Hybrid Beamforming and Adaptive RF Chain Activation for Uplink Cell-Free Millimeter-Wave Massive MIMO Systems

Nhan Thanh Nguyen, *Member, IEEE*, Kyungchun Lee, *Senior Member, IEEE*, and Huaiyu Dai, *Fellow, IEEE*

***Abstract*—In this work, we investigate hybrid analog–digital beamforming (HBF) architectures for uplink cell-free (CF) millimeter-wave (mmWave) massive multiple-input multiple-output (MIMO) systems. We first propose two HBF schemes, namely, decentralized HBF (D-HBF) and semi-centralized HBF (SC-HBF). In the former, both the digital and analog beamformers are generated independently at each AP based on the local channel state information (CSI). In contrast, in the latter, only the digital beamformer is obtained locally at the access point (AP), whereas the analog beamforming matrix is generated at the central processing unit (CPU) based on the global CSI received from all APs. We show that the analog beamformers generated in these two HBF schemes provide approximately the same achievable rates despite the lower complexity of D-HBF and its lack of CSI requirement. Furthermore, to reduce the power consumption, we propose a novel adaptive radio frequency (RF) chain-activation (ARFA) scheme, which dynamically activates/deactivates RF chains and their connected analog-to-digital converters (ADCs) and phase shifters (PSs) at the APs based on the CSI. For the activation of RF chains, low-complexity algorithms are proposed, which can achieve significant improvement in energy efficiency (EE) with only a marginal loss in the total achievable rate.**

***Index Terms*—Cell-free massive MIMO, mmWave communication, hybrid beamforming, RF chain activation.**

I. Introduction

Recently, many attempts have been made to utilize millimeter waves (mmWaves) for high-data-rate mobile broadband communications. The main challenge of mmWave communication is the large path loss due to high carrier frequencies, which significantly limits the system performance and cell coverage [1]–[3]. Fortunately, the short wavelength of mmWave systems facilitates the deployment of massive multiple-input multiple-output (MIMO) systems, which can provide large beamforming gains to compensate for the path loss in mmWave channels. Furthermore, Ngo *et al.* in [4], [5] introduce a cell-free (CF) massive MIMO architecture, in which a very large number of distributed access points (APs) connected to a single central processing unit (CPU) simultaneously serve a much smaller number of users over the same time/frequency resources. Therefore, the CF massive MIMO system is capable of providing good quality of service (QoS) uniformly to all served users, regardless of their locations in the coverage area, by using simple linear signal processing schemes [4]–[6]. From these aspects, the combination of mmWave and CF massive MIMO systems with the deployment of large numbers of antennas at the APs could be a symbiotic convergence of technologies that can significantly improve the performance of next-generation wireless communication systems [6].

A. Related works

The performance CF massive MIMO systems in conventional sub-6-GHz frequency bands have been analyzed intensively [4], [5], [7]–[21]. In particular, the closed-form expression of the achievable rate of CF massive MIMO systems is derived in [4], [5], [7], [8]. Furthermore, comparisons of CF massive MIMO and conventional small-cell MIMO systems in [4], [5], [8]–[10] show that the former is more robust to shadow fading correlation and significantly outperforms the latter in terms of throughput and coverage probability. In [11], a low-complexity power control technique with zero-forcing (ZF) precoding design is introduced for CF massive MIMO systems. In [12], an optimal power-allocation algorithm is proposed to maximize the total EE, which can double the total EE compared to the equal power control scheme. Furthermore, an AP selection (APS) scheme is proposed in [12], in which each user chooses and connects to only a subset of APs to reduce the power consumption caused by the backhaul links. Meanwhile, in [13], the EE maximization problem is considered under the effect of quantization distortion of the weighted received signals at the APs. In [14], [15], the authors propose AP-clustering approaches to overcome limitations upon the scalability of CF massive MIMO systems. Specifically, the APs are grouped in clusters and multiple CPUs are used to manage these clusters, leading to a reduction in the data distribution and computational complexity involved in channel estimation, power control, and beamforming.

Another line of work has attempted to investigate the performance of CF massive MIMO systems in mmWave

This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1A6A1A03032119), by Basic Science Research Program through the NRF, funded by the Ministry of Education (NRF-2016R1D1A1B03933122), by the Academy of Finland, 6G Flagship program under Grant 346208, EERA Project under grant 332362, and Infotech Program funded by University of Oulu Graduate School. (Corresponding author: Kyungchun Lee.)

N. T. Nguyen was with Seoul National University of Science and Technology, Seoul, Republic of Korea. He is currently with Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland (e-mail: nhan.nguyen@oulu.fi).

K. Lee is with the Department of Electrical and Information Engineering and the Research Center for Electrical and Information Technology, Seoul National University of Science and Technology, 232 Gongneung-ro, Nowon-gu, Seoul, 01811, Republic of Korea (e-mail: kclee@seoultech.ac.kr).

H. Dai is with the Department of Electrical and Computer Engineering, North Carolina State University, NC, USA. (e-mail: Huaiyu_Dai@ncsu.edu).

channels [6], [22]–[24]. In particular, Femenias et al. introduced a hybrid beamforming (HBF) framework for CF mmWave massive MIMO systems with limited fronthaul capacity in [6], and the eigen-beamforming scheme is applied to generate precoders/combiners. Specifically, the phases of analog beamformers are obtained by quantizing those of dominant eigenvectors of the channel covariance matrix known at the APs. In [22], a hybrid precoding algorithm leveraging antenna-array response vectors is applied to distributed MIMO systems with partially-connected HBF architecture. Although the partially-connected HBF structure has lower power consumption than the fully-connected one, it cannot fully exploit the beamforming gains [25]. In [23] and [24], Alonzo et al. introduce an uplink multi-user estimation scheme along with low-complexity HBF architectures. Specifically, the baseband and analog precoders at each AP are generated by decomposing the fully digital ZF precoding matrix using the block-coordinate descent algorithm. Moreover, the problems of pilot assignment and channel estimation are considered in [26] and [27], respectively.

In mmWave communications, a large number of antennas at the APs not only provide beamforming gains to compensate for the propagation loss in mmWave channels, but also enhance favorable propagation [28]. In particular, it is shown in [28] that given a fixed total number of antennas at the APs, employing more antennas at fewer APs is more beneficial than deploying more APs with fewer antennas in terms of favorable propagation and channel hardening. Furthermore, large antenna arrays at the APs can achieve high beamforming gains to overcome the severe path loss in mmWave communication. However, an excessively high power consumption is required in this deployment because signal processing in the conventional digital domain requires a dedicated RF chain and analog-to-digital converter (ADC) for each antenna. Therefore, energy-efficient HBF schemes dedicated to CF mmWave massive MIMO systems are required, but only limited works exist in the literature focusing on the optimization of HBF for CF mmWave massive MIMO systems. Specifically, in [6], [22]–[24], the analog beamformers are all separately generated at the APs based on the local CSI, which are similar to HBF schemes for small-cell mmWave massive MIMO systems when each AP in the CF mmWave massive MIMO system is considered as a base station in the small-cell system. The antenna-selection (AS) schemes, which are proposed for the transceiver with a limited number of RF chains in [29], [30], can be applied to CF mmWave massive MIMO systems to reduce power consumption. However, they can cause performance degradation, especially for HBF in the highly correlated channels of mmWave communication [31]. Conversely, several studies [32]–[36] show that low and/or variable resolution ADCs/DACs can offer a good tradeoff between the rate and power consumption in conventional cellular (mmWave) massive MIMO systems. Motivated by this, Xiong *et al.* [37] propose the use of variable resolution ADCs in CF massive MIMO systems. Through analytical and numerical analysis, Xiong *et al.* show that more quantization bits should be allocated to the AP with larger aggregated large-scale fading and lower channel correlation [37]. Moreover, in CF mmWave massive MIMO systems, partially connected hybrid beamforming architectures [25], [38]–[40] with a reduced number of phase shifters in the analog beamformer can be leveraged to significantly reduce the power consumption.

B. Contributions

In this work, we investigate the HBF for uplink CF mmWave massive MIMO systems in two scenarios: the global CSI of all APs is available or unavailable at the CPU. Then, we propose an adaptive RF chain-activation (ARFA) scheme, which provides considerable power reduction while nearly maintaining the system's total achievable rate; thus, the EE is remarkably improved. Our specific contributions can be summarized as follows:

- We first propose the decentralized HBF (D-HBF) and semi-centralized HBF (SC-HBF) schemes. Both have digital beamformers generated at the AP, but their difference lies in the analog beamformer. Specifically, in SC-HBF, the analog beamforming matrices for all APs are generated at the CPU based on the global CSI. In contrast, that of the D-HBF is obtained at each AP based only on the local CSI. By exploiting the global CSI to jointly optimize the analog combiners at the CPU, SC-HBF is expected to outperform D-HBF. However, our analytical and numerical results show that D-HBF can perform approximately the same as SC-HBF while requiring substantially lower computational complexity and no global CSI.
- In CF mmWave massive MIMO systems with L APs and N RF chains at each AP, the power consumption is approximately proportional to LN. Because L is large in CF massive MIMO systems, the total power consumption can be excessively high. To overcome this challenge, we propose an ARFA scheme. In this scheme, the RF chains are selectively activated at the APs based on partial CSI, and the number of active RF chains at the APs is optimized so that the proposed scheme can significantly reduce the total power consumption while causing only marginal performance loss. Our numerical analysis reveals that in CF mmWave massive MIMO systems, the proposed ARFA scheme with a relatively small number of active RF chains can exhibit performance comparable to that of the conventional fixed-activation HBF scheme, which activates all the available RF chains. As a result, a considerable improvement in the EE is achieved.
- In the proposed ARFA scheme, high computational complexity is required to find the optimal numbers of active RF chains at numerous APs with an exhaustive search. To reduce complexity, we propose a low-complexity near-optimal algorithms for the ARFA with SC-HBF. Furthermore, ARFA is incorporated with D-HBF in the proposed D-ARFA schemes, creating a singular value-based and path loss-based D-ARFA, wherein the CPU requires a very limited amount of information from the APs. Our simulation results show that the proposed algorithms perform very close to the conventional HBF scheme, in which all the available RF chains are turned on.

We note that an RF chain-selection (RFS) scheme is introduced in [41]–[43] for the conventional cellular mmWave massive MIMO systems. The RFS scheme and our proposed ARFA scheme are similar in exploiting a reduced number of RF chains for power reduction. However, they have the following differences. First, in [41], the point-to-point transmission in a conventional cellular mmWave massive MIMO system is considered. Thus, the number of active RF chains at the transmitter (or receiver) is a single integer whose optimal value is optimized via maximizing the EE performance [41]. In contrast, we consider the CF mmWave massive MIMO system, and the numbers of active RF chains at numerous APs are jointly optimized by maximizing the system's total achievable throughput. This results in unequal numbers of active RF chains at the APs, and an AP can even deactivate all RF chains. Second, [41] focuses on optimizing the number of RF chains at the transmitter through an optimal power allocation performed in the baseband domain, as part of the digital processor. This approach is, however, not applicable at the receiver because the RFS at the transmitter is formulated as a non-binary power allocation problem [41], while a binary switching algorithm is required for the RFS at the receivers. Furthermore, in the considered CF MIMO system, a receiver (an AP in this work) receives signals from multiple transmitters (UEs) simultaneously. When the UEs are equipped with small-sized arrays, the simple digital beamforming is generally employed. Therefore, in our considered system, the RF chains at the UEs are not required to be optimized, and the number of RF chains at the two sides cannot be the same for the purpose of dynamic RF chain activation. Similar to [41], the work [42] considers a point-to-point downlink mmWave mMIMO system. Conversely, in [43], a downlink multi-cell mmWave mMIMO system is considered. Assuming that the number of RF chains activated in each BS is equal to the number of users served by that BS, the RF chain activation problem is equivalent to the BS-user association problem with binary variables and constraints. In contrast, each AP serves all the users simultaneously in our considered uplink CF mMIMO systems. Therefore, such AP-user association problems will not occur. Thus, the approach proposed in [43] is not valid in our considered system. Because of these systematic differences, the algorithm presented in [41]–[43] cannot be leveraged for our proposed ARFA scheme, which thus requires novel algorithms as presented in Section IV.

The remainder of this paper is organized as follows: Section II introduces the system and channel models, whereas Section III describes the D-HBF and SC-HBF schemes. In Section IV, low-complexity ARFA algorithms are presented, and the power consumption of the proposed ARFA scheme is analyzed in Section V. Section VI presents simulation results, and the conclusion follows in Section VII.

II. System and channel model

A. System model

We consider the uplink of a CF mmWave massive MIMO system, where L APs and K user equipments (UEs) are distributed in a large area. Each AP is equipped with N_r receive antennas and $N(\leq N_r)$ RF chains, whereas each UE is equipped with N_t transmit antennas. All APs, which are connected to a CPU via fronthaul links, simultaneously and jointly serve K UEs. In the case where numerous antennas are used at the UEs, i.e., N_t is large, a hybrid precoding architecture can be employed to reduce the power consumption and hardware cost. However, in this work, to focus on the design of the hybrid combiner at the receiver side, we assume that N_t is small, and a fully digital precoder is used at the UEs. Furthermore, a fully-connected architecture is considered for analog combining, in which N RF chains are connected to N_r receive antennas via a network of NN_r PSs [44], [45]. We adopt a narrowband block-fading channel model [44], [45]. Let $\boldsymbol{H}_{kl} \in \mathbb{C}^{N_r \times N_t}$ denote the channel matrix representing channels between the kth UE and lth AP. With the presence of channel estimation errors, a noisy channel is known to the AP. Following the channel estimation error model in [46] for mmWave channels, we have $\boldsymbol{H}_{kl} = \hat{\boldsymbol{H}}_{kl} + \boldsymbol{\Delta}_{kl}$, where $\boldsymbol{\Delta}_{kl}$ represents the channel estimation error that is uncorrelated with $\boldsymbol{H}_{kl}$ based on the minimum mean square error (MMSE) estimation property [12]. The entries of $\boldsymbol{\Delta}_{kl}$ are independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance ε_{kl}^2. The analog combined signal at the lth AP can be expressed as

$$\boldsymbol{y}_l = \sqrt{\rho} \sum_{k=1}^{K} \boldsymbol{F}_l^H \left(\hat{\boldsymbol{H}}_{kl} + \boldsymbol{\Delta}_{kl} \right) \boldsymbol{x}_k + \boldsymbol{F}_l^H \boldsymbol{z}_l, \qquad (1)$$

where $\boldsymbol{x}_k \in \mathbb{C}^{N_t \times 1}$ is the vector of symbols sent from the kth UE such that $\mathbb{E}\left\{\boldsymbol{x}_k^H \boldsymbol{x}_k\right\} = 1, \forall k$, and $\boldsymbol{z}_l \sim \mathcal{CN}(0, \sigma_n^2 \boldsymbol{I}_{N_r})$ is the i.i.d. additive white Gaussian noise (AWGN) vector, where $\boldsymbol{I}_{N_r}$ denotes the identity matrix of size $N_r \times N_r$. Furthermore, ρ represents the average transmit power. The analog combining matrix at the lth AP, i.e., $\boldsymbol{F}_l \in \mathbb{C}^{N_r \times N}$, is given by $\boldsymbol{F}_l = [\boldsymbol{f}_{l1}, \ldots, \boldsymbol{f}_{lN}]$, where $\boldsymbol{f}_{ln} = [f_{ln}^{(1)}, \ldots, f_{ln}^{(N_r)}]^T$ is the analog weight vector corresponding to the nth RF chain at the lth AP, and $f_{ln}^{(i)}$ is the ith element of $\boldsymbol{f}_{ln}$, which has the constant amplitude $1/\sqrt{N_r}$ but different phases, i.e., $f_{ln}^{(i)} = \frac{1}{\sqrt{N_r}} e^{j\theta_{ln}^{(i)}}, \forall l, n, i$.

To detect the symbols transmitted from the UEs, the lth AP multiplies the received signal with a digital combining matrix $\boldsymbol{W}_l \in \mathbb{C}^{N \times KN_t}$, which leads to $\boldsymbol{r}_l = \boldsymbol{W}_l^H \boldsymbol{y}_l = \boldsymbol{W}_l^H \boldsymbol{F}_l^H \sum_{k=1}^{K} \left(\hat{\boldsymbol{H}}_{kl} + \boldsymbol{\Delta}_{kl} \right) \boldsymbol{x}_k + \boldsymbol{F}_l^H \boldsymbol{z}_l$. As a result, the signal input-output relationship at the lth AP can be expressed as

$$\boldsymbol{r}_l = \sqrt{\rho} \sum_{k=1}^{K} \boldsymbol{W}_l^H \boldsymbol{F}_l^H \left(\hat{\boldsymbol{H}}_{kl} + \boldsymbol{\Delta}_{kl} \right) \boldsymbol{x}_k + \boldsymbol{W}_l^H \boldsymbol{F}_l^H \boldsymbol{z}_l, \qquad (2)$$

Then, the so-obtained locally detected signals $\boldsymbol{r}_l, \forall l$ are sent to the CPU via a fronthaul network to perform the final signal detection. In this work, we assume a simple centralized decoding scheme at the CPU, which requires minimal information exchange between the APs and CPU. In this scheme, the final decoded signal at the CPU is given as the average of the local estimates, that is, $\frac{1}{L} \sum_{l=1}^{L} \boldsymbol{r}_l$ [10].

The composite received signal available at the CPU can be expressed as

$$\begin{bmatrix} \boldsymbol{r}_1 \\ \vdots \\ \boldsymbol{r}_L \end{bmatrix} = \sqrt{\rho} \sum_{k=1}^{K} \begin{bmatrix} \boldsymbol{W}_1^H \boldsymbol{F}_1^H \left(\hat{\boldsymbol{H}}_{k1} + \boldsymbol{\Delta}_{k1} \right) \\ \vdots \\ \boldsymbol{W}_L^H \boldsymbol{F}_L^H \left(\hat{\boldsymbol{H}}_{kL} + \boldsymbol{\Delta}_{kL} \right) \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{W}_1^H \boldsymbol{F}_1^H \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{W}_L^H \boldsymbol{F}_L^H \boldsymbol{z}_L \end{bmatrix} . \tag{3}$$

Let $\boldsymbol{F} = \text{diag}\left\{ \boldsymbol{F}_1, \ldots, \boldsymbol{F}_L \right\} \in \mathbb{C}^{LN_r \times LN}$ and $\boldsymbol{W} = \text{diag}\left\{ \boldsymbol{W}_1, \ldots, \boldsymbol{W}_L \right\} \in \mathbb{C}^{LN \times LKN_t}$ be block-diagonal matrices containing the analog and digital combiners for all L APs. In this work, we refer to $\boldsymbol{F}$ and $\boldsymbol{W}$ as *global combiners*, whereas $\{\boldsymbol{F}_1, \ldots, \boldsymbol{F}_L\}$ and $\{\boldsymbol{W}_1, \ldots, \boldsymbol{W}_L\}$ for the signal combined at the APs $\{1, \ldots, L\}$ are referred to as the *local combiners*. Furthermore, let $\hat{\boldsymbol{H}}_l = \left[\hat{\boldsymbol{H}}_{1l}, \ldots, \hat{\boldsymbol{H}}_{Kl} \right] \in \mathbb{C}^{N_r \times KN_t}$ and $\boldsymbol{\Delta}_l = \left[\boldsymbol{\Delta}_{1l}, \ldots, \boldsymbol{\Delta}_{Kl} \right] \in \mathbb{C}^{N_r \times KN_t}$ denote the estimated channel matrix between the K UEs and lth AP and its corresponding estimation error. Define

$$\boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_1 \\ \vdots \\ \boldsymbol{r}_L \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_K \end{bmatrix}, \boldsymbol{z} = \begin{bmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_L \end{bmatrix}, \hat{\boldsymbol{H}} = \begin{bmatrix} \hat{\boldsymbol{H}}_1 \\ \vdots \\ \hat{\boldsymbol{H}}_L \end{bmatrix}, \boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\Delta}_1 \\ \vdots \\ \boldsymbol{\Delta}_L \end{bmatrix},$$

where $\boldsymbol{r} \in \mathbb{C}^{LKN_t \times 1}$, $\boldsymbol{x} \in \mathbb{C}^{KN_t \times 1}$, $\boldsymbol{z} \in \mathbb{C}^{LN_r \times 1}$, and $\hat{\boldsymbol{H}}, \boldsymbol{\Delta} \in \mathbb{C}^{LN_r \times KN_t}$. Then, (3) can be rewritten in a more compact form as

$$\begin{aligned} \boldsymbol{r} &= \sqrt{\rho} \boldsymbol{W}^H \boldsymbol{F}^H \hat{\boldsymbol{H}} \boldsymbol{x} + \sqrt{\rho} \boldsymbol{W}^H \boldsymbol{F}^H \boldsymbol{\Delta} \boldsymbol{x} + \boldsymbol{W}^H \boldsymbol{F}^H \boldsymbol{z} \\ &= \sqrt{\rho} \boldsymbol{W}^H \boldsymbol{F}^H \hat{\boldsymbol{H}} \boldsymbol{x} + \boldsymbol{W}^H \boldsymbol{F}^H \left(\sqrt{\rho} \boldsymbol{\Delta} \boldsymbol{x} + \boldsymbol{z} \right), \\ &= \sqrt{\rho} \boldsymbol{W}^H \boldsymbol{F}^H \hat{\boldsymbol{H}} \boldsymbol{x} + \boldsymbol{W}^H \boldsymbol{F}^H \hat{\boldsymbol{z}}, \end{aligned} \tag{4}$$

where $\hat{\boldsymbol{z}} = \sqrt{\rho} \boldsymbol{\Delta} \boldsymbol{x} + \boldsymbol{z} \sim \mathcal{CN}(0, \sigma_n^2 \boldsymbol{I}_{LN_r} + \rho \boldsymbol{\Psi})$, with $\boldsymbol{\Psi} = \text{diag}\left\{ N_t \sum_{k=1}^{K} \varepsilon_{k1} \boldsymbol{I}_{N_r}, \ldots, N_t \sum_{k=1}^{K} \varepsilon_{kL} \boldsymbol{I}_{N_r} \right\}$. By assuming that $\varepsilon^2 = \sum_{k=1}^{K} \varepsilon_{kl}, \forall l$, we have that $\boldsymbol{\Psi} = N_t \varepsilon^2 \boldsymbol{I}_{LN_r}$.

We note that the analog processing is separately performed at the APs because the ADCs, RF chains, and PSs are installed at the APs. However, $\{\boldsymbol{F}_1, \ldots, \boldsymbol{F}_L\}$ can be generated either at the APs based on their local CSI or at the CPU based on the global CSI. Once the analog combiner is obtained, digital processing can be carried out at the corresponding AP. We follow the common assumption in [5], [20] that the digital combining is performed at the APs individually. Therefore, in this work, the D-HBF scheme refers to the HBF with analog combiners generated at each AP separately, whereas SC-HBF implies that the analog combiners are computed at the CPU based on the global CSI.

B. Channel model

The channels between the UEs and APs are modeled based on the narrow-band geometric Saleh–Valenzuela channel model, which is widely adopted for mmWave systems [41], [44]–[46]. Specifically, the channel matrix between the lth AP and kth UE can be expressed as [45]

$$\boldsymbol{H}_{kl} = \sqrt{\frac{G_{\text{a}}}{\beta_{kl}} \frac{N_r N_t}{P_{kl}}} \sum_{p=1}^{P_{kl}} \alpha_{kl}^{(p)} \boldsymbol{a}_r(\phi_{kl}^{(p)}) \boldsymbol{a}_t^H(\psi_{kl}^{(p)}), \tag{5}$$

where P_{kl} is the number of effective channel paths corresponding to a limited number of scatters between the kth UE and the lth AP, $\alpha_{kl}^{(p)}$ is the gain of the pth path. Furthermore, $\phi_{kl}^{(p)}$ and $\psi_{kl}^{(p)}$ are the azimuth angles of arrival (AoA) and departure (AoD), respectively. All channel path gains $\alpha_{kl}^{(p)}$ are assumed to be i.i.d. Gaussian random variables with zero mean and unit variance, i.e., $\alpha_{kl}^{(p)} \sim \mathcal{CN}(0,1), \forall l, k$. Furthermore, $\boldsymbol{a}_r(\cdot)$ and $\boldsymbol{a}_t(\cdot)$ represent the normalized receive and transmit array response vectors at an AP and a UE, respectively, which depend on the structure of the antenna array. In this work, we consider a uniform linear array (ULA), where $\boldsymbol{a}_r(\cdot)$ is given by $\boldsymbol{a}_r(\phi) = \frac{1}{\sqrt{N_r}} [1, e^{j \frac{2\pi}{\lambda} d_s \sin(\phi)}, \ldots, e^{j(N_r - 1) \frac{2\pi}{\lambda} d_s \sin(\phi)}]^T$ with λ denoting the wavelength of the signal and d_s being the antenna spacing in the antenna array [45]. For the transmitter, $\boldsymbol{a}_t(\cdot)$ can be written in a similar fashion. In (5), G_{a} is the antenna gain. Furthermore, β_{kl} represents the path loss between the kth UE and lth AP, which is given in dB as [47], [48]

$$\beta_{kl}[\text{dB}] = \beta_0 + 10\vartheta \log_{10} \left(\frac{d_{kl}}{d_0} \right) + A_{\xi}. \tag{6}$$

Here, $\beta_0 = 10 \log_{10} \left(\frac{4\pi d_0}{\lambda} \right)^2$, where $d_0 = 1$ m, ϑ is the average path loss exponent over distance, and A_{ξ} is a zero-mean Gaussian random variable with a standard deviation ξ in dB representing the effect of shadow fading.

III. SC-HBF and D-HBF

Because $\hat{\boldsymbol{H}}$ and $\boldsymbol{\Delta}$ are independent, the combiners $\boldsymbol{W}$ and $\boldsymbol{F}$ are independent of $\boldsymbol{\Delta}$. The AP treats the channel estimate as the true channel, and the last term in (4) is considered as the effective noise. Therefore, the total achievable rate R can be expressed as [44]

$$R = \log_2 \left| \boldsymbol{I}_{LKN_t} + \rho \boldsymbol{R}^{-1} \boldsymbol{W}^H \boldsymbol{F}^H \hat{\boldsymbol{H}} \hat{\boldsymbol{H}}^H \boldsymbol{F} \boldsymbol{W} \right|, \tag{7}$$

where $\boldsymbol{R} = \left(\sigma_n^2 + \rho N_t \varepsilon^2 \right) \boldsymbol{W}^H \boldsymbol{F}^H \boldsymbol{F} \boldsymbol{W}$. We aim to design hybrid combiners that maximize R. However, the joint design of $\boldsymbol{F}$ and $\boldsymbol{W}$ are significantly challenging. Therefore, we adopted the scheme in [49]. Specifically, we decoupled the design of $\boldsymbol{F}$ and $\boldsymbol{W}$ to first design the analog combiner assuming an optimal digital combiner and then determine the optimal digital combiner for the derived analog one [49]. Note that there is no constraint on the digital combining coefficients. Therefore, with an optimal and fixed digital beamformer, the analog beamforming design problem is formulated as

$$(\text{P}_{\text{a}}) \quad \max_{\boldsymbol{F}_1, \ldots, \boldsymbol{F}_L} \quad \log_2 \left| \boldsymbol{I}_{LN} + \gamma \left(\boldsymbol{F}^H \boldsymbol{F} \right)^{-1} \boldsymbol{F}^H \hat{\boldsymbol{H}} \hat{\boldsymbol{H}}^H \boldsymbol{F} \right|, \tag{8a}$$

$$\text{s.t.} \quad \boldsymbol{F} = \text{diag}\left\{ \boldsymbol{F}_1, \ldots, \boldsymbol{F}_L \right\}. \tag{8b}$$

$$\boldsymbol{f}_{ln} \in \mathcal{F}, \forall l, n, \tag{8c}$$

where $\gamma = \frac{\rho}{\sigma_n^2 + N_t \rho \varepsilon^2}$, and $\mathcal{F}$ is the set of feasible analog combining coefficients $f_{ln}^{(i)} = \frac{1}{\sqrt{N_r}} e^{j\theta_{ln}^{(i)}}, \forall l, n, i$, which have constant modulus $\left| f_{ln}^{(i)} \right| = \frac{1}{\sqrt{N_r}}$. To simplify the objective function in (P_{a}), we assume $\boldsymbol{F}_l^H \boldsymbol{F}_l \approx \boldsymbol{I}_N$ [44], [49], which is tight in the considered CF mmWave massive MIMO system with a sufficiently large number of antennas

Algorithm 1 SC-HBF scheme

```
Output: ${F_1^\star, \dots, F_L^\star}$, ${W_1^\star, \dots, W_L^\star}$, and ${R_1^{\mathrm{a}}, \dots, R_L^{\mathrm{a}}}$.
1:  At the CPU: $\boldsymbol{Q}_0 = \boldsymbol{I}_K$
2:  for $l = 1 \to L$ do
3:      for $n = 1 \to N$ do
4:          Set $\boldsymbol{u}_{ln}^\star$ to the singular vector corresponding to the
            $n$th largest singular value of $\hat{\boldsymbol{H}}_l \boldsymbol{Q}_{l-1}^{-1} \hat{\boldsymbol{H}}_l^H$.
5:          $\boldsymbol{f}_{ln}^\star = \frac{1}{\sqrt{N_r}} \mathcal{Q}(\boldsymbol{u}_{ln}^\star)$
6:      end for
7:      $\boldsymbol{F}_l^\star = [\boldsymbol{f}_{l1}^\star, \dots, \boldsymbol{f}_{lN}^\star]$
8:      $\boldsymbol{G}_l = \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l^\star \boldsymbol{F}_l^{\star H} \hat{\boldsymbol{H}}_l$
9:      $\boldsymbol{Q}_l = \boldsymbol{Q}_{l-1} + \gamma \boldsymbol{G}_l$
10:     $R_l^{\mathrm{a}} = \log_2 \left( \boldsymbol{I}_N + \gamma \boldsymbol{F}_l^{\star H} \hat{\boldsymbol{H}}_l \boldsymbol{Q}_{l-1}^{-1} \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l^\star \right)$
11: end for
12: At the $l$th AP: compute $\boldsymbol{W}_l^\star$ based on (14).
```

deployed at each AP. Consequently, we have $\boldsymbol{F}^H \boldsymbol{F} \approx \boldsymbol{I}_{LN}$, and the objective function in (8a), which is the sum rate achieved by analog combining, can be approximated by $\log_2 \left| \boldsymbol{I}_{LN} + \gamma \boldsymbol{F}^H \hat{\boldsymbol{H}} \hat{\boldsymbol{H}}^H \boldsymbol{F} \right| \triangleq R^{\mathrm{a}}$. Therefore, the optimal analog combiners can be solved approximately in

$$(\mathrm{P}_{\mathrm{a}}') \quad \max_{\boldsymbol{F}_1, \dots, \boldsymbol{F}_L} R^{\mathrm{a}}, \text{ s.t. } (8\mathrm{b}), (8\mathrm{c}). \tag{9}$$

The objective function R^{a} of $(\mathrm{P}_{\mathrm{a}}')$ is further investigated in the following theorem.

Theorem 1: In a CF mmWave massive MIMO system with L APs, we have $R^{\mathrm{a}} = \sum_{l=1}^{L} R_l^{\mathrm{a}}$, where

$$R_l^{\mathrm{a}} = \log_2 \det \left(\boldsymbol{I}_N + \gamma \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l \boldsymbol{Q}_{l-1}^{-1} \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l \right), \tag{10}$$

with $\boldsymbol{Q}_0 = \boldsymbol{I}_{KN_t}$ and

$$\boldsymbol{Q}_{l-1} = \boldsymbol{Q}_{l-2} + \gamma \hat{\boldsymbol{H}}_{l-1}^H \boldsymbol{F}_{l-1} \boldsymbol{F}_{l-1}^H \hat{\boldsymbol{H}}_{l-1}. \tag{11}$$

Proof: See Appendix A. □

Based on Theorem 1, R^{a} can be maximized by optimizing $\{R_1^{\mathrm{a}}, \dots, R_L^{\mathrm{a}}\}$ corresponding to APs $\{1, \dots, L\}$. As a result, finding $\boldsymbol{F}_1^\star, \dots, \boldsymbol{F}_L^\star$ in $(\mathrm{P}_{\mathrm{a}}')$ can be done by maximizing R_l^{a}:

$$\boldsymbol{F}_l^\star = \arg \max_{\boldsymbol{F}_l} R_l^{\mathrm{a}}, \forall l, \text{ s.t. } \boldsymbol{f}_{l1}, \dots, \boldsymbol{f}_{lN} \in \mathcal{F}. \tag{12}$$

Let $\{\boldsymbol{u}_{l1}^\star, \dots, \boldsymbol{u}_{lN}^\star\}$ be the N singular vectors corresponding to N largest singular values of $\hat{\boldsymbol{H}}_l \boldsymbol{Q}_{l-1}^{-1} \hat{\boldsymbol{H}}_l^H$, which are in decreasing order. Then, columns $\{\boldsymbol{f}_{l1}^\star, \dots, \boldsymbol{f}_{lN}^\star\}$ of a near-optimal solution to (12) can be obtained by quantizing $\{\boldsymbol{u}_{l1}^\star, \dots, \boldsymbol{u}_{lN}^\star\}$, respectively, to the nearest vector in $\mathcal{F}$ [22], i.e.,

$$\boldsymbol{f}_{ln}^\star = \arg \min_{\boldsymbol{f}_{ln} \in \mathcal{F}} \left\| \boldsymbol{u}_{ln}^\star - \boldsymbol{f}_{ln} \right\|^2, \forall n. \tag{13}$$

At the lth AP, once the analog combiner $\boldsymbol{F}_l^\star$ is found, the optimal digital combiner is given as the MMSE solution, i.e.,

$$\boldsymbol{W}_l^\star = \boldsymbol{J}^{-1} \boldsymbol{F}_l^{\star H} \hat{\boldsymbol{H}}_l, \tag{14}$$

where $\boldsymbol{J} = \boldsymbol{F}_l^{\star H} \hat{\boldsymbol{H}}_l \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l^\star + \frac{1}{\gamma} \boldsymbol{F}_l^{\star H} \boldsymbol{F}_l^\star$ [49]. In the following subsections, we propose two HBF schemes in which the analog combiners are derived based on different assumptions for CSI.

A. SC-HBF

It is evident from (10) and (11) that R_l^{a} depends not only on $\hat{\boldsymbol{H}}_l$, but also on $\hat{\boldsymbol{H}}_{l-1}, \hat{\boldsymbol{H}}_{l-2}, \dots, \hat{\boldsymbol{H}}_1$. Therefore, from (10) and (12), it is observed that finding $\boldsymbol{F}_l^\star$ requires not only $\hat{\boldsymbol{H}}_l$ but also $\hat{\boldsymbol{H}}_{l-1}, \hat{\boldsymbol{H}}_{l-2}, \dots, \hat{\boldsymbol{H}}_1$. This is similar to the requirements for determining analog beamformers for sub-arrays in the partially-connected HBF architecture [25], [40]. As a result, solving $\{\boldsymbol{F}_1^\star, \boldsymbol{F}_2^\star, \dots, \boldsymbol{F}_L^\star\}$ requires the CSI of the channels between all L APs and K UEs, i.e., $\left\{ \hat{\boldsymbol{H}}_1, \hat{\boldsymbol{H}}_2, \dots, \hat{\boldsymbol{H}}_L \right\}$, which can be available at the CPU; hence, finding $\boldsymbol{F}^\star$ based on (12) requires a SC-HBF scheme.

Algorithm 1 presents the proposed SC-HBF scheme to obtain $\{\boldsymbol{F}_1^\star, \boldsymbol{F}_2^\star, \dots, \boldsymbol{F}_L^\star\}$. In particular, in steps 3–6, the combining vector $\boldsymbol{f}_{ln}^\star$ is obtained by quantizing $\boldsymbol{u}_{ln}^\star$ based on (13), which ensures that the resultant analog combiners belong to the feasible set $\mathcal{F}$. Then, $\boldsymbol{F}_l^\star$ is found in step 7 and $\boldsymbol{G}_l$ is computed in step 8, followed by $\boldsymbol{Q}_l$ being updated in step 9 based on (11). In step 10, R_l^{a} corresponding to the lth AP is computed. Furthermore, the digital combiner is computed at each AP, as in step 12. We note that in Algorithm 1, steps 1–11 are performed at the CPU, whereas step 12 is performed at the APs.

B. D-HBF

Let $\{\tilde{\boldsymbol{u}}_{l1}^\star, \dots, \tilde{\boldsymbol{u}}_{lN}^\star\}$ be the N singular vectors corresponding to the N largest singular values of $\hat{\boldsymbol{H}}_l$, which are in decreasing order. Furthermore, define

$$\tilde{\boldsymbol{f}}_{ln}^\star = \arg \min_{\boldsymbol{f}_{ln} \in \mathcal{F}} \left\| \tilde{\boldsymbol{u}}_{ln}^\star - \boldsymbol{f}_{ln} \right\|^2, \forall n. \tag{15}$$

Then, in the D-HBF scheme, the optimal local analog combiner generated at the lth AP based on $\hat{\boldsymbol{H}}_l$ can be given as $\tilde{\boldsymbol{F}}_l^\star = \left[\tilde{\boldsymbol{f}}_{l1}^\star, \dots, \tilde{\boldsymbol{f}}_{lN}^\star \right]$ [22]. Let $\tilde{\boldsymbol{F}}^\star = \operatorname{diag}\left\{ \tilde{\boldsymbol{F}}_1^\star, \dots, \tilde{\boldsymbol{F}}_L^\star \right\}$. In the following theorem, we show that the total achievable rate achieved by analog combining in the D-HBF scheme is approximately equal to that in SC-HBF.

Theorem 2: In CF mmWave massive MIMO systems with large L and low SNRs due to the significant pathloss in the mmWave channels, the total achievable rate achieved by the analog combining in the D-HBF scheme, denoted by $\tilde{R}^{\mathrm{a}}$, is approximately the same as that of the SC-HBF scheme, i.e.,

$$\tilde{R}^{\mathrm{a}} = \log_2 \det \left(\boldsymbol{I}_{KN_t} + \gamma \hat{\boldsymbol{H}}^H \tilde{\boldsymbol{F}}^\star \tilde{\boldsymbol{F}}^{\star H} \hat{\boldsymbol{H}} \right) \approx R^{\mathrm{a}}, \tag{16}$$

where R^{a} is given in Theorem 1.

Proof: See Appendix B. □

It is observed that D-HBF can be performed with considerably lower computational complexity than SC-HBF. Specifically, only N singular vectors corresponding to the N largest singular values of the channel matrix are required to form the analog combiner. In contrast, additional matrix inversions, multiplications, and additions are performed in steps 4, 8, and 9 of Algorithm 1 for the SC-HBF scheme. Notably, despite the simpler implementation and lower complexity of the D-HBF scheme, it can approximately achieve the performance of SC-HBF, as stated in Theorem 2. Furthermore, the D-HBF scheme requires less information exchange between the APs and CPU. Specifically, only KN_t complex numbers in $\boldsymbol{r}_l$ are sent to the CPU on the fronthaul link to perform the final soft detection [5], [20], whereas the exchange of the CSI and analog combining matrix is not required, in contrast to the SC-HBF scheme

Theorem 2 indicates that the global CSI at the CPU is not very helpful in improving the achievable rate in CF mmWave massive MIMO systems. However, this does not mean that further information exchange via the fronthaul links is completely useless. Indeed, the information exchange between the APs and CPU in CF massive MIMO systems can also be exploited to improve the EE. In the next section, it is discussed that by adaptively activating RF chains based on global CSI in SC-HBF or on limited information in D-HBF, the power consumption can be reduced, which leads to improved EE.

IV. Adaptive RF chain activation

A. Problem formulation and basic ideas

The global analog combiner $\boldsymbol{F}$ can be expressed as

$$\boldsymbol{F} = \mathrm{diag}\{\underbrace{[\boldsymbol{f}_{11},\ldots,\boldsymbol{f}_{1N}]}_{\text{analog combiner at AP }1},\ldots,\underbrace{[\boldsymbol{f}_{L1},\ldots,\boldsymbol{f}_{LN}]}_{\text{analog combiner at AP }L}\},$$

and the following facts are noted:

- Based on Theorem 1, the total achievable rate obtained by analog combining can be expressed as a sum of $\{R_1^{\mathrm{a}},\ldots,R_L^{\mathrm{a}}\}$ corresponding to APs $\{1,\ldots,L\}$. In CF mmWave massive MIMO systems, the APs are distributed in a large area, and their communication channels experience different path losses and shadowing effects. Therefore, the contributions of the local analog combiners at different APs to the total achievable rate are of different significances.
- In a local analog combiner $\boldsymbol{F}_l$, combining vectors $\{\boldsymbol{f}_{l1},\ldots,\boldsymbol{f}_{lN}\}$ have different contributions to the sub-rate R_l^{a} given in (10). Specifically, the contribution of $\boldsymbol{f}_{ln}$ is more significant than that of $\boldsymbol{f}_{lm}$ if $n < m$ because n and m are the indices of the ordered singular values of $\hat{\boldsymbol{H}}_l\boldsymbol{Q}_{l-1}^{-1}\hat{\boldsymbol{H}}_l^H$ in SC-HBF and of $\hat{\boldsymbol{H}}_l$ in D-HBF.

As a result, it is likely that a subset of analog combining vectors in $\{\boldsymbol{f}_{11},\ldots,\boldsymbol{f}_{LN}\}$ are insignificant and can be removed from the global combiner $\boldsymbol{F}$ without causing considerable performance loss. We note that at an AP, an analog combining vector represents the effect of N_r PSs connected to an RF chain, followed by an ADC. Therefore, an insignificant analog combining vector can be removed from signal combining by turning off its corresponding RF chain, ADC, and PSs, which results in a reduction in the total power consumption. Motivated by this, we propose an ARFA scheme that selectively activates RF chains at the APs. Let $\boldsymbol{n} = \{n_1,\ldots,n_L\}$, where n_l is the number of turned-on RF chains out of N RF chains installed at the lth AP, $0 \le n_l \le N$. We note that for $n_l = 0$, all the RF chains at the lth AP are turned off, and this AP does not consume any power for signal combining. The optimal activation of RF chains at the APs can be performed based on the following remark.

Remark 1: Because $\boldsymbol{f}_{ln}$ is always more important at the lth AP than $\boldsymbol{f}_{lm}$ with $n < m$ in terms of achievable rate, the problem of optimal activation of RF chains at an AP is equivalent to finding an optimal number of turned-on RF chains at that AP. Specifically, for the lth AP, if the ARFA scheme suggests using $n_l^\star$ RF chains, then the first $n_l^\star$ RF chains corresponding to $\{\boldsymbol{f}_{l1},\ldots,\boldsymbol{f}_{ln_l^\star}\}$ are selected for analog signal combining, whereas the others are deactivated to save power.

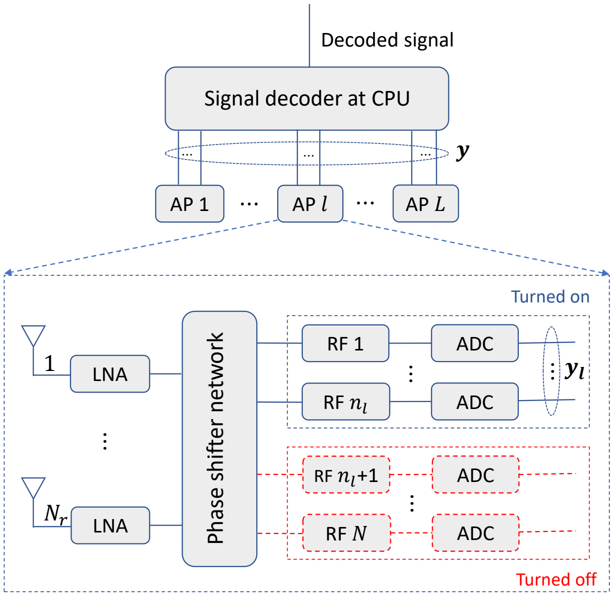


Fig. 1. HBF architecture with the ARFA scheme. In the phase-shifter network, n_lN_r out of NN_r PSs are turned on at the lth AP.

The HBF architecture with the proposed ARFA scheme is illustrated in Fig. 1. As an example, at the lth AP, only n_l out of N RF chains are turned on. Furthermore, the ADCs and PSs connected to the inactive RF chains are also turned off. Consequently, the local combiner $\boldsymbol{F}_l$ consists of only n_l analog combining vectors, i.e., $\boldsymbol{F}_l = \{\boldsymbol{f}_1,\ldots,\boldsymbol{f}_{n_l}\}$.

Unlike the conventional fixed-activation HBF, in the proposed ARFA scheme, the global combiner $\boldsymbol{F}$, R_l^{a}, and R depend on $\boldsymbol{n}$. Therefore, in this section, they are expressed as functions of $\boldsymbol{n}$, i.e., $\boldsymbol{F}(\boldsymbol{n})$, $R_l^{\mathrm{a}}(\boldsymbol{n})$, and $R(\boldsymbol{n})$, respectively. We limit the total number of turned-on RF chains in the ARFA scheme to $L\bar{n}$, i.e., $\sum_{l=1}^{L} n_l = L\bar{n}$, where $\bar{n}(\le N)$ is the average number of activated RF chains at each AP. Based on Remark 1, the optimal activation of RF chains at the APs in the ARFA scheme can be performed by solving

$$\boldsymbol{n}^\star = \arg\max_{\boldsymbol{n}\in\mathcal{S}} R(\boldsymbol{n}), \tag{17}$$

where $\mathcal{S} = \left\{\boldsymbol{n} : 0 \le n_l \le N, \sum_{l=1}^{L} n_l = L\bar{n}\right\}$ is the feasible set of $\boldsymbol{n}$. The optimal $\boldsymbol{n}^\star$ in (17) can be found by exhaustive search over the entire feasible set $\mathcal{S}$. However, in CF mmWave massive MIMO systems, L is large; thus, an excessively large number of candidates for $\boldsymbol{n}$ need to be examined in the exhaustive search, which is almost computationally prohibitive. In the following subsections, we propose three low-complexity algorithms to find $\boldsymbol{n}^\star$. By abuse of notation, we use $\boldsymbol{F}^\star$ for the global combiner found by the proposed ARFA algorithms. Furthermore, we note that Algorithm 1 can be easily modified for the ARFA scheme by replacing N in steps 3, 7, and 10

Algorithm 2 HBF with SC-ARFA

```
Output: $\boldsymbol{F}^\star$
1: Initialize $\boldsymbol{n} = [n_1, \ldots, n_L]$ with $n_l = \bar{n}, \forall l$.
2: Use Algorithm 1 to find $\boldsymbol{F}(\boldsymbol{n})$, $R^{\mathrm{a}}_l(\boldsymbol{n}), \forall l$, and $R(\boldsymbol{n})$.
3: $\boldsymbol{F}^\star = \boldsymbol{F}(\boldsymbol{n})$, $R^\star = R(\boldsymbol{n})$.
4: Obtain $\{n_{[1]}, \ldots, n_{[L]}\}$ s.t. $R^{\mathrm{a}}_{[1]}(\boldsymbol{n}) > \ldots > R^{\mathrm{a}}_{[L]}(\boldsymbol{n})$.
5: $i = 1$, $k = L$
6: while $i < k$ do
7:     while $n_{[i]} = N$ do
8:         $i = i + 1$
9:     end while
10:    while $n_{[k]} = 0$ do
11:        $k = k - 1$
12:    end while
13:    Update $\boldsymbol{n}$: $n_{[i]} = n_{[i]} + 1$, $n_{[k]} = n_{[k]} - 1$.
14:    Use Algorithm 1 to find $\boldsymbol{F}(\boldsymbol{n})$ and $R(\boldsymbol{n})$ with the updated $\boldsymbol{n}$.
15:    Update $\boldsymbol{F}^\star = \boldsymbol{F}(\hat{\boldsymbol{n}}^\star)$ if $R(\hat{\boldsymbol{n}}^\star) > R^\star$.
16: end while
17: At the $l$th AP, compute $\boldsymbol{W}_l^\star$ based on (14), $\forall l$.
```

with n_l, reflecting a dynamic number of analog combining vectors for the local combiner $\boldsymbol{F}_l$ at the lth AP.

B. ARFA with SC-HBF (SC-ARFA)

In SC-ARFA, the ARFA scheme is incorporated with SC-HBF, and the optimal numbers of active RF chains at the APs are found at the CPU based on the global CSI. The idea to find $\boldsymbol{n}^\star$ is to turn on/off as many RF chains as possible at the APs corresponding to the largest/smallest R^{a}_l, as presented in Algorithm 2. In steps 1–3, all elements of $\boldsymbol{n}$ are set to $\bar{n}$, then $\boldsymbol{F}(\boldsymbol{n}), R^{\mathrm{a}}_l(\boldsymbol{n}), \forall l$, and $R(\boldsymbol{n})$ are computed. In step 4, the elements of $\boldsymbol{n}$ are ordered to obtain $\{n_{[1]}, \ldots, n_{[L]}\}$ in the decreasing order of sub-rates $\{R^{\mathrm{a}}_1(\boldsymbol{n}), \ldots, R^{\mathrm{a}}_L(\boldsymbol{n})\}$. Therefore, $n_{[i]}$ is the number of turned-on RF chains at the AP with the ith largest sub-rates, i.e., $R^{\mathrm{a}}_{[i]}(\boldsymbol{n})$.

In step 5, we initialize $i = 1$ and $k = L$. In steps 6–16, $n_{[i]}$ is increased by one, whereas $n_{[k]}$ is decreased by one, in each iteration. We note that in step 13, $n_{[i]}$ and $n_{[k]}$ are updated simultaneously to guarantee $\sum_{l=1}^{L} n_{[l]} = L\bar{n}$. The updates of $n_{[i]}$ and $n_{[k]}$ result in a new candidate $\boldsymbol{n}$. Hence, $\boldsymbol{F}(\boldsymbol{n})$ and $R(\boldsymbol{n})$ are found in step 14, and $\boldsymbol{F}^\star$ is updated if the performance is improved, as shown in step 15. Once $n_{[i]}$ reaches the maximum, i.e., N, the number of turned-on RF chains at the AP associated with the $(i+1)$th largest sub-rate, i.e., $n_{[i+1]}$, is considered, as shown in steps 7–9. In contrast, if $n_{[k]}$ reaches the minimum, i.e., zero, $n_{[k-1]}$ is considered next, as shown in steps 10–12. This iterative process is terminated if $i \geq k$, for which we have $R^{\mathrm{a}}_{[i]}(\boldsymbol{n}) \leq R^{\mathrm{a}}_{[k]}(\boldsymbol{n})$ and the increase (decrease) in $n_{[i]}$ $(n_{[k]})$ is unlikely to provide performance improvement. Once all the analog combiners are determined and sent to the APs, the digital combiner at each AP is determined, as in step 17.

We note that the ARFA process needs to be performed at the CPU to jointly optimize the numbers of RF chains at all APs. In the SC-ARFA schemes, the global CSI is exploited to evaluate the candidates for $\boldsymbol{n}$. However, the employment of SC-HBF in these schemes requires high computational complexity and a large amount of information exchanged between the CPU and APs, as discussed in Section III-B. This motivates us to propose an ARFA scheme incorporated with D-HBF in the next subsection.

Algorithm 3 HBF with SV-based D-ARFA

```
Output: $\boldsymbol{F}^\star$
1: Each AP finds the $N$ largest singular values of its channel matrix and sends them to the CPU. Specifically, the $l$th AP finds and sends a vector $\boldsymbol{e}_l = \left[\lambda_1^{(l)}, \ldots, \lambda_N^{(l)}\right]$, where $\lambda_n^{(l)}$ is the $n$th largest singular value of $\hat{\boldsymbol{H}}_l$.
2: The CPU finds $\lambda_{L\bar{n}}$, which is the $(L\bar{n})$th largest element in the singular value set $\{\lambda_1^{(1)}, \ldots, \lambda_N^{(1)}, \ldots, \lambda_1^{(L)}, \ldots, \lambda_N^{(L)}\}$ received from all APs.
3: for $l = 1 \to L$ do
4:     The CPU sets $n_l^\star$ to the number of elements in $\boldsymbol{e}_l$ that are not smaller than $\lambda_{L\bar{n}}$ and sends $n_l^\star$ to the $l$th AP.
5:     The $l$th AP determines its local analog combiner $\boldsymbol{F}_l^\star$ for $n_l^\star$ RF chains, i.e., $\boldsymbol{F}_l^\star = [\tilde{\boldsymbol{f}}_{l1}^\star, \ldots, \tilde{\boldsymbol{f}}_{ln_l^\star}^\star]$, where $\tilde{\boldsymbol{f}}_{ln}^\star$ is given by (15), and determines $\boldsymbol{W}_l^\star$ based on (14).
6: end for
```

C. ARFA with D-HBF (D-ARFA)

Without global CSI, the ARFA scheme can be performed if the CPU knows the qualities of the available combining vectors or the path loss corresponding to each AP. The former idea relies on the fact that a combining vector is obtained by quantizing a singular vector of the channel matrix, as shown in (15). Therefore, the quality of a combining vector can be evaluated based on its corresponding singular value. In contrast, the latter idea for D-ARFA is motivated by the observation that the AP with more significant path loss should have fewer activated RF chains because it is more likely to have a low sub-rate.

1) Singular values-based D-ARFA (SV-based D-ARFA): The SV-based D-ARFA scheme is summarized in Algorithm 3. Specifically, in step 1, each AP finds and sends the N largest singular values of the channel matrix to the CPU. Here, only the N largest singular values are sent because, in the proposed ARFA scheme, only n_l out of N combining vectors are selected for signal combining at the lth AP. As a result, the set of LN singular values $\left\{\lambda_1^{(1)}, \ldots, \lambda_N^{(1)}, \ldots, \lambda_1^{(L)}, \ldots, \lambda_N^{(L)}\right\}$ is available at the CPU, where $\lambda_n^{(l)}$ is the nth largest singular value associated with the lth AP. Then, the numbers of active RF chains at the APs are determined in steps 3–6. Specifically, an RF chain at an AP is suggested for activation if its corresponding singular value is not smaller than $\lambda_{L\bar{n}}$ found in step 2. In other words, the number of active RF chains at the lth AP, that is, $n_l^\star$, is set as the number of elements in $\left\{\lambda_1^{(l)}, \ldots, \lambda_N^{(l)}\right\}$ that are not smaller than $\lambda_{L\bar{n}}$. Finally, the CPU sends the value $n_l^\star$ back to the lth AP, which is then used for signal combining based on the D-HBF scheme.

2) Path loss-based D-ARFA (PL-based D-ARFA): In the SV-based D-ARFA scheme, the largest singular values of the channel matrices are required to find $\boldsymbol{n}^\star$. This entails high computational complexity, especially when L and N_r are

Algorithm 4 HBF with PL-based D-ARFA

```
Output: $\boldsymbol{F}^\star$
1:  Find $\boldsymbol{n} = \{n_1, \dots, n_L\}$ based on (18).
2:  Obtain $\{n_{[1]}, \dots, n_{[L]}\}$ s.t. $\alpha_{[1]} > \dots > \alpha_{[L]}$.
3:  $t = 1$
4:  while $\boldsymbol{n} \notin \mathcal{S}$ do
5:      if $\sum_{l=1}^{L} n_{[l]} < L\bar{n}$ and $n_{[t]} < N$ then
6:          $n_{[t]} = n_{[t]} + 1$
7:      end if
8:      if $\sum_{l=1}^{L} n_{[l]} > L\bar{n}$ and $n_{[L-t+1]} > 0$ then
9:          $n_{[L-t+1]} = n_{[L-t+1]} - 1$
10:     end if
11:     $t = t + 1$
12:     Reset $t = 1$ if $t > L$.
13: end while
14: Obtain $\boldsymbol{n}^\star$ by reordering $\{n_{[1]}, \dots, n_{[L]}\}$ to the original order.
15: for $l = 1 \to L$ do
16:     The CPU sends $n_l^\star$ to the $l$th AP.
17:     The $l$th AP determines its local analog combiner $\boldsymbol{F}_l^\star$ for $n_l^\star$ RF chains, i.e., $\boldsymbol{F}_l^\star = [\tilde{\boldsymbol{f}}_{l1}^\star, \dots, \tilde{\boldsymbol{f}}_{ln_l^\star}^\star]$, where $\tilde{\boldsymbol{f}}_{ln}^\star$ is given by (15) and determine $\boldsymbol{W}_l^\star$ based on (14).
18: end for
```

large. To avoid this, we herein propose the PL-based D-ARFA scheme, in which $\boldsymbol{n}^\star$ is obtained based on the total path losses associated with the APs. In CF massive MIMO systems, the APs are distributed in a large area. Therefore, the contribution of an AP to the total achievable rate considerably depends on its path loss.

Let $\beta_l = \sum_{k=1}^{K} \beta_{kl}$ be the sum of path loss of the lth AP, with β_{kl} given in (6), and let $\alpha_l = \frac{1}{\beta_l}, \forall l$. The number of activated RF chains at the lth AP can be set to

$$n_l = \min \left\{ N, \left\lfloor L\bar{n} \frac{\alpha_l}{\sum_{i=1}^{L} \alpha_i} \right\rceil \right\}, \forall l, \tag{18}$$

where $\min\{N, \cdot\}$ is used to guarantee $n_l \leq N$, and $\lfloor \cdot \rceil$ rounds a real number to its nearest integer. However, because of rounding, it is possible to obtain $\sum_{l=1}^{L} n_l \neq L\bar{n}$, which leads to $\boldsymbol{n} \notin \mathcal{S}$. To solve this problem, we propose Algorithm 4.

In Algorithm 4, the elements of $\boldsymbol{n}$ found in step 1 based on (18) are sorted in step 2 in decreasing order of $\{\alpha_1, \dots, \alpha_L\}$, i.e., in increasing order of the sums of path loss $\{\beta_1, \dots, \beta_L\}$, to generate $\{n_{[1]}, \dots, n_{[L]}\}$. Here, the order index $[t]$ indicates that $n_{[t]}$ RF chains are chosen to be activated at the AP associated with the tth-smallest path loss. Therefore, if $\sum_{l=1}^{L} n_{[l]} < L\bar{n}$ and $n_{[t]} < N$, $n_{[t]}$ is increased by one. In contrast, if $\sum_{l=1}^{L} n_{[l]} > L\bar{n}$ and $n_{[L-t+1]} > 0$, $n_{[L-t+1]}$ is decreased by one. This process is repeated until $\boldsymbol{n} \in \mathcal{S}$ is satisfied, as shown in steps 3–13. In this procedure, by initializing $t = 1$ and gradually increasing t, the numbers of turned-on RF chains for the APs with path loss are chosen to increase first, whereas those for the APs with larger path loss are chosen to decrease first. In step 14, $\{n_{[1]}, \dots, n_{[L]}\}$ are reordered into the original order. In steps 15–18, the numbers of active RF chains at the APs are determined, which are then fed back to the APs for SC-HBF, as in steps 3–6 of Algorithm 3.

Table I. Comparison of computational complexities.

Schemes	Complexities
SC-HBF	$\mathcal{C}_{\text{SC-HBF}} = \mathcal{O}(L(N\mathcal{C}_{\text{a}} + \mathcal{C}_{\text{u}} + \mathcal{C}_{\text{d}}))$
D-HBF	$\mathcal{C}_{\text{D-HBF}} = \mathcal{O}(L(NKN_tN_r + \mathcal{C}_{\text{d}}))$
SC-ARFA	$\mathcal{C}_{\text{SC-ARFA}} = \mathcal{O}((I_2+1)\mathcal{C}_{\text{SC-HBF}} + L\mathcal{C}_{\text{d}})$
SV-based D-ARFA	$\mathcal{C}_{\text{SV-D-ARFA}} = \mathcal{O}(L(NKN_tN_r + \mathcal{C}_{\text{d}}))$
PL-based D-ARFA	$\mathcal{C}_{\text{PL-D-ARFA}} = \mathcal{O}\left(L(KN_tN_r\bar{n} + \mathcal{C}_{\text{d}})\right)$
Exhaustive search	$> \mathcal{O}\left(LNKN_tN_r + p_L(L\bar{n})\mathcal{C}_{\text{d}}\right)$

D. Complexity Analysis

In this section, we analyze the complexity of the proposed HBF and ARFA schemes. Considering the SC-HBF scheme in Algorithm 1, to compute $\hat{\boldsymbol{H}}_l \boldsymbol{Q}_{l-1}^{-1} \hat{\boldsymbol{H}}_l^H \triangleq \boldsymbol{T}_l \in \mathbb{C}^{N_r \times N_r}$, a complexity of $\mathcal{O}(K^3N_t^3 + 2K^2N_t^2N_r)$ is required for matrix inversion and multiplications. In steps 3–6, N singular vectors of $\boldsymbol{T}_l$ can be obtained by performing the rank-N thin SVD of $\boldsymbol{T}_l$ with a complexity of $\mathcal{O}(NN_r^2)$ [50], [51]. As a result, the total complexity for analog precoding (in steps 3–7) is $\mathcal{C}_{\text{a}} = \mathcal{O}(K^3N_t^3 + 2K^2N_t^2N_r + NN_r^2)$. The updates in steps 8–10 have the complexity $\mathcal{C}_{\text{u}} = \mathcal{O}(NKN_tN_r + 2NK^2N_t^2 + (N^2+1)KN_t)$, and that to obtain the digital combiner $\boldsymbol{W}_l^\star$ in step 12 is $\mathcal{C}_{\text{d}} = \mathcal{O}(2K^2N_t^2N_r)$. From the above analysis, the total complexity of Algorithm 1 is $\mathcal{C}_{\text{SC-HBF}} = \mathcal{O}(L(N\mathcal{C}_{\text{a}} + \mathcal{C}_{\text{u}} + \mathcal{C}_{\text{d}}))$. In the D-HBF scheme, analog beamformers are obtained based on (15) by performing the rank-N thin SVD of $\hat{\boldsymbol{H}}_l, l = 1, \dots, L$ with a complexity of $\mathcal{O}(LNKN_tN_r)$. Thus, its total complexity is only $\mathcal{C}_{\text{D-HBF}} = \mathcal{O}(L(NKN_tN_r + \mathcal{C}_{\text{d}}))$.

Most of the complexity of Algorithm 2 originates from steps 2 and 14, which are to perform Algorithm 1. Its complexity can be estimated as $\mathcal{C}_{\text{SC-ARFA}} = \mathcal{O}((I_2+1)\mathcal{C}_{\text{SC-HBF}} + L\mathcal{C}_{\text{d}})$, where I_2 is the number of iterations performing steps 6–16. In Algorithm 3, the complexity of step 1 is $\mathcal{O}(LNKN_tN_r)$, which is to obtain N largest singular values/vectors of $\hat{\boldsymbol{H}}_l, l = 1, \dots, L$. We note that in step 5, $n_l^\star$ singular vectors are not computed; they are selected from those obtained in step 1. As a result, the complexity of Algorithm 3 is $\mathcal{C}_{\text{SV-D-ARFA}} = \mathcal{O}(L(NKN_tN_r + \mathcal{C}_{\text{d}}))$. Compared to the SV-based D-ARFA scheme, $n_l^\star$ in the PL-based D-ARFA scheme can be obtained without performing SVD, resulting in a much lower complexity. Specifically, only the rank-$n_l^\star$ thin SVD of $\hat{\boldsymbol{H}}_l$ is performed with a complexity of $\mathcal{O}(n_l^\star KN_tN_r)$. Thus, Algorithm 4 has the total complexity as $\mathcal{C}_{\text{PL-D-ARFA}} = \mathcal{O}(KN_tN_r \sum_{l=1}^{L} n_l^\star + L\mathcal{C}_{\text{d}}) = \mathcal{O}\left(L(KN_tN_r\bar{n} + \mathcal{C}_{\text{d}})\right)$, with the note that $\sum_{l=1}^{L} n_l^\star = L\bar{n}$.

The complexities of Algorithms 1–4 and the D-HBF schemes are summarized in Table I. In general, the decentralized (D-HBF and D-ARFA) schemes have much lower complexities compared to their semi-centralized counterparts (SC-HBF and SC-ARFA). Among the proposed ARFA schemes, the PL-based D-ARFA has the lowest complexity because $n_l^\star$ is obtained based only on the large-scale fading channels. To illustrate the complexity reduction of the proposed schemes, we now consider the case that $n_l^\star, \forall l$ are obtained with exhaustive search. The problem of searching for the best $n_l^\star, \forall l$ is equivalent to selecting the best partitions of $L\bar{n}$ into L parts $\{n_1, \dots, n_L\}$ with $n_l \in \{0, 1, \dots, N\}, \forall l$. Denote the number of possible partitions as $p_L(L\bar{n})$, which increases with L. Note that the rank-N thin SVD can be performed once and the resultant singular vectors are used for all permutations. The complexity of the exhaustive search scheme is lower bounded as $\mathcal{O}\left(LNKN_tN_r + p_L(L\bar{n})\mathcal{C}_{\text{d}}\right)$, which is prohibitive because L is large in CF MIMO systems.

V. Power consumption analysis

In the considered uplink CF mMIMO system, the total power consumption is modeled as [12], [13], [50], [52], [53]

$$P_{\text{total}} = \sum_{k=1}^{K} \left(P_{\text{TX},k} + P_{\text{UE},k}\right) + \sum_{l=1}^{L} \left(P_{\text{fix},l} + P_{\text{BF},l} + P_{\text{FH},l}\right), \tag{19}$$

where $P_{\text{TX},k}$ and $P_{\text{UE},k}$ represent the transmit power and the required power to run circuit components at the kth UE, respectively; $P_{\text{fix},l}$, $P_{\text{BF},l}$, and $P_{\text{FH},l}$ respectively denote the fixed power consumption term, the variable power consumption for the beamforming structure, and the fronthaul power consumption for the lth AP. $P_{\text{TX},k}$ is given as

$$P_{\text{TX},k} = \rho\sigma_n^2 \sum_{k=1}^{K} \frac{1}{\eta_k} \mathbb{E}\left\{\|\boldsymbol{x}_k\|^2\right\} = \sum_{k=1}^{K} \frac{\rho\sigma_n^2}{\eta_k}, \tag{20}$$

where $\eta_k \in (0, 1]$ denotes the power amplifier efficiency of the UE k, and the last equality is obtained by $\mathbb{E}\left\{\|\boldsymbol{x}_k\|^2\right\} = 1, \forall k$. In an HBF architecture, each antenna requires a low-noise amplifier (LNA) and two mixers, and each RF chain requires one ADC and N_r PSs, as illustrated in Fig. 1 [25], [54], [55]. Therefore, $P_{\text{BF},l}$ linearly depends on the numbers of antennas (N_r) and active RF chains at the lth AP (n_l) as follows:

$$P_{\text{BF},l} = N_r p_{\text{BF},1} + n_l p_{\text{BF},2}, \tag{21}$$

where $p_{\text{BF},1} = p_{\text{LNA}} + 2p_{\text{M}}$, $p_{\text{BF},2} = N_r p_{\text{PS}} + p_{\text{RF}} + p_{\text{ADC}}$, with p_{LNA}, p_{M}, p_{PS}, p_{RF}, and p_{ADC} respectively denoting the power consumed by an LNA, mixer, PS, RF chain, and ADC.$P_{\text{FH},l}$ can be obtained by [13], [52]

$$P_{\text{FH},l} = P_{\text{FH,max}} \frac{R_{\text{FH},l}}{C_{\text{FH},l}} = \kappa_l R_{\text{FH},l}, \tag{22}$$

where $P_{\text{FH,max}}$ is the maximum power required for the fronthaul traffic at the full capacity $C_{\text{FH},l}$, $R_{\text{FH},l}$ is the actual fronthaul rate between the lth AP and the CPU, and $\kappa_l = \frac{P_{\text{FH,max}}}{C_{\text{FH},l}}$. In the considered decentralized signal processing scheme, $2KN_t\tau_d\alpha_l$ bits are required to quantize the signal vector $\boldsymbol{r}_l \in \mathbb{C}^{KN_t \times 1}$ during each coherence interval [13], [56] at the lth AP before being sent to the CPU. Here, α_l is the number of quantization bits at the lth AP, and τ_d is the length (in symbols) of the uplink data. As a result, $R_{\text{FH},l}$ is given by [13], [56]

$$R_{\text{FH},l} = \frac{2KN_t\tau_d\alpha_l}{T_c}, \tag{23}$$

where T_c is the coherence time (in seconds). Assume that all the UEs have the same power amplifier efficiency and circuit power consumption, i.e., $\eta_k = \eta$, $P_{\text{UE},k} = P_{\text{UE}}, \forall k$, and that all APs have the same fixed power consumption, number of quantization bits, and capacity, i.e., $P_{\text{fix},l} = P_{\text{fix}}$, $\alpha_l = \alpha$, $C_{\text{FH},l} = C_{\text{FH}}$, $\kappa_l = \kappa$, $\forall l$. Then, we have $P_{\text{FH},l} = P_{\text{FH}}$ and $R_{\text{FH},l} = R_{\text{FH}}$, $\forall l$. Furthermore, we note that AP l requires P_{fix}, even when it is in sleep mode; in contrast, P_{FH} and $P_{\text{BF},l}$ are only consumed when it is in the active mode. Let $\mathbb{A}$ be the set of APs in active mode and $|\mathbb{A}|$ be the number of active APs. Then, from (19)–(23), the total power consumption can be expressed as

$$\begin{aligned} P_{\text{total}} &= \frac{K\rho\sigma_n^2}{\eta} + KP_{\text{UE}} + LP_{\text{fix}} + |\mathbb{A}|\, P_{\text{FH}} \\ &\qquad + \sum_{l\in\mathbb{A}} \left(N_r p_{\text{BF},1} + n_l p_{\text{BF},2}\right), \\ &= P_0 + |\mathbb{A}|\, P_{\text{FH}} + |\mathbb{A}|\, N_r p_{\text{BF},1} + p_{\text{BF},2} \sum_{l\in\mathbb{A}} n_l, \end{aligned} \tag{24}$$

where $P_0 = \frac{K\rho\sigma_n^2}{\eta} + KP_{\text{UE}} + LP_{\text{fix}}$, a fixed term in P_{total}, for simple exposition.

It is observed from (24) that P_{total} varies depending on the number of active APs, i.e., $|\mathbb{A}|$; the total number of turned on RF chains, i.e., $\sum_{l\in\mathbb{A}} n_l$; and the number of antennas N_r. More specifically, it is a linearly increasing function of these factors. Therefore, P_{total} can be minimized by using only a subset of APs in the APS scheme [12], using a reduced number of antennas in the AS scheme [29], or optimizing both $\sum_{l\in\mathbb{A}} n_l$ and $|\mathbb{A}|$ in the proposed ARFA scheme. Next, we compare these schemes in terms of the total power consumption. Furthermore, conventional fixed-activation HBF schemes are also considered as benchmarks.

- *ARFA scheme:* When the ARFA scheme is employed, n_l is different among the APs; however, the total number of RF chains is fixed to $L\bar{n}$, i.e., $\sum_{l\in\mathbb{A}} n_l = L\bar{n}$. By inserting this into (24), we obtain

$$P_{\text{total}}^{\text{ARFA}} = P_0 + |\mathbb{A}|\, P_{\text{FH}} + |\mathbb{A}|\, N_r p_{\text{BF},1} + L\bar{n} p_{\text{BF},2}, \tag{25}$$

where $\mathbb{A}$ contains only the APs with at least one activated RF chain. Therefore, we have $|\mathbb{A}| = \sum_{l\in\mathbb{A}} \delta_l$ with $\delta_l = 1$ if $n_l > 0$, and $\delta_l = 0$ if $n_l = 0$. We note that the proposed ARFA algorithms have different operations, which can result in different $\mathbb{A}$. Therefore, they can have different power consumption.

- *Fixed-activation HBF:* We refer to the SC-HBF and D-HBF without the ARFA as the *fixed-activation HBF*. In this scheme, the same number of RF chains are activated at all L APs. For comparison with the proposed ARFA schemes, we consider two deployments: $n_l = N, \forall l$ and $n_l = \bar{n}, \forall l$. We note that with fixed activation HBF, all the APs are in active mode because they have a fixed nonzero number of RF chains for signal processing, i.e., $|\mathbb{A}| = L$. By inserting $n_l = N, \forall l$, and $n_l = \bar{n}, \forall l$ into (24), we obtain

$$P_{\text{total}}^{\text{fix},N} = P_0 + LP_{\text{FH}} + LN_r p_{\text{BF},1} + LN p_{\text{BF},2}, \tag{26}$$

$$P_{\text{total}}^{\text{fix},\bar{n}} = P_0 + LP_{\text{FH}} + LN_r p_{\text{BF},1} + L\bar{n} p_{\text{BF},2}. \tag{27}$$

- *APS scheme:* In this scheme, only a subset of the APs is selected based on received power [12], whereas the others are put into the sleep mode. For comparison with the proposed schemes in mmWave systems, we assume the conventional deployment of RF chains at the APs, i.e., each AP is equipped with N RF chains, all of which are used for analog signal combining, i.e., $n_l = N, \forall l$. For a fair comparison, the number of APs in active mode in this scheme is assumed to be $\frac{L\bar{n}}{N}$. This guarantees that a total of $L\bar{n}$ RF chains are used at the selected APs, which is equal to the number of activated RF chains in the proposed ARFA scheme and fixed-activation HBF scheme with $n_l = \bar{n}, \forall l$. By inserting $n_l = N, \forall l$, and $|\mathbb{A}| = \frac{L\bar{n}}{N}$ into (24), we have

$$P_{\text{total}}^{\text{APS}} = P_0 + \frac{L\bar{n}}{N} P_{\text{FH}} + \frac{L\bar{n}}{N} N_r p_{\text{BF},1} + L\bar{n} p_{\text{BF},2}. \tag{28}$$

- *AS scheme:* In the AS scheme, at each AP, only N_r^{AS} of N_r antennas are activated, corresponding to N_r^{AS} received signals put through the digital signal combining [29]. In other words, analog signal combining is conducted by a network of N_r^{AS} switches rather than NN_r PSs, in contrast to the other compared schemes. Therefore, at each AP, N_r switches are required, whereas the numbers of antennas, RF chains, and ADCs are the same and as small as N_r^{AS}, and the number of mixers is $2N_r^{\text{AS}}$. Furthermore, in the AS scheme, all the APs are in the active mode, i.e., $|\mathbb{A}| = L$. Let p_{SW} be the power consumed by a switch. The total power consumption in this scheme is given as

$$P_{\text{total}}^{\text{AS}} = P_0 + LP_{\text{FH}} + LN_r p_{\text{SW}} + LN_r^{\text{AS}} (p_{\text{RF}} + p_{\text{ADC}} + p_{\text{BF},1}). \tag{29}$$

By comparing (25) to (26)–(28), we observe that:

- The proposed ARFA scheme requires no higher power consumption than the fixed-activation HBF schemes with $n_l = N$ and $n_l = \bar{n}, \forall l$, because $\bar{n} < N$ and $|\mathbb{A}| \leq L$. Furthermore, we note that a dominant part of the power consumed for beamforming is created by the RF chains and ADC. Therefore, from (25) and (26), it is clear that a considerable reduction in power consumption can be obtained by the ARFA if $\bar{n} \ll N$ is chosen.
- Both the power consumption and total achievable rate of the proposed ARFA scheme significantly depend on $\bar{n}$. Specifically, a smaller $\bar{n}$ leads to a reduction in both power consumption and total achievable rate with respect to the fixed-activation HBF scheme with $n_l = N, \forall l$. This tradeoff is discussed further in the next section.
- It is observed from (28) and (25) that the APS and proposed ARFA schemes have a difference of $\left|\frac{L\bar{n}}{N} - |\mathbb{A}|\right| (P_{\text{FH}} + N_r p_{\text{BF},1})$ in power consumption, even though they have the same total number of active RF chains. Specifically, the APS scheme requires slightly lower power consumption, but its achievable rate is much lower than that of the ARFA scheme, as is shown in the next section. It is not certain from (29) and (25) which of AS and ARFA schemes has the lower power consumption, which will be determined based on the simulation results in the next section.

Furthermore, compared to the dynamic HBF architectures with switches introduced in [57], our proposed ARFA scheme requires no switches. Thus, it has no additional power consumption and switch delays. In particular, it is worth noting that the introduction of the switching network in [57] only offers the dynamic connections between the RF chain/phase shifters and antennas; it does not provide the dynamic RF chain activation as does our proposed scheme. As a result, the scheme in [57] provides a good achievable rate because all the RF chains are activated, but its total power consumption would be high if applied in CF mmWave massive MIMO systems, where numerous APs are deployed. These are also the key differences between our proposed scheme and the existing switch-based HBF architectures [58]–[60].

Table II. Simulation parameters [12], [13], [25], [52]

Parameters	Values
Power amplifier efficiency	$\eta = 0.3$
Coherent time and data length	$T_c = 2$ ms, $\tau_d = 180$ symbols
No. of quantization bits	$\alpha = 2$ bits
UE and fixed power term	$P_{\text{UE}} = 1$ W, $P_{\text{fix}} = 0.825$ W
Fronthaul capacity	$C_{\text{FH}} = 100$ Mbps
Component power	$p_{\text{LNA}} = 20$ mW, $p_{\text{ADC}} = 200$ mW, $p_{\text{RF}} = 40$ mW, $p_{\text{PS}} = 30$ mW, $p_{\text{M}} = 0.3$ mW, $p_{\text{SW}} = 5$ mW, $P_{\text{FH,max}} = 50$ W

VI. Simulation results

A. Simulation parameters

Simulations are performed to evaluate the total achievable rates, power consumption, EEs, and computational complexities of the proposed SC-HBF, D-HBF, and ARFA schemes. In simulations, K UEs and L APs are uniformly distributed at random within a square coverage area of size $D \times D$ m^2, where D is set to 1000 m [5]. The large-scale fading coefficients are computed based on (6) with $\vartheta = 4.1$, $\xi = 7.6$, and the antenna gain is set to $G_{\text{a}} = 15$ dBi [47]. Furthermore, we assume $f_c = 28$ GHz, $B = 100$ MHz, and NF $= 9$ dB for the carrier frequency, system bandwidth, and noise figure, respectively. As a result, the noise power is given as $\sigma_n^2 = -174$ dBm/Hz $+ 10\log_{10}(B) + \text{NF}$.

The channel coefficients between each UE and AP are generated based on the geometric Saleh–Valenzuela channel model given in (5). For simplicity, we assume an identical number of effective channel paths between each UE and AP, which is set to $P_{kl} = 3, \forall l, k$ [1], [45], [61], reflecting the limited scattering in mmWave channels. The AoDs and AoAs are uniformly distributed in $\left[-\frac{\pi}{6}, \frac{\pi}{6}\right]$ and $\left[-\frac{\pi}{12}, \frac{\pi}{12}\right]$, respectively. The ULA model is employed for the antenna arrays at the APs and UEs with antenna spacing of half a wavelength, i.e., $\frac{d_s}{\lambda} = \frac{1}{2}$ [40], [57]. The phases in the analog combiner are selected from $\Theta = \left\{0, \frac{2\pi}{2^b}, \frac{4\pi}{2^b}, \ldots, \frac{2(2^b-1)\pi}{2^b}\right\}$, where $b = 4$ is set, implying 4-bit quantization of the PSs. The parameters in Table II are assumed to compute the total power consumption [12], [13], [25], [52], [62]. Furthermore, in the simulations, the variance of the CSI error is set to $\varepsilon^2 = \epsilon \times \bar{g}$, where $\epsilon \in \{0.1, 0.01, 0.001, 0\}$, and $\bar{g} = \frac{G_a}{KL} \sum_l^L \sum_{k=1}^K \frac{1}{\beta_{kl}}$ is the average large-scale fading coefficient.

B. Performance of the C-HBF and SC-HBF schemes

We numerically evaluate the total achievable rates of the C-HBF and SC-HBF schemes, which are analyzed in Section III. We assume the conventional RF chain deployment in this section, i.e., all the N available RF chains are active for analog combining. For comparison, we consider the beam steering scheme, in which the PSs are used to optimally orient the array response in space [22], [44]. In other words, the phases of the analog combining coefficients are obtained by quantizing

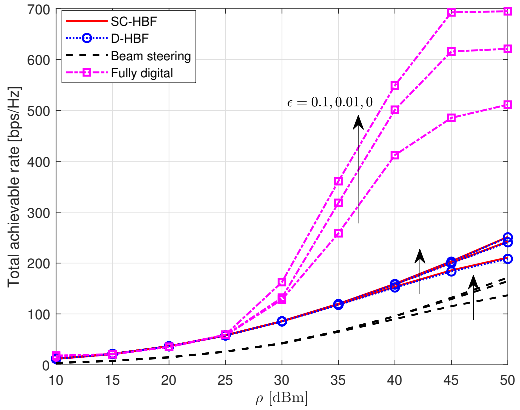


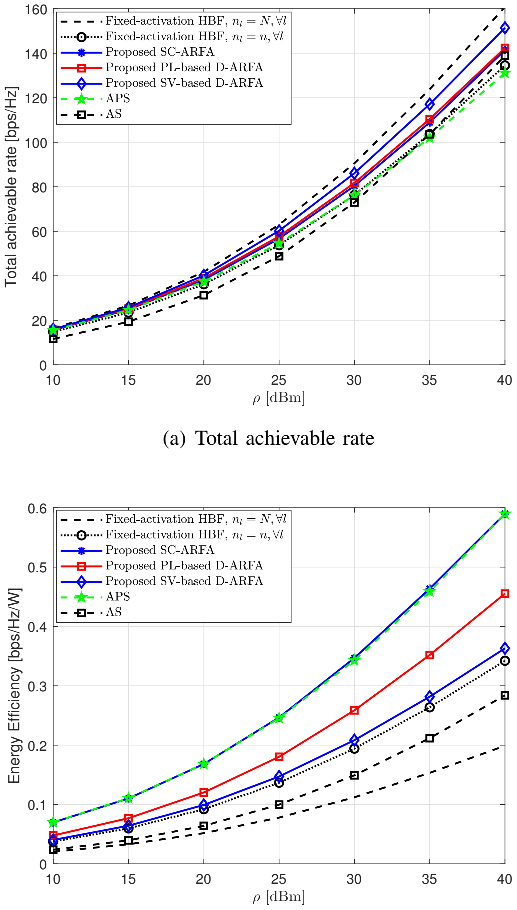
Fig. 2. Total achievable rates of the C-HBF and SC-HBF schemes compared to those of the beam steering scheme with $L = 32$, $K = 8$, $N_t = 4$, $N_r = 64$, $N = 8$, and $\epsilon = \{0.5, 0.1, 0.01, 0\}$.

those of the array response vectors that match best with the dominant eigenmode of the channel matrix. We also show the performance of the digital beamformer, which is the optimal eigen beamforming scheme.

Fig. 2 shows the total achievable rates of the SC-HBF, D-HBF, beam steering, and fully digital schemes with $N_r = 64$, $N_t = 4$, $N = 8$, $L = 32$, $K = 8$ [22], and $\epsilon = \{0.1, 0.01, 0\}$. It is clearly seen in Fig. 2 that the D-HBF and SC-HBF schemes have almost the same total achievable rates because they employ the same digital beamformer and their analog beamformers perform approximately the same, as stated in Theorem 2. However, as ϵ increases, D-HBF is slightly outperformed by SC-HBF. This is caused by the centralized generation of the analog beamformer in SC-HBF, which utilizes the global CSI. It is also observed in Fig. 2 that the proposed SC-HBF and D-HBF schemes outperform the beam steering approach since beam steering alone cannot perfectly capture the channel's dominant eigenmodes [44]. However, the hybrid beamformers are outperformed by the fully digital one for both cases of perfect and imperfect CSI. This is obvious because in the fully digital beamforming architecture, $N_r (\gg N)$ RF chains are used to fully exploit the channel eigen modes.

C. Performance of the proposed ARFA scheme

The total achievable rates, power consumption, and EEs of the proposed ARFA schemes, namely, SC-ARFA, PL-based D-ARFA, and SV-based D-ARFA, are compared to those of the fixed-activation HBF with $n_l = N$ and $n_l = \bar{n}, \forall l$, APS, and AS schemes discussed in Section V. In our simulations, SC-HBF is used for the fixed-activation HBF and APS schemes. We note that the SC-HBF and D-HBF provides almost identical performance, as shown in Fig. 2, and for the same RF chain deployment, they have the same power consumption. We consider a CF mmWave massive MIMO system with $L = 32$, $K = 8$, $N_r = 64$, $N_t = 4$, $N = 8$ [6], [22], $\bar{n} = 2$, and $\epsilon = 0.01$. In the AS scheme, the number of selected antennas at each AP is set to $N_r^{\text{AS}} = 32$, which ensures that the AS scheme achieves comparable total achievable rates with respect to the proposed schemes, allowing us to compare them in terms of EE. In the simulations, the power consumption of the fixed-activation HBF schemes with $n_l = N$, $n_l = \bar{n}$, the APS, and AS scheme is computed based on (26)–(29), whereas that of the proposed ARFA schemes is obtained through simulations because it depends on δ_{n_l}, as indicated in (25). The EE of a scheme is calculated as the ratio between the total achievable rate and the total power consumption.



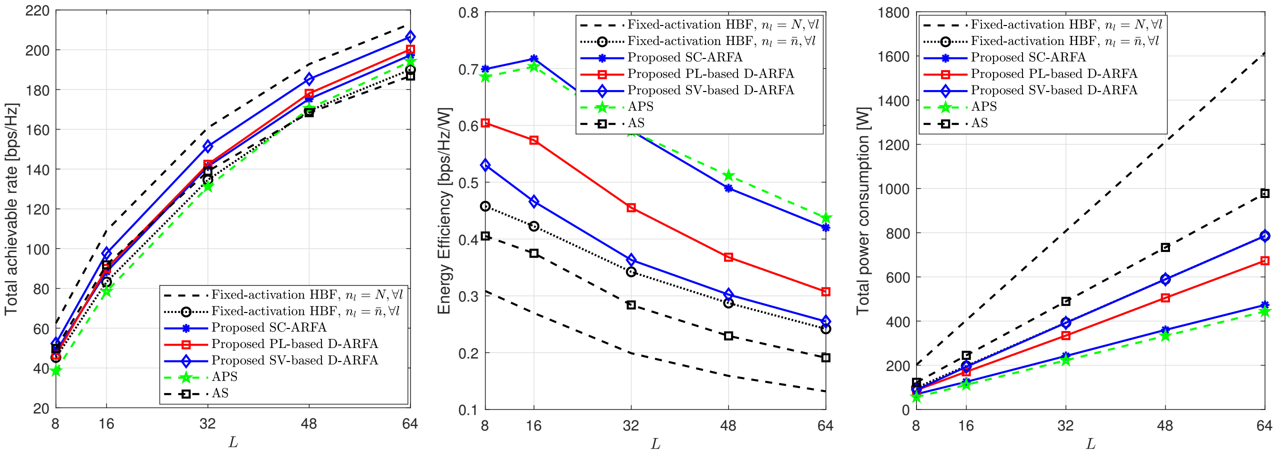
(a) Total achievable rate

(b) Energy efficiency

Fig. 3. Total achievable rates and EEs of the proposed ARFA schemes compared to those of the fixed-activation HBF with $n_l = N$, $n_l = \bar{n}, \forall l$, APS, and AS schemes. Simulation parameters are $L = 32$, $K = 8$, $N_t = 4$, $N_r = 64$, $N = 8$, and $\bar{n} = 2$.

In Fig. 3, we show the total achievable rates and EEs of the considered schemes versus the average transmit power ρ for $L = 32$, $K = 8$, $N_t = 4$, $N_r = 64$, $N = 8$, and $\bar{n} = 2$. From Fig. 3, the following observations are noted:

- As shown in Fig. 3(a), the fixed-activation HBF scheme with $n_l = N$ achieves the highest total achievable rate because it activates all the available APs and RF chains. However, in this scheme, power consumption by the LN RF chains is high. Therefore, its EE is significantly lower than those of the other considered schemes, wherein only $L\bar{n}$ $(\ll LN)$ RF chains are turned on, as shown in Fig. 3(b). Despite the reduced number of active RF chains, the proposed ARFA schemes perform close to the fixed-activation HBF with $n_l = N$.
- Among the proposed ARFA schemes, the SV-based D-ARFA achieves the highest achievable rate, but it is outperformed by the SC-ARFA and PL-based ARFA in terms of EE. However, all these proposed schemes



(a) Total achievable rate

(b) EE

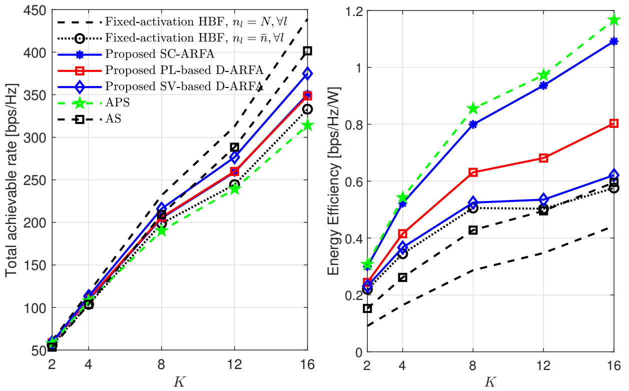
(c) Total power consumption

Fig. 4. Total achievable rates, EEs, and power consumption of the proposed ARFA schemes compared to those of the fixed-activation HBF with $n_l = N$, $n_l = \bar{n}, \forall l$, APS, and AS schemes. Simulation parameters are $L = \{8, 16, 32, 48, 64\}$, $K = 8$, $N_t = 4$, $N_r = 64$, $N = 8$, $\bar{n} = 2$, and $\rho = 50$ dBm.

achieve remarkable improvement in EE with a marginal rate loss with respect to the fixed-activation HBF scheme with $n_l = N$. For example, at $\rho = 40$ dBm, the SV-based D-ARFA exhibits a performance loss of approximately 5.3%, whereas an approximately 85.2% improvement in EE is attained compared to that obtained by the fixed-activation HBF scheme with $n_l = N$.

- It is also observed that the proposed ARFA schemes outperform the APS and AS schemes in terms of the total achievable rate. The fixed-activation HBF scheme with $n_l = \bar{n}$ is outperformed by the proposed ARFA schemes in both achievable rate and EE. It is also clear that the AS scheme is not an energy-efficient scheme for the CF mmWave massive MIMO system.

In Fig. 4, we show the total achievable rates, EEs, and power consumption of the considered schemes versus the number of APs. In this figure, we use the same simulation parameters as in Fig. 3, except for the varying numbers of APs, i.e., $L = \{8, 16, 32, 48, 64\}$, and $\rho = 40$ dBm. In Figs. 4(a) and 4(b), the observations on the achievable rates and EEs of the considered schemes are similar to those from Fig. 3. In particular, it is seen that in the entire range of L, the proposed ARFA schemes have small losses in total achievable rate but significant improvement in EE with respect to the fixed-activation HBF scheme with $n_l = N$. Furthermore, the proposed ARFA schemes perform better than or comparable to the APS scheme in terms of both achievable rate and EE. In particular, it is clear that the AS and fixed-activation schemes with $n_l = \bar{n}$ are less efficient in both the spectral and energy compared to the proposed schemes. To further explain the EEs, we consider the total power consumption of these schemes in Fig. 4(c). It can be seen that the total power consumption of the fixed-activation schemes quickly increases with L. Therefore, activating all the N RF chains at all the APs causes an extremely high power consumption for the CF mmWave massive MIMO system, motivating the ARFA in this work. Among the other schemes, the AS scheme consumes the highest power while achieving the lowest rates, making it energy-inefficient, as seen in Fig. 4(b). The proposed ARFA and APS schemes have comparable and low power consumption.



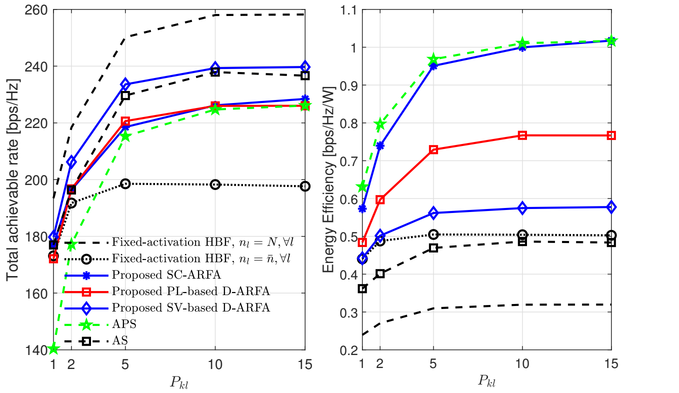
(a) Total achievable rate

(b) Energy Efficiency

Fig. 5. Performance of the proposed ARFA schemes with $K = [2, 16]$, $L = 32$, $N_t = 4$, $N_r = 64$, $N = 8$, $\bar{n} = 2$, and $\rho = 50$ dBm.

In Fig. 5, we indicate the total achievable rate and EE of the proposed ARFA schemes with $K = \{2, 4, 8, 12, 16\}$, $L = 32$. The other simulation parameters are the same as those in Fig. 4. Because the total achievable rate of the uplink system is not affected by inter-user interference, those of all the considered schemes significantly increase with K, as seen in Fig. 5(a). However, when K increases, the total power consumption also increases. Thus, the improvement in the EE (Fig. 5(b)) is less significant than that in the total rate. In particular, it indicates that for large K the AS scheme performs better than the ARFA and APS schemes in terms of rate. This is because with more received data streams, the chance that the optimal antennas are selected increases. However, it is still outperformed by the ARFA and APS schemes in terms of EE. Similar to the observations in previous figures, the fixed-activation HBF with $n_l = N$ provides the best achievable rate but the worst EE.

We demonstrate the total achievable rate and EE of the proposed ARFA schemes compared to those of the conventional schemes for different numbers of effective paths, i.e., $P_{kl} = \{1, 2, 5, 10, 15\}$ in Fig. 6 [44]. The simulation parameters are set the same as those in Fig. 5 with $K = 8$. It is seen that as P_{kl} increases, all the considered schemes benefit



(a) Total achievable rate (b) Energy Efficiency

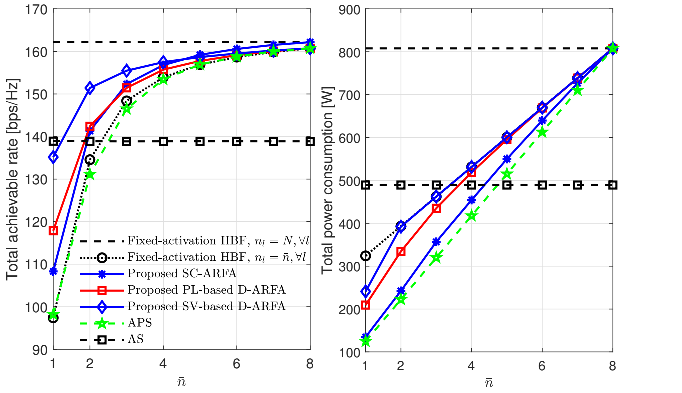
Fig. 6. Performance of the proposed ARFA schemes with $P_{kl} = [1, 20]$, $K = 8$, $L = 32$, $N_t = 4$, $N_r = 64$, $N = 8$, $\bar{n} = 2$, and $\rho = 50$ dBm.

from muti-path propagation. However, when P_{kl} becomes sufficiently large, the performance improvement increases slowly or remains unchanged with P_{kl}. This is reasonable because the channel ranks are limited by $\min\{N_r, N_t\} = 4$, and for $P_{kl} > 4$, increasing P_{kl} no longer results in channel rank enhancement. In particular, the fixed activation HBF scheme with $n_l = \bar{n}_l, \forall l$ has the worst performance at large P_{kl} because in this scheme, the ranks of the effective channels, i.e., $\boldsymbol{F}_l^H \boldsymbol{H}_l$, are limited by $\bar{n}_l = 2, \forall l$. However, this phenomenon is not seen for the proposed ARFA schemes even though they have the same number of active RF chains on average, i.e., $\bar{n}_l$. This is because the ARFA schemes enable an adaptive activation to select the best RF chains to exploit the multipath fading as well as the rank enhancement.

D. Tradeoff between achievable rates and power consumption

The total achievable rate and power consumption of the considered schemes versus $\bar{n}$ are evaluated numerically in Fig. 7 for $L = 32$, $K = 8$, $N_r = 64$, $N_t = 4$, $N = 8$, $\bar{n} = \{1, 2, \ldots, 8\}$, and $\rho = 40$ dBm. From Fig. 7, the following observations can be noted:

- The total achievable rate and power consumption of the fixed-activation HBF with $n_l = N$ and those of the AS scheme remain unchanged with $\bar{n}$ because N and N_r^{AS} RF chains, respectively, are always active at every AP. In contrast, those of the other schemes depend on $\bar{n}$. Specifically, as $\bar{n}$ increases, both the total achievable rate and power consumption of the fixed-activation HBF scheme with $n_l = \bar{n}$, the proposed ARFA, and the APS schemes increase to approach those of the fixed-activation HBF with $n_l = N$.
- In Fig. 7(a), the proposed ARFA schemes perform closest to the fixed-activation HBF scheme with $n_l = N$, even for a small $\bar{n}$. In terms of power consumption, they require slightly higher power than the APS scheme. However, their EEs are comparable, as shown in Figs. 3 and 4, owing to the efficient use of RF chains. Furthermore, for $\bar{n} \geq 4$, the AS scheme has the lowest power consumption at the cost of the smallest total achievable rate.
- For the optimal performance–power consumption tradeoff in the assumed environment, $\bar{n} \in [2, 4]$ can be chosen in the proposed ARFA schemes to achieve 34.2–69.8% power reduction with marginal performance loss. In particular, for $\bar{n} = 4$, the performance loss is only $2.7-3.6\%$. With $\bar{n} = 1$, only a single RF chain on average is turned on at each AP, and a significant loss in the achievable rate is observed for the proposed ARFA with respect to the fixed-activation HBF with $n_l = N$.



(a) Total achievable rate (b) Power consumption

Fig. 7. Total achievable rates and EEs of the proposed ARFA schemes compared to those of the fixed-activation HBF with $n_l = N$, $n_l = \bar{n}, \forall l$, and APS schemes. Simulation parameters are $L = 32$, $K = 8$, $N_t = 4$, $N_r = 64$, $N = 8$, $\bar{n} = \{1, 2, \ldots, 8\}$, and $\rho = 50$ dBm.

E. Analysis of fronthauling and computations loads

In this section, we evaluate the amount of information exchange between the CPU and APs, which is presented in Table III. In the SC-HBF scheme, the CSI for $\hat{\boldsymbol{H}}_l$ of size $N_r \times KN_t$ is sent from the lth AP to the CPU, which is used at the CPU to generate $\boldsymbol{F}_l$ of size $N_r \times N$. However, we note that all the entries of $\boldsymbol{F}_l$ have constant amplitudes of $\frac{1}{\sqrt{N_r}}$. Therefore, only $N_r N$ real numbers representing the phases are fed back on the reverse link. A similar analysis is valid for SC-ARFA with the note that only an average of $N_r \bar{n}$ real numbers need to be fed back from the APs to the CPU because in these schemes, only an average of $\bar{n}$ RF chains are activated. It is observed that the amount of information exchange between the CPU and APs is relatively large in the SC-HBF and SC-ARFA schemes.

In contrast, those for the decentralized schemes are small. Specifically, in the D-HBF schemes, only KN_t complex numbers representing the estimate of the transmitted signal, i.e., $\boldsymbol{r}_l$, are sent to the CPU for the final soft estimation. In the SV-based D-ARFA scheme, an addition of N real numbers for the N singular values are sent from an AP to the CPU for each channel variation to perform ARFA. In contrast, the transmission of path loss values in the PL-based D-ARFA scheme can be ignored because of their slow variations. On the reverse link from the CPU to an AP, only a single real number, which is the number of active RF chains, is fed back in the SV- and PL-based D-ARFA schemes, as demonstrated in Algorithms 3 and 4, respectively. Given that $N \ll N_r$, the decentralized schemes require much less information exchange between the CPU and each AP compared to the semi-centralized schemes.

In Table III, we also show the run-time complexities of the proposed schemes. The results are obtained for $K = 8$, $L = 32$, $N_t = 4$, $N_r = 64$, $N = 8$, $\bar{n} = 2$, and $\rho = 50$

Table III. Comparison of the decentralized and semi-centralized schemes in terms of fronthauling load and run time.

Schemes	Fronthauling load: AP to CPU	Fronthauling load: CPU to AP	Run-time (s)
SC-HBF	$N_r K N_t$ complex numbers	$N_r N$ real numbers	0.801
SC-ARFA	$N_r K N_t$ complex numbers	$N_r \bar{n}$ real numbers	10.142
D-HBF	$K N_t$ complex numbers	0	0.692
SV-based D-ARFA	$K N_t$ complex numbers and N real numbers	1 real number	0.766
PL-based D-ARFA	$K N_t$ complex numbers	1 real number	0.792

dBm. It is observed that the decentralized HBF and ARFA schemes have the lowest run time, whereas that of the SC-ARFA scheme is the highest and approximately 12 times higher than the other schemes. The SV- and PL-based D-ARFA schemes have slightly longer run times than the D-HBF owing to the optimization of $\boldsymbol{n}^{\star}$. The complexities in terms of run time in this table align well with those in Table I for the number of operations.

VII. Conclusion

In this work, we propose two HBF schemes for CF mmWave massive MIMO systems, including SC-HBF and D-HBF, in which the analog combiners are generated at the CPU based on the global CSI and at each AP based on the local CSI, respectively. Notably, although the D-HBF requires substantially lower computational complexity and no information exchange between the CPU and APs, it achieves approximately the same total achievable rate as that obtained by the SC-HBF scheme. Furthermore, to reduce the power consumption in the CF mmWave massive MIMO system, we propose adaptive activation of RF chains at the APs. Low-complexity algorithms are developed to select the number of active RF chains at the APs such that the system's power consumption is significantly reduced with only a marginal loss in the total achievable rate. The efficiency of the proposed schemes is justified by the simulation results, which show that the proposed ARFA scheme achieves significant improvement in EE while leading to a loss of only small loss in total achievable rate. For future studies, the optimization of the proposed analog combination schemes for wideband systems will be considered. Furthermore, the proposed ARFA scheme can be incorporated with low-resolution ADCs [63], [64] to further reduce the power consumption.

Appendix A
Proof of Theorem 1

Let $\boldsymbol{Q} = \boldsymbol{I}_{LN} + \gamma \boldsymbol{F}^H \hat{\boldsymbol{H}} \hat{\boldsymbol{H}}^H \boldsymbol{F}$. Because $\boldsymbol{F}$ is a block-diagonal matrix, we have

$$\hat{\boldsymbol{H}}^H \boldsymbol{F} = \left[\hat{\boldsymbol{H}}_1^H \boldsymbol{F}_1, \hat{\boldsymbol{H}}_2^H \boldsymbol{F}_2, \ldots, \hat{\boldsymbol{H}}_L^H \boldsymbol{F}_L\right],$$

leading to $\hat{\boldsymbol{H}}^H \boldsymbol{F}\boldsymbol{F}^H \hat{\boldsymbol{H}} = \sum_{l=1}^{L} \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l$. Therefore, $\boldsymbol{Q}$ can be expressed as $\boldsymbol{Q} = \boldsymbol{I}_{KN_t} + \gamma \sum_{l=1}^{L} \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l$. By letting $\boldsymbol{G}_l = \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l$, $\boldsymbol{Q}$ can be further expanded as

$$\begin{aligned}
\boldsymbol{Q} &= \underbrace{\boldsymbol{I}_{KN_t} + \gamma \boldsymbol{G}_1}_{\triangleq \boldsymbol{E}_1} + \gamma \boldsymbol{G}_2 + \ldots + \gamma \boldsymbol{G}_L \\
&= \boldsymbol{E}_1 (\underbrace{\boldsymbol{I}_{KN_t} + \gamma \boldsymbol{E}_1^{-1} \boldsymbol{G}_2}_{\triangleq \boldsymbol{E}_2} + \ldots + \gamma \boldsymbol{E}_1^{-1} \boldsymbol{G}_L) \\
&= \boldsymbol{E}_1 \boldsymbol{E}_2 (\underbrace{\boldsymbol{I}_{KN_t} + \gamma \boldsymbol{E}_2^{-1} \boldsymbol{E}_1^{-1} \boldsymbol{G}_2}_{\triangleq \boldsymbol{E}_3} + \ldots + \gamma \boldsymbol{E}_2^{-1} \boldsymbol{E}_1^{-1} \boldsymbol{G}_L) = \ldots \\
&= \boldsymbol{E}_1 \boldsymbol{E}_2 \ldots \boldsymbol{E}_L, \qquad (30)
\end{aligned}$$

where $\boldsymbol{E}_l = \boldsymbol{I}_{KN_t} + \gamma (\boldsymbol{E}_1 \ldots \boldsymbol{E}_{l-1})^{-1} \boldsymbol{G}_l$, $l = 2, 3, \ldots, L$. As a result, R^{a} can be expressed as

$$\begin{aligned}
R^{\text{a}} &= \log_2 \det \boldsymbol{Q} = \sum_{l=1}^{L} \log_2 \det(\boldsymbol{E}_l) \\
&= \sum_{l=1}^{L} \log_2 \det(\boldsymbol{I}_{KN_t} + \gamma (\underbrace{\boldsymbol{E}_1 \ldots \boldsymbol{E}_{l-1}}_{\triangleq \boldsymbol{Q}_{l-1}})^{-1} \boldsymbol{G}_l) \qquad (31) \\
&= \sum_{l=1}^{L} \log_2 \det \left(\boldsymbol{I}_{KN_t} + \gamma \boldsymbol{Q}_{l-1}^{-1} \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l\right) \\
&= \sum_{l=1}^{L} \log_2 \det \left(\boldsymbol{I}_N + \gamma \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l \boldsymbol{Q}_{l-1}^{-1} \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l\right), \qquad (32)
\end{aligned}$$

as given in Theorem 1. The last equality in (32) follows from $\det(\boldsymbol{I}_{KN_t} + \boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{I}_N + \boldsymbol{B}\boldsymbol{A})$ with $\boldsymbol{A} = \boldsymbol{Q}_{l-1}^{-1} \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l \in \mathbb{C}^{KN_t \times N}$ and $\boldsymbol{B} = \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l \in \mathbb{C}^{N \times KN_t}$. Furthermore, from the definition of $\boldsymbol{Q}_{l-1}$ in (31), we have $\boldsymbol{E}_{l-1} = \boldsymbol{I}_{KN_t} + \gamma (\boldsymbol{E}_1 \ldots \boldsymbol{E}_{l-2})^{-1} \boldsymbol{G}_{l-1} = \boldsymbol{I}_{KN_t} + \gamma \boldsymbol{Q}_{l-2}^{-1} \boldsymbol{G}_l$. Finally, recalling that $\boldsymbol{G}_l = \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l$, we obtain the expression of $\boldsymbol{Q}_{l-1}$ in (11), i.e.,

$$\begin{aligned}
\boldsymbol{Q}_{l-1} &= \boldsymbol{Q}_{l-2} \boldsymbol{E}_{l-1} = \boldsymbol{Q}_{l-2} (\boldsymbol{I}_{KN_t} + \gamma \boldsymbol{Q}_{l-2}^{-1} \boldsymbol{G}_l) \\
&= \boldsymbol{Q}_{l-2} + \gamma \hat{\boldsymbol{H}}_{l-1}^H \boldsymbol{F}_{l-1} \boldsymbol{F}_{l-1}^H \hat{\boldsymbol{H}}_{l-1}, \qquad (33)
\end{aligned}$$

with $\boldsymbol{Q}_0 = \boldsymbol{I}_{KN_t}$, which completes the proof.

Appendix B
Proof of Theorem 2

From (33), $\boldsymbol{Q}_{l-1}$ in (10) can be expressed as

$$\boldsymbol{Q}_{l-1} = \boldsymbol{I}_{KN_t} + \gamma \sum_{i=1}^{l-1} \hat{\boldsymbol{H}}_i^H \boldsymbol{F}_i \boldsymbol{F}_i^H \hat{\boldsymbol{H}}_i, l = 2, \ldots, L. \qquad (34)$$

1) When l is small: With the assumption of very low SNRs in CF mmWave massive MIMO, we have $\boldsymbol{Q}_{l-1} \approx \boldsymbol{I}_{KN_t}$ for small l, leading to

$$R_l^{\mathrm{a}} \approx \tilde{R}_l^{\mathrm{a}} = \log_2 \det\left(\boldsymbol{I}_N + \gamma \boldsymbol{F}_l^H \hat{\boldsymbol{H}}_l \hat{\boldsymbol{H}}_l^H \boldsymbol{F}_l\right), \qquad (35)$$

where R_l^{a} is the sub-rate associated with the lth AP in SC-HBF, given in (10). The unconstrained combiner that maximizes $\tilde{R}_l$ in (35) is the matrix with columns being the N singular vectors corresponding to the N largest singular values of $\hat{\boldsymbol{H}}_l$. As a result, the analog combining vectors in the D-HBF scheme can be determined as in (15).

2) As l increases: Because $\boldsymbol{F}_i$ only depends on $\hat{\boldsymbol{H}}_i$ for small l, $\{\hat{\boldsymbol{H}}_i^H \boldsymbol{F}_i \boldsymbol{F}_i^H \hat{\boldsymbol{H}}_i\}, i = 1, \ldots, l-1$ are independent of each other. As l grows and becomes sufficiently large, by the law of large numbers, we have $\sum_{i=1}^{l-1} \hat{\boldsymbol{H}}_i^H \boldsymbol{F}_i \boldsymbol{F}_i^H \hat{\boldsymbol{H}}_i \rightarrow (l-1)\bar{\boldsymbol{E}}$, where $\bar{\boldsymbol{E}} = \mathbb{E}\left\{\hat{\boldsymbol{H}}_i^H \boldsymbol{F}_i \boldsymbol{F}_i^H \hat{\boldsymbol{H}}_i\right\}$ has constant diagonal elements and zeros for off-diagonal elements. Therefore, $\boldsymbol{Q}_{l-1}$ in (34) becomes approximately diagonal even when l is large, as does $\boldsymbol{Q}_{l-1}^{-1}$.

Based on the ordered singular value decomposition, $\hat{\boldsymbol{H}}_l$ can be factorized as $\hat{\boldsymbol{H}}_l = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$, where $\boldsymbol{\Sigma}$ is an $N_r \times KN_t$ rectangular diagonal matrix with the singular values of $\hat{\boldsymbol{H}}_l$ on the main diagonal in decreasing order, whereas $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices of size $N_r \times N_r$ and $KN_t \times KN_t$, whose columns are the left- and right-singular vectors of $\hat{\boldsymbol{H}}_l$, respectively. Then, R_l^{a} in (10) can be expressed as

$$R_l^{\mathrm{a}} = \log_2 \det(\boldsymbol{I}_N + \gamma \boldsymbol{F}_l^H \boldsymbol{U} \underbrace{\boldsymbol{\Sigma}\boldsymbol{V}^H \boldsymbol{Q}_{l-1}^{-1} \boldsymbol{V} \boldsymbol{\Sigma}^H}_{\triangleq \boldsymbol{\Lambda}} \boldsymbol{U}^H \boldsymbol{F}_l). \qquad (36)$$

Because $\boldsymbol{Q}_{l-1}^{-1}$ is approximately a diagonal matrix with constant diagonal elements, as shown above, $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}\boldsymbol{V}^H \boldsymbol{Q}_{l-1}^{-1} \boldsymbol{V} \boldsymbol{\Sigma}^H$ becomes approximately a diagonal matrix, and its diagonal elements are in decreasing order. Therefore, the optimal solution of $\max_{\boldsymbol{F}_l} R_l^{\mathrm{a}} = \log_2 \det(\boldsymbol{I}_N + \gamma \boldsymbol{F}_l^H \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^H \boldsymbol{F}_l)$ is approximately the matrix with columns being the first N columns of $\boldsymbol{U}$, which are the singular vectors corresponding to the N largest singular values of $\hat{\boldsymbol{H}}_l$, implying the analog combining vectors given in (15) for D-HBF. This completes the proof.

References

[1] N. T. Nguyen and K. Lee, "Coverage and Cell-Edge Sum-Rate Analysis of mmWave Massive MIMO Systems With ORP Schemes and MMSE Receivers," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5349–5363, 2018.

[2] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.

[3] N. T. Nguyen and K. Lee, "Cell coverage extension with orthogonal random precoding for massive MIMO systems," *IEEE Access*, vol. 5, pp. 5410–5424, 2017.

[4] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *IEEE 16th Int. Workshop Signal Process. Advances in Wireless Commun. (SPAWC)*, 2015, pp. 201–205.

[5] ——, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, 2017.

[6] G. Femenias and F. Riera-Palou, "Cell-Free Millimeter-Wave Massive MIMO Systems With Limited Fronthaul Capacity," *IEEE Access*, vol. 7, pp. 44 596–44 612, 2019.

[7] G. Interdonato, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Downlink training in cell-free massive MIMO: A blessing in disguise," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5153–5169, 2019.

[8] A. Papazafeiropoulos, P. Kourtessis, M. Di Renzo, S. Chatzinotas, and J. M. Senior, "Performance Analysis of Cell-Free Massive MIMO Systems: A Stochastic Geometry Approach," *IEEE Trans. Veh. Tech.*, vol. 69, no. 4, pp. 3523–3537, 2020.

[9] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and H. Yang, "Cell-free massive MIMO systems," in *IEEE Asilomar Conf. Signals, Systems and Computers*, 2015, pp. 695–699.

[10] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, 2019.

[11] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy efficiency in cell-free massive MIMO with zero-forcing precoding design," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1871–1874, 2017.

[12] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. and Network.*, vol. 2, no. 1, pp. 25–39, 2017.

[13] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "Energy efficiency of the cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Trans. Green Commun. Networking*, vol. 3, no. 4, pp. 971–987, 2019.

[14] G. Interdonato, P. Frenger, and E. G. Larsson, "Scalability aspects of cell-free massive MIMO," in *IEEE Int. Conf. Commun. (ICC)*, 2019, pp. 1–6.

[15] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, 2020.

[16] S. E. Hajri, J. Denis, and M. Assaad, "Enhancing Favorable Propagation in Cell-Free Massive MIMO Through Spatial User Grouping," in *IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, 2018, pp. 1–5.

[17] X. Hu, C. Zhong, X. Chen, W. Xu, H. Lin, and Z. Zhang, "Cell-free massive MIMO systems with low resolution ADCs," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6844–6857, 2019.

[18] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, 2017.

[19] J. Zhang, Y. Wei, E. Björnson, Y. Han, and S. Jin, "Performance analysis and power control of cell-free massive MIMO systems with hardware impairments," *IEEE Access*, vol. 6, pp. 55 302–55 314, 2018.

[20] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 706–709, 2017.

[21] S.-H. Park, O. Simeone, Y. C. Eldar, and E. Erkip, "Optimizing Pilots and Analog Processing for Channel Estimation in Cell-Free Massive MIMO with One-Bit ADCs," in *IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, 2018, pp. 1–5.

[22] J. Li, D.-W. Yue, and Y. Sun, "Performance analysis of millimeter wave massive MIMO systems in centralized and distributed schemes," *IEEE Access*, vol. 6, pp. 75 482–75 494, 2018.

[23] M. Alonzo and S. Buzzi, "Cell-free and user-centric massive MIMO at millimeter wave frequencies," in *IEEE Annual Int. Symposium Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–5.

[24] M. Alonzo, S. Buzzi, A. Zappone, and C. D'Elia, "Energy-Efficient Power Control in Cell-Free and User-Centric Massive MIMO at Millimeter Wave," *IEEE Trans. Green Commun. Network.*, 2019.

[25] N. T. Nguyen and K. Lee, "Unequally Sub-connected Architecture for Hybrid Beamforming in Massive MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1127–1140, Feb. 2020.

[26] H. Liu, J. Zhang, X. Zhang, A. Kurniawan, T. Juhana, and B. Ai, "Tabu-search-based pilot assignment for cell-free massive MIMO systems," *IEEE Trans. Veh. Tech.*, vol. 69, no. 2, pp. 2286–2290, 2019.

[27] Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel estimation for cell-free mmWave massive MIMO through deep learning," *IEEE Trans. Veh. Tech.*, vol. 68, no. 10, pp. 10 325–10 329, 2019.

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

[29] T.-H. Tai, W.-H. Chung, and T.-S. Lee, "A low complexity antenna selection algorithm for energy efficiency in massive mimo systems," in *IEEE Int. Conf. Data Science and Data Intensive Systems*, 2015, pp. 284–289.

[30] Z. Liu, W. Du, and D. Sun, "Energy and spectral efficiency tradeoff for massive MIMO systems with transmit antenna selection," *IEEE Trans. Veh. Tech.*, vol. 66, no. 5, pp. 4453–4457, 2016.

[31] X. Gao, L. Dai, and A. M. Sayeed, "Low RF-complexity technologies to enable millimeter-wave MIMO with large antenna array for 5G wireless communications," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 211–217, 2018.

[32] A. Kaushik, E. Vlachos, C. Tsinos, J. Thompson, and S. Chatzinotas, "Joint bit allocation and hybrid beamforming optimization for energy efficient millimeter wave MIMO systems," *IEEE Trans. Green Commun. Network.*, vol. 5, no. 1, pp. 119–132, 2020.

[33] S. Dey, E. Sharma, and R. Budhiraja, "Dynamic resolution ADC/DAC massive MIMO FD relaying system over correlated rician channel," in *European Signal Process. Conf. (EUSIPCO)*. IEEE, 2021, pp. 1653–1657.

[34] J.-C. Chen, "Spectral- and energy-efficient hybrid receivers for millimeter-wave massive multiuser MIMO uplink systems with variable-resolution ADCs," *IEEE Systems Journal*, 2020.

[35] I. Z. Ahmed, H. Sadjadpour, and S. Yousefi, "Capacity analysis and bit allocation design for variable-resolution ADC in massive MIMO," in *Military Commun. Conf. (MILCOM)*. IEEE, 2018, pp. 1–6.

[36] Y. Dong and L. Qiu, "Spectral efficiency of massive MIMO systems with low-resolution ADCs and MMSE receiver," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1771–1774, 2017.

[37] Y. Xiong, S. Sun, L. Qin, N. Wei, L. Liu, and Z. Zhang, "Performance analysis on cell-free massive MIMO with capacity-constrained fronthauls and variable-resolution ADCs," *IEEE Systems Journal*, 2021.

[38] X. Song, T. Kühne, and G. Caire, "Fully-/partially-connected hybrid beamforming architectures for mmwave mu-mimo," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1754–1769, 2019.

[39] G. M. Gadiel, N. T. Nguyen, and K. Lee, "Dynamic unequally sub-connected hybrid beamforming architecture for massive mimo systems," *IEEE Trans. Veh. Tech.*, vol. 70, no. 4, pp. 3469–3478, 2021.

[40] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave mimo systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, 2016.

[41] A. Kaushik, J. Thompson, E. Vlachos, C. Tsinos, and S. Chatzinotas, "Dynamic RF Chain Selection for Energy Efficient and Low Complexity Hybrid Beamforming in Millimeter Wave MIMO Systems," *IEEE Trans. Green Commun. Networking*, vol. 3, no. 4, pp. 886–900, 2019.

[42] E. Vlachos, J. Thompson, A. Kaushik, and C. Masouros, "Radio-frequency chain selection for energy and spectral efficiency maximization in hybrid beamforming under hardware imperfections," *Proceedings of the Royal Society A*, vol. 476, no. 2244, p. 20200451, 2020.

[43] M. Feng, S. Mao, and T. Jiang, "Dynamic base station sleep control and RF chain activation for energy-efficient millimeter-wave cellular systems," *IEEE Trans. Veh. Tech.*, vol. 67, no. 10, pp. 9911–9921, 2018.

[44] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.

[45] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.

[46] Z. Luo, H. Liu, Y. Li, H. Wang, and L. Zhang, "Robust hybrid transceiver design for af relaying in millimeter wave systems under imperfect CSI," *IEEE Access*, vol. 6, pp. 29 739–29 746, 2018.

[47] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, 2015.

[48] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, "38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications," in *IEEE ICC*, 2012, pp. 4568–4573.

[49] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, 2016.

[50] S. Payami, N. M. Balasubramanya, C. Masouros, and M. Sellathurai, "Phase shifters versus switches: An energy efficiency perspective on hybrid beamforming," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 13–16, 2018.

[51] M. Brand, "Fast low-rank modifications of the thin singular value decomposition," *Linear algebra and its applications*, vol. 415, no. 1, pp. 20–30, 2006.

[52] B. Dai and W. Yu, "Energy efficiency of downlink transmission strategies for cloud radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 1037–1050, 2016.

[53] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, 2015.

[54] K. Roth and J. A. Nossek, "Achievable rate and energy efficiency of hybrid and digital beamforming receivers with low resolution ADC," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2056–2068, 2017.

[55] K. Roth, H. Pirzadeh, A. L. Swindlehurst, and J. A. Nossek, "A comparison of hybrid beamforming and digital beamforming with low-resolution ADCs for multiple users and imperfect CSI," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 484–498, 2018.

[56] M. Bashar, H. Q. Ngo, K. Cumanan, A. G. Burr, P. Xiao, E. Björnson, and E. G. Larsson, "Uplink Spectral and Energy Efficiency of Cell-Free Massive MIMO with Optimal Uniform Quantization," *IEEE Trans. Commun.*, 2020.

[57] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of rf chains and phase shifters, and scheduling design with hybrid analog–digital beamforming," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3311–3326, 2016.

[58] S. S. Ioushua and Y. C. Eldar, "A family of hybrid analog–digital beamforming methods for massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 67, no. 12, pp. 3243–3257, 2019.

[59] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.

[60] M. Ma, N. T. Nguyen, and M. Juntti, "Switch-based hybrid beamforming for wideband multi-carrier communications," in *IEEE Int. ITG Workshop Smart Antennas (WSA)*, 2021.

[61] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, 2015.

[62] N. T. Nguyen, D. Vu, K. Lee, and M. Juntti, "Hybrid relay-reflecting intelligent surface-assisted wireless communications," *IEEE Trans. Veh. Technol.*, 2022.

[63] 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, 2020.

[64] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 205–211, 2018.