Cooperative Beamforming for RIS-Aided Cell-Free Massive MIMO Networks

Xinying Ma, Deyou Zhang, Ming Xiao, Senior Member, IEEE, Chongwen Huang, Member, IEEE, and Zhi Chen, Senior Member, IEEE

Abstract—The combination of cell-free massive multiple-input multiple-output (CF-mMIMO) and reconfigurable intelligent surface (RIS) is envisioned as a promising paradigm to improve network capacity and enhance coverage capability. However, to reap full benefits of RIS-aided CF-mMIMO, the main challenge is to efficiently design cooperative beamforming (CBF) at base stations (BSs), RISs, and users. Firstly, we investigate the fractional programing to convert the weighted sum-rate (WSR) maximization problem into a tractable optimization problem. Then, the alternating optimization framework is employed to decompose the transformed problem into a sequence of sub-problems, i.e., hybrid BF (HBF) at BSs, passive BF at RISs, and combining at users. In particular, the alternating direction method of multipliers algorithm is utilized to solve the HBF subproblem at BSs. Concretely, the analog BF design with unit-modulus constraints is solved by the manifold optimization (MO) while we obtain a closed-form solution to the digital BF design that is essentially a convex least-square problem.

I. INTRODUCTION

The deployment of ultra-dense network (UDN) has been considered as an indispensable technology to meet the requirements of massive connectivity and seamless coverage for sixth generation (6G) wireless networks [1]. Intriguingly, the basic idea of UDN can be well achieved by an emerging embodiment, namely cell-free massive multiple-input multiple-output (CF-mMIMO) [2], [3], [4], which benefits from the integration of distributed networks and massive MIMO.

With regard to CF-mMIMO networks, a considerable number of distributed base stations (BSs) coherently serve a small number of users on the same time-frequency resources and all the BSs are connected to a central processing unit (CPU) via backhaul wireless links [5]. There is no channel state information (CSI) exchange among these BSs, and thus the CPU acts an important role in coordination and computational assistance for BSs. Under such a setup, CF-mMIMO entails some distinctive advantages as follows: i) thanks to a large number of BSs, CF-mMIMO provides much higher coverage probability and stronger macro-diversity than conventional small-cell networks [6]; ii) since CF-mMIMO networks possess no cell boundaries, the inter-cell interference (ICI) can be eliminated by jointly cooperating the distributed BSs [2]; iii) favorable propagation is another potential characteristic existing in CF-mMIMO networks that is able to alleviate inter-user interferences (IUIs) [7]. Nevertheless, in addition to these advantages, CF-mMIMO also faces some challenges in practice.

A. Prior Works

One of the main challenges is that high frequency bands adopted by CF-mMIMO in 6G wireless networks are more vulnerable to blockage, and thus much denser deployment...
of BSs is demanded to compensate for the coverage hole and the serious propagation attenuation, e.g., millimeter wave (mmWave) and terahertz (THz) bands [8], [9]. However, more distributed BSs definitely result in much higher network overhead. To balance the network capacity and energy consumption, the reconfigurable intelligent surface (RIS) is regarded as a promising paradigm to assist mmWave CF-mMIMO networks in an energy-efficient way [10], [11], [12]. Specifically, RIS consists of a large number of passive reflecting elements and is capable of manipulating the propagation direction of impinging electromagnetic waves by adjusting the amplitude and phase shift without active radio frequency (RF) chains [13], [14], [15]. With the assistance of RISs, the wireless environment becomes controllable and reconfigurable, and CF-mMIMO networks can enhance the signal coverage capability with low energy consumption [16]. Therefore, the combination of RIS and CF-mMIMO is cast as a potential approach to improve the network capability. To fully reap the benefits, the cooperative beamforming (CBF) design (e.g., at BSs, RISs and users) is a crucial performance indicator for RIS-aided mmWave CF-mMIMO networks. In [17], authors first estimate the CSI by sending uplink pilots, and then design the conjugate BF for downlink data transmission with estimated CSI. More importantly, considering the communication scenarios with high mobility and short coherence time, it is more appropriate to design the CBF schemes based on the statistical CSI, which can greatly reduce feedback overhead and computational complexity [18]. To maximize the energy efficiency, the work in [19] proposes a joint design of transmit beamformers at BSs and reflecting coefficients at RISs in the case of limited backhaul capacity. In [20], a fully decentralized BF scheme is originally developed, which greatly decreases the backhaul signaling compared to centralized approaches. However, these aforementioned works just consider single-antenna users that are unable to meet the requirements of high frequency wireless networks. Although the case with multi-antenna users is considered in [21], it neglects the combining design at users and only considers the fully digital BF (DBF) architecture that endures heavy power consumption and huge hardware complexity at mmWave or THz bands. To this end, it is imperative and practical for RIS-aided mmWave CF-mMIMO networks to concurrently explore the hybrid BF (HBF) structure at BSs, the combining at multi-antenna users and passive BF at RISs.

In addition, it is impractical for conventional mmWave CF-mMIMO networks that all users are served by all BSs. This is because fully-connected mode endures serious communication costs, including but not limited to backhaul signaling overhead, energy consumption and intensive computation [22]. Fortunately, one of the most promising solutions is to consider the partially-connected CF-mMIMO (P-CF-mMIMO), also called user-centric approach, where each user is only served by a subset of BSs [23], [24], [25], [26]. By properly cutting down the number of connected links among users and BSs, the P-CF-mMIMO network is capable of suppressing communication costs with negligible performance degradation compared to the conventional CF-mMIMO network. For instance, considering the P-CF-mMIMO network, the work in [25] considers that each BS only selects a part of users that own stronger channel gains (e.g., larger norms). This simple selection criterion brings obvious performance penalty in the case of perfect CSI and erodes fairness for users with weak channel conditions. In [26], a newly scalable CF-mMIMO system is proposed by exploiting the dynamic cooperation cluster, but neglects the diagonal matrix optimization in terms of BS selection. In addition, the work in [27] proposes a structured massive access framework for CF-mMIMO systems, where each user selects its desired BSs by a competitive mechanism. To minimize the network power consumption, authors propose a low-complexity user selection approach by treating the selection problem as an alternative problem [28]. These heuristic BS selection strategies cannot always guarantee the optimal network performance, especially for complex and diverse 6G application scenarios. To this end, except the existing works [25], [26], [27], [28], solving the BS selection problem from the perspective of integer programming is another promising alternative.

B. Our Contributions

Against the above background, this paper takes into account a more practical RIS-aided mmWave CF-mMIMO scenarios that adopts HBF architecture and deploys multiple antennas at users, and models the intractable BS selection problem as a solvable optimization problem, which are greatly different from prior works. Based on the above discussion, we first focus on the joint design of passive BF at RISs and active BF at transceivers by maximizing the weighted sum-rate (WSR) for the RIS-aided mmWave CF-mMIMO system. Then, the second implementation is to address the P-CF-mMIMO network and solve the BS selection problem from the perspective of integer programming. The main contributions can be listed as follows.

- We propose a more general RIS-aided mmWave CF-mMIMO scenario, where multiple BSs with multiple antennas serve multiple users with multiple antennas assisted by multiple RISs. Then we transform the non-convex and multiple-ratio WSR maximization problem into a sequence of tractable subproblems by leveraging the fractional programming (FP).
- We utilize an improved alternating direction method of multipliers (ADMM) algorithm to optimize the augmented Lagrangian objective with non-convex constraints, aiming to achieve the HBF design at BSs. Specifically, we use the manifold optimization (MO) technique to design the analog BF (ABF) with unit-modulus constraints and derive a closed-form solution for the DBF design.
- We formulate the passive BF design at RISs as a convex quadratically constrained quadratic program (QCQP) problem and directly employ the primal-dual subgradient (PDS) algorithm to optimize reflecting coefficients of RISs. Since the combining design at users is recast as a non-convex QCQP problem with unit-modulus constraints, the MO algorithm is utilized to obtain a high-quality solution.
In addition, we originally model the BS selection problem as a binary integer quadratic programming (BIQP) problem in the RIS-aided P-CF-mMIMO network, and develop a relaxed linear approximation (RLA) algorithm to solve this formulated BIQP problem by substituting the quadratic term with new variables and linear constraints.

Simulation results reveal that RISs can enhance the WSR performance of both CF-mMIMO and P-CF-mMIMO networks compared to the case without RISs and the case with random phase of RISs. Remarkably, the RLA network compared to the case without RISs and the case without RISs and the RLA or CBF design. Section IV discusses the RLA based BS selection and CBF design. Section V and Section VI, respectively.

The remainder of this paper is organized as follows. The RIS-enabled CF-mMIMO system model and problem formulation are presented in Section II. In Section III, we explore the CBF design. Section IV discusses the RLA based BS selection for the P-CF-mMIMO system. Finally, simulation results and conclusion are presented in Section V and Section VI, respectively.

Notations: $A^H$, $A^\dagger$, $A^T$, $A^{-1}$, and rank ($A$) are conjugate transpose, conjugate transpose, inverse, pseudo-inverse and the rank of $A$, respectively. $\|A\|_F$ is the Frobenius norm. $\text{diag} (a)$ is a diagonal matrix with elements of $a$ on its diagonal. $\text{tr} (A)$ is the trace of matrix $A$. $\text{vec} (A)$ is the column-ordered vectorization. $1_N$ and $0_N$ denote $N$-dimensional all-ones and all-zeros vector. $\Re (\cdot)$ and $\Im (\cdot)$ denote the real and imaginary part of its argument. $\otimes$ and $\circ$ denote the Kronecker and Hadamard products. $\mathbb{E} [\cdot]$ denotes the expectation, and $O (\cdot)$ indicates the number of complex multiplications.

II. System Model of RIS-AIDED CF-mMIMO Networks

In this section, we introduce the RIS-aided CF-mMIMO system model as well as the formulation of the WSR maximization problem.

A. System Model

We consider a downlink RIS-aided CF-mMIMO network, as shown in Fig. 1, where a set of BSs $B = \{1,2,\ldots,B\}$ serve a group of multi-antenna users $K = \{1,2,\ldots,K\}$ via the aid of a set of RISs $\mathcal{R} = \{1,2,\ldots,R\}$. Each BS with $N_{\text{RF}}$ RF chains, each user with one RF chain and each RIS without RF chain are equipped with $N_t$, $N_r$ and $M$ array elements, respectively. Besides, all BSs and RISs are connected to the CPU to implement signal processing. Assume that BSs employ the HBF architecture and each BS transmits $K$ data streams to the users. In the downlink mode, the transmitted signal from the $b$th BS can be expressed as

$$x_b = F_{\text{RF},b} F_{\text{BB},b} s = \sum_{k=1}^{K} F_{\text{RF},b} f_{\text{BB},b,k} s_k,$$

where $s \in [s_1, s_1, \ldots, s_K]^T \in \mathbb{C}^{N \times 1}$ indicates the symbol vector that satisfies $\mathbb{E} [ss^H] = I_K$. $F_{\text{RF},b} \in \mathbb{C}^{N_t \times N_{\text{RF}}}$. $F_{\text{BB},b} \in [f_{\text{BB},b,1}, \ldots, f_{\text{BB},b,K}] \in \mathbb{C}^{N_{\text{RF}} \times K}$ denote AFB and DBF matrices at the $b$th BS, respectively. The power constraint for the $b$th BS is denoted by $\|F_{\text{RF},b} f_{\text{BB},b} s\|_2^2 \leq P_b$ ($\forall b \in B$), where $P_b$ is the maximum transmit power.

Due to the deployment of RISs, the entire channel between the $b$th BS and the $k$th user consists of BS-user link $H_{b,k} \in \mathbb{C}^{N_t \times N_r}$, BS-RIS link $G_{b,r} \in \mathbb{C}^{M \times N_r}$ and RIS-user link $V_{r,k} \in \mathbb{C}^{N_r \times M}$. Hence, the equivalent channel from the $b$th BS to the $k$th user can be expressed as

$$H_{b,k} = \tilde{H}_{b,k} + \sum_{r=1}^{R} V_{r,k} \Theta_r G_{b,r},$$

where $\Theta_r = \text{diag}(\theta_{r,1}, \ldots, \theta_{r,M}) \in \mathbb{C}^{M \times M}$ denotes the phase shift matrix of the $r$th BS and each diagonal entry can be further defined as $\theta_{r,m} = \phi_{r,m} + j \omega_{r,m}, r \in \mathcal{R}, m \in \mathcal{M}$. In practice, the reflecting amplitude $u_{r,m}$ and phase shift $\phi_{r,m}$ can be controlled separately, where $|u_{r,m}| \leq 1$. Considering the sparse nature of mmWave channel, the typical Saleh-Valenzuela channel model with limited scattering paths is adopted [29]. Thus, $\tilde{H}_{b,k}, G_{b,r}$ and $V_{r,k}$ can be represented as

$$H_S = \sqrt{\frac{N_1 N_2}{L}} \sum_{l=1}^{L} \beta_l a_2 (\omega^l_1, \chi^l_1) a_1^H (\omega^l_1, \chi^l_1),$$

where $H_S \in \{\tilde{H}_{b,k}, G_{b,r}, V_{r,k}\}$. $L$ is the number of paths that contains one LoS path and $(L - 1)$ NLoS paths, and $\beta_l$ denotes the complex path gain of $l$th path. $N_1$ and $N_2$ respectively denote the numbers of array elements at transmitter and receiver. $\omega^l_1 (\chi^l_1)$ and $\omega^l_2 (\chi^l_2)$ denote the associated azimuth (elevation) angles of arrival (AoAs) and angles of departure (AoDs) of $l$th path, respectively. In addition, the uniform planar array (UPA) structure is considered for BSs, RISs and users. The array response for the UPA with $N_x N_y$-elements can be given by

$$a (\omega, \varsigma) = \frac{1}{\sqrt{N_x N_y}} \left[1, e^{j 2 \pi \omega d_x (n_x \sin(\omega) + n_y \cos(\varsigma))}, \ldots \right]^T,$$

where $0 \leq n_x \leq N_x - 1$ and $0 \leq n_y \leq N_y - 1$. $\lambda_s$ and $d_s$ represent the signal wavelength and the antenna spacing, respectively.
where $\tilde{H}_{b,k}^{\text{ideal}}$ is the perfect CSI, $\tilde{H}_{b,k}$ is the estimated CSI and $\Delta \tilde{H}_{b,k}$ denotes estimated CSI error that follows the circularly symmetric complex Gaussian (CSCG) distribution, i.e.,

$$\text{vec} \left( \Delta \tilde{H}_{b,k} \right) \sim \mathcal{CN} \left( 0, \Sigma_{b,k}^{1} \right),$$

where $\Sigma_{b,k}^{1} \in \mathbb{C}^{N_c \times N_c}$ is the error covariance matrix. Similarly, $\Sigma_{b,r}^{2}$ and $\Sigma_{r,k}^{3}$ are the error covariance matrices of cascaded BS-RIS-user channels (e.g., $G_{b,r}$ and $V_{r,k}$), respectively. Given the estimated $\tilde{H}_{b,k}$, $G_{b,r}$ and $V_{r,k}$, the main goal aims to design passive and active beamforming.

By combining (1) and (2), the downlink received signal from the $b$th BS to the $k$th user can be expressed as

$$y_{b,k} = \sum_{j=1}^{K} H_{b,k} F_{RF,b} f_{BB,b,j} s_j + n_k,$$

where $n_k$ denotes the noise at the $k$th user following Gaussian distribution of $\mathcal{CN} \left( 0, \sigma_k^2 I \right)$ that corrupts the received signal. After being processed by a combiner $w_k \in \mathbb{C}^{N_c \times 1}$, the received signal of the $k$th user can be further written as (9), shown at the bottom of the page, where $V_k = [V_{1,k}, \ldots, V_{R,k}] \in \mathbb{C}^{N_c \times R}$, $G_b = [G_{b,1}^T, \ldots, G_{b,R}^T]^T \in \mathbb{C}^{RM \times N_c}$ and $\Phi = \text{diag} (\Theta_1, \ldots, \Theta_R) \in \mathbb{C}^{RM \times RM}$, respectively.

According to the system model, the overall protocol of the RIS-aided CF-mMIMO network is presented in Fig. 2, where each frame contains multiple subframes and each subframe begins with the synchronization operation to enhance bandwidth efficiency. Moreover, the first subframe is utilized to estimate CSI and build new wireless links, while the rest subframes are able to implement the data transmission.

**B. Problem Formulation**

To maximize the WSR of the RIS-aided CF-mMIMO system, the signal-to-interference-plus-noise ratio (SINR) needs to be calculated for each user. Based on (9), the SINR of the $k$th user can be expressed as

$$\gamma_k = \frac{\sum_{b=1}^{B} w_k^H \tilde{H}_{b,k}^{\text{ideal}} F_{RF,b} f_{BB,b,k} s_k}{\sum_{b=1,j \neq k}^{B} w_k^H \tilde{H}_{b,k}^{\text{ideal}} F_{RF,b} f_{BB,b,j} s_j + \sigma_k^2},$$

where $\sigma_k^2 = N_c \sigma_{B,k}^2$ denotes the effective noise variance. Subsequently, the WSR of the RIS-aided CF-mMIMO system can be written as

$$R_{\text{sum}} = \sum_{k=1}^{K} \omega_k \log (1 + \gamma_k),$$

where $\omega_k > 0 (\forall k \in \mathcal{K})$ denotes the WSR weight for the $k$th user with $\sum_{k \in \mathcal{K}} \omega_k = 1$.

Finally, the WSR maximization problem can be formulated as

$$\begin{align*}
\max_{\mathbf{F}, \mathbf{\Phi}, \mathbf{w}} & \quad R_{\text{sum}} \\
n\text{s.t.} & \quad ||F_{RF,b} F_{BB,b}||_F^2 \leq P_b, \forall b \in \mathcal{B}, \\
& \quad ||F_{RF,b} (i_1, j_1)||_2^2 = 1, \forall i_1, j_1, b \in \mathcal{B}, \\
& \quad ||\Theta_r (i_2, i_3)||_2^2 \leq 1, \forall i_2, i_3, r \in \mathcal{R}, \\
& \quad ||w_k (i_3)||_2^2 = 1, \forall i_3, k \in \mathcal{K}, \\
& \quad \mathbf{F} = \{F_b \mid \forall b \in \mathcal{B}\}, \quad \mathbf{F}_b = F_{RF,b} F_{BB,b} \quad \text{and} \quad \mathbf{w} = [w_1^T, w_2^T, \ldots, w_K^T]^T, \quad \text{respectively.}
\end{align*}$$

**III. COOPERATIVE BEAMFORMING FOR RIS-AIDED CF-MMIMO NETWORKS**

The purpose of this section is to realize the CBF design for the RIS-aided CF-mMIMO network. To convert problem (12) into a tractable problem, we use the Lagrangian dual transform and FP transformation [31] to deal with the sum-logarithm term, which can be given by (13), shown at the bottom of the next page, where $\mu_k = \omega_k (1 + \lambda_k)$ and $\lambda = \{\lambda_k \mid \forall k \in \mathcal{K}\}$ refers to a collection of auxiliary variables.

By using the quadratic transform on the fractional term, we further recast $f_D (\mathbf{F}, \mathbf{\Phi}, \mathbf{w}, \lambda)$ as (14), shown at the bottom of the next page, where $\xi = \{\xi_k \mid \forall k \in \mathcal{K}\}$ denotes a set of auxiliary variables.

Subsequently, we define $f (\mathbf{F}, \mathbf{\Phi}, \mathbf{w}, \lambda, \xi) = -f_Q (\mathbf{F}, \mathbf{\Phi}, \mathbf{w}, \lambda, \xi)$, and problem (12) can be equivalently rewritten as

$$\begin{align*}
\min_{\mathbf{F}, \mathbf{\Phi}, \mathbf{w}} & \quad f (\mathbf{F}, \mathbf{\Phi}, \mathbf{w}, \lambda, \xi) \\
n\text{s.t.} & \quad (12b), (12c), (12d), (12e).
\end{align*}$$

It can be observed that constraints (12b)-(12e) involve multiple variables and include non-convex sets,
e.g., (12c), (12e). To tackle this issue, the alternating optimization (AO) is widely treated as an efficient approach [32]. More specifically, we optimize BF variables (F, Φ, w) and auxiliary variables (λ, ξ) in an iterative manner until the objective function converges.

A. Fix (F, Φ, w) and Solve (λ, ξ)

According to the FP transform process in [31], the optimal solutions to (λ, ξ) can be respectively calculated by solving $\partial f_D (F, \Phi, w, \lambda, \xi)/\partial \lambda_k = 0$ and $\partial f_Q (F, \Phi, w, \lambda, \xi)/\partial \xi_k = 0, \forall k \in \mathcal{K}$. Then, the optimal variables ($\lambda^*, \xi^*$) can be respectively given by

$$\lambda^*_k = \gamma_k, \quad (16a)$$

$$\xi^*_k = \frac{1}{\sqrt{\mu_k}} \left( \sum_{b=1}^{B} w^H_b H_{b,k} F_{RF,b} F_{BB,b,k} \right)^2 + \sigma^2, \quad (16b)$$

With the optimized ($\lambda^*, \xi^*$), the CBF optimization problem is divided into three subproblems that can be settled alternately.

B. Fix (Φ, w, λ, ξ) and Solve F

To begin with, the considered CF-mMIMO network belongs to a centralized mode, which implies that all the related information can be exchanged and acquired by the CPU. Therefore, the HBF design for B BSs can be executed in parallel. For $b$th ($\forall b \in \mathcal{B}$) BS, the reformulated WSR optimization problem can be given by

$$\min_{F_{RF,b}, F_{BB,b}} f_1 (F_{RF,b}, F_{BB,b}) \quad \text{s.t.} \quad (12b), (12c), \quad (17a)$$

where the new objective function $f_1 (F_{RF,b}, F_{BB,b})$ is given by

$$f_1 (F_{RF,b}, F_{BB,b}) = \sum_{k=1}^{K} \sum_{b=1}^{B} |\xi_k|^2 w^H_k H_{b,k} F_{RF,b} F_{BB,b,k} + C_{b,k,j} |^2$$

$$-2\Re \left( \sum_{k=1}^{K} \sqrt{\mu_k} \xi_k^* w^H_k H_{b,k} F_{RF,b} F_{BB,b,k} \right) + D_{b,k,j}, \quad (18)$$

where $C_{b,k,j}$ and $D_{b,k,j}$ are irrelevant terms for the $b$th BS, which can be respectively given by

$$C_{b,k,j} = \sum_{p \neq b}^{B} w^H_p H_{p,k} F_{RF,p} F_{BB,p,j}$$

$$D_{b,k,j} = -2\Re \left( \sum_{k=1}^{B} \sum_{p \neq b}^{B} \sqrt{\mu_k} \xi_k^* w^H_k H_{p,k} F_{RF,p} F_{BB,p,j} \right), \quad (19)$$

To solve the constrained problem (17), ADMM is an efficient tool that blends the benefits of dual decomposition and augmented Lagrangian method [33]. By considering $F_b = F_{RF,b} F_{BB,b}$ into problem (17), the augmented Lagrangian function can be formulated as

$$L_1 (F_b, F_{RF,b}, F_{BB,b}, \Delta) = f_1 (F_{RF,b}, F_{BB,b})$$

$$+ \frac{\rho_1}{2} \left( \left| |F_b| - F_{RF,b} F_{BB,b} + \Delta \right|^2 \right) + \text{E} (\Delta), \quad (20)$$

where $\rho_1 > 0$ indicate the dual variable and the penalty parameter and $\text{E} (\Delta) = -\text{tr}(\Delta^H \Delta)/2\rho_1$ is a constant term when $\Delta \in \mathbb{C}^{N_t \times \mathcal{K}}$ remains fixed.

Based on the working principle of ADMM [34], [35], the iterative process for optimizing $L_1 (F_b, F_{RF,b}, F_{BB,b}, \Delta)$ can be expressed as

$$F_{RF,b}^{t+1} = \arg \min_{F_{RF,b}} L_1 (F_b, F_{RF,b}^{t}, F_{BB,b}, \Delta^t)$$

$$\text{s.t.} \quad |F_{BB,b}|^2 \leq P_b, \quad (21a)$$

$$F_{BB,b}^{t+1} = \arg \min_{F_{BB,b}} L_1 (F_b, F_{RF,b}^{t}, F_{BB,b}^{t+1}, \Delta^t)$$

$$\text{s.t.} \quad |F_{RF,b} (i_1, j_1)|^2 = 1, \forall i_1, j_1, \quad (21b)$$

$$\Delta^{t+1} = \Delta^t + \rho_1 \left( F_{RF,b}^t - F_{RF,b}^{t+1} \right), \quad (21c)$$

where superscript $t$ denotes the iteration index for proceeding ADMM algorithm.

1) Optimize $F_b$: Given fixed ($F_{RF,b}^t, F_{BB,b}^t, \Delta^t$), the minimization problem with regard to $F_b$ can be written as

$$\min_{F_b} \frac{\rho_1}{2} \left( \left| |F_b| - F_{RF,b}^t F_{BB,b}^t + \Delta^t \right|^2 \right)$$

$$\text{s.t.} \quad |F_b|^2 \leq P_b, \quad (22)$$

$$f_D (F, \Phi, w, \lambda) = \sum_{k=1}^{K} \omega_k \log (1 + \lambda_k) - \lambda_k \omega_k + \frac{\mu_k}{2} \left( \sum_{b=1}^{B} w^H_b H_{b,k} F_{RF,b} F_{BB,b,k} \right)^2$$

$$\sum_{j=1}^{B} w^H_j H_{b,k} F_{RF,b} F_{BB,b,j}$$

$$f_Q (F, \Phi, w, \lambda, \xi) = \sum_{k=1}^{K} \omega_k \log (1 + \lambda_k) + \Re \left( \sum_{k=1}^{K} \left( \sum_{b=1}^{B} \sqrt{\mu_k} \xi_k^* w^H_k H_{b,k} F_{RF,b} F_{BB,b,k} \right) \right)$$

$$- \sum_{k=1}^{K} \omega_k \lambda_k - \sum_{k=1}^{K} |\xi_k|^2 \sigma^2 - \sum_{k=1}^{K} \sum_{j=1}^{B} \left( \sum_{b=1}^{B} w^H_b H_{b,k} F_{RF,b} F_{BB,b,j} \right)^2, \quad (14)$$
Taking into account the power constraint, the augmented Lagrangian for problem (22) can be formulated as

$$
L_2 (F_b, \mu_t) = \frac{\rho_1}{2} \left\| F_b - F_{RF,b}^t F_{BB,b} + \frac{\Delta^t}{\rho_1} \right\|^2_F + \tilde{\mu}_F \left( \left\| F_b \right\|^2_F - P_b \right),
$$

(23)

where $\tilde{\mu}_F$ indicates the Lagrange multiplier. By solving Karush-Kuhn-Tucker (KKT) conditions for (23), given by

$$
\frac{\partial L_2 (F_b, \tilde{\mu}_F)}{\partial F_b} = 0, \quad \left\| F_b \right\|^2_F \leq P_b,
$$

$$
\tilde{\mu}_F \left( \left\| F_b \right\|^2_F - P_b \right) = 0, \quad \tilde{\mu}_F \geq 0,
$$

(24)
a closed-form solution to $F_b$ can be expressed as

$$
F_{t+1}^b = \sqrt{\frac{\rho_1}{2}} F_{RF,b}^t F_{BB,b}^t - \frac{\Delta^t}{\rho_1}.
$$

(25)

2) Optimize $F_{RF,b}$: With fixed $(F_{t+1}^b, F_{BB,b}^t | \Delta^t)$, the optimization problem (21b) for $F_{RF,b}$ can be reformulated as

$$
\mathcal{F}_{RF,b}^{t+1} = \arg \min_{F_{RF,b}} f_2 (F_{RF,b})
$$

; s.t. $\left\| F_{RF,b} \right\|^2_F (i_1, j_1) = 1, \forall i_1, j_1,

(26)

where

$$
f_2 (F_{RF,b}) = \sum_{k=1}^{K} \sum_{j=1}^{K} \left\| \xi_k^b \parallel w_k^T H_{b,k} F_{RF,b} F_{BB,b} + C_{b,k,j}^t \right\|^2_F
$$

$$
-2 \Re \left( \sum_{k=1}^{K} \sqrt{\tilde{\mu}_F} \xi_k^b w_k^T H_{b,k} F_{RF,b} F_{BB,b} \right)
$$

$$
+ \frac{\rho_1}{2} \left\| F_{t+1}^b - F_{RF,b} F_{BB,b} + \frac{\Delta^t}{\rho_1} \right\|^2_F.
$$

(27)

To reap the benefits of high performance and complexity promised by the MO method [36], we convert a constrained problem (26) into an unconstrained problem over smooth Riemannian manifolds [37]. Assuming that $F_{RF,b} = \text{vec}(F_{RF,b})$, problem (26) can be rewritten as

$$
\mathcal{F}_{RF,b}^{t+1} = \arg \min_{\mathcal{F}_{RF,b}} f_3 (\mathcal{F}_{RF,b})
$$

(28)

where

$$
f_3 (\mathcal{F}_{RF,b}) = f_{RF,b}^T \left( \frac{\rho_1}{2} B_b^T F_{BB,b}^t + \xi_b^t \right) \mathcal{F}_{RF,b}
$$

$$
+ 2 \Re \left( \mathcal{F}_{RF,b}^t H_{bb} \right),
$$

$$
\mathcal{A}_{b,k,j} = \mathcal{F}_{BB,b}^T \otimes \left( w_k^T H_{b,k} \right),
$$

$$
\mathcal{E}_b = \sum_{k=1}^{K} \sum_{j=1}^{K} \xi_k^b \parallel a_{b,k,j}^t \parallel a_{b,k,j}^t,\n$$

$$
\mathcal{B}_b = F_{BB,b}^T \otimes I_{N_t},
$$

$$
\mathcal{B}_b^t = \mathcal{F}_{BB,b}^T \otimes I_{N_t},
$$

$$
\mathcal{B}_b = \sum_{k=1}^{K} \sqrt{\mathcal{F}_{RF,b}} \left( \xi_k^b e_k^T \otimes \left( w_k^T H_{b,k} \right) \right)^H,
$$

$$
\mathcal{M}_b^t = \mathcal{B}_b^t + \mathcal{B}_b + \mathcal{B}_b
$$

$$
\Pi_b^t = \sum_{k=1}^{K} \sum_{j=1}^{K} \xi_k^b a_{b,k,j}^t H_{b,k,j} - \frac{\rho_1}{2} B_b^T \mathcal{M}_b^t - b_b^t,
$$

(30)

and the proof for problem (28) is provided in Appendix A.

In addition, $M_b$ represents the complex circle Riemannian manifold that can be specifically expressed as

$$
M_b = \{ f_{RF,b} \in \mathbb{C}^{N_t N_{RF}} : |f_{RF,b}| = 1 \}
$$

(31)

where $z$ indicates the tangent vector at the point $f_{RF,b} \in M_b$. More directly, $T_{M_b}$ can be further regarded as the set of all tangent vectors orthogonal to the given point, and one of these tangent vectors is the negative Riemannian gradient that implies the fastest descent direction [38]. To this end, we define a unique tangent vector $\nabla f_3 (f_{RF,b})$ as the Riemannian gradient at $f_{RF,b} \in M_b$, given by the orthogonal projection of the Euclidean gradient to the tangent spaces as follow

$$
\nabla f_3 (f_{RF,b}) = \nabla f_3 (f_{RF,b}) - \Re \left\{ \nabla f_3 (f_{RF,b}) \circ f_{RF,b} \right\} \circ f_{RF,b},
$$

(32)

where

$$
\nabla f_3 (f_{RF,b}) = \left( \frac{\rho_1}{2} B_b^T B_b + \xi_b^t \right) f_{RF,b} + \Pi_b^t.
$$

To guarantee that $\nabla f_3 (f_{RF,b})$ is mapped from $T_{M_b}$ onto the manifold itself, the retraction operation is proposed to determine the following point remaining on the prescribed manifold while moving along a tangent vector. Thus, the retraction of $\nabla f_3 (f_{RF,b})$ at $f_{RF,b} \in M_b$ can be expressed as

$$
f_{RF,b}^{t+1} = \text{Retr} (f_{RF,b}, \ell \nabla f_3 (f_{RF,b}))
$$

$$
= \text{vec} \left( f_{RF,b}^{t+1} - \ell \nabla f_3 (f_{RF,b}) \right),
$$

(33)

where superscript $i$ denotes the iteration index, $\ell$ is the step-size, and $\text{Retr} (\cdot)$ is the retraction operation from $T_{M_b}$ to $M_b$, respectively. Particularly, the step-size selection strategy for the MO algorithm is Armijo backtracking line search. Based on the above analysis, problem (28) can be well solved by the MO method.

3) Optimize $F_{BB,b}$: Given the set $(F_{t+1}^b, F_{BB,b}^{t+1}, \Delta^t)$, the DBF design $F_{BB,b}$ for problem (21c) can be expressed as

$$
F_{BB,b}^{t+1} = \arg \min_{F_{BB,b}} f_2 (F_{BB,b}),
$$

(34)

where $f_2 (F_{BB,b})$ is presented in (27). By defining $f_{BB,b} = \text{vec}(F_{BB,b})$, problem (34) can be reformulated as

$$
F_{BB,b}^{t+1} = \arg \min_{f_{BB,b}} f_4 (f_{BB,b})
$$

(35)
where
\[ f_4(f_{\text{BB},b}) = f_{\text{BB},b}^H \left( C_b^T + \frac{\rho_1}{2} D_b^{H} D_b^T \right) f_{\text{BB},b} \]
\[ + 2 \Re \left( \left( r_b^{(T)} + d_b^{(H)} - \frac{\rho_1}{2} m_b^{H} D_b^T \right) f_{\text{BB},b} \right) , \]
\[ \tilde{d}_{b,j} = \sum_{k=1}^{K} [\xi_k^2] c_{b,k}^T c_{b,j}^{* T}, d_{b} = [\tilde{d}_{b,1}^{T}, \ldots, \tilde{d}_{b,K}^{T}]^T, \]
\[ c_{b,k}^T = w_{k}^H H_{b,k} F_{\text{RF},b}^{T+1} D_b^T \]
\[ C_b^T = I_K \otimes \left( \sum_{k=1}^{K} [\xi_k^2] c_{b,k}^T c_{b,k}^{* T} \right), \]
\[ r_b^T = [F_{b,1}^T, \ldots, F_{b,K}^T]^T, \]
\[ \tilde{f}_{b,k}^{T+1} = -\sqrt{\mu_k} \xi_k^H w_{k}^H H_{b,k} F_{\text{RF},b}^{T+1} + E_{k,j}. \]

Since problem (35) with quadratic objective function is convex and unconstrained, the optimal solution can be obtained directly by setting \( \partial f_4(f_{\text{BB},b})/\partial f_{\text{BB},b} = 0 \). Thus, \( f_{\text{BB},b} \) can be given by
\[ f_{\text{BB},b}^{T+1} = \left( C_b^T + \frac{\rho_1}{2} D_b^{H} D_b^T \right)^{-1} \left( \frac{\rho_1}{2} D_b^{H} m_b^T - r_b^{T+1} \right). \]
(37)

C. Fix (F, Φ, λ, ξ) and Solve w

To design the combining vectors for K users, the original problem (15) with regard to w can be rewritten as
\[ \min_w f_5(w) \]
\[ \text{s.t. } |w(i)| = 1, \forall i = 1, 2, \ldots, K N_r, \]
(38)

where
\[ f_5(w) = \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 \sum_{b=1}^{B} w_{b,k}^H H_{b,k} F_{\text{RF},b} f_{\text{BB},b} - \Re \left( \sum_{b=1}^{B} \sum_{k=1}^{K} 2 \sqrt{\mu_k} \xi_k^H w_{b,k}^H H_{b,k} F_{\text{RF},b} f_{\text{BB},b,k} \right). \]
(39)

After some algebraic transformations, \( f_5(w) \) can be simplified as a more compact form. Consequently, problem (38) can be equivalently written as
\[ \min_w w^H P w - \Re \left( 2 w^H q \right) \]
\[ \text{s.t. } |w(i)| = 1, \forall i = 1, 2, \ldots, K N_r, \]
(40)

where
\[ \tilde{q}_k = \sum_{b=1}^{B} \sqrt{\mu_k} \xi_k^H H_{b,k} F_{\text{RF},b} f_{\text{BB},b,k}, \]
\[ \tilde{p}_k = \sum_{j=1}^{K} |\xi_j|^2 p_{k,j}^H, \]
\[ p_{k,j} = \sum_{b=1}^{B} H_{b,k} F_{\text{RF},b} f_{\text{BB},b,j}, \]
\[ q = [\tilde{q}_1^T, \tilde{q}_2^T, \ldots, \tilde{q}_K^T]^T, \]
\[ P = \text{diag}(\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_K). \]
(41)

Notably, problem (40) is a non-convex QCQP problem with quadratic objective function and unit-modulus constraints, which is similar to problem (28). In this regard, the MO algorithm can be performed to seek a high-quality solution.

D. Fix (F, w, Λ, ξ) and Solve Φ

Given the optimized set (F, w, Λ, ξ), the remaining compound terms in (14) is only related to Φ. Hence, the passive BF design for R RISs can be written as
\[ \min_{\Phi} f_6(\Phi) \]
\[ \text{s.t. } |\Phi(\cdot)| = 1 \forall i_2 = 1, 2, \ldots, RM, \]
(42)

where
\[ f_6(\Phi) = -\Re \left( \sum_{k=1}^{K} \sum_{b=1}^{B} 2 \sqrt{\mu_k} \xi_k^H w_{k}^H V_k \Phi G_b F_{\text{RF},b} f_{\text{BB},b,k} \right) \]
\[ + \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 \left( w_{k}^H V_k \Phi F_{\text{RF},b} f_{\text{BB},b,k} + E_{k,j} \right)^2 \]
and \( E_{k,j} = \sum_{k=1}^{B} w_{k}^H H_{b,k} F_{\text{RF},b} f_{\text{BB},b,k} \).

Base on problem (42), \( f_6(\Phi) \) is still laborious to be handled. By introducing \( \varphi = \Phi I_{RM} \), problem (42) can be further simplified as
\[ \min_{\varphi} \varphi^H Z \varphi + \varphi^H \kappa + \kappa^H \varphi \]
\[ \text{s.t. } |\varphi(\cdot)| = 1 \forall i_2 = 1, 2, \ldots, RM, \]
(44)

where
\[ u_{b,k} = \sqrt{\mu_k} \xi_k^H \text{diag}(w_{b}^H V_k) G_b F_{\text{RF},b} f_{\text{BB},b,k}, \]
\[ v_{b,k,j} = \text{diag}(w_{b}^H V_k) G_b F_{\text{RF},b} f_{\text{BB},b,j}, \]
\[ \kappa = \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 E_{k,j} v_{b,k,j} - \sum_{k=1}^{K} u_{b,k}^H, \]
\[ Z = \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 \left( \sum_{b=1}^{B} v_{b,k,j}^H \right) \left( \sum_{b=1}^{B} v_{b,k,j}^T \right). \]
(45)

To solve the convex problem (44) with convex constraints, the PDS method can be leveraged to get a desired solution in a computation-efficient way [21], [39]. Furthermore, we continue the iterative process by optimizing the variable set (F, Φ, w, Λ, ξ) until the AO algorithm converges. The overall procedure for CBF design is summarized in Algorithm 1.

IV. BASE STATION SELECTION FOR RIS-AIDED P-CF-mMIMO SYSTEM

The fully-connected CF-mMIMO architecture results in inevitably extravagant network costs due to a vast number of communication links. To suppress resultant communication costs, we develop a novel P-CF-mMIMO network architecture that selects partial communication links among BSs and users in this section. Concretely, the CPU is able to determine
where one BS can connect each user based on channel conditions. Moreover, we introduce the integer programming to solve the BS selection problem with binary integer constraints so as to assist the CPU in making decisions.

A. Problem Formulation for BS Selection

With respect to the RIS-aided P-CF-mMIMO system, the transmitted signal from the $b$th BS can be expressed as

$$\tilde{x}_b = \mathbf{F}^{RF}_b \mathbf{F}^{BB}_b \mathbf{A}_b \mathbf{s} = \sum_{k=1}^{K} \mathbf{F}^{RF}_b \mathbf{F}^{BB}_b \mathbf{f}_{b,k} \tau_{b,k} s_k,$$

(46)

where $\mathbf{A}_b = \text{diag}(\tau_{b,1}, \tau_{b,2}, \ldots, \tau_{b,K})$ is a diagonal matrix determining network connections among users and the $b$th BS and the $k$th entry satisfies $\tau_{b,k} \in \{0, 1\}$. More precisely, $\tau_{b,k} = 1$ if the $b$th BS communicates with the $k$th user, otherwise $\tau_{b,k} = 0$, which can be dynamically controlled by the CPU. We define $\mathcal{B}_b = \{1, 2, \ldots, B_b\}$ ($\forall b \in \mathcal{B}$) as the subset of users that is served by the $b$th BS. Similarly, $\mathcal{B}_k = \{1, 2, \ldots, K\}$ ($\forall k \in \mathcal{K}$) is defined as the subset of BSs connecting to the $k$th user. Based on the number of connected links between users and BSs, we can further get the condition of $\sum_{b \in \mathcal{B}} K_b = \sum_{k \in \mathcal{K}} B_k$.

To better evaluate the influence brought by the number of communication links, we define a new metric for the proposed P-CF-mMIMO system as network connection ratio (NCR), which can be further expressed as

$$\alpha = \frac{\sum_{b \in \mathcal{B}} K_b}{BK},$$

(47)

where $\alpha \in (0, 1)$. This metric plays a significant role in characterizing communication costs, including the integration of latency, computational complexity, signaling overhead, backhaul overhead and energy consumption. In fact, NCR determines the total number of communication links among BSs and users. For instance, when $\alpha$ is smaller, the communication costs decline as well as WSR performance. Given larger $\alpha$, both communication costs and WSR performance raise accordingly. Hence, by adjusting $\alpha$, the P-CF-mMIMO can make a better tradeoff between performance and communication costs.

For the RIS-aided P-CF-mMIMO network, the SINR of the $k$th user can be calculated as

$$\gamma_k = \frac{\left| \sum_{b=1}^{B} \mathbf{w}_b^H \mathbf{H}_{b,k} \mathbf{F}^{RF}_b \mathbf{F}^{BB}_b \mathbf{f}_{b,k} \tau_{b,k} \right|^2}{\sum_{b=1}^{B} \sum_{j \neq k} \mathbf{w}_b^H \mathbf{H}_{b,j} \mathbf{F}^{RF}_b \mathbf{F}^{BB}_b \mathbf{f}_{b,j} \tau_{b,j} + \sigma^2}.$$  

(48)

Then, the WSR for $K$ users can be written as

$$\hat{R}_{sum} = \sum_{k=1}^{K} \omega_k \log_2 (1 + \gamma_k).$$

(49)

Similar to the previous section, by using the FP technique [31], the objective function (49) of the WSR optimization problem can be rewritten as (49).

Based on the transformed function (50), shown at the bottom of the page, the joint optimization of CBF design and BS selection for RIS-aided P-CF-mMIMO systems can be formulated as

$$\min_{\mathbf{F}, \hat{\mathbf{F}}, \hat{\mathbf{w}}, \hat{\mathbf{A}}} \hat{f} \left( \hat{\mathbf{F}}, \hat{\mathbf{w}}, \hat{\mathbf{A}} \right),$$

(51a)

s.t. $\left| \mathbf{F}^{RF}_b \mathbf{F}^{BB}_b \mathbf{A}_b \right|_F \leq P_b, \forall b \in \mathcal{B},$

(51b)

$\left| \mathbf{F}^{RF}_b \mathbf{f}_{i,b} \right|^2 = 1, \forall i, b \in \mathcal{B},$

(51c)

$\left| \hat{\mathbf{w}} \mathbf{f}_{i,k} \right|^2 \leq 1, \forall i, k \in \mathcal{K},$

(51d)

$\left| \mathbf{A}_b \right|_F = \alpha BK, \mathbf{A}_b(i_4, i_4) \in \{0, 1\}, \forall i_4,$

(51f)

where $\mathbf{A} = \{\mathbf{A}_b| \forall b \in \mathcal{B}\}, \hat{\mathbf{F}} = \{\hat{\mathbf{F}}_b| \forall b \in \mathcal{B}\}$ and $\hat{\mathbf{b}} = \mathbf{F}^{RF}_b \mathbf{F}^{BB}_b \mathbf{A}_b$. The main difference between problem (15) and problem (51) is extra integer constraints (51b) and (51f). Hence, we pay more attention to the BS selection problem that incorporates integer constraints.

B. Fix ($\hat{\mathbf{F}}, \hat{\mathbf{w}}, \hat{\mathbf{A}}$) and Solve ($\hat{\lambda}_k, \hat{\xi}_k$)

Similar to (16), the auxiliary variables ($\hat{\lambda}_k, \hat{\xi}_k$) for the $k$th user can be respectively given by

$$\hat{\lambda}_k^{*} = \hat{\gamma}_k,$$

(52a)

$$\hat{\xi}_k = \frac{1}{2} \left( \sum_{b=1}^{B} \mathbf{w}_b^H \mathbf{H}_{b,k} \mathbf{F}^{RF}_b \mathbf{F}^{BB}_b \mathbf{f}_{b,k} \tau_{b,k} - \sum_{b=1}^{B} \sum_{j \neq k} \mathbf{w}_b^H \mathbf{H}_{b,j} \mathbf{F}^{RF}_b \mathbf{F}^{BB}_b \mathbf{f}_{b,j} \tau_{b,j} \right)^2.$$  

(50)
\[
\hat{\xi}_k = \sqrt{\mu_k} \left( \sum_{b=1}^{B} w_k^H H_{b,k} F_{RF,b} f_{BB,b,k} T_b,k \right) + \frac{\sigma^2}{\sum_{j=1}^{K} \left( \sum_{b=1}^{B} w_k^H H_{b,k} F_{RF,b} f_{BB,b,j} T_b,j \right)^2} \right) + \sigma^2.
\] (52b)

C. Fix (\(\hat{\Phi}, \hat{\w}, \hat{\lambda}, \hat{\xi}\)) and Solve \(\hat{F}\)

According to the P-CF-MIMO system, the BS selection matrix \(\Lambda_b\) affects the power allocation and the HBF design. Thereby, the HBF design at the \(b\)th BS can be formulated as

\[
\begin{align*}
\min_{F_{RF,b}, F_{BB,b}, \Lambda_b} \quad & f_1(F_{RF,b}, F_{BB,b}, \Lambda_b) \quad (53a) \\
\text{s.t.} \quad & (51b), (51c), (51f), \\
& (53b)
\end{align*}
\]

where

\[
\begin{align*}
\hat{f}_1(F_{RF,b}, F_{BB,b}, \Lambda_b) &= \hat{D}_{b,k,j} - 2\Re \left( \sum_{k=1}^{K} \sqrt{\mu_k} \xi_k w_k^H H_{b,k} F_{RF,b} f_{BB,b,k} T_b,k \right) \\
&+ \sum_{k,j=1}^{K} \left| \xi_k \right|^2 w_k^H H_{b,k} F_{RF,b} f_{BB,b,j} T_b,j + \hat{C}_{b,k,j}^2 \right), \\
& (54)
\end{align*}
\]

and \(\hat{C}_{b,k,j}\) and \(\hat{D}_{b,k,j}\) are irrelevant terms for optimizing \(F_{RF,b}\) and \(F_{BB,b}\).

By considering an auxiliary variable \(\tilde{F}_b = F_{RF,b} F_{BB,b} \Lambda_b\), the augmented Lagrangian function for problem (55) can be formulated as

\[
L_3 \left( \tilde{F}_b, F_{RF,b}, F_{BB,b}, \Lambda_b, \Delta \right) = \hat{f}_1(F_{RF,b}, F_{BB,b}, \Lambda_b) \\
+ \frac{\rho_2}{2} \left\| \tilde{F}_b - F_{RF,b} F_{BB,b} \Lambda_b + \frac{\Delta}{\rho_2} \right\|^2_F - \frac{1}{2} (\Delta^H \Delta) \right) \right) + \frac{\rho_2}{2} \left\| F_{BB,b} \right\|^2_F \\
(55)
\]

Based on (55), the ADMM algorithm can be utilized to solve \(\left( \tilde{F}_b, F_{RF,b}, F_{BB,b}, \Lambda_b \right)\) in an iterative manner, which can be specifically expressed as

\[
\begin{align*}
\hat{F}_{RF,b}^{t+1} &= \arg \min_{F_{RF,b}} L_3 \left( \tilde{F}_b, F_{RF,b}, F_{BB,b}^{t+1}, \Lambda_b^{t+1}, \Delta^{t+1} \right) \\
&\quad \text{s.t. } \left\| \tilde{F}_b \right\|_F \leq P_b, \\
(56a) \\
F_{BB,b}^{t+1} &= \arg \min_{F_{BB,b}} L_3 \left( \tilde{F}_b, F_{RF,b}, F_{BB,b}^{t+1}, \Lambda_b^{t+1}, \Delta^{t+1} \right) \\
&\quad \text{s.t. } \left| F_{BB,b} \right| (i_1, j_1) = 1, \forall i_1, j_1, \\
(56b) \\
\Lambda_b^{t+1} &= \arg \min_{\Lambda_b} L_3 \left( \tilde{F}_b, F_{RF,b}, F_{BB,b}^{t+1}, \Lambda_b, \Delta^{t+1} \right) \\
&\quad \text{s.t. } \left\| \Lambda_b \right\|_F = K_b, \Lambda_b(i_4, i_4) \in \{0, 1\}, \forall i_4, \\
(56c) \\
\Delta^{t+1} &= \Delta^t + \rho_2 \left( \tilde{F}_b - F_{RF,b} F_{BB,b}^{t+1} \right), \\
(56d)
\end{align*}
\]

The subproblems (56a)-(56c) in the P-CF-MIMO system are similar to aforementioned subproblems (21a)-(21c) existing in the CF-MIMO system. Hence, the detailed steps for optimizing \(\hat{F}_{RF,b}, F_{BB,b}^{t+1}, \Lambda_b^{t+1}\) are omitted for brevity, and we concentrate on the BS selection problem.

D. Relaxed Linear Approximation Method

Given the optimized \(\hat{F}_{RF,b}, F_{BB,b}^{t+1}, \Lambda_b^{t+1}\), the BS selection problem (56d) related to \(\Lambda_b\) can be reformulated as

\[
\begin{align*}
\Lambda_b^{t+1} &= \arg \min_{\Lambda_b} f_2(\Lambda_b) \\
&\quad \text{s.t. } \left\| \Lambda_b \right\|_F = K_b, \\
&\quad \Lambda_b(i_4, i_4) \in \{0, 1\}, \forall i_4, \\
(57)
\end{align*}
\]

and

\[
\begin{align*}
\hat{f}_2(\Lambda_b) &= \sum_{k=1}^{K} \sum_{j=1}^{J} \left| \xi_k \right|^2 w_k^H H_{b,k} F_{RF,b}^{t+1} F_{BB,b}^{t+1} T_b,j + \hat{C}_{b,k,j}^2 \\
&- 2\Re \left( \sum_{k=1}^{K} \sqrt{\mu_k} \xi_k w_k^H H_{b,k} F_{RF,b}^{t+1} F_{BB,b}^{t+1} T_b,k \right) \\
+ \rho_2 \left\| F_{BB,b}^{t+1} - F_{RF,b}^{t+1} \right\|^2_F \\
(58)
\end{align*}
\]

To simplify problem (57), we define a new vector \(\tau_b = \Lambda_b 1_K\). As a consequence, the original problem (57) can be recast as

\[
\begin{align*}
\min_{\tau_b} \tau_b^T U_b \tau_b + \left\| r_b \right\|_T^2 \\
&\quad \text{s.t. } I_K \tau_b \geq z_b, \\
&\quad \tau_b \in \{0, 1\}^K, \\
(59a) \\
\end{align*}
\]

where

\[
\begin{align*}
\tau_b^T = 2\Re (1_H^T - g_b^T T + \rho_2 \hat{\tau}_b^T T M_b^T), \\
M_b^T = [\ldots, M_b^T(:, (k-1)K + k), \ldots], k \in K, \\
z_b = [z_{b,1}, \ldots, z_{b,K}]^T, z_{b,k} = 1 - \sum_{p \in B \setminus \{k\}} \tau_p, \\
U_b^T = L_b^T + \frac{\rho_2}{2} \hat{M}_b^T H_b^T, \\
\hat{M}_b^T = [\hat{M}_b^T(:, 1)], \\
g_b = [g_{b,1}, \ldots, g_{b,K}]^T, \\
\hat{L}_b^T = \sum_{k=1}^{K} \left[ \xi_k \right]^2 \Psi_{b,k,j} \Psi_{b,k,j}^T, \\
L_b^T = \text{diag}(\hat{L}_{b,1}, \ldots, \hat{L}_{b,K}), \\
(60)
\end{align*}
\]

and the specific derivation process for problem (59) is provided in Appendix B.

Problem (59) is a BIQP problem with linear integer constraints (59b) and (59c). In detail, (59b) stems from the fact that each user is served by at least one BS while (59c) implies that \(K_b\) users are served by the \(b\)th BS concurrently. To this end, the RLA method is proposed to solve problem (59), which
replaces each quadratic term in the objective function by a new binary variable and two inequality constraints. Then, problem (59) can be equivalently written as

\[
\min_{\tau_b} \sum_{i=1}^{K} \sum_{j=1, j \neq i}^{K} u_{b,i,j}^t \tau_{b,i} \tau_{b,j} + \sum_{i=1}^{K} r_{b,i}^t \tau_{b,i} \\
\text{s.t. } \tau_b \in G_b,
\]  
(61)

where \(u_{b,i,j}^t\) denotes \((i,j)\)th entry of \(U_b^t\), \(r_{b,i}^t\) denotes \(i\)th entry of \(r_b^t\), and

\[
G_b = \left\{ \tau_b \in \{0, 1\}^K : \sum_{i=1}^{K} \tau_{b,i} = K_b, \tau_{b,i} \geq z_{b,i}, i \in \mathcal{K} \right\}.
\]  
(62)

To eliminate the quadratic terms in problem (61), we use a new variable \(\zeta_{b,i,j}\) to substitute for each product \(\tau_{b,i} \tau_{b,j}\). For the special case \(i = j\), we can get \(\zeta_{b,i,i} = \tau_{b,i} \tau_{b,i} = \tau_{b,i}\) due to the constraint \(\tau_{b,i} \in \{0, 1\}, i \in \mathcal{K}\). By introducing \(\zeta_{b,i,j} = \tau_{b,i} \tau_{b,j}\), the extra constraints can be expressed as

\[
\zeta_{b,i,j} = \max \left\{ \tau_{b,i} + \tau_{b,j} - 1, 0 \right\}, \quad (63a)
\]

\[
\zeta_{b,i,j} = \min \left\{ \tau_{b,i}, \tau_{b,j} \right\}, \quad (63b)
\]

Considering that \(U_b \in \mathbb{C}^{K \times K}\) is symmetric and positive semidefinite, problem (61) can be reformulated as

\[
\min_{\tau_b} \sum_{i=1}^{K} \sum_{j=1, j \neq i}^{K} 2u_{b,i,j}^t \zeta_{b,i,j} + \sum_{i=1}^{K} \left( u_{b,i,i}^t + r_{b,i}^t \right) \tau_{b,i} \\
\text{s.t. } \tau_b \in G_b,
\]  
(64a)

\[
\zeta_{b,i,j} \leq \tau_{b,i}, i, j \in \mathcal{K}, i < j, \quad (64b)
\]

\[
\zeta_{b,i,j} \leq \tau_{b,j}, i, j \in \mathcal{K}, i < j, \quad (64c)
\]

\[
\zeta_{b,i,j} \geq \tau_{b,i} + \tau_{b,j} - 1, i, j \in \mathcal{K}, i < j, \quad (64d)
\]

\[
\zeta_{b,i,j} \in \{0, 1\}, i, j \in \mathcal{K}, i < j. \quad (64f)
\]

The standard linearization process in problem (64) neglects some intrinsic constraints [40], so we propose an incremental RLA method that considers stronger constraints in the following. Specifically, given \(\tau_{b,i} = 1 (\forall i \in \mathcal{K})\), we can get \(\zeta_{b,i,j} = \tau_{b,j} \forall j \in \mathcal{K}\), which further implies that \(\zeta_{b,i,1}, \zeta_{b,i,2}, \ldots, \zeta_{b,i,K} \in \mathcal{G}_b\) in the case of \(\tau_{b,i} = 1\). Conversely, if \(\tau_{b,i} = 0\), then \(\zeta_{b,i,j} = 0 (\forall j \in \mathcal{K})\). Hence, the constraint related to \(\tau_{b,i}\) (\(\forall i \in \mathcal{K}\)) can be represented as

\[
D_b (\tau_{b,i}) = \begin{cases} 
0_K, & \tau_{b,i} = 0 \\
\tau_{b,i}, & \tau_{b,i} = 1 
\end{cases}
\]  
(65)

where \(D_b\) in the case of \(\tau_{b,i} = 1\) can be given by

\[
D_b (\tau_{b,i}) = \left\{ (\zeta_{b,i,1}, \ldots, \zeta_{b,i,K}) \in \{0, 1\}^K : \sum_{j \in \mathcal{K}} \zeta_{b,i,j} = K_b \tau_{b,i}, \zeta_{b,i,j} \geq z_{b,j} \tau_{b,i} \right\}. \quad (66)
\]

Eventually, the BS selection problem for the P-CF-mMIMO system can be formulated as

\[
\min_{\tau_b} \sum_{i=1}^{K} \sum_{j=1, j \neq i}^{K} 2u_{b,i,j}^t \zeta_{b,i,j} + \sum_{i=1}^{K} \left( u_{b,i,i}^t + r_{b,i}^t \right) \tau_{b,i} \\
\text{s.t. } (\zeta_{b,i,1}, \zeta_{b,i,2}, \ldots, \zeta_{b,i,K}) \in D_b (\tau_{b,i}), \forall i \in \mathcal{K},
\]  
(67a)

As discussed above, our proposed RLA algorithm transforms the BIQP problem into binary integer linear programming (BILP) problem. Remarkably, the BILP problem can be easily handled by existing general-purpose solvers [41], including but not limited to branch-and-bound, cutting plane and Lagrangean duality.

E. Fix \((\hat{\Phi}, \hat{\lambda}, \hat{\xi})\) and Solve \((\hat{\omega}, \hat{\Phi})\)

Due to the utilization of BS selection mechanism in the RIS-aided P-CF-mMIMO system, the BF optimization problems for solving \(\hat{\Phi}\) and \(\hat{\omega}\) also need to be reformulated. According to problem (51), the combiner design at users can be formulated as

\[
\min_{\hat{\omega}} \hat{\omega}^H \hat{\Phi} \hat{\omega} - \Re \left( 2\hat{\omega}^H \hat{\xi} \right) \\
\text{s.t. } |\hat{\omega} (i)|^2 = 1, \forall i = 1, 2, \ldots, KN_r. \quad (68)
\]

where

\[
\hat{\Phi}_k = \sum_{b=1}^{B} \sqrt{H_{b,k}^H H_{b,k}} F_{RF,b} F_{BB,b,k} \tau_{b,k} ,
\]

\[
\hat{\Phi}_{b,k} = \sum_{b=1}^{B} H_{b,k} F_{RF,b} F_{BB,b,k} \tau_{b,k} , \hat{\Phi} = \left[ \hat{\Phi}_1^T, \ldots, \hat{\Phi}_{K}^T \right]^T ,
\]

\[
\hat{\Phi}_k = \sum_{b=1}^{B} \left| \xi_b \right|^2 \hat{\Phi}_{b,k} \hat{\Phi}_{b,k}^H , \hat{\Phi} = \text{diag} \left( \hat{\Phi}_1, \ldots, \hat{\Phi}_K \right). \quad (69)
\]

Problem (68) is a QCQP problem with unit-modulus constraints and we employ the MO algorithm to solve such a constrained problem. To optimize \(\hat{\Phi}\) in the P-CF-mMIMO system, we extract the related terms from problem (51). By defining \(\hat{\phi} = \Phi I_{RM}\), the optimization problem for designing \(\hat{\Phi}\) at RISs can be expressed as

\[
\min_{\hat{\phi}} \hat{\phi}^H \hat{\Phi} \hat{\phi} + \hat{\phi}^H \hat{\kappa} + \hat{\kappa}^H \hat{\phi} \\
\text{s.t. } |\hat{\phi} (i_2)|^2 \leq 1, \forall i_2 = 1, 2, \ldots, RM, \quad (70)
\]

where

\[
\hat{E}_{b,k,j} = \sum_{b=1}^{B} w_k^H \hat{H}_{b,k} F_{RF,b} F_{BB,b,k} \tau_{b,k},
\]

\[
\hat{u}_{b,k} = \sqrt{\mu \xi_b^H} \text{diag} (w_k^H V_k) G_{b,k} F_{RF,b} F_{BB,b,k} \tau_{b,k},
\]

\[
\hat{v}_{b,k,j} = \text{diag} (w_k^H V_k) G_{b,k} F_{RF,b} F_{BB,b,k} \tau_{b,k},
\]

\[
\hat{\kappa} = \sum_{k=1}^{K} \sum_{j=1}^{K} \left| \xi_b \right|^2 \hat{E}_{b,k,j} \hat{v}_{b,k,j}^* - \sum_{k=1}^{K} \sum_{b=1}^{B} \hat{u}_{b,k}^* ,
\]

\[
\hat{\Phi} = \sum_{k=1}^{K} \sum_{j=1}^{K} \left| \xi_b \right|^2 \left( \sum_{b=1}^{B} \hat{v}_{b,k,j}^* \right)^2 \left( \sum_{b=1}^{B} \hat{v}_{b,k,j} \right). \quad (71)
\]

It is worth noting that \(\hat{\Phi}\) is positive semidefinite and the constraint set is convex. Therefore, similar to problem (44), the convex problem (71) can be well solved by the PDS method.
F. Complexity and Convergence Analyses

The overall computational complexity of the AO algorithm in the RIS-aided CF-mMIMO system stems from updating the variable set \((\mathbf{F}, \Phi, \mathbf{w}, \lambda, \xi)\). Firstly, the complexity of updating auxiliary variables \((\lambda, \xi)\) is determined by (16). Secondly, the complexity of HBF design at BSs comes from the ADMM algorithm, e.g., (21a)-(21d). Specifically, since \(\mathbf{F}_b\), \(\mathbf{F}_{BB,b}\) and \(\Delta\) have close-form formulas, the complexity of updating \((\mathbf{F}_b, \mathbf{F}_{BB,b}, \Delta)\) can be directly calculated by (25), (37) and (21d), respectively. The complexity of updating \(\mathbf{F}_{RF,b}\) comes from the MO method as well as the combining design of \(\mathbf{w}\) at users, and the complexity of optimizing \(\Phi\) at RISs is caused by the PDS method. The extra complexity in the P-CF-mMIMO system is from the RLA method, which adopts the cutting plane to solve the BILP problem [42]. To sum up, the detailed complexity of CF-mMIMO and P-CF-mMIMO systems is shown in Table I, where \(t_{\text{max}}\) and \(I_{\text{max}}\) denote the maximum iteration number of ADMM and AO, respectively.

To prove the convergence of the AO algorithm in RIS-aided CF-mMIMO systems, we denote the obtained solution of problem (51) at \(i\)th iteration as \(\hat{\mathbf{F}}^{(i)}, \hat{\Phi}^{(i)}, \hat{\mathbf{w}}^{(i)}\). Thus, the iterative process of the proposed AO algorithm can be expressed as

\[
\hat{f}\left(\hat{\mathbf{F}}^{(i)}, \hat{\Phi}^{(i)}, \hat{\mathbf{w}}^{(i)}\right) \geq \hat{f}\left(\hat{\mathbf{F}}^{(i)+1}, \hat{\Phi}^{(i)+1}, \hat{\mathbf{w}}^{(i)+1}\right) \geq \hat{f}\left(\hat{\mathbf{F}}^{(i)+1}, \hat{\Phi}^{(i)+1}, \hat{\mathbf{w}}^{(i)+1}\right),
\]

where the inequalities (a), (b) and (c) come from executing ADMM, MO and PDS methods, whose convergence properties have been demonstrated in [33], [38], and [39], respectively. Hence, the WSR is non-increasing over the iterations. Since the WSR of the practical P-CF-mMIMO systems is finite, the optimized solution has a lower bound. Consequently, the sequence \(\hat{f}\left(\hat{\mathbf{F}}^{(i)}, \hat{\Phi}^{(i)}, \hat{\mathbf{w}}^{(i)}\right)\) at least converges to a locally optimal solution of original problem (51), but cannot guarantee a globally optimal one.

V. Simulation Results

In this section, simulation results are provided to demonstrate the effectiveness of proposed CBF algorithms in the RIS-aided CF-mMIMO network. Meanwhile, simulation results also reveals that, by using the BS selection strategy, the proposed P-CF-mMIMO network makes a better tradeoff between communication costs and WSR performance compared with the fully-connected CF-mMIMO network.

A. Simulation Setup

We consider a practical RIS-aided mmWave CF-mMIMO scenario as shown in Fig. 3, where the physical positions of BSs, RISs and users are deployed within the \(300 \times 300\) \(m^2\) square area following the 2D Poisson distribution. The heights of BSs, RISs and users are set as 6 m, 4 m and 1.8 m, respectively. Given the 3D coordinate information, the LoS distance between any two points can be obtained. The number of array elements for each BS, each RIS and each user is set as \(N_c = 32\), \(M = 64\) and \(N_r = 8\), respectively. Considering the sparse feature of mmWave channel, each communication link (e.g., \(\tilde{\mathbf{H}}_{b,k}\), \(\tilde{\mathbf{G}}_{b,r}\) and \(\tilde{\mathbf{V}}_{r,k}\)) is composed of \(L = 4\) propagation paths, including \(1\) LoS path and \(3\) NLoS paths. More specifically, the complex gain of the LoS path is calculated by free space loss and the complex gains of the NLoS paths are calculated by multipath propagation loss [43]. The AoAs and AoDs of the mmWave channels are selected from \((0, 2\pi)\) following the uniform distribution. In addition, the working frequency of the RIS-aided CF-mMIMO system is set as \(f_c = 28\) GHz and the received noise power is \(\sigma^2 = -85\) dBm. For simplification, the maximum transmit power for all \(B\) BSs is set as \(P_b = P_{\text{max}}, \forall b \in \mathcal{B}\).

B. Robustness Performance Evaluation

Considering the statistical CSI error model, the diverse variances of \(\text{vec}(\Delta \tilde{\mathbf{H}}_{b,k})\), \(\text{vec}(\Delta \tilde{\mathbf{G}}_{b,r})\) and \(\text{vec}(\Delta \tilde{\mathbf{V}}_{r,k})\) are defined as \(\delta_{b,k}^2 = \|\tilde{\mathbf{H}}_{b,k}\|^2_F\), \(\delta_{b,r}^2 = \|\tilde{\mathbf{G}}_{b,r}\|^2_F\) and \(\delta_{r,k}^2 = \|\tilde{\mathbf{V}}_{r,k}\|^2_F\), respectively.
the random BS selection, indicating that the integer programming couples well with the BS selection problem. Remarkably, the RLA empowered RIS-aided P-CF-mMIMO system with $\alpha = 0.75$ is able to achieve basically the same WSR compared with the RIS-aided CF-mMIMO system (e.g., 0.72 bps/Hz gap), which proves the effectiveness of our proposed RLA based BS selection scheme.

Fig. 6 explores the cumulative distribution function (CDF) of rate per user with different user densities in both CF-mMIMO and P-CF-mMIMO systems, where all the users are randomly distributed in the predefined area. As a whole, Fig. 6 (a) and Fig. 6 (b) have a similar curve tendency. In Fig. 6 (a), when it comes to the CF-mMIMO system, the FD-BF scheme achieves the best per user rate performance, and outstrips the proposed CBF scheme with ideal RISs in terms of 95%-likely per user rate (2.29 bps/Hz versus 1.98 bps/Hz) and basically the same median per user rate, implying that these performance gaps can be negligible. Moreover, the curves with random RISs in Fig. 6 (a) endure obvious performance loss compared to the case with ideal RISs, which indicates that RISs produce a dramatic impact on enhancing the network capacity. Another important observation in Fig. 6 (a) is that the CDF of the RLA empowered P-CF-mMIMO system with $\alpha = 0.75$ is basically consistent with the CF-mMIMO system, where the performance gap of 95%-likely per user rate and median rate are 0.17 bps/Hz and 0.12 bps/Hz, respectively. The random BS selection scheme has the worst per user rate performance, and in turn verifies the effectiveness of our proposed RLA based BS selection scheme. Moreover, from Fig. 6 (a) and Fig. 6 (b), we can observe that smaller user density case with $K = 4$ owns much better per user rate than the large user density case with $K = 8$. The intuition is that more users certainly bring higher IUIs and worsen the per user rate.

### C. WSR Performance Comparisons

Fig. 5 investigates WSR comparisons versus the transmit power $P_{\text{max}}$ for both CF-mMIMO and P-CF-mMIMO systems, where $B = 6$, $R = 3$ and $K = 4$. From Fig. 5 (a), it can be observed that the WSR of all the considered schemes grows rapidly with the increase of $P_{\text{max}}$ at BSs. Particularly, the WSR of our proposed CBF with ideal RISs approaches the fully-digital BF (FD-BF) that owns the optimal performance. With regard to the non-ideal RIS cases, the proposed CBF using RISs with 2 bits achieves the similar WSR performance compared with the ideal RIS case, while deploying RISs with 1 bit suffers from about 1.14 bps/Hz performance penalty at $P = -10$ dBm. To validate the benefit of RISs, we also present the simulation curve without RISs that has worse WSR performance than RIS-aided schemes. Fig. 5 (a) also reveals that LoS path is essential for high frequency communications. Fig. 5 (b) depicts the WSR performance of the RIS-aided P-CF-mMIMO system that can be controlled by adjusting the parameter $\alpha$. More precisely, as the value of $\alpha$ increases, the WSR performance of the RIS-aided P-CF-mMIMO system becomes better, and the communication costs grow accordingly. Under the same setting of $\alpha$, our proposed RLA based BS selection scheme greatly outperforms

### D. Simulation Setup

Fig. 7 illustrates the sum-rate comparisons of considered schemes with the increasing number of RIS elements. Except for the CF-mMIMO system with random RISs and the P-CF-mMIMO system with random BS selection, the WSR of all the remaining schemes continuously grows as the number of reflecting elements rises. In particular, RISs have a greater impact on the CF-mMIMO system without LoS path. The reason is that the case without LoS path mainly depends on the BS-RIS channel and RIS-user channel to achieve high WSR performances, and thus RISs become more essential for the case without LoS path. For instance, the WSR gap between the CBF with ideal RISs and the case without LoS is around 8.01 bps/Hz at $N_{\text{RIS}} = 32$ and 4.35 bps/Hz at $N_{\text{RIS}} = 96$, respectively. Furthermore, the performance gap between RLA based BS selection scheme and the random BS selection scheme grows larger with the increasing number of RIS elements, which also proves the significance of our proposed RLA based scheme for the RIS-aided P-CF-mMIMO system.

Fig. 8 presents the convergence performance of considered schemes versus the number of iterations in both CF-mMIMO
Fig. 5. WSR versus transmit power $P_{\text{max}}$: (a) CF-mMIMO; (b) P-CF-mMIMO.

Fig. 6. CDF versus rate per user: (a) $K = 4$; (b) $K = 8$.

Fig. 7. WSR versus the number of reflecting elements $N_{\text{RIS}}$.

Fig. 8. Sum-rate versus the number of iterations $I_{\text{max}}$.

and P-CF-mMIMO systems, where the parameter setup is $B = 6$, $R = 3$ and $K = 4$. It can be easily observed that the WSR performance of all the considered schemes rises along with the increasing number of iterations, and then gradually slows down until convergence. Particularly, the converged number of iterations for both CF-mMIMO and P-CF-mMIMO systems is around $I_{\text{max}} = 60$. Under the condition of $I_{\text{max}} = 60$, the WSR of the RLA based P-CF-mMIMO system with $\alpha = 0.75$ approaches the CBF scheme of the CF-mMIMO system with ideal RISs (about 0.42 bps/Hz gap). Nevertheless, at the beginning of the iterative process, the RLA based scheme converges slower than the FD-BF and proposed CBF schemes. This phenomenon stems from that the P-CF-mMIMO system has to optimize the extra BS selection matrix $\Lambda_b$ and thus results in slower network convergence. In addition, compared with random BS selection scheme, the RLA based BS selection scheme reaps obvious WSR improvement. By choosing desired $\alpha$, our proposed RLA...
based P-CF-mMIMO system is able to balance communication costs and WSR performances.

VI. CONCLUSION

This paper firstly investigated the active and passive BF design for the RIS-aided mmWave CF-mMIMO system. To decrease communication costs brought by fully-connected CF-mMIMO systems, we proposed a novel P-CF-mMIMO framework that a limited number of communication links were connected among users and BSs. The BS selection problem was modeled as a BIQP problem and then we developed an incremental RLA scheme to solve this integer programming problem. By selecting appropriate NCR, our proposed P-CF-mMIMO systems achieved a better compromise between performance and communication costs, which is able to meet diverse requirements in 6G wireless networks. Ultimately, considering the high network delay and severe signaling overhead brought by the centralized manner, our research work will concentrate on developing the fully decentralized CBF frameworks in the near future.

APPENDIX A

THE DETAILED PROOF FOR PROBLEM (28)

From problem (26) we note that the objective function is composed of three parts. To obtain problem (28), the first term of \( f_2(F_{RF,b}) \) can be reformulated as

\[
\sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 |a_{b,k,j}^t| f_{RF,b} + C_{b,k,j}^t
\]

\[
= \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 |a_{b,k,j}^t| f_{RF,b} + C_{b,k,j}^t
\]

\[
= f_{RF,b}^H \left( \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 a_{b,k,j}^t a_{b,k,j}^H \right) f_{RF,b} + f_{RF,b}^H \left( \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 a_{b,k,j}^t C_{b,k,j}^t \right) f_{RF,b} + f_{RF,b}^H \left( \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 C_{b,k,j}^t a_{b,k,j}^t \right) f_{RF,b} + f_{RF,b}^H \left( \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 C_{b,k,j}^t C_{b,k,j}^t \right)
\]
To handle the quadratic term, the last part in $\hat{f}_2(\mathbf{A}_b)$ can be rewritten as

$$
\sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 \left| \mathbf{w}_b^H \mathbf{H}_{b,k} \mathbf{F}_{j+1} \mathbf{b}_{j}^H \mathbf{b}_{b,k,j} \mathbf{H} \right|^2
$$

$$
= \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 \left( \mathbf{r}_{b,k,j}^H \mathbf{r}_{b,k,j} + \hat{C}_{b,k,j} \right)^2
$$

$$
= \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 \left( \mathbf{r}_{b,k,j}^H \mathbf{r}_{b,k,j} + \hat{C}_{b,k,j} \right) \left( \mathbf{r}_{b,k,j}^H \mathbf{r}_{b,k,j} + \hat{C}_{b,k,j} \right)
$$

$$
= \sum_{k=1}^{K} \sum_{j=1}^{K} \sum_{l=1}^{K} \sum_{m=1}^{K} |\xi_k|^2 \hat{C}_{b,k,j}^2 \mathbf{r}_{b,k,j}^H \mathbf{r}_{b,k,j} + \sum_{k=1}^{K} \sum_{j=1}^{K} \sum_{l=1}^{K} \sum_{m=1}^{K} \hat{C}_{b,k,j} \mathbf{r}_{b,k,j}^H \mathbf{r}_{b,k,j}
$$

$$
= \sum_{k=1}^{K} \sum_{j=1}^{K} |\xi_k|^2 \mathbf{r}_{b,k,j}^H \mathbf{r}_{b,k,j} + \sum_{k=1}^{K} \sum_{j=1}^{K} \hat{C}_{b,k,j} \mathbf{r}_{b,k,j}^H \mathbf{r}_{b,k,j}
$$

(80)

By merging congeners of (78), (79) and (80), we accomplish the proof for problem (59).

REFERENCES

[1] X. You et al., “Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts,” Sci. China Inf. Sci., vol. 64, no. 1, Jan. 2021, Art. no. 110301.

[2] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, “Cell-free massive MIMO versus small cells,” IEEE Trans. Wireless Commun., vol. 16, no. 3, pp. 1834–1850, Jan. 2017.

[3] 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, Jul. 2017.

[4] H. A. Anmar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, “User-centric cell-free massive MIMO networks: A survey of opportunities, challenges and solutions,” IEEE Commun. Surveys Tuts., vol. 24, no. 1, pp. 611–652, 1st Quart., 2022.

[5] T. C. Mai, H. Q. Ngo, and T. Q. Duong, “Downlink spectral efficiency of cell-free massive MIMO systems with multi-antenna users,” IEEE Trans. Commun., vol. 68, no. 8, pp. 4803–4815, Aug. 2020.

[6] J. Zhang, J. Zhang, D. W. K. Ng, S. Jin, and B. Ai, “Improving sum-rate of cell-free massive MIMO with expanded compute-and-forward,” IEEE Trans. Signal Process., vol. 70, pp. 202–215, 2022.

[7] E. Björnson and L. Sanguinetti, “Making cell-free massive MIMO competitive with MMSE processing and centralized implementation,” IEEE Trans. Wireless Commun., vol. 19, no. 1, pp. 77–90, Jan. 2020.

[8] M. Xiao et al., “Millimeter wave communications for future mobile networks,” IEEE J. Sel. Areas Commun., vol. 35, no. 9, pp. 1909–1935, Sep. 2017.

[9] Z. Chen et al., “A survey on terahertz communications,” China Commun., vol. 16, no. 2, pp. 1–35, Feb. 2019.

[10] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for wireless energy efficiency,” IEEE Trans. Wireless Commun., vol. 18, no. 8, pp. 4803–4815, Aug. 2019.

[11] X. Ma et al., “Intelligent reflecting surface enhanced indoor terahertz communication systems,” Nano Commun. Netw., vol. 24, May 2020, Art. no. 100284.

[12] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” IEEE Commun. Mag., vol. 58, no. 1, pp. 106–112, Nov. 2020.

[13] C. Pan et al., “Multicell MIMO communications relying on intelligent reflecting surfaces,” IEEE Trans. Wireless Commun., vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[14] L. Dai et al., “Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results,” IEEE Access, vol. 8, pp. 45913–45923, 2020.

[15] C. Huang et al., “Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends,” IEEE Wireless Commun., vol. 27, no. 5, pp. 118–125, Oct. 2020.

[16] E. Shi et al., “Wireless energy transfer in RIS-aided cell-free massive MIMO systems: Opportunities and challenges,” IEEE Commun. Mag., vol. 58, no. 3, pp. 26–32, Mar. 2020.
[36] C. Pan et al., “An overview of signal processing techniques for RIS/IRS-aided wireless systems,” IEEE J. Sel. Topics Signal Process., vol. 16, no. 5, pp. 883–917, Aug. 2022.

[37] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, “Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems,” IEEE J. Sel. Topics Signal Process., vol. 10, no. 3, pp. 485–500, Feb. 2016.

[38] B. Nicolas, An Introduction to Optimization on Smooth Manifolds. Cambridge, U.K.: Cambridge Univ. Press, Apr. 2022.

[39] S. Boyd, Subgradient Methods. May 2014, pp. 1–39. [Online]. Available: https://web.stanford.edu/class/ee392o/subgrad_method.pdf

[40] L. J. Watters, “Reduction of integer polynomial programming problems to zero-one linear programming problems,” Oper. Res., vol. 15, no. 6, pp. 1171–1174, 1967.

[41] L. Wolsey, Integer Programming. Hoboken, NJ, USA: Wiley, 1998.

[42] A. Basu et al., “Complexity of branch-and-bound and cutting planes in mixed-integer optimization,” Math. Program., vol. 198, pp. 787–810, Mar. 2022.

[43] M. R. Akdeniz et al., “Millimeter wave channel modeling and cellular capacity evaluation,” IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

Xinying Ma received the B.S. degree from Southwest Jiaotong University, Chengdu, China, in 2015, and the M.S. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 2018, where he is currently pursuing the Ph.D. degree with the National Key Laboratory of Science and Technology on Communications. From October 2021 to October 2022, he was a Visiting Student with the Department of Electrical Engineering and Computer Science, KTH Royal Institute of Technology. His research works have been cited more than 1600 times by Google Scholar. His current research interests include terahertz communications, millimeter wave communications, intelligent reflecting surface, and machine learning. He has served as a Reviewer for IEEE COMMUNICATIONS LETTER, IEEE ACCESS, IEEE TRANSACTIONS ON NANOTECHNOLOGY, IEEE TRANSACTIONS ON TERAHERTZ SCIENCE AND TECHNOLOGY, IEEE Wireless Communications Magazine, and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.

Deyou Zhang received the B.S. and M.S. degrees from the Harbin Institute of Technology, Harbin, China, in 2012 and 2014, respectively, and the Ph.D. degree from The University of Sydney, Sydney, Australia, in 2020. He is currently a Post-Doctoral Researcher with the Division of Information Science and Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. His research interests include millimeter wave communications, intelligent reflecting surface, and wireless federated learning.

Ming Xiao (Senior Member, IEEE) received the bachelor’s and master’s degrees in engineering from the University of Electronic Science and Technology of China, Chengdu, in 1997 and 2002, respectively, and the Ph.D. degree from the Chalmers University of Technology, Sweden, in November 2007. From 1997 to 1999, he was a Network and Software Engineer with China Telecom. From 2000 to 2002, he also held a position at Sichuan Communications Administration. Since November 2007, he has been with the Department of Information Science and Engineering, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Sweden, where he is currently an Associate Professor. He was an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS from 2012 to 2017 and IEEE WIRELESS COMMUNICATIONS LETTERS from 2012 to 2016. He has been a Senior Editor of IEEE COMMUNICATIONS LETTERS since January 2015, an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since 2018, and an Area Editor of IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY since 2019.

Chongwen Huang (Member, IEEE) received the B.Sc. degree from Nankai University, Tianjin, China, in 2010, the M.Sc. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2013, and the Ph.D. degree from the Singapore University of Technology and Design (SUTD) in 2019. From October 2019 to September 2020, he was a Post-Doctoral Researcher with SUTD. Since September 2020, he has been a tenure-track young Professor with Zhejiang University. His main research interests include holographic MIMO surface/reconfigurable intelligent surface, B5G/6G wireless communication, mmWave/THz communications, and deep learning technologies for wireless communications. He was a recipient of the IEEE Marconi Prize Paper Award in Wireless Communications and the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2021. He has been serving as an Editor for IEEE COMMUNICATIONS LETTERS, Signal Processing (Elsevier), EURASIP Journal on Wireless Communications and Networking, and Physical Communication since 2021.

Zhi Chen (Senior Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electrical engineering from the University of Electronic Science and Technology of China (UESTC) in 1997, 2000, and 2006, respectively. In April 2006, he joined the National Key Laboratory of Science and Technology on Communications, UESTC, where he has been a Professor since 2013. He was a Visiting Scholar with the University of California at Riverside, Riverside, CA, USA, from 2010 to 2011. His current research interests include 5G mobile communications, tactile internet, and terahertz communication. He has served as a Reviewer for various international journals and conferences, including IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and IEEE TRANSACTIONS ON SIGNAL PROCESSING.