Two-timescale clustering with signaling analysis for indoor RIS-aided Cell-free systems

Paulo Tealdi, André Almeida Souza Coelho, Rodrigo Pinto Lemos

***Abstract*— This paper proposes a user-centric clustering formation algorithm for both AP and RIS based on the greedy algorithm in the weighted sum-rate maximization problem for indoor RIS-aided cell-free scenarios, where each user is matched with a group of APs and a group of RISs. For this, a non-convex joint precoding problem is solved following methods available in the literature, where we add a two-timescale approach to reduce overall signaling overhead.**

***Keywords*— Precoding, Optimization, Cluster, Cell-free, RIS.**

I. Introduction

Cell-free systems enhance user capacity by deploying multiple Access Points (APs) to serve User Equipment (UEs) under the control of a centralized processing unit. Given the potential for large coverage areas with numerous APs and UEs, scalability is a critical design consideration [1], [6]. Specifically for indoor environments, some unique challenges arise due to high-density blockages, leading to significant multi-path effects and attenuation of direct links. Cell-free systems address these issues by leveraging multiple APs to serve users from various angles and positions, mitigating deep fading and maintaining high spectral efficiency [1].

Similarly, Reconfigurable Intelligent Surfaces (RISs) offer an additional solution by controlling reflection patterns through individual phase shifts, enabling beamforming of reflected signals in desired directions [5] and, when combined with cell-free systems, RISs can significantly enhance capacity while being a low-cost and power-efficient solution [2].

The problem with network densification, as proposed with cell-free [1], is that the higher amount of APs causes the amount of estimated channels to grow and the overhead to increase due to backhaul signaling, meaning worse latencies. With the implementation of RISs, the problem gets worse because they are generally built with high amounts of reflecting elements, with each representing new channels to be estimated.

While prior studies have taken steps to mitigate signaling overhead in RIS-aided cell-free networks — [2] by employing a two-timescale centralized precoding scheme and [3] by adopting a distributed precoding approach — none has explored the joint clustering of APs and RISs as a means to further reduce this overhead. In this work, we address that gap by introducing a novel user-centric scheme that jointly clusters APs and RISs under a centralized framework, thereby minimizing the required channel estimation and backhaul signaling load. To the best of our knowledge, this is the first study to propose such a joint AP–RIS clustering strategy in cell-free massive MIMO. This approach directly supports our overarching objective of improving network scalability via reduced signaling overhead. We demonstrate the effectiveness of the proposed method by comparing its signaling overhead against both the baseline centralized solution from [2] and the distributed solution from [3].

Notations: $\mathbb{C}$ is the set of complex numbers, $[\cdot]^{-1}$, $[\cdot]^T$ and $[\cdot]^H$ denote the inverse, transpose and conjugate-transpose (hermitian) operations, respectively; $\mathbf{w}_{i,j}$ is the element at row i and column j of the matrix $\mathbf{w}$; $||\cdot||$ is the Euclidean norm of its argument; diag() is the diagonal operation; $\mathbf{0}_{X,Y}$ is the zero matrix with dimensions X and Y.

II. System Model

The complete channel between user k and AP l includes direct and RIS paths and can be written as [2]

$$\mathbf{h}_{l,k}^H = \mathbf{H}_{l,k}^H + \sum_{r=1}^{R} \mathbf{F}_{r,k}^H \mathbf{\Theta}_r^H \mathbf{G}_{l,r} \tag{1}$$

Where $\mathbf{H}_{l,k} \in \mathbb{C}^{N_t \times N_r}$ is the direct channel from AP l to UE k, $\mathbf{F}_{r,k} \in \mathbb{C}^{M \times N_r}$ is the RIS to UE component of the indirect channel, $\mathbf{G}_{l,r} \in \mathbb{C}^{N_t \times M}$ is the AP to RIS component of the indirect channel and $\mathbf{\Theta}_r = \text{diag}(\theta_{r,1}, \cdots, \theta_{r,M})$ represents the phase shift matrix of RIS r, being $\theta_{r,m}$ the individual phase shift of element m of RIS r where $|\theta_{r,m}| \leq 1$. Channel State Information (CSI) is assumed to be perfectly known.

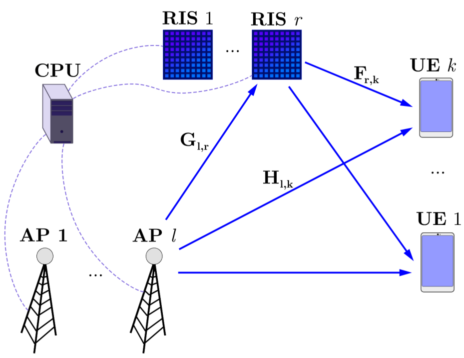

Fig. 1. RIS-aided cell-free system diagram

We denote N_t, N_r and M as the number of antennas of each AP, the number of antennas of each UE and the number

Paulo Tealdi (paulotealdi@discente.ufg.br), André Almeida Souza Coelho (andreproj1@gmail.com) and Rodrigo Pinto Lemos (lemos@ufg.br) are with the Center of Excellence in Wireless Intelligent Networks and Advanced Services (CERISE), at the School of Electrical, Mechanical and Computer Engineering, Universidade Federal de Goiás (EMC-UFG). The authors thank CERISE and the Goiás State Research Support Foundation (FAPEG) for their support and funding of this research.

of elements of each RIS, respectively. Let $\mathcal{K} = \{1, \cdots, K\}$, $\mathcal{R} = \{1, \cdots, R\}$, $\mathcal{L} = \{1, \cdots, L\}$, $\mathcal{M} = \{1, \cdots, M\}$ denote the index sets of users, RISs, APs and RIS elements.

Further simplification can be made by defining $\mathbf{F}_k = [\mathbf{F}_{1,k}^{\mathrm{T}}, \cdots, \mathbf{F}_{R,k}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{MR \times N_r}$, $\mathbf{\Theta} = \mathrm{diag}\left(\mathbf{\Theta}_1, \cdots, \mathbf{\Theta}_R\right) \in \mathbb{C}^{MR \times MR}$ and $\mathbf{G}_l = [\mathbf{G}_{l,1}^{\mathrm{T}}, \cdots, \mathbf{G}_{l,R}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{MR \times N_t}$, so the complete channel in (1) can now be written as

$$\mathbf{h}_{l,k}^H = \mathbf{H}_{l,k}^H + \mathbf{F}_k^H \mathbf{\Theta}^H \mathbf{G}_l \tag{2}$$

A transmitted symbol to user k denoted by $s_k \in \mathbb{C}$ is initially precoded with a precoding vector $\mathbf{w}_{l,k} \in \mathbb{C}^{N_t \times 1}$ in the downlink transmission, so the precoded symbols $\mathbf{w}_{l,k} s_k$ are summed up for all k in AP l to generate the AP's precoded symbol $\mathbf{x}_l$. Then, user k receives every transmitted symbol $\mathbf{x}_l$ destined to every user plus an additive white Gaussian noise (AWGN) distributed as $\mathcal{CN}(\mathbf{0}, \sigma_k^2 \mathbf{I}_{N_r})$ and denoted by $\mathbf{z}_k \in \mathbb{C}^{N_r \times 1}$. Received signal from user k can be written as

$$\mathbf{y}_k = \underbrace{\mathbf{h}_k^H \mathbf{w}_k s_k}_{\text{Desired signal}} + \underbrace{\sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{h}_k^H \mathbf{w}_i s_i}_{\text{Interference}} + \mathbf{z}_k \tag{3}$$

where the equality holds defining $\mathbf{h}_k = [\mathbf{h}_{1,k}^{\mathrm{T}}, \cdots, \mathbf{h}_{L,k}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{LN_t \times N_r}$ and $\mathbf{w}_k = [\mathbf{w}_{1,k}^{\mathrm{T}}, \cdots, \mathbf{w}_{L,k}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{LN_t \times 1}$. Signal-to-interference-plus-noise ratio (SINR) can be calculated as

$$\mathrm{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{\substack{i=1 \\ i \neq k}}^{K} |\mathbf{h}_k^H \mathbf{w}_i|^2 + \sigma_k^2}, \forall k \in \mathcal{K} \tag{4}$$

Furthermore, $\mathrm{SINR}_{l,k}$ and $\mathrm{SINR}_{r,k}$ are the SINR values with only AP l and RIS r being active, respectively.

A. Indoor Channel Model

The channels from AP to RIS and from RIS to UE are assumed to be on LoS (Line of Sight) conditions since both AP and RIS are positioned at high altitude, and RISs are strategically positioned to be in LoS condition with UEs. Large-scale fading is modeled as follows using InH-Shopping Malls-LOS and InH-Shopping Malls-NLOS single slope (FFS) parameters from [4]

$$\mathrm{PL}(f, d) = 20 \log_{10}\left(\frac{4\pi f}{c}\right) + 10n \log_{10}\left(\frac{d}{1\text{ m}}\right) + X_\sigma \tag{5}$$

where f is the frequency in Hz, n the pathloss exponent, d the distance in meters, X_σ the shadow fading with σ in dB and c is the speed of light. We use Rician fading to account for small-scale fading.

$$\mathbf{H} = \sqrt{\frac{\kappa}{1+\kappa}} \mathbf{H}^{\mathrm{LoS}} + \sqrt{\frac{1}{1+\kappa}} \mathbf{H}^{\mathrm{NLoS}} \tag{6}$$

B. Joint Clustering Algorithm

To reduce channel estimation overhead, selection matrices $\mathcal{D}_{r,k}$ and $\mathcal{D}_{l,k}$ are generated by SINR matrices $\mathrm{SINR}_{r,k}$ and $\mathrm{SINR}_{l,k}$ with a greedy approach by choosing the RISs/APs and UEs pairs with bigger SINR values. Complete channel $\mathbf{h}_{l,k}$ is modified by a selection matrix $\mathcal{D}_{l,k}$, resulting in $\mathcal{D}_{l,k} \mathbf{h}_{l,k}$, where

$$\mathcal{D}_{l,k} \mathbf{h}_{l,k} = \begin{cases} \mathbf{h}_{l,k}, & \text{if AP } l \text{ is selected to user } k \\ \mathbf{0}_{N_t \times N_r}, & \text{if AP } l \text{ is not selected to user } k \end{cases} \tag{7}$$

Similarly, RIS channels $\mathbf{F}_k^H \mathbf{\Theta}^H \mathbf{G}_l$ can be modified by $\mathcal{D}_{r,k}$, generating the new complete channel. Since it relies on SINR values, it needs optimized $\mathbf{W}$ and $\mathbf{\Theta}$, and therefore, the matrices are built at the end of the first timescale to be used in the next small timescales.

C. Optimization Problem

With (4), the WSR (weighted sum-rate) R_{sum} can be calculated by

$$R_{\text{sum}} = \sum_{k=1}^{K} \eta_k \log_2(1 + \mathrm{SINR}_k) \tag{8}$$

where η_k is the weight of each user k, but we assume that $\eta_k = 1$, $\forall k \in \mathcal{K}$. Now, WSR can be maximized by optimal precoding vectors and selection matrices, given as

$$\begin{aligned} \max_{\mathbf{\Theta}, \mathbf{W}, \mathcal{D}_{l,k}, \mathcal{D}_{r,k}} \quad & R_{\text{sum}}(\mathbf{\Theta}, \mathbf{W}, \mathcal{D}_{l,k}, \mathcal{D}_{r,k}) \\ \text{s.t.} \quad & C_1 : \sum_{k=1}^{K} ||\mathbf{w}_{l,k}||^2 \leq P_t, \forall l \in \mathcal{L} \\ & C_2 : |\theta_{r,m}| \leq 1, \forall r \in \mathcal{R}, \forall m \in \mathcal{M} \\ & C_3 : \sum_{l \in \mathcal{L}} \mathcal{D}_{l,k} = \mathbf{I}_{N_t} AP_{\text{cluster}}, \forall k \in \mathcal{K} \\ & C_4 : \sum_{r \in \mathcal{R}} \mathcal{D}_{r,k} = \mathbf{I}_{N_t} R_{\text{cluster}}, \forall k \in \mathcal{K} \\ & C_5 : \mathcal{D}_{l,k} \in \{\mathbf{0}_{N_t}, \mathbf{I}_{N_t}\}, \forall l \in \mathcal{L}, \forall k \in \mathcal{K} \\ & C_6 : \mathcal{D}_{r,k} \in \{\mathbf{0}_{N_t}, \mathbf{I}_{N_t}\}, \forall r \in \mathcal{R}, \forall k \in \mathcal{K} \end{aligned} \tag{P1}$$

which is separated into subproblems, where $\mathbf{\Theta}$ and $\mathbf{W}$ are optimized as [2] employing Alternating Optimization, and the maximization with respect to $\mathcal{D}_{l,k}$ and $\mathcal{D}_{r,k}$ is as in section B and follows Algorithm 1.

III. Simulation and Results

RISs are installed on the ceiling at $(20, 0)$, $(60, 0)$, $(100, 0)$ and $(140, 0)$ with height $h_R = 4$ m, and the APs are installed on the wall, evenly distributed from $(0, -20)$ to $(160, -20)$, with height $h_A = 3$ m. We consider K = 4 active UEs, $L = 40$ APs, R = 4 available RIS, operating frequency of $f = 15$ GHz, transmit power $P_t = 1$ mW, AP and UE have one antenna each $N_t = N_r = 1$. For all simulations, we assume a noise level of $\sigma_k^2 = -80$ dBm$, \forall k \in \mathcal{K}$.

Algorithm 1 Proposed Algorithm

Input: $\mathbf{H}_{l,k}$, $\mathbf{G}_{l,r}$ and $\mathbf{F}_{r,k}$ where $l \in \mathcal{L}$, $k \in \mathcal{K}$, $r \in \mathcal{R}$.
Output: $\mathbf{W}_{\text{opt}}$ and $\mathbf{\Theta}_{\text{opt}}$.

```
Initialize W and Θ;
if not first timescale then
    Update h_{l,k} and F_k^H Θ^H G_l as in section B
end if
Update W and Θ as in [2]
if first timescale then
    Update D_{l,k} and D_{r,k} with SINR_{l,k} and SINR_{r,k}
end if
return W_opt and Θ_opt.
```

A. Cluster sizes effect

As shown in [5], RIS does not make much improvement in system performance while the direct link is strong, but it can hold the performance while the direct link is weak. Here we explore with different values of R_{cluster} and AP_{cluster} with a NLoS only direct link ($\kappa = 0$). AP to RIS channels are configured with $\kappa = 10$, and from RIS to UE with $\kappa = 8$.

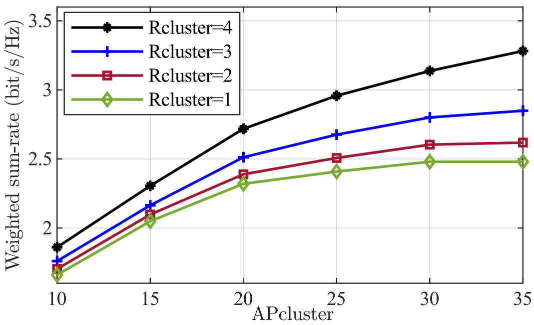

Fig. 2. WSR results varying cluster sizes with a blocked direct channel

Figure 2 shows that selecting a greater number of RISs per cluster leads to improved overall system performance. The increasing gap between the curves is attributed to the growing number of RIS elements, which follows a quadratic relationship with the channel gains — a well-known characteristic of RISs [5]. The tendency of the curves to stabilize with increasing AP_{cluster} values arises from the algorithm's preference to select the more distant APs last, as these typically contribute smaller average improvements to WSR.

B. Signaling analysis and comparison with literature

With clusters overlap, some RISs/APs can be part of two or more clusters and a number of them can be unused. Let L_a be the number of active APs (APs being used) and R_a the number of active RISs. T is the total number of timescales and X is the number of timescales with partial CSI. For the first timescale, system will perform full CSI, so the proposed algorithm has a signaling overhead after I_o iterations of $N_r N_t LK + I_o(KN_r + RM + 2LKN_t)$, which is the same as the centralized one, against $N_r N_t LK + I_o(KN_r + RM + (2L-1)KN_t$ of the distributed [3]. However, for the next X timescales, only partial CSI is performed due to clustering, so the proposed algorithm for C clusters has a signaling overhead of $CN_r N_t L_c K + I_o(KN_r + R_a M + 2L_a KN_t)$, while it stays the same for centralized and distributed frameworks.

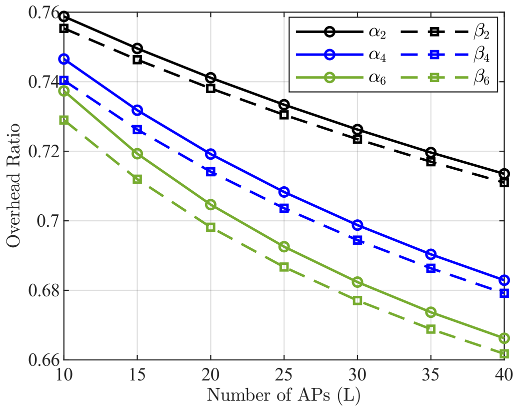

Fig. 3. Signaling overhead ratios α_K = Proposed/Distributed and β_K = Proposed/Centralized for $K = 2,\ 4$ or 6 (number of UEs)

Parameters include $C = K$, $L_c = 0.25 \cdot L$, $I_o = 30$, $R_a = 3$, $L_a = 0.5 \cdot L$, and $T = 11 = X + 1$. For instance, when $L = 40$, Figure 3 shows that $\alpha_4 = 0.6829$ and $\beta_4 = 0.6792$, which means that the proposed algorithm provided overhead reductions of 31.71% and 32.08% when compared to distributed and centralized algorithms, respectively. This result is flexible and depends on the framework parameters, where, for example, the number of partial CSI timescales could be tuned to meet performance requirements, since not updating the clusters for long enough can degrade system performance.

IV. Conclusions

The proposed algorithm proved to be a flexible solution towards making RIS-aided cell-free systems scalable by significantly reducing overall signaling overhead when compared to the conventional centralized cell-free framework and to the distributed framework from [3]. Further investigation will be made into algorithm complexity, clustering in distributed cell-free, pilot contamination, and more complex yet efficient clustering algorithms.

References

[1] Özlem Tuğfe Demir, Emil Björnson and Luca Sanguinetti (2021), “Foundations of User-Centric Cell-Free Massive MIMO”, Foundations and Trends® in Signal Processing: Vol. 14, No. 3-4, pp 162–472. DOI: 10.1561/2000000109.

[2] Z. Zhang and L. Dai, "A Joint Precoding Framework for Wideband Reconfigurable Intelligent Surface-Aided Cell-Free Network," in IEEE Transactions on Signal Processing, vol. 69, pp. 4085-4101, 2021, doi: 10.1109/TSP.2021.3088755.

[3] Zhang, Peng, et al. "Joint distributed precoding and beamforming for RIS-aided cell-free massive MIMO systems." IEEE Transactions on Vehicular Technology (2023).

[4] K. Haneda et al., "Indoor 5G 3GPP-like channel models for office and shopping mall environments," 2016 IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, Malaysia, 2016, pp. 694-699, doi: 10.1109/ICCW.2016.7503868.

[5] Emil Björnson and Özlem Tuğfe Demir. (2024). Introduction to Multiple Antenna Communications and Reconfigurable Surfaces. Boston–Delft: Now Publishers

[6] E. Björnson and L. Sanguinetti, "Scalable Cell-Free Massive MIMO Systems," in IEEE Transactions on Communications, vol. 68, no. 7, pp. 4247-4261, July 2020, doi: 10.1109/TCOMM.2020.2987311.