Low-Resolution Hybrid Beamforming in Millimeter-wave Multi-user Systems

H. Yu[1,2], H. D. Tuan[2], E. Dutkiewicz[2], H. V. Poor[3], and L. Hanzo[4]

Abstract—**A new hybrid amalgam of analog and baseband digital beamforming is conceived for millimeter-wave (mmWave) multi-user networks. The base station is equipped with a large-scale antenna-array, which however relies on a limited number of radio frequency chains for mitigating the severe path-loss of mmWave band transmission. Each user is also equipped with a multi-antenna array. The analog beamformer's weight-resolution is set low for the sake of power efficient implementation. The hybrid beamformer design relies on the novel optimization criterion of directly maximizing the geometric means of users' rates, which is shown to result in fair rate-distributions for the users without imposing a minimum user-rate constraint. Furthermore, new computationally efficient algorithms are developed, which are purely based on closed-form low-complexity expressions. Hence these low-complexity solutions are eminently suitable for large-scale mmWave arrays. Numerical examples are provided for demonstrating their efficiency.**

Index Terms—**Millimeter-wave communications, hybrid beamforming, analog beamforming, baseband beamforming, nonlinear discrete optimization, nonconvex optimization algorithms**

I. INTRODUCTION

Since the treatises [1], [2], the millimeter-wave (mmWave) frequency band spanning from 30 GHz to 300 GHz has received considerable interest from both academia and industry. Hence numerous special issues have been devoted to its coevolutionary development [3]–[6]. Given that its physical channel modeling and compact antenna array configurations are well-understood [3], [7], [8], one of the main issues becomes the development of signal processing techniques for approaching its capacity potential [4], [9]. The hybrid combination of analog beamforming (ABF) and digital beamforming (DBF) has been accepted as the most power-efficient signal processing technique for mmWave communications [4], [6], [9]–[11]. However, the design of hybrid beamforming (HBF) remains challenging owing to the following factors: (i) the HBF matrix is a product of the ABF and DBF matrices, so both the transmit power constraint and its rate functions are complex and hard

[1]School of Communication and Information Engineering, Shanghai University Shanghai 200444, China (email: hw_yu@shu.edu.cn); [2]School of Electrical and Data Engineering, University of Technology Sydney, Broadway, NSW 2007, Australia (email:Hongwen.Yu@uts.edu.au, Tuan.Hoang@uts.edu.au, Eryk.Dutkiewicz@uts.edu.au); [3]Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA (email: poor@princeton.edu); [4]School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK (email: lh@ecs.soton.ac.uk)

to formally optimize; (ii) The entries of the ABF matrix are subject to the unit modulus constraint. As for single-user mmWave systems, most contributions aim for designing a HBF matrix to approximate fully digital beamforming matrix by alternately optimizing the ABF and DBF matrices [12]–[14]. It should be noted however that the alternating optimization in the ABF matrix still remains challenging due to the unit modulus constraints on its entries [15].

The challenges in HBF design are further aggravated in multi-user mmWave systems. To ease the computation at complexity, the design has typically been based on the users' sum rate (SR) maximization [16], [17], which however allocates a large fraction of the total SR to a few users having good channel conditions, while leaving the rest of the users with almost zero rates. For alternating optimization in the ABF matrix, both [16], [17] simply optimized each of its entries with all other entries held fixed. Our recent work [18] has shown that the HBF design based on a simple ABF relying on the closed-form expression proposed in [19] outperforms that HBF design of [16] or [20] by a high margin. Max-min rate optimization, which aims for maximizing the users' minimum rate, is more appropriate for multi-user communications [21]–[24]. In principle, the joint design of ABF and DBF for max-min rate optimization may be tackled by approximately adjusting the exact-penalty algorithms developed in [25], which invoke convex problems at each iteration for generating better points. However, as the ABF and DBF matrices tend to be high-dimensional, these convex problems are of large-scale and thus are still computationally intractable. Making things worse, each ABF weight/entry of the ABF matrix has to be from a discrete set of its finite-resolution counterpart for practical implementation in the mmWave frequency range [26]. Since such alternating optimization of the ABF matrix involves extremely high-dimensional combinatoric problems, [16] and [17] tackled such problem by checking the value of the objective function used at all 2^b possible points of the b-resolution set for each entry with all other entries held fixed. Needless to say, both nonlinear manifold optimization based approach [15] and exactly-penalized optimization based approach [25] are not applicable to such mixed discrete continuous optimization problems.

Against the above background, this paper develops a new systematic design approach for HBF in multi-user mmWave communications. Explicitly, the contributions of the paper are as follows:

- We improve the spectral efficiency of multi-user HBF by maximizing the novel objective function constituted by the geometric mean of the users' rates (GM-rate). Our

GM-rate based optimization has substantial advantages over its conventional SR-based counterpart, because the latter results in assigning excessive rates to the privileged users having high channel quality, while only granting meagre rates for the rest. This implement may be circumvented by setting certain minimum user-rate constraints, but at the unaffordable cost of exacerbating the computational complexity for large-scale antenna-arrays. By contrast, our GM-rate based regime tends to distribute the rates to all users in a fair and equitable manner without imposing any minimum user-rate constraints, while the overall SR is still high. The rate-fairness is quantified in terms of the users' rate deviation (RD) from their mean and by the rate ratio (RR) of the users' maximal and minimal rates.

- We develop low-complexity optimization algorithms for the joint design of ABF and DBF to maximize the GM-rate, which iterate by evaluating low-complexity closed-form expressions and thus are eminently suitable for massive antenna-arrays. The key point is to exploit the unit modulus constraints to develop tight concave minorants [27] for the objective functions. Optimization of the former results in better points for the latter. Most importantly, it admits closed-form solutions of scalable complexity, regardless whether the ABF has infinite or finite resolution. As such, better ABF is generated by closed-form expressions of scalable complexity, which is achieved upon bypassing any exhaustive search over the entire discrete set of solutions. This result is novel even from an optimization perspective [27].

To sum up, we boldly and explicitly contrast our novel contributions to the literature in Table I.

The paper is organized as follows. Section II and Section III respectively develop approximation optimization (AO) and penalty optimization (PO) algorithms for the HBF design to maximize the GM-rate, which iterate by purely relying on closed-form expressions. Section IV evaluates the performance of all the algorithms. By comparing them to the existing algorithms in terms of the achievable SR. Section VI concludes the paper.

Notation. Only the optimization variables are boldfaced. For a complex number x, $\angle x$ is its argument, so its trigonometric representation is $x = |x|e^{j\angle x}$. As such we have $x = e^{j\angle x}$ for $|x| = 1$, i.e. a unit modulus complex number is completely characterized by its argument. The inner product between vectors x and y is defined as $\langle x, y \rangle = x^H y$. Analogously, $\langle X, Y \rangle = \text{trace}(X^H Y)$ for the matrices X and Y. We also use $\langle X \rangle$ for the trace of X when X is a square matrix. $X \succeq 0$ ($X \succ 0$, resp.) means that X is Hermitian symmetric and positive semi-definite (positive definite, resp.). Accordingly, $X \succeq Y$ ($X \succ Y$, resp.) means that $X - Y \succeq 0$ ($X - Y \succ 0$, resp.). $||X||$ is the Frobenius norm of the matrix X, which is defined by $\sqrt{\langle X^H X \rangle}$. $[X]^2$ stands for $XX^H \succeq 0$, so $||X||^2 = \langle [X]^2 \rangle$. For Hermitian symmetric matrix X, denote by $\lambda_{\max}(X)$ its largest eigenvalue. Thus, it is true that $\lambda_{\max}(X)I \succeq X$, so $\langle \lambda_{\max}(X)I - X, Y \rangle = \langle (\lambda_{\max}(X)I - X)Y \rangle \geq 0$ for all $Y \succeq 0$, where I is the identity matrix of the same size as X. However, when the size of the identity matrix is not clear

in expressions, we use the notation I_N to emphasize its size of $N \times N$. Lastly, $\mathbb{R}_+^N$ is the set of N-dimensional real vectors with positive entries.

Ingredient. According to [27, p. 366], a function $\bar{f}$ is said to be a tight minorant of a function f over the domain $\text{dom}(f)$ at a point $z^{(\kappa)} \in \text{dom}(f)$ if it satisfies the conditions of global bounding

$$f(z) \geq \bar{f}(z) \ \forall \ z \in \text{dom}(f), \tag{1}$$

and matching at $z^{(\kappa)}$:

$$f(z^{(\kappa)}) = \bar{f}(z^{(\kappa)}). \tag{2}$$

Lemma 1: The optimal solution $z^{(\kappa+1)}$ of the tight minorant maximization problem

$$\max_{z \in \text{dom}(f)} \bar{f}(z) \tag{3}$$

provides a better feasible point than $z^{(\kappa)}$ upon maximizing f over $\text{dom}(f)$ as far as $\bar{f}(z^{(\kappa+1)}) \neq \bar{f}(z^{(\kappa)})$, i.e.

$$f(z^{(\kappa+1)}) > f(z^{(\kappa)}). \tag{4}$$

Proof. In fact, according to (1) we have $f(z^{(\kappa+1)}) \geq \bar{f}(z^{(\kappa+1)})$ and then $\bar{f}(z^{(\kappa+1)}) > \bar{f}(z^{(\kappa)})$, because $z^{(\kappa+1)}$ and $z^{(\kappa)}$ are the optimal solution and a feasible point of the the problem (3), so $f(z^{(\kappa+1)}) > \bar{f}(z^{(\kappa)})$ which is (4) because $f(z^{(\kappa)}) = \bar{f}(z^{(\kappa)})$ by (2).

The following matrix inequality [21] for all $\mathbf{V}$, $\bar{V}$, and positive definite $\mathbf{Y}$ and $\bar{Y}$ of appropriate dimension is frequently used in the paper

$$
\begin{aligned}
\ln|I + [\mathbf{V}]^2 \mathbf{Y}^{-1}| \geq \ &\ln|I + [\bar{V}]^2 \bar{Y}^{-1}| - \langle [\bar{V}]^2 \bar{Y}^{-1} \rangle \\
&+ 2\Re\{\langle \bar{V}^H \bar{Y}^{-1} \mathbf{V} \rangle\} - \langle \bar{Y}^{-1} \\
&- \left([\bar{V}]^2 + \bar{Y}\right)^{-1}, [\mathbf{V}]^2 + \mathbf{Y} \rangle.
\end{aligned}
\tag{5}
$$

One can check that the right hand side (RHS) of (5) matches with its left hand side (LHS) at $(\bar{V}, \bar{Y})$ so the former provides a tight minorant of the latter at $(\bar{V}, \bar{Y})$.

II. APPROXIMATION OPTIMIZATION BASED HBF DESIGN

Consider the mmWave downlink (DL) of a base station (BS) serving N_u users indexed by $i \in \mathcal{N}_u \triangleq \{1, 2, \ldots, N_u\}$.[1] The BS is equipped with a massive N-antenna array, while each user (UE) i is equipped with an N_R-antenna array. For $\mathcal{N} \triangleq \{1, \ldots, N\}$ and $\mathcal{N}_{RF} \triangleq \{1, \ldots, N_{RF}\}$, where N_{RF} is the number of radio frequency (RF) chains that the BS uses for HBF, let $\boldsymbol{\theta} \triangleq [\boldsymbol{\theta}_{n,j}]_{(n,j) \in \mathcal{N} \times \mathcal{N}_{RF}} \in [0, 2\pi)^{N \times N_{RF}}$ be the phase shift matrix. Define the following one-to-one mapping from $[0, 2\pi)^{N \times N_{RF}}$ to $\mathbb{U}^{N \times N_{RF}} \triangleq \{\mathbf{U} = [\mathbf{u}(n,j)]_{(n,j) \in \mathcal{N} \times \mathcal{N}_{RF}} : |\mathbf{u}_{n,j}| = 1, (n,j) \in \mathcal{N} \times \mathcal{N}_{RF}\}$:

$$
\begin{aligned}
V_{RF}(\boldsymbol{\theta}) &\triangleq \begin{bmatrix} V_{RF}^1(\boldsymbol{\theta}) & \ldots & V_{RF}^{N_{RF}}(\boldsymbol{\theta}) \end{bmatrix} \\
&= [e^{j\boldsymbol{\theta}_{n,j}}]_{(n,j) \in \mathcal{N} \times \mathcal{N}_{RF}} \in \mathbb{U}^{N \times N_{RF}}
\end{aligned}
$$

---

[1] Each RF chain serves a group of users will be considered in our future research [28].

TABLE I: Contrasting our novel contributions to the related literature.

Literature / Contents	This work	[16]	[17]	[18]	[15]	[25]	[21]–[24]
SR maximization	√	√	√			√	
max-min rate optimization					√	√	√
GM-rate maximization	√						
zero rate issue		√	√		√		
computational tractability in ABF	√						
computational efficiency for massive antenna-arrays	√						
combinatoric optimization	√						
multi-antenna users	√		√				√

to represent the phase shift based ABF matrix. In this paper, we always set $N_{RF} \leq N_u$. Since ABF relying on infinite resolution weights is not practical for mmWave communication [26], we opt for b-bit resolution, i.e.

$$\boldsymbol{\theta}_{n,j} \in \mathcal{B} \triangleq \{\nu \frac{2\pi}{2^b}, \nu = 0, 1, \ldots, 2^b - 1\}, (n,j) \in \mathcal{N} \times \mathcal{N}_{RF}. \quad (6)$$

In what follows, the projection of $\alpha \in [0, 2\pi)$ into $\mathcal{B}$ denoted by $\lfloor \alpha \rfloor_b$ is referred to as its b-bit rounded version:

$$\lfloor \alpha \rfloor_b = \nu_\alpha \frac{2\pi}{2^b} \quad (7)$$

with

$$\nu_\alpha \triangleq \arg \min_{\nu=0,1,\ldots,2^b-1} \left| \nu \frac{2\pi}{2^b} - \alpha \right|, \quad (8)$$

which can be readily found, because we have $\nu_\alpha \in \{\nu, \nu+1\}$ for $\alpha \in [\nu \frac{2\pi}{2^b}, (\nu+1)\frac{2\pi}{2^b}]$. When $b = \infty$, it is true that

$$\alpha = \lfloor \alpha \rfloor_\infty. \quad (9)$$

Upon denoting the mmWave channel between the BS and UE $i \in \mathcal{N}_u$ by $H_i \in \mathbb{C}^{N_R \times N}$, the signal received at UE i is formulated as:

$$y_i = H_i V_{RF}(\boldsymbol{\theta}) x + n_i, \quad (10)$$

where $x \in \mathbb{C}^{N_{RF}}$ is the baseband signal, and n_i is the background noise of power σ.

Let $s_i \in \mathbb{C}^{N_R}$ with $\mathbb{E}(s_i s_i^H) = I_{N_R}$ be the information intended for UE i, which is "beamformed" by $\mathbf{V}_i^B \in \mathbb{C}^{N_{RF} \times N_R}$ before the BS's transmission. For $\mathbf{s} \triangleq (s_1, \ldots, s_{N_u})^T$ and

$$\mathbf{V}^B = \begin{bmatrix} \mathbf{V}_1^B & \cdots & \mathbf{V}_{N_u}^B \end{bmatrix} \in \mathbb{C}^{N_{RF} \times (N_u N_R)}, \quad (11)$$

which is the DBF matrix, the baseband signal x in (10) is $x = \mathbf{V}^B \mathbf{s} = \sum_{\ell=1}^{N_u} \mathbf{V}_\ell^B s_\ell$, i.e. (10) is in fact the following multi-input multi-output (MIMO) equation

$$\begin{aligned} y_i &= H_i V_{RF}(\boldsymbol{\theta}) \sum_{\ell=1}^{N_u} \mathbf{V}_\ell^B s_\ell + n_i & (12) \\ &= \mathcal{H}_i(\boldsymbol{\theta}) \sum_{\ell=1}^{N_u} \mathbf{V}_\ell^B s_\ell + n_i, & (13) \end{aligned}$$

for

$$\mathcal{H}_i(\boldsymbol{\theta}) \triangleq H_i V_{RF}(\boldsymbol{\theta}) \in \mathbb{C}^{N_R \times N_{RF}}, i = 1, \ldots, N_u \quad (14)$$

The product of the ABF matrix $V_{RF}(\boldsymbol{\theta})$ and beamformer $\mathbf{V}_i^B$, $i \in \mathcal{N}_u$ is given by

$$\mathbf{V}_i^{HD} \triangleq V_{RF}(\boldsymbol{\theta}) \mathbf{V}_i^B \in \mathbb{C}^{N \times N_R}, \quad (15)$$

and accordingly,

$$\mathbf{V}^{HD} \triangleq \begin{bmatrix} \mathbf{V}_1^{HD} & \cdots & \mathbf{V}_{N_u}^{HD} \end{bmatrix} \in \mathbb{C}^{N \times (N_u N_R)}, \quad (16)$$

is referred to as the HBF matrix, for distinguishing it from the fully digital beamforming matrix of

$$\mathbf{V}^{FD} \in \mathbb{C}^{N \times (N_u N_R)}, \quad (17)$$

which does not rely on the matrix product structure (15).

It follows from the equations (13) that the achievable rate of UE i' is

$$\rho_i(\boldsymbol{\theta}, \mathbf{V}^B) \triangleq \ln \left| I_{N_R} + [\mathcal{H}_i(\boldsymbol{\theta}) \mathbf{V}_i^B]^2 \Psi_i^{-1}(\boldsymbol{\theta}, \mathbf{V}^B) \right|, \quad (18)$$

with

$$\Psi_i(\boldsymbol{\theta}, \mathbf{V}^B) \triangleq \sum_{\ell \neq i}^{N_u} [H_i V_{RF}(\boldsymbol{\theta}) \mathbf{V}_\ell^B]^2 + \sigma I_{N_R}. \quad (19)$$

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

$$\sum_{i=1}^{N_u} ||V_{RF}(\boldsymbol{\theta}) \mathbf{V}_i^B||^2 = \sum_{i=1}^{N_u} \langle V_{RF}^H(\boldsymbol{\theta}) V_{RF}(\boldsymbol{\theta}), [\mathbf{V}_i^B]^2 \rangle \leq P. \quad (20)$$

As mentioned in the Introduction, the authors of [15]–[17] considered the SR maximization problem

$$\max_{\boldsymbol{\theta}, \mathbf{V}^B} \sum_{i=1}^{N_u} \rho_i(\boldsymbol{\theta}, \mathbf{V}^B) \quad \text{s.t.} \quad (6), (20). \quad (21)$$

Here, particularly, [17] addressed (21) via the following PO:

$$\max_{\mathbf{U}, \mathbf{V}^B, \mathbf{V}^{HD}} \sum_{i=1}^{N_u} \hat{r}_i(\mathbf{V}^{HD}) - c ||\mathbf{V}^{HD} - \mathbf{U} \mathbf{V}^B||^2 \quad (22a)$$

$$\text{s.t.} \quad \mathbf{u}(n,j) \in \{e^{j2\pi\nu/2^b}, \nu = 0, 1, \ldots, 2^b - 1\},$$
$$(n,j) \in \mathcal{N} \times \mathcal{N}_{RF}, \quad (22b)$$
$$||\mathbf{V}^{HD}||^2 \leq P, \quad (22c)$$

with the penalty parameter $c > 0$ to be updated, where by definition we have

$$\hat{r}_i(\mathbf{V}^{HD}) \triangleq \ln \left| I_{N_R} + [H_i \mathbf{V}_i^{HD}]^2 \widetilde{\Psi}_i^{-1}(\mathbf{V}^{HD}) \right|, \quad (23)$$

along with

$$\widetilde{\Psi}_i(\mathbf{V}^{HD}) \triangleq \sum_{\ell \neq i}^{N_u} [H_i \mathbf{V}_\ell^{HD}]^2 + \sigma I_{N_R}. \quad (24)$$

The constraint (22b) in $\mathbf{U}$ is equivalent to the constraint (6) in $\boldsymbol{\theta}$, and the constraint (22c) in $\mathbf{V}^{HD}$ is equivalent to the

constraint (20) under the equality constraint $\mathbf{V}^{HD} = \mathbf{U}\mathbf{V}^B$, which is enforced by the penalty term in the objective function of (22a). The algorithm in [17] is based on alternating optimization between each of the three sets of variables $\mathbf{V}^{HD}$, $\mathbf{U}$, and $\mathbf{V}^B$ with the other two held fixed. As such, alternating optimization in $\mathbf{U}$ with $(\mathbf{V}^{HD}, \mathbf{V}^B)$ held fixed at $(\bar{V}^{HD}, \bar{V}^B)$ is provided by the problem

$$\min_{\mathbf{U}} ||\bar{V}^{HD} - \mathbf{U}\bar{V}^B||^2 \quad \text{s.t.} \quad (22b), \tag{25}$$

which was addressed by optimizing each single variable $\mathbf{u}(n, j)$ with all $NN_{RF}-1$ other variables $\mathbf{u}(n', j')$, $(n', j') \neq (n, j)$ held fixed, to avoid the combinatoric complexities calculation. For $b < \infty$, like [16] it simply checks at all 2^b possible points $e^{j2\pi\nu/2^b}$, $\nu = 0, 1, \ldots, 2^b-1$. The simulations in [17] showed its superior performance over that achieved by [15], [16].

The well-known drawback of the SR maximization problem (21) is that it results in assigning excessive rates to the privileged users having high channel quality, while only granting meagre rates for the rest. As such it cannot serve as a metric of multi-user spectral efficiency. A more appropriate optimization problem for multi-user spectral efficiency is the following max-min rate optimization problem

$$\max_{\boldsymbol{\theta}, \mathbf{V}^B} \min_{i=1,\ldots,N_u} \rho_i(\boldsymbol{\theta}, \mathbf{V}^B) \quad \text{s.t.} \quad (6), (20), \tag{26}$$

with unknown computational solution. For $b = \infty$, it may be computed by adjusting the exact-penalty algorithms developed in [25], which invoke convex problems at each iteration to generate better points. The computational complexity of these problems is on the order of $\mathcal{O}[(NN_{RF})^6]$ as it involves $\mathcal{O}(NN_{RF})$ decision variables and $\mathcal{O}(NN_{RF})$ convex constraints, which is still intractable when N is very large.

We now propose to exploit the following HBF design problem of maximizing the GM of user-rates (GM-rate) for multi-user spectral efficiency:

$$\max_{\boldsymbol{\theta}, \mathbf{V}^B} \left(\prod_{i=1}^{N_u} \rho_i(\boldsymbol{\theta}, \mathbf{V}^B) \right)^{1/N_u} \quad \text{s.t.} \quad (6), (20). \tag{27}$$

Our recent works [29]–[31] show that the GM-rate optimization results in fair distributions of user-rates while keeping their sum high without setting the threshold on the user rates that causes very difficult nonconvex constraints. Such attractive accomplishment will be shown by our simulation in Section IV.

The remainder of this section is devoted to an AO approach for computing (27).

A. Approximation optimization

The power constraint (20) is computationally intractable because the ABF matrix $V_{RF}(\boldsymbol{\theta})$ and the DBF matrix $\mathbf{V}^B$ are coupled. Our first contribution is that of decoupling them as follows. Using the law of large numbers, it was shown in [12] that for large N, we have

$$V_{RF}^H(\boldsymbol{\theta})V_{RF}(\boldsymbol{\theta}) \approx NI_{N_{RF}}, \tag{28}$$

so

$$\sum_{i=1}^{N_u} \langle V_{RF}^H(\boldsymbol{\theta})V_{RF}(\boldsymbol{\theta}), [\mathbf{V}_i^B]^2 \rangle \approx \sum_{i=1}^{N_u} \langle NI_{N_{RF}}, [\mathbf{V}_i^B]^2 \rangle$$
$$= N||\mathbf{V}^B||^2. \tag{29}$$

Therefore the power constraint (20) is approximated by the following convex constraint

$$||\mathbf{V}^B||^2 \leq P/N, \tag{30}$$

which is independent of $V_{RF}(\boldsymbol{\theta})$. For $\rho(\boldsymbol{\theta}, \mathbf{V}^B) \triangleq (\rho_1(\boldsymbol{\theta}, \mathbf{V}^B), \ldots, \rho_{N_u}(\boldsymbol{\theta}, \mathbf{V}^B))^T \in \mathbb{R}^{N_u}$, which is a nonlinear mapping of $(\boldsymbol{\theta}, \mathbf{V}^B)$, we thus consider the following problem of AO associated with (27):

$$\max_{\boldsymbol{\theta}, \mathbf{V}^B} F(\rho(\boldsymbol{\theta}, \mathbf{V}^B)) \triangleq \left(\prod_{i=1}^{N_u} \rho_i(\boldsymbol{\theta}, \mathbf{V}^B) \right)^{1/N_u} \quad \text{s.t.} \quad (6), (30). \tag{31}$$

Initialized by a $(V^{B,(0)}, \theta^{(0)})$ feasible for (31), for $\kappa = 1, \ldots$, let $(\theta^{(\kappa)}, V^{B,(\kappa)})$ be a feasible point for (31) that is found from the $(\kappa - 1)$-st round. To resolve the high nonlinearity of the objective function in (31), which is a composition of the GM function $F(\rho) \triangleq \left(\prod_{i=1}^{N_u} \rho_i \right)^{1/N_u}$ and the nonlinear mapping $\rho(\boldsymbol{\theta}, \mathbf{V}^B)$, we recast (31) in the following maximin problem[2]

$$\max_{\boldsymbol{\theta}, \mathbf{V}^B} \min_{\boldsymbol{\gamma} \in \mathbb{R}^{N_u}+, \prod_{i=1}^{N_u} \gamma_i = 1} \langle \boldsymbol{\gamma}, \rho(\boldsymbol{\theta}, \mathbf{V}^B) \rangle \quad \text{s.t.} \quad (6), (30). \tag{32}$$

Initialized by a $(V^{B,(0)}, \theta^{(0)})$ feasible for (31), for $\kappa = 0, 1, \ldots$, we optimize in $\boldsymbol{\gamma}$ to have

$$\gamma_i^{(\kappa)} \triangleq \frac{\max_{i' \in \mathcal{N}_u} \rho_{i'}(\theta^{(\kappa)}, V^{B,(\kappa)})}{\rho_i(\theta^{(\kappa)}, V^{B,(\kappa)})}, i \in \mathcal{N}_u. \tag{33}$$

For $\gamma^{(\kappa)} \triangleq (\gamma_1^{(\kappa)}, \ldots, \gamma_{N_U}^{(\kappa)})^T$, we iterate $(V^{B,(\kappa+1)}, \theta^{(\kappa+1)})$ by solving the following problem of mixed discrete continuous optimization problem:

$$\max_{\boldsymbol{\theta}, \mathbf{V}^B} F^{(\kappa)}(\boldsymbol{\theta}, \mathbf{V}^B) \triangleq \langle \gamma^{(\kappa)}, \rho(\boldsymbol{\theta}, \mathbf{V}^B) \rangle \quad \text{s.t.} \quad (6), (30). \tag{34}$$

As discussed in [29]–[32], the problem (34) is the same as the following problem

$$\max_{\boldsymbol{\theta}, \mathbf{V}^B} \mathcal{L}^{(\kappa)}(\rho(\boldsymbol{\theta}, \mathbf{V}^B)) \quad \text{s.t.} \quad (6), (30), \tag{35}$$

where $\mathcal{L}^{(\kappa)}(\rho(\boldsymbol{\theta}, \mathbf{V}^B)$ is the linearized function of $F(\rho)$ at $\rho(\theta^{(\kappa)}, V^{B,(\kappa)})$:

$$\mathcal{L}^{(\kappa)}(\rho(\boldsymbol{\theta}, \mathbf{V}^B) = \frac{F(\rho(\theta^{(\kappa)}, V^{B,(\kappa)}))}{N_u} \sum_{i=1}^{N_u} \frac{\rho_i(\boldsymbol{\theta}, \mathbf{V}^B)}{\rho_i(\theta^{(\kappa)}, V^{B,(\kappa)})}. \tag{36}$$

One should not confuse $\mathcal{L}^{(\kappa)}(\rho(\boldsymbol{\theta}, \mathbf{V}^B))$, which is still a nonlinear function in $(\boldsymbol{\theta}, \mathbf{V}^B)$, with the linear approximation of the function $\tilde{F}(\boldsymbol{\theta}, \mathbf{V}^B) \triangleq F(\rho(\boldsymbol{\theta}, \mathbf{V}^B))$ at $(\theta^{(\kappa)}, V^{B,(\kappa)})$, which is used for the standard gradient ascent algorithms [33]. In fact, $\mathcal{L}^{(\kappa)}(\rho(\boldsymbol{\theta}, \mathbf{V}^B))$ provides a nonlinear approximation of $\tilde{F}(\boldsymbol{\theta}, \mathbf{V}^B)$ at $(\theta^{(\kappa)}, V^{B,(\kappa)})$.

[2] By Cauchy's inequality, we have $\frac{1}{N_u} \langle \boldsymbol{\gamma}, \rho(\boldsymbol{\theta}, \mathbf{V}^B) \rangle \geq [\prod_{i=1}^{N_u} \gamma_i \rho_i(\boldsymbol{\theta}, \mathbf{V}^B)]^{1/N_u} = [\prod_{i=1}^{N_u} \rho_i(\boldsymbol{\theta}, \mathbf{V}^B)]^{1/N_u}$ with the equality sign at $\gamma_1 \rho_1(\boldsymbol{\theta}, \mathbf{V}^B) = \cdots = \gamma_{N_u} \rho_{N_u}(\boldsymbol{\theta}, \mathbf{V}^B)$

1) Alternating ascent in the DBF matrix: We seek DBF $V^{B,(\kappa+1)}$ ensuring that

$$F^{(\kappa)}(\theta^{(\kappa)}, V^{B,(\kappa+1)}) > F^{(\kappa)}(\theta^{(\kappa)}, V^{B,(\kappa)}), \quad (37)$$

by considering the following problem:

$$\max_{\mathbf{V}^B} F^{(\kappa)}(\theta^{(\kappa)}, \mathbf{V}^B) \quad \text{s.t.} \quad (30). \quad (38)$$

which is nonconvex because its objective function is nonconcave.

By using the inequality (5), we obtain the following tight concave quadratic minorant of $\rho_i(\theta^{(\kappa)}, \mathbf{V}^B)$ at $V^{B,(\kappa)}$:

$$\rho_i^{(\kappa)}(\mathbf{V}^B) \triangleq a_i^{(\kappa)} + 2\Re\{\langle A_i^{(\kappa)} \mathbf{V}_i^B \rangle\}$$
$$- \langle \mathcal{H}_i^H(\theta^{(\kappa)}) B_i^{(\kappa)} \mathcal{H}_i(\theta^{(\kappa)}), \sum_{\ell=1}^{N_u} [\mathbf{V}_\ell^B]^2 \rangle (39)$$

with

$$a_i^{(\kappa)} \triangleq \rho_i(\theta^{(\kappa)}, V^{B,(\kappa)})$$
$$- \langle [\mathcal{H}_i(\theta^{(\kappa)}) V_i^{B,(\kappa)}]^2 \Psi_i^{-1}(\theta^{(\kappa)}, V^{B,(\kappa)}) \rangle - \sigma \langle B_i^{(\kappa)} \rangle,$$
$$A_i^{(\kappa)} \triangleq [\mathcal{H}_i(\theta^{(\kappa)}) V_i^{B,(\kappa)}]^H \Psi_i^{-1}(\theta^{(\kappa)}, V^{B,(\kappa)}) \mathcal{H}_i(\theta^{(\kappa)}),$$

and

$$B_i^{(\kappa)} \triangleq \Psi_i^{-1}(\theta^{(\kappa)}, V^{B,(\kappa)}) - \left(\Psi_i(\theta^{(\kappa)}, V^{B,(\kappa)}) + [\mathcal{H}_i(\theta^{(\kappa)}) V_i^{B,(\kappa)}]^2\right)^{-1}.$$

Thus, the following concave quadratic function provides a tight minorant of the objective function in (38):

$$\tilde{F}^{(\kappa)}(\mathbf{V}^B) \triangleq \sum_{i=1}^{N_u} \gamma_i^{(\kappa)} \rho_i^{(\kappa)}(\mathbf{V}^B)$$
$$= \sum_{i=1}^{N_u} \gamma_i^{(\kappa)} a_i^{(\kappa)} + 2\sum_{i=1}^{N_u} \gamma_i^{(\kappa)} \Re\{\langle A_i^{(\kappa)} \mathbf{V}_i^B \rangle\}$$
$$- \sum_{\ell=1}^{N_u} \langle \mathcal{Q}^{(\kappa)}, [\mathbf{V}_\ell^B]^2 \rangle, \quad (40)$$

for $\mathcal{Q}^{(\kappa)} \triangleq \sum_{\ell=1}^{N_u} \gamma_\ell^{(\kappa)} (\mathcal{H}_\ell(\theta^{(\kappa)}))^H B_\ell^{(\kappa)} \mathcal{H}_i(\theta^{(\kappa)})$. We solve the following problem of tight minorant maximization for (38) to generate $V^{B,(\kappa+1)}$

$$\max_{\mathbf{V}^B} \tilde{F}^{(\kappa)}(\mathbf{V}^B) \triangleq \sum_{i=1}^{N_u} \gamma_i^{(\kappa)} \rho_i^{(\kappa)}(\mathbf{V}^B) \quad \text{s.t.} \quad (30). \quad (41)$$

This convex quadratic problem admits the following closed-form solution

$$V_i^{B,(\kappa+1)} = \begin{cases} \gamma_i^{(\kappa)} (\mathcal{Q}^{(\kappa)})^{-1} (A_i^{(\kappa)})^H \\ \text{if } \sum_{i=1}^{N_u} ||(\mathcal{Q}^{(\kappa)})^{-1} \gamma_i^{(\kappa)} (A_i^{(\kappa)})^H||^2 \leq P/N \\ \left(\mathcal{Q}^{(\kappa)} + \mu I_{N_{RF}}\right)^{-1} \gamma_i^{(\kappa)} (A_i^{(\kappa)})^H \quad \text{otherwise,} \end{cases} \quad (42)$$

where $\mu > 0$ is chosen such that

$$\sum_{i=1}^{N_u} ||(\mathcal{Q}^{(\kappa)} + \mu I_{N_{RF}})^{-1} \gamma_i^{(\kappa)} (A_i^{(\kappa)})^H||^2 = P/N. \quad (43)$$

The computational complexity of (43) is on the order of $\mathcal{O}(N_{RF} N_u N_R)$, i.e. it is linearly scalable in $N_{RF} N_u N_R$. By Lemma 1, we thus have (37) as claimed.

2) Alternating ascent in the ABF: We seek $\theta^{(\kappa+1)}$ feasible for (6) for ensuring that

$$F^{(\kappa)}(\theta^{(\kappa+1)}, V^{B,(\kappa+1)}) > F^{(\kappa)}(\theta^{(\kappa)}, V^{B,(\kappa+1)}), \quad (44)$$

by considering the following problem

$$\max_{\theta} F^{(\kappa)}(\theta, V^{B,(\kappa+1)}) = \langle \gamma^{(\kappa)}, \rho(\theta, V^{B,(\kappa+1)}) \rangle \quad \text{s.t.} \quad (6), \quad (45)$$

This is a combinatoric problem because its objective function is nonlinear, while its constraint (6) is discrete.

Recalling from (18) that $\rho_i(\theta, V^{B,(\kappa+1)}) \triangleq \ln \left| I_{N_R} + [\mathcal{H}_i(\theta) V_i^{B,(\kappa+1)}]^2 \Psi_i^{-1}(\theta, V^{B,(\kappa+1)}) \right|$, and using the inequality (5), we obtain the following tight minorant of $\rho_i(\theta, V^{B,(\kappa+1)})$ at $\theta^{(\kappa)}$:

$$\tilde{\rho}_i^{(\kappa)}(\theta) \triangleq \tilde{a}_i^{(\kappa)} + 2\Re\{\langle \tilde{A}_i^{(\kappa)} V_{RF}(\theta) \rangle\}$$
$$- \langle H_i^H \tilde{B}_i^{(\kappa)} H_i, V_{RF}(\theta) (\sum_{\ell=1}^{N_u} [V_\ell^{B,(\kappa+1)}]^2) V_{RF}^H(\theta) \rangle, (46)$$

with

$$\tilde{a}_i^{(\kappa)} \triangleq \rho_i(\theta^{(\kappa)}, V^{B,(\kappa+1)})$$
$$- \langle [\mathcal{H}_i(\theta^{(\kappa)}) V_i^{B,(\kappa+1)}]^2 \Psi_i^{-1}(\theta^{(\kappa)}, V^{B,(\kappa+1)}) \rangle - \sigma \langle \tilde{B}_i^{(\kappa)} \rangle,$$
$$\tilde{A}_i^{(\kappa)} \triangleq V_i^{B,(\kappa+1)} (\mathcal{H}_i(\theta^{(\kappa)}) V_i^{B,(\kappa+1)})^H \Psi_i^{-1}(\theta^{(\kappa)}, V^{B,(\kappa+1)}) H_i,$$

and

$$\tilde{B}_i^{(\kappa)} \triangleq \Psi_i^{-1}(\theta^{(\kappa)}, V^{B,(\kappa+1)}) - \left(\Psi_i(\theta^{(\kappa)}, V^{B,(\kappa+1)}) + [\mathcal{H}_i(\theta^{(\kappa)}) V_i^{B,(\kappa+1)}]^2\right)^{-1}.$$

Thus, a tight minorant of the objective function in (45) is

$$\sum_{i=1}^{N_u} \gamma_i^{(\kappa)} \tilde{\rho}_i^{(\kappa)}(\theta) =$$
$$\tilde{a}^{(\kappa)} + 2\Re\{\langle \tilde{A}^{(\kappa)} V_{RF}(\theta) \rangle\}$$
$$- \langle \tilde{B}^{(\kappa)}, V_{RF}(\theta) \tilde{C}^{(\kappa)} V_{RF}^H(\theta) \rangle = (47)$$
$$\tilde{a}^{(\kappa)} + 2\Re\{\langle \tilde{A}^{(\kappa)} V_{RF}(\theta) \rangle\} - \lambda^{(\kappa)} \langle [V_{RF}(\theta)]^2 \rangle$$
$$+ \left(\lambda^{(\kappa)} \langle [V_{RF}(\theta)]^2 - \langle \tilde{B}^{(\kappa)}, V_{RF}(\theta) \tilde{C}^{(\kappa)} V_{RF}^H(\theta) \rangle \right) = (48)$$
$$\tilde{a}^{(\kappa)} + 2\Re\{\langle \tilde{A}^{(\kappa)} V_{RF}(\theta) \rangle\} - \lambda^{(\kappa)} N N_{RF}$$
$$+ \left(\lambda^{(\kappa)} \langle [V_{RF}(\theta)]^2 \rangle - \langle \tilde{B}^{(\kappa)}, V_{RF}(\theta) \tilde{C}^{(\kappa)} V_{RF}^H(\theta) \rangle \right) \quad (49)$$

for

$$\tilde{a}^{(\kappa)} \triangleq \sum_{i=1}^{N_u} \gamma_i^{(\kappa)} \tilde{a}_i^{(\kappa)}, \tilde{A}^{(\kappa)} \triangleq \sum_{i=1}^{N_u} \gamma_i^{(\kappa)} \tilde{A}_i^{(\kappa)},$$
$$\tilde{B}^{(\kappa)} \triangleq \sum_{i=1}^{N_u} \gamma_i^{(\kappa)} H_i^H \tilde{B}_i^{(\kappa)} H_i, \tilde{C}^{(\kappa)} \triangleq \sum_{\ell=1}^{N_u} [V_\ell^{B,(\kappa+1)}]^2,$$

and

$$\lambda^{(\kappa)} \triangleq \lambda_{\max}(\tilde{B}^{(\kappa)}) \lambda_{\max}(\tilde{C}^{(\kappa)}),$$

where for deriving (49) we exploit the following constant modulus of $V_{RF}(\theta)$:

$$\langle [V_{RF}(\theta)]^2 \rangle = ||V_{RF}(\theta)||^2 \equiv N N_{RF} \ \forall \theta. \quad (50)$$

Regarding the last term in (49), note that

$$\lambda^{(\kappa)} \langle [V_{RF}(\theta)]^2 \rangle - \langle \tilde{B}^{(\kappa)}, V_{RF}(\theta) \tilde{C}^{(\kappa)} V_{RF}^H(\theta) \rangle \geq 0 \ \forall V_{RF}(\theta)$$

because

$$\langle \tilde{B}^{(\kappa)}, V_{RF}(\boldsymbol{\theta})\tilde{C}^{(\kappa)}V_{RF}^H(\boldsymbol{\theta})\rangle \leq$$
$$\lambda_{\max}(\tilde{B}^{(\kappa)})\lambda_{\max}(\tilde{C}^{(\kappa)})\langle V_{RF}(\boldsymbol{\theta})V_{RF}^H(\boldsymbol{\theta})\rangle =$$
$$\lambda^{(\kappa)}\langle [V_{RF}(\boldsymbol{\theta})]^2\rangle.$$

Thus, $\lambda^{(\kappa)}\langle [V_{RF}(\boldsymbol{\theta})]^2\rangle - \langle \tilde{B}^{(\kappa)}, V_{RF}(\boldsymbol{\theta})\tilde{C}^{(\kappa)}V_{RF}^H(\boldsymbol{\theta})\rangle$ is nonnegative quadratic function of $V_{RF}(\boldsymbol{\theta})$, so it must be convex [27]. Hence, its linearization at $V_{RF}(\theta^{(\kappa)})$ provides its tight minorant [27]:

$$2\Re\{\langle \lambda^{(\kappa)}V_{RF}^H(\theta^{(\kappa)})V_{RF}(\boldsymbol{\theta})\rangle\}$$
$$-2\Re\{\langle \tilde{C}^{(\kappa)}V_{RF}^H(\theta^{(\kappa)})\tilde{B}^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\}$$
$$-\lambda^{(\kappa)}NN_{RF} + \langle \tilde{C}^{(\kappa)}V_{RF}^H(\theta^{(\kappa)})\tilde{B}^{(\kappa)}V_{RF}(\theta^{(\kappa)})\rangle. \quad (51)$$

As $V_{RF}(\boldsymbol{\theta})$ is a nonlinear mapping, (51) provides a nonlinear function of $\boldsymbol{\theta}$. It follows from (49) and (51) that a tight minorant of the objective function in (45) is the following trigonometric function:

$$\hat{F}^{(\kappa)}(\boldsymbol{\theta}) \triangleq \tilde{a}_p^{(\kappa)} + 2\Re\{\langle \tilde{D}^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\}, \quad (52)$$

for $\tilde{a}_p^{(\kappa)} \triangleq \tilde{a}^{(\kappa)} - 2\lambda^{(\kappa)}NN_{RF} + \langle \tilde{C}^{(\kappa)}V_{RF}^H(\theta^{(\kappa)})\tilde{B}^{(\kappa)}V_{RF}(\theta^{(\kappa)})\rangle$, and $\tilde{D}^{(\kappa)} \triangleq \tilde{A}^{(\kappa)} + \lambda^{(\kappa)}V_{RF}^H(\theta^{(\kappa)}) - \tilde{C}^{(\kappa)}V_{RF}^H(\theta^{(\kappa)})\tilde{B}^{(\kappa)}$. We thus solve the following problem of tight minorant maximization for (45) to generate $\theta^{(\kappa+1)}$:

$$\max_{\boldsymbol{\theta}} \hat{F}^{(\kappa)}(\boldsymbol{\theta}) \quad \text{s.t.} \quad (6). \quad (53)$$

This trigonometric discrete problem admits the following closed-form solution

$$\theta^{(\kappa+1)} = \left[2\pi - \lfloor \angle \tilde{D}^{(\kappa)}(j,n)\rfloor_b\right]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}. \quad (54)$$

The computational complexity of (54) is on the order of $\mathcal{O}(NN_uN_R)$, i.e. it is also linearly scalable in NN_uN_R. By Lemma 1, we have (44) as claimed, provided that $\theta^{(\kappa+1)} \neq \theta^{(\kappa)}$.

As the first booster, we use (28) to obtain the following approximated minorant of the objective function in (47)

$$\text{RHS of (47)}$$
$$= \tilde{a}^{(\kappa)} + 2\Re\{\langle \tilde{A}^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\}$$
$$-\langle \lambda_{\max}(\tilde{B}^{(\kappa)})I_N, V_{RF}(\boldsymbol{\theta})\tilde{C}^{(\kappa)}V_{RF}^H(\boldsymbol{\theta})\rangle$$
$$+\langle \lambda_{\max}(\tilde{B}^{(\kappa)})I_N - \tilde{B}^{(\kappa)}, V_{RF}(\boldsymbol{\theta})\tilde{C}^{(\kappa)}V_{RF}^H(\boldsymbol{\theta})\rangle$$
$$\approx \tilde{a}^{(\kappa)} + 2\Re\{\langle \tilde{A}^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\} - \lambda_{\max}(\tilde{B}^{(\kappa)})N\langle \tilde{C}^{(\kappa)}\rangle$$
$$+\langle \lambda_{\max}(\tilde{B}^{(\kappa)})I_N - \tilde{B}^{(\kappa)}, V_{RF}(\boldsymbol{\theta})\tilde{C}^{(\kappa)}V_{RF}^H(\boldsymbol{\theta})\rangle$$
$$\geq \tilde{\tilde{a}}^{(\kappa)} + 2\Re\{\langle \tilde{\tilde{D}}^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\}\rangle, \quad (55)$$

with $\tilde{a}^{(\kappa)}$, $\tilde{A}^{(\kappa)}$, $\tilde{B}^{(\kappa)}$, and $\tilde{C}^{(\kappa)}$ defined from (47), and then

$$\tilde{\tilde{a}}^{(\kappa)} \triangleq \tilde{a}^{(\kappa)} - \lambda_{\max}(\tilde{B}^{(\kappa)})N\langle \tilde{C}^{(\kappa)}\rangle$$
$$-\langle \lambda_{\max}(\tilde{B}^{(\kappa)})I_N - \tilde{B}^{(\kappa)}, V_{RF}(\theta^{(\kappa)})\tilde{C}^{(\kappa)}V_{RF}^H(\theta^{(\kappa)})\rangle,$$

and

$$\tilde{\tilde{D}}^{(\kappa)} \triangleq \tilde{A}^{(\kappa)} + \tilde{C}^{(\kappa)}V_{RF}^H(\theta^{(\kappa)})\left(\lambda_{\max}(\tilde{B}^{(\kappa)})I_N - \tilde{B}^{(\kappa)}\right),$$

based on which we generate $\theta^{(\kappa+1)}$ by solving the following problem of approximate minorant maximization for (45):

$$\max_{\boldsymbol{\theta}} \tilde{\tilde{a}}^{(\kappa)} + 2\Re\{\langle \tilde{\tilde{D}}^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\} \quad \text{s.t.} \quad (6), \quad (56)$$

which admits the following closed-form solution

$$\theta^{(\kappa+1)} = \left[2\pi - \lfloor \angle \tilde{\tilde{D}}^{(\kappa)}(j,n)\rfloor_b\right]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}. \quad (57)$$

As the second booster, we also generate $\theta^{(\kappa+1)}$ by

$$\theta^{(\kappa+1)} = \left[2\pi - \lfloor \angle \hat{D}^{(\kappa)}(j,n)\rfloor_b\right]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}. \quad (58)$$

with $\hat{D}^{(\kappa)} \triangleq \tilde{A}^{(\kappa)} - \tilde{C}^{(\kappa)}V_{RF}^H(\theta^{(\kappa)})\tilde{B}^{(\kappa)}$, which is the optimal solution by maximizing the linearized function of the tight minorant defined by (47). Note that this function and that defined by (55) are not necessarily minorants of the objective in (45) so $\theta^{(\kappa+1)}$ generated by (57) and (58) are not necessarily better than $\theta^{(\kappa)}$.

3) Algorithm and its convergence: Algorithm 1 is the pseudo-code of implementing the alternating ascent iterations (42) and (54), (57), and (58) in addressing the problem (31). Note that

$$\langle \partial F(\rho(\theta^{(\kappa)}, V^{B,(\kappa)}))/\partial\rho, \rho(\boldsymbol{\theta}, \mathbf{V}^B)\rangle = \delta^{(\kappa)}F^{(\kappa)}(\boldsymbol{\theta}, \mathbf{V}^B) \quad (59)$$

with $\delta^{(\kappa)} \triangleq F(\rho(\theta^{(\kappa)}, V^{B,(\kappa)}))/(N_u\max_{i'\in\mathcal{N}_u}\rho_{i'}(\theta^{(\kappa)}, V^{B,(\kappa)})) > 0$, and as such the sequence $\{(V^{B,(\kappa)}, \theta^{(\kappa)})\}$ converges to $(V^{B,(\infty)}, \theta^{(\infty)})$, which satisfies the first-order optimality condition in $\mathbf{V}^B$ ($\boldsymbol{\theta}$, resp. with other variable held fixed at $\theta^{(\infty)}$ ($V^{B,(\infty)}$, resp.) [27]. Moreover, as it will shown by simulations in Section IV, the following inequalities are achieved after a few iterations

$$F(\theta^{(\kappa+1)}, V^{B,(\kappa+1)}) > F(\theta^{(\kappa)}, V^{B,(\kappa+1)}) > F(\theta^{(\kappa)}, V^{B,(\kappa)}), \quad (60)$$

so $V^{B,(\kappa+1)}$ generated by (42) is an ascent of the full-length step, while $\theta^{(\kappa+1)}$ generated by (54) provides a new ascent for discrete optimization.

Algorithm 1 AO algorithm for locating the ABF matrix.

1: **Initialization:** Initialize $(\theta^{(0)}, V^{B,(0)})$.
2: **Repeat until convergence of the objective function in (31):** Generate $V^{B,(\kappa+1)}$ by (42) and $\theta^{(\kappa+1)}$ by taking the best of three generated by (54), (57), and (58). Reset $\kappa := \kappa + 1$.
3: **Output** $(\theta^{opt}, \tilde{V}^{B,opt}) = (\theta^{(\kappa)}, V^{B,(\kappa)})$.

B. Performance booster by DBF

Until now, Algorithm 1 has not addressed the original problem (27), only its approximation problem (31). This subsection considers the problem of optimizing the DBF matrix $\mathbf{V}^B$ in (27) with $\boldsymbol{\theta}$ held fixed at θ^{opt} that was found by Algorithm 1.

Under $\boldsymbol{\theta} = \theta^{opt}$, the rate of UE i defined by (18) is $r_i(\mathbf{V}^B) \triangleq \rho_i(\theta^{opt}, \mathbf{V}^B) = \ln\left|I_{N_R} + [\tilde{H}_i\mathbf{V}_i^B]^2\Phi_i^{-1}(\mathbf{V}^B)\right|$, with $\tilde{H}_i = \mathcal{H}_i(\theta^{opt})$, and $\Phi_i(\mathbf{V}^B) \triangleq \Psi_i(\theta^{opt}, \mathbf{V}^B) = \sum_{\ell\neq i}^{N_u}[\tilde{H}_i\mathbf{V}_\ell^B]^2 + \sigma I_{N_R}$. Thus, the problem (27) with $\boldsymbol{\theta}$ held fixed at θ^{opt} is formulated as:

$$\max_{\mathbf{V}^B} f_{GM}(\mathbf{V}^B) \triangleq \left(\prod_{i=1}^{N_u} r_i(\mathbf{V}^B)\right)^{1/N_u} \quad (61a)$$

$$\text{s.t.} \quad \sum_{i=1}^{N_u} ||V_{RF}(\theta^{opt})\mathbf{V}_i^B||^2 \le P. \tag{61b}$$

Initialized from

$$V^{B,0} = t_0 \tilde{V}^{B,opt}, t_0 = \sqrt{P/||V_{RF}(\theta^{(\kappa)})\tilde{V}^{B,opt}||^2} \tag{62}$$

with $\tilde{V}^{B,opt}$ found by Algorithm 1, for $\kappa = 0,1,\ldots$, we iterate $V^{B,(\kappa+1)}$ based on the problem

$$\max_{\mathbf{V}^B} f^{(\kappa)}(\mathbf{V}^B) \triangleq \langle \gamma^{(\kappa)}, r(\mathbf{V}^B) \rangle \quad \text{s.t.} \quad (61b), \tag{63}$$

where $r(\mathbf{V}^B) \triangleq (r_1(\mathbf{V}^B),\ldots,r_{N_u}(\mathbf{V}^B))^T$ and $\gamma^{(\kappa)} \triangleq (\gamma_1^{(\kappa)},\ldots,\gamma_{N_u}^{(\kappa)})^T$ with

$$\gamma_i^{(\kappa)} = \frac{\max_{i'\in\mathcal{N}_u} r_{i'}(V^{B,(\kappa)})}{r_i(V^{B,(\kappa)})}, i \in \mathcal{N}_u. \tag{64}$$

Upon using the inequality (5), we obtain the following tight minorant of $r_i(\mathbf{V}^B)$:

$$\begin{aligned} r_i^{(\kappa)}(\mathbf{V}^B) &\triangleq \tilde{a}_i^{(\kappa)} + 2\Re\{\langle A_i^{(\kappa)}\mathbf{V}_i^B\rangle\} \\ &\quad - \langle B_i^{(\kappa)}\left(\sum_{\ell=1}^{N_u}[\tilde{H}_i\mathbf{V}_\ell^B]^2 + \sigma I_{N_R}\right)\rangle \tag{65} \\ &= a_i^{(\kappa)} + 2\Re\{\langle A_i^{(\kappa)}\mathbf{V}_i^B\rangle\} \\ &\quad - \langle \tilde{H}_i^H B_i^{(\kappa)}\tilde{H}_i, \sum_{\ell=1}^{N_u}[\mathbf{V}_\ell^B]^2\rangle, \tag{66} \end{aligned}$$

with

$$\tilde{a}_i^{(\kappa)} \triangleq r_i(V^{B,(\kappa)}) - \langle [\tilde{H}_iV_i^{B,(\kappa)}]^2\Phi_i^{-1}(V^{B,(\kappa)})\rangle,$$

$$a_i^{(\kappa)} \triangleq \tilde{a}_i^{(\kappa)} - \sigma\langle B_i^{(\kappa)}\rangle,$$

$$A_i^{(\kappa)} \triangleq (V_i^{B,(\kappa)})^H \tilde{H}_i^H \Phi_i^{-1}(V^{B,(\kappa)})\tilde{H}_i,$$

and

$$B_i^{(\kappa)} \triangleq \Phi_i^{-1}(V^{B,(\kappa)}) - \left(\Phi_i(V^{B,(\kappa)}) + [\tilde{H}_iV_i^{B,(\kappa)}]^2\right)^{-1}.$$

For $r^{(\kappa)}(\mathbf{V}^B) \triangleq (r_1^{(\kappa)}(\mathbf{V}^B),\ldots,r_{N_u}^{(\kappa)}(\mathbf{V}^B))^T$, we then have the following tight minorant of the objective function in (63):

$$\begin{aligned} \tilde{f}^{(\kappa)}(\mathbf{V}^B) &\triangleq \langle \gamma^{(\kappa)}, r^{(\kappa)}(\mathbf{V}^B)\rangle \\ &= a^{(\kappa)} + 2\sum_{i=1}^{N_u}\Re\{\langle \gamma_i^{(\kappa)}A_i^{(\kappa)}\mathbf{V}_i^B\rangle\} \\ &\quad - \sum_{i=1}^{N_u}\langle \mathcal{Q}^{(\kappa)},[\mathbf{V}_i^B]^2\rangle \tag{67} \end{aligned}$$

where $a^{(\kappa)} \triangleq \sum_{i=1}^{N_u} a_i^{(\kappa)}$, and $\mathcal{Q}^{(\kappa)} \triangleq \sum_{\ell=1}^{N_u}\gamma_\ell^{(\kappa)}\tilde{H}_\ell^H B_\ell^{(\kappa)}\tilde{H}_\ell$.

We solve the following problem of tight minorant maximization for (63) to generate $V^{B,(\kappa+1)}$:

$$\max_{\mathbf{V}^B} \tilde{f}^{(\kappa)}(\mathbf{V}^B) \quad \text{s.t.} \quad (61b). \tag{68}$$

This convex quadratic problem admits the following closed-form solution

$$V_i^{B,(\kappa+1)} = \begin{cases} \gamma_i^{(\kappa)}(\mathcal{Q}^{(\kappa)})^{-1}(A_i^{(\kappa)})^H \\ \quad \text{if } \sum_{i=1}^{N_u}|V_{RF}(\theta^{opt})(\mathcal{Q}^{(\kappa)})^{-1}\gamma_i^{(\kappa)}(A_i^{(\kappa)})^H|^2 \le P \\ (\mathcal{Q}^{(\kappa)} + \mu[V_{RF}^H(\theta^{opt})]^2)^{-1}\gamma_i^{(\kappa)}(A_i^{(\kappa)})^H \\ \quad \text{otherwise,} \end{cases} \tag{69}$$

where $\mu > 0$ is chosen for ensuring that

$$\sum_{i=1}^{N_u}||V_{RF}(\theta^{opt})\left(\mathcal{Q}^{(\kappa)}+\mu[V_{RF}^H(\theta^{opt})]^2\right)^{-1}\gamma_i^{(\kappa)}(A_i^{(\kappa)})^H||^2 = P. \tag{70}$$

By Lemma 1

$$f^{(\kappa)}(V^{B,(\kappa+1)}) > f^{(\kappa)}(V^{B,(\kappa)}), \tag{71}$$

provided that $f^{(\kappa)}(V^{B,(\kappa+1)}) \ne f^{(\kappa)}(V^{B,(\kappa)})$, i.e. $V^{B,(\kappa+1)}$ is a better feasible point than $V^{B,(\kappa)}$. Like Algorithm 1, Algorithm 2 provided below generates a sequence $\{V^{B,(\kappa)}\}$ of gradually improved feasible points for (61), which converges at least to a locally optimal solution.

Algorithm 2 Booster algorithm

1: **Initialization:** Initialize $V^{B,(0)}$ by (62).
2: **Repeat until convergence of the objective function in (61):** Generate $V^{B,(\kappa+1)}$ by (69). Reset $\kappa := \kappa + 1$.
3: **Output** $V^{B,(\kappa)}$.

III. PENALTY OPTIMIZATION APPROACH

Recalling the definition (15) of HBF and (23) with $\widetilde{\Psi}_i(\mathbf{V}^{HD})$ defined from (24) as the achievable rate of UE, our HBF design problem of maximizing the GM-rate is formulated as:

$$\max_{\boldsymbol{\theta},\mathbf{V}^B,\mathbf{V}^{HD}} f(\mathbf{V}^{HD}) \triangleq \left(\prod_{i=1}^{N_u}\hat{r}_i(\mathbf{V}^{HD})\right)^{1/N_u} \quad \text{s.t.} \quad (7),(15),(22c). \tag{72}$$

We will address (72) via the following PO:

$$\begin{aligned} \max_{\boldsymbol{\theta},\mathbf{V}^B,\mathbf{V}^{HD}} &F_p(\boldsymbol{\theta},\mathbf{V}^B,\mathbf{V}^{HD}) \\ &\triangleq f(\mathbf{V}^{HD}) - c||\mathbf{V}^{HD} - V_{RF}(\boldsymbol{\theta})\mathbf{V}^B||^2 \\ &\quad \text{s.t.} \quad (6),(22c) \tag{73} \end{aligned}$$

with the penalty parameter $c > 0$ to be updated.

Initialized by $(\theta^{(0)}, V^{B,(0)}, V^{HD,(0)})$ which is feasible for (72), for $\kappa = 0,1,\ldots$, we let $(\theta^{(\kappa)}, V^{B,(\kappa)}, V^{HD,(\kappa)})$ be a feasible point for (73) that is found from the $(\kappa-1)$-st round and $\gamma^{(\kappa)} \triangleq (\gamma_1^{(\kappa)},\ldots,\gamma_{N_u}^{(\kappa)})^T$ with

$$\gamma_i^{(\kappa)} \triangleq \frac{\max_{\ell\in\mathcal{N}_u}\hat{r}_\ell(V^{HD,(\kappa)})}{\hat{r}_i(V^{HD,(\kappa)})}, i \in \mathcal{N}_u. \tag{74}$$

For $\hat{r}(\mathbf{V}^{HD}) \triangleq (\hat{r}_1(\mathbf{V}^{HD}), \ldots, \hat{r}_{N_u}(\mathbf{V}^{HD}))^T$, we iterate $(\theta^{(\kappa+1)}, V^{B,(\kappa+1)}, V^{HD,(\kappa+1)})$ by solving the following problem

$$\max_{\boldsymbol{\theta}, \mathbf{V}^B, \mathbf{V}^{HD}} F_p^{(\kappa)}(\boldsymbol{\theta}, \mathbf{V}^B, \mathbf{V}^{HD})$$
$$\triangleq \langle \gamma^{(\kappa)}, \hat{r}(\mathbf{V}^{HD}) \rangle - c||\mathbf{V}^{HD} - V_{RF}(\boldsymbol{\theta})\mathbf{V}^B||^2$$
$$\text{s.t. } (6), (22c). \quad (75)$$

A. Alternating ascent in $\mathbf{V}^{HD}$

We seek $V^{HD,(\kappa+1)}$ such that

$$F_p^{(\kappa)}(\theta^{(\kappa)}, V^{B,(\kappa)}, V^{HD,(\kappa+1)}) > F_p^{(\kappa)}(\theta^{(\kappa)}, V^{B,(\kappa)}, V^{HD,(\kappa)}) \quad (76)$$

by solving the problem

$$\max_{\mathbf{V}^{HD}} F_p^{(\kappa)}(\theta^{(\kappa)}, V^{B,(\kappa)}, \mathbf{V}^{HD}) \quad \text{s.t. } (22c). \quad (77)$$

By exploiting the inequality (5), we obtain the following tight concave quadratic minorant of the function $\hat{r}_i(\mathbf{V}^{HD})$ at $V^{HD,(\kappa)}$:

$$\hat{r}_i^{(\kappa)}(\mathbf{V}^{HD}) \triangleq a^{(\kappa)} + 2\Re\{\langle A_i^{(\kappa)} \mathbf{V}_i^{HD} \rangle\}$$
$$- \langle H_i^H B_i^{(\kappa)} H_i, \sum_{\ell=1}^{N_u} [\mathbf{V}_\ell^{HD}]^2 \rangle, \quad (78)$$

with

$$a_i^{(\kappa)} \triangleq r_i(V^{HD,(\kappa)}) - \langle [H_i V_i^{HD,(\kappa)}]^2 \widetilde{\Psi}_i^{-1}(V^{HD,(\kappa)}) \rangle$$
$$- \sigma \langle B_i^{(\kappa)} \rangle,$$

$$A_i^{(\kappa)} \triangleq (V_i^{HD,(\kappa)})^H H_i^H \widetilde{\Psi}_i^{-1}(V^{HD,(\kappa)}) H_i,$$

and

$$B_i^{(\kappa)} \triangleq \widetilde{\Psi}_i^{-1}(V^{HD,(\kappa)}) - \left(\widetilde{\Psi}_i(V^{HD,(\kappa)}) + [H_i V_i^{HD,(\kappa)}]^2\right)^{-1}.$$

A tight minorant of the objective function in (77) is the following convex function:

$$\tilde{f}^{(\kappa)}(\mathbf{V}^{HD}) \triangleq \sum_{i=1}^{N_u} \gamma_i^{(\kappa)} \hat{r}_i^{(\kappa)}(\mathbf{V}^{HD})$$
$$- c||\mathbf{V}^{HD} - V_{RF}(\theta^{(\kappa)})V^{B,(\kappa)}||^2 \quad (79)$$
$$= a^{(\kappa)} + 2\sum_{i=1}^{N_u} \Re\{\langle A_{c,i}^{(\kappa)} \mathbf{V}_i^{HD} \rangle\}$$
$$- \sum_{i=1}^{N_u} \langle \mathcal{Q}^{(\kappa)}, [\mathbf{V}_i^{HD}]^2 \rangle \quad (80)$$

for

$$a^{(\kappa)} \triangleq \sum_{i=1}^{N_u} a_i^{(\kappa)} - c||V_{RF}(\theta^{(\kappa)})V^{B,(\kappa)}||^2,$$

$$A_{c,i}^{(\kappa)} \triangleq \gamma_i^{(\kappa)} A_i^{(\kappa)} + c(V_i^{B,(\kappa)})^H V_{RF}^H(\theta^{(\kappa)}),$$

and

$$\mathcal{Q}^{(\kappa)} \triangleq \sum_{\ell=1}^{N_u} \gamma_\ell^{(\kappa)} H_\ell^H B_\ell^{(\kappa)} H_\ell + cI_N.$$

We generate $V^{HD,(\kappa+1)}$ as the optimal solution of the following convex quadratic problem of tight minorant maximization for (77),

$$\max_{\mathbf{V}^{HD}} \tilde{f}^{(\kappa)}(\mathbf{V}^{HD}) \quad \text{s.t. } (22c). \quad (81)$$

which admits the closed-form solution

$$V_i^{HD,(\kappa+1)} = \begin{cases} (\mathcal{Q}^{(\kappa)})^{-1} A_{c,i}^{(\kappa)} \\ \quad \text{if } \sum_{i=1}^{N_u} ||(\mathcal{Q}^{(\kappa)})^{-1} A_{c,i}^{(\kappa)}||^2 \le P \\ (\mathcal{Q}^{(\kappa)} + \mu I_N)^{-1} A_{c,i}^{(\kappa)} \quad \text{otherwise,} \end{cases} \quad (82)$$

where $\mu > 0$ is chosen such that

$$\sum_{i=1}^{N_u} ||(\mathcal{Q}^{(\kappa)} + \mu I_N)^{-1} A_{c,i}^{(\kappa)}||^2 = P. \quad (83)$$

By Lemma 1, we have (76) as claimed, provided that $\tilde{f}^{(\kappa)}(V^{HD,(\kappa+1)}) \ne \tilde{f}^{(\kappa)}(V^{HD,(\kappa)})$.

B. Alternating ascent in $\boldsymbol{\theta}$

We seek $\theta^{(\kappa+1)}$ such that

$$F_p(\theta^{(\kappa+1)}, V^{B,(\kappa)}, V^{HD,(\kappa+1)}) > $$
$$F_p(\theta^{(\kappa)}, V^{B,(\kappa)}, V^{HD,(\kappa+1)}), \quad (84)$$

which is equivalent to

$$-||V^{HD,(\kappa+1)} - V_{RF}(\theta^{(\kappa+1)})V^{B,(\kappa)}||^2 > $$
$$-||V^{HD,(\kappa+1)} - V_{RF}(\theta^{(\kappa)})V^{B,(\kappa)}||^2. \quad (85)$$

To this end, we solve the following combinatoric problem

$$\max_{\boldsymbol{\theta}} \varphi^{(\kappa)}(\boldsymbol{\theta}) \triangleq -||V^{HD,(\kappa+1)} - V_{RF}(\boldsymbol{\theta})V^{B,(\kappa)}||^2$$
$$\text{s.t. } (6), \quad (86)$$

where we have:

$$\varphi^{(\kappa)}(\boldsymbol{\theta})$$
$$= -||V^{HD,(\kappa+1)}||^2 - \langle [V^{B,(\kappa)}]^2 V_{RF}^H(\boldsymbol{\theta}) V_{RF}(\boldsymbol{\theta}) \rangle$$
$$+ 2\Re\{\langle V^{B,(\kappa)} (V^{HD,(\kappa+1)})^H V_{RF}(\boldsymbol{\theta}) \rangle\}$$
$$= -||V^{HD,(\kappa+1)}||^2 - \lambda^{(\kappa)} NN_{RF}$$
$$+ 2\Re\{\langle V^{B,(\kappa)} (V^{HD,(\kappa+1)})^H V_{RF}(\boldsymbol{\theta}) \rangle\}$$
$$+ \left[\lambda^{(\kappa)}||V_{RF}(\boldsymbol{\theta})||^2 - \langle [V^{B,(\kappa)}]^2 V_{RF}^H(\boldsymbol{\theta}) V_{RF}(\boldsymbol{\theta}) \rangle\right]$$
$$\ge -||V^{HD,(\kappa+1)}||^2 - \lambda^{(\kappa)} NN_{RF}$$
$$+ 2\Re\{\langle V^{B,(\kappa)} (V^{HD,(\kappa+1)})^H V_{RF}(\boldsymbol{\theta}) \rangle\}$$
$$+ \left[2\Re\{\lambda^{(\kappa)} \langle V_{RF}^H(\theta^{(\kappa)}) V_{RF}(\boldsymbol{\theta}) \rangle\}\right.$$
$$- 2\Re\{\langle [V^{B,(\kappa)}]^2 V_{RF}^H(\theta^{(\kappa)}) V_{RF}(\boldsymbol{\theta}) \rangle\} - \lambda^{(\kappa)} NN_{RF}$$
$$\left. + \langle [V^{B,(\kappa)}]^2 V_{RF}^H(\theta^{(\kappa)}) V_{RF}(\theta^{(\kappa)}) \rangle\right] \quad (87)$$
$$= a^{(\kappa)} + 2\Re\{\langle A^{(\kappa)} V_{RF}(\boldsymbol{\theta}) \rangle\}, \quad (88)$$

with

$$\lambda^{(\kappa)} \triangleq \lambda_{\max}([V^{B,(\kappa)}]^2),$$

$$a^{(\kappa)} \triangleq -||V^{HD,(\kappa+1)}||^2 - 2\lambda^{(\kappa)} NN_{RF}$$
$$+ \langle [V^{B,(\kappa)}]^2 V_{RF}^H(\theta^{(\kappa)}) V_{RF}(\theta^{(\kappa)}) \rangle,$$

and

$$A^{(\kappa)} \triangleq V^{B,(\kappa)}(V^{HD,(\kappa+1)})^H + \lambda^{(\kappa)}\langle V_{RF}^H(\theta^{(\kappa)}) \\ -[V^{B,(\kappa)}]^2 V_{RF}^H(\theta^{(\kappa)}).$$

Note that (87) follows from (50), while (87) is derived by replacing the last term $\langle [V^{B,(\kappa)}]^2 V_{RF}^H(\boldsymbol{\theta})V_{RF}(\boldsymbol{\theta})\rangle$ in (87), which is a convex quadratic function of $V_{RF}(\boldsymbol{\theta})$, by its linearized function at $V_{RF}(\theta^{(\kappa)})$. One can easily check that the RHS of (88) constitutes a tight minorant of $\varphi^{(\kappa)}(\boldsymbol{\theta})$.

We thus solve the following discrete trigonometric problem of tight minorant maximization for (86) to generate $\theta^{(\kappa+1)}$:

$$\min_{\boldsymbol{\theta}} \; \tilde{\varphi}^{(\kappa)}(\boldsymbol{\theta}) \triangleq a^{(\kappa)} + 2\Re\{\langle A^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\} \quad \text{s.t.} \quad (6), \quad (89)$$

which admits the following closed-form solution:

$$\theta^{(\kappa+1)} = [2\pi - \lfloor \angle A^{(\kappa)}(j,n) \rceil_b]_{(n,j)\in\mathcal{N}\times\mathcal{N}_u}. \quad (90)$$

By Lemma 1, we have (85)/(84) as claimed, provided that $\varphi^{(\kappa)}(\theta^{(\kappa+1)}) \neq \varphi^{(\kappa)}(\theta^{(\kappa)})$.

As its booster, we use (28) and (87) to obtain

$$\varphi^{(\kappa)}(\boldsymbol{\theta}) \approx -||V^{HD,(\kappa+1)}||^2 - N||V^{B,(\kappa)}||^2 \\ + 2\Re\{\langle V^{B,(\kappa)}(V^{HD,(\kappa+1)})^H V_{RF}(\boldsymbol{\theta})\rangle\} \quad (91)$$

$$= \hat{a}^{(\kappa)} + 2\Re\{\langle \hat{A}^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\}, \quad (92)$$

for $\hat{a}^{(\kappa)} \triangleq -||V^{HD,(\kappa+1)}||^2 - N||V^{B,(\kappa)}||^2$ and $\hat{A}^{(\kappa)} \triangleq V^{B,(\kappa)}(V^{HD,(\kappa+1)})^H$, based on which we generate $\theta^{(\kappa+1)}$ by solving the following problem of AO for (86):

$$\max_{\boldsymbol{\theta}} \; \hat{a}^{(\kappa)} + 2\Re\{\langle \hat{A}^{(\kappa)}V_{RF}(\boldsymbol{\theta})\rangle\} \quad \text{s.t.} \quad (6), \quad (93)$$

which admits the closed-form solution:

$$\theta^{(\kappa+1)} = [2\pi - \lfloor \angle \hat{A}^{(\kappa)}(j,n) \rceil_b]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}. \quad (94)$$

C. Alternating optimization in $\mathbf{V}^B$

We seek $V^{B,(\kappa+1)}$ such that

$$V^{B,(\kappa+1)}$$
$$= \arg\max_{\mathbf{V}^B} F_p(\theta^{(\kappa+1)}, \mathbf{V}^B, V^{HD,(\kappa+1)})$$
$$= \arg\min_{\mathbf{V}^B} ||V^{HD,(\kappa+1)} - V_{RF}(\theta^{(\kappa+1)})\mathbf{V}^B||^2$$
$$= (V_{RF}^H(\theta^{(\kappa+1)})V_{RF}(\theta^{(\kappa+1)}))^{-1}V_{RF}^H(\theta^{(\kappa+1)})V^{HD,(\kappa+1)}. \quad (95)$$

D. Algorithm

Algorithm 3 formulates the pseudo-code of implementing the alternating ascent iterations (82), (90) and (94), and (95) in addressing the problem (73). Note that (73) is not exactly the penalized formulation of (72) in terms of the definition in [34]. As such, its performance is sensitive to the specific choice of the penalty parameter c and also to the initial feasible points. Another point is that the dimension of $\mathbf{V}^{HD}$ is $N \times (N_u N_R)$, which is very high compared to the dimension $N_{RF} \times (N_u N_R)$ of the DBF matrix and the closed-form (82) involves the inversion of high-dimensional matrices.

Algorithm 3 PO Algorithm

1: **Initialization:** Initialize $\theta^{(0)}$, $V^{B,(0)}$, and $V^{B,(0)}$. Set $\kappa = 0$ and $c = c_0$.
2: **Repeat until** $||V^{HD,(\kappa)} - V_{RF}(\theta^{(\kappa)})V^{B,(\kappa)}||^2 \leq N_u 10^{-3}$: Generate $V^{HD,(\kappa+1)}$ by (82). Generate $\theta^{(\kappa+1)}$ by taking the best of two generated by (90) and (94), and $V^{B,(\kappa+1)}$ by (95). If $||V^{HD,(\kappa+1)} - V_{RF}(\theta^{(\kappa+1)})V^{B,(\kappa+1)}||^2 > 0.9||V^{HD,(\kappa)} - V_{RF}(\theta^{(\kappa)})V^{B,(\kappa)}||^2$, reset $c := 1.2c$ in (73), and accordingly in (77), (79), (80), (82), (83). Reset $\kappa := \kappa + 1$.
3: **Output** $V^{HD,(\kappa)}$, $\theta^{(\kappa)}$ and $V^{B,(\kappa)}$. Reset

$$V^{B,(\kappa)} \to t_0 V^{B,(\kappa)}, t_0 = \sqrt{P/||V_{RF}(\theta^{(\kappa)})V^{B,(\kappa)}||^2}$$

and then output $f(V_{RF}(\theta^{(\kappa)})V^{B,(\kappa)})$ as well as $\hat{r}_i(V_{RF}(\theta^{(\kappa)})V^{B,(\kappa)})$.

IV. NUMERICAL RESULTS

With the users randomly located within the cell radius of 200 meters, so that half of the users are in the cell-center while the remaining users are placed near the cell-edge. The path-loss of UE i experienced at a distance d_i from the BS is set to $\rho_i = 36.72 + 35.3 \log 10(d_i)$dB taking into account a 16.5 dB gain due to multiple-antenna mmWave transmission [7], [8], [35]. The mmWave channel $H_i \in \mathbb{C}^{N_R \times N}$ between the BS and UE i in (10) is modelled by [12]

$$H_i = F\sqrt{10^{-\rho_i/10}} \sum_{c=1}^{N_c} \sum_{\ell=1}^{N_{sc}} \alpha_{i,c,\ell} a_r\left(\phi_{i,c,\ell}^r\right) a_t^H\left(\phi_{i,c,\ell}^t, \theta_{i,c,\ell}^t\right), \quad (96)$$

where $F = \sqrt{\frac{NN_R}{N_c N_{sc}}}$, N_c is the number of scattering clusters, N_{sc} is the number of scatterers within each cluster, and $\alpha_{i,c,\ell} \sim \mathcal{CN}(0,1)$ is the complex gain of the ℓth path in the cth cluster between the BS and UE i. Under a uniform planar array antenna configuration having half wavelength antenna spacing with N_1 and N_2 elements on horizon and vertical, respectively, the normalized transmit and receive antenna array response vectors $a_t\left(\phi_{i,c,\ell}^t, \theta_{i,c,\ell}^t\right)$ and $a_r\left(\phi_{i,c,\ell}^r\right)$ are defined as [12]:

$$a_t\left(\phi_{i,c,\ell}^t, \theta_{i,c,\ell}^t\right)$$
$$= \frac{1}{\sqrt{N}}\left[1, e^{j\pi(x\sin(\phi_{i,c,\ell}^t)\sin(\theta_{i,c,\ell}^t) + y\cos(\theta_{i,c,\ell}^t))}, \ldots, \right.$$
$$\left. e^{j\pi((N_1-1)\sin(\phi_{i,c,\ell}^t)\sin(\theta_{i,c,\ell}^t) + (N_2-1)\cos(\theta_{i,c,\ell}^t))}\right]^T, \quad (97)$$

and

$$a_r\left(\phi_{i,c,\ell}^r\right) = \frac{1}{\sqrt{N_R}}\left[1, e^{j\pi\sin(\phi_{i,c,\ell}^r)}, \ldots, e^{j(N_R-1)\pi\sin(\phi_{i,c,\ell}^r)}\right]^T, \quad (98)$$

where we have $0 \leq x \leq (N_1 - 1)$ and $0 \leq y \leq (N_2 - 1)$. $\phi_{i,c,\ell}^t$ and $\theta_{i,c,\ell}^t$ are the azimuth angle and elevation angle of departure for the ℓth path in the cth cluster arriving from the BS to the UE i, respectively, $\phi_{k,c,\ell}^r$ is the azimuth angle of arrival for the ℓth path in the cth cluster from the BS to

UE i, the angles are generated according to the Laplacian distribution in conjunction with random mean cluster angles in the interval $[0, 2\pi)$ and spreads of 10 degrees within each cluster. As in [7], we set $N_c = 5$ and $N_{sc} = 10$. Note that H_i can be readily estimated by exploiting the sparsity of the channel in the angular domain [36]–[38].

The carrier frequency is set to 28 GHz, the noise power density is set to -174 dBm/Hz, while the bandwidth is set to B = 100 MHz. Unless otherwise stated, $b = 3$, $P = 15$ dBm, $N_{RF} = 8$, $N_u = 8$, $N = 64$ and $N_1 = 8$ are used. The results are multiplied by $\log_2 e$ to convert the unit nats/sec into the unit bps/Hz. The convergence tolerance of the proposed algorithms is set to 10^{-3}.

Below, we use the following legends to specify the proposed implementations:

- "AO" and "3-bit AO" refer to the performance of the inner approximation algorithm 1 used for determining the ABF matrix having infinite resolution and 3-bit resolution, respectively, and then implementing the baseband DBF Algorithm 2.
- "PO" and "3-bit PO" refer to the results of implementing the penalty optimization Algorithm 3 with $\theta^{(\kappa+1)}$ generated by (90).

A. Comparison with existing algorithms

We use the legends "Shi-Hong" and "3-bit Shi-Hong" to specify the performance of the Algorithm [17] for computing (21).

It is plausible that our proposed Algorithms 1, 2, and 3 are eminently suitable for addressing the problem (21) by setting $\gamma_i^{(\kappa)} \equiv 1$ in (34), (63) and (75), which improve the sum-rate objective function in each iteration as illustrated by Fig. 1.
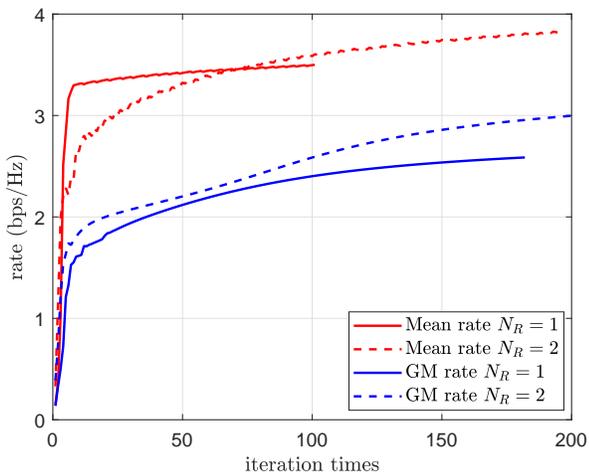

Fig. 1: Convergence of Algorithm 1 in optimizing the sum rate/mean rate and GM rate

Fig. 2 plots the achievable SR of the proposed algorithms AO and PO, and "Shi-Hong" for different number of RF chains. Fig. 2(a) shows that "PO" has the best SR, AO outperforms "Shi-Hong", and their 3-bit resolution solutions follow the same trend under $N_R = 1$. For $N_R = 2$, Fig. 2(b)

shows that "AO" has the best performance for $N_{RF} \leq 6$, while "PO" has the best performance for $N_{RF} \geq 7$, and among all 3-bit resolution algorithms, "3-bit AO" has the best performance for $N_{RF} \leq 5$, while "3-bit PO" has the best performance for $N_{RF} \geq 6$. It is not surprising that AO, PO and "Shi-Hong" have difference performance under $N_R = 1$ and $N_R = 2$ since they enjoy the spatial diversity associated with the increased number of receive antennas at the UE in varying degrees.
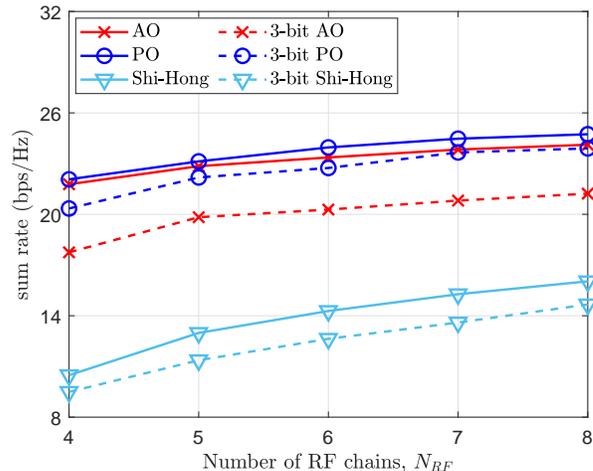
(a)
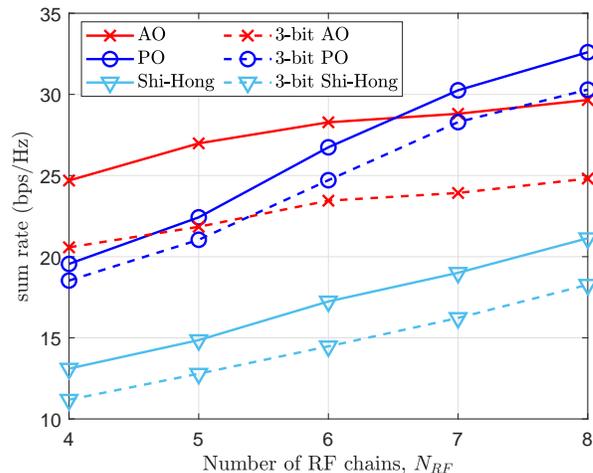
(b)

Fig. 2: Achievable sum rate vs the number of RF chains: (a) $N_R = 1$, (b) $N_R = 2$

To substantiate the fact that SR-based optimization cannot avoid zero rate allocation, Table II provides the average number of zero-rate users (ZR-UEs) in producing Fig. 2. As expected, Table II shows that the number of ZR-UEs increases when N_{RF} is reduced for all proposed algorithms because the less data streams are transmitted. Table II also shows that there are always ZR-UEs in SR maximization.

B. GM-rate performance index and qualitative analysis

Fig. 1 illustrates the convergence behaviour of Algorithm 1, which terminates within 200 iterations, confirming (60) on

TABLE II: The average number of ZR-UEs versus N_{RF}

(a) $N_R = 1$

	$N_{RF} = 4$	$N_{RF} = 5$	$N_{RF} = 6$	$N_{RF} = 7$	$N_{RF} = 8$
AO	4.10	3.81	3.63	3.22	3.07
PO	3.97	3.03	2.31	1.93	1.43
Shi-Hong	2.72	2.03	1.67	1.27	1.07
3-bit AO	4.48	4.27	3.93	3.77	3.69
3-bit PO	3.97	3.03	2.21	1.63	1.42
3-bit Shi-Hong	2.63	1.96	1.48	1.02	0.98

(b) $N_R = 2$

	$N_{RF} = 4$	$N_{RF} = 5$	$N_{RF} = 6$	$N_{RF} = 7$	$N_{RF} = 8$
AO	4.66	4.22	4.13	4.05	3.99
PO	4.17	3.96	3.73	3.55	2.97
Shi-Hong	3.44	2.88	2.30	1.74	1.65
3-bit AO	4.40	4.14	4.13	3.96	3.81
3-bit PO	4.11	3.76	3.48	3.23	2.63
3-bit Shi-Hong	3.14	2.57	2.00	1.77	1.56

steady increase of the GM-rate objective. Algorithm 2 follows a similar pattern.

Fig. 3 plots the achievable GM under different numbers of RF chains for $N_R = 1$ and $N_R = 2$. For $N_R = 1$, "AO" has the best performance except for $N_{RF} = 8$, under which "PO" outperforms "AO". Similarity, "3-bit AO" outperforms "3-bit PO" for $N_{RF} \leq 7$, and "3-bit PO" outperforms "3-bit AO" when $N_{RF} = 8$. For $N_R = 2$, "AO" has the best performance, "3-bit AO" and "3-bit PO" follows the same trend of Fig. 3(a). It can also be observed that AO based algorithms are better benefited from increasing the number of receive antennas than PO based algorithms. As expected, all the algorithms benefit from increasing the number of RF chains. Furthermore, by comparing Fig. 3(a) and Fig. 3(b) it can be found "PO" based algorithms benefit a less extend from the increasing number of receive antennas at the UEs than "AO" based algorithms.

Fig. 4(a) and Fig. 4(b) portray the rate distribution for $N_R = 1$ and $N_R = 2$, respectively, where all the proposed algorithms are capable of avoiding the assignment of zero rate, hence demonstrating superiority.

Table III and Table IV show the min-rate/max-rate and the rate variance versus N_{RF} for $N_R = 1$ and $N_R = 2$, respectively. Table III shows that "PO" has best performance, "3-bit PO" outperforms "AO" and "3-bit AO". Table IV shows that "PO" the best resultant rate distribution, "AO" has the worst resultant rate distribution despite it has the best overall GM rate.

The sum rates of the proposed algorithms are examined in Table V. Table V shows "AO" and "3-bit AO" has better performance than "PO" and "3-bit PO" except when ($N_R = 1, N_{RF} = 8$). Table V also shows "PO" and "3-bit PO" can not benefit from the increasing number of receive antennas at the UEs except for $N_{RF} = 8$.

Fig. 5, which plots the achievable GM under different numbers of BS antennas N for $N_R = 1$ and $N_R = 2$. For $N_R = 1$, "PO" has best GM rate, while for $N_R = 2$ "AO" has the best performance. Among their 3-bit resolutions algorithms, "3-bit PO" outperforms "3-bit AO" all the time. Fig. 5 also shows the performance of all the proposed algorithms is improved by using more resources to exploit beamforming flexibility.
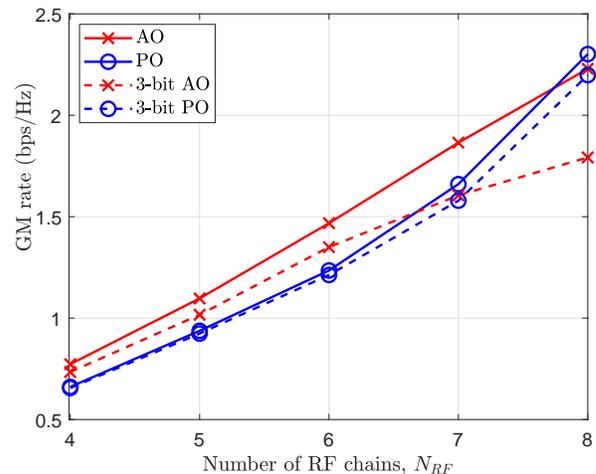
(a)
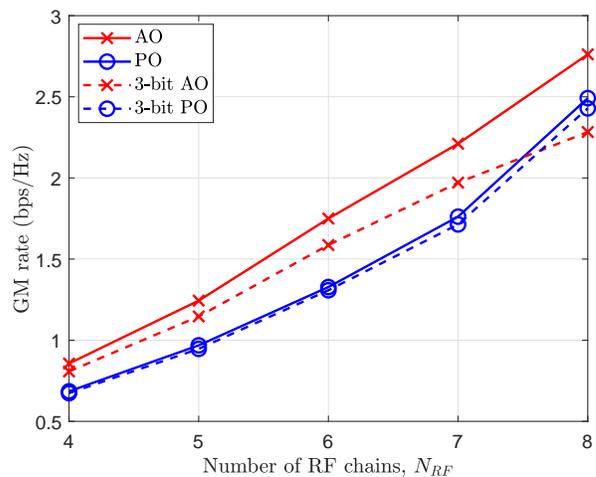
(b)

Fig. 3: Achievable GM vs the number of RF chains: (a) $N_R = 1$, (b) $N_R = 2$

TABLE III: The min-rate/max-rate versus N_{RF}

(a) $N_R = 1$

	$N_{RF} = 4$	$N_{RF} = 5$	$N_{RF} = 6$	$N_{RF} = 7$	$N_{RF} = 8$
AO	0.08	0.08	0.09	0.12	0.21
PO	0.16	0.14	0.16	0.18	0.20
3-bit AO	0.08	0.07	0.07	0.09	0.12
3-bit PO	0.16	0.11	0.13	0.15	0.20

(b) $N_R = 2$

	$N_{RF} = 4$	$N_{RF} = 5$	$N_{RF} = 6$	$N_{RF} = 7$	$N_{RF} = 8$
AO	0.08	0.08	0.09	0.11	0.25
PO	0.21	0.20	0.21	0.23	0.25
3-bit AO	0.12	0.08	0.08	0.11	0.17
3-bit PO	0.21	0.14	0.17	0.20	0.25

TABLE IV: The rate variance versus N_{RF}

(a) $N_R = 1$

	$N_{RF} = 4$	$N_{RF} = 5$	$N_{RF} = 6$	$N_{RF} = 7$	$N_{RF} = 8$
AO	2.52	3.80	3.34	2.73	2.14
PO	0.33	0.82	0.98	1.46	2.34
3-bit AO	1.81	2.78	2.99	2.78	2.73
3-bit PO	0.42	1.13	1.52	1.76	2.17

(b) $N_R = 2$

	$N_{RF} = 4$	$N_{RF} = 5$	$N_{RF} = 6$	$N_{RF} = 7$	$N_{RF} = 8$
AO	3.51	4.85	4.51	3.32	2.22
PO	0.22	0.50	0.73	1.05	1.77
3-bit AO	2.10	3.34	3.54	3.16	2.52
3-bit PO	0.25	0.72	1.02	1.22	1.76

TABLE V: The sum rate (bps/Hz) versus N_{RF}

(a) $N_R = 1$

	$N_{RF} = 4$	$N_{RF} = 5$	$N_{RF} = 6$	$N_{RF} = 7$	$N_{RF} = 8$
AO	10.13	15.42	18.09	19.95	21.18
PO	6.65	9.80	12.32	16.20	21.89
3-bit AO	9.23	13.31	16.61	18.19	19.13
3-bit PO	6.61	9.74	12.29	15.71	20.96

(b) $N_R = 2$

	$N_{RF} = 4$	$N_{RF} = 5$	$N_{RF} = 6$	$N_{RF} = 7$	$N_{RF} = 8$
AO	11.51	17.72	21.58	23.17	25.03
PO	6.45	9.31	12.44	16.16	22.47
3-bit AO	9.77	14.84	19.14	21.07	22.15
3-bit PO	6.37	9.14	12.23	15.66	22.03

We then examine the achievable GM under different power budgets P for $N_R = 1$ and $N_R = 2$ in Fig. 6. As expected, the GM increases upon increasing the available power budget due to the availability of more power for information delivery. And the trend follows Fig. 5.

Furthermore, Fig. 7 allows us to compare the performance achieved by the b-bit solution under different values of b. As expect, all the b-bit resolution algorithms benefit from the increasing b. Fig. 7 also shows an approximately up to 30% and 10% reduction in the GM-rate for a 3-bit quantization compared to the ∞-resolution for AO based algorithms and PO based algorithms, respectively. Furthermore, Fig. 7 also shows "b-bit PO" has similar performance with "PO" for $b \geq 4$.

Finally, Fig. 8 plots the GM attained for different numbers

of RF chains under $N_R = 1$ and $N_R = 2$ in serving more UEs ($N_u = 16$ and $N = 128$). It can be observed that Fig. 8 has the same trend as Fig. 3, namely that "AO" has the best performance except for ($N_R = 1$, $N_{RF} = 8$), "3-bit PO" quickly outperforms "3-bit AO" for $N_{RF} = 8$. The results also confirm the efficiency of the proposed algorithms in serving more users.

V. CONCLUSIONS

We have considered a mmWave communication network, which consists of a base station equipped with a massive antenna array for serving multiple downlink users. We have proposed different hybrid beamforming algorithms optimized by maximizing the GM of the users' rates, which result in a
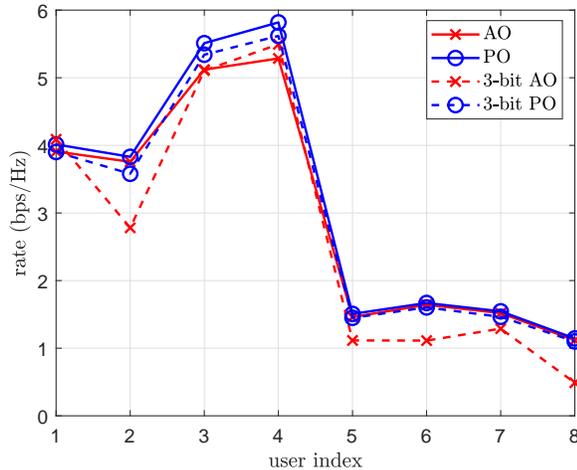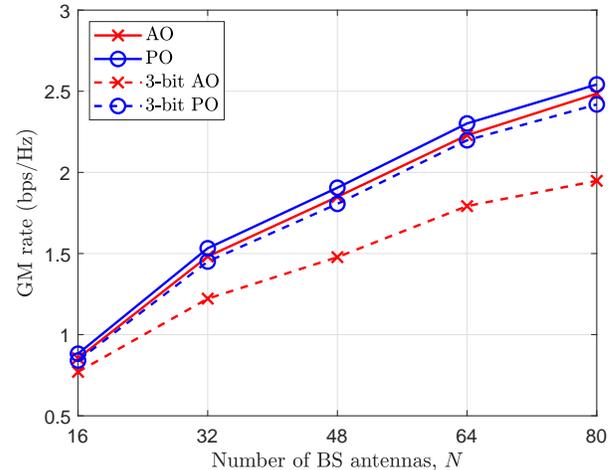
(a)
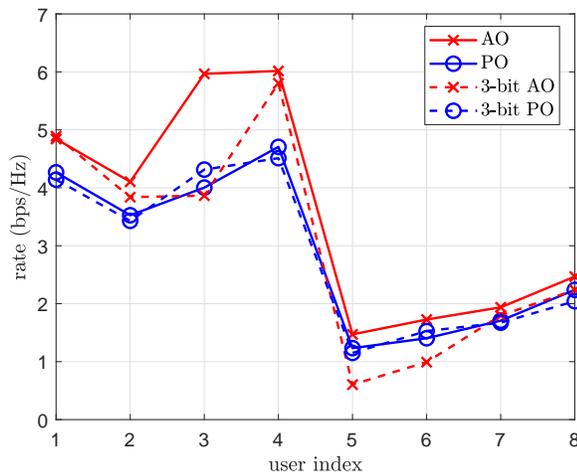
(b)

Fig. 4: Rate distribution: (a) $N_R = 1$, (b) $N_R = 2$

(a)

(b)

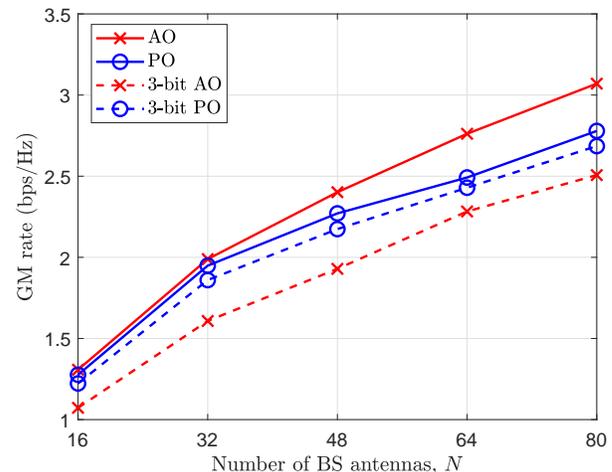Fig. 5: Achievable GM vs the number of BS antennas: (a) $N_R = 1$, (b) $N_R = 2$

fair rate distribution, while maintaining reasonable sum rates. The analog beamforming component was of a low resolution, facilitating its practical implementation in high mmWave bandwidth scenarios. Extension of the GM maximization-based approach to multi-carrier mmWave communication is under our current study.

REFERENCES

[1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, pp. 101–107, Jun. 2011.

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

[3] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, pp. 870–913, Secondquarter 2018.

[4] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Select. Topics Signal Process.*, vol. 10, pp. 436–453, April 2016.

[5] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon, and G. Chen, "Millimeter wave communication: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1616–1653, 2018.

[6] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 836–869, 2018.

[7] 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.

[8] 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.

[9] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, 2014.

[10] M. Li, Z. Wang, H. Li, Q. Liu, and L. Zhou, "A hardware-efficient hybrid beamforming solution for mmwave MIMO systems," *IEEE Wirel. Commun.*, vol. 26, pp. 137–143, February 2019.

[11] K. Satyanarayana, M. El-Hajjar, P. Kuo, A. Mourad, and L. Hanzo, "Hybrid beamforming design for full-duplex millimeter wave communication," *IEEE Trans. Veh. Techn.*, vol. 68, pp. 1394–1404, Feb 2019.
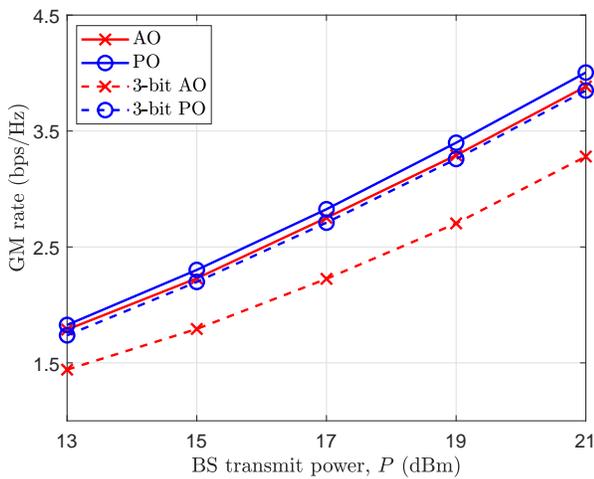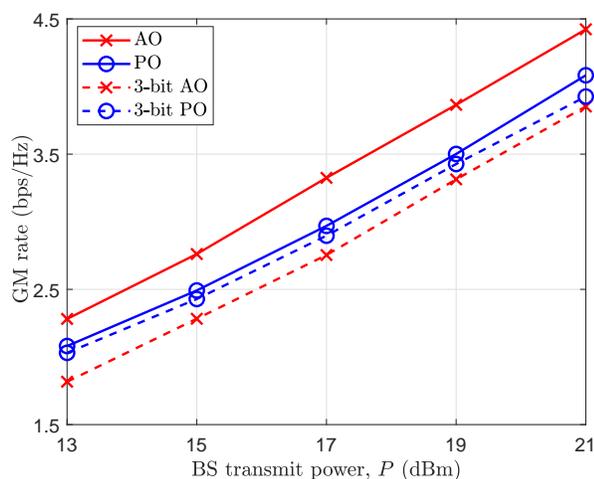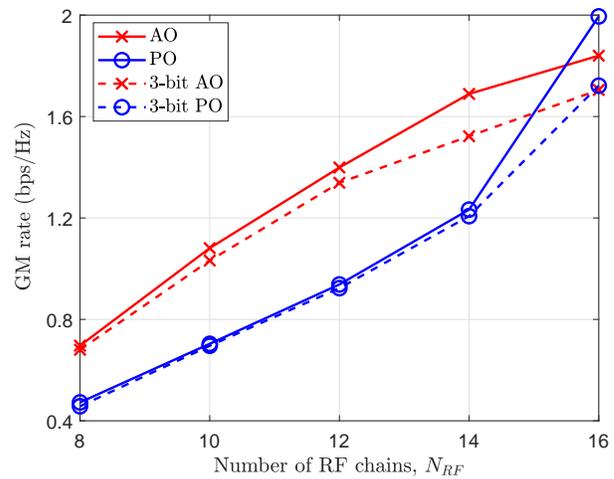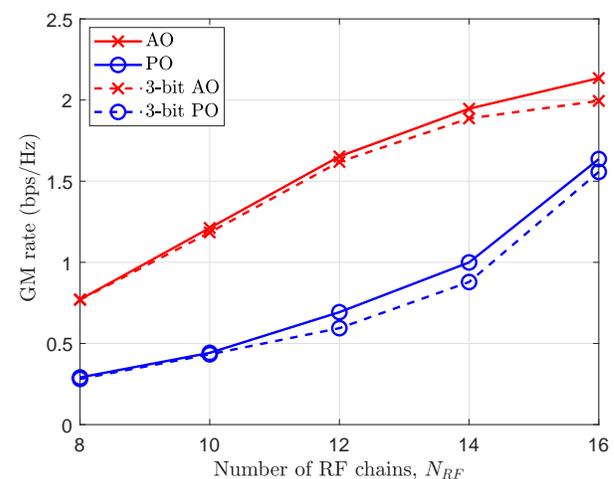
Fig. 6: Achievable GM vs power budget: (a) $N_R = 1$, (b) $N_R = 2$

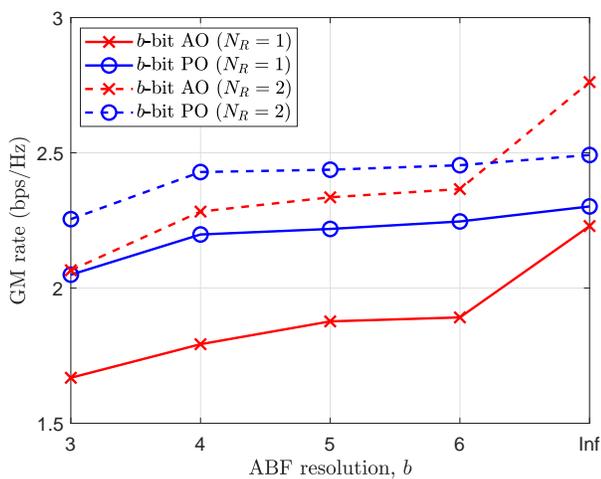Fig. 8: Achievable GM vs the number of RF chains: (a) $N_R = 1$, (b) $N_R = 2$

Fig. 7: Achievable GM vs the value of b

[12] 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, pp. 1499–1513, Mar. 2014.

[13] R. Rajashekar and L. Hanzo, "Iterative matrix decomposition aided block diagonalization for mm-wave multiuser MIMO systems," *IEEE Trans. Wirel. Commun.*, vol. 16, pp. 1372–1384, Mar. 2017.

[14] W. Ni, X. Dong, and W.-S. Lu, "Near-optimal hybrid processing for massive MIMO systems via matrix decomposition," *IEEE Trans. Signal Process.*, vol. 65, pp. 3922–3933, Aug. 2017.

[15] T. Lin, J. Cong, Y. Zhu, J. Zhang, and K. B. Letaief, "Hybrid beam-forming for millimeter wave systems using the MMSE criterion," *IEEE Trans. Commun.*, vol. 67, pp. 3693–3708, May 2019.

[16] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Selec. Topics Signal Process.*, vol. 10, pp. 501–513, Mar. 2016.

[17] Q. Shi and M. Hong, "Spectral efficiency optimization for millimeter wave multiuser MIMO systems," *IEEE J. Select. Topics Signal Process.*, vol. 12, pp. 455–468, Jun. 2018.

[18] 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, pp. 2943–2956, Mar. 2020.

[19] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wirel. Commun. Lett.*, vol. 3, pp. 653–656, Dec. 2014.

[20] 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, pp. 3818–3832, Jul. 2017.

[21] 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, pp. 2340–2353, June 2016.

[22] L. D. Nguyen, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Multi-user regularized zero forcing beamforming," *IEEE Trans. Signal Process.*, vol. 67, pp. 2839–2853, Jun. 2019.

[23] H. T. Nguyen, H. D. Tuan, T. Q. Duong, H. V. Poor, and W.-J. Hwang, "Nonsmooth optimization algorithms for multicast beamforming in content-centric fog radio access networks," *IEEE Trans. Signal Process.*, vol. 68, pp. 1455–1469, 2020.

[24] H. Yu, H. D. Tuan, T. Q. Duong, Y. Fang, and L. Hanzo, "Improper Gaussian signaling for integrated data and energy networking," *IEEE Trans. Commun.*, vol. 68, pp. 3922–3934, Jun. 2020.

[25] H. Yu, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "Joint design of reconfigurable intelligent surfaces and transmit beamforming under proper and improper Gaussian signaling," *IEEE J. Select. Areas Commun.*, vol. 38, pp. 2589–2603, Nov. 2020.

[26] J. D. Krieger, C.-P. Yeang, and G. W. Wornell, "Dense delta-sigma phased arrays," *IEEE Trans. Antenn. Propag.*, vol. 61, pp. 1825–1837, Apr. 2013.

[27] H. Tuy, *Convex Analysis and Global Optimization (second edition)*. Springer International, 2017.

[28] H. Yuan, N. Yang, X. Ding, C. Han, K. Yang, and J. An, "Cluster-based multi-carrier hybrid beamforming for massive device terahertz communications," *IEEE Transactions on Communications*, vol. 70, no. 5, pp. 3407–3420, 2022.

[29] H. Yu, H. D. T. E. Dutkiewicz, H. V. Poor, and L. Hanzo, "Maximizing the geometric mean of user-rates to improve rate-fairness: Proper vs. improper Gaussian signaling," *IEEE Trans. Wirel. Commun.*, vol. 21, pp. 295–309, Jan. 2022.

[30] W. Zhu, H. D. Tuan, E. Dutkiewicz, and L. Hanzo, "Collaborative beamforming aided fog radio access networks," *IEEE Trans. Veh. Techn.*, vol. 71, pp. 7805–7820, Jul. 2022.

[31] H. D. Tuan, A. A. Nasir, H. Q. Ngo, E. Dutkiewicz, and H. V. Poor, "Scalable user rate and energy-efficiency optimization in cell-free massive MIMO," *IEEE Trans. Commun. (early access)*.

[32] A. A. Nasir, H. D. Tuan, E. Dutkiewicz, and L. Hanzo, "Finite-resolution digital beamforming for multi-user millimeter-wave networks," *IEEE Trans. Veh. Techn. (early access)*.

[33] R. Fletcher, *Practical Methods of Optimization (second ed.)*. John Wiley & Sons, 1987.

[34] J. F. Bonnans, J. C. Gilbert, C. Lemarechal, and C. Sagastigabal, *Numerical Optimization-Theoretical and Practical Aspects (second edition)*. Springer, 2006.

[35] "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.

[36] C. Huang, L. Liu, and C. Yuen, "Asymptotically optimal estimation algorithm for the sparse signal with arbitrary distributions," *IEEE Trans. Veh. Technol.*, vol. 67, pp. 10070–10075, Oct 2018.

[37] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using LSE and sparse message passing for mmwave MIMO systems," *IEEE Trans. Signal Process.*, vol. 67, pp. 245–259, Jan 2019.

[38] M. A and A. P. Kannu, "Channel estimation strategies for multi-user mm wave systems," *IEEE Trans. Wirel. Commun.*, vol. 66, pp. 5678–5690, Nov 2018.