Analysis of RRU Correlation Performance in Full Spectrum Uplink Cell-free RAN Systems

Ziyang Zhang*†, Dongming Wang*†, Yunxiang Guo*†, Yanfeng Hu*†, Jie ling*†, and Baiping Xiong*†
*National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
†Purple Mountain Laboratories, Nanjing 211111, China
Emails:{ziyangzhang, wangdm, , huyanfeng, , xiongbp}@seu.edu.cn

2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall) | 979-8-3503-2928-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/VTC2023-Fall60731.2023.10333796

***Abstract*—In recent years, with the improvement of mobile communication network performance, cell-free mMIMO with fully integrated and distributed processing has been widely studied, and wireless access networks have also become a widely studied topic in the academic community. This article analyzes the spectral efficiency of a new full spectrum cell-free wireless access network architecture with multiple EDUs and derives the performance upper and lower bounds of its traversal achievable rate. The full set formula and distributed processing can be used as two special cases; Secondly, the article further considers the distribution of users and large-scale fading models and studies the location distribution of RRUs. It is concluded that a uniform distribution of RRUs is beneficial for user traversal and speed improvement, and RRUs in multiple EDUs need to be as intertwined as possible, which is different from traditional multi-node clustering centralized collaborative processing; The article further proposes an modified genetic algorithms (GA), which simulates the performance of a new full spectrum cell-free wireless network with pilot pollution. The performance is compared with that of multi-RRU clustering group collaboration processing through Monte Carlo simulation.**

***Index Terms*—Cell-free, massive MIMO, radio access network, full spectrum, RRU location analysis, genetic algorithms**

I. Introduction

With the continuous development of communication technology and the increasing demand for data transmission, the fifth-generation mobile communication system (5G) is being widely promoted, and the academic and industrial communities are also starting to study the sixth-generation (6G) technology to meet higher communication needs in the future [1]. In 6G technology, cell-free massive multiple-input multiple-output (CF-mMIMO) will be an important research direction.

There have been many studies on cell-free systems [2], [3]. The article [4] points out that mMIMO is a more general extension of traditional Shannon information theory. If channel information is globally known, channel models and capacity analysis in complex application scenarios such as single-user, multi-user, or multi-base station joint processing can be described in a unified form. The spatial freedom of MIMO channels can be artificially increased by increasing the antenna configuration at both ends of the transmitter and receiver [5]. There is no so-called performance limit for future mobile communication systems under specific theoretical assumptions. CF-mMIMO technology is a more revolutionary technology that will break traditional cellular structures and achieve more flexible and efficient network coverage. In the research of CF-mMIMO, the scalable implementation of joint transceivers may break the cellular architecture and achieve cell-free communication. This will further improve the performance of the network and lay the foundation for the development of 6G. However, in the research of CF-mMIMO, some technical challenges still need to be addressed: firstly, CF-mMIMO requires more antennas to achieve higher transmission rates, and traditional centralized processing complexity is very high for multi-user scenarios. Secondly, due to the complexity of cell-free communication structures, more efficient architecture implementation and more advanced algorithms are needed to support it. Therefore, the research on CF-mMIMO is still in the theoretical research stage, and further research is needed on its architecture implementation and location association deployment.

At present, there is a lot of work on the scalability of CF-mMIMO and RRU location distribution. [6] proposes a scalable CF-MIMO framework, which adopts a user-centered dynamic cooperative cluster (DCC) association method, provides a new algorithm for joint initial access, pilot allocation, and cluster formation, and proposes a scalable method for standard channel estimation, precoding, and combining. [7] considers the CF-MIMO system in which the AP location satisfies the Poisson Point process (PPP), and deduces the downlink coverage probability and achievable rate. Based on the Poisson Point process and assuming that the distribution of AP is random, [8] simulates the real AP behavior on mobile networks. Considering maximum ratio combining (MRC) and minimum mean square error (MMSE) combined at the AP receiver to estimate channel statistics, CF-MIMO uplink spectrum efficiency (SE) is derived, and the relationship between AP density and distribution and system spectrum efficiency is proved. [9] considers the approximation of realizable rates of several linear precoders and detectors in uplink (UL) and downlink (DL) of non-cooperative multi-cell time division systems. [10] studies the performance of channel hardening and confidence propagation in real random AP deployment in CF-mMIMO systems. [11] proposes a scalable CF-mMIMO implementation architecture for CF-mMIMO systems through distributed transceivers and scalable cooperative transmission.

On this basis, the potential key technologies such as channel information acquisition, transceiver design, dynamic user and access point association, and new duplex are introduced,

and the performance of distributed receiver design is evaluated. [12] derives two approximations of uplink rates with perfect/non-perfect channel state information (CSI) in CF-mMIMO systems, and all these approximations are not only simple but also converge to the classical boundary implemented in traditional large-scale MIMO systems where the base station (BS) antennas are located in the same location. [13] compares the asymptotic rate performance of downlink multiuser systems with multiple base station antennas, which are either located in the same location or uniformly distributed in the cell. Two representative linear precoding schemes, maximum ratio transmission (MRT) and zero-forcing beamforming (ZFBF), are considered and used to characterize the effect of BS antenna layout on the rate performance. [14] derives the lower bounds of the capacity of MRC, zero-forcing (ZF), and MMSE detection in centralized processing scenarios.

However, the existing work is to analyze the spectral efficiency of distributed CF-mMIMO systems based on stochastic geometry architecture, and most of the work of RRU and user location analysis is based on MRT. This scheme does not make full use of the cooperative characteristics of cell-free, but there is little research on cooperative transceivers based on interference suppression, and there is no theoretical research on the strategy based on the association between RRU and EDU. Therefore, in view of the above problems, under the new architecture, the theory explains the optimal RRU distribution of uplink EDU in this paper. The modified GA algorithm is used to Improve system performance, and compared with the clustering scheme Considering the influence of non-perfect channel state information and multi-antenna correlation, the simulation results and analysis are provided.

Notation: In this paper, bold uppercase $\mathbf{A}$ (bold lowercase $\mathbf{a}$) denotes a matrix (a vector). $\mathbf{I}_N$ denotes the $N \times N$ dimensional identity matrix. $(\cdot)^H$, $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^{-1}$ and $(\cdot)^\dagger$ stand for the conjugate transpose, transpose, conjugate, inverse and pseudo-inverse, respectively. $\mathrm{diag}\{\mathbf{a}\}$, $\mathrm{diag}\{\mathbf{A}\}$ and $\mathrm{blkdiag}\{\mathbf{A}_1, \cdots, \mathbf{A}_N\}$ represent a diagonal matrix with $\mathbf{a}$ along its main diagonal, a vector constructed by the main diagonal of the matrix $\mathbf{A}$, a block diagonal matrix , respectively. $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ denotes the complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$. $\otimes$ denotes the Kronecker product of two matrices. Finally, $\mathbb{E}\{\cdot\}$ is the expectation operator.

II. System model

This section introduces the implementation of the CF-mMIMO system. Under the new architecture, we analyze the uplink SE of CF-mMIMO. The combining strategy and procoding scheme adopt a unified representation. We consider CF-mMIMO with L RRU each with N antennas and K single antenna user equipment (UE). Suppose LN is very large and $L \gg K$. At the lth RRU, the received signal $\mathbf{y}_{\mathrm{UL},l}$ can be expressed as [15]:

$$\mathbf{y}_{\mathrm{UL},l} = \sum_{k=1}^{K} \mathbf{h}_{l,k} \sqrt{p_k} s_k + \mathbf{z}_l. \tag{1}$$

where s_k denotes the transmit symbol of the k user, p_k represents the uplink transmit power of the kth UE, and $\mathbf{h}_{l,k}$ represents the $N \times 1$ channel state information (CSI) from the kth UE to the lth RRU, and $\mathbf{z}_l \sim \mathcal{CN}\left(0, \sigma_{\mathrm{UL}}^2 \mathbf{I}_N\right)$ represents additive white Gaussian noise.

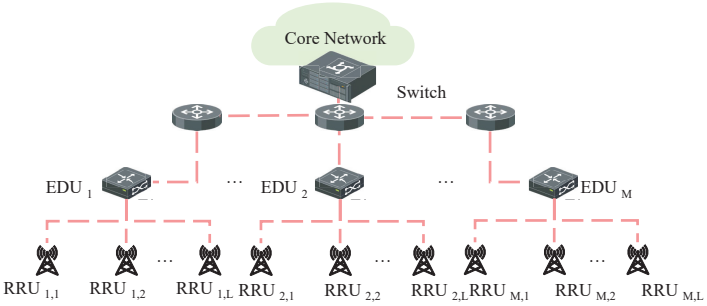

Fig. 1. System model, a total of M EDU.

It is assumed that there are M EDUs in the system, as shown in the Fig. 1. Let the channel vector from the kth UE to all RRUs be:

$$\mathbf{h}_k = \left[\mathbf{h}_{k,1}^{\mathrm{T}}, \mathbf{h}_{k,2}^{\mathrm{T}} \cdots \mathbf{h}_{k,L}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{LN}, \tag{2}$$

where $\mathbf{h}_k = \mathbf{B}_k^{1/2} \mathbf{g}_k$ and the large-scale fading matrix is $\mathbf{B}_k = \mathrm{diag}\left(\beta_{k,1}, \beta_{k,2}, \cdots, \beta_{k,L}\right) \otimes \mathbf{I}_N$. Correlated Rayleigh fading channel vector $\mathbf{g}_k$ has the following distribution:

$\mathbf{g}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$, where $\mathbf{R}_k = \mathrm{diag}\left(\mathbf{R}_{k,1}, \ldots, \mathbf{R}_{k,L}\right) \in \mathbb{C}^{LN \times LN}$ represents the kth Block diagonal space correlation matrix.

According to [16], the uplink signal-to-noise ratio (SINR) of the ktH UE is

$$\gamma_k^{(\mathrm{UL,d})} = \frac{p_k \left| \sum_{m=1}^{M} \mathbf{v}_{k,m}^{\mathrm{H}} \mathbf{h}_{k,m} \right|^2}{\sum_{i=1, i \neq k}^{K} p_i \left| \sum_{m=1}^{M} \mathbf{v}_{k,m}^{\mathrm{H}} \mathbf{h}_{i,m} \right|^2 + \sigma_{\mathrm{UL}}^2 \left(\sum_{m=1}^{M} \left\| \mathbf{v}_{k,m} \right\|^2 \right)}. \tag{3}$$

where $\mathbf{v}_{k,m}$ is the precoding vectors and combining vectors for user k. For the convenience of analysis, we adopt perfect channel state information (CSI), the user is completely associated with the RRU, and consider the RRU single antenna scenario, namely: $N = 1$. Obviously, the complete centralized and fully distributed case can be introduced as special cases of $M = 1$ and $M = L$ [6]. In the next section, we will analyze the performance relationship between the ergodic achievable rate and the RRU position distribution as an example.

III. Association Analysis of collaborative RRU

A. Upper and lower bounds on achievable rate

For the ZF combining with perfect CSI [17], the original centralized combining matrix is $\mathbf{V} = \mathbf{H}\left(\mathbf{H}^{\mathbf{H}}\mathbf{H}\right)^{-1} \in \mathbb{C}^{K \times LN}$ and it is assumed that RRU is a single antenna, so $\mathbf{H} \triangleq [\mathbf{h}_1, \ldots, \mathbf{h}_k, \ldots, \mathbf{h}_K] \in \mathbb{C}^{L \times K}$. Due to the zero-forcing of interference, each EDU is subjected to interference of 0,

and the matrix dimension decreases as the number of EDUs increases. We have:

$$\gamma_k^{(\mathrm{UL,d})} = \frac{p_k M^2}{\sum\limits_{m=1}^{M} \left\|\mathbf{v}_{k,m}\right\|^2}, \tag{4}$$

where $\mathbf{v}_{k,m}$ is the kth column of $\mathbf{V}_m$ and $\mathbf{V}_m = \mathbf{H}_m\left(\mathbf{H}_m^{\mathbf{H}}\mathbf{H}_m\right)^{-1}$, the precoding condition needs to meet the requirement that the number of antennas in the EDU L_m is much greater than the number of users K, and the user set is $\mathcal{K}$. Therefore, the achievable rate of user K is:

$$R_k = \mathbb{E}_{\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_M}\left\{\log_2\left(1+\frac{p_k M^2}{\sum\limits_{m=1}^{M}\left[\left(\mathbf{H}_m^{\mathrm{H}}\mathbf{H}_m\right)^{-1}\right]_{kk}}\right)\right\}, \tag{5}$$

Among them, different channel matrices $\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_M$ are independent, where $\mathbf{H}_m \in {}^{L_m \times K}$. The number of RRUs for EDUs satisfies the following two expressions: $L_1 + L_2 + \cdots + L_M = L$, $\mathcal{L}_1 \cup \mathcal{L}_2 \cup \cdots \cup \mathcal{L}_M = \mathcal{L}$, $\mathcal{L} = \{l | \forall l = 1, \ldots, L\}$.

So using Jensen's inequality, the upper bound of the achievable rate is:

$$R_k \le \log_2\left(1+\mathbb{E}_{\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_M}\left\{\frac{p_k M^2}{\sum\limits_{m=1}^{M}\left\|\mathbf{v}_{k,m}\right\|^2}\right\}\right) \triangleq R_k^{\mathrm{ub}}. \tag{6}$$

The lower bound of achievable rate is:

$$R_k \ge \log_2\left(1+\frac{p_k M^2}{\mathbb{E}_{\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_M}\left\{\sum\limits_{m=1}^{M}\left\|\mathbf{v}_{k,m}\right\|^2\right\}}\right) \triangleq R_k^{\mathrm{lb}}. \tag{7}$$

For the upper bound, it is necessary to analyze the expected fraction and use the matrix inversion formula:

$$\left(\mathbf{H}_m^{\mathrm{H}}\mathbf{H}_m\right)^{-1} = \frac{\left(\mathbf{H}_m^{\mathrm{H}}\mathbf{H}_m\right)^*}{\det\left(\mathbf{H}_m^{\mathrm{H}}\mathbf{H}_m\right)}, \tag{8}$$

where $*$ represents the adjugate matrix of the matrix, let $\mathbf{G_m} = \mathbf{H_m^H H_m}$,$\mathbf{G_{m,k}}$ to be the k,kth element of algebraic Cofactor of $\mathbf{G_m}$, so the k,k element of inverse matrix $\mathbf{H}_m^{\mathrm{H}}\mathbf{H}_m$ can be expressed as:

$$\left[\left(\mathbf{H}_m^{\mathrm{H}}\mathbf{H}_m\right)^{-1}\right]_{kk} = \frac{\det\left(\mathbf{G}_{m,k}\right)}{\det\left(\mathbf{G}_m\right)}. \tag{9}$$

Therefore,

$$\frac{1}{\sum\limits_{m=1}^{M}\left\|\mathbf{v}_{k,m}\right\|^2} = \frac{1}{\sum\limits_{m=1}^{M}\left[\left(\mathbf{H}_m^{\mathrm{H}}\mathbf{H}_m\right)^{-1}\right]_{kk}} = \frac{1}{\sum\limits_{m=1}^{M}\left[\frac{\det(\mathbf{G}_{m,k})}{\det(\mathbf{G}_m)}\right]}. \tag{10}$$

Using Jensen's inequality, it can be obtained:

$$\frac{1}{\frac{1}{M}\left[\sum\limits_{i=1}^{M}\frac{\det(\mathbf{G}_{i,k})}{\det(\mathbf{G}_i)}\right]} \le \frac{1}{M}\sum_{i=1}^{M}\frac{1}{\frac{\det(\mathbf{G}_{i,k})}{\det(\mathbf{G}_i)}} = \frac{1}{M}\sum_{i=1}^{M}\frac{\det\left(\mathbf{G}_i\right)}{\det\left(\mathbf{G}_{i,k}\right)}. \tag{11}$$

According to the (11) further scaling, the upper bound (6) can be expressed as:

$$R_k^{\mathrm{ub}} \le \log_2\left(1+p_k M_{\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_M}\left\{\frac{1}{M}\sum_{i=1}^{M}\frac{\det\left(\mathbf{G}_i\right)}{\det\left(\mathbf{G}_{i,k}\right)}\right\}\right). \tag{12}$$

Using Schur's Complementary Lemma [13]:

$$\begin{aligned}\frac{\det\left(\mathbf{G}_i\right)}{\det\left(\mathbf{G}_{i,k}\right)} &\overset{L_m \gg K}{\approx} \left\|\mathbf{h}_{k,m}\right\|^2 - \sum_{j\neq k, j\in\mathcal{K}}\frac{\mathbf{h}_{k,m}\mathbf{h}_{j,m}^{\mathrm{H}}\mathbf{h}_{j,m}\mathbf{h}_{k,m}^{\mathrm{H}}}{\left\|\mathbf{h}_{j,m}\right\|^2} \\ &= \left\|\mathbf{h}_{k,m}\right\|^2 - \sum_{j\neq k, j\in\mathcal{K}}\left|h_{k,l_{j,m}^*}\right|^2 = \sum_{l\in\widetilde{\mathcal{L}}_{m,k}}\left|h_{k,l}\right|^2,\end{aligned} \tag{13}$$

where $\widetilde{\mathcal{L}}_{m,k} = \mathcal{L}_m - \left\{l_{j,m}^*\right\}_{j\in\mathcal{K}, j\neq k}$, $l_{j,m}^*$ represents the the jth UE in the mth EDU closest RRU in distance, each EDU needs to complete this operation. Following the representation, we define $\mathcal{A}_{m,k} \triangleq \mathrm{Unique}\left(\left\{l_{m,n}^\star = \arg\max\limits_{l\in\mathcal{L}_m}\beta_{l,n} | \forall n \neq k\right\}\right)$, $\widetilde{\mathcal{L}}_{m,k}$ can be expressed as $\widetilde{\mathcal{L}}_{m,k} = \mathcal{L}_m / \mathcal{A}_{m,k}$.

Lemma 1:If the random variable Y_i follows an independent gamma distribution, i.e: $Y_i \sim \Gamma\left(k_i, \theta_i\right)$, so $f_{Y_i}\left(y\right) = \frac{1}{{\theta_i}^{k_i}\Gamma(k_i)} y^{k_i-1} e^{-\frac{y}{\theta_i}}$, [18] provides an expression for gamma approximation, with multiple gamma distributed random variables Y_i. The distribution of $\sum\limits_i Y_i$ can be approximated as:

$$\sum_i Y_i \sim \Gamma\left(\frac{\left(\sum\limits_i k_i\theta_i\right)^2}{\sum\limits_i k_i\theta_i^2}, \frac{\sum\limits_i k_i\theta_i^2}{\sum\limits_i k_i\theta_i}\right). \tag{14}$$

We can obtain:

$$\frac{\det\left(\mathbf{G}_i\right)}{\det\left(\mathbf{G}_{i,k}\right)} \approx \sum_{l\in\widetilde{\mathcal{L}}_{m,k}}\left|h_{k,l}\right|^2 = \Lambda_{k,m} \sim \Gamma(\Psi_{k,m}, \Phi_{k,m}). \tag{15}$$

where $\Phi_{k,m} \triangleq \frac{\sum\limits_{\tilde{l}\in\widetilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}^2}{\sum\limits_{\tilde{l}\in\widetilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}}$,$\Psi_{k,m} \triangleq \frac{\left(\sum\limits_{\tilde{l}\in\widetilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}\right)^2}{\sum\limits_{\tilde{l}\in\widetilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}^2}$.

Therefore, due to the independence of the channel and the gamma approximation Lemma 1, the sum of multiple gamma approximations is still a gamma distribution, and the form of the sum of multiple expected values should theoretically be consistent with the form of direct approximation:

$$R_k \le R_k^{\mathrm{ub}} = \log_2\left(1 + p_k \mathbb{E}_{\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_M}\left\{\sum_{m=1}^{M}\Lambda_{k,m}\right\}\right) = \log_2\left(1 + p_k\left\{\sum_{m=1}^{M}\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}\right\}\right). \tag{16}$$

For the lower bound of achievable rate(7), we introduce the Inverted Gamma distribution [12], $X_{k,m} = \frac{1}{\Lambda_{k,m}} \sim \mathbf{Inverted\ Gamma}(\Psi_{k,m}, \Phi_{k,m})$,where $\Phi_{k,m} \triangleq \frac{\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}^2}{\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}}$,$\Psi_{k,m} \triangleq \frac{\left(\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}\right)^2}{\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}^2}$. And the expectation of this random variable meets:$\mathbb{E}\{X_{k,m}\} = \frac{1}{\Phi_{k,m}(\Psi_{k,m}-1)}$

So based on the above derivation, it can be concluded that:

$$\frac{1}{\sum_{m=1}^{M}\|\mathbf{v}_{k,m}\|^2} = \frac{1}{\sum_{m=1}^{M}\frac{\det(\mathbf{Z}_{m,k})}{\det(\mathbf{Z}_m)}} \overset{L_m\gg K}{\approx} \frac{1}{\sum_{m=1}^{M}\left(\frac{1}{\sum_{l\in\tilde{\mathcal{L}}_{m,k}}|h_{k,l}|^2}\right)} = \frac{1}{\sum_{m=1}^{M}\left(\frac{1}{\Lambda_{k,m}}\right)} = \frac{1}{\sum_{m=1}^{M}X_{k,m}}. \tag{17}$$

Therefore, the lower bound (7) can be expressed as:

$$R_k^{\mathrm{lb}} = \log_2\left(1 + \frac{p_k M^2}{\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_M\left\{\sum_{m=1}^{M}X_{k,m}\right\}}\right) = \log_2\left(1 + \frac{p_k M^2}{\sum_{m=1}^{M}\frac{1}{\Phi_{k,m}(\Psi_{k,m}-1)}}\right). \tag{18}$$

Obviously, for the scenario of a single EDU, (16) and (18) are consistent with the form of the article [12] and the upper and lower bound results are consistent with the form of traditional large-scale MIMO when all RRUs are located in the same position.

B. Large-scale fading analysis

For the convenience of analysis and not generality, we choose a free space fading model, assuming that the large-scale model of the channel is:

$$\beta_{k,l} = d_{k,l}^{-\alpha}, \tag{19}$$

where $d_{k,l}$ is the distance between user k and RRU l, and α is the path fading coefficient. We note that $d_{k,l} > 0$is valid in any case.

For large-scale fading coefficients, we aim to maximize the achievable rate of user k. This problem is represented by its upper or lower bound on the rate (16), (18), and we use the upper bound to transform the problem into:

$$\max\left\{\log_2\left(1 + p_k\left\{\sum_{m=1}^{M}\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}\right\}\right)\right\}. \tag{20}$$

It is difficult to deal with directly by optimizing the form of (20). Consider the function $f(x) = x^{-\alpha}$, which is a convex function, because $f''(x) = \alpha(\alpha+1)x^{-\alpha-2} > 0$ for all $x > 0$, we can use Jensen inequality to get:

$$\frac{1}{n}\sum_{i=1}^{n}d_i{}^{-\alpha} = \frac{1}{n}\sum_{i=1}^{n}f(d_i) \ge f\left(\frac{1}{n}\sum_{i=1}^{n}d_i\right) = \left(\frac{1}{n}\sum_{i=1}^{n}d_i\right)^{-\alpha}. \tag{21}$$

For ease of processing, we will use the square of the distance to represent the above equation:

$$\frac{1}{n}\sum_{i=1}^{n}d_i{}^{-\alpha} = \frac{1}{n}\sum_{i=1}^{n}\left(d_i{}^2\right)^{-\frac{\alpha}{2}} = \frac{1}{n}\sum_{i=1}^{n}f\left(d_i{}^2\right) \ge f\left(\frac{1}{n}\sum_{i=1}^{n}d_i{}^2\right) = \left(\frac{1}{n}\sum_{i=1}^{n}d_i{}^2\right)^{-\frac{\alpha}{2}}. \tag{22}$$

According to Jensen inequality (22), we can approximate the (20) problem and obtain its lower bound as follows:

$$R_k^{\mathrm{ub}} \ge \log_2\left(1 + p_k\left\{\sum_{m=1}^{M}\left(\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}d_{k,l}{}^2\right)^{-\frac{\alpha}{2}}\right\}\right). \tag{23}$$

By maximizing the design of its lower bound, the original problem of maximizing SINR is transformed into the problem of minimizing the distance between user k and RRU, as follows:

$$\max\left\{\sum_{m=1}^{M}\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}\beta_{\tilde{l},k}\right\} \to \min\left\{\sum_{m=1}^{M}\left(\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}d_{k,l}{}^2\right)\right\}. \tag{24}$$

Assuming there are L RRUs on the plane, the coordinates of these RRUs are $(x_1, y_1), (x_2, y_2), \cdots, (x_L, y_L)$, and the coordinates of user k are (x_k, y_k). The distance from user k to RRU l is represented as follows:

$$d_{k,l} = \sqrt{(x_k - x_l)^2 + (y_k - y_l)^2}. \tag{25}$$

The distance and square of user k to all RRUs are represented as follows:

$$d_k^2 = \left(\sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}d_{k,l}{}^2\right) = \sum_{\tilde{l}\in\tilde{\mathcal{L}}_{m,k}}\left[(x_k - x_l)^2 + (y_k - y_l)\right]^2. \tag{26}$$

Assuming there is an EDU in the scene, which is traditional centralized processing, we need to minimize d_k^2, we assume that users follow a uniform distribution in the plane. Therefore, according to Concavity and convexity of associative function, taking partial derivatives of x and y for the sum of squared distances (26) can obtain the following form,

$$\frac{\partial d_k^2}{\partial x_k} = \frac{\partial d_k^2}{\partial y_k} = 0. \tag{27}$$

Based on the necessary conditions for extreme values, we require d_k^2 's artial derivatives x_k and y_k is equal to 0, so:

$$(x_k, y_k) = \frac{1}{n} \sum_{\tilde{l} \in \tilde{\mathcal{L}}_{m,k}} (x_l, y_l). \tag{28}$$

That is to say, the distance from user k to all RRUs is the smallest, and only if user k is located at the center of gravity of all points, then the total distance sum (26) is the smallest, that is, the lower bound (24) is the largest.

According to (23), we consider the case of multiple EDUs. At this point, we need users to be at the center of gravity of RRUs in multiple groups of EDUs, and the center of gravity of users is at the midpoint of the plane. Therefore, each group of RRUs needs to satisfy a uniform distribution as much as possible, with better performance. Obviously, as the number of EDUs increases, this premise becomes less easily satisfied. That is, as the number of EDUs increases, the difficulty of achieving statistically uniform distribution in a space increases. In response to the above problem, In the next chapter, we will design a modified graph coloring algorithm that minimizes the overall distance. Due to the randomness of the RRU distribution, we will try to interweave the RRUs between groups as much as possible to ensure that they are uniformly distributed and that the centers of gravity overlap as much as possible

C. RRU correlation method

A genetic algorithm based on the distance relationship between RRUs was proposed to associate EDUs and RRUs [16]. The designed fitting function is:

$$f(x) = \frac{1}{\sum\limits_{p \in \mathcal{P}} \sum\limits_{q \in \mathcal{Q}} \cdots \sum\limits_{u \in \mathcal{U}} \sum\limits_{v \in \mathcal{V}} \left(d_{p,q}^2 + \cdots + d_{p,v}^2 + \cdots + d_{u,v}^2 \right)^{1/2}}, \tag{29}$$

where $d_{p,v}$ is the distance between the pth RRU and the vth RRU, $\mathcal{P}, \mathcal{Q}, \cdots, \mathcal{U}, \mathcal{V}$ is the $p, q, \cdots, u, v$'s corresponding RRU grouping of EDUs, and $\mathcal{T}$ is a complete set of all RRUs, and constraints can be described as:

$$\mathcal{P} \cup \mathcal{Q} \cup \cdots \mathcal{U} \cup \mathcal{V} = \mathcal{T}, \tag{30a}$$

$$(\mathcal{P} \cap \mathcal{Q}) \cup (\mathcal{P} \cap \mathcal{U}) \cup (\mathcal{P} \cap \mathcal{V}) \cup \cdots \cup (\mathcal{U} \cap \mathcal{V}) = \varnothing, \tag{30b}$$

$$\|\mathcal{P}| - |\mathcal{Q}\| \leq 1, \|\mathcal{P}| - |\mathcal{V}\| \leq 1, \cdots, \|\mathcal{U}| - |\mathcal{V}\| \leq 1. \tag{30c}$$

The specific implementation process of the genetic algorithm in this article is shown in the algorithm 1. However, based on the derivation in the previous section, we will modify the fitness function to the following form, which will significantly reduce computational complexity.

$$f(x) = \frac{1}{\left(\bar{d}_{\mathcal{P}}^2 + \bar{d}_{\mathcal{Q}}^2 + \cdots + \bar{d}_{\mathcal{V}}^2 \right)^{1/2}}, \tag{31}$$

where $\bar{d}_{\mathcal{V}}$ represented as the distance between the center of gravity of all RRUs in RRU group $\mathcal{V}$ and the origin point.

Algorithm 1 GA-based interlacing Algorithm

Input: Number of RRU L, Distance matrix between RRUs $\mathbf{D}$, Crossover rate c, Mutation rate m, population size n_p

Pop = makingPopulation(n_p) under constraint of (30);
$i = 0$;
while $i < gen$ **do**
  Pop.scores = fitnessFunction(Pop) using (31);
  newPop = [ ];
  $j = 0$;
  while $j < numberofchildP$ **do**
    [child1,child2] = selection(Pop)
    Crossover(child1,child2,c);
    Mutation(child1,child2,m);
    if child meet system constraints (30) **then**
      newPop.append(child1)
      newPop.append(child2)
      $j = j + 2$
    end if
  end while
  newPop.scores = fitnessFunction(newPop) using (31)
  Pop.append(newPop)
  SortBasedOnScroes(Pop)
  Pop = Pop$(0 : N - 1)$
  $i = i + 1$.
end while

Output: Best RRU Group $\mathcal{L}$ = minimumScore(Pop).

IV. Results and discussions

In the Fig. 2, we verified that in the perfect CSI, in order to be as consistent as possible with the centralized processing of Cell-free, we referred to the simulation settings in [12]. There are a total of L single antenna RRUs and K users uniformly distributed within a certain range in the system, and we chose the path loss factor $\alpha = 4$ and number of users $K = 10$.

It is shown that the comparison between the Monte Carlo simulation results using ZF combining and the upper and lower bound performance derived from theory. Through the difference between the EDUs of different groups and the theoretical values, we can see that the more EDU groups, the worse the system performance. At the same time, it also verifies the effectiveness of the derived performance upper and lower bounds (16), (18).

The Fig. 3 shows the tightness between the Monte Carlo simulation results of the average spectral efficiency using ZF combining and the theoretically derived upper and lower bounds. As the number of EDUs increases, the difference between the derived lower bound and the Monte Carlo simulation

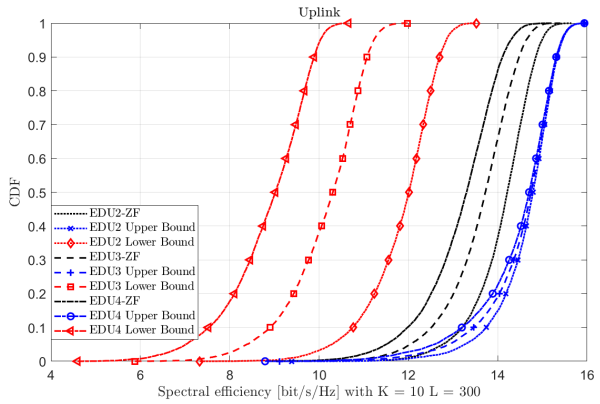

Fig. 2. Verification of upper and lower bounds on the spectral efficiency of uplink users as a function of the number of EDUs, $L = 300$, $N = 1$, $K = 10$.

results increases, while the upper bound can always maintain good closeness This also demonstrates the effectiveness of using upper bounds to describe the spectral efficiency performance of frequency systems.

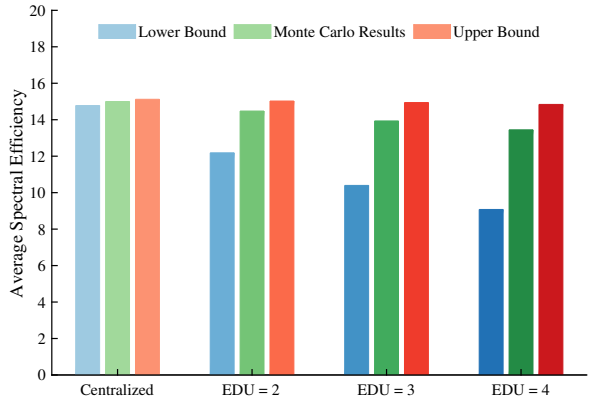

Fig. 3. Comparison of the upper and lower bounds of the average spectral efficiency of uplink users with the number of EDUs, $L = 300$, $N = 1$, $K = 10$.

We adopt a more complex channel combining and precoding scheme simultaneously. Due to the problem of noise power amplification in the ZF merging method, as shown in (4), we consider the influence of multiple antennas and assume $N = 4$. The simulation parameters of the article are shown in Table (I).

In this section, we consider more complex scenarios for simulation. We consider non ideal channel CSI for uplink and downlink, and the system modeling adopts more complex large-scale fading as follows:

$$\beta_{k,l}(\mathrm{dB}) = -30.5 - 36.7 \log_{10}(d_{k,l}) + F_{k,l}, \qquad (32)$$

where $F_{k,l}$ represents the shadow fading effect.

At the same time, due to the existence of channel estimation error and the limitation of the number of pilots, we will

TABLE I
SIMULATION PARAMETERS

Simulation Parameters	Values
Uplink transmission power	200 mW
Antenna height	10 m
Area size	200×200 m^2
The number of RRUs, L	100
Number of antennas per RRU, N	4
Number of UEs, K	24
Number of orthogonal pilots,L_P	24
Transmission bandwidth, B	20 MHz
Carrier frequency, f_c	2 GHz
Noise power spectral density, N_0	-174 dBm/Hz

consider the non-perfect CSI with the existence of pilot pollution. At this time, using MMSE combining method the combination vector at the EDU can be expressed as follows:

$$\mathbf{v}_{k,m} = p_k \left(\sum_{i=1}^{K} p_i \left(\widehat{\mathbf{h}}_{i,m} \widehat{\mathbf{h}}_{i,m}^{\mathrm{H}} + \mathbf{C}_{i,m} \right) + \sigma_{\mathrm{UL}}^2 \mathbf{I}_{mN} \right)^{-1} \widehat{\mathbf{h}}_{k,m}. \qquad (33)$$

where $\mathbf{h}_{i,k} = \hat{\mathbf{h}}_{i,k} + \tilde{\mathbf{h}}_{i,k}$ represent as channel estimation information and channel estimation error, respectively.

At pilot number $L_p = 24$ and user number $K = 48$, the uplink simulation results are shown in Figs. 4. We can see that as the number of EDUs increases, the system and spectral efficiency continue to decrease. The performance of RRU clustering association to EDUs is significantly weaker than the interleaving strategy, which verifies the performance analysis of collaboration in section III.

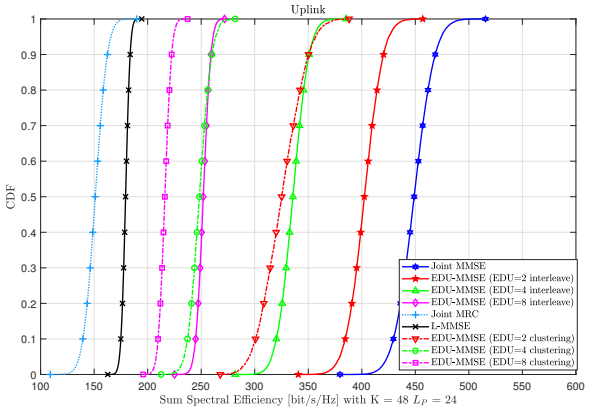

Fig. 4. Uplink sum spectral efficiency v.s. EDU number, $L = 100$, $N = 4$, $K = 48$, $L_p = 24$.

V. CONCLUSIONS

In this paper, we analyzed the spectral efficiency performance of the new full spectrum cell-free wireless access network architecture, and derived its performance upper and lower bounds for the case of multiple EDUs. Besides,

considering the uniform distribution of users for large-scale fading models, we studied the location distribution of RRUs and found that a uniform distribution of RRUs is beneficial for user achievable rate. At the same time, the centers of gravity of multiple groups of RRUs should overlap with the centers of gravity of users as much as possible, which is different from traditional multi-node clustering centralization processing; Finally, with genetic algorithms the superiority of the new architecture and algorithm is demonstrated by comparing the performance of Monte Carlo simulation with multi-RRU clustering and group collaboration processing.

REFERENCES

[1] X. You, C.-X. Wang, J. Huang, X. Gao, Z. Zhang, M. Wang, Y. Huang, C. Zhang, Y. Jiang, J. Wang *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Scientia Sinica Informationis*, vol. 64, no. 1, pp. 1–74, 2021.

[2] 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.

[3] X. You, D. Wang, and J. Wang, *Distributed MIMO and Cell-Free Mobile Communication*. Springer, 2021.

[4] Y. Xiaohu, "Shannon theory and future 6G′s technique potentials," *Scientia Sinica Informationis*, vol. 50, no. 9, pp. 1377–1394, 2020.

[5] M. Matthaiou, H. Q. Ngo, P. J. Smith, H. Tataria, and S. Jin, "Massive MIMO with a generalized channel model: Fundamental aspects," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2019, pp. 1–5.

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

[7] A. Papazafeiropoulos, P. Kourtessis, M. Di Renzo, S. Chatzinotas, and J. M. Senior, "Performance analysis of cell-free massive MIMO systems: A stochastic geometry approach," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 3523–3537, 2020.

[8] M. Zbairi, I. Ez-Zazi, and M. Arioua, "Uplink spectral efficiency of cell free massive MIMO based on stochastic geometry approach," in *2021 4th International Conference on Advanced Communication Technologies and Networking (CommNet)*. IEEE, 2021, pp. 1–6.

[9] 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, 2013.

[10] Z. Chen and E. Björnson, "Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5205–5219, 2018.

[11] D. Wang, X. You, Y. Huang, W. Xu, J. Li, P. Zhu, Y. Jiang, Y. Cao, X. Xia, Z. Zhang *et al.*, "Full-spectrum cell-free RAN for 6G systems: system design and experimental results," *Science China Information Sciences*, vol. 66, no. 3, pp. 1–14, 2023.

[12] P. Liu, K. Luo, D. Chen, and T. Jiang, "Spectral efficiency analysis of cell-free massive MIMO systems with zero-forcing detector," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 795–807, 2019.

[13] J. Wang and L. Dai, "Asymptotic rate analysis of downlink multi-user systems with co-located and distributed antennas," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3046–3058, 2015.

[14] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, 2013.

[15] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, 2019.

[16] Y. Cao, Z. Zhang, X. Xia, P. Xin, D. Liu, K. Zheng, M. Lou, J. Jin, Q. Wang, D. Wang *et al.*, "From ORAN to Cell-Free RAN: Architecture, performance analysis, testbeds and trials," *arXiv preprint arXiv:2301.12804*, 2023.

[17] J. Zhang, J. Zhang, E. Björnson, and B. Ai, "Local partial zero-forcing combining for cell-free massive MIMO systems," *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8459–8473, 2021.

[18] L. Zhang, P. Zhu, J. Li, and J. Cao, "Downlink ergodic rate analysis of das with linear beamforming under pilot contamination," in *2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)*. IEEE, 2017, pp. 1–6.