IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 December 2023, accepted 22 January 2024, date of publication 1 February 2024, date of current version 8 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3360716

RESEARCH ARTICLE

RF Chain-Wise Clustering Schemes for Millimeter Wave Cell-Free Massive MIMO With Centralized Hybrid Beamforming

SHUNSUKE KAMIWATARI, ISSEI KANNO, TAKAHIRO HAYASHI, AND YOSHIAKI AMANO
KDDI Research Inc., Fujimino-shi, Saitama 356-8502, Japan

Corresponding author: Shunsuke Kamiwatari (sh-kamiwatari@kddi.com)

This work was supported by the National Institute of Information and Communications Technology (NICT), Japan, under Grant JPJ012368C00401.

ABSTRACT In this paper, we propose clustering schemes appropriate for a cell-free massive multi-input multi-output (CF mMIMO) system with centralized hybrid beamforming (BF) in the millimeter wave (mmWave) band, with the aim of reducing computational complexity and achieving scalability at the central processing unit (CPU) while ensuring system performance. Conventional access point (AP)-wise clustering schemes have been proposed for a decentralized architecture that mitigates inter-user interference at each AP through local digital BF and analog BF of each radio frequency (RF) chain to efficiently improve performance. However, in a centralized architecture, in which the digital BF weights is designed at the CPU while considering all RF chains of all APs together. In cases where each AP employs multiple RF chains forming different analog beams, the coupling loss (the sum of the path loss and BF gain of the analog beam) of each RF chain may differ even between the same AP and UE due to differences in the analog BF gain. AP-wise clustering is not well suited because RF chains with a high coupling loss (the sum of the path loss and BF gain of the analog beam) may be included among the APs of the cluster assigned to each UE. In the method proposed in this paper, clusters are formed on a per-RF-chain basis (RF chain-wise clustering). In this way, RF chains with lower coupling loss can be selected for each UE regardless of to which AP an RF chain belongs when forming a cluster for each UE. Moreover, the existing clustering schemes do not account for the amount of inter-cluster interference. Accordingly, we propose a cluster recombination scheme to effectively mitigate inter-cluster interference. Then, by combining RF chain-wise clustering and cluster recombination, a higher received signal power can be achieved while mitigating inter-cluster interference. Through simulation-based evaluations, we show that the proposed RF chain-wise clustering and cluster recombination schemes can achieve superior spectral efficiency while effectively reducing the complexity of centralized digital BF.

INDEX TERMS Cell-free massive MIMO, centralized architecture, hybrid beamforming, millimeter wave, clustering.

I. INTRODUCTION

Cell-free massive multiple-input multiple-output (CF mMIMO) technology is a candidate foundation for sixth generation (6G) networking [1], [2], [3]. In CF mMIMO, a large number of access points (APs) distributed in an area are employed by a central processing unit (CPU) to communicate with user equipment devices (UEs). Since the distribution of the APs and their cooperation, CF mMIMO can provide communications with uniform quality and high reliability regardless of the locations of the UEs [4], [5].

Recently, the millimeter wave (mmWave) band has been utilized in commercial mobile communication systems such as fifth generation (5G) new radio (NR) to enhance

The associate editor coordinating the review of this manuscript and approving it for publication was Walid Al-Hussaibi.

communication throughput and capacity through the availability of a wide spectrum [6], [7]. For 6G, an even wider spectrum is expected to be utilized, including the mmWave band [8]. However, the short wavelength of mmWave signals causes significant propagation and blockage losses [9]. On the other hand, the short wavelength also allows a base station to deploy a large number of antennas with a small antenna array [10]. Hence, in the mmWave system, a beamforming (BF) scheme is typically employed at the base station to address the above issues [11]. In the BF architecture, there are two types of BF schemes: analog BF at the antenna array and digital BF at the baseband unit. Although digital BF can provide higher BF gain, it increases the cost and power consumption of the base station because it requires radio frequency (RF) chains for the number of antenna elements. Therefore, a hybrid BF architecture, combining digital and analog BF, has been employed in mmWave communication systems, which typically have a large number of antenna elements at the base station, in order to reduce computational complexity while obtaining the effects of the hybrid BF [11], [12].

Systems employing such hybrid BF schemes for CF mMIMO in the mmWave band has been under investigated [13], [14], [15]. In CF mMIMO, by cooperating among distributed APs, the effect of the blockage loss can be reduced [16]. Typically, each AP has a large number of antenna elements connected to multiple RF chains. Each RF chain has an analog BF function that determines the analog beam for each RF chain. In most cases, each AP has a predefined analog BF codebook, and steers analog beams according to the codebook in order to find the analog beam that maximizes the signal power to the UEs [11], [17]. Considering that the number of available analog beams at the same time for an AP is limited by the number of RF chains, and it is typically assumed to be smaller than the number of spatially multiplexed UEs. Therefore, the CPU decides the assignment of analog beams to each RF chain based on the analog beam sweeps. According analog BF at each AP, the digital BF weights are calculated in order to maximize the signal-to-interference-plus-noise ratio (signal to interference-plus-noise ratio (SINR)) and reduce the interference between multiplexed UEs. Combining analog BF at the APs and digital BF, CF mMIMO with hybrid BF can enhance the signal levels of the UEs and mitigate the inter-UEinterference at the same time.

There are two approaches to splitting the responsibility for hybrid BF processing among the APs and the CPU in CF mMIMO. One is a decentralized architecture [15]. In decentralized architecture, each AP has a digital BF functionality and mitigates inter-UE interference [18]. Therefore, decentralized architecture can reduce the amount of information needed to be exchange through the fronthaul between the CPU and the APs. Moreover, implementing the digital BF function on APs can reduce the computational load on the CPU [15]. However, the interference mitigation effect of digital BF is limited to the number of RF chains in an AP. Therefore, in hybrid BF architecture. where the number of RF chains is smaller than the number of antenna elements, the performance in the decentralized architecture tends to decrease as the number of spatial multiplexed UEs increases [19].

The other approach is a centralized architecture. In the centralized architecture, the digital BF function is deployed at the CPU, and processed through RF chains in all APs connected to the CPU [17]. Digital BF at the CPU maximizes the SINR with other UEs by means of minimum mean square error (MMSE)-based digital BF weights. When digital BF is implemented at the CPU, more UEs can be spatially multiplexed simultaneously than is possible under a decentralized architecture because all RF chains of all APs can be used at the same time for interference suppression of the spatially multiplexed UEs. On the other hand, from the perspective of computational complexity at the CPU, an increase in the number of users increases the computational load imposed by digital BF on the CPU to a prohibitive degree. Also, the amount of information to be exchanged through fronthaul is larger than the decentralized architecture. This is because the CPU needs to estimate the channel between all APs and all UEs, also the CPU calculates the digital BF weights to mitigate all inter-UE interference according to the estimated channel. As a result, the CF mMIMO system lacks the scalability.

From the above discussion, there is a trade-off between decentralized and centralized architectures in terms of overall system computational load and system performance. When using hybrid BF in the mmWave band, the number of RF chains per AP is assumed to be less than the number of UEs which simultaneously connected [17], so the effects of digital BF can be degraded due to decentralized architecture. Therefore, to address the above issues, this paper focuses on the reduction of the computational complexity at the CPU with centralized architecture while ensuring the system performance to strike this trade-off.

In order to reduce the computational load at the CPU and maintain scalability in the decentralized architecture, AP-wise clustering scheme has been considered in which multiple APs are selected for each UE, thereby forming clusters, and signal processing is performed based on these clusters [18]. Each UE selects the subset of the APs and forms an AP cluster. The channel estimation and the digital BF weights calculation are processed by AP cluster basis, and it enables CF mMIMO systems to maintain scalability even with the increase of APs and UEs. In AP-wise clustering, all RF chains in the APs selected to the cluster are utilized to transmit signals for the UEs that forming the clusters. In most cases, each UE selects the APs with higher received signal power that are considered to be in relatively good communication condition for the cluster. On the other hand, when each AP employs hybrid BF and multiple RF chains, in analog BF, each RF chain in the AP utilizes one analog

beam for a certain UE. Therefore, with hybrid BF, even the coupling loss (the sum of the path loss and beamforming gain of the analog beam assigned to the RF chain) between the same AP and UE is significantly different for each RF chain. When hybrid BF is applied, AP-wise clustering may cause the signal power for a UE to degrade due to the inclusion of RF chains with high coupling losses in the cluster assigned to that UE.

In order to ensure more received signal power for each UEduring the clustering, we propose a RF chain-wise clustering scheme. RF chain-wise clustering allows the UEs to select the RF chains with higher received signal power and include them in their clusters, regardless of to which AP an RF chain belongs. This enables the UEs to form clusters with higher received signal power while achieving superior spectral efficiency. In our conference paper [20], we introduced RF chain-wise clustering and reported simulation results obtained with the proposed scheme. In RF chain-wise clustering, each UEcan effectively select RF chains with high received signal power, and at the CPU, partial MMSE (P-MMSE) [19]-based digital BF can be performed based on the RF chains with higher received signal power for each UE so as to sufficiently suppress interference. Notably, although the effectiveness of the existing clustering schemes has been shown for scenarios in which inter-cluster interference effects are relatively small, these schemes do not consider the amount of interference between clusters. In particular, the P-MMSE scheme mitigates inter-cluster interference only between UEs that share RF chains between their selected clusters. However, the amount of inter-cluster interference depends not only on the combination of RF chains selected for the cluster associated with each UE but also on the combinations of RF chains selected for the clusters associated with other UEs. Therefore, the spectral efficiency may degrade when the effects of inter-cluster interference are large.

To solve this problem, in this paper, we propose combining cluster recombination with RF chain-wise clustering. For cluster recombination, we consider that more inter-cluster interference is expected to be suppressed by digital BF with appropriate modifications to the RF chains belonging to the clusters for UEs with large interference effects. Through simulation-based evaluations, we simulate the impact of the clustering scheme on the spectral efficiency (SE) of the downlink (DL). In addition, we clarify scenarios in which RF chain-wise clustering and cluster recombination will be effective.

The contribution and findings of this paper are as follows:

- We propose a per-RF-chain clustering scheme for mm-wave CF mMIMO with centralized hybrid BF. Numerical results show that the proposed scheme can improve the received signal power by selecting RF chains with high received signal power for each cluster.
- We clarify the impact of inter-cluster interference on CF mMIMO in various simulation scenarios. In particular, in the Open Square scenario, in which radio waves tend to spread over a wide area, the SE degrades even with RF chain-wise clustering due to inter-cluster interference.
- We propose a cluster recombination scheme for refining the clustering results based on the received signal power. Through simulation-based evaluations, we show that a higher received signal power can be achieved while mitigating inter-cluster interference.

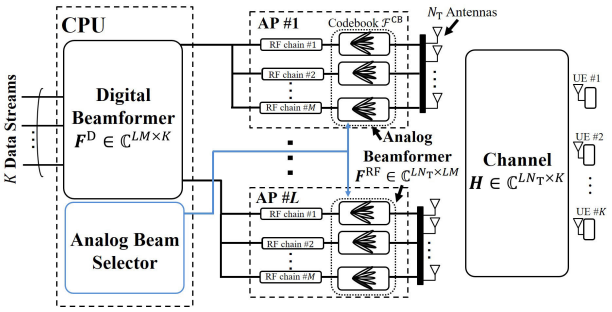


FIGURE 1. A mmWave CF mMIMO system with hybrid BF.

The following notations are used through this paper. $\boldsymbol{v}$ and $\boldsymbol{V}$ denote column vectors and matrices, respectively. In addition, $\boldsymbol{V}_{(i,k)}$, and $\boldsymbol{V}_{(i)}$ denotes (i,k)-th element, and (i)-th row vector of a matrix $\boldsymbol{V}$, respectively. $|\cdot|$, $\|\cdot\|_{\mathrm{F}}$, and $\|\cdot\|_0$ denote absolute values, Frobenius norm, and L-0 norm respectively. Furthermore, the transpose, complex conjugate transpose, and pseudo-inverse matrix are denoted by $\cdot^{\mathrm{T}}$, $\cdot^{\mathrm{H}}$, and $\cdot^{\dagger}$, respectively. We define the imaginary unit as j and the Napier number as e, and $\mathcal{CN}\left(\mu_{\mathrm{n}}, \sigma_{\mathrm{n}}^2\right)$ is a complex Gaussian distribution with mean μ_{n} and variance σ_{n}^2. $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix and $\mathrm{diag}[a_1, \ldots, a_N]$ denotes the $N \times N$ diagonal matrix with $a_1, \ldots, a_N$ diagonal components. The operation $\min(p, q)$ compares real numbers p and q and returns the smaller value. The operation $\arg(\boldsymbol{A})$ returns a matrix containing the phases of the entries of $\boldsymbol{A}$. The operation $\arg\max_{x \in \mathcal{A}} f(x)$ represents an operation that returns the argument x in the set $\mathcal{A}$ that maximizes $f(x)$, and the operation $\arg\min_{x \in \mathcal{A}} f(x)$ represents an operation that returns the argument x that minimizes $f(x)$ in the elements of the set $\mathcal{A}$.

The rest of the paper is organized as follows. The system model of CF mMIMO, the channel model in mmWave band and the procedures of hybrid BF are described in Section II. The conventional clustering scheme is described in Section III-A. The proposed clustering scheme is described in Section III-B. Furthermore, we show the evaluation results in Section IV. Finally, Section V concludes the paper.

II. SYSTEM MODEL

We assume a mmWave CF mMIMO system with centralized hybrid beamforming as shown in Fig. 1. In this paper, we consider a DL data communication. There is one CPU,

and the CPU employs L APs. Through the APs, the CPU transmit signals to K UEs. The CPU does the baseband signal processing including digital precoding, then the CPU sends the baseband signals to each RF chain of each AP via fronthaul. Each RF chain of the APs has an analog BF function and the CPU controls analog beams of each RF chain at the APs. Hence the system completely controls the hybrid BF with digital BF at the CPU and analog BF at APs.

Each AP has N_{T} antenna elements and M RF chains, and each RF chain is connected to N_{T} antenna elements. We assume that the number of RF chains per AP M is less than the number of UEs K, but the number of APs is greater than the number of UEs K, $LM > K > M$ [17], and each UE has an single antenna. The system bandwidth is set to B and divided into N_{RB} resource block (RB)s. Each RB consists of N_{SC} subcarriers. In this paper, we assume flat fading within each RB, and use the subcarrier located in the center of each RB as a representative. The DL received signal of UE k at the u-th RB is written as follows:

$$y_k[u] = \sum_{l=1}^{L} \boldsymbol{h}_{l,k}^{\mathrm{H}}[u]\boldsymbol{x}_l[u] + n_k[u], \tag{1}$$

where $\boldsymbol{x}_l[u] \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ is the signal transmitted from the l-th AP at the u-th RB and $n_k[u] \sim \mathcal{CN}(0, \sigma_k^2)$ denotes the Gaussian noise at the k-th UE at the u-th RB. $\boldsymbol{h}_{l,k}[u]$ is the propagation channel between the l-th AP and the k-th UE at the u-th RB.

A. CHANNEL MODEL

In this paper, we assume a wideband fading channel model in which the propagation model includes a line-of-sight (LoS) propagation path and non-line-of-sight (NLoS) propagation paths [21], [22], [23]. Additionally, we use a probability variable depending on the distance between each AP and UE to distinguish between the case in which the propagation channel includes the LoS propagation path and the case in which the propagation channel includes only NLoS propagation paths. The channel between the l-th AP and the k-th UE at the u-th RB, $\boldsymbol{h}_{l,k}[u]$ in (1), can be expressed as follows:

$$\boldsymbol{h}_{l,k}[u] = \sqrt{\frac{1}{N_{\mathrm{path}}}} \sum_{i=1}^{N_{\mathrm{path}}} \frac{\beta_{l,k,u}^{i}\delta_{l,k,u}}{\sqrt{\alpha_{l,k,u}^{\mathrm{NLoS}}}} \boldsymbol{a}_u(\theta_{l,k}^{i}) + \omega_{l,k}\boldsymbol{h}_{l,k}^{\mathrm{LoS}}[u], \tag{2}$$

where $\omega_{l,k} \in \{0, 1\}$ is the probability variable which follows a Bernoulli distribution, and $\omega_{l,k}$ takes 1 with probability $P_{l,k}^{\mathrm{LoS}}$. $P_{l,k}^{\mathrm{LoS}}$ is the probability of the existence of the LoS propagation path between the l-th AP and the k-th UE. $P_{l,k}^{\mathrm{LoS}}$ is based on the distance between the l-th AP and the k-th UE, which is denoted by $d_{l,k}$, and is given by the following equation:

$$P_{l,k}^{\mathrm{LoS}} = \min\left(\frac{20}{d_{l,k}}, 1\right)\left(1 - e^{-\frac{d_{l,k}}{39}}\right) + e^{-\frac{d_{l,k}}{39}}. \tag{3}$$

In other words, we assume the Rician fading channel model between the l-th AP and the k-th UE in the case of $\omega_{l,k} = 1$ or the Rayleigh fading channel in the case of $\omega_{l,k} = 0$. Additionally, N_{path} is the number of the propagation path. $\beta_{l,k,u}^{i} \sim \mathcal{CN}(0, 1)$ and $\theta_{l,k,u}^{i}$ are random channel gain and angle of departure of the i-th path, respectively. $\delta_{l,k,u} = e^{-j2\pi\tau_{l,k}^{i}f_{\mathrm{c}}^{u}}$ denotes the coefficient for delay, in which $\tau_{l,k}^{i}$ and f_{c}^{u} are the delay of the i-th path and the the carrier frequency of the u-th RB, respectively. $\alpha_{l,k,u}^{\mathrm{NLoS}}$ is the path loss component, and $\alpha_{l,k,u}^{\mathrm{NLoS}}$ is given by

$$\alpha_{l,k,u}^{\mathrm{NLoS}} = 20\log_{10}\frac{4\pi f_{\mathrm{c}}^{u}}{c} + 10\gamma^{\mathrm{NLoS}}\log_{10}(d_{l,k}) + X_{\sigma}^{\mathrm{NLoS}}[\mathrm{dB}], \tag{4}$$

where γ_{NLoS} and $X_{\sigma}^{\mathrm{NLoS}}$ are the path loss exponent of NLoS path and the shadow fading term of NLoS path, respectively. $\boldsymbol{a}_u(\theta_{l,k}^{i})$ is the array response vector for each AP. We assume that APs are each equipped with one uniform linear array (ULA) and that $\boldsymbol{a}_u(\theta_{l,k}^{i})$ is given by

$$\boldsymbol{a}_u(\theta_{l,k}^{i}) = \left[1, e^{j2\pi\left(\frac{f_{\mathrm{c}}^{u}}{c}d_{\mathrm{ant}}\right)\theta_{l,k}^{i}}, \cdots, e^{j(N_{\mathrm{T}}-1)2\pi\left(\frac{f_{\mathrm{c}}^{u}}{c}d_{\mathrm{ant}}\right)\theta_{l,k}^{i}}\right], \tag{5}$$

where d_{ant} denotes the distance between antenna elements. $\boldsymbol{h}_{l,k}^{\mathrm{LoS}}[u]$ is the LoS propagation path vector and can be written as

$$\boldsymbol{h}_{l,k}^{\mathrm{LoS}}[u] = \frac{\beta_{l,k,u}^{\mathrm{LoS}}}{\sqrt{\alpha_{l,k,u}^{\mathrm{LoS}}}}\boldsymbol{a}_u(\theta_{l,k}^{\mathrm{LoS}}), \tag{6}$$

$\beta_{l,k,u}^{\mathrm{LoS}} \sim \mathcal{CN}(0, 1)$, $\alpha_{l,k,u}^{\mathrm{LoS}}$ and $\theta_{l,k,u}^{\mathrm{LoS}}$ denote the random gain, the path loss component and the AoD of the LoS propagation path between the l-th AP and the k-th UE at the u-th RB. Similar to (4), $\alpha_{l,k,u}^{\mathrm{LoS}}$ is given by follows:

$$\alpha_{l,k,u}^{\mathrm{LoS}} = 20\log_{10}\frac{4\pi f_{\mathrm{c}}^{u}}{c} + 10\gamma^{\mathrm{LoS}}\log_{10}(d_{l,k}) + X_{\sigma}^{\mathrm{LoS}}[\mathrm{dB}], \tag{7}$$

where γ^{LoS} and $X_{\sigma}^{\mathrm{LoS}}$ denote the path loss exponent and the shadow fading term for the LoS propagation path, respectively.

To clarify the scenarios in which the proposed schemes would be effective, in our evaluation, we consider the Street Canyon and Open Square scenarios and use different path loss exponents γ^{NLoS} and γ^{LoS} and shadow fading terms $X_{\sigma}^{\mathrm{NLoS}}$ and $X_{\sigma}^{\mathrm{LoS}}$ in (4) and (7).

B. HYBRID BF

In this section, we formulate a hybrid BF scheme for the transmitted signal $\boldsymbol{x}_l$ based on (1). First, the digital baseband signals are combined through digital BF at the CPU. In this paper, we assume that the CPU processes digital BF for each RB. The baseband signal of the u-th RB at the l-th AP after digital BF, $\boldsymbol{x}_l^{\mathrm{D}}[u] \in \mathbb{C}^{M\times 1}$, is given by

$$\boldsymbol{x}_l^{\mathrm{D}}[u] = \sum_{k=1}^{K} \boldsymbol{F}_{l,k}^{\mathrm{D}}[u]s_k[u], \tag{8}$$

where s_k is the transmitted symbol at the u-th RB for the k-th UE and $\boldsymbol{F}^{\mathrm{D}}_{l,k} \in \mathbb{C}^{M\times 1}$ denotes the digital BF vector from the l-th AP to the k-th UE at the u-th RB. Then, the CPU sends these baseband signals to the APs. At the AP, the baseband signals $\boldsymbol{x}^{\mathrm{D}}_l$ are combined with analog BF by RF chains. We assume that the analog BF vector of the m-th RF chain at the l-th AP $\boldsymbol{f}^{\mathrm{RF}}_{l,m} \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ is selected from the predefined codebook set $\mathcal{F}^{\mathrm{CB}} = \left[\boldsymbol{f}^{\mathrm{CB}}_1, \ldots, \boldsymbol{f}^{\mathrm{CB}}_{N_{\mathrm{bm}}}\right]$, and the same analog beams are applied to all RBs. After analog BF, the transmitted signal from the l-th AP $\boldsymbol{x}_l$ can be written by:

$$\boldsymbol{x}_l[u] = \boldsymbol{F}^{\mathrm{RF}}_l \boldsymbol{x}^{\mathrm{D}}_l[u], \tag{9}$$

where $\boldsymbol{F}^{\mathrm{RF}}_l = [\boldsymbol{f}^{\mathrm{RF}}_{l,1}, \cdots, \boldsymbol{f}^{\mathrm{RF}}_{l,M}] \in \mathbb{C}^{N_{\mathrm{T}}\times M}$ denotes the matrix of analog BF in the l-th AP.

From (1), (8) and (9), it is found that the received signal at the k-th UE, $y_k[u]$, is given by

$$\begin{aligned} y_k[u] = &\sum_{l=1}^{L} \boldsymbol{h}^{\mathrm{H}}_{l,k}[u] \boldsymbol{F}^{\mathrm{RF}}_l \boldsymbol{F}^{\mathrm{D}}_{l,k}[u] s_k[u] \\ &+ \sum_{l=1}^{L} \boldsymbol{h}^{\mathrm{H}}_{l,k}[u] \boldsymbol{F}^{\mathrm{RF}}_l \sum_{j\neq k}^{K} \boldsymbol{F}^{\mathrm{D}}_{l,j}[u] s_j[u] + n_k[u]. \end{aligned} \tag{10}$$

The SINR of the k-th UE, denoted by $\mathrm{SINR}_k[u]$, and its achievable SE, $\mathrm{SE}_k[u]$, are given by

$$\mathrm{SINR}_k[u] = \frac{P_{\mathrm{tx}} \left| \sum_{l=1}^{L} \boldsymbol{h}^{\mathrm{H}}_{l,k}[u] \boldsymbol{F}^{\mathrm{RF}}_l \boldsymbol{F}^{\mathrm{D}}_{l,k}[u] \right|^2}{\sum_{j\neq k}^{K} P_{\mathrm{tx}} \left| \sum_{l=1}^{L} \boldsymbol{h}^{\mathrm{H}}_{l,k}[u] \boldsymbol{F}^{\mathrm{RF}}_l \boldsymbol{F}^{\mathrm{D}}_{l,j}[u] \right|^2 + \sigma^2_k}, \tag{11a}$$

$$\mathrm{SE}_k[u] = \log_2 \left(1 + \mathrm{SINR}_k[u]\right), \tag{11b}$$

where P_{tx} denotes the AP transmit power, which we assume to be the same for all APs.

C. ANALOG BEAM SELECTION

As mentioned above, the CPU selects the analog beams for each RF chain from a predefined codebook. To select effective analog beams that can provide higher BF gain to the UEs, CPU and APs decide analog beams with the following procedure presented in [17].

First, each AP transmits pilot signals with N_{bm} analog beams in accordance with the predefined codebook $\mathcal{F}^{\mathrm{CB}}$. In this paper, we assume that each AP uses the same codebook $\mathcal{F}^{\mathrm{CB}}$, and that the pilot signal is transmitted over a number of RBs equivalent to the primary synchronization signal in 5G NR, using contiguous RBs in the system bandwidth [24]. The pilot signal is transmitted over the set of RBs $\mathcal{U}^{\mathrm{pilot}} = [u_1, \ldots, u_{N_{\mathrm{pilot}}}]$, where N_{pilot} is the number of RBs used for pilot signal. RBs of $\mathcal{U}^{\mathrm{pilot}}$ are located contiguous N_{pilot} RBs in the center of the system bandwidth. Each UE measures the received signal power $p_{l,k,i}$ for each pilot signal sent by an analog beam from the APs, where $p_{l,k,i}$ is given as follows:

$$p_{l,k,i} = \sum_{u\in\mathcal{U}^{\mathrm{pilot}}} \left| \boldsymbol{h}^{\mathrm{H}}_{l,k}[u] \boldsymbol{f}^{\mathrm{RF}}_i \right|^2. \tag{12}$$

Then, each UE determines the analog beam $\tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k}$ with the highest received signal power for each AP, where $\tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k}$ is given as follows:

$$\tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k} = \underset{\boldsymbol{f}^{\mathrm{RF}}_i \in \mathcal{F}^{\mathrm{CB}}}{\arg\max}\, p_{l,k,i}. \tag{13}$$

After measurement, each UE reports the received signal power and the beam index of $\tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k}$ to the CPU through the APs.

The number of analog beams that an AP can use simultaneously is limited by M, the number of RF chains in an AP. We assume $M < K$, so the AP cannot use all the analog beams with the highest analog BF gain to the UEs $\tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k}$ calculated in (13) simultaneously. Therefore, the CPU needs to determine the combination of analog beams of each AP. In this paper, the CPU associates M UEs with each AP, and each AP allocates analog beams to its RF chains directed to the M UEs allocated to the AP. The CPU decide association between APs and UEs by the following algorithm, which was presented in [17].

By using feedback from the UEs, the CPU calculates the total received signal power ξ_k from analog beam $\tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k}$ of all APs for each UE based on (13). ξ_k is given by

$$\xi_k = \sum_{l=1}^{L} \sum_{u\in\mathcal{U}^{\mathrm{pilot}}} \left| \boldsymbol{h}^{\mathrm{H}}_{l,k}[u] \tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k} \right|^2. \tag{14}$$

The CPU identifies the AP $\hat{l}_k$ with the lowest received signal power for UE k. $\hat{l}_k$ is calculated as follows:

$$\hat{l}_k = \underset{l=1,L}{\arg\min} \sum_{u\in\mathcal{U}^{\mathrm{pilot}}} \left| \boldsymbol{h}^{\mathrm{H}}_{l,k} \tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k} \right|^2, \forall k. \tag{15}$$

Because the effect of analog BF is considered to be small, the association between AP $\hat{l}_k$ and the k-th UE is deleted. The CPU starts this procedure from the UE with the highest ξ_k, i.e., the UE whose communication conditions are considered the best. By starting this procedure with UEs with large ξ_k, the CPU can assign the APs with large BF effects to the UEs where the communication conditions are considered to be relatively worse. The CPU repeats this procedure until the number of UEs associated with the AP, N_{UE}, satisfies the constraint $N_{\mathrm{UE}} \leq M$ for all APs. After this algorithm is run, each AP is associated with M different UEs, and the APs select analog beams for these M UEs, that is, AP l uses $\tilde{\boldsymbol{f}}^{\mathrm{RF}}_{l,k}$ to send a signal. Through this procedure, the CPU can select the analog beams to obtain a high analog BF effect while considering UE fairness.

D. DIGITAL BF AT THE CPU

Once the analog beams for all RF chains have been determined, the CPU processes digital BF. As aforementioned, digital BF is processed for each RB. The CPU estimates the effective channel matrix of each RB including the analog

beam patterns, $\boldsymbol{H}_{\mathrm{CPU}}[u]\in\mathbb{C}^{LM\times K}$. $\boldsymbol{H}_{\mathrm{CPU}}[u]$ is given by

$$\boldsymbol{H}_{\mathrm{CPU}} = [\boldsymbol{H}_1[u], \cdots, \boldsymbol{H}_K[u]], \tag{16a}$$

$$\boldsymbol{H}_k[u] = \left(\boldsymbol{h}_{1,k}^{\mathrm{H}}[u]\boldsymbol{f}_{1,1}^{\mathrm{RF}}, \cdots, \boldsymbol{h}_{L,k}^{\mathrm{H}}[u]\boldsymbol{f}_{L,M}^{\mathrm{RF}}\right)^{\mathrm{T}}. \tag{16b}$$

The CPU calculates the weights for digital BF by using $\boldsymbol{H}_{\mathrm{CPU}}[u]$ in (16). Suppose that the MMSE-based scheme is used for digital BF; then, the weights are expressed by the following equation:

$$\boldsymbol{W}_{\mathrm{MMSE}}[u]=\left(\boldsymbol{H}_{\mathrm{CPU}}^{\mathrm{H}}[u]\boldsymbol{H}_{\mathrm{CPU}}[u] + \sigma^2\boldsymbol{I}_{LM}\right)^{\dagger}\boldsymbol{H}_{\mathrm{CPU}}^{\mathrm{H}}[u]. \tag{17}$$

In the DL transmission, the CPU multiplies the digital BF weight $\boldsymbol{W}_{\mathrm{MMSE}}[u]$ and the transmit data symbol, then sends to the APs via fronthaul. At the APs, each RF chain is assigned analog beam by the aforementioned algorithm, and APs send the transmit signals to the UEs. The computational load at calculation of the digital BF weights in the MMSE is given as follow [19]:

$$C_{\mathrm{MMSE}} = \sum_{k=1}^{K}\left(\frac{(ML)^2+ML}{2}K + (ML)^2 + \frac{(ML)^3-ML}{3}\right). \tag{18}$$

As seen from (18), the computational load at calculation of the digital BF weights under the MMSE scheme increases as the number of APs L and the number of UEs K increase significantly.

III. P-MMSE-BASED DIGITAL BF AND CLUSTERING

In the system described above, all APs are employed for transmitting signals to each UE and the CPU computes the MMSE-based digital beamforming weights based on the effective channel between all APs and UEs. However, as the number of APs (L) and UEs (K) increases, the computational load for channel estimation and weight calculation at the CPU also increases, and this method is not scalable in practice. Moreover, the APs are distributed and use analog beams to the specific direction. If an AP is situated far away from a UE or if the analog beam of an AP is not directed towards the UE, the AP's impact on the UE's performance gain becomes negligible. To effectively reduce the complexity, a P-MMSE-based digital BF scheme, which is applied in conjunction with AP-wise clustering, is considered to achieve scalability [19]. In AP-wise clustering, each UE selects subset of APs with the highest received signal power to the AP cluster. The CPU mitigates inter-UE interference through APs selected for the cluster by using P-MMSE scheme.

In the following subsections, we first describe a scheme in which the P-MMSE approach is applied in combination with the conventional AP clustering scheme. Then, we propose a clustering scheme that is suitable for the centralized architecture.

A. AP-WISE CLUSTERING

The UEs measure the channel power between each AP and themselves in order to determine subset of the APs for inclusion in their own clusters. As in the analog beam selection, we assume that channel power is measured from the pilot signal transmitted over the set of RBs $\mathcal{U}^{\mathrm{pilot}}$. The channel power between AP l and UE k is given by

$$P_{l,k}^{\mathrm{AP}} = \sum_{m=1}^{M}\sum_{u\in\mathcal{U}^{\mathrm{pilot}}}\left|\boldsymbol{h}_{l,k}^{\mathrm{H}}[u]\boldsymbol{f}_{l,m}^{\mathrm{RF}}\right|^2, \tag{19}$$

where $\boldsymbol{f}_{l,m}$ is the analog BF vector of the m-th RF chain of the l-th AP. In this paper, each UE selects N_{cl} APs with high channel power for its own cluster. Then The UEs send the index of APs which selected for their cluster. With AP-wise clustering, the received signal of the k-th UE expressed in (10) becomes

$$y_k[u] = \sum_{l=1}^{L}\boldsymbol{h}_{l,k}^{\mathrm{H}}[u]\boldsymbol{F}_l^{\mathrm{RF}}\sum_{j=1}^{K}\boldsymbol{D}_{j,l}^{\mathrm{AP}}\boldsymbol{F}_{l,j}^{\mathrm{D}}[u]s_j[u] + n_k, \tag{20}$$

where $\boldsymbol{D}_{k,l}^{\mathrm{AP}} \in \mathbb{C}^{M\times}$ is a diagonal matrix that can be written as

$$\boldsymbol{D}_{k,l}^{\mathrm{AP}} = \mathrm{diag}(d_{k,l}^{\mathrm{AP}}, \ldots, d_{k,l}^{\mathrm{AP}}). \tag{21}$$

Here, $d_{k,l}^{\mathrm{AP}} = 1$ if AP l is included in the AP cluster of UE k, and $d_{k,l}^{\mathrm{AP}} = 0$ otherwise. As seen from (21), the AP-wise clustering involves the decision of whether or not to include all of the RF chains of an AP in a cluster. The P-MMSE-based weight vector for the k-th UE at the u-th RB $\bar{\boldsymbol{w}}_k^{\mathrm{AP}}[u]$ can be written as the follows:

$$\bar{\boldsymbol{w}}_k^{\mathrm{AP}}[u] = \frac{\boldsymbol{v}_k^{\mathrm{AP}}[u]}{\sqrt{\boldsymbol{v}_k^{\mathrm{AP}\mathrm{H}}[u]\boldsymbol{D}_k^{\mathrm{AP}}\boldsymbol{v}_k^{\mathrm{AP}}[u]}}, \tag{22}$$

$$\boldsymbol{v}_k^{\mathrm{AP}}[u] = p_k^{\mathrm{AP}}\left(\sum_{i\in\mathcal{P}_k^{\mathrm{AP}}}p_i^{\mathrm{AP}}\boldsymbol{D}_k^{\mathrm{AP}}\boldsymbol{H}_k[u]\boldsymbol{H}_k^{\mathrm{H}}[u]\boldsymbol{D}_k^{\mathrm{AP}} + \boldsymbol{Z}_k^{\mathrm{AP}}\right)^{\dagger} \times \boldsymbol{D}_k^{\mathrm{AP}}\boldsymbol{H}_k[u], \tag{23}$$

$$\text{with } \boldsymbol{Z}_k^{\mathrm{AP}} = \boldsymbol{D}_k^{\mathrm{AP}}\left(\sigma_k^2\boldsymbol{I}_{LM}\right)\boldsymbol{D}_k^{\mathrm{AP}},$$

where $\boldsymbol{D}_k^{\mathrm{AP}} \in \mathbb{C}^{LM\times LM}$ is the clustering indicator of the k-th UE and $\boldsymbol{D}_k^{\mathrm{AP}} = \mathrm{diag}(\boldsymbol{D}_{k,1}^{\mathrm{AP}}, \cdots, \boldsymbol{D}_{k,L}^{\mathrm{AP}})$. $\mathcal{P}_k^{\mathrm{AP}}$ denotes the set of UEs that share a selected AP with the k-th UE in their clusters and p_k^{AP} is the transmit power for UE k. From (22) and (23), the SINR of UE k is given as follows [19]:

$$\mathrm{SINR}_k^{\mathrm{AP}}[u] = \frac{p_k^{\mathrm{AP}}\left|\boldsymbol{H}_k^{\mathrm{H}}[u]\boldsymbol{D}_k^{\mathrm{AP}}\bar{\boldsymbol{w}}_k^{\mathrm{AP}}[u]\right|^2}{\sum_{i\neq k}^{K}p_i^{\mathrm{AP}}\left|\boldsymbol{H}_k^{\mathrm{H}}[u]\boldsymbol{D}_i^{\mathrm{AP}}\bar{\boldsymbol{w}}_i^{\mathrm{AP}}[u]\right|^2 + \sigma_k^2}. \tag{24}$$

Equations (22) - (24) show that AP-wise clustering with P-MMSE suppresses inter-user interference between UEs that select the same AP for their cluster. This is because interference between UEs that select the same AP for the cluster is expected to have a significant impact.

The computational complexity of calculating the digital BF weights under the P-MMSE scheme for all UEs is given as follows [19]:

$$C_{\mathrm{P-MMSE}}^{\mathrm{AP}} = \sum_{k=1}^{K}\Bigg(\frac{(M|\mathcal{M}_k^{\mathrm{AP}}|)^2 + M|\mathcal{M}_k^{\mathrm{AP}}|}{2}|\mathcal{P}_k^{\mathrm{AP}}| + (M|\mathcal{M}_k^{\mathrm{AP}}|)^2 + \frac{(M|\mathcal{M}_k^{\mathrm{AP}}|)^3 - M|\mathcal{M}_k^{\mathrm{AP}}|}{3}\Bigg), \tag{25}$$

where $|\mathcal{P}_k^{\mathrm{AP}}| \leq K$ denotes the number of elements in the set $\mathcal{P}_k^{\mathrm{AP}}$ and $\mathcal{M}_k^{\mathrm{AP}}$ denotes the set of APs transmitting signals and suppressing interference with UE k, where $|\mathcal{M}_k^{\mathrm{AP}}| \leq L$. As seen from (25) and (18), P-MMSE-based BF requires only channel information between the UEs and the APs selected by the UEs for the cluster. Therefore, the P-MMSE-based BF and AP clustering can effectively reduce the computational complexity of digital BF and the amount of information to be exchanged through fronthaul while ensuring effective interference mitigation, which is crucial for improving system performance.

B. RF CHAIN-WISE CLUSTERING

In the algorithm described above, each UE measures the channel power for each AP and forms its own AP-wise cluster. However, when a CF mMIMO employs hybrid BF, especially when the CPU assigns analog beams to each of its multiple RF chains independently, the channel power of each RF chain may differ significantly even in the same APs. This is because each RF chain in the AP may be assigned analog beams with different directions. Therefore, with AP-wise clustering, RF chains with low channel power for the UE may be included in the cluster, resulting in system performance degradation. Therefore, in RF chain-wise clustering, power measurement and cluster formation are performed not for each AP but for each RF chain. With the proposed scheme, the UEs can use the RF chains with the highest received power among the APs for cluster formation regardless of the APs to which they belong.

1) RF CHAIN-WISE CLUSTER SELECTION

In the proposed scheme, the channel power measurement is processed in each analog beam units (i.e., RF chain units), and the effective channel power including the analog BF gain between the m-th RF chain of AP l and UE k is given by

$$P_{k,l,m}^{\mathrm{RF}} = \sum_{u\in\mathcal{U}^{\mathrm{pilot}}}\left|\boldsymbol{h}_{l,k}^{\mathrm{H}}[u]\boldsymbol{f}_{l,m}^{\mathrm{RF}}\right|^2. \tag{26}$$

Regarding the channel power measurement, UEs select RF chains to the cluster. In the AP-wise clustering, each UE selects N_{cl} APs to be included in the cluster, and each AP has M RF chains. In other words, a cluster contains $N_{\mathrm{cl}}M$ RF chains. In this paper, to unify the number of RF chains in a cluster, the UEs select $N_{\mathrm{cl}}M$ RF chains for the cluster with highest channel power $P_{k,m,l}^{\mathrm{RF}}$. The indicator matrix of clustering between the l-th AP and the k-th UE $\boldsymbol{D}_{k,l}^{\mathrm{RF}} \in \mathbb{C}^{M\times M}$ can be written as follows:

$$\boldsymbol{D}_{k,l}^{\mathrm{RF}} = \mathrm{diag}(d_{k,l,1}^{\mathrm{AP}}, \cdots, d_{k,l,m}^{\mathrm{AP}}, \cdots, d_{k,l,M}^{\mathrm{AP}}), \tag{27}$$

where $d_{k,l,m}^{\mathrm{RF}} = 1$ if the m-th RF chain of the l-th AP is included in the cluster of the k-th UE, and $d_{k,l,m}^{\mathrm{RF}} = 0$ otherwise. Comparing (21) and (27), in the AP-wise clustering, the diagonal components of $\boldsymbol{D}_{k,l}^{\mathrm{AP}}$ have the same value $d_{k,l}^{\mathrm{AP}}$, while in the RF chain-wise clustering, the diagonal components of $\boldsymbol{D}_{k,l}^{\mathrm{RF}}$ can have different value. From (22) and (23), the P-MMSE weight vector $\bar{\boldsymbol{w}}_k^{\mathrm{RF}}$ can be written as follows:

$$\bar{\boldsymbol{w}}_k^{\mathrm{RF}}[u] = \frac{\boldsymbol{v}_k^{\mathrm{RF}}[u]}{\sqrt{\boldsymbol{v}_k^{\mathrm{RF}\mathrm{H}}[u]\boldsymbol{D}_k^{\mathrm{RF}}\boldsymbol{v}_k^{\mathrm{RF}}[u]}}, \tag{28}$$

$$\boldsymbol{v}_k^{\mathrm{RF}}[u] = p_k^{\mathrm{RF}}\left(\sum_{i\in\mathcal{P}_k^{\mathrm{RF}}} p_i^{\mathrm{RF}}\boldsymbol{D}_k^{\mathrm{RF}}\boldsymbol{H}_k[u]\boldsymbol{H}_k^{\mathrm{H}}[u]\boldsymbol{D}_k^{\mathrm{RF}} + \boldsymbol{Z}_k^{\mathrm{RF}}\right)^{\dagger} \times \boldsymbol{D}_k^{\mathrm{RF}}\boldsymbol{H}_k[u],$$
$$\text{with } \boldsymbol{Z}_k^{\mathrm{RF}} = \boldsymbol{D}_k^{\mathrm{RF}}\left(\sigma_k^2\boldsymbol{I}_{LM}\right)\boldsymbol{D}_k^{\mathrm{RF}}, \tag{29}$$

where $\boldsymbol{D}_k^{\mathrm{RF}} \in \mathbb{C}^{LM\times LM}$ is the clustering indicator of the k-th UE and $\boldsymbol{D}_k^{\mathrm{RF}} = \mathrm{diag}(\boldsymbol{D}_{k,1}^{\mathrm{RF}}, \cdots, \boldsymbol{D}_{k,L}^{\mathrm{RF}})$. $\mathcal{P}_k^{\mathrm{RF}}$ denotes the set of UEs that share a selected RF chain with the k-th UE in their clusters and p_k^{RF} is the transmit power for UE k. Similar to (24), (28) and (29), the SINR of the k-th UE can be written as follows:

$$\mathrm{SINR}_k^{\mathrm{RF}}[u] = \frac{p_k^{\mathrm{RF}}\left|\boldsymbol{H}_k^{\mathrm{H}}[u]\boldsymbol{D}_k^{\mathrm{RF}}\bar{\boldsymbol{w}}_k^{\mathrm{RF}}[u]\right|^2}{\sum_{i\neq k}^{K} p_i^{\mathrm{RF}}\left|\boldsymbol{H}_k^{\mathrm{H}}[u]\boldsymbol{D}_i^{\mathrm{RF}}\bar{\boldsymbol{w}}_i^{\mathrm{RF}}[u]\right|^2 + \sigma_k^2}. \tag{30}$$

From (25), the computational complexity of calculating the digital BF weights under the P-MMSE scheme for all UEs is given as follows [19]:

$$C_{\mathrm{P-MMSE}}^{\mathrm{RF}} = \sum_{k=1}^{K}\Bigg(\frac{|\mathcal{M}_k^{\mathrm{RF}}|^2 + |\mathcal{M}_k^{\mathrm{RF}}|}{2}|\mathcal{P}_k^{\mathrm{RF}}| + |\mathcal{M}_k^{\mathrm{RF}}|^2 + \frac{|\mathcal{M}_k^{\mathrm{RF}}|^3 - |\mathcal{M}_k^{\mathrm{RF}}|}{3}\Bigg), \tag{31}$$

where $|\mathcal{P}_k^{\mathrm{RF}}| \leq K$ denotes the number of elements in the set $\mathcal{P}_k^{\mathrm{RF}}$ and $\mathcal{M}_k^{\mathrm{RF}}$ denotes the set of RF chains transmitting signals to and suppressing interference with the k-th UE, where$|\mathcal{M}_k^{\mathrm{RF}}| \leq ML$. Similar to AP-wise clustering, Each UE informs the CPU of the index of the RF chains selected for the cluster, and the CPU identifies the combination of AP and UE that requires channel information. RF chain-wise clustering in combination with P-MMSE-based digital BF can effectively reduce the computational complexity of digital BF and the information to be exchanged on fronthaul.

2) CLUSTER RECOMBINATION

The clustering schemes described above select the APs/RF chains with the highest received signal power to increase the

received signal power for each UE. However, these schemes do not consider the amount of interference between clusters. The P-MMSE scheme mitigates inter-cluster interference between UEs that select the same AP for inclusion in their clusters. However, the amount of inter-cluster interference depends not only on the combination of APs selected by each UE for its cluster but also on the combinations of APs selected by other UEs for their clusters. If clustering is based solely on the received signal power, then the system capacity might degrade in environments with strong inter-cluster interference. Therefore, we propose cluster recombination to effectively suppress inter-cluster interference while maintaining a high received signal power.

In the proposed scheme, the UEs reselect APs for their clusters to refine the clustering results based on the received signal power. As seen from (28), (29) and (31), the larger the number of other UEs $|\mathcal{P}_k^{\mathrm{RF}}|$ that share a selected AP as UE k for their clusters, the more inter-cluster interference is mitigated by the P-MMSE scheme, while the computational load imposed on the CPU for digital BF increases. Therefore, the proposed method recombines clusters only for UEs whose $|\mathcal{P}_k^{\mathrm{RF}}|$ values are small, i.e., UEs that are considered to be strongly affected by inter-cluster interference. The steps of the proposed scheme are as follows.

First, the CPU determines the UE l_k with the lowest $|\mathcal{P}_{L_k}^{\mathrm{RF}}|$. Then, the CPU finds the RF chain i_{l_k} with the largest received signal power among the RF chains that UE l_k has not selected for its cluster. Next, the CPU identifies the RF chain $\hat{i_{l_k}}$ with the lowest received signal power among the RF chains selected by UE l_k for its cluster and not selected by the other UEs. UE l_k adds RF chain i_{l_k} to its cluster and removes RF chain $\hat{i_{l_k}}$ from its cluster. Finally, the CPU calculates $|\mathcal{P}^{\mathrm{RF}_k}|$ for all k. The CPU repeats these steps until $|\mathcal{P}_k^{\mathrm{RF}}| \geq P_{\min}$ for all UEs. The computational complexity of digital BF, $C_{\mathrm{P-MMSE}}^{\mathrm{ReComb}}$, can be calculated as shown in (31) using $|\mathcal{P}_k^{\mathrm{ReComb}}|$ and $|\mathcal{M}_k^{\mathrm{ReComb}}|$, which are the newly obtained values of $|\mathcal{P}_k^{\mathrm{RF}}|$ and $|\mathcal{M}_k^{\mathrm{RF}}|$ from the clusters after recombination. $C_{\mathrm{P-MMSE}}^{\mathrm{ReComb}}$ can be written as follows [19]:

$$C_{\mathrm{P-MMSE}}^{\mathrm{ReComb}} = \sum_{k=1}^{K}\Bigg(\frac{|\mathcal{M}_k^{\mathrm{ReComb}}|^2 + |\mathcal{M}_k^{\mathrm{ReComb}}|}{2}|\mathcal{P}_k^{\mathrm{ReComb}}| + |\mathcal{M}_k^{\mathrm{ReComb}}|^2 + \frac{|\mathcal{M}_k^{\mathrm{ReComb}}|^3 - |\mathcal{M}_k^{\mathrm{ReComb}}|}{3}\Bigg). \tag{32}$$

While cluster recombination can effectively mitigate inter-cluster interference, the computational complexity of clustering and digital BF becomes high. Additionally, cluster recombination requires information on the $P_{l_k,i,m}$ values between all RF chains and UEs. Therefore, for cluster recombination, the UEs should send feedback on $P_{l_k,i,m}$ for all RF chains.

IV. NUMERICAL EVALUATION

In this section, we present numerical evaluation results to clarify the effectiveness of the proposed RF chain-wise clustering and cluster recombination schemes through comparisons with the AP-wise clustering scheme. The parameters in the simulation are shown in Table 1 [24]. In AP-wise clustering, each UE selects N_{cl} APs with higher received signal power to include in its cluster. To unify the number of RF chains per cluster, in RF chain-wise clustering, each UE selects $N_{\mathrm{cl}}M$ RF chains with higher received signal power for its cluster. We evaluate and compare the DL SEs of each UE under the proposed scheme (RF chain-wise clustering and cluster recombination), AP-wise clustering and optimal CF mMIMO without clustering in II-D as a benchmark.

To clarify the scenarios in which the proposed schemes would be effective, in this evaluation, we consider the Street Canyon and Open Square scenarios and use different path loss exponents γ_{NLoS} and γ_{LoS} and shadow fading terms $X_{\sigma_{\mathrm{NLoS}}}$ and $X_{\sigma_{\mathrm{LoS}}}$ in (4) and (7). The values of γ_{NLoS}, γ_{LoS}, $X_{\sigma_{\mathrm{NLoS}}}$ and $X_{\sigma_{\mathrm{LoS}}}$ in each scenario are given in Table 2 [21].

Fig. 2 shows the cumulative distribution of the average DL SE in all resource blocks for each user with $M = 2$ in the street canyon scenario. As seen from Fig. 2, clustering causes the SE to degrade depending on the cluster size. During the clustering, the CPU considers and mitigates intra-cluster interference but does not take into account inter-cluster interference. It can be seen that when the cluster size is small, the effect of inter-cluster interference is greater.

Algorithm 1 Cluster Recombination

while $\forall|\mathcal{P}_k^{\mathrm{RF}}| < P_{\min}$ **do**
  $l_k = \arg\min_{l=1,\cdots,K} |\mathcal{P}_l^{\mathrm{RF}}|$
  $i_{l_k}, m_{l_k} = \arg\max_{i,m\notin\mathcal{M}_{l_k}} P_{l_k,i,m}$
  $\hat{i_{l_k}}, \hat{m_{l_k}} = \arg\min_{i,m\in\mathcal{M}_{l_k}\cap i,m\notin\mathcal{M}_k,\forall k\neq l_k} P_{l_k,i,m}$
  add i_{l_k} and m_{l_k} to the cluster
  remove $\hat{i_{l_k}}$ and $\hat{m_{l_k}}$ from the cluster
end while

TABLE 1. Simulation parameters.

Meaning	Character	Value
Evaluation area	-	1 km × 1 km
Number of APs	L	256
Number of UEs	K	20
Transmit power of each AP	-	43 dBm
Noise power	-	-174 dBm/Hz
Bandwidth	-	400 MHz
Number of RBs in the bandwidth	N_{RB}	264
Number of RBs for the pilot signal	N_{pilot}	10
Number of analog beams per AP	N_{bm}	36
Number of antennas per AP	N_{T}	16
Number of RF chains per AP	M	2,4
Number of APs per cluster	N_{cl}	4, 16

TABLE 2. Propagation model parameters.

Scenario	Path loss γ	Shadow fading X_σ
Street Canyon, NLoS	3.19	8.2
Street Canyon, LoS	1.98	3.1
Open Square, NLoS	2.89	7.1
Open Square, LoS	1.85	4.2

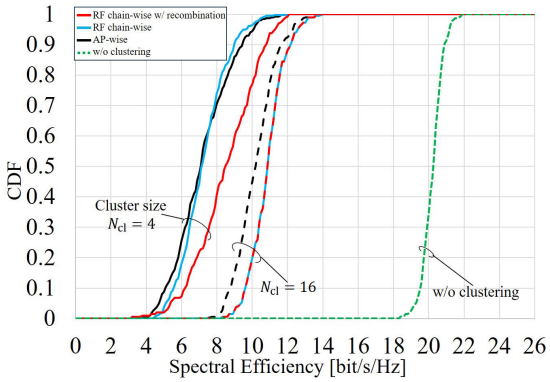


FIGURE 2. Downlink SEs per UE under different clustering schemes; Street Canyon, $K = 20$, $M = 2$.

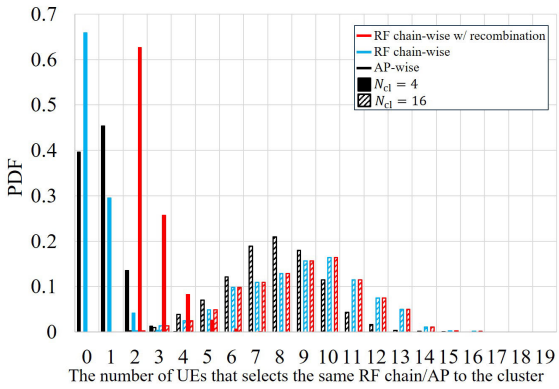


FIGURE 3. $|\mathcal{P}_k^{\mathrm{AP}}|$ and $|\mathcal{P}_k^{\mathrm{RF}}|$, the numbers of UEs that select the same APs for their clusters under different clustering schemes; Street Canyon, $K = 20$, $M = 2$.

A comparison of the clustering schemes shows that especially with $N_{\mathrm{cl}} = 16$, RF chain-wise clustering improves the SE compared to AP-wise clustering. With RF chain-wise clustering, clusters can be formed of RF chains with high signal power for the corresponding UEs, resulting in an improved SE.

Although the proposed scheme can improve SE in the case of $N_{\mathrm{cl}} = 16$, the SE becomes worse with RF chain-wise clustering for $N_{\mathrm{cl}} = 4$. As the cluster size N_{cl} decreases, the impact of inter-cluster interference becomes larger. This is because there is a decrease not only in the number of APs in each cluster but also in the number of other UEs that select the same APs for their clusters, i.e., $|\mathcal{P}_k^{\mathrm{AP}}|$ and $|\mathcal{P}_k^{\mathrm{RF}}|$. Fig. 3 shows the probability density functions of $|\mathcal{P}_k^{\mathrm{AP}}|$ and $|\mathcal{P}_k^{\mathrm{RF}}|$. The results indicate that in the case of $N_{\mathrm{cl}} = 4$, $|\mathcal{P}_k^{\mathrm{RF}}|$ in RF chain-wise clustering tends to be small. In AP-wise clustering, each AP uses several analog beams. Therefore, an AP tends to be selected by various UEs that are located in multiple directions relative to the AP. On the other hand, in RF chain-wise clustering, each RF chain has one analog beam. Therefore, an RF chain tends to be selected by UEs that are in a particular direction relative to the AP. As a result, $|\mathcal{P}_k^{\mathrm{RF}}|$ in RF chain-wise clustering becomes small compared to $|\mathcal{P}_k^{\mathrm{AP}}|$ in AP-wise clustering.

As seen from Fig. 2, by combining RF chain-wise clustering and cluster recombination, the SE can be improved compared to AP-wise clustering. When $N_{\mathrm{cl}} = 4$, $|\mathcal{P}_k^{\mathrm{AP}}|$ and $|\mathcal{P}_k^{\mathrm{RF}}|$ (the numbers of UEs that select the same APs for their clusters) decrease, and the inter-cluster interference tends to be stronger, especially in RF chain-wise clustering. In cluster recombination, the CPU recombines the clusters to increase $|\mathcal{P}_k^{\mathrm{RF}}|$ for UEs with smaller $|\mathcal{P}_k^{\mathrm{RF}}|$ values to mitigate inter-cluster interference, and this process is repeated until $|\mathcal{P}_k^{\mathrm{RF}}|$ is larger than P_{min} for all UEs. Fig. 2 indicates that in the case of $N_{\mathrm{cl}} = 4$, $|\mathcal{P}_k^{\mathrm{RF}}|$ increases with cluster recombination, resulting in an increase in the SE. In the case of $N_{\mathrm{cl}} = 16$, $|\mathcal{P}_k^{\mathrm{RF}}|$ is already sufficiently large for each UE when RF chain-wise clustering is applied. Therefore, the inter-cluster interference can be suppressed even when cluster recombination is not applied. When the proposed scheme is used, the SE at the 5-th percentile level increases by 14.04% with $N_{\mathrm{cl}} = 4$ and by 10.99% with $N_{\mathrm{cl}} = 16$.

Figs. 4 and 5 show the SE and $|\mathcal{P}_k^{\mathrm{RF}}|$, respectively, in the Open Square scenario. Similar to the Street Canyon scenario, RF chain-wise clustering can improve the SE when $N_{\mathrm{cl}} = 16$, and the difference in SE between the clustering schemes becomes small when $N_{\mathrm{cl}} = 4$ due to inter-cluster interference. A comparison between the evaluation environments shows that in Open Square, the SE is degraded when clustering is applied in the CF mMIMO system. In the Open Square scenario, the effects of path loss and shadowing are smaller than in the Street Canyon scenario. When clustering is applied, inter-cluster interference occurs. In the Open Square case, the interference from distant APs is not strongly affected by path loss and shadowing, resulting in an increased interference power and a degraded SE. On the other hand, in the case of CF mMIMO without clustering, the SE becomes larger in an Open Square scenario compared to a Street Canyon scenario. In the case of CF mMIMO without clustering, all inter-user interference is mitigated by digital BF. Therefore, in an Open Square scenario where the effects of path loss and shadowing are small and the desired signal power is large, higher SE can be achieved.

To clarify the impact of the number of RF chains per AP on the proposed schemes, we compare the DL SE per UE in a case in which each AP has more RF chains. Fig. 6 shows the DL SE with $M = 4$ in the Street Canyon scenario. Similar to the case with $M = 2$, the proposed schemes improve the DL SE. Compared to Fig. 2, the difference in the DL SE between AP-wise clustering and RF chain-wise clustering is larger. Especially in the case of $N_{\mathrm{cl}} = 4$, RF chain-wise clustering can yield a superior DL SE compared to AP-wise clustering, even without cluster recombination. Under the assumption of a constant AP transmission power, the transmission power per analog beam tends to decrease as the number of RF chains per AP (M) increases. This is because, with an increasing M, the power of each chain required for analog BF is reduced. In addition, the directions

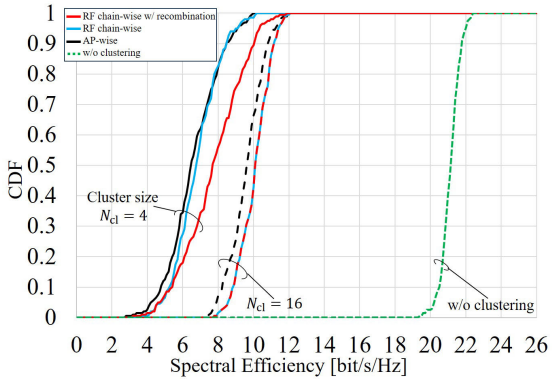


FIGURE 4. Downlink SEs per UE under different clustering schemes; Open Square, $K = 20$, $M = 2$.

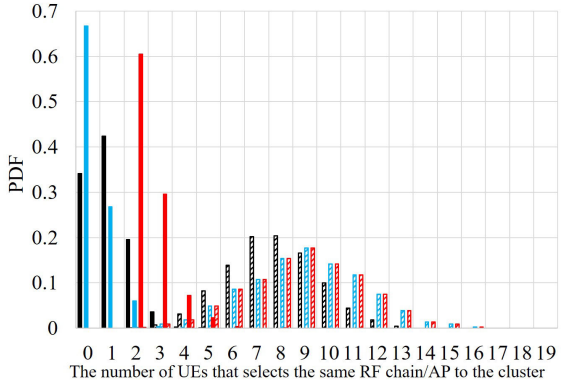


FIGURE 5. $|\mathcal{P}_k^{\mathrm{AP}}|$ and $|\mathcal{P}_k^{\mathrm{RF}}|$ under different clustering schemes; Open Square, $K = 20$, $M = 2$.

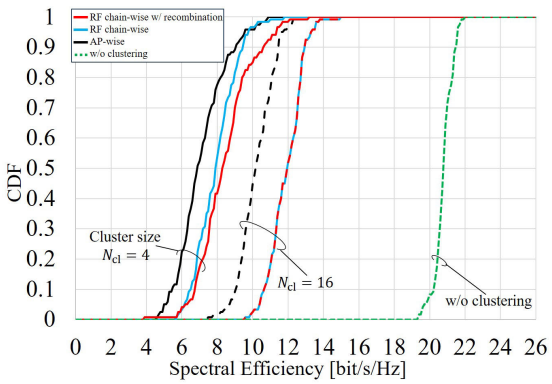


FIGURE 6. DL SEs per UE under different clustering schemes; Street Canyon, $K = 20$, $M = 4$.

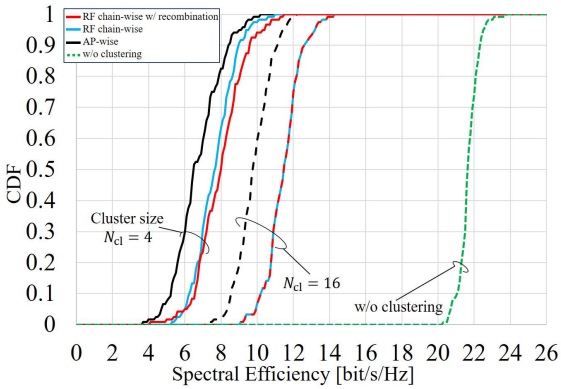


FIGURE 7. DL SEs per UE under different clustering schemes; Open Square, $K = 20$, $M = 4$.

of the analog beams are independent between RF chains at an AP. Therefore, in AP-wise clustering, the received signal power of the UEs may decrease. On the other hand, RF chain-wise clustering can select RF chains for a cluster regardless of the APs to which they belong. Additionally, increasing the total number of RF chains is expected to increase the number of RF chains for which analog beams with high analog BF gains for the UEs are selected. Therefore, even if M increases, In RF chain-wise clustering, a UE can select RF chains with high received signal power for its cluster.

Finally, we compare the computational complexity under these clustering schemes. Table 3 shows the indicators of computational complexity for each clustering scheme, C_{MMSE}, $C_{\mathrm{P-MMSE}}^{\mathrm{AP}}$, $C_{\mathrm{P-MMSE}}^{\mathrm{RF}}$ and $C_{\mathrm{P-MMSE}}^{\mathrm{ReComb}}$, defined in (18), (25), (31) and (32). Table 3 shows that all clustering schemes can reduce the computational complexity compared with the case without clustering. As seen by comparing the clustering schemes, cluster recombination causes the computational complexity to increase as more intercluster interference is mitigated by digital BF at the CPU. In the case of $N_{\mathrm{cl}} = 16$, however, the computational complexity of RF chain-wise clustering alone is almost the same as that of RF chain-wise clustering with cluster recombination. This is because when $N_{\mathrm{cl}} = 16$, intercluster interference can be mitigated even without cluster recombination because each UE selects a relatively large number of RF chains for its cluster. Therefore, the clusters after cluster recombination show little difference from the clusters before recombination.

TABLE 3. Indicators of computational complexity: C_{MMSE}, $C_{\mathrm{P-MMSE}}^{\mathrm{AP}}$, $C_{\mathrm{P-MMSE}}^{\mathrm{RF}}$ and $C_{\mathrm{P-MMSE}}^{\mathrm{ReComb}}$.

M	$M = 2$		$M = 4$	
N_{cl}	16	4	16	4
RF chain-wise (10^6)	130.5	0.02494	2124	0.3253
RF chain-wise w/ recombination (10^6)	130.5	1.402	2124	3.212
AP-wise (10^6)	75.75	0.01276	826.2	0.1541
w/o clustering (10^6)	952.6		7389	

V. CONCLUSION

We have proposed an RF chain-wise clustering scheme for a mm-wave CF mMIMO system with centralized hybrid BF in which each AP has multiple RF chains with an analog BF functionality. In RF chain-wise clustering, a cluster consisting of the RF chains with the highest received signal power is formed for each UE, regardless of the APs to which the RF chains belong. Additionally, to effectively

mitigate inter-cluster interference, we have proposed a cluster recombination scheme to be applied in combination with RF chain-wise clustering. With this proposed scheme, a higher received signal power can be achieved while mitigating inter-cluster interference.

Through simulation-based evaluations, we have shown that hybrid BF with the proposed RF chain-wise clustering scheme can achieve a superior SE while effectively reducing the complexity of centralized digital BF. On the other hand, in the case of a small cluster size and the Open Square scenario, in which radio waves tend to spread over a wide area, the SE degrades even with RF chain-wise clustering due to inter-cluster interference. However, by combining RF chain-wise clustering and cluster recombination, the DL SE can be improved even in such scenarios. In particular, RF chain-wise clustering is effective in scenarios with a large number of RF chains per AP, and cluster recombination is effective for clustering with a small number of RF chains per cluster.

Moreover, the proposed scheme requires only the received signal power of the signals transmitted from each RF chain and does not depend on the design method of the hybrid BF. Therefore, the proposed scheme has the potential to achieve high-quality communication while reducing the complexity of CF mMIMO systems that use low-complexity hybrid BF design techniques.

ACKNOWLEDGMENT

An earlier version of this paper was presented at the 2022 IEEE Global Communications Conference [DOI: 10.1109/GLOBECOM48099.2022.10001553].

REFERENCES

[1] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[2] 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.

[3] Next G Alliance. (2022). *Next G Alliance Report: 6G Technologies*. [Online]. Available: https://www.nextgalliance.org/wp-content/uploads/dlm_uploads/2022/07/TWG-report-6G-technologies.pdf

[4] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive MIMO: A survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 492–523, 1st Quart., 2022.

[5] 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, Jan. 2017.

[6] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.

[7] Hexa-X. (2021). *Deliverable D2.2 Initial Radio Models and Analysis Towards Ultra-high Data Rate Links in 6G*. [Online]. Available: https://hexa-x.eu/wp-content/uploads/2022/01/Hexa-X-D2_2.pdf

[8] J.-C. Guo, Q.-Y. Yu, W.-B. Sun, and W.-X. Meng, "Robust efficient hybrid pre-coding scheme for mmWave cell-free and user-centric massive MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 8006–8022, Dec. 2021.

[9] I. K. Jain, R. Kumar, and S. S. Panwar, "The impact of mobile blockers on millimeter wave cellular systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 854–868, Apr. 2019.

[10] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[11] C. M. Yetis, E. Bjornson, and P. Giselsson, "Joint analog beam selection and digital beamforming in millimeter wave cell-free massive MIMO systems," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1647–1662, 2021.

[12] Y. Chen, D. Chen, Y. Tian, and T. Jiang, "Spatial lobes division-based low complexity hybrid precoding and diversity combining for mmWave IoT systems," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3228–3239, Apr. 2019.

[13] G. Femenias and F. Riera-Palou, "Cell-free millimeter-wave massive MIMO systems with limited fronthaul capacity," *IEEE Access*, vol. 7, pp. 44596–44612, 2019.

[14] A. A. Avidh, Y. Sambo, S. Ansari, and M. A. Imran, "Hybrid beamforming with fixed phase shifters for uplink cell-free millimetre-wave massive MIMO systems," in *Proc. Joint Eur. Conf. Netw. Commun. 6G Summit (EuCNC/6G Summit)*. Porto, Portugal: IEEE, Jun. 2021, pp. 19–24.

[15] J. Kassam, D. Castanheira, A. Silva, R. Dinis, and A. Gameiro, "Distributed hybrid equalization for cooperative millimeter-wave cell-free massive MIMO," *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5300–5316, Aug. 2022.

[16] M. Alonzo, S. Buzzi, A. Zappone, and C. Delia, "Energy-efficient power control in cell-free and user-centric massive MIMO at millimeter wave," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 3, pp. 651–663, Sep. 2019.

[17] Z. Wang, R. Liu, H. Li, M. Li, and Q. Liu, "Hybrid beamforming design for C-RAN based mmWave cell-free systems," in *Proc. IEEE 92nd Veh. Technol. Conf. (VTC-Fall)*. Victoria, BC, Canada: IEEE, Nov. 2020, pp. 1–5.

[18] M. Alonzo and S. Buzzi, "Cell-free and user-centric massive MIMO at millimeter wave frequencies," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–5.

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

[20] S. Kamiwatari, I. Kanno, T. Ohseki, K. Yamazaki, and Y. Kishi, "RF chain-wise clustering for centralized mm-wave cell-free massive MIMO with hybrid beamforming," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2022, pp. 764–769.

[21] K. Haneda et al., "5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments," in *Proc. IEEE 83rd Veh. Technol. Conf.*, May 2016, pp. 1–7.

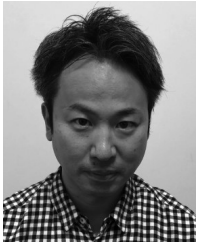
[22] Y. Chen, D. Chen, T. Jiang, and L. Hanzo, "Channel-covariance and angle-of-departure aided hybrid precoding for wideband multiuser millimeter wave MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8315–8328, Sep. 2019.

[23] Y. Chen, Y. Xiong, D. Chen, T. Jiang, S. X. Ng, and L. Hanzo, "Hybrid precoding for WideBand millimeter wave MIMO systems in the face of beam squint," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1847–1860, Mar. 2021.

[24] *Physical Channels and Modulation*, document TS 38.211, Version 18.0.0, 3rd Gener. Partnership Project (3GPP), Sep. 2023.

SHUNSUKE KAMIWATARI received the B.E. and M.E. degrees in communication engineering from Kyushu University, Fukuoka, Japan, in 2018 and 2020, respectively. Then, he joined KDDI Corporation, in 2020, where he has been engaged in research on signal processing and wireless communication systems. Since 2021, he has been engaged in research on wireless communication systems, including 5G and 6G technology with KDDI Research Inc. His current research interests include signal processing and beamforming schemes for wireless communication systems, especially in cell-free massive MIMO.

ISSEI KANNO received the Ph.D. degree from the Tokyo Institute of Technology, Tokyo, Japan, in 2008. Then, he joined KDDI Corporation, where he has been engaged in research on software-defined radio, antennas, and propagation in mobile communication systems. From 2013 to 2015, he was engaged in research on cognitive radio with the Advanced Telecommunication Research Institute International (ATR). Since 2015, he has been engaged in research on wireless communication systems, including 5G and 6G technology with KDDI Research Inc. His current research interests include signal processing and resource utilization for wireless communication systems. He received the Best Paper Award from IEEE at IEEE WCNC 2010, the Young Researchers' Award from IEICE, in 2011, and the Distinguished Service Award from IEICE, in 2012 and 2020.



TAKAHIRO HAYASHI received the B.E. and M.E. degrees in information and communication engineering from Yokohama National University, Japan, in 2002 and 2004, respectively. In 2004, he joined KDDI Corporation, where he became engaged in telecommunication network planning and optimization. Since 2010, he has been engaged in research and development on mobile communication systems with KDDI Research Inc. He is currently involved in the development of new frequency bands and radio propagation prediction using machine learning. He received the Young Researchers' Award from IEICE, in 2011.

YOSHIAKI AMANO received the B.E. and M.E. degrees in electrical and electronic engineering from Nagoya University, Aichi, Japan, in 1995 and 1999, respectively. He joined KDD Corporation (currently KDDI Corporation), in 1999, where he became involved in research and development on CDMA cellular systems, smart antennas, and wireless performance evaluation and improvement of mobile terminals. He also served as a member of the Energy Business Planning Division, KDDI Corporation, where he helped to launch the company's electric power retail business, from 2015 to 2018. In 2018, he launched research and development on dynamic spectrum sharing systems and intelligent reflecting surfaces with KDDI Research Inc., before leading research on wireless communication systems with beyond 5G/6G technology. He received the Young Researcher's Award from IEICE and the Meritorious Award on Radio from the Association of Radio Industries and Businesses (ARIB), in 2005 and 2011, respectively.