] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *EURASIP Eurasip J. Wirel. Commun.*, vol. 2019, no. 1, Aug. 2019.

[8] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Massive MIMO with spatially correlated rician fading channels," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3234–3250, May 2019.

[9] C. D'Andrea, A. Garcia-Rodriguez, G. Geraci, L. G. Giordano, and S. Buzzi, "Analysis of UAV communications in cell-free massive MIMO systems," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 133–147, 2020.

[10] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, 2020.

[11] A. Eiben and J. Smith, *Introduction to Evolutionary Computing*. Springer Berlin Heidelberg, 2015.

[12] M. M. M. Freitas *et al.*, "Scalable user-centric distributed massive MIMO systems with limited processing capacity," in *Proc. IEEE Int. Conf. Comm. (ICC)*, 2023, pp. 1–7.