

Improving Fairness for Cell-Free Massive MIMO Through Interference-Aware Massive Access

Shuaifei Chen ¹, Graduate Student Member, IEEE,

Jiayi Zhang ², Senior Member, IEEE,

Emil Björnson ³, Fellow, IEEE, and Bo Ai ⁴, Fellow, IEEE

Abstract—Cell-free massive multiple-input multiple-output (CF mMIMO) provides good interference management by coordinating many more access points (APs) than user equipments (UEs). It becomes challenging to determine which APs should serve which UEs with which pilots when the number of UEs approximates the number of APs and far exceeds the number of pilots. Compared to the previous work, a better compromise between spectral efficiency (SE) and implementation simplicity is needed in such massive access scenarios. This paper proposes an interference-aware massive access (IAMA) scheme realizing joint AP-UE association and pilot assignment for CF mMIMO by exploiting the large-scale interference features. We propose an interference-aware reward as a novel performance metric and use it to develop two iterative algorithms to optimize the association and pilot assignment. The numerical results show a prominent advantage of our IAMA scheme over the benchmark schemes in terms of the user fairness and the average SE.

Index Terms—Massive access, interference-aware, cell-free massive MIMO, pilot assignment, user scheduling.

I. INTRODUCTION

Cell-free massive multiple-input multiple-output (CF mMIMO) is recognized as a promising paradigm for the sixth-generation (6G) networks [1]. The core idea is to coordinate a large number of distributed access points (APs) with a central processing unit (CPU) to provide an almost uniform service quality for the user equipments (UEs) in the coverage area [2], [3]. However, the enormous wireless devices collaborating not only improves the network throughput, but also introduces

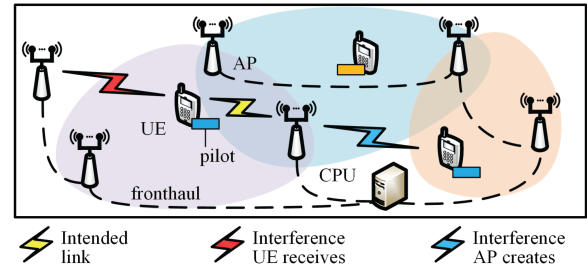


Fig. 1. Interference relationships in scalable CF mMIMO systems.

a huge challenge on the interference management, especially in the massive access scenarios where the pilot reuse ratio is high.

User access comprises *pilot assignment* and *AP-UE association*. The former assigns each UE a pilot to acquire the channel state information (CSI) and the latter associates each UE with at least one AP to perform coherent transmission. User access becomes challenging as the network gets more crowded since the increasing UE density implies inevitable pilot reuse among the UEs, which causes substantial inter-user interference, namely *pilot contamination* [4]. Also, a high UE density aggravates the competition among UEs for accessing their associated APs, which motivates more delicate AP-UE association schemes. Moreover, user access should maintain the signal processing operated at each AP under limited complexity and resource requirements to make the system *scalable* [5]. The simplest but naive scheme is to assign the pilots and APs at random [6]. There exist greedy schemes that refine the spectral efficiency (SE) of the weakest UE iteratively [6], but cannot guarantee convergence to the global optimal pilot assignment results. The graph-based schemes formulate the assignment problems as graph problems (such as the graph coloring [7], the weighted matching [8], and the Max k -Cut [9]), and solve them with the corresponding algorithms. Scalable schemes are proposed in [5] and [10], whereof the former performs joint AP-UE association and pilot assignment, and the latter clusters UEs such that the UEs in the same cluster share the same pilot.

In a wireless network, being aware of the interference features is critical to the transmission design. This can be characterized by the treating-interference-as-noise (TIN) optimality conditions from an information-theoretic perspective [11], [12], [13], which reflect the interference relationship between an intended link and the two *most* influential interfering links corresponding to the intended UE and AP, respectively (as illustrated in Fig. 1). Especially, [14] investigated the probability that the TIN conditions hold in a CF mMIMO system using stochastic geometry, which is directly related to the interference relationships between the APs and the UEs. Due to the implementation simplicity, robustness to channel uncertainty, and good characterization of interference, the TIN optimality conditions are exploited for designing interference-aware schemes for scheduling in cellular systems in [15] and power control which further improves the SE in [12].

Motivated by the discussion above, we propose an interference-aware massive access (IAMA) scheme for scalable CF mMIMO, where the interference features are transformed into a useful performance metric for scheme design. Two assignment algorithms are developed for joint AP-UE association and pilot assignment to maximize the user fairness or the average SE. The SE improvements achieved by our IAMA scheme are demonstrated by the numerical results.

Manuscript received 9 September 2022; revised 14 November 2022; accepted 1 December 2022. Date of publication 7 December 2022; date of current version 18 April 2023. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 2021YJS001, in part by the National Key R&D Program of China under Grant 2020YFB1807201, in part by the National Natural Science Foundation of China under Grants 61971027 and 62221001, in part by Beijing Natural Science Foundation under Grant L202013, in part by the Natural Science Foundation of Jiangsu Province, Major Project under Grant BK2021002, in part by the Fundamental Research Funds for the Central Universities under Grant 2022JBQY004, in part by ZTE Corporation, and in part by the State Key Laboratory of Mobile Network and Mobile Multimedia Technology. The work of Emil Björnson was supported by the Swedish Foundation for Strategic Research under Grant FFL18-0277. The review of this article was coordinated by Dr. Nghi H. Tran. (Corresponding authors: Jiayi Zhang; Bo Ai.)

Shuaifei Chen and Jiayi Zhang are with the School of Electronic and Information Engineering and the Frontiers Science Center for Smart Highspeed Railway System, Beijing Jiaotong University, Beijing 100044, China (e-mail: shuaifeichen@bjtu.edu.cn; jiayizhang@bjtu.edu.cn).

Emil Björnson is with the Department of Computer Science, KTH Royal Institute of Technology, SE-16440 Kista, Sweden (e-mail: emilbj@kth.se).

Bo Ai is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China, also with the Frontiers Science Center for Smart High-Speed Railway System and the Henan Joint International Research Laboratory of Intelligent Networking and Data Analysis, Zhengzhou University, Zhengzhou 450001, China, and also with the Research Center of Networks and Communications, Peng Cheng Laboratory, Shenzhen 518066, China (e-mail: boai@bjtu.edu.cn).

Digital Object Identifier 10.1109/TVT.2022.3227408

Notation: Boldface lowercase letters, \mathbf{x} , denote column vectors, boldface uppercase letters, \mathbf{X} , denote matrices, and calligraphic uppercase letters, \mathcal{A} , denote sets. Superscript H denotes the conjugate transpose. The $n \times n$ identity matrix and zero matrix are \mathbf{I}_n and $\mathbf{0}_n$, respectively. $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R})$ denotes the multi-variate circularly symmetric complex Gaussian distribution with correlation \mathbf{R} . $\mathbb{E}\{\cdot\}$ denotes the expected value.

II. CF mMIMO SYSTEM MODEL

We consider a CF mMIMO system consisting of K single-antenna UEs and L APs, each equipped with N antennas. As illustrated in Fig. 1, the APs are connected via fronthaul connections to a CPU, which is responsible for coordinating and processing the signals of all UEs. The user-centric CF architecture is adopted [5], where each UE is associated with a subset of the APs. This procedure is elaborated in Section III. For now, we let $\mathcal{M}_k \subset \{1, \dots, L\}$ denote the subset of APs associated with UE k and let $a_{kl} = 1$ if $l \in \mathcal{M}_k$ and $a_{kl} = 0$ otherwise, $\forall k, l$.

We adopt the standard block fading model where the channel between UE k and AP l , denoted by $\mathbf{h}_{kl} \in \mathbb{C}^N$, is constant in time-frequency coherence blocks of τ_c channel uses [4]. In each block, the channels are assumed to be subject to spatially correlated Rayleigh fading, i.e., $\mathbf{h}_{kl} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R}_{kl})$, where $\mathbf{R}_{kl} \in \mathbb{C}^{N \times N}$ is the spatial correlation matrix and $\beta_{kl} \triangleq \text{tr}(\frac{\mathbf{R}_{kl}}{N})$ is the large-scale fading coefficient (LSFC) that describes pathloss and shadowing. We assume that AP l knows the correlation matrices $\{\mathbf{R}_{kl} : k = 1, \dots, K\}$, which represent the long-term channel statistics. These correlation matrices can be accurately estimated using classical methods [4]. We consider the downlink operation, where each block dedicates τ_p channel uses for pilots and the remaining $\tau_c - \tau_p$ channel uses for payload data.

A. Data Transmission and Spectral Efficiency

During the channel estimation, $\tau_p < K$ holds in massive access scenarios due to the coherence block length limitation caused by natural channel variations in the time and frequency domain. We adopt that the pilots are selected from a pool of τ_p orthogonal sequences, and thus, some UEs have to share the same pilot. We let \mathcal{P}_t denote the set of UEs sharing pilot t and refer to these UEs as *co-pilot UEs*. The pilot assignment is elaborated in Section III. For now, we denote by $t_k \in \{1, \dots, \tau_p\}$ the index of the pilot assigned to UE k . When the UEs in \mathcal{P}_{t_k} transmit pilot t_k , the pilot signal $\mathbf{y}_{t_k l}^p \in \mathbb{C}^N$ received at AP l is [4, Sec. 3]

$$\mathbf{y}_{t_k l}^p = \sum_{i \in \mathcal{P}_{t_k}} \sqrt{\tau_p \rho_p} \mathbf{h}_{il} + \mathbf{n}_{t_k l}, \quad (1)$$

where ρ_p represents the pilot transmit power and $\mathbf{n}_{t_k l} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ is the thermal noise. The *minimum mean-squared-error (MMSE)* estimate of \mathbf{h}_{kl} is [4, Sec. 3]

$$\hat{\mathbf{h}}_{kl} = \sqrt{\tau_p \rho_p} \mathbf{R}_{kl} \Psi_{t_k l}^{-1} \mathbf{y}_{t_k l}^p \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \tau_p \rho_p \mathbf{R}_{kl} \Psi_{t_k l}^{-1} \mathbf{R}_{kl}), \quad (2)$$

where $\Psi_{t_k l} = \sum_{i \in \mathcal{P}_{t_k}} \tau_p \rho_p \mathbf{R}_{il} + \sigma^2 \mathbf{I}_N$ is the correlation matrix of $\mathbf{y}_{t_k l}^p$ in (1).

Let $\mathbf{w}_{kl} = \bar{\mathbf{w}}_{kl} / \sqrt{\mathbb{E}\{\|\bar{\mathbf{w}}_{kl}\|^2\}}$ denote the normalized precoder that AP l selects for transmission to UE k such that $\mathbb{E}\{\|\mathbf{w}_{kl}\|^2\} = 1$. Then the received signal at UE k is

$$y_k^{\text{dl}} = \sum_{l=1}^L \mathbf{h}_{kl}^H \sum_{i=1}^K a_{il} \mathbf{w}_{il} s_i + n_k, \quad (3)$$

where $s_k \in \mathbb{C}$ is the independent unit-power payload signal intended for UE k , $\rho_{il} \geq 0$ is the transmit power that AP l assigns to UE i , and

$n_k \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ is the receiver noise. The total transmission power of each AP is upper bounded by the maximum power ρ_{dl} .

We employ the widely used *hardening bound* [4, Th. 4.6],

$$\text{SE}_k = (1 - \tau_p / \tau_c) \log_2(1 + \text{SINR}_k), \quad (4)$$

to compute the achievable downlink SE, where the signal-to-interference-and-noise ratio (SINR) is given by

$$\text{SINR}_k = \frac{\left| \mathbb{E} \left\{ \sum_{l=1}^L a_{kl} \mathbf{h}_{kl}^H \mathbf{w}_{kl} \right\} \right|^2}{\sum_{i=1}^K \left| \mathbb{E} \left\{ \sum_{l=1}^L a_{il} \mathbf{h}_{kl}^H \mathbf{w}_{il} \right\} \right|^2 - \left| \mathbb{E} \left\{ \sum_{l=1}^L a_{kl} \mathbf{h}_{kl}^H \mathbf{w}_{kl} \right\} \right|^2}. \quad (5)$$

The SE expression in (4) holds for any scalable precoding scheme, e.g., the local partial MMSE (LP-MMSE) precoding [5] or the classical maximum ratio (MR) precoding [4].

B. Interference-Aware Rewards (IARs) for CF mMIMO

The TIN optimality conditions indicate when the boundary of the capacity region is approached within a constant gap and power control is essential to do that [11]. In cellular networks, interference-aware scheduling schemes are designed based on the following TIN condition [15]:

$$\kappa \text{SNR}^\mu \geq \max \text{INR}_{\text{ap}} \cdot \max \text{INR}_{\text{ue}}, \quad (6)$$

where SNR denotes the signal-to-noise ratio (SNR) of the intended link, INR_{ap} denotes the interference-to-noise ratio (INR) of the link between the intended AP and the interfering UE and INR_{ue} denotes the INR of the link between the intended UE and the interfering AP. In (6), SNR, INR_{ap} , and INR_{ue} rely on statistical knowledge, which is averaged concerning the fast fading. Parameters $\kappa \geq 1$ and $1 \leq \mu \leq 2$ are introduced in [15] for system optimization.

The condition in (6) will not directly be applied to our considered CF user access since a UE is served by multiple APs. In our case, the interference comes from the imperfect CSI caused by the pilot reuse, which will both reduce the channel estimation quality and make it harder to suppress interference among the co-pilot UEs. Similar to [14], we denote by

$$\mathcal{S}_{t,k}^{\text{ue}} = \mathcal{P}_t \setminus \{k\} \quad (7)$$

the set of UEs sharing pilot t except UE k , which are referred to as the *interfering UEs* of UE k when UE k is assigned with pilot t . Also, we denote by

$$\mathcal{S}_{t,k}^{\text{ap}} = \bigcup_{i \in \mathcal{S}_{t,k}^{\text{ue}}} \mathcal{M}_i \setminus \mathcal{M}_k \quad (8)$$

the set of the *interfering APs* of UE k when UE k is assigned with pilot t . Unlike the scheduling in [15] where only the links fulfilling the TIN condition in (6) are established, we propose a novel performance metric tailored for CF mMIMO, as

$$\text{iar}_{kl}^t \triangleq \kappa(\beta_{kl})^\mu - \max_{i \in \mathcal{S}_{t,k}^{\text{ue}}} \beta_{il} \cdot \max_{j \in \mathcal{S}_{t,k}^{\text{ap}}} \beta_{kj}, \quad \forall k, l, t, \quad (9)$$

which reflects the strength of an intended link compared to the interference by only utilizing the LSFCs. We refer to iar_{kl}^t defined in (9) as the *interference-aware reward (IAR)* of a tuple (k, l, t) , which corresponds to the strength of the link between AP l and UE k when pilot t is used.

III. INTERFERENCE-AWARE MASSIVE ACCESS (IAMA)

In this section, we propose an IAMA scheme that uses the proposed IARs in (9) with the goal of improving the SE of the majority of UEs, e.g., the 90%-likely SE, which is a well-used performance criterion representing the SE that can be provided to 90% of all UEs [2]. Since a

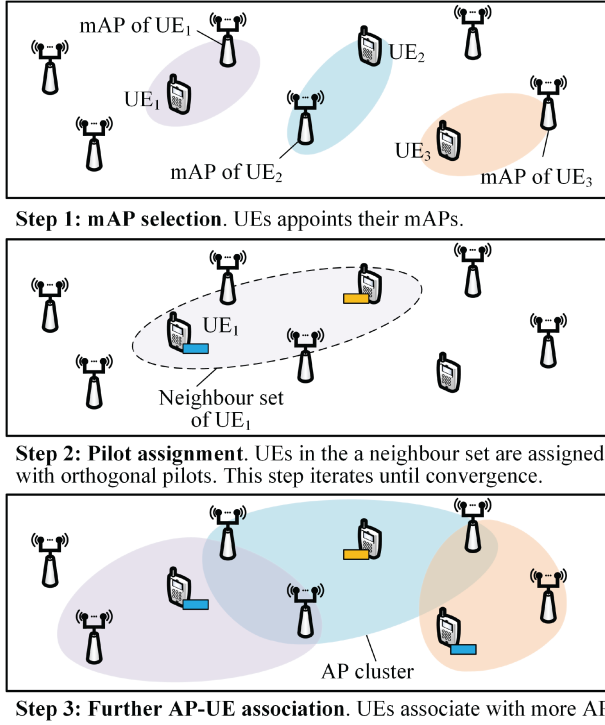


Fig. 2. The proposed IAMA scheme for joint pilot assignment and AP-UE association operates through three main steps.

new UE needs to be assigned a pilot to perform coherent transmission with its associated APs when it accesses the network, two constraints should be met during the AP-UE-pilot association:

- Each UE is associated with at least one AP to not being dropped from service inadvertently;
- Each AP serves at most one UE per pilot to avoid causing substantial pilot contamination.

As illustrated in Fig. 2, our proposed IAMA scheme operates through three steps: 1) Master AP (mAP) selection; 2) pilot assignment; and 3) further AP-UE association. Since only the LSFCs among the APs and the UEs are employed at the CPU, the IAMA scheme works for many coherence blocks.

Further details on above steps are provided later in this section. Before that, recall that we use $a_{kl} \in \{0, 1\}$, $\forall k, l$, to indicate the AP-UE association, we introduce another binary notation $b_{kl} \in \{0, 1\}$, $\forall k, l$, to ensure the association algorithm convergence. More precisely, during the access procedure, $b_{kl} = 1$ prevents the association between AP l and UE k from being considered again, and $b_{kl} = 0$ otherwise.

A. mAP Selection

Similar to [5], each UE first appoints a mAP assisting in the following pilot assignment and AP-UE association. One well-used approach is to let each UE select the AP with largest LSFC as its mAP, however, with the risk that an AP is selected as the mAP by more than τ_p UEs. This motivates us to develop the Multiple-UEs Single-AP-each (MUSA) algorithm to assign each UE to one mAP while each AP is assigned to at most τ_p UEs as their common mAP. The MUSA algorithm tries to maximize the total LSFC of all associated AP-UE pairs, i.e., $\sum_{k,l} \beta_{kl} \cdot a_{kl}$, and consists of the following steps:

- 1) Each UE initially appoints the AP that it has the largest LSFC to as its *candidate* mAP.

Algorithm 1. MUSA Assignment.

Input: $\{\beta_{kl} : \forall k, l\}$, τ_p
Output: $\{a_{kl} : \forall k, l\}$

- 1 **Initiation:** $a_{kl} = 0, b_{kl} = 0, \forall k, l$;
- 2 **for** UE $k = 1, \dots, K$ **do**
- 3 Select AP $\ell = \arg \max_l \beta_{kl}$ as its *candidate* mAP and let $a_{k\ell} \leftarrow 1$;
- 4 Include the *overburdened* APs in $\mathcal{C}_{ap} = \{l : \sum_k a_{kl} > \tau_p\}$;
- 5 **while** $\mathcal{C}_{ap} \neq \emptyset$ and $\sum_k b_{kl} < K, \exists l \in \mathcal{C}_{ap}$ **do**
- 6 **for** overburdened AP $l \in \mathcal{C}_{ap}$ **do**
- 7 Include the UEs selecting AP l in $\mathcal{C}_{ue} = \{k : a_{kl} = 1, l \in \mathcal{C}_{ap}\}$;
- 8 Let $\Delta_k = 0, k \in \mathcal{C}_{ue}$;
- 9 **for** UE $k \in \mathcal{C}_{ue}$ **do**
- 10 Find *alternative* mAP $\ell_k = \arg \max_{j \neq l, b_{kj} \neq 1} \beta_{kj}$ and compute LSFC loss $\Delta_k = \beta_{kl} - \beta_{k\ell_k}$;
- 11 Find UE $i = \arg \min_{k \in \mathcal{C}_{ue}} \Delta_k$, replace its mAP as AP ℓ_i by $a_{i\ell} \leftarrow 0$ and $a_{i\ell_i} \leftarrow 1$, and mark the association between AP l and UE i by $b_{il} \leftarrow 1$;
- 12 Update cluster $\mathcal{C}_{ap} = \{l : \sum_k a_{kl} > \tau_p\}$;

- 2) Find the *overburdened* APs selected by more than τ_p UEs and include them in cluster $\mathcal{C}_{ap} = \{l : \sum_k a_{kl} > \tau_p\}$.
- 3) For each AP $l \in \mathcal{C}_{ap}$, include the UEs selecting AP l as the mAP in cluster $\mathcal{C}_{ue} = \{k : a_{kl} = 1, l \in \mathcal{C}_{ap}\}$. For each UE $k \in \mathcal{C}_{ue}$, compute the *LSFC loss* as $\Delta_k \triangleq \beta_{kl} - \beta_{k\ell_k}$, where AP $\ell_k = \arg \max_{j \neq l, b_{kj} \neq 1} \beta_{kj}$ is the *alternative* mAP having the largest LSFC with UE k except AP l . Find UE $i = \arg \min_{k \in \mathcal{C}_{ue}} \Delta_k$ with the smallest LSFC loss, replace the mAP of UE i by its AP ℓ_i , and mark the association between AP l and UE i by letting $b_{il} = 1$.
- 4) Repeat steps 2) and 3) until $\mathcal{C}_{ap} = \emptyset$ or $\sum_k b_{kl} = K, \forall l \in \mathcal{C}_{ap}$. The pseudo-code of MUSA is summarized in Algorithm 1.

B. Pilot Assignment

To reduce the pilot contamination, the UEs prefer to be assigned pilots that are orthogonal to their neighbouring UEs. With this consideration in mind, we propose the following IAR-based pilot assignment scheme:

- 1) Each UE selects a pilot from the τ_p pilots at random.
- 2) Consider a generic UE k and include it and the $\tau_p - 1$ neighbouring UEs having the largest LSFCs with the mAP of UE k (i.e., AP ℓ_k) in UE k 's *neighbour set* \mathcal{N}_k , where $|\mathcal{N}_k| = \tau_p$. By using (9), compute all *potential* IARs of the tuples (i, ℓ_k, t) , i.e., $\{\text{iar}_{i\ell_k}^t : t = 1, \dots, \tau_p, i \in \mathcal{N}_k\}$.
- 3) Update the pilot assignment in \mathcal{N}_k by performing the MUSA assignment in Algorithm 1, where the inputs are replaced by $\{\text{iar}_{i\ell_k}^t\}$ and integer 1, and the output indicates which pilot is assigned to which UE in \mathcal{N}_k .
- 4) Repeat steps 2) and 3) until the maximum number of allowed iterations is reached or convergence, measured by the change in the sum IAR $\sum_{i=1}^K \text{iar}_{i\ell_k}^t$.

Step 3) implies that MUSA assigns each UE one pilot while each pilot is assigned to at most one UE, based on the IARs $\{\text{iar}_{i\ell_k}^t : t = 1, \dots, \tau_p, i \in \mathcal{N}_k\}$, $\forall k$.

C. Further AP-UE Association

Given the assigned mAP and pilot, each UE prefers to access more serving APs to improve the diversity gain. This motivates our Multiple-UEs Multiple-APs-each (MUMA) algorithm that optimizes the AP-UE association, based on the *potential* IARs of the tuples (k, l, t_k) , i.e.,

$$\{\text{iar}_{kl}^{t_k} : k = 1, \dots, K, l = 1, \dots, L\}. \quad (10)$$

Recall that although our IAMA scheme aims to improve the SE of most UEs, an option for improving the average SE is also provided. More precisely, the MUMA algorithm either maximizes the user fairness as $\max_{\{a_{kl}\}} \min_k \Sigma_k$, where $\Sigma_k \triangleq \sum_l \text{iar}_{kl}^{t_k} a_{kl}$ denotes the per-UE sum IAR, or maximizes the total sum IAR as $\max_{\{a_{kl}\}} \sum_{k,l} \text{iar}_{kl}^{t_k} a_{kl}$. We use the association between the UEs and their mAPs to initialize $\{b_{kl}, \forall k, l\}$. The MUMA algorithm operates as follows:

- 1) By using (9), compute all potential IARs in (10).
- 2) Each AP associates τ_p UEs with the largest IARs.
If the goal is $\max_{\{a_{kl}\}} \sum_{k,l} \text{iar}_{kl}^{t_k} a_{kl}$, stop algorithm and return $\{a_{kl}\}$; otherwise, continue.
- 3) Find the *weakest* UE $k' = \arg \min_k \Sigma_k$ with the smallest per-UE sum IAR.
- 4) Find the *closest* AP $l' = \arg \max_{l, b_{kl'} \neq 1} \text{iar}_{kl'}^{t_{k'}}$ for UE k' , satisfying $b_{k'l} \neq 1$ with the largest IAR.
- 5) Find the *most distant* UE $k^* = \arg \min_{k, b_{kl'} \neq 1} \text{iar}_{kl'}^{t_k}$ for AP l' , satisfying $b_{kl'} \neq 1$ with the smallest IAR.
- 6) Mark the association between AP l' and UE k' by $b_{k'l'} \leftarrow 1$. If UE k' still has a smaller per-UE sum IAR than UE k^* after taking AP l' from UE k^* , then UE k' takes AP l' from UE k^* ; otherwise, keep status quo.
- 7) Repeat step from 3) to 6) until $\sum_l b_{kl} = L, \exists k$.

The pseudo-code of MUMA is summarized in Algorithm III-C.

D. Benchmark Schemes and Complexity Analysis

Three existing user access schemes are considered as benchmarks, they are “Scalable” with the complexity of $\mathcal{O}(KL + K\tau_p + L\tau_p)$ [5], “Greedy” with the complexity of $\mathcal{O}(3KL + L)$ [6], and “Graph” with the complexity of $\mathcal{O}(KL + K^2/2 + K/2 + \tau_p)$ [9]. The Hungarian scheme proposed in [8] is not considered since Graph offers better SE [9]. For fair comparison in the scalable scenario, each AP serves at most τ_p UEs and allocates its transmit power with the fractional power allocation policy [5].

We consider two use cases of our proposed IAMA scheme: a) “IARsum” for maximizing the sum SE and b) “IARmin” for maximizing the user fairness. The complexity of MUMA depends on computing Δ_k for UE $k \in \mathcal{C}_{ue}$ of APs in \mathcal{C}_{ap} , from line 5 to line 12 in Algorithm 1, which traverses all entries in $\{\beta_{kl}, \forall k, l\}$ at most K times, and, thus, with the complexity of $\mathcal{O}(K^2 L)$ [16]. The complexity of MUMA depends on comparing the IARs of the weak UEs, from line 7 to line 17 in Algorithm III-C, which traverses all entries in $\{\text{iar}_{kl}^{t_k} : \forall k, l\}$ at most L times, and, thus, with the complexity of $\mathcal{O}(KL^2)$ [16]. To sum up, the total complexity of the IAMA scheme is $\mathcal{O}(K^2 L + K\tau_p^3 + KL^2)$. For the considered massive access scenario where $K \approx L \gg \tau_p$, the complexity of the IAMA scheme is dominated by $\mathcal{O}(K^3)$.

IV. NUMERICAL RESULTS

In this section, we quantify the SE performance of the proposed IAMA scheme, in which the results regarding LP-MMSE are obtained from Monte Carlo simulation while the results regarding MR are analytically computed by [5, Cor. 3]. We consider a 0.5×0.5 km² coverage area and use the wrap-around technique to approximate an infinitely massive access scenario, where $L = 50$ APs are deployed at

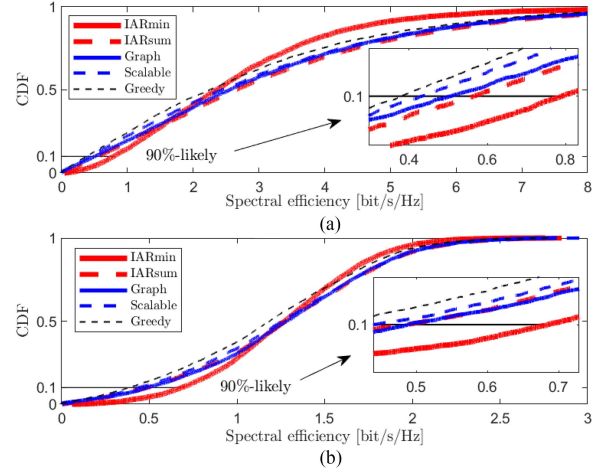


Fig. 3. Downlink SE per UE with different access schemes ($K = 50$). (a) LP-MMSE precoding. (b) MR precoding.

Algorithm 2. MUMA Assignment.

Input: $\{\text{iar}_{kl}^{t_k} : \forall k, l\}, \{a_{kl} : \forall k, l\}, \{b_{kl} = a_{kl} : \forall k, l\}, \tau_p$
Output: $\{a_{kl} : \forall k, l\}$

```

1 for AP  $l = 1, \dots, L$  do
2   Sort  $\{\text{iar}_{1l}^{t_1}, \dots, \text{iar}_{Kl}^{t_K}\}$  in descending order and
   include  $\tau_p$  UEs with the largest values in cluster  $\mathcal{C}_{ue}$ ;
3   Associate the UEs in  $\mathcal{C}_{ue}$  by letting  $a_{il} = 1, i \in \mathcal{C}_{ue}$ ;
4 if The goal is  $\max_{\{a_{kl}\}} \sum_{k,l} \text{iar}_{kl}^{t_k} a_{kl}$  then
5   Return;
6 else if The goal is  $\max_{\{a_{kl}\}} \min_k \Sigma_k$  then
7   while  $\sum_l b_{kl} < L, \forall k$  do
8     Compute the per-UE sum IARs  $\Sigma_k, \forall k$ ;
9     Find UE  $k' = \arg \min_k \Sigma_k$ ;
10    Find AP  $l' = \arg \max_{l, b_{kl'} \neq 1} \text{iar}_{kl'}^{t_{k'}}$  for UE  $k'$ ;
11    Find UE  $k^* = \arg \min_{k, b_{kl'} \neq 1} \text{iar}_{kl'}^{t_k}$  for AP  $l'$ ;
12     $b_{k'l'} \leftarrow 1$ ;
13    if  $\Sigma_{k^*} - \text{iar}_{k^*l'}^{t_{k^*}} \leq \Sigma_{k'}$  then
14       $b_{k^*l} \leftarrow 1$ , where  $l \in \{j : a_{k^*j} = 1\}$ ;
15      Continue;
16     $b_{k'l} \leftarrow 1$ , where  $l \in \{j : a_{k'j} = 1\}$ ;
17     $a_{k^*l'} \leftarrow 0, a_{k'l'} \leftarrow 1$ ;

```

random and each is equipped with a half-wavelength-spaced uniform linear array with $N = 4$ antennas. Two different numbers of UEs are considered: a) $K = 50$ corresponding to 200 UEs/km² and b) $K = 100$ corresponding to 400 UEs/km². The 3GPP Urban Microcell model is used to compute the large-scale propagation conditions, such as pathloss and shadow fading. Unless otherwise specified, other system parameters are referred to [5], [14], which are $\rho_p = 0.1$ W, $\rho_{dl} = 1$ W, $\sigma^2 = -94$ dBm, $\tau_p = 5$, $\tau_c = 200$, $\kappa = 10$, and $\mu = 1.8$.

Fig. 3 shows the cumulative distribution function (CDF) of the downlink SE per UE with $K = 50$, which gives a macro perspective comparison among the schemes. The most prominent observation is that IARmin significantly outperforms IARsum and other schemes on the user fairness (quantified by the 90%-likely SE value) with both precoders. There are two reasons. On the one hand, MUMA ensures

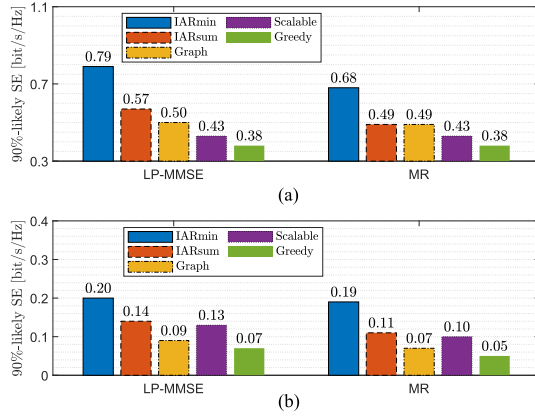


Fig. 4. 90%-likely SE with different numbers of UEs and precoders. (a) $K = 50$. (b) $K = 100$.

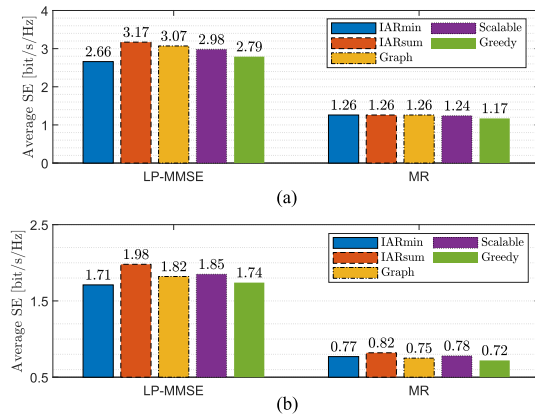


Fig. 5. Average SE with different numbers of UEs and precoders. (a) $K = 50$. (b) $K = 100$.

that each weak UE with poor channel condition accesses at least one AP and satisfies that each AP serves at most one UE per pilot. On the other hand, MUMA tries to associate the weak UEs with more serving APs which promotes the user fairness. Without the max-min association readjustment in MUMA (i.e., from line 6 to line 17 in Algorithm III-C), IARsum falls behind IARmin, but still provides a higher 90%-likely SE than the other considered schemes due to the accurate interference characterization of the IARs in (9) such that each UE is assigned the best pilot causing the least pilot contamination.

In Fig. 4, we elaborate the 90%-likely SE with different numbers of UEs and precoders. Fig. 4(a) quantifies the 90%-likely SE in Fig. 3, from which we can see that IARmin achieves 58% and 39% higher 90%-likely SE than the best benchmark (i.e., Graph) when using LP-MMSE precoding and MR precoding, respectively. Fig. 4(b) compares the considered schemes in a denser scenario with $K = 100$. When compared to Fig. 4(a), it is clear that the 90%-likely SE is deteriorated by the severe pilot contamination, while the advantage of IARmin grows. More precisely, we observe that IARmin achieves 67% and 90% higher 90%-likely SE than the best benchmark (i.e., Scalable) when using LP-MMSE precoding and MR precoding, respectively.

The average SE of the considered schemes is evaluated in Fig. 5. Although improving the average SE is not the main goal of our IAMA scheme, IARsum still outperforms the benchmarks in all considered cases. When comparing Fig. 5(a) and (b), we can see the advantage

of IARsum grows when the number of UEs increases, benefiting from the protection mechanism for the weak UEs of MUSA. For example, when using LP-MMSE precoding, the average SE of IARsum is slightly higher than the best benchmark Graph with $K = 50$ while the improvement compared to the best benchmark Scalable becomes 7% with $K = 100$. IARmin loses the average SE for improving the 90%-likely SE, but still slightly outperforms some benchmarks in the MR cases, where the weak UEs rely on accessing more serving APs for avoiding interference.

V. CONCLUSION

We proposed an interference-aware massive access scheme for scalable CF mMIMO in this paper. We proposed an iterative procedure for joint AP-UE association and pilot assignment by exploiting the proposed IARs, where two assignment algorithms are developed to maximize the user fairness or the average SE. The numerical results showed that our IAMA scheme significantly improved the user fairness compared to the state-of-the-art benchmark schemes, especially when using LP-MMSE precoding in a denser scenario.

REFERENCES

- [1] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.
- [2] Ö. T. Demir, E. Björnson, and L. Sanguinetti, "Foundations of user-centric cell-free massive MIMO," *Foundations Trends Signal Process.*, vol. 14, no. 3/4, pp. 162–472, 2021.
- [3] S. Chen, J. Zhang, E. Björnson, and B. Ai, "A survey on user-centric cell-free massive MIMO systems," *Digit. Commun. Netw.*, vol. 8, pp. 695–719, 2022.
- [4] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signal Process.*, vol. 11, no. 3/4, pp. 154–655, 2017.
- [5] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.
- [6] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.
- [7] H. Liu, J. Zhang, S. Jin, and B. Ai, "Graph coloring based pilot assignment for cell-free massive MIMO systems," *IEEE Trans. Veh. Tech.*, vol. 69, no. 8, pp. 9180–9184, Aug. 2020.
- [8] S. Buzzi, C. D'Andrea, M. Fresia, Y.-P. Zhang, and S. Feng, "Pilot assignment in cell-free massive MIMO based on the hungarian algorithm," *IEEE Wireless Commun. Lett.*, vol. 10, no. 1, pp. 34–37, Jan. 2021.
- [9] W. Zeng, Y. He, B. Li, and S. Wang, "Pilot assignment for cell-free massive MIMO systems using a weighted graphic framework," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6190–6194, Jun. 2021.
- [10] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai, "Structured massive access for scalable cell-free massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1086–1100, Apr. 2021.
- [11] C. Geng, N. Naderializadeh, A. S. Avestimehr, and S. A. Jafar, "On the optimality of treating interference as noise," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1753–1767, Apr. 2015.
- [12] C. Geng and S. A. Jafar, "On the optimality of treating interference as noise: Compound interference networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4630–4653, Aug. 2016.
- [13] X. Yi and G. Caire, "Optimality of treating interference as noise: A combinatorial perspective," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4654–4673, Aug. 2016.
- [14] S. Chen, J. Zhang, Z. Chen, and B. Ai, "Treating interference as noise in cell-free massive MIMO networks," in *Proc. IEEE Int. Conf. Commun.*, 2022, pp. 1385–1390.
- [15] M. Bacha, M. Di Renzo, and B. Clerckx, "Treating interference as noise in cellular networks: A stochastic geometry approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1918–1932, Mar. 2020.
- [16] D. Jungnickel, *Graphs, Networks and Algorithms*, vol. 3. Berlin, Germany: Springer, 2005.