

Max-min Rate Optimization of Low-Complexity Hybrid Multi-User Beamforming Maintaining Rate-Fairness

W. Zhu^{1,2}, H. D. Tuan², E. Dutkiewicz², H. V. Poor³, and L. Hanzo⁴

Abstract—A wireless network serving multiple users in the millimeter-wave or the sub-terahertz band by a base station is considered. High-throughput multi-user hybrid-transmit beamforming is conceived by maximizing the minimum rate of the users. For the sake of energy-efficient signal transmission, the array-of-subarrays structure is used for analog beamforming relying on low-resolution phase shifters. We develop a convex-solver based algorithm, which iteratively invokes a convex problem of the same beamformer size for its solution. We then introduce the soft max-min rate objective function and develop a scalable algorithm for its optimization. Our simulation results demonstrate the striking fact that soft max-min rate optimization not only approaches the minimum user rate obtained by max-min rate optimization but it also achieves a sum rate similar to that of sum-rate maximization. Thus, the soft max-min rate optimization based beamforming design conceived offers a new technique of simultaneously achieving a high individual quality-of-service for all users and a high total network throughput.

Index Terms—Millimeter-wave and sub-THz bands, hybrid beamforming, analog beamforming of low resolution, baseband beamforming, max-min rate optimization, nonconvex optimization algorithms

I. INTRODUCTION

The millimeter-wave (mMwave) band ranging from 30 to 300 GHz and the sub-Terahertz (sTHz) band ranging from 0.1 to 1 THz [1] have emerged as the leading candidates for spectrum exploitation in addressing the forthcoming spectrum scarcity and facilitating high-volume data delivery. These bands offer explicit advantages due to their rapidly developing advanced circuit design [2]–[6].

To mitigate the significant path loss experienced in the mMwave and sTHz bands, as well as to manage power consumption in circuitry, it is necessary to utilize a large number of transmit antennas (TAs) while limiting the number of radio frequency (RF) chains used for signal transmission. Hybrid

beamforming (HBF) modelled by the matrix-vector product of analog and digital (baseband) beamforming is considered the most promising signal processing technique for addressing these challenges.

Initially, analog beamforming (ABF) was based on a fully-connected (FC) architecture, where each RF chain was connected to all antennas. However, it necessitated an excessive number of phase shifters, even for a low number of RF chains, and thus still consumed considerable power. Recently, the array-of-subarrays structure (AOSA) [7] has emerged as a much more practical low-power solution for HBF [8], [9], where each RF chain is connected to a subset of antennas. The AOSA also enables the utilization of more RF chains, thereby improving the spatial diversity attained.

The HBF design has been the subject of extensive research [10]–[28], with single user HBF being considered in [12], [15], [22], [24], [26] and multiuser (MU) HBF being considered in [13], [16], [18]–[21], [27], [28]. Due to the computationally challenging unit modulus constraints imposed on each entry of the ABF matrix, all these papers have only developed heuristic procedures, which do not guarantee convergence or predictable performance. For instance, the authors of [20], [21] assumed that there was no MU interference in their ABF alternating optimization and utilized semi-definite relaxation (SDR) in their baseband beamforming (BBF) solution. Similarly, the authors of [13] utilize SDR in both their ABF and BBF alternating optimization. It should be mentioned that SDR is based on convex problems of excessive dimensions. For instance, for alternating optimization of the ABF matrix of size 64×4 having 256 decision variables as considered in [13], the resultant SDR involves $256 \times 257/2 = 32,896$ decision variables. Such a complex computation is clearly beyond the capacity of existing convex-solvers. Moreover, SDR cannot be used in alternating optimization, since it cannot generate a feasible point.¹

Another issue of MU beamforming is that it is often based on sum rate (SR) maximization [11], [14], which results in zero rates for many users [30, Table II]. To improve the rates of all users while maintaining computational tractability, our previous treatise [30] proposed maximizing the geometric mean of the users' rates. However, the ratio of the minimum and maximum rates [30, Table III] is still well below 0.25, instead of approaching unity for the sake of rate-fairness. The

The work was supported in part by the Australian Research Council's Discovery Projects under Grant DP190102501, in part by the U.S National Science Foundation under Grants CNS-2128448 and ECCS-2335876, in part by the Engineering and Physical Sciences Research Council projects EP/W016605/1, EP/X01228X/1 and EP/Y026721/1 as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028)

¹School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China (email: wenbozhu@shu.edu.cn); ²School of Electrical and Data Engineering, University of Technology Sydney, Broadway, NSW 2007, Australia (email: wenbo.zhu@student.uts.edu.au, tuan.hoang@uts.edu.au, eryk.dutkiewicz@uts.edu.au); ³Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA (email: poor@princeton.edu); ⁴School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K (email: lh@ecs.soton.ac.uk)

¹SDR is only efficient in very limited cases, namely when the SDR problem has only a single solution of rank-one. Otherwise, it does not perform better than a very trivial technique [29]

authors of [31] aim for maximizing the sum dirty paper coding (DPC) rate, which is capable of providing fairer rate distributions than conventional SR maximization [32]. However, DPC is a strictly information-theoretic concept, which cannot be implemented in practice.

Against the above background, this is the first piece of work that considers the HBF design problem of providing uniformly high throughputs for all users. In contrast to other studies, we also restrict the phase shifters to have low resolution for practical implementation. In a nutshell, our contributions are three-fold:

- We develop a convex-solver based algorithm for HBF design by maximizing the users' minimum rate (MR), which iteratively invokes a convex problem of the same beamformer size to generate a gradually improved feasible point;
- We propose a new optimization formulation, termed as soft max-min rate optimization for addressing the computational issues encountered in high-dimensional nonconvex problems. Accordingly, a scalable algorithm is developed for their solution, which is based on a closed-form expression for gradually generating an improved feasible point;
- The extensive simulations show the striking benefits of soft max-min optimization based beamforming: its minimum rate (MR) is almost as high as that of max-min rate optimization based beamforming, and its SR performance approaches that of SR maximization based beamforming. Hence, utilizing soft max-min optimization yields valuable insights into identifying beneficial near-optimal solutions for the concurrent SR and MR objectives.

In Table I, we boldly contrast our contributions to the related literature.

The remainder of the paper is organized as follows. Section II is devoted to the development of a convex-solver based algorithm for the max-min rate optimization of HBFs. Section III is dedicated to the conception of a scalable algorithm for the soft max-min rate optimization of HBFs. Section IV considers similar designs for the case of ABFs under the FC structure. Section V provides our simulations, while Section VI concludes the paper. The Appendix provides mathematical tools for the algorithmic derivations.

Notation. Only the optimization variables are boldfaced; $C(0, a)$ for $a > 0$ represents the set of Gaussian distributions with zero mean and power a ; $\angle x$ is the argument of a complex number x ; The inner product between the matrices X and Y is defined by $\langle X, Y \rangle = \text{trace}(X^H Y)$; We also use $\langle X \rangle$ for the trace of X when X is a long matrix expression. Furthermore, $[X]^2$ refers to XX^H , so we have $[X^H]^2 = X^H X$, and $\|AX\|^2 = \langle [A^H]^2, [X]^2 \rangle$ for the matrices X and A . For a real vector $\theta = (\theta_1, \dots, \theta_n)^T$, we define $e^{j\theta}$ as the complex vector $(e^{j\theta_1}, \dots, e^{j\theta_n})^T$, $\text{diag}[X_i]_{i \in \mathcal{I}}$ forms a matrix arranging X_i , $i \in \mathcal{I}$ in diagonal format. For instance, $\text{diag}[A_i]_{i=1,2} \triangleq \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$.

Ingredient. According to [33, p. 366], a function \bar{f} is said to be a tight minorant (majorant, resp.) of a function f over the domain $\text{dom}(f)$ at a point $\bar{x} \in \text{dom}(f)$ if $f(x) \geq \bar{f}(x) \forall x \in$

$\text{dom}(f)$ ($f(x) \leq \bar{f}(x) \forall x \in \text{dom}(f)$, resp.) and $f(z^{(\kappa)}) = \bar{f}(z^{(\kappa)})$. When \bar{f} is a tight minorant, $f(x_{opt}) \geq \bar{f}(\bar{x})$ holds for $x_{opt} = \arg \max_{x \in \text{dom}(f)} \bar{f}(x)$. When \bar{f} is a tight majorant, we have $f(x_{opt}) \leq \bar{f}(\bar{x})$ for $x_{opt} = \arg \min_{x \in \text{dom}(f)} \bar{f}(x)$.

II. MAX-MIN RATE OPTIMIZATION BASED HBF DESIGN

We consider the downlink (DL) of a base station (BS) serving K users indexed by $k \in \mathcal{K} \triangleq \{1, 2, \dots, K\}$. The BS is equipped with a massive N -antenna array, while each user equipment (UE) k has a single antenna.

For $\mathcal{N} \triangleq \{1, \dots, N\}$ and $\mathcal{N}_{RF} \triangleq \{1, \dots, N_{RF}\}$, where N_{RF} is the number of RF chains that the BS uses for HBF, let us assume that each RF chain is connected to only $L = N/N_{RF}$ antennas, so the phase shift based AB matrix $V_{RF}(\boldsymbol{\theta})$ has the following AOSA structure [7]:

$$V_{RF}(\boldsymbol{\theta}) \triangleq \text{diag}[v_{RF}^j(\boldsymbol{\theta}_j)]_{j=1, \dots, N_{RF}}, \quad (1)$$

with $v_{RF}^j(\boldsymbol{\theta}_j) = e^{j\boldsymbol{\theta}_j} \in \mathbb{C}^L$, for $\boldsymbol{\theta}_j \triangleq (\boldsymbol{\theta}_{1,j}, \dots, \boldsymbol{\theta}_{L,j})^T \in \mathbb{R}^L$, which satisfy the following discrete constraints of b -bit resolution for their practical implementation [34]:

$$\boldsymbol{\theta}_{\ell,j} \in \mathcal{B} \triangleq \left\{ \nu \frac{2\pi}{2^b}, \nu = 0, 1, \dots, 2^b - 1 \right\}, (\ell, j) \in \mathcal{L} \times \mathcal{N}_{RF}, \quad (2)$$

with $\mathcal{L} \triangleq \{1, \dots, L\}$. This AOSA only needs N phase shifters, so the circuit power consumption (in mW unit) is

$$N_{RF} \times 118 + N \times 20 = \frac{N}{L} \times 118 + N \times 20, \quad (3)$$

where 118 mW is the power consumption per RF chain [35], and 20 mW is the power consumption per phase shifter [36].

Let $h_k \triangleq [h_{k,1} \ \dots \ h_{k,N_{RF}}] \in \mathbb{C}^{1 \times N}$ along with $h_{k,j} \in \mathbb{C}^{1 \times L}$ represent the channel between the BS and UE $k \in \mathcal{K}$, which is assumed to be known.²

For $s_k \in C(0, 1)$ being the information intended for UE k , which is "beamformed" by $\mathbf{v}_k^B \in \mathbb{C}^{N_{RF}}$, the signal received at UE k is

$$y_k = h_k V_{RF}(\boldsymbol{\theta}) \sum_{\ell=1}^K \mathbf{v}_\ell^B s_\ell + n_k \quad (4)$$

$$= \tilde{h}_k(\boldsymbol{\theta}) \sum_{\ell=1}^K \mathbf{v}_\ell^B s_\ell + n_k, \quad (5)$$

where $n_k \in C(0, \sigma)$ is the background noise, and

$$\begin{aligned} \tilde{h}_k(\boldsymbol{\theta}) &\triangleq h_k V_{RF}(\boldsymbol{\theta}) \\ &= [h_{k,1} v_{RF}^1(\boldsymbol{\theta}_1) \ \dots \ h_{k,N_{RF}} v_{RF}^{N_{RF}}(\boldsymbol{\theta}_{N_{RF}})] \\ &\in \mathbb{C}^{1 \times N_{RF}}, k \in \mathcal{K}. \end{aligned} \quad (6)$$

We will also use the following representations:

$$\begin{aligned} \tilde{h}_k(\boldsymbol{\theta}) \mathbf{v}_\ell^B &= \sum_{j=1}^{N_{RF}} h_{k,j} v_{RF}^j(\boldsymbol{\theta}_j) \mathbf{v}_\ell^B(j) \\ &= \tilde{h}_k(\mathbf{v}_\ell^B) v_{RF}(\boldsymbol{\theta}) \end{aligned} \quad (7)$$

²The reader is referred e.g. to [37] and to the references therein for its efficient estimation.

TABLE I: Our novel contributions

Contents \ Literature	This work	[11], [14]	[13]	[20], [21]	[30]	[31]
Energy-efficiency	✓					
Zero rates		✓				
Low complexity	✓					
Scalable computations	✓				✓	
Algorithmic convergence	✓				✓	
Low-resolution ABF	✓				✓	
Uniformly high rates	✓					
High sum rate	✓					

for

$$\boldsymbol{\theta} \triangleq \begin{bmatrix} \boldsymbol{\theta}_1 \\ \dots \\ \boldsymbol{\theta}_{N_{RF}} \end{bmatrix} \in \mathbb{R}^N, v_{RF}(\boldsymbol{\theta}) \triangleq \begin{bmatrix} v_{RF}^1(\boldsymbol{\theta}_1) \\ \dots \\ v_{RF}^{N_{RF}}(\boldsymbol{\theta}_{N_{RF}}) \end{bmatrix} \in \mathbb{C}^N, \\ \tilde{h}_k(\mathbf{v}_\ell^B) \triangleq [\mathbf{v}_\ell^B(1)h_{k,1} \quad \dots \quad \mathbf{v}_\ell^B(N_{RF})h_{k,N_{RF}}]. \quad (8)$$

Let the BBF matrix be defined by

$$\mathbf{v}^B = [\mathbf{v}_1^B \quad \dots \quad \mathbf{v}_K^B] \in \mathbb{C}^{N_{RF} \times K}. \quad (9)$$

From (5), the achievable rate of UE k is defined by

$$r_k(\boldsymbol{\theta}, \mathbf{v}^B) \triangleq \ln \left(1 + \frac{|\tilde{h}_k(\boldsymbol{\theta})\mathbf{v}_k^B|^2}{\psi_k(\boldsymbol{\theta}, \mathbf{v}^B)} \right), \quad (10)$$

with

$$\psi_k(\boldsymbol{\theta}, \mathbf{v}^B) \triangleq \sum_{\ell \neq k} |\tilde{h}_k(\boldsymbol{\theta})\mathbf{v}_\ell^B|^2 + \sigma. \quad (11)$$

Given the power budget P , the BS's transmit power is constrained as

$$\begin{aligned} \sum_{k=1}^K \|V_{RF}(\boldsymbol{\theta})\mathbf{v}_k^B\|^2 &= \sum_{k=1}^K \sum_{j=1}^{N_{RF}} \|v_{RF}^j(\boldsymbol{\theta}_j)\mathbf{v}_k^B(j)\|^2 \\ &= L \sum_{k=1}^K \|\mathbf{v}_k^B\|^2 \leq P \\ \Leftrightarrow \sum_{k=1}^K \|\mathbf{v}_k^B\|^2 &\leq P/L, \end{aligned} \quad (12)$$

which is independent of $\boldsymbol{\theta}$. We consider the following problem of max-min rate optimization:

$$\max_{\boldsymbol{\theta}, \mathbf{v}^B} f(\boldsymbol{\theta}, \mathbf{v}^B) \triangleq \min_{k \in \mathcal{K}} r_k(\boldsymbol{\theta}, \mathbf{v}^B) \quad \text{s.t.} \quad (2), (12), \quad (13)$$

which is computationally challenging due to the following complications: (a) the rate function $r_k(\boldsymbol{\theta}, \mathbf{v}^B)$ is nonconcave, making the optimization objective function (OF) in (13) both nonsmooth (nondifferentiable) and also nonconcave; (b) The constraint (2) is discrete, having as many as 2^{bN} discrete feasible points for optimization in $\boldsymbol{\theta}$.

For circumventing the issue (b), we introduce a new continuous variable

$$\boldsymbol{\phi} \triangleq (\boldsymbol{\phi}_1^T, \dots, \boldsymbol{\phi}_{N_{RF}}^T)^T \in \mathbb{C}^N, \boldsymbol{\phi}_j \triangleq (\boldsymbol{\phi}_{1,j}, \dots, \boldsymbol{\phi}_{L,j})^T \in \mathbb{C}^L. \quad (14)$$

We then define

$$r_k(\boldsymbol{\phi}, \mathbf{v}^B) \triangleq \ln \left(1 + \frac{|\tilde{h}_k(\boldsymbol{\phi})\mathbf{v}_k^B|^2}{\psi_k(\boldsymbol{\phi}, \mathbf{v}^B)} \right), \quad (15)$$

for

$$\begin{aligned} \tilde{h}_k(\boldsymbol{\phi}) &= h_k \text{diag}[\boldsymbol{\phi}_j]_{j=1, \dots, N_{RF}} \\ &= [h_{k,1}\boldsymbol{\phi}_1 \quad \dots \quad h_{k,N_{RF}}\boldsymbol{\phi}_{N_{RF}}], \end{aligned} \quad (16)$$

and

$$\psi_k(\boldsymbol{\phi}, \mathbf{v}^B) \triangleq \sum_{\ell \neq k} |\tilde{h}_k(\boldsymbol{\phi})\mathbf{v}_\ell^B|^2 + \sigma. \quad (17)$$

By defining $f(\boldsymbol{\phi}, \mathbf{v}^B) \triangleq \min_{k \in \mathcal{K}} r_k(\boldsymbol{\phi}, \mathbf{v}^B)$, we address the following penalized optimization for solving the problem in (13):

$$\begin{aligned} \max_{\boldsymbol{\theta}, \mathbf{v}^B, \boldsymbol{\phi}} f_\gamma(\boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{v}^B) &\triangleq f(\boldsymbol{\phi}, \mathbf{v}^B) - \gamma \|\boldsymbol{\phi} - v_{RF}(\boldsymbol{\theta})\|^2 \\ \text{s.t.} &\quad (2), (12), \end{aligned} \quad (18)$$

where $\gamma > 0$ is a penalty parameter. Note that no constraint is imposed on $\boldsymbol{\phi}$, while the discrete variable $\boldsymbol{\theta}$ is decoupled from the max-min rate OF. The motivated reader is referred to [38, Chapter 16] for discussions on the efficacy of the penalty optimization methodology.

We now propose an alternating optimization-based procedure for the solution of (18). Initialized by the triplet $(v^{B,(0)}, \boldsymbol{\phi}^{(0)}, \boldsymbol{\theta}^{(0)})$ feasible for (18), let $(v^{B,(\kappa)}, \boldsymbol{\phi}^{(\kappa)}, \boldsymbol{\theta}^{(\kappa)})$ be a feasible point for (18) that is found from the $(\kappa - 1)$ -st iteration.

A. Alternating optimization in BBF

We seek BBF $v^{B,(\kappa+1)}$ ensuring that $f_\gamma(\boldsymbol{\theta}^{(\kappa)}, \boldsymbol{\phi}^{(\kappa)}, v^{B,(\kappa+1)}) > f_\gamma(\boldsymbol{\theta}^{(\kappa)}, \boldsymbol{\phi}^{(\kappa)}, v^{B,(\kappa)})$, which is the same as

$$f(\boldsymbol{\phi}^{(\kappa)}, v^{B,(\kappa+1)}) > f(\boldsymbol{\phi}^{(\kappa)}, v^{B,(\kappa)}), \quad (19)$$

by considering the following problem:

$$\max_{\mathbf{v}^B} f(\boldsymbol{\phi}^{(\kappa)}, \mathbf{v}^B) \triangleq \min_{k \in \mathcal{K}} r_k(\boldsymbol{\phi}^{(\kappa)}, \mathbf{v}^B) \quad \text{s.t.} \quad (12). \quad (20)$$

Recalling from (10) and (11) that $r_k(\boldsymbol{\phi}^{(\kappa)}, \mathbf{v}^B) = \ln [1 + |\tilde{h}_k(\boldsymbol{\theta}^{(\kappa)})\mathbf{v}_k^B|^2 / \psi_k(\boldsymbol{\theta}^{(\kappa)}, \mathbf{v}^B)]$ with

$$\psi_k(\boldsymbol{\theta}^{(\kappa)}, \mathbf{v}^B) = \sum_{\ell \neq k} |\tilde{h}_k(\boldsymbol{\theta}^{(\kappa)})\mathbf{v}_\ell^B|^2 + \sigma, \quad (21)$$

as well as by applying the inequality (89) for $\bar{x} = x_k^{(\kappa)} \triangleq \tilde{h}_k(\boldsymbol{\theta}^{(\kappa)})v_k^{B,(\kappa)}$ and $\bar{y} = y_k^{(\kappa)} \triangleq \psi_k(\boldsymbol{\phi}^{(\kappa)}, v^{B,(\kappa)})$, we obtain

the following tight concave quadratic minorant of $r_k(\phi^{(\kappa)}, \mathbf{v}^B)$ at $v^{B,(\kappa)}$:

$$r_k^{(\kappa)}(\mathbf{v}^B) \triangleq \alpha_k^{(\kappa)} + 2\Re\{a_k^{(\kappa)} \mathbf{v}_k^B\} - \beta_k^{(\kappa)} \langle [\tilde{h}_k^H(\phi^{(\kappa)})]^2, \sum_{\ell=1}^K [\mathbf{v}_\ell^B]^2 \rangle, \quad (22)$$

with $\alpha_k^{(\kappa)} \triangleq r_k(\phi^{(\kappa)}, v^{B,(\kappa)}) - |x_k^{(\kappa)}|^2/y_k^{(\kappa)} - \sigma\beta_k^{(\kappa)}$, $a_k^{(\kappa)} \triangleq (x_k^{(\kappa)})^* \tilde{h}_k(\phi^{(\kappa)})/y_k^{(\kappa)}$, $\beta_k^{(\kappa)} \triangleq 1/y_k^{(\kappa)} - 1/(y_k^{(\kappa)} + |x_k^{(\kappa)}|^2)$.

We thus solve the following convex problem of minorant maximization of (20) to generate $v^{B,(\kappa+1)}$ ensuring (19):

$$\max_{\mathbf{v}^B} f_B^{(\kappa)}(\mathbf{v}^B) \triangleq \min_{k \in \mathcal{K}} r_k^{(\kappa)}(\mathbf{v}^B) \quad \text{s.t.} \quad (12). \quad (23)$$

B. Alternating optimization in ϕ

We seek $\phi^{(\kappa+1)}$ for ensuring that

$$f_\gamma(\theta^{(\kappa)}, \phi^{(\kappa+1)}, v^{B,(\kappa+1)}) > f_\gamma(\theta^{(\kappa)}, \phi^{(\kappa)}, v^{B,(\kappa+1)}) \quad (24)$$

by considering the following problem:

$$\begin{aligned} & \max_{\phi} f_\gamma(\theta^{(\kappa)}, \phi, v^{B,(\kappa+1)}) \\ & \triangleq \min_{k \in \mathcal{K}} r_k(\phi, v^{B,(\kappa+1)}) - \gamma \|\phi - v_{RF}(\theta^{(\kappa)})\|^2. \end{aligned} \quad (25)$$

By recalling from (15) and (7) that $r_k(\phi, v^{B,(\kappa+1)}) \triangleq \ln \left[1 + |\tilde{h}_k(v_k^{B,(\kappa+1)})\phi|^2 / \psi_k(\phi, v^{B,(\kappa+1)}) \right]$ with

$$\psi_k(\phi, v^{B,(\kappa+1)}) = \sum_{\ell \neq k} |\tilde{h}_k(v_\ell^{B,(\kappa+1)})\phi|^2 + \sigma, \quad (26)$$

and by applying the inequality (89) of the Appendix for $\bar{x} = \tilde{x}_k^{(\kappa)} \triangleq \tilde{h}_k(v_k^{B,(\kappa+1)})\phi^{(\kappa)}$ and $\bar{y} = \tilde{y}_k^{(\kappa)} \triangleq \psi_k(\phi^{(\kappa)}, v^{B,(\kappa+1)})$, we obtain the following tight minorant of $r_k(\phi, v^{B,(\kappa+1)})$ at $\phi^{(\kappa)}$:

$$\tilde{r}_k^{(\kappa)}(\phi) \triangleq \tilde{\alpha}_k^{(\kappa)} + 2\Re\{\tilde{a}_k^{(\kappa)} \phi\} - \tilde{\beta}_k^{(\kappa)} \langle \sum_{\ell=1}^K [\tilde{h}_k^H(v_\ell^{B,(\kappa+1)})]^2, [\phi]^2 \rangle, \quad (27)$$

with $\tilde{\alpha}_k^{(\kappa)} \triangleq r_k(\phi^{(\kappa)}, v^{B,(\kappa+1)}) - |\tilde{x}_k^{(\kappa)}|^2/\tilde{y}_k^{(\kappa)} - \sigma\tilde{\beta}_k^{(\kappa)}$, $\tilde{a}_k^{(\kappa)} \triangleq \frac{(\tilde{x}_k^{(\kappa)})^*}{\tilde{y}_k^{(\kappa)}} \tilde{h}_k(v_k^{B,(\kappa+1)})$, $\tilde{\beta}_k^{(\kappa)} \triangleq 1/\tilde{y}_k^{(\kappa)} - 1/(\tilde{y}_k^{(\kappa)} + |\tilde{x}_k^{(\kappa)}|^2)$.

We thus solve the following convex problem of minorant maximization formulated in (25) for generating $\phi^{(\kappa+1)}$ satisfying (24):

$$\max_{\phi} f_{\gamma,z}^{(\kappa)} \triangleq \min_{k \in \mathcal{K}} \tilde{r}_k^{(\kappa)}(\phi) - \gamma \|\phi - v_{RF}(\theta^{(\kappa)})\|^2. \quad (28)$$

C. Alternating optimization in ABF

To seek $\theta^{(\kappa+1)}$ for ensuring that $f_\gamma(\theta^{(\kappa+1)}, \phi^{(\kappa+1)}, v^{B,(\kappa+1)}) > f_\gamma(\theta^{(\kappa)}, \phi^{(\kappa+1)}, v^{B,(\kappa+1)})$, which is the same as

$$\|\phi^{(\kappa+1)} - v_{RF}(\theta^{(\kappa+1)})\|^2 < \|\phi^{(\kappa+1)} - v_{RF}(\theta^{(\kappa)})\|^2, \quad (29)$$

we consider the problem $\min_{\theta} \|\phi^{(\kappa+1)} - v_{RF}(\theta)\|^2$ s.t. (2), which admits the following closed-form solution:

$$\theta_{\ell,j}^{(\kappa+1)} = \lfloor \angle \phi_{\ell,j}^{(\kappa+1)} \rfloor_b, (\ell, j) \in \mathcal{L} \times \mathcal{N}_{RF}. \quad (30)$$

Here and after, $\lfloor \alpha \rfloor_b$ is the b -bit rounded version of $\alpha \in [0, 2\pi)$ defined by $\lfloor \alpha \rfloor_b = \nu_\alpha \frac{2\pi}{2^b}$ with $\nu_\alpha \triangleq \arg \min_{\nu' \in \{\nu, \nu+1\}} |\nu' \frac{2\pi}{2^b} - \alpha|$, where ν is selected for satisfying that $\alpha \in [\nu \frac{2\pi}{2^b}, (\nu+1) \frac{2\pi}{2^b}]$. If $\nu_\alpha = 2^b$ we reset it to $\nu_\alpha = 0$.

Algorithm 1 Max-min rate optimization-based HBF algorithm

- 1: **Initialization:** Initialize a feasible point $(\theta^{(0)}, \phi^{(0)}, v^{B,(0)})$ for (18).
- 2: **Repeat until convergence of the objective function in (18):** Generate $v^{B,(\kappa+1)}$ by solving the convex problem (23), $\phi^{(\kappa+1)}$ by solving the convex problem (28), and $\theta^{(\kappa+1)}$ by (30). Reset $\kappa := \kappa + 1$.
- 3: **Output** $(\theta^{opt}, \tilde{v}^{B,opt}) = (\theta^{(\kappa)}, v^{B,(\kappa)})$.

D. Max-min rate optimization and its convergence

Algorithm 1 summarizes the computational procedure iterating by solving the convex problems (23) as well as (28), and the closed-form (30) to generate a sequence $\{(\phi^{(\kappa)}, \theta^{(\kappa)}, v^{B,(\kappa)})\}$ of improved feasible points for (18), because we have $f_\gamma(\phi^{(\kappa+1)}, \theta^{(\kappa+1)}, v^{B,(\kappa+1)}) > f_\gamma(\phi^{(\kappa)}, \theta^{(\kappa)}, v^{B,(\kappa)})$ by (19), (24), and (29). This sequence is convergent according to Cauchy's theorem. Moreover, for a sufficient large γ , we have $\|\phi^{(\kappa)} - v_{RF}(\theta^{(\kappa)})\| \rightarrow 0$, so $(\theta^{(\kappa)}, v^{B,(\kappa)})$ represents an optimized solution of the max-min rate optimization problem (13).

E. SR maximization based HBF algorithm

Instead of the problem (13) of max-min optimization, we now consider the following problem of SR maximization:

$$\max_{\theta, \mathbf{v}^B} g(\theta, \mathbf{v}^B) \triangleq \sum_{k=1}^K r_k(\theta, \mathbf{v}^B) \quad \text{s.t.} \quad (2), (12), \quad (31)$$

which is addressed based on the following problem of penalized optimization:

$$\begin{aligned} & \max_{\theta, \mathbf{v}^B, \phi} g_\gamma(\theta, \phi, \mathbf{v}^B) \triangleq \sum_{k=1}^K r_k(\phi, \mathbf{v}^B) - \gamma \|\phi - v_{RF}(\theta)\|^2 \\ & \quad \text{s.t.} \quad (2), (12). \end{aligned} \quad (32)$$

Initialized by $(z^{(0)}, \phi^{(0)}, \theta^{(0)})$ feasible for (32), let $(z^{(\kappa)}, \phi^{(\kappa)}, \theta^{(\kappa)})$ be a feasible point for (32) that is found from the $(\kappa - 1)$ -st iteration. The alternating optimization at the κ -th iteration proceeds as follows.

1) *Alternating optimization in BBF:* Similarly to (23), we generate $v^{B,(\kappa+1)}$ by solving the problem

$$\max_{\mathbf{v}^B} \sum_{k=1}^K r_k^{(\kappa)}(\mathbf{v}^B) \quad \text{s.t.} \quad (12), \quad (33)$$

where $r_k^{(\kappa)}$ is defined from (22). By taking into account that $\sum_{k=1}^K r_k^{(\kappa)}(\mathbf{v}^B) = \alpha^{(\kappa)} + \sum_{k=1}^K 2\Re\{a_k^{(\kappa)} \mathbf{v}_k^B\} - \sum_{k=1}^K \langle \Xi^{(\kappa)}, [\mathbf{v}^B]^2 \rangle$ with $\alpha^{(\kappa)} \triangleq \sum_{k=1}^K \alpha_k^{(\kappa)}$ and $\Xi^{(\kappa)} = \sum_{k=1}^L \beta_k^{(\kappa)} [\tilde{h}_k^H(\phi^{(\kappa)})]^2$, the problem (33) admits the closed-form solution of

$$v_k^{B,(\kappa+1)} = \begin{cases} (\Xi^{(\kappa)})^{-1} (a_k^{(\kappa)})^H & \\ \text{if } \sum_{k=1}^K \|(\Xi^{(\kappa)})^{-1} (a_k^{(\kappa)})^H\|^2 \leq P/L, & \\ (\Xi^{(\kappa)} + \lambda I_{N_{RF}})^{-1} (a_k^{(\kappa)})^H & \\ \text{otherwise,} & \end{cases} \quad (34)$$

where $\lambda > 0$ is found by bisection, so that $\sum_{k=1}^K \|(\Xi^{(\kappa)} + \lambda I_{N_{RF}})^{-1} (a_k^{(\kappa)})^H\|^2 = P/L$.

Algorithm 2 Scalable SR maximization-based HBF algorithm

- 1: **Initialization:** Initialize $(\theta^{(0)}, \phi^{(0)}, v^{B,(0)})$.
 - 2: **Repeat until convergence of the objective function in (32):** Generate $v^{B,(\kappa+1)}$ by (34), and $\phi^{(\kappa+1)}$ by (37), and $\theta^{(\kappa+1)}$ by (30). Reset $\kappa := \kappa + 1$.
 - 3: **Output** $(\theta^{opt}, \tilde{v}^{B,opt}) = (\theta^{(\kappa)}, v^{B,(\kappa)})$.
-

2) *Alternating optimization in ϕ :* Like in (28), we generate $\phi^{(\kappa+1)}$ by solving the problem

$$\max_{\phi} \sum_{k=1}^K \tilde{r}_k^{(\kappa)}(\phi) - \gamma \|\phi - v_{RF}(\theta^{(\kappa)})\|^2, \quad (35)$$

where $\tilde{r}_k^{(\kappa)}(\phi)$ is defined in (27). By presenting $\sum_{k=1}^K \tilde{r}_k^{(\kappa)}(\phi) - \gamma \|\phi - v_{RF}(\theta^{(\kappa)})\|^2 = \tilde{\alpha}^{(\kappa)} + 2\Re\{\tilde{a}^{(\kappa)}\phi\} - \langle \tilde{\Xi}^{(\kappa)}, [\phi]^2 \rangle$ with $\tilde{\alpha}^{(\kappa)} \triangleq \sum_{k=1}^K \tilde{a}_k^{(\kappa)} - \gamma N$, and

$$\begin{aligned} \tilde{a}^{(\kappa)} &\triangleq \sum_{k=1}^K \tilde{a}_k^{(\kappa)} + \gamma [v_{RF}(\theta^{(\kappa)})]^H, \\ \tilde{\Xi}^{(\kappa)} &\triangleq \sum_{k=1}^K \tilde{\beta}_k^{(\kappa)} \sum_{\ell=1}^K [\tilde{h}_k^H(v_\ell^{B,(\kappa+1)})]^2 + \gamma I_N, \end{aligned} \quad (36)$$

the problem (35) admits the closed-form solution of

$$\phi^{(\kappa+1)} = (\tilde{\Xi}^{(\kappa)})^{-1} (\tilde{a}^{(\kappa)})^H. \quad (37)$$

3) *Algorithm :* It may now be seen that alternating optimization in θ is based on the closed-form (30). As such, Algorithm 2 constructed for solving problem (32) is of scalable complexity, with the total computational complexity of each iteration being on the order of $\mathcal{O}(N_{RF}K) + \mathcal{O}(N)$.

III. SOFT MAX MIN RATE OPTIMIZATION BASED HBF DESIGN

The total computational complexity of the convex problems (23) and (28) that are solved at each iteration of Algorithm 1 is on the order of $\mathcal{O}[(N_{RF}K)^3] + \mathcal{O}(N^3)$, which is high, because N is large. This motivates us in this section to develop another technique of finding the best MR by scalable computation.

One has

$$\begin{aligned} &\max_{\theta, \mathbf{v}^B} \min_{k \in \mathcal{K}} r_k(\theta, \mathbf{v}^B) \\ \Leftrightarrow &\max_{\theta, \mathbf{v}^B} \min_{k \in \mathcal{K}} \ln \left(1 + \frac{1}{c} \frac{|\tilde{h}_k(\theta) \mathbf{v}_k^B|^2}{\psi_k(\theta, \mathbf{v}^B)} \right) \end{aligned} \quad (38)$$

$$\Leftrightarrow \max_{\theta, \mathbf{v}^B} \left[-\max_{k \in \mathcal{K}} \ln \left(1 + \frac{|\tilde{h}_k(\theta) \mathbf{v}_k^B|^2}{c \psi_k(\theta, \mathbf{v}^B)} \right)^{-1} \right], \quad (39)$$

while

$$\begin{aligned} &\max_{k \in \mathcal{K}} \ln \left(1 + \frac{|\tilde{h}_k(\theta) \mathbf{v}_k^B|^2}{c \psi_k(\theta, \mathbf{v}^B)} \right)^{-1} \\ &\leq \ln \left(\sum_{k=1}^K \left(1 + \frac{|\tilde{h}_k(\theta) \mathbf{v}_k^B|^2}{c \psi_k(\theta, \mathbf{v}^B)} \right)^{-1} \right) \end{aligned} \quad (40)$$

$$= \ln \left(\frac{\sum_{k=1}^K \left(1 + \frac{|\tilde{h}_k(\theta) \mathbf{v}_k^B|^2}{c \psi_k(\theta, \mathbf{v}^B)} \right)^{-1}}{K} \right) + \ln K \quad (41)$$

$$\leq \max_{k \in \mathcal{K}} \ln \left(1 + \frac{|\tilde{h}_k(\theta) \mathbf{v}_k^B|^2}{c \psi_k(\theta, \mathbf{v}^B)} \right)^{-1} + \ln K. \quad (42)$$

Note that for sufficiently small c , $\ln K$ is very small compared to the absolute value of the LHS of (40). In other words, by choosing small enough c , the LHS of (40) can be approximated with arbitrary tolerance by the right-hand side (RHS) of (41), which is $\ln \pi_c(\theta, \mathbf{v}^B)$ for

$$\pi_c(\theta, \mathbf{v}^B) \triangleq \sum_{k=1}^K \left(1 - \frac{|\tilde{h}_k(\theta) \mathbf{v}_k^B|^2}{|\tilde{h}_k(\theta) \mathbf{v}_k^B|^2 + c \psi_k(\theta, \mathbf{v}^B)} \right). \quad (43)$$

Instead of the max-min optimization problem (13), we thus consider the following problem referred to as the soft max-min optimization problem:

$$\max_{\theta, \mathbf{v}^B} [-\ln \pi_c(\theta, \mathbf{v}^B)] \quad \text{s.t.} \quad (2), (12), \quad (44)$$

which is equivalent to the problem

$$\min_{\theta, \mathbf{v}^B} \ln \pi_c(\theta, \mathbf{v}^B) \quad \text{s.t.} \quad (2), (12). \quad (45)$$

We then address its solution by the following problem of penalized optimization:

$$\begin{aligned} \min_{\theta, \mathbf{v}^B, \phi} f_{\gamma, c}(\theta, \phi, \mathbf{v}^B) &\triangleq \ln \pi_c(\theta, \mathbf{v}^B) + \gamma \|\phi - v_{RF}(\theta)\|^2 \\ \text{s.t.} &\quad (2), (12), \quad (46) \end{aligned}$$

where $\gamma > 0$ is a penalty parameter, and

$$\pi_c(\phi, \mathbf{v}^B) \triangleq \sum_{k=1}^K \left(1 - \frac{|\tilde{h}_k(\phi) \mathbf{v}_k^B|^2}{|\tilde{h}_k(\phi) \mathbf{v}_k^B|^2 + c \psi_k(\phi, \mathbf{v}^B)} \right), \quad (47)$$

with $\tilde{h}_k(\phi)$ and $\psi_k(\phi, \mathbf{v}^B)$ defined from (16) and (17).

We now propose an alternating optimization-based procedure for the solution of (46). Initialized by $(v^{B,(0)}, \phi^{(0)}, \theta^{(0)})$ feasible for (18), let $(v^{B,(\kappa)}, \phi^{(\kappa)}, \theta^{(\kappa)})$ be a feasible point for (46) that is found from the $(\kappa - 1)$ -st iteration.

A. Alternating optimization in BBF

We seek BBF $v^{B,(\kappa+1)}$ ensuring that $f_{\gamma, c}(\theta^{(\kappa)}, \phi^{(\kappa)}, v^{B,(\kappa+1)}) < f_{\gamma, c}(\theta^{(\kappa)}, \phi^{(\kappa)}, v^{B,(\kappa)})$, which is the same as

$$\ln \pi_c(\phi^{(\kappa)}, v^{B,(\kappa+1)}) < \ln \pi_c(\phi^{(\kappa)}, v^{B,(\kappa)}), \quad (48)$$

by considering the following problem:

$$\min_{\mathbf{v}^B} \ln \pi_c^{(\kappa)}(\mathbf{v}^B) \quad \text{s.t.} \quad (12), \quad (49)$$

where we have

$$\begin{aligned} \pi_c^{(\kappa)}(\mathbf{v}^B) &\triangleq \pi_c(\phi^{(\kappa)}, \mathbf{v}^B) \\ &= \sum_{k=1}^K \left(1 - \frac{|\tilde{h}_k(\phi^{(\kappa)}) \mathbf{v}_k^B|^2}{|\tilde{h}_k(\phi^{(\kappa)}) \mathbf{v}_k^B|^2 + c \psi_k(\phi^{(\kappa)}, \mathbf{v}^B)} \right) \end{aligned}$$

with $\psi_k(\phi^{(\kappa)}, \mathbf{v}^B)$ defined from (21). Applying the inequality (88) of the Appendix for $\bar{x}_k = x_k^{(\kappa)} \triangleq \tilde{h}_k(\phi^{(\kappa)}) v_k^{B,(\kappa)}$ and $\bar{y}_k = y_k^{(\kappa)} \triangleq \psi_k(\phi^{(\kappa)}, v^{B,(\kappa)})$ yields the following tight majorant of $\ln \pi_c^{(\kappa)}(\mathbf{v}^B)$ at $v^{B,(\kappa)}$:

$$\rho^{(\kappa)}(\mathbf{v}^B) \triangleq a^{(\kappa)} - 2 \sum_{k=1}^K d_k^{(\kappa)} \Re\{x_k^{(\kappa)*} \tilde{h}_k(\phi^{(\kappa)}) \mathbf{v}_k^B\}$$

$$\begin{aligned}
& + \sum_{k=1}^K c_k^{(\kappa)} \left(c \sum_{\ell \neq k} |\tilde{h}_k(\phi^{(\kappa)}) \mathbf{v}_\ell^B|^2 \right. \\
& \left. + |\tilde{h}_k(\phi^{(\kappa)}) \mathbf{v}_k^B|^2 \right) \\
& = a_b^{(\kappa)} - 2 \sum_{k=1}^K \Re\{b_k^{(\kappa)} \mathbf{v}_k^B\} + \sum_{k=1}^K \langle C_k^{(\kappa)}, [\mathbf{v}_k^B]^2 \rangle,
\end{aligned} \tag{50}$$

where

$$a^{(\kappa)} \triangleq f_{sf,b}^{(\kappa)}(v^{B,(\kappa)}) + \sum_{k=1}^K d_k^{(\kappa)} |x_k^{(\kappa)}|^2 + c\sigma \sum_{k=1}^K c_k^{(\kappa)}, \tag{51}$$

$$d_k^{(\kappa)} \triangleq \frac{(cy_k^{(\kappa)} + |x_k^{(\kappa)}|^2)^{-1}}{\pi_c^{(\kappa)}(v^{B,(\kappa)})}, c_k^{(\kappa)} \triangleq d_k^{(\kappa)} \frac{|x_k^{(\kappa)}|^2}{cy_k^{(\kappa)} + |x_k^{(\kappa)}|^2}, \tag{52}$$

and

$$\begin{aligned}
b_k^{(\kappa)} & \triangleq d_k^{(\kappa)} (\tilde{h}_k(\phi^{(\kappa)}) \mathbf{v}_k^{B,(\kappa)})^* \tilde{h}_k(\phi^{(\kappa)}), \\
C_k^{(\kappa)} & \triangleq c \sum_{\ell \in \mathcal{K} \setminus \{k\}} c_\ell^{(\kappa)} [\tilde{h}_\ell^H(\phi^{(\kappa)})]^2 + c_k^{(\kappa)} [\tilde{h}_k^H(\phi^{(\kappa)})]^2.
\end{aligned} \tag{53}$$

We thus solve the following problem of majorant minimization to generate $v^{B,(\kappa+1)}$ ensuring (48):

$$\min_{\mathbf{v}^B} \rho^{(\kappa)}(\mathbf{v}^B) \quad \text{s.t.} \quad (12), \tag{54}$$

which admits the closed-form solution of

$$v_k^{B,(\kappa+1)} = \begin{cases} (C_k^{(\kappa)})^{-1} (b_k^{(\kappa)})^H & \text{if } \sum_{k=1}^K \|(C_k^{(\kappa)})^{-1} (b_k^{(\kappa)})^H\|^2 \leq P/L, \\ (C_k^{(\kappa)} + \lambda I_{N_{RF}})^{-1} (b_k^{(\kappa)})^H & \text{otherwise,} \end{cases} \tag{55}$$

where $\lambda > 0$ is found by bisection, so that $\sum_{k=1}^K \|(C_k^{(\kappa)} + \lambda I_{N_{RF}})^{-1} (b_k^{(\kappa)})^H\|^2 = P/L$.

B. Alternating optimization in ϕ

We seek $\phi^{(\kappa+1)}$ for ensuring that

$$f_{\gamma,c}(\theta^{(\kappa)}, \phi^{(\kappa+1)}, v^{B,(\kappa+1)}) < f_{\gamma,c}(\theta^{(\kappa)}, \phi^{(\kappa)}, v^{B,(\kappa+1)}), \tag{56}$$

by considering the following problem:

$$\min_{\phi} [\ln \tilde{\pi}_c^{(\kappa)}(\phi) + \gamma \|\phi - v_{RF}(\theta^{(\kappa)})\|^2], \tag{57}$$

where

$$\begin{aligned}
\tilde{\pi}_c^{(\kappa)}(\phi) & \triangleq \pi_c(\phi, v^{B,(\kappa+1)}) \\
& = \sum_{k=1}^K \left(1 - \frac{|\tilde{h}_k(v_k^{B,(\kappa+1)}) \phi|^2}{|\tilde{h}_k(v_k^{B,(\kappa+1)}) \phi + c\psi_k(\phi, v^{B,(\kappa+1)})|} \right)
\end{aligned}$$

with $\psi_k(\phi, v^{B,(\kappa+1)})$ defined from (26). Applying the inequality (88) for $\bar{x}_k = \tilde{x}_k^{(\kappa)} \triangleq \tilde{h}_k(v_k^{B,(\kappa+1)}) \phi^{(\kappa)}$, and $\bar{y}_k = \tilde{y}_k^{(\kappa)} \triangleq \psi_k(\phi^{(\kappa)}, v^{B,(\kappa+1)})$ yields the following tight majorant of $\ln \tilde{\pi}_c^{(\kappa)}(\phi)$ at $\phi^{(\kappa)}$:

$$\tilde{\rho}^{(\kappa)}(\phi) \triangleq \tilde{a}^{(\kappa)} - 2 \sum_{k=1}^K \tilde{d}_k^{(\kappa)} \Re\{(\tilde{x}_k^{(\kappa)})^* \tilde{h}_k(v_k^{B,(\kappa+1)}) \phi\}$$

Algorithm 3 Scalable soft max-min rate optimization based HBF algorithm

- 1: **Initialization:** Initialize $(\theta^{(0)}, \Phi^{(0)}, v^{B,(0)})$.
- 2: **Repeat until convergence of the objective function in (46):** Generate $v^{B,(\kappa+1)}$ by (55), $\phi^{(\kappa+1)}$ by (63), and $\theta^{(\kappa+1)}$ by (30). Reset $\kappa := \kappa + 1$.
- 3: **Output** $(\theta^{opt}, v^{B,opt}) = (\theta^{(\kappa)}, v^{B,(\kappa)})$.

$$\begin{aligned}
& + \sum_{k=1}^K \tilde{c}_k^{(\kappa)} \left(c \sum_{\ell \neq k} |\tilde{h}_k(v_\ell^{B,(\kappa+1)}) \phi|^2 \right. \\
& \left. + |\tilde{h}_k(v_k^{B,(\kappa+1)}) \phi|^2 \right) \\
& = \tilde{a}^{(\kappa)} - 2 \Re\{\tilde{b}^{(\kappa)} \phi\} + \langle \tilde{C}^{(\kappa)}, [\phi]^2 \rangle,
\end{aligned} \tag{58}$$

where we have

$$\tilde{a}^{(\kappa)} \triangleq f_{sf,z}^{(\kappa)}(\phi^{(\kappa)}) - \sum_{k=1}^K \tilde{d}_k^{(\kappa)} |\tilde{x}_k^{(\kappa)}|^2 - c\sigma \sum_{k=1}^K \tilde{c}_k^{(\kappa)}, \tag{59}$$

$$\tilde{d}_k^{(\kappa)} \triangleq \frac{(c\tilde{y}_k^{(\kappa)} + |\tilde{x}_k^{(\kappa)}|^2)^{-1}}{\tilde{\pi}_c^{(\kappa)}(\phi)}, \tilde{c}_k^{(\kappa)} \triangleq \tilde{d}_k^{(\kappa)} \frac{|\tilde{x}_k^{(\kappa)}|^2}{c\tilde{y}_k^{(\kappa)} + |\tilde{x}_k^{(\kappa)}|^2}, \tag{60}$$

and

$$\begin{aligned}
\tilde{b}^{(\kappa)} & \triangleq \sum_{k=1}^K \left(\tilde{d}_k^{(\kappa)} (\tilde{h}_k(v_k^{B,(\kappa+1)}) \phi^{(\kappa)})^* \tilde{h}_k(v_k^{B,(\kappa+1)}) \right), \\
\tilde{C}^{(\kappa)} & \triangleq \sum_{k=1}^K \tilde{c}_k^{(\kappa)} \left(c \sum_{\ell \in \mathcal{K} \setminus \{k\}} [\tilde{h}_\ell^H(v_\ell^{B,(\kappa+1)})]^2 \right. \\
& \left. + [\tilde{h}_k^H(v_k^{B,(\kappa+1)})]^2 \right).
\end{aligned} \tag{61}$$

We thus solve the following problem of majorant minimization of (57) to generate $\phi^{(\kappa+1)}$ ensuring (56):

$$\min_{\phi} \tilde{\rho}^{(\kappa)}(\phi) + \gamma \|\phi - v_{RF}(\theta^{(\kappa)})\|^2 \tag{62}$$

which admits the closed-form solution of

$$\phi^{(\kappa+1)} = (\tilde{C}^{(\kappa)} + \gamma I_N)^{-1} \left[(\tilde{b}^{(\kappa)})^H + \gamma v_{RF}(\theta^{(\kappa)}) \right]. \tag{63}$$

C. Alternating optimization in ABF

Generate $\theta^{(\kappa+1)}$ according (30).

D. Soft max-min rate optimization and its convergence

Algorithm 3 summarizes the computational procedure iterating by evaluating the closed-form expressions of (55), (63), and (30) to generate a sequence $\{(\phi^{(\kappa)}, \theta^{(\kappa)}, v^{B,(\kappa)})\}$ of improved feasible points for (46), because $f_{\gamma,c}(\phi^{(\kappa+1)}, \theta^{(\kappa+1)}, v^{B,(\kappa+1)}) < f_{\gamma,c}(\phi^{(\kappa)}, \theta^{(\kappa)}, v^{B,(\kappa)})$ by (48), (56), and (29). This sequence is convergent by Cauchy's theorem. Moreover, for a sufficient large γ , we have $\|\phi^{(\kappa)} - v_{RF}(\theta^{(\kappa)})\| \rightarrow 0$, so $(\theta^{(\kappa)}, v^{B,(\kappa)})$ represents an optimized solution for the soft max-min rate optimization problem (44)/(45). The total computational complexity of each iteration is on the order of $\mathcal{O}(N_{RF}K) + \mathcal{O}(N)$.

IV. BASELINE PERFORMANCE OF FULLY-CONNECTED RF CHAINS

To show the advantage of AOSA we have to compare its performance to that of HBF using FC-based ABF. For the full connection of each RF chain, let $\boldsymbol{\theta} \triangleq [\theta_{n,j}]_{(n,j) \in \mathcal{N} \times \mathcal{N}_{RF}} \in [0, 2\pi]^{N \times N_{RF}}$ be the phase shift matrix. Instead of the diagonal structure (1), the FC ABF matrix is ‘structure-free’, formulated as:

$$V_{RF}(\boldsymbol{\theta}) \triangleq [e^{j\theta_{n,j}}]_{(n,j) \in \mathcal{N} \times \mathcal{N}_{RF}}. \quad (64)$$

For

$$\tilde{h}_k(\boldsymbol{\theta}) \triangleq h_k V_{RF}(\boldsymbol{\theta}) \in \mathbb{C}^{1 \times N_{RF}}, k \in \mathcal{K}, \quad (65)$$

the achievable rate of UE k is defined by (10)-(11), while the transmit constraint is

$$\sum_{k=1}^K \|V_{RF}(\boldsymbol{\theta}) \mathbf{v}_k^B\|^2 = \sum_{k=1}^K \langle [V_{RF}^H(\boldsymbol{\theta})]^2, [\mathbf{v}_k^B]^2 \rangle \leq P, \quad (66)$$

which is dependent on $\boldsymbol{\theta}$, unlikely (12). Our AOSA-related discussions of the previous sections are still relevant for FC, albeit with some more transforms involved in deriving the analytical forms of $\text{vect}[V_{RF}(\boldsymbol{\theta})]$ to find closed-form based solutions.

A. Max-min rate optimization based design

Similarly to (18), we address the problem of max-min rate optimization via the following problem of penalized optimization:

$$\max_{\boldsymbol{\Phi}, \mathbf{v}^B} \left[\min_{k=1, \dots, K} r_k(\boldsymbol{\Phi}, \mathbf{v}^B) - \gamma \|\boldsymbol{\Phi} - V_{RF}(\boldsymbol{\theta})\|^2 \right] \text{ s.t. } (2), (67a)$$

$$\sum_{k=1}^K \langle [\boldsymbol{\Phi}^H]^2, [\mathbf{v}_k^B]^2 \rangle \leq P \quad (67b)$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{N \times N_{RF}}$ is the new variable, and then we define $r_k(\boldsymbol{\Phi}, \mathbf{v}^B) \triangleq \ln \left(1 + \frac{\tilde{h}_k(\boldsymbol{\Phi}) \mathbf{v}_k^B}{\psi_k(\boldsymbol{\Phi}, \mathbf{v}^B)} \right)$ for $\tilde{h}_k(\boldsymbol{\Phi}) \triangleq h_k \boldsymbol{\Phi} \in \mathbb{C}^{N_{RF} \times N_{RF}}$, and $\psi_k(\boldsymbol{\Phi}, \mathbf{v}^B) \triangleq \sum_{\ell \neq k} \tilde{h}_k(\boldsymbol{\Phi}) \mathbf{v}_\ell^B + \sigma$, and $\gamma > 0$ is a penalty parameter.

We briefly present an alternating optimization-based procedure for the solution of (67). Initialized by $(v^{B,(0)}, \Phi^{(0)}, \theta^{(0)})$ feasible for (67), let $(v^{B,(\kappa)}, \Phi^{(\kappa)}, \theta^{(\kappa)})$ be a feasible point for (67) that is found from the $(\kappa - 1)$ -st iteration.

1) *Alternating optimization in BBF*: $v^{B,(\kappa+1)}$ is generated by solving the convex problem of

$$\max_{\mathbf{v}^B} \min_{k=1, \dots, K} r_k^{(\kappa)}(\mathbf{v}^B) \text{ s.t. } \sum_{k=1}^K \langle [(\Phi^{(\kappa)})^H]^2, [\mathbf{v}_k^B]^2 \rangle \leq P, \quad (68)$$

where $r_k^{(\kappa)}(\mathbf{v}^B)$ is a tight concave quadratic minorant of $r_k(\Phi^{(\kappa)}, \mathbf{v}^B)$ at $v^{B,(\kappa)}$ defined by

$$r_k^{(\kappa)}(\mathbf{v}^B) \triangleq \alpha_k^{(\kappa)} + 2\Re\{a_k^{(\kappa)} \mathbf{v}_k^B\} - \beta_k^{(\kappa)} \langle [h_k^H(\Phi^{(\kappa)})]^2, \sum_{\ell=1}^K [\mathbf{v}_\ell^B]^2 \rangle, \quad (69)$$

with $\alpha_k^{(\kappa)} \triangleq r_k(\Phi^{(\kappa)}, v^{B,(\kappa)}) - |x_k^{(\kappa)}|^2 / y_k^{(\kappa)} - \sigma \beta_k^{(\kappa)}$, $a_k^{(\kappa)} \triangleq (x_k^{(\kappa)})^* \tilde{h}_k(\Phi^{(\kappa)}) / y_k^{(\kappa)}$, $\beta_k^{(\kappa)} \triangleq 1 / y_k^{(\kappa)} - 1 / (y_k^{(\kappa)} + |x_k^{(\kappa)}|^2)$ for $x_k^{(\kappa)} \triangleq \tilde{h}_k(\Phi^{(\kappa)}) v_k^{B,(\kappa)}$, and $y_k^{(\kappa)} \triangleq \psi_k(\Phi^{(\kappa)}, v^{B,(\kappa)})$.

Algorithm 4 Max-min rate optimization based FC HFB algorithm

- 1: **Initialization**: Initialize $(\theta^{(0)}, \Phi^{(0)}, v^{B,(0)})$.
- 2: **Repeat until convergence of the objective function in (67)**: Generate $v^{B,(\kappa+1)}$ by solving the convex problem (68), and $\Phi^{(\kappa+1)}$ by solving the convex problem (70), and $\theta^{(\kappa+1)}$ by (72). Reset $\kappa := \kappa + 1$.
- 3: **Output** $(\theta^{opt}, \tilde{v}^{B,opt}) = (\theta^{(\kappa)}, v^{B,(\kappa)})$.

2) *Alternating optimization in Φ* : $\Phi^{(\kappa+1)}$ is generated by solving the following convex problem:

$$\max_{\boldsymbol{\Phi}} \min_{k=1, \dots, K} \tilde{r}_k^{(\kappa)}(\boldsymbol{\Phi}) - \gamma \|\boldsymbol{\Phi} - V_{RF}(\theta^{(\kappa)})\|^2 \text{ s.t. } \langle [\boldsymbol{\Phi}^H]^2, \sum_{k=1}^K [v_k^{B,(\kappa+1)}]^2 \rangle \leq P, \quad (70)$$

where $\tilde{r}_k^{(\kappa)}(\boldsymbol{\Phi})$ is a tight concave minorant of $r_k(\boldsymbol{\Phi}, v^{B,(\kappa+1)})$ at $\Phi^{(\kappa)}$ defined by

$$\tilde{r}_k^{(\kappa)}(\boldsymbol{\Phi}) \triangleq \tilde{\alpha}_k^{(\kappa)} + 2\Re\{\langle \tilde{A}_k^{(\kappa)} \boldsymbol{\Phi} \rangle\} - \tilde{\beta}_k^{(\kappa)} \langle h_k^H h_k, \boldsymbol{\Phi} \left(\sum_{\ell=1}^K [v_\ell^{B,(\kappa+1)}]^2 \right) \boldsymbol{\Phi}^H \rangle, \quad (71)$$

with $\tilde{\alpha}_k^{(\kappa)} \triangleq r_k(\Phi^{(\kappa)}, v^{B,(\kappa+1)}) - |\tilde{x}_k^{(\kappa)}|^2 / \tilde{y}_k^{(\kappa)} - \sigma \tilde{\beta}_k^{(\kappa)}$, $\tilde{A}_k^{(\kappa)} \triangleq (\tilde{x}_k^{(\kappa)})^* v_k^{B,(\kappa+1)} h_k$, $\tilde{\beta}_k^{(\kappa)} \triangleq 1 / \tilde{y}_k^{(\kappa)} - 1 / (\tilde{y}_k^{(\kappa)} + |\tilde{x}_k^{(\kappa)}|^2)$, for $\tilde{y}_k^{(\kappa)} \triangleq \tilde{h}_k(\Phi^{(\kappa)}) v_k^{B,(\kappa+1)}$ and $\tilde{x}_k^{(\kappa)} \triangleq \psi_k(\Phi^{(\kappa)}, v^{B,(\kappa+1)})$.

3) *Alternating optimization in ABF*: $\theta^{(\kappa+1)}$ is generated according to the following formula:

$$\theta_{n,j}^{(\kappa+1)} = \lfloor \angle \Phi^{(\kappa+1)}(n, j) \rfloor_b, (n, j) \in \mathcal{N} \times \mathcal{N}_{RF}. \quad (72)$$

4) *Algorithm*: Like Algorithm 1, Algorithm 4 also generates a sequence of gradually improved feasible points for (67), so its convergence is guaranteed by Cauchy’s theorem. The total computational complexity of the convex problems (68) and (70) is $\mathcal{O}(N_{RF}^3 K^3) + \mathcal{O}(N_{RF}^3 N^3)$.

B. SR maximization-based design

Similarly to (32), the problem of SR maximization is addressed via the following problem of penalized optimization:

$$\max_{\boldsymbol{\Phi}, \mathbf{v}^B} \left[\sum_{k=1}^K r_k(\boldsymbol{\Phi}, \mathbf{v}^B) - \gamma \|\boldsymbol{\Phi} - V_{RF}(\boldsymbol{\theta})\|^2 \right] \text{ s.t. } (2), (66). \quad (73)$$

1) *Alternating optimization in BBF*: $v^{B,(\kappa+1)}$ is generated by solving the problem of

$$\max_{\mathbf{v}^B} \sum_{k=1}^K r_k^{(\kappa)}(\mathbf{v}^B) \text{ s.t. } \sum_{k=1}^K \langle [(\Phi^{(\kappa)})^H]^2, [\mathbf{v}_k^B]^2 \rangle \leq P, \quad (74)$$

with $r_k^{(\kappa)}(\mathbf{v}^B)$ defined from (69). By expressing

$$\sum_{k=1}^K r_k^{(\kappa)}(\mathbf{v}^B) = \alpha^{(\kappa)} + \sum_{k=1}^K 2\Re\{a_k^{(\kappa)} \mathbf{v}_k^B\} - \sum_{k=1}^K \langle \Xi^{(\kappa)}, [\mathbf{v}_k^B]^2 \rangle,$$

with $\alpha^{(\kappa)} \triangleq \sum_{k=1}^K \alpha_k^{(\kappa)}$ and $\Xi^{(\kappa)} \triangleq \sum_{k=1}^K \beta_k^{(\kappa)} [h_k^H(\Phi^{(\kappa)})]^2$, the problem (74) admits the closed-form solution of

$$v_k^{B,(\kappa+1)} = \begin{cases} (\Xi^{(\kappa)})^{-1} (a_k^{(\kappa)})^H \\ \text{if } \sum_{k=1}^K \|\Phi^{(\kappa)} (\Xi^{(\kappa)})^{-1} (a_k^{(\kappa)})^H\|^2 \leq P, \\ (\Xi^{(\kappa)} + \lambda[(\Phi^{(\kappa)})^H]^2)^{-1} (a_k^{(\kappa)})^H \\ \text{otherwise,} \end{cases} \quad (75)$$

where $\lambda > 0$ is found by bisection, so that $\|\Phi^{(\kappa)} (\Xi^{(\kappa)} + \lambda[(\Phi^{(\kappa)})^H]^2)^{-1} (a_k^{(\kappa)})^H\|^2 = P$.

2) *Alternating optimization in Φ* : $\Phi^{(\kappa+1)}$ is generated by solving the following problem:

$$\begin{aligned} \max_{\Phi} \sum_{k=1}^K \tilde{r}_k^{(\kappa)}(\Phi) - \gamma \|\Phi - V_{RF}(\theta^{(\kappa)})\|^2 \\ \text{s.t. } \langle \Phi^H \Phi, \sum_{k=1}^K [v_k^{B,(\kappa+1)}]^2 \rangle \leq P, \end{aligned} \quad (76)$$

with $\tilde{r}_k^{(\kappa)}(\Phi)$ defined from (71). For $\phi = \text{vec}(\Phi)$, by using the identity $h_k \Phi v_\ell^{B,(\kappa+1)} = h_{k,\ell}^{(\kappa+1)} \phi$ with $h_{k,\ell}^{(\kappa+1)} \triangleq (v_\ell^{B,(\kappa+1)})^T \otimes h_k \in \mathbb{C}^{1 \times (NN_{RF})}$, we formulate

$$\tilde{r}_k^{(\kappa)}(\Phi) = \tilde{a}_k^{(\kappa)} + 2\Re\{\tilde{a}_k^{(\kappa)} \phi\} - \langle \tilde{\Xi}^{(\kappa)}, [\phi]^2 \rangle,$$

for $\tilde{a}_k^{(\kappa)} = \text{vec}[(\tilde{A}_k^{(\kappa)})^T]^T$ and $\tilde{\Xi}^{(\kappa)} \triangleq \beta_k^{(\kappa)} \sum_{\ell=1}^K [(h_{k,\ell}^{(\kappa+1)})^H]^2$. Then

$$\begin{aligned} \sum_{k=1}^K \tilde{r}_k^{(\kappa)}(\Phi) - \gamma \|\Phi - V_{RF}(\theta^{(\kappa)})\|^2 \\ = \tilde{\alpha}^{(\kappa)} + 2\Re\{\tilde{a}^{(\kappa)} \phi\} - \langle \tilde{\Xi}^{(\kappa)}, [\phi]^2 \rangle, \end{aligned}$$

for $\tilde{\alpha}^{(\kappa)} \triangleq \sum_{k=1}^K \tilde{a}_k^{(\kappa)} - \gamma NN_{RF}$ and $\tilde{a}^{(\kappa)} \triangleq \sum_{k=1}^K \tilde{a}_k^{(\kappa)} + \gamma(\text{vec}(V_{RF}(\theta^{(\kappa)})))^H$, $\tilde{\Xi}^{(\kappa)} \triangleq \sum_{k=1}^K \tilde{\Xi}_k^{(\kappa)} + \gamma I_{NN_{RF}}$.

Furthermore, we have $\Phi \sum_{k=1}^K v_k^{B,(\kappa+1)} = \mathcal{A}^{(\kappa+1)} \phi$, for $\mathcal{A}^{(\kappa+1)} \triangleq \left(\sum_{k=1}^K v_k^{B,(\kappa+1)} \right)^T \otimes I_N$, so the problem (76) is reformulated as

$$\begin{aligned} \max_{\phi} \tilde{\alpha}^{(\kappa)} + 2\Re\{\tilde{a}^{(\kappa)} \phi\} - \phi^H \tilde{\Xi}^{(\kappa)} \phi \\ \text{s.t. } \langle [\mathcal{A}^{(\kappa+1)}]^H, [\phi]^2 \rangle \leq P, \end{aligned} \quad (77)$$

which admits the closed-form solution of

$$\phi^{(\kappa+1)} = \begin{cases} (\tilde{\Xi}^{(\kappa)})^{-1} (\tilde{a}^{(\kappa)})^H \\ \text{if } \|\mathcal{A}^{(\kappa+1)} (\tilde{\Xi}^{(\kappa)})^{-1} (\tilde{a}^{(\kappa)})^H\|^2 \leq P \\ (\tilde{\Xi}^{(\kappa)} + \lambda[(\mathcal{A}^{(\kappa+1)})^H]^2)^{-1} (\tilde{a}^{(\kappa)})^H \\ \text{otherwise,} \end{cases} \quad (78)$$

where $\lambda > 0$ is found by bisection, so that $\|\mathcal{A}^{(\kappa+1)} (\tilde{\Xi}^{(\kappa)} + \lambda[(\mathcal{A}^{(\kappa+1)})^H]^2)^{-1} (\tilde{a}^{(\kappa)})^H\|^2 = P$.

3) *Algorithm*: Thus in parallel to Algorithm 2, Algorithm 5 presents a scalable computational procedure for the solution of (73), with the total computational complexity of each iteration being on the order of $\mathcal{O}(N_{RF}K) + \mathcal{O}(N_{RF}N)$.

Algorithm 5 Scalable SR maximization-based FC HBF algorithm

- 1) **Initialization**: Initialize $(\theta^{(0)}, \Phi^{(0)}, v^{B,(0)})$.
- 2) **Repeat until convergence of the objective function (73)**: Generate $v^{B,(\kappa+1)}$ by (75), and $\Phi^{(\kappa+1)}$ by (78), and $\theta^{(\kappa+1)}$ by (72). Reset $\kappa := \kappa + 1$.
- 3) **Output** $(\theta^{opt}, \tilde{v}^{B,opt}) = (\theta^{(\kappa)}, v^{B,(\kappa)})$.

C. *Soft max-min rate optimization-based design*

Similarly to (46), the soft max-min rate problem is addressed via the following problem of penalized optimization:

$$\min_{\theta, \mathbf{v}^B, \Phi} [\ln \pi_c(\Phi, \mathbf{v}^B) + \gamma \|\Phi - V_{RF}(\theta)\|^2] \quad \text{s.t.} \quad (2), (67b). \quad (79)$$

where $\gamma > 0$ is a penalty parameter, and

$$\pi_c(\Phi, \mathbf{v}^B) \triangleq \sum_{k=1}^K \left(1 - \frac{|\tilde{h}_k(\Phi) \mathbf{v}_k^B|^2}{|\tilde{h}_k(\Phi) \mathbf{v}_k^B|^2 + c \psi_k(\Phi, \mathbf{v}^B)} \right).$$

1) *Alternating optimization in BBF*: $v^{B,(\kappa+1)}$ is generated by solving the following problem:

$$\min_{\mathbf{v}^B} \rho^{(\kappa)}(\mathbf{v}^B) \quad \text{s.t.} \quad \sum_{k=1}^K \langle (\Phi^{(\kappa)})^H \Phi^{(\kappa)}, [\mathbf{v}_k^B]^2 \rangle \leq P, \quad (80)$$

where $\rho^{(\kappa)}(\mathbf{v}^B)$ is a tight majorant of $\ln \pi_c(\Phi^{(\kappa)}, \mathbf{v}^B)$ defined by

$$\rho^{(\kappa)}(\mathbf{v}^B) \triangleq a^{(\kappa)} + 2 \sum_{k=1}^K \Re\{b_k^{(\kappa)} \mathbf{v}_k^B\} - \sum_{k=1}^K \langle C_k^{(\kappa)}, [\mathbf{v}_k^B]^2 \rangle$$

with

$$\begin{aligned} a^{(\kappa)} &\triangleq \ln \pi_c(\Phi^{(\kappa)}, v^{B,(\kappa)}) + \sum_{k=1}^K d_k^{(\kappa)} |\bar{x}_k|^2 + c\sigma \sum_{k=1}^K c_k^{(\kappa)}, \\ b_k^{(\kappa)} &\triangleq d_k^{(\kappa)} (\tilde{h}_k(\Phi^{(\kappa)}) v_k^{B,(\kappa)})^* \tilde{h}_k(\Phi^{(\kappa)}), \\ C_k^{(\kappa)} &\triangleq c \sum_{\ell \in \mathcal{K} \setminus \{k\}} c_\ell^{(\kappa)} [h_\ell^H(\Phi^{(\kappa)})]^2 + c_k^{(\kappa)} [h_k^H(\Phi^{(\kappa)})]^2, \end{aligned}$$

for $d_k^{(\kappa)} \triangleq \frac{(c\bar{y}_k + |\bar{x}_k|^2)^{-1}}{\pi_c(\Phi^{(\kappa)}, v^{B,(\kappa)})}$, $c_k^{(\kappa)} \triangleq d_k^{(\kappa)} \frac{|\bar{x}_k|^2}{c\bar{y}_k + |\bar{x}_k|^2}$, $\bar{x}_k \triangleq \tilde{h}_k(\Phi^{(\kappa)}) v^{B,(\kappa)}$, and $\bar{y}_k \triangleq \psi_k(\Phi^{(\kappa)}, v^{B,(\kappa)})$. The problem (80) admits the closed-form solution of

$$v_k^{B,(\kappa+1)} = \begin{cases} (C_k^{(\kappa)})^{-1} (b_k^{(\kappa)})^H \\ \text{if } \sum_{k=1}^K \langle [(\Phi^{(\kappa)})^H]^2, [(C_k^{(\kappa)})^{-1} (b_k^{(\kappa)})^H]^2 \rangle \leq P, \\ (C_k^{(\kappa)} + \lambda[(\Phi^{(\kappa)})^H]^2)^{-1} (b_k^{(\kappa)})^H \\ \text{otherwise,} \end{cases} \quad (81)$$

where $\lambda > 0$ is found by bisection, so that $\sum_{k=1}^K \langle [(\Phi^{(\kappa)})^H]^2, [(C_k^{(\kappa)} + \lambda[(\Phi^{(\kappa)})^H]^2)^{-1} (b_k^{(\kappa)})^H]^2 \rangle = P$.

2) *Alternating optimization in Φ* : $\phi^{(\kappa+1)} \triangleq \text{vect}(\Phi^{(\kappa+1)})$ is generated by solving the problem

$$\begin{aligned} \max_{\phi \triangleq \text{vect}(\Phi)} \tilde{\rho}^{(\kappa)}(\phi) - \gamma \|\phi - v_{RF}(\theta^{(\kappa)})\|^2 \\ \text{s.t. } \|\mathcal{A}^{(\kappa+1)} \phi\|^2 \leq P, \end{aligned} \quad (82)$$

Algorithm 6 Scalable soft max-min optimization-based FC HBF algorithm

- 1: **Initialization:** Initialize $(\theta^{(0)}, \Phi^{(0)}, v^{B,(0)})$.
 - 2: **Repeat until convergence of the objective function in (79):** Generate $v^{B,(\kappa+1)}$ by (81), $\Phi^{(\kappa+1)}$ by (83), and $\theta^{(\kappa+1)}$ by (30). Reset $\kappa := \kappa + 1$.
 - 3: **Output** $(\theta^{opt}, v^{B,opt}) = (\theta^{(\kappa)}, v^{B,(\kappa)})$.
-

where $\tilde{\rho}^{(\kappa)}(\phi)$ is a tight majorant of $\ln \pi_c(\Phi, v^{B,(\kappa+1)})$ defined by

$$\tilde{\rho}^{(\kappa)}(\phi) \triangleq \tilde{a}^{(\kappa)} + 2\Re\{\tilde{b}^{(\kappa)}\phi\} - \phi^H \tilde{C}^{(\kappa)}\phi$$

with

$$\begin{aligned} \tilde{a}^{(\kappa)} &\triangleq \ln \pi_c(\Phi^{(\kappa)}, v^{B,(\kappa+1)}) - \sum_{k=1}^K \tilde{d}_k^{(\kappa)} |\bar{x}_k|^2 \\ &\quad - c\sigma \sum_{k=1}^K \tilde{c}_k^{(\kappa)}, \\ \tilde{b}^{(\kappa)} &\triangleq \sum_{k=1}^K \left(\tilde{d}_k^{(\kappa)} (h_{k,k}^{(\kappa+1)} \varphi^{(\kappa)})^* h_{k,k}^{(\kappa+1)} \right), \\ \tilde{C}^{(\kappa)} &\triangleq \sum_{k=1}^K \tilde{c}_k^{(\kappa)} \left(c \sum_{\ell \in \mathcal{K} \setminus \{k\}} [(h_{k,\ell}^{(\kappa+1)})^H]^2 + [(h_{k,k}^{(\kappa+1)})^H]^2 \right), \end{aligned}$$

for $\tilde{d}_k^{(\kappa)} \triangleq \frac{(c\bar{y}_k + |\bar{x}_k|^2)^{-1}}{\pi_c(\Phi^{(\kappa)}, v^{B,(\kappa+1)})}$, $\tilde{c}_k^{(\kappa)} \triangleq \tilde{d}_k^{(\kappa)} \frac{|\bar{x}_k|^2}{c\bar{y}_k + |\bar{x}_k|^2}$, and $\bar{x}_k = \bar{h}_k(\Phi^{(\kappa)}) v_k^{B,(\kappa+1)}$, $\bar{y}_k = \psi_k(\Phi^{(\kappa)}, v^{B,(\kappa+1)})$.

The problem (82) admits the closed-form solution of

$$\phi^{(\kappa+1)} = \begin{cases} (\tilde{C}^{(\kappa)} + \gamma I_{NN_{RF}})^{-1} ((b^{(\kappa)})^H + \gamma v_{RF}(\theta^{(\kappa)})) \\ \text{if } \sum_{k=1}^K \|\mathcal{A}^{(\kappa+1)} (\tilde{C}^{(\kappa)} + \gamma I_{NN_{RF}})^{-1} \\ \quad ((b^{(\kappa)})^H + \gamma v_{RF}(\theta^{(\kappa)}))\|^2 \leq P, \\ (C_k^{(\kappa)} + \gamma I_{NN_{RF}} + \lambda [(\mathcal{A}^{(\kappa+1)})^H]^2)^{-1} \\ \quad ((b^{(\kappa)})^H + \gamma v_{RF}(\theta^{(\kappa)})) \\ \text{otherwise,} \end{cases} \quad (83)$$

where $\lambda > 0$ is found by bisection, so that $\|\mathcal{A}^{(\kappa+1)} (C_k^{(\kappa)} + \gamma I_{NN_{RF}} + \lambda [(\mathcal{A}^{(\kappa+1)})^H]^2)^{-1} ((b^{(\kappa)})^H + \gamma v_{RF}(\theta^{(\kappa)}))\|^2 = P$.

3) *Alternating optimization in ABF:* Generate $\theta^{(\kappa+1)}$ according (30).

4) *Algorithm:* Thus in parallel to Algorithm 3, Algorithm 6 presents a scalable computational procedure for the solution of (79), with the total computational complexity of each iteration being the same as that of Algorithm 5.

V. NUMERICAL RESULTS

This section analyzes the performance of the proposed algorithms along with their computational convergence. The number of downlink transmit antennas (DL-TA) at the BS is $N = 72$, and that of UEs is $K = 8$. All the UEs are randomly placed in a cell of radius 200 meters. The path-loss of UE k experienced at a distance d_k from the BS is set to $\rho_k = 36.72 + 35.3 \log_{10}(d_k)$ (in dB), taking into account a 16.5 dB gain provided by multiple-antenna aided mmWave transmission [39]–[41]. The channel

$h_k \in \mathbb{C}^{1 \times N}$ between the BS and UE k is modelled by [42] $h_k = F \sqrt{10^{-\rho_k/10}} \sum_{c=1}^{N_c} \sum_{\ell=1}^{N_{sc}} \alpha_{k,c,\ell} a(\phi_{k,c,\ell}, \theta_{k,c,\ell})$, where $F = \sqrt{\frac{N}{N_c N_{sc}}}$, N_c is the number of scattered clusters, N_{sc} is the number of scatterers within each cluster, and $\alpha_{k,c,\ell} \sim \mathcal{CN}(0, 1)$ is the complex gain of the ℓ th path in the c th cluster between the BS and UE k . We set $N_c = 5$ and $N_{sc} = 10$ as in [40]. Assuming a uniform planar antenna array configuration having half-wavelength antenna spacing with N_1 and N_2 elements in the horizontal and vertical dimensions, respectively, the normalized antenna array response vectors $a(\phi_{k,c,\ell}, \theta_{k,c,\ell})$ is defined as

$$\begin{aligned} a(\phi_{k,c,\ell}, \theta_{k,c,\ell}) &= \frac{1}{\sqrt{N}} \left(1, e^{j\pi(x \sin(\phi_{k,c,\ell}^t) \sin(\theta_{k,c,\ell}^t) + y \cos(\theta_{k,c,\ell}^t))}, \dots, \right. \\ &\quad \left. e^{j\pi((N_1-1) \sin(\phi_{k,c,\ell}^t) \sin(\theta_{k,c,\ell}^t) + (N_2-1) \cos(\theta_{k,c,\ell}^t))} \right)^T, \end{aligned}$$

where $0 \leq x \leq (N_1 - 1)$ and $0 \leq y \leq (N_2 - 1)$, $\phi_{k,c,\ell}$ and $\theta_{k,c,\ell}$ are the azimuth angle and elevation angle of departure for the ℓ th path in the c th cluster arriving from the BS to the UE k , respectively. The angles are generated using the Laplacian distribution in combination with random mean cluster angles in the interval of $[0, 2\pi)$ and a 10-degree spread for each cluster. Assuming a carrier frequency of 28 GHz, the noise power density is -174 dBm/Hz. The results are multiplied by $\log_2 e$ to convert the unit nats/sec into the unit bps/Hz. The algorithm terminates, when the penalty term falls below 10^{-2} .

We use the following legends to specify the proposed implementations:

- For the AOSA based algorithms, “max-min” and “3-bit max-min” refer to the convex-solver-based Algorithm 1 with the ABF matrix having ∞ resolution and 3-bit resolution, respectively; “soft max-min” and “3-bit soft max-min” refer to the scalable Algorithm 3 with the ABF matrix having ∞ resolution and 3-bit resolution, respectively; “SR” and “3-bit SR” refer to Algorithm 2 with the ABF matrix having ∞ resolution and 3-bit resolution, respectively.
- For the FC HBF algorithms, “FC max-min” refers to Algorithm 4 with the ABF matrix having ∞ resolution; “FC soft max-min” refers to Algorithm 6 with the ABF matrix having ∞ resolution; “FC SR” refers to Algorithm 5 with the ABF matrix having ∞ resolution; “FC RZFB max-min” refers to the regularized zero-forcing beamforming (RZFB)-aided max-min rate algorithm proposed in [43]. The superior performance of FC RZFB over the algorithms in [11] and [17] has been demonstrated by [43, Fig. 1] and [43, Figs. 3, 6, 8], respectively. Moreover, the superior performance of the algorithm in [17] over other existing algorithms [11], [44]–[46] has been demonstrated by [17, Fig. 14]. Thus FC RZFB max-min serves as the baseline algorithm, since it outperforms the existing algorithms [11], [44]–[46];³ “FC Shi-Hong” refers to the

³It should be noted that in this paper we only consider cases of low $N_{RF} \leq K$, for which the zero-forcing beamforming-aided algorithm proposed in [27] is not applicable.

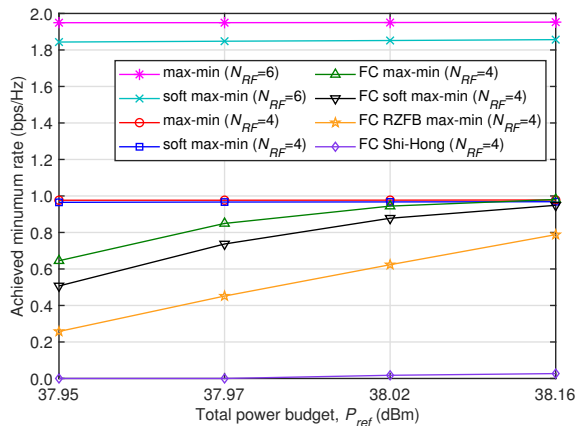


Fig. 1: The achievable MR vs. the total power P_{ref} .

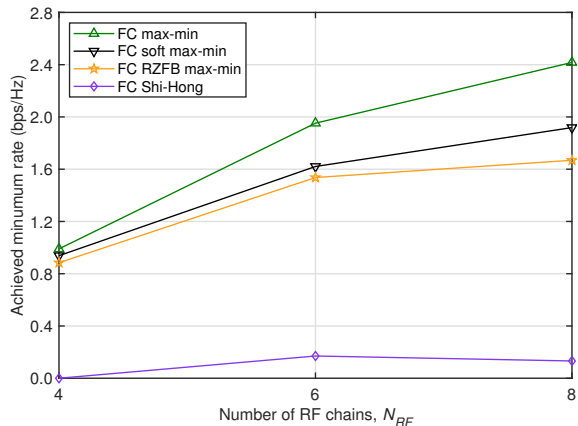


Fig. 2: MR of max-min and soft max-min vs. existing algorithms under different number of RF chains N_{RF} .

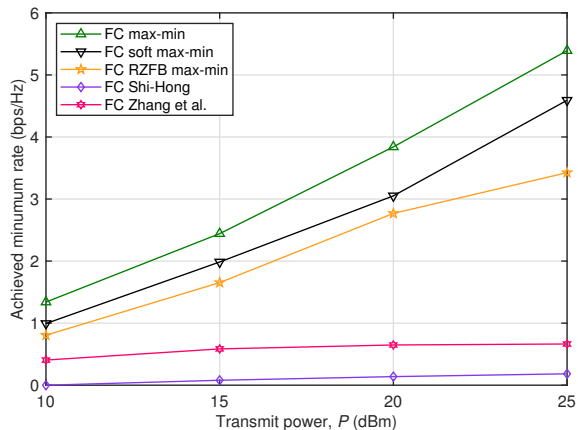


Fig. 3: MR of max-min and soft max-min vs. existing algorithms under $N_{RF} = K$ and different transmit power P .

SR algorithm proposed in [14]; “FC Zhang et al.” refers to the algorithm proposed in [18], [19], which only works for the case of $N_{RF} = K$.

A. AOSA vs. FC

We start by evaluating the performance of the AOSA and FC structures. To ensure a fair comparison, we use the total

power in the case of FC as the reference, which is defined as

$$P_{ref} = \bar{P} + \bar{N}_{RF} \times 118 + N\bar{N}_{RF} \times 20, \quad (84)$$

where \bar{P} is the transmit power in mW, \bar{N}_{RF} is the original number of RF chains used for FC, 118 mW is the circuit power consumption per RF chain, and 20 mW is the circuit power consumption per phase shifter [10]. To determine the total power budget P_{ref} , we fix \bar{N}_{RF} to 4 and vary the transmit power \bar{P} from 10 dBm (10 mW) to 25 dBm (316 mW) at 5 dBm (3 mW) intervals. Then the relationship between the transmit power P and N_{RF} for the AOSA structure is

$$P + N_{RF} \times 118 + N \times 20 = P_{ref}. \quad (85)$$

Table II compares the SR maximized by the SR maximization-based Algorithms 2 and 5 at $P_{ref} = 38.02$ dBm ($\bar{P} = 20$ dBm). The AOSA achieves a higher SR than the FC, because the former enables us to allocate much more transmit power and exploit more RF chains, hence resulting in more effective spatial DL beamforming under the same total power budget.

We define the negligible rate of less than 0.001 bps/Hz as zero rate (ZR). Table III displays the number of ZR users under maximizing the SR. The results demonstrate that SR maximization is not suitable for MU services, although increasing the number of RF chains for transmitting more data streams also goes some way towards reducing the average number of ZR UEs. Moreover, the average number of ZR UEs is 0.2 when $N_{RF} = 8$, indicating that increasing the number of RF chains mitigates the problem prevent them in most cases. However, the transmit powers P required by the AOSA-based SR algorithms with $N_{RF} = 4$, $N_{RF} = 6$, and $N_{RF} = 8$ are 36.45 dBm, 36.22 dBm, and 35.96 dBm, respectively, which are impractically high.

In Fig. 1, we present a performance comparison between the two structures using our max-min-based algorithms and the existing algorithms of [14], [43]. AOSA outperforms FC in terms of the achieved MR and the performance gap becomes wider, when more RF chains are utilized. Furthermore, our proposed soft max-min algorithm is capable of achieving MR that is comparable to those obtained by the convex-solver-based max-min algorithm. It follows from (84) and (85) that under the same P_{ref} , the transmit power P of AOSA in (85) is very high compared to that of its FC counterpart \bar{P} in (84). The former is not sensitive to the value of the latter in the interval of 10 dBm to 25 dBm. This is why the performance of AOSA is seen to be flat in Fig. 1. Additionally, both the FC RZFB max-min and FC Shi-Hong algorithms were implemented under the FC structure associated with $N_{RF} = 4$, as dictated by the specific total power budget P_{ref} , and they were also characterized in Fig. 1. We can observe that both the FC RZFB max-min and FC Shi-Hong algorithms are outperformed by our proposed max-min and soft max-min algorithms in terms of their MR.

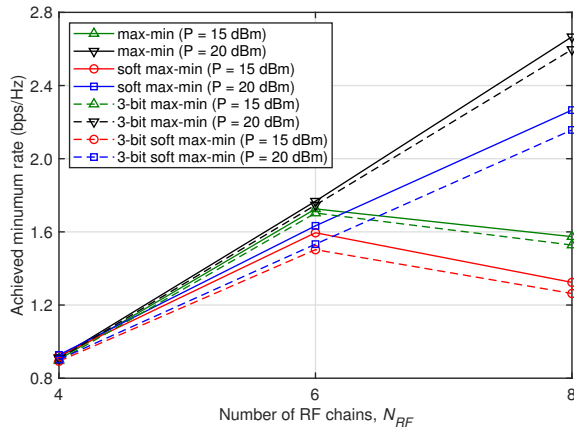
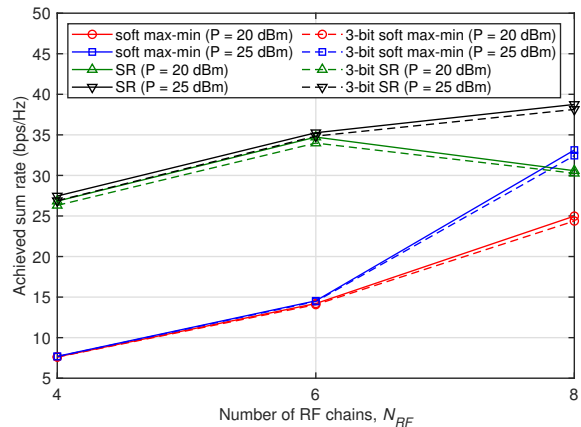
Fig. 2 facilitates a comprehensive analysis by comparing our proposed max-min and soft max-min algorithms to the FC RZFB max-min and FC Shi-Hong algorithms, under a fixed total transmit power P_{ref} of 40.97 dBm upon varying the number of RF chains. In this context, P_{ref} is calculated

TABLE II: The SR achieved upon maximizing it at $P_{ref} = 38.02$ dBm

	SR ($N_{RF} = 8$)	SR ($N_{RF} = 6$)	SR ($N_{RF} = 4$)	FC SR ($N_{RF} = 4$)
Achieved SR (bps/Hz)	63.9	51.4	35.8	33.2

TABLE III: The average number of ZRs in maximizing SR at $P_{ref} = 38.02$ dBm

	SR ($N_{RF} = 8$)	SR ($N_{RF} = 6$)	SR ($N_{RF} = 4$)	FC SR ($N_{RF} = 4$)
The average # of ZRs	0.2	2.0	4.0	4.0

Fig. 4: The achieved MR vs. the number of RF chains N_{RF} .Fig. 5: The achieved SR vs. the number of RF chains N_{RF} .TABLE V: Min-rate/max-rate and Jains fairness index of user-rate allocation with $N_{RF} = 8$ and $P = 20$ dBm

	soft max-min	3-bit soft max-min
Min-rate/max-rate	0.46	0.45
Jain's fairness index	0.91	0.90

by setting the number of RF chains to 8 and P to 15 dBm. Notably, both the max-min and soft max-min algorithms outperform the FC RZFB max-min algorithm for all the values of N_{RF} considered. Furthermore, they exhibit a significant performance advantage over the SR maximization-based FC Shi-Hong algorithm.

In Fig. 3, we compare our proposed max-min and soft max-min algorithms to the existing algorithms of [14], [18], [19], [43] while varying the transmit power P . To simulate the FC Zhang algorithm [18], [19], we set $N_{RF} = K = 8$. It is observed that our proposed max-min and soft max-min algorithms exhibit superior performance compared to the others. This is because the FC Zhang algorithm focuses on enhancing the signal energy, rather than effectively mitigating the multi-user interference, hence resulting in a lower minimum rate.

Hence, from now on, we will utilize AOSA instead of the FC structure. Additionally, we will focus on the max-min-based algorithms because of the deficiency of the ZR SR maximization based algorithms.

B. The AOSA performance under low-resolution ABF

In this subsection, we initially evaluate the performance by varying the numbers of RF chains. To ensure a fair comparison, we maintain a constant total power budget P_{ref} , while increasing the number of RF chains. This results in additional “virtual antennas” for increased angular flexibility,

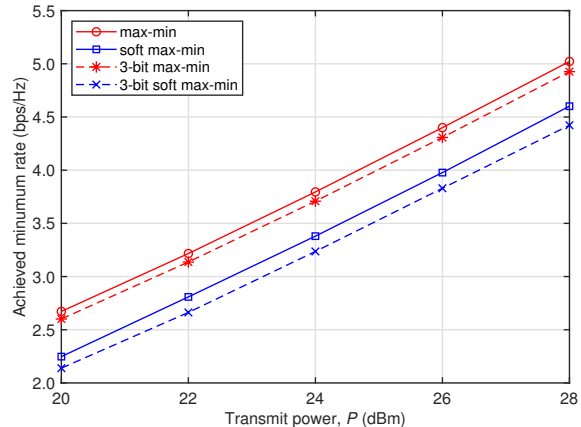
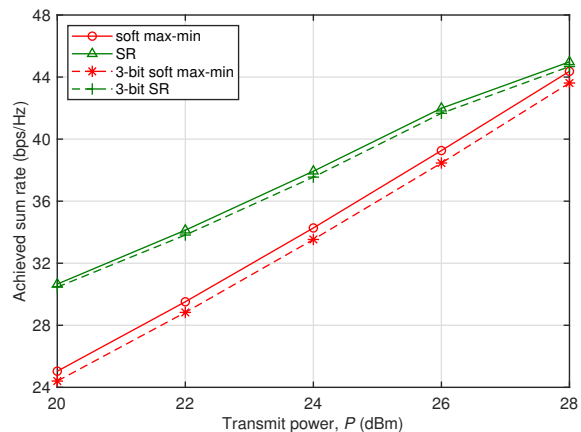
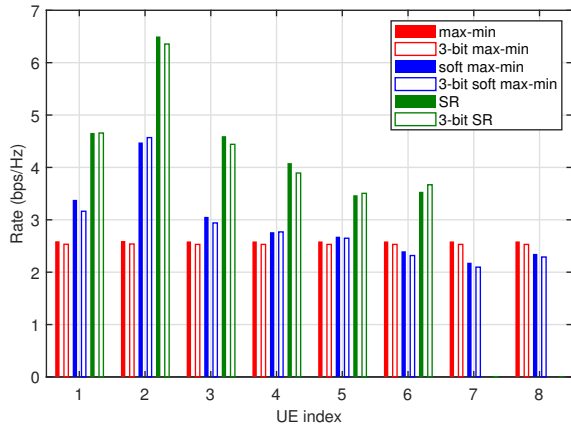
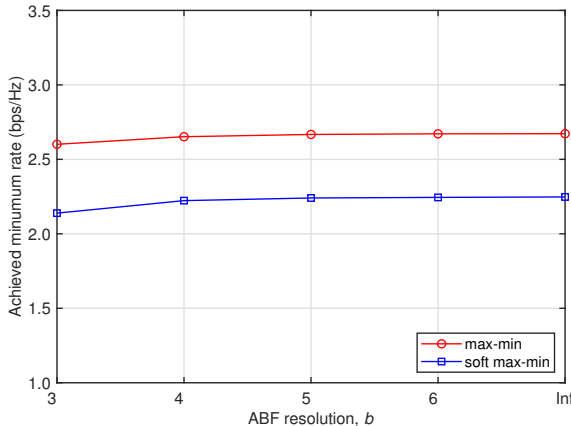
Fig. 6: The achieved MR vs. the transmit power P .Fig. 7: The achieved SR vs. the transmit power P .

TABLE IV: The average number of ZRs in maximizing SR with moderate power allocation

	SR ($N_{RF} = 8$)	SR ($N_{RF} = 6$)	SR ($N_{RF} = 4$)
The average # of ZRs	1.6	2.1	4.0

Fig. 8: The distribution of user-rates at $N_{RF} = 8$ and $P = 20$ dBm.Fig. 9: The achieved MR vs. ABF resolution b .

but at a reduced transmit power budget P . Thus, we have to find a balance between the number of RF chains N_{RF} and the transmit power budget P . For calculating the total reference power P_{ref} , we set the number of RF chains to 8 and P to 15 dBm. The resultant total power budget is $P_{ref} = 33.83$ dBm. Fig. 4 shows the MR achieved by the max-min-based algorithms. When we reduce the number of RF chains from 8 to 6, more transmit power is allocated at the baseband, which results in a higher MR achieved by $N_{RF} = 6$ than by $N_{RF} = 8$. However, despite the higher transmit power budget at $N_{RF} = 4$, using 8 RF chains still outperforms using 4 RF chains, demonstrating the advantages of utilizing more RF chains for enhancing the digital beamforming part. We then increase the transmit power P to investigate the conditions under which digital beamforming using 8 RF chains can be fully leveraged. At $P = 20$ dBm, where $P_{ref} = 33.95$ dBm, the achievable MR associated with $N_{RF} = 8$ substantially outperforms the configurations using fewer RF chains but higher transmit power allocation. Essentially, when the total

power budget is limited, we opt for utilizing a smaller number of RF chains to allow for having an adequate transmission power. However, to fully exploit digital beamforming utilizing a larger number of RF chains, a higher transmit power is necessary.

Then in Fig. 5, we compare the SRs obtained by the soft max-min algorithm and the SR maximization-based algorithm. It can be observed that increasing the number of RF chains and the BS transmit power also leads to an increase in SR for the soft max-min algorithm. However, the SR maximization-based algorithm requires a higher transmit power P to fully exploit the benefits of utilizing 8 RF chains. As the transmit power budget P is increased to 25 dBm, the benefits of 8 RF chains become more substantial than those of 6 RF chains for the SR maximization-based algorithm. It is noteworthy that when the transmit power budget P is set to 25 dBm and N_{RF} is increased from 6 to 8, the performance gap between the soft max-min algorithm and the SR maximization-based algorithm narrows significantly. This inspires us to improve the performance of the soft max-min algorithm in terms of its SR by utilizing a larger number of RF chains and a higher transmit power budget, allowing it to match the performance of the SR maximization-based algorithm. Additionally, Fig. 4 and Fig. 5 illustrate that the performance of MR and SR using 3-bit resolution is comparable to that of the ∞ resolution scheme, which is also evident from Fig. 9.

Furthermore, we present Table IV to summarize the average number of ZR UEs resulting from maximizing the SR under moderate BS transmit power in Fig. 5. Compared to the observation made concerning Table III, we can see that when the transmit power P of the SR algorithms using $N_{RF} = 4$, $N_{RF} = 6$, and $N_{RF} = 8$ reduced to 27.57 dBm, 25.26 dBm, and 20 dBm, respectively, maximizing the SR using $N_{RF} = 8$ still results in low-rate connections or even in ZR UEs in energy-efficient signal transmission scenarios.

Fig. 6 depicts the MR achieved by the max-min-based algorithms as the transmit power P at $N_{RF} = 8$. Similar to the observation in Fig. 4, the MR achieved using 3-bit resolution is comparable to that of the ∞ resolution, and the soft max-min algorithm achieves a slightly lower MR than to the max-min algorithm. In Fig. 7, we compare the SR obtained by the max-min-based algorithms to that of the SR maximization-based Algorithm 2 for $N_{RF} = 8$. The soft max-min algorithm achieves much higher SR than the max-min algorithm. It is worth noting that as the transmit power P increases, the SR obtained by the soft max-min algorithm approaches that of the SR maximization-based algorithm. In other words, we can approach the optimal MR and SR by the proposed soft max-min algorithm. This is a surprise, because it is commonly maintained that the MR and SR performances are conflicting, i.e. one of them must be sacrificed to improve the other.

To demonstrate the ability of our proposed soft max-min algorithm to achieve a fair rate allocation, Fig. 8 portrays

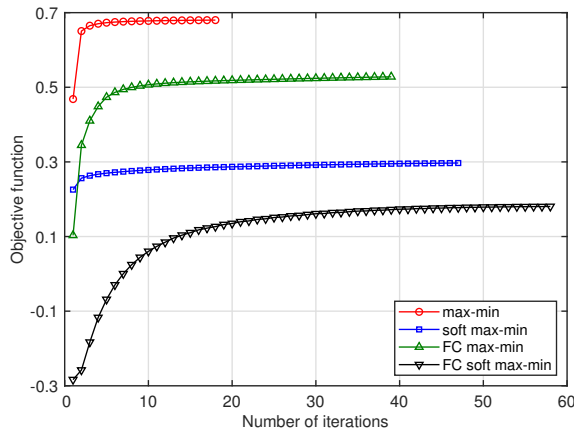


Fig. 10: Convergence of the algorithms

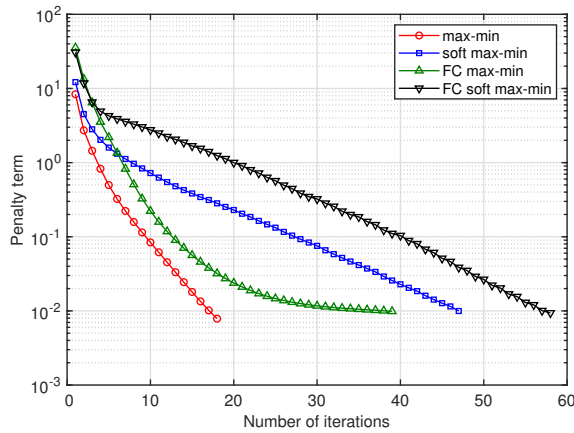


Fig. 11: Convergence of penalty terms

the user-rate distribution pattern obtained for $N_{RF} = 8$ and $P = 20$ dBm. We can observe that the soft max-min algorithm achieves a MR that is comparable to that of the max-min algorithm, while maintaining a good SR. By contrast, maximizing the SR results in the allocation of ZR, thereby literally disconnecting certain UEs. To provide a more detailed analysis, Table V quantifies the fairness of the user-rate distribution by evaluating both the ratio of min-rate to max-rate and Jain's fairness index of user-rate allocation [47] for $N_{RF} = 8$ and $P = 20$ dBm. The results show that the soft max-min algorithm yields a Jain's fairness index that is closer to one, indicating that it achieves a distribution of user-rates that is nearly uniform.

Fig. 9 shows the MR achieved for different resolutions of the ABF, given $N_{RF} = 8$ and $P = 20$ dBm. The performance achieved using 5-bit and 6-bit resolutions is nearly indistinguishable from that of the ∞ resolution case. However, for 3-bit resolution, the max-min algorithm shows an approximate reduction of 3% compared to the ∞ resolution, while the soft max-min algorithm shows an approximate reduction of 5% compared to the ∞ resolution.

C. Algorithmic convergence

Finally, we characterize the convergence of the proposed algorithms. In our simulations, the coefficient c used in soft

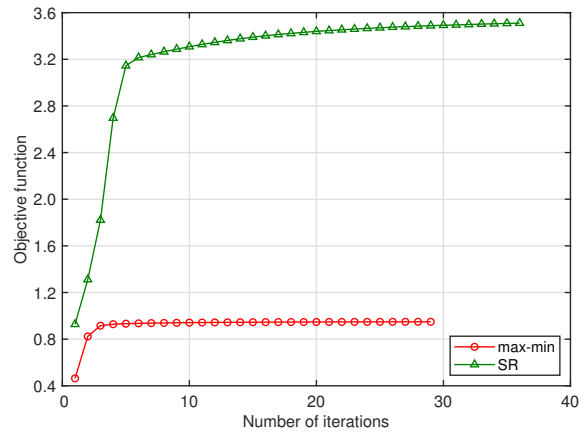


Fig. 12: Convergence of the objective functions in (13) and (31)

TABLE VI: The achieved MR vs. c .

	$c = 1$	$c = 0.1$	$c = 0.01$
soft max-min	2.1118	2.2351	1.5217
3-bit soft max-min	2.0440	2.1388	1.4468

max-min based algorithms is set to 0.1. It should be noted that the setting of c is not constant and should be selected appropriately based on the specific scenario. We present Table VI to illustrate the impact of c on the MR achieved. The table reveals that the highest MR is attained when $c = 0.1$. To ensure a reasonable convergence speed with the penalty parameter γ , we begin by selecting an initial value for γ so that the penalty term's magnitude is comparable to that of the objective. As the iterations progress, we gradually increase the value of γ . For instance, let's consider the penalty parameter γ for implementing Algorithm 1. We randomly generate $\phi^{(0)}$ with the modulus of its entries being less than 1 and $v^{B,(0)}$ satisfying the power constraint (12). Then the triplet $(\theta^{(0)}, \phi^{(0)}, v^{B,(0)})$ with $\theta_{\ell,j}^{(0)} = \lfloor \angle \phi_{\ell,j}^{(0)} \rfloor_b$, $(\ell, j) \in \mathcal{L} \times \mathcal{N}_{RF}$ (see (30)) is clearly a feasible point for the problem (18). For implementing the first iteration, we set $\gamma = f(\phi^{(0)}, v^{B,(0)}) / \|\phi^{(0)} - v_{RF}(\theta^{(0)})\|^2$ ensuring that the objective function $f(\phi^{(0)}, v^{B,(0)})$ is of the same magnitude as the penalty term $\gamma \|\phi^{(0)} - v_{RF}(\theta^{(0)})\|^2$. As the iterative process continues, we update $\gamma \rightarrow 1.2\gamma$, whenever $\|\phi^{(\kappa+1)} - v_{RF}(\theta^{(\kappa+1)})\|^2 > 0.9 \|\phi^{(\kappa)} - v_{RF}(\theta^{(\kappa)})\|^2$.

Fig. 10 depicts the convergence pattern of the proposed algorithms in generating Fig. 1 with $P_{ref} = 38.02$ dBm and $N_{RF} = 4$, while Fig. 11 depicts the convergence to zero of the penalty terms. The FC algorithms require more iterations than their AOSA counterparts, because they involve many more decision variables of the phase shifters. Finally, the convergence patterns of objective function (13) in generating Fig. 4, and of objective function (31) in generating Fig. 5, are depicted in Fig. 12 both for $N_{RF} = 4$ and $P_{ref} = 33.95$ dBm, illustrating the efficacy of the proposed algorithms in resolving these two problems. To offer a concise depiction of the convergence behaviors of the MR and SR, we use the mean rate value for SR maximization.

VI. CONCLUSIONS

A communication network relying on a massive antenna-array at the BS was considered, which supported multiple users. For energy-efficient delivery of high bit rates over mmWave and sub-Terahertz frequency bands, hybrid beamforming was harnessed, which relied on the concatenation of an array-of-subarrays structured analog beamformer and a baseband beamformer. The analog beamformer had a low resolution for the sake of a low-complexity practical implementation. To offer a uniform quality-of-service to all users, we maximized the users' minimum rate. Furthermore, we have shown that our new soft max-min rate based design is computationally attractive, since it is based on scalable algorithms, and it succeeds in attaining an attractive minimum rate and sum rate.

APPENDIX: INEQUALITY INGREDIENTS

Define the function

$$\pi(\mathbf{x}, \mathbf{y}) \triangleq \sum_{k=1}^K \left(1 - \frac{\|\mathbf{x}_k\|^2}{\mathbf{y}_k}\right)$$

for $\mathbf{x} \triangleq (\mathbf{x}_1, \dots, \mathbf{x}_K)$ and $\mathbf{y} \triangleq (\mathbf{y}_1, \dots, \mathbf{y}_K)$, with $\mathbf{x}_k \in \mathbb{C}^{N_k}$ and $\mathbf{y}_k \in \mathbb{R}$, $k = 1, \dots, K$ over the domain

$$\{(\mathbf{x}, \mathbf{y}) : \|\mathbf{x}_k\|^2 < \mathbf{y}_k, k = 1, \dots, K\}. \quad (86)$$

Theorem 1: In the domain constrained by (86) the function $\chi(\mathbf{x}, \mathbf{y}) \triangleq \ln \pi(\mathbf{x}, \mathbf{y})$ is concave. Then the following inequality holds true for all (\mathbf{x}, \mathbf{y}) and $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ in the domain constrained by (86):

$$\begin{aligned} \ln \pi(\mathbf{x}, \mathbf{y}) &\leq \ln \pi(\bar{\mathbf{x}}, \bar{\mathbf{y}}) + \frac{1}{\pi(\bar{\mathbf{x}}, \bar{\mathbf{y}})} \sum_{k=1}^K \frac{\|\bar{\mathbf{x}}_k\|^2}{\bar{\mathbf{y}}_k} \\ &+ \frac{1}{\pi(\bar{\mathbf{x}}, \bar{\mathbf{y}})} \sum_{k=1}^K \left(-2 \frac{\Re\{\bar{\mathbf{x}}_k^H \mathbf{x}_k\}}{\bar{\mathbf{y}}_k} + \frac{\|\bar{\mathbf{x}}_k\|^2}{\bar{\mathbf{y}}_k^2} \mathbf{y}_k \right). \end{aligned} \quad (87)$$

Proof: By [48, Appendix], the function $\pi(\mathbf{x}, \mathbf{y})$ is concave. Therefore, $\chi(\mathbf{x}, \mathbf{y})$ is concave as the composition of the concave and monotonically increased function $\ln \pi$ and concave function $\pi(\mathbf{x}, \mathbf{y})$ [33]. Note that the RHS of (87) is the linearized function of the concave function $\ln \pi(\mathbf{x}, \mathbf{y})$ at $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$, so it is a tight majorant of the LHS [33]. \square

For $c > 0$ define

$$\pi_c(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^K \left(1 - \frac{\|\mathbf{x}_k\|^2}{\|\mathbf{x}_k\|^2 + c\mathbf{y}_k}\right).$$

It follows from (87) that the following inequality holds true for all $\mathbf{x}_k \in \mathbb{C}^{N_k}$, $\bar{\mathbf{x}}_k \in \mathbb{C}^{N_k}$, and $\mathbf{y}_k > 0$, $\bar{\mathbf{y}}_k > 0$, $k = 1, \dots, K$:

$$\begin{aligned} \ln \pi_c(\mathbf{x}, \mathbf{y}) &\leq \ln \pi_c(\bar{\mathbf{x}}, \bar{\mathbf{y}}) + \frac{1}{\pi_c(\bar{\mathbf{x}}, \bar{\mathbf{y}})} \sum_{k=1}^K \frac{\|\bar{\mathbf{x}}_k\|^2}{c\bar{\mathbf{y}}_k + \|\bar{\mathbf{x}}_k\|^2} \\ &+ \frac{1}{\pi_c(\bar{\mathbf{x}}, \bar{\mathbf{y}})} \sum_{k=1}^K \left(-2 \frac{\Re\{\bar{\mathbf{x}}_k^H \mathbf{x}_k\}}{c\bar{\mathbf{y}}_k + \|\bar{\mathbf{x}}_k\|^2} \right. \\ &\left. + \frac{\|\bar{\mathbf{x}}_k\|^2}{(c\bar{\mathbf{y}}_k + \|\bar{\mathbf{x}}_k\|^2)^2} (c\mathbf{y}_k + \|\mathbf{x}_k\|^2) \right). \end{aligned} \quad (88)$$

Note that the particular case $K = 1$ of (88) is the following inequality, which was derived in [49]:

$$\begin{aligned} \ln \left(1 + \frac{\|\mathbf{x}\|^2}{\mathbf{y}}\right) &\geq \ln \left(1 + \frac{\|\bar{\mathbf{x}}\|^2}{\bar{\mathbf{y}}}\right) - \frac{\|\bar{\mathbf{x}}\|^2}{\bar{\mathbf{y}}} + 2 \frac{\Re\{\bar{\mathbf{x}}^H \mathbf{x}\}}{\bar{\mathbf{y}}} \\ &- \frac{\|\bar{\mathbf{x}}\|^2}{\bar{\mathbf{y}}(\|\bar{\mathbf{x}}\|^2 + \bar{\mathbf{y}})} (\|\mathbf{x}\|^2 + \mathbf{y}), \end{aligned} \quad (89)$$

for all $\mathbf{x} \in \mathbb{C}^N$, $\mathbf{y} > 0$, and $\bar{\mathbf{x}} \in \mathbb{C}^N$, $\bar{\mathbf{y}} > 0$.

REFERENCES

- [1] S. Ju, Y. Xing, O. Kanhere, and T. S. Rappaport, "Millimeter wave and sub-terahertz spatial statistical channel model for an indoor office building," *IEEE J. Select. Areas Commun.*, vol. 39, no. 6, pp. 1561–1575, Jun. 2021.
- [2] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proc. IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug. 2011.
- [3] J. Federici and L. Moeller, "Review of terahertz and subterahertz wireless communications," *J. Appl. Phys.*, vol. 107, no. 11, p. 6, Jun. 2010.
- [4] K. Okada *et al.*, "Full four-channel 6.3-Gb/s 60-GHz CMOS transceiver with low-power analog and digital baseband circuitry," *IEEE J. Solid State Circuits*, vol. 48, no. 1, pp. 46–64, Jan. 2013.
- [5] I. F. Akyildiz and J. M. Jornet, "Realizing ultra-massive MIMO (1024x1024) communication in the (0.06-10) Terahertz band," *Nano Commun. Networks*, vol. 8, pp. 46–54, June 2016.
- [6] J. M. J. M. Polese, T. Melodia, and M. Zorzi, "Toward end-to-end, full-stack 6G terahertz networks," *IEEE Commun. Mag.*, vol. 58, no. 11, pp. 48–54, Nov. 2020.
- [7] O. E. Ayach, R. W. Heath, S. Rajagopal, and Z. Pi, "Multimode precoding in millimeter wave MIMO transmitters with multiple antenna sub-arrays," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 3476–3480.
- [8] J. Du, W. Xu, H. Shen, X. Dong, and C. Zhao, "Hybrid precoding architecture for massive multiuser MIMO with dissipation: Sub-connected or fully connected structures?" *IEEE Trans. Wirel. Commun.*, vol. 17, no. 8, pp. 5465–5479, Aug. 2018.
- [9] O. Kanhere, H. Poddar, Y. Xing, D. Shakya, S. Ju, and T. S. Rappaport, "A power efficiency metric for comparing energy consumption in future wireless networks in the millimeter-wave and terahertz bands," *IEEE Wirel. Commun.*, vol. 29, no. 6, pp. 56–63, 2022.
- [10] C. Lin and G. Y. Li, "Energy-efficient design of indoor mmWave and Sub-THz systems with antenna arrays," *IEEE Trans. Wirel. Commun.*, vol. 15, no. 7, pp. 4660–4672, Jul. 2016.
- [11] F. Sahrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Select. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Mar. 2016.
- [12] J.-C. Chen, "Hybrid beamforming with discrete phase shifters for millimeter-wave massive MIMO systems," *IEEE Trans. Vehic. Techn.*, vol. 66, no. 8, pp. 7604–7608, Aug. 2017.
- [13] L. Kong, S. Han, and C. Yang, "Hybrid precoding with rate and coverage constraints for wideband massive MIMO systems," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 7, pp. 4634–4647, Jul. 2018.
- [14] Q. Shi and M. Hong, "Spectral efficiency optimization for millimeter wave multiuser MIMO systems," *IEEE J. Select. Topics Signal Process.*, vol. 12, no. 3, pp. 455–468, Jun. 2018.
- [15] T.-H. Tsai, M.-C. Chiu, and C.-C. Chao, "Sub-system SVD hybrid beamforming design for millimeter wave multi-carrier systems," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 1, pp. 518–531, Jan. 2019.
- [16] Z. Ni, J. A. Zhang, K. Yang, F. Gao, and J. An, "Low-complexity subarray-based RF precoding for wideband multiuser millimeter wave systems," *IEEE Trans. Vehic. Techn.*, vol. 69, no. 7, pp. 8028–8033, Jul. 2020.
- [17] A. A. Nasir, H. D. Tuan, T. Q. Duong, H. V. Poor, and L. Hanzo, "Hybrid beamforming for multi-user millimeter-wave networks," *IEEE Trans. Vehic. Techn.*, vol. 69, no. 3, pp. 2943–2956, Mar. 2020.
- [18] Y. Zhang, J. Du, Y. Chen, X. Li, K. M. Rabie, and R. Kharel, "Near-optimal design for hybrid beamforming in mmwave massive multi-user MIMO systems," *IEEE Access*, vol. 8, pp. 129 153–129 168, 2020.
- [19] —, "Dual-iterative hybrid beamforming design for millimeter-wave massive multi-user MIMO systems with sub-connected structure," *IEEE Trans. Veh. Techn.*, vol. 69, no. 11, pp. 13 482–13 496, Nov. 2020.

- [20] W. Hao *et al.*, “Robust design for intelligent reflecting surface-assisted MIMO-OFDMA terahertz IoT networks,” *IEEE Internet of Things J.*, vol. 8, no. 16, pp. 13 052–13 064, 2021.
- [21] C. Fang, B. Makki, J. Li, and T. Svensson, “Hybrid precoding in cooperative millimeter wave networks,” *IEEE Trans. Wirel. Commun.*, vol. 20, no. 8, pp. 5373–5388, Aug. 2021.
- [22] F. Gao, B. Wang, C. Xing, J. An, and G. Y. Li, “Wideband beamforming for hybrid massive MIMO terahertz communications,” *IEEE J. Select. Areas Commun.*, vol. 39, no. 6, pp. 1725–1740, Jun. 2021.
- [23] J.-C. Guo, Q.-Y. Yu, W.-B. Sun, and W.-X. Meng, “Robust efficient hybrid pre-coding scheme for mmwave cell-free and user-centric massive MIMO communications,” *IEEE Trans. Wirel. Commun.*, vol. 20, no. 12, pp. 8006–8022, Dec. 2021.
- [24] C. Qi, Q. Liu, X. Yu, and G. Y. Li, “Hybrid precoding for mixture use of phase shifters and switches in mmWave massive MIMO,” *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4121–4133, Jun. 2022.
- [25] X. Wang, Z. Lin, F. Lin, and L. Hanzo, “Joint hybrid 3D beamforming relying on sensor-based training for reconfigurable intelligent surface aided terahertz-based multiuser massive MIMO systems,” *IEEE Sensors J.*, vol. 22, no. 14, pp. 14 540–14 552, 2022.
- [26] Y. Liu, Y. Cai, A. Liu, M. Zhao, and L. Hanzo, “Latency minimization for mmwave D2D mobile edge computing systems: Joint task allocation and hybrid beamforming design,” *IEEE Trans. Vehic. Techn.*, vol. 71, no. 11, pp. 12 206–12 221, Nov. 2022.
- [27] X. Cui and Q. Li, “Hybrid beamforming with finite-resolution phase shifters for multiuser millimeter-wave downlink,” *IEEE Wirel. Commun. Lett.*, vol. 9, no. 2, pp. 219–222, Feb. 2020.
- [28] L. You *et al.*, “Hybrid analog/digital precoding for downlink massive MIMO LEO satellite communications,” *IEEE Trans. Wirel. Commun.*, vol. 21, no. 8, pp. 5962–5976, Aug. 2022.
- [29] A. H. Phan, H. D. Tuan, H. H. Kha, and H. H. Nguyen, “Beamforming optimization in multi-user amplify-and-forward wireless relay networks,” *IEEE Trans. Wirel. Commun.*, vol. 11, pp. 1510–1520, Apr. 2012.
- [30] H. Yu, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, “Low-resolution hybrid beamforming in millimeter-wave multi-user systems,” *IEEE Trans. Vehic. Techn.*, vol. 72, no. 7, pp. 8941–8955, Jul. 2023.
- [31] K. Wu, J. A. Zhang, X. Huang, Y. J. Guo, and L. Hanzo, “Simultaneous beam and user selection for the beamspace mmWave/THz massive MIMO downlink,” *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1785–1797, Mar. 2023.
- [32] V.-D. Nguyen, H. D. Tuan, T. Q. Duong, H. V. Poor, and O.-S. Shin, “Precoder design for signal superposition in MIMO-NOMA multicell networks,” *IEEE J. Select. Areas Commun.*, vol. 35, no. 12, pp. 2681–2695, Dec. 2017.
- [33] H. Tuy, *Convex Analysis and Global Optimization (second edition)*. Springer International, 2017.
- [34] J. D. Krieger, C.-P. Yeang, and G. W. Wornell, “Dense delta-sigma phased arrays,” *IEEE Trans. Antenn. Propag.*, vol. 61, no. 4, pp. 1825–1837, Apr. 2013.
- [35] N. Deferm and P. Reynaert, *CMOS Front Ends for Millimeter Wave Wireless Communication Systems*. Springer New York, NY, USA, 2015.
- [36] W. Shin, O. Inac, Y.-C. Ou, B. Ku, and G. M. Rebeiz, “A 108-114 GHz 4x4 wafer-scale phased array transmitter with high-efficiency on-chip antennas,” *IEEE J. Solid-State Circuits*, vol. 48, no. 9, pp. 2041–2055, May 2013.
- [37] Z. Ni, J. A. Zhang, K. Yang, F. Gao, and J. An, “Estimation of multiple angle-of-arrivals with localized hybrid subarrays for millimeter wave systems,” *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1897–1910, Mar. 2020.
- [38] J. F. Bonnans, J. C. Gilbert, C. Lemarechal, and C. Sagastigabal, *Numerical Optimization-Theoretical and Practical Aspects (second edition)*. Springer, 2006.
- [39] “3GPP technical specification group radio access network evolved universal terrestrial radio access (E-UTRA): Further advancements for E-UTRA physical layer aspects (release 9),” 2010.
- [40] 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. Select. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.
- [41] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, “Overview of millimeter wave communications for fifth-generation (5G) wireless networks with a focus on propagation models,” *IEEE Trans. Antenn. Propag.*, vol. 65, no. 12, pp. 6213–6230, 2017.
- [42] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, “Spatially sparse precoding in millimeter wave MIMO systems,” *IEEE Tran. Wirel. Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [43] H. Yu, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, “Regularized zero-forcing aided hybrid beamforming for millimeter-wave multiuser MIMO systems,” *IEEE Trans. Wirel. Commun.*, vol. 22, no. 5, pp. 3280–3295, May 2023.
- [44] S. Park, J. Park, A. Yazdan, and R. W. Heath, “Exploiting spatial channel covariance for hybrid precoding in massive MIMO systems,” *IEEE Trans. Signal Process.*, vol. 65, no. 14, pp. 3818–3832, Jul. 2017.
- [45] J. Noh, T. Kim, J. Seol, and C. Lee, “Zero-forcing based hybrid beamforming for multi-user millimeter wave systems,” *IET Communications*, vol. 10, no. 18, pp. 2670–2677, 2016.
- [46] W. Hao, M. Zeng, Z. Chu, and S. Yang, “Energy-efficient power allocation in millimeter wave massive MIMO with non-orthogonal multiple access,” *IEEE Wirel. Commun. Lett.*, vol. 6, no. 6, pp. 782–785, Jun. 2017.
- [47] R. Jain, D.-M. Chiu, and W. R. Hawe, “A quantitative measure of fairness and discrimination for resource allocation in shared computer systems,” *Digital Equipment, Tech. Rep. DEC-TR-301*, Sept. 1984.
- [48] U. Rashid, H. D. Tuan, H. H. Kha, and H. H. Nguyen, “Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks,” *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 488–499, Feb. 2014.
- [49] H. H. M. Tam, H. D. Tuan, and D. T. Ngo, “Successive convex quadratic programming for quality-of-service management in full-duplex MU-MIMO multicell networks,” *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2340–2353, June 2016.