# Regularized Zero-Forcing Aided Hybrid Beamforming for Millimeter-Wave Multi-user MIMO Systems

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

***Abstract*—This paper considers hybrid beamforming consisting of analog beamforming (ABF) coupled with digital baseband beamforming (DBF) which is designed for multi-user (MU) multiple input multiple output (MIMO) millimeter-wave (mmWave) communications. ABF uses a limited number of radio frequency (RF) chains and finite-resolution phase-shifters to alleviate the power consumption at the base station (BS), while DBF uses either zero-forcing beamforming (ZFB) or regularized zero forcing beamforming (RZFB) to restrain MU interference. The joint design of ABF and DBF constitutes a computationally challenging mixed discrete continuous optimization problem. The paper develops efficient algorithms for its solution, which iterate scalable-complex expressions. Furthermore, we conceive a new class of MU RZFB for attaining higher rates. Simulations are provided to demonstrate the viability of the proposed algorithms and the advantages of the conceived RZFB.**

***Index Terms*—Multi-user multiple-input-multiple-output millimeter-wave communications, hybrid beamforming, analog beamforming with b-bit resolution, zero-forcing beamforming, regularized zero-forcing beamforming, Brunn-Minkowski geometry, mixed discrete continuous optimization, scalable complexity**

## I. Introduction

Millimeter-wave (mmWave) communication relies on the frequency range spanning from 30 GHz to 300 GHz to deliver gigabit/s rates [1]–[4]. Since the path-loss at these frequencies tends to be high [5], [6], mitigating the power consumption becomes a critical issue for mmWave communication [7]–[9]. Hybrid beamforming (HBF) consisting of analog beamforming (ABF) using a limited number of radio frequency (RF) chains coupled with digital baseband beamfoming (DBF) has been proposed for addressing this issue [10], [11]. However, the design of HBF is challenging, not only because the entries of the ABF matrix are subject to the unit modulus constraint but also because the matrix is also of a large scale, which imposes high-dimensional nonlinear constraints on the joint ABF and DBF matrix optimization. In single-user mmWave systems, the HBF design tends to rely on the product of ABF and DBF matrices to approximate a fully digital beamforming matrix [12]–[16].

As a further development, multi-user (MU) mmWave HBF has been considered in [17]–[24], for example. More particularly, the problem of sum-rate (SR) maximization was addressed both in [23] and [24] by invoking computationally tractable iterative processes, which avoid convex solvers. The MU interference was not considered in [20] in the alternating optimization of the ABF matrix. To elaborate a little further, SR maximization has the weakness that it assigns high rates/powers to a few selected users having the best channel conditions and thus leaving other users with near-to-zero rates. Unfortunately, this impediment cannot be eliminated by imposing a specific minimum user-rate constraint for ensuring fair rate distributions, because this potentially makes the optimization problems computationally intractable.

Both zero forcing beamforming (ZFB) and regularized zero forcing beamforming (RZFB) are suitable for orthogonal or quasi-orthogonal massive multiple input multiple output (MIMO) systems [25]–[27]. MmWave communication also benefits from massive antenna-array, but having a limited number of RF chains for ABF destroys the orthogonality.[1] Hence the design of ZFB and RZFB in HBF is much more challenging than its spatial multiplexing based massive MIMO counterpart hence requiring more research [28]–[31]. As the effective mmWave channels are dependent on the ABF matrices, the works [29]–[31] aim for designing the ABF and DBF separately. The ABF design of [29], [30] aims for maximizing the so-called sum signal-to-leakage-plus-noise-ratio (SLNR) based on the channel covariance under identical fixed-power transmission, which is not directly related to the users' rate. Moreover, this problem is quite complex and is thus simplified by setting equal SLNRs. The ABF design of [31], which matches the ABF to the phase of the channel, results in RZFB that approaches the performance of the optimal fully digital ZFB in terms of the user's worst (minimal) rate as long as the number of RF chains is not lower than the number of users. All results in [28]–[31] are applicable to single-antenna users only, who are served by single information streams.

Against the above background, this paper offers the follow-

This work was supported in part by Technology Key Project of Guangdong Province, China (HZJBGS-2021001), in part by the Australian Research Council's Discovery Projects under Grant DP190102501, in part by the U.S. National Science Foundation under Grant CNS-2128448, and in part by the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/P003990/1 (COALESCE), and the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028).

[1] Shanghai Institute for Advanced Communication and Data Science, Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, 200444 Shanghai, 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@student.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)

[1]Given a limited number of RFs chains, the dimension of the effective channel vectors of the links spanning from the BS to the users is small.

ing contributions:

- The joint design of finite resolution ABF and MIMO ZFB to improve the users' rates via maximizing their geometric mean (GM-rate) is proposed. The users are equipped with multiple antennas to receive multiple information streams. In contrast to conventional SR maximization, GM-rate maximization is capable of improving the users' rate-fairness without explicitly imposing rate constraints.[2] The design problem of maximizing the users' worst rate is also addressed;
- Based on the Brunn-Minkowski geometry of positive definite matrices, the joint design of finite resolution ABF and MIMO RZFB to improve the GM-rate is also developed. A new MIMO RZFB is proposed for improving the GM-rate and thus the users' rates;
- Computationally efficient algorithms, which iterate scalable-complex expressions, are developed for solving the resultant problems of mixed discrete continuous optimization. The discrete constraints imposed on finite-resolution ABF are also efficiently dealt with.

In a nutshell, we boldly contrast our novel contributions to the related literature in Table I.

The paper is organized as follows. Section II is devoted to the joint design of ABF and MIMO ZFB, while Section III is dedicated to the joint design of ABF and MIMO RZFB. Section IV proposes a new MIMO RZFB solution for improving all users' rates. The performance of the proposed designs is evaluated by simulations in Section V, while Section VI concludes the paper.

*Notation.* Only the optimization variables are boldfaced. The inner product between vectors $x$ and $y$ is defined as $\langle x, y\rangle = x^H y$. Analogously, $\langle X, Y\rangle = \mathsf{trace}(X^H Y)$ for matrices $X$ and $Y$. We also use $\langle X\rangle$ for the trace of $X$ when $X$ is a square matrix. $||X||$ is the Frobenius norm of the matrix $X$, which is defined by $\sqrt{\mathsf{trace}(X^H X)}$. $[X]^2$ stands for $XX^H$ so $||X||^2 = \langle [X]^2\rangle$. $X \succeq 0$ ($X \succ 0$, resp.) means that $X$ is Hermitian symmetric ($X^H = X$) and positive semi-definite (definite, resp.). Denote by $\lambda_{\max}$ its maximal eigenvalue. Accordingly, $X \succeq Y$ ($X \succeq Y$, resp.) means $X - Y \succeq 0$ ($X - Y \succ 0$, resp.). diag$[A_k]_{k\in\mathcal{K}}$ is the diagonal matrix with the matrices $A_k$, $k \in \mathcal{K}$ on its diagonal. $I_n$ is the identity matrix of size $n \times n$. For a complex number $x$, we denote the argument by $\angle x$. $\mathcal{CN}_N(0)$ is the set of proper (circular) Gaussian variables in $\mathbb{C}^N$ having zero means. Note that $\mathbb{E}(ss^T) = 0\ \forall s \in \mathcal{CN}_N(0)$.

The following matrix inequalities [34], [35], which hold for all matrices $\mathbf{V} \in \mathbb{C}^{n\times m}$, $\bar{V} \in \mathbb{C}^{n\times m}$, and positive definite matrices $\mathbf{Y} \in \mathbb{C}^{n\times n}$ and $\bar{Y} \in \mathbb{C}^{n\times n}$, are frequently used in the paper:

$$\mathbf{V}^H\mathbf{Y}^{-1}\mathbf{V} \succeq \bar{V}^H\bar{Y}^{-1}\mathbf{V} + \mathbf{V}^H\bar{Y}^{-1}\bar{V} - \bar{V}^H\bar{Y}^{-1}\mathbf{Y}\bar{Y}^{-1}\bar{V}, \tag{1}$$

and

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

[2]The GM-rate based fairness has also been interpreted as a manifestation of proportional fairness (see e.g. [32], [33] and references therein).

One can see that the left hand side (LHS) of the matrix inequality (1) is a nonlinear form of $(\mathbf{V}, \mathbf{Y})$ while the right hand side (RHS) is a linear form of $(\mathbf{V}, \mathbf{Y})$, and they match at $(\bar{V}, \bar{Y})$. The LHS of (2) is a log-determinant function of $(\mathbf{V}, \mathbf{Y})$ while the RHS of (2) is a concave quadratic function of $(\mathbf{V}, \mathbf{Y})$, and they match at $(\bar{V}, \bar{Y})$. Thus, according to [36] the RHS of (2) provides a tight concave quadratic minorant for the LHS of (2).

## II. Joint design of ABF and ZFB

Consider a mmWave communication network of a single base station (BS) serving $K$ downlink users (UEs) having indices as $k \in \mathcal{K} \triangleq \{1, \ldots, K\}$, $K_N$ UEs are located near the BS with indices $k_N \in \mathcal{K}_N \triangleq \{1, \ldots, K_N\}$, and the remaining UEs are located far from the BS with indices $k_F$, $k_F \in \mathcal{K}_F \triangleq \{K_N + 1, \ldots, K\}$. The BS is equipped with a massive $N$-antenna array and $N_{RF}$ RF chains. As such

$$N \gg N_{RF}. \tag{3}$$

Each UE $k$ is equipped with a moderate $N_R$-antenna array. Let $H_k \in \mathbb{C}^{N_R\times N}$ be the channel's impulse response (CIR) spanning from the BS to UE $k$, which is modelled by [12], [37]

$$H_k = \tau\sqrt{10^{-\rho_k/10}}\sum_{c=1}^{N_c}\sum_{\ell=1}^{N_{sc}}\alpha_{k,c,\ell}a_r\left(\phi^r_{k,c,\ell}\right)a_t^H\left(\phi^t_{k,c,\ell}, \theta^t_{k,c,\ell}\right), \tag{4}$$

where $\tau = \sqrt{\frac{NN_R}{N_cN_{sc}}}$, $N_c$ and $N_{sc}$ respectively are the number of scattering clusters and that of scatterers within each cluster. Furthermore, $\alpha_{k,c,\ell}$ is the complex gain of the $\ell$th path in the $c$th cluster between the BS and UE $k$, $\phi^t_{k,c,\ell}$ and $\theta^t_{k,c,\ell}$ are the azimuth angle and elevation angle of departure for the $\ell$th path in the $c$th cluster arriving from the BS at the UE $k$, respectively, $\phi^r_{k,c,\ell}$ is the azimuth angle of arrival for the $\ell$th path in the $c$th cluster from the BS to UE $k$, and $\rho_k$ is the path-loss (in dB) experienced by UE $k$. Under a uniform planar array antenna configuration having half wavelength antenna spacing with $N_1$ and $N_2$ elements in horizon and vertical, respectively, the normalized transmit and receive antenna array response vectors $a_t\left(\phi^t_{k,c,\ell}, \theta^t_{k,c,\ell}\right)$ and $a_r\left(\phi^r_{k,c,\ell}\right)$ are defined as [12]

$$\begin{aligned}&a_t\left(\phi^t_{k,c,\ell}, \theta^t_{k,c,\ell}\right) \\ =\ & \frac{1}{\sqrt{N}}\Big[1, e^{j\pi(x\sin(\phi^t_{k,c,\ell})\sin(\theta^t_{k,c,\ell}) + y\cos(\theta^t_{k,c,\ell}))}, \ldots, \\ & e^{j\pi((N_1-1)\sin(\phi^t_{k,c,\ell})\sin(\theta^t_{k,c,\ell}) + (N_2-1)\cos(\theta^t_{k,c,\ell}))}\Big]^T,\end{aligned} \tag{5}$$

and

$$a_r\left(\phi^r_{k,c,\ell}\right) = \frac{1}{\sqrt{N_R}}\left[1, e^{j\pi\sin(\phi^r_{k,c,\ell})}, \ldots, e^{j(N_R-1)\pi\sin(\phi^r_{k,c,\ell})}\right]^T, \tag{6}$$

where we have $0 \leq x \leq (N_1 - 1)$ and $0 \leq y \leq (N_2 - 1)$. Similar to the HBF design of [14], [15], [38], in this paper, it is assumed that perfect channel state information (CSI) is available at both the transmitter and receiver and that there is perfect synchronization between them. The CIR can be

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

| Contents \ Literature | **This work** | [20], [23], [24] | [28] | [29], [30] | [31] |
|---|---|---|---|---|---|
| MIMO mmWave | √ | | | | |
| finite resolution ABF | √ | | √ | | |
| digital ZFB | √ | | √ | √ | |
| digital RZFB | √ | | | | √ |
| SR maximization | √ | √ | √ | | |
| max-min rate optimization | | | | | √ |
| GM-rate maximization | √ | | | | |
| zero rate issue | | √ | √ | | |
| computational tractability in finite resolution ABF | √ | | | | |
| computational tractability for massive antenna-arrays | √ | √ | √ | | |

readily estimated by exploiting the sparsity of the channel in the angular domain [39]–[41]. Its knowledge is assumed in the paper.

For $\mathcal{N} \triangleq \{1,\dots,N\}$ and $\mathcal{N}_{RF} \triangleq \{1,\dots,N_{RF}\}$, $\boldsymbol{\theta} \triangleq [\boldsymbol{\theta}_{n,j}]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}} \in [0,2\pi)^{N\times N_{RF}}$, let

$$\boldsymbol{\Theta} \triangleq [e^{\jmath\boldsymbol{\theta}_{n,j}}]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}} \in \mathbb{C}^{N\times N_{RF}}$$

represent the phase shift based ABF matrix. Since having an infinite resolution for $\boldsymbol{\theta}_{n,j}$ is not practical for the implementation of mmWave communications [42], we focus our attention on finite resolution, [3]which is represented by the constraint:

$$\boldsymbol{\theta}_{n,j} \in \mathcal{B} \triangleq \{\iota\frac{2\pi}{2^b}, \iota = 0,1,\dots,2^b-1\}, n\in\mathcal{N}; j\in\mathcal{N}_{RF}. \tag{7}$$

In what follows, for $x\in[0,2\pi)$, its projection into $\mathcal{B}$ denoted by $\lfloor x\rceil_b$ is termed as its $b$-bit rounded value, i.e. we have:

$$\lfloor x\rceil_b = \iota_x\frac{2\pi}{2^b} \tag{8}$$

in conjunction with

$$\iota_x \triangleq \arg\min_{\iota=0,1,\dots,2^b-1}\left|\iota\frac{2\pi}{2^b}-x\right|, \tag{9}$$

which can be readily found as $\iota_x\in\{\iota,\iota+1\}$ for $x\in[\iota\frac{2\pi}{2^b},(\iota+1)\frac{2\pi}{2^b}]$. For $b=\infty$, we have

$$x = \lfloor x\rceil_\infty. \tag{10}$$

Upon denoting the baseband signal by $x\in\mathbb{C}^{N_{RF}}$, the received signal of at UE $k$ becomes

$$y_k = H_k\boldsymbol{\Theta}x+\nu_k, \tag{11}$$

where $\nu_k$ is noise having the covariance of $\sigma$, which incorporates both the background noise and other uncertainties, such as the channel estimation error. The development of robust designs that rely on imperfect channel state estimation is an interesting topic for future research.

Let $s_k\in\mathcal{CN}_{N_R}(0)$ with $\mathbb{E}([s_k]^2)=I_{N_R}$ be the information intended for UE $k$, which is processed by a MIMO beamformer $\mathbf{V}_k^B\in\mathbb{C}^{N_{RF}\times N_R}$ before the BS's transmission. For $s\triangleq(s_1^T,\dots,s_K^T)^T$ and

$$\mathbf{V}^B = \begin{bmatrix}\mathbf{V}_1^B & \dots & \mathbf{V}_K^B\end{bmatrix}\in\mathbb{C}^{N_{RF}\times(KN_R)}, \tag{12}$$

[3]The following comment of an anonymous reviewer is gratefully acknowledged: The phase shifter's power consumption is determined by that of the input voltage biasing network as well as by the bandwidth over which the phase shift is desired.

which is termed as the baseband beamformer, the baseband signal $x$ in (11) becomes $x=\mathbf{V}^Bs$. Based on (11), the corresponding MIMO equation becomes:

$$y = H\boldsymbol{\Theta}\mathbf{V}^Bs+\nu, \tag{13}$$

where we have:

$$y\triangleq\begin{bmatrix}y_1\\ \dots\\ y_K\end{bmatrix}\in\mathbb{C}^{KN_R}, H\triangleq\begin{bmatrix}H_1\\ \dots\\ H_K\end{bmatrix}\in\mathbb{C}^{(KN_R)\times N},$$

$$\nu\triangleq\begin{bmatrix}\nu_1\\ \dots\\ \nu_K\end{bmatrix}\in\mathbb{C}^{KN_R}.$$

This section deals with the scenario of $N_{RF}>KN_R$, under which ZFB exists:

$$\mathbf{V}^B = \boldsymbol{\Theta}^HH^H([H\boldsymbol{\Theta}]^2)^{-1}\mathsf{diag}[\mathbf{P}_k]_{k\in\mathcal{K}}, \mathbf{P}_k\in\mathbb{C}^{N_R\times N_R}, \tag{14}$$

leading from (13) to

$$y = \mathsf{diag}[\mathbf{P}_ks_k]_{k\in\mathcal{K}}+\nu. \tag{15}$$

For $\mathbf{P}\triangleq\mathsf{diag}[\mathbf{P}_k]_{k\in\mathcal{K}}$, the rate of UE $k$ is

$$r_k(\mathbf{P}_k)\triangleq\ln\left|I_{N_R}+\frac{1}{\sigma}[\mathbf{P}_k]^2\right|, \tag{16}$$

while the transmit power is

$$\pi(\boldsymbol{\theta},\mathbf{P}) \triangleq ||\boldsymbol{\Theta}\mathbf{V}^B||^2 \tag{17}$$
$$= \langle H\boldsymbol{\Theta}\boldsymbol{\Theta}^H\boldsymbol{\Theta}\boldsymbol{\Theta}^HH^H\left[([H\boldsymbol{\Theta}]^2)^{-1}\mathsf{diag}[\mathbf{P}_k]_{k\in\mathcal{K}}\right]^2\rangle. \tag{18}$$

Given the power budget $P$, the power constraint is

$$\pi(\boldsymbol{\theta},\mathbf{P})\le P. \tag{19}$$

We consider the following problems:

$$\max_{\mathbf{P},\boldsymbol{\theta}}\ r_{GM}(\mathbf{P})\triangleq\left(\prod_{k\in\mathcal{K}}r_k(\mathbf{P}_k)\right)^{1/K}\quad\text{s.t.}\quad(7),(19), \tag{20}$$

and

$$\max_{\mathbf{P},\boldsymbol{\theta}}\ r_{MM}(\mathbf{P})\triangleq\min_{k\in\mathcal{K}}r_k(\mathbf{P}_k)\quad\text{s.t.}\quad(7),(19), \tag{21}$$

where (20) is a geometric mean rate (GM-rate) optimization problem having an objective function given by the GM of users' rates, while (21) is a max-min rate optimization problem. Our recent result [43] shows that GM-rate optimization helps to improve all users' rates in a very fair manner. This feature of

GM-rate maximization will be underlined in the simulations in Section V. Both (20) and (21) are computationally challenging optimization problems of mixed discrete continuous nature [36].

The next two subsections are devoted to the computational solution of the GM-rate maximization problem (20), while the last subsection dedicated to that of the max-min rate optimization problem (21).

### *A. ABF design by joint optimization*

To address the problem (20), we first approximate the function $\pi(\boldsymbol{\theta}, \mathbf{p})$ in (17) as follows:

$$\begin{aligned}\pi(\boldsymbol{\theta}, \mathbf{P}) &\approx N\langle H\boldsymbol{\Theta}\boldsymbol{\Theta}^H H^H \left[([H\boldsymbol{\Theta}]^2)^{-1}\mathsf{diag}[\mathbf{P}_k]_{k\in\mathcal{K}}\right]^2\rangle \quad (22)\\ &= N\langle([H\boldsymbol{\Theta}]^2)^{-1}, \mathsf{diag}\left[[\mathbf{P}_k]^2\right]_{k\in\mathcal{K}}\rangle, \quad (23)\end{aligned}$$

where for the approximation (22) we used [12] $\boldsymbol{\Theta}^H\boldsymbol{\Theta} \approx NI_{N_{RF}}$. The power constraint (19) is thus approximated by the following constraint

$$\langle([H\boldsymbol{\Theta}]^2)^{-1}, \mathsf{diag}\left[[\mathbf{P}_k]^2\right]_{k\in\mathcal{K}}\rangle \le P/N. \quad (24)$$

We thus consider the following approximation problem for (21):

$$\max_{\mathbf{P},\boldsymbol{\theta}} \ r_{GM}(\mathbf{P}) \quad \text{s.t.} \quad (7), (24). \quad (25)$$

To minimize the nonlinearity of the objective function in (25) as the GM of nonlinear functions, we use the following equivalent formulation of the max-min optimization[4]

$$\max_{\mathbf{P},\boldsymbol{\theta}} \min_{\prod_{k\in\mathcal{K}}\boldsymbol{\gamma}_k=1, \boldsymbol{\gamma}_k>0} [\sum_{k\in\mathcal{K}} \boldsymbol{\gamma}_k r_k(\mathbf{P}_k)] \quad \text{s.t.} \quad (7), (24). \quad (26)$$

In what follows, for nonnegative integer $\kappa$ we use the notations $\theta^{(\kappa)} \triangleq [\theta^{(\kappa)}_{n,j}]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}$, $\Theta^{(\kappa)} \triangleq [e^{\jmath\theta^{(\kappa)}_{n,j}}]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}$, and $P^{(\kappa)} = \text{diag}[P^{(\kappa)}_k]_{k\in\mathcal{K}}$.

After initialization by $(P^{(0)}, \theta^{(0)})$, for $\kappa = 0, 1, \dots$, we optimize $\boldsymbol{\gamma}$ to have

$$\gamma^{(\kappa)}_k = \frac{\max_{k'\in\mathcal{K}} r_{k'}(P^{(\kappa)}_{k'})}{r_k(P^{(\kappa)}_k)}, k \in \mathcal{K}. \quad (27)$$

We then iterate $(P^{(\kappa+1)}, \theta^{(\kappa+1)})$ at the $\kappa$-th round by solving the following problem

$$\max_{\mathbf{P},\boldsymbol{\theta}} \ r^{(\kappa)}_{GM}(\mathbf{P}) \triangleq \sum_{k\in\mathcal{K}} \gamma^{(\kappa)}_k r_k(\mathbf{P}_k) \quad \text{s.t.} \quad (7), (24). \quad (28)$$

*1) Alternating optimization in* $\mathbf{P}$*:* We will seek $P^{(\kappa+1)}$ such that

$$r^{(\kappa)}_{GM}(P^{(\kappa+1)}) > r^{(\kappa)}_{GM}(P^{(\kappa)}). \quad (29)$$

Applying the inequality (2) yields

$$\begin{aligned}r_k(\mathbf{P}_k) &\ge r^{(\kappa)}_k(\mathbf{P}_k) \quad (30)\\ &\triangleq a^{(\kappa)}_k + 2\Re\{\langle A^{(\kappa)}_k\mathbf{P}_k\rangle\} - \langle B^{(\kappa)}_k, [\mathbf{P}_k]^2\rangle, \quad (31)\end{aligned}$$

[4] $\frac{1}{K}[\sum_{k\in\mathcal{K}}\boldsymbol{\gamma}_k r_k(\mathbf{P}_k)] \ge [\prod_{k\in\mathcal{K}}\boldsymbol{\gamma}_k r_k(\mathbf{P}_k)]^{1/K} = [\prod_{k\in\mathcal{K}} r_k(\mathbf{P}_k)]^{1/K}$ with the equality sign at $\boldsymbol{\gamma}_1 r_1(\mathbf{P}_k) = \dots = \boldsymbol{\gamma}_K r_K(\mathbf{P}_K)$ according to Cauchy's inequality.

for

$$\begin{aligned}a^{(\kappa)}_k &\triangleq r_k(P^{(\kappa)}_k) - \tfrac{||P^{(\kappa)}_k||^2}{\sigma} - \sigma\langle B^{(\kappa)}_k\rangle, A^{(\kappa)}_k \triangleq \tfrac{1}{\sigma}(P^{(\kappa)}_k)^H,\\ B^{(\kappa)}_k &\triangleq \tfrac{1}{\sigma}I_{N_R} - \left([P^{(\kappa)}_k]^2 + \sigma I_{N_R}\right)^{-1}.\end{aligned} \quad (32)$$

Therefore, we have:

$$\begin{aligned}r^{(\kappa)}_{GM}(\mathbf{P}) &\ge \tilde{r}^{(\kappa)}_{GM}(\mathbf{P}) \quad (33)\\ &\triangleq \sum_{k\in\mathcal{K}} \gamma^{(\kappa)}_k r^{(\kappa)}_k(\mathbf{P}_k)\\ &= a^{(\kappa)} + 2\sum_{k\in\mathcal{K}}\gamma^{(\kappa)}_k\Re\{\langle A^{(\kappa)}_k\mathbf{P}_k\rangle\}\\ &\quad -\sum_{k\in\mathcal{K}}\langle\gamma^{(\kappa)}_k B^{(\kappa)}_k, [\mathbf{P}_k]^2\rangle, \quad (34)\end{aligned}$$

with $a^{(\kappa)} \triangleq \sum_{k\in\mathcal{K}}\gamma^{(\kappa)}_k a^{(\kappa)}_k$. Note that we have $r^{(\kappa)}_{GM}(P^{(\kappa)}) = \tilde{r}^{(\kappa)}_{GM}(P^{(\kappa)})$, so $\tilde{r}^{(\kappa)}_{GM}(\mathbf{P})$ provides a tight minorant for $r^{(\kappa)}_{GM}(\mathbf{P})$ [36].

We solve the following convex problem of minorant maximization to generate $P^{(\kappa+1)}$:

$$\max_{\mathbf{P}} \ \tilde{r}^{(\kappa)}_{GM}(\mathbf{P}) \quad \text{s.t.} \quad \langle([H\Theta^{(\kappa)}]^2)^{-1}, [\mathbf{P}]^2\rangle \le P/N. \quad (35)$$

Assuming that $\mathcal{R}^{(\kappa)}_k$ are diagonal blocks of size $N_R \times N_R$ of the positive definite matrix $([H\Theta^{(\kappa)}]^2)^{-1}$, which are positive definite too [44]. Then (35) admits the following closed-form solution

$$P^{(\kappa+1)}_k = \begin{cases} (B^{(\kappa)}_k)^{-1}(A^{(\kappa)}_k)^H \\ \text{if } \sum_{k\in\mathcal{K}}||(\mathcal{R}^{(\kappa)}_k)^{1/2}(B^{(\kappa)}_k)^{-1}(A^{(\kappa)}_k)^H||^2 \le P/N, \\ \left(\gamma^{(\kappa)}_k B^{(\kappa)}_k + \mu\mathcal{R}^{(\kappa)}_k\right)^{-1}\gamma^{(\kappa)}_k(A^{(\kappa)}_k)^H \text{otherwise,}\end{cases}$$

where $\mu > 0$ is found by bisection such that $\sum_{k\in\mathcal{K}}||(\mathcal{R}^{(\kappa)}_k)^{1/2}\left(\gamma^{(\kappa)}_k B^{(\kappa)}_k + \mu\mathcal{R}^{(\kappa)}_k\right)^{-1}\gamma^{(\kappa)}_k(A^{(\kappa)}_k)^H||^2 = P/N$.

It follows from (33) that $r^{(\kappa)}_{GM}(P^{(\kappa+1)}) \ge \tilde{r}^{(\kappa)}_{GM}(P^{(\kappa+1)})$ and then $\tilde{r}^{(\kappa)}_{GM}(P^{(\kappa+1)}) > \tilde{r}^{(\kappa)}_{GM}(P^{(\kappa)}) = r^{(\kappa)}_{GM}(P^{(\kappa)})$ because $P^{(\kappa+1)}$ and $P^{(\kappa)}$ are the optimal solution and a feasible point for the problem (35). Therefore, (29) is verified, provided that $r^{(\kappa)}_{GM}(P^{(\kappa+1)}) \neq r^{(\kappa)}_{GM}(P^{(\kappa)})$.

*2) Alternating optimization in* $\boldsymbol{\theta}$*:* Note that the objective function in (28) is independent of $\boldsymbol{\theta}$, hence the alternating optimization in $\boldsymbol{\theta}$ aims for minimizing the power consumption defined by the right hand side (RHS) of the power constraint (24):

$$\min_{\boldsymbol{\theta}} \ \langle[P^{(\kappa+1)}]^2, ([H\boldsymbol{\Theta}]^2)^{-1}\rangle \quad \text{s.t.} \quad (7), \quad (36)$$

which is a complex problem of discrete optimization. We elaborate further as follows:

$$\begin{aligned}&\langle[P^{(\kappa+1)}]^2, ([H\boldsymbol{\Theta}]^2)^{-1}\rangle\\ =\ &\alpha\langle[\boldsymbol{\Theta}]^2\rangle - \left(\alpha\langle[\boldsymbol{\Theta}]^2\rangle - \langle[P^{(\kappa+1)}]^2, ([H\boldsymbol{\Theta}]^2)^{-1}\rangle\right)\\ =\ &\alpha NN_{RF} - \varphi(\boldsymbol{\theta}),\end{aligned}$$

where $\alpha > 0$ is chosen for ensuring that the function $\varphi(\boldsymbol{\theta}) \triangleq \alpha\langle[\boldsymbol{\Theta}]^2\rangle - \langle[P^{(\kappa+1)}]^2, ([H\boldsymbol{\Theta}]^2)^{-1}\rangle$ is convex in $\boldsymbol{\Theta}$ [36, Prop. 4.2]. The problem (36) is equivalent to the following problem

$$\max_{\boldsymbol{\theta}} \ \varphi(\boldsymbol{\theta}) \quad \text{s.t.} \quad (7). \quad (37)$$

We now derive a closed-form for the Frank-and-Wolf iteration (FWI) for concave programming [45]–[47]. In Appendix A, we show that

$$\begin{aligned}\varphi(\boldsymbol{\theta}) &\geq 2\sum_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}\varphi_{n,j}^{(\kappa)}(\boldsymbol{\theta}_{n,j})-a^{(\kappa)} &(38)\\ &\triangleq \varphi^{(\kappa)}(\boldsymbol{\theta}), &(39)\end{aligned}$$

for

$$\begin{gathered}a^{(\kappa)}\triangleq\alpha NN_{RF}+3\langle[P^{(\kappa+1)}]^2,([H\Theta^{(\kappa)}]^2)^{-1}\rangle,\\ B^{(\kappa)}\triangleq(\Theta^{(\kappa)})^H H^H([H\Theta^{(\kappa)}]^2)^{-1}[P^{(\kappa+1)}]^2([H\Theta^{(\kappa)}]^2)^{-1}H\\ B^{(\kappa)}\in\mathbb{C}^{N_{RF}\times N},\end{gathered} \quad (40)$$

and

$$\begin{aligned}\varphi_{n,j}^{(\kappa)}(\boldsymbol{\theta}_{n,j}) \triangleq\ & \cos\left(\angle\left(\alpha e^{-\jmath\theta_{n,j}^{(\kappa)}}+B^{(\kappa)}(j,n)\right)+\boldsymbol{\theta}_{n,j}\right)\\ & \left|\alpha e^{-\jmath\theta_{n,j}^{(\kappa)}}+B^{(\kappa)}(j,n)\right|, \in\mathcal{N}\times\mathcal{N}_{RF}.\end{aligned}$$

Moreover, $\varphi^{(\kappa)}$ is a tight minorant [36] of $\varphi$ because $\varphi(\theta^{(\kappa)})=\varphi^{(\kappa)}(\theta^{(\kappa)})$.

The FWI generates $\theta^{(\kappa+1)}$ by solving the following problem of discrete optimization

$$\max_{\boldsymbol{\theta}}\ \varphi^{(\kappa)}(\boldsymbol{\theta})\quad\text{s.t.}\quad(7), \quad (41)$$

which is losslessly decomposed into $NN_{RF}$ independent subproblems

$$\max_{\boldsymbol{\theta}_{n,j}}\ \varphi_{n,j}^{(\kappa)}(\boldsymbol{\theta}_{n,j})\quad\text{s.t.}\quad(7), \quad (42)$$

Each subproblem (42) admits the following closed-form solution

$$\theta_{n,j}^{(\kappa+1)}=[2\pi-\lfloor\angle\left(\alpha e^{-\jmath\theta_{n,j}^{(\kappa)}}+B^{(\kappa)}(j,n)\right)\rfloor_b],(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}. \quad (43)$$

For $\theta^{(\kappa+1)}\triangleq[\theta_{n,j}^{(\kappa+1)}]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}$, it follows from (39) that $\varphi(\theta^{(\kappa+1)})\geq\varphi^{(\kappa)}(\theta^{(\kappa+1)})$ and moreover $\varphi^{(\kappa)}(\theta^{(\kappa+1)})>\varphi^{(\kappa)}(\theta^{(\kappa)})=\varphi^{(\kappa)}(\theta^{(\kappa)})$, because the former and the latter are the optimal value and a feasible value for (41). We thus have

$$\varphi(\theta^{(\kappa+1)})>\varphi(\theta^{(\kappa)}), \quad (44)$$

whenever $\varphi(\theta^{(\kappa+1)})\neq\varphi(\theta^{(\kappa)})$, i.e. $\theta^{(\kappa+1)}$ is a better feasible point than $\theta^{(\kappa)}$ for the problem (37). As such, Algorithm 1 generates a sequence $\{\theta^{(\kappa)}\}$ of improved feasible points for the discrete set defined by (7) and converges after a finite number of iterations. As the computational complexity of (36) and (43) increases linearly with $K$ and $N_{RF}$, Algorithm 1 provides scalable-complex iterations for the computational solution of (25).

---

**Algorithm 1** Scalable-complex iterations for AFB

---

1: **Initialization:** Initialize a feasible $(P^{(0)},\theta^{(0)})$. Set $\kappa=0$.
2: **Repeat until convergence of** $\theta^{(\kappa)}$**:** Generate $P^{(\kappa+1)}$ by (36) and $\theta^{(\kappa+1)}$ by FWI (43). Reset $\kappa:=\kappa+1$.
3: **Output** $\theta^{opt}=\theta^{(\kappa)}$ and $P^{(\kappa)}$.

---

### B. MIMO ZFB design for GM-rate maximization

As stated above, the problem (25) is only an approximation of the problem (20), where the power constraint (19) in (20) is approximated by the constraint (24). Of course, we can scale the DBF solution of (25) to satisfy the constraint (19). Following [48], we can achieve a much better GM-rate by solving the following optimal baseband beamformers problem

$$\max_{\mathbf{P}}\ r_{GM}(\mathbf{P})\quad\text{s.t.}\quad\sum_{k\in\mathcal{K}}\langle\mathcal{R}_k,[\mathbf{P}_k]^2\rangle\leq P, \quad (45)$$

where $\theta^{opt}$ is found from Algorithm 1, while $\mathcal{R}_k$ are diagonal blocks of size $N_R$ of the matrix $\left[([H\Theta^{opt}]^2)^{-1}H\Theta^{opt}(\Theta^{opt})^H\right]^2$. It should be noted that unlike (24), which is an approximated power constraint, (45) provides the exact power constraint.

Let $P^{(\kappa)}$ be a feasible point for (45) that is found from the $(\kappa-1)$-st iteration and then $r_k^{(\kappa)}(\mathbf{P}_k)$ and $\tilde{r}_{GM}^{(\kappa)}(\mathbf{P})$ are defined from (31) and (33). We solve the following convex problem to generate $P^{(\kappa+1)}$:

$$\max_{\mathbf{P}}\ \tilde{r}_{GM}^{(\kappa)}(\mathbf{P})\quad\text{s.t.}\quad\sum_{k\in\mathcal{K}}\langle\mathcal{R}_k,[\mathbf{P}_k]^2\rangle\leq P, \quad (46)$$

which admits the closed-form solution of

$$P_k^{(\kappa+1)}=\begin{cases}(B_k^{(\kappa)})^{-1}(A_k^{(\kappa)})^H\\ \text{if }\sum_{k\in\mathcal{K}}||(\mathcal{R}_k)^{1/2}(B_k^{(\kappa)})^{-1}(A_k^{(\kappa)})^H||^2\leq P,\\ \left(\gamma_k^{(\kappa)}B_k^{(\kappa)}+\mu\mathcal{R}_k\right)^{-1}\gamma_k^{(\kappa)}(A_k^{(\kappa)})^H\text{otherwise,}\end{cases}$$

where $\mu>0$ is found by bisection such that $\sum_{k\in\mathcal{K}}||(\mathcal{R}_k)^{1/2}\left(\gamma_k^{(\kappa)}B_k^{(\kappa)}+\mu\mathcal{R}_k\right)^{-1}\gamma_k^{(\kappa)}(A_k^{(\kappa)})^H||^2=P$. The computational complexity of (47) is linear in $K$. Moreover, it can be readily shown that (29) holds, so Algorithm 2 below provides scalable-complex iterations for designing ZFB by maximizing the GM-rate.

---

**Algorithm 2** Scalable-complex iterations for MIMO ZFB for maximizing the GM-rate.

---

1: **Initialization:** Scale $P^{(\kappa)}$ found by Algorithm 1 to satisfy the power constraint in (45) and reset it as the initial point $P^{(0)}$. Set $\kappa=0$.
2: **Repeat until convergence of** $P^{(\kappa)}$**:** Generate $P^{(\kappa+1)}$ by (47). Reset $\kappa:=\kappa+1$.
3: **Output** $P^{opt}=P^{(\kappa)}$.

---

In summary, the joint design of ABF and DBF to maximize the GM-rate in (20) consists of two steps:

- Step 1: implement Algorithm 1 for solving the approximation problem (25) to decide the optimal ABF. Also use its optimal ZFB to generate an initial point for Step 2.
- Step 2: implement Algorithm 2 for solving the problem (45) to decide the optimal ZFB.

### C. Max-min rate MIMO ZFB

The two previous subsections have addressed the problem (20) of GM-rate maximization. By contrast, this subsection addresses the problem (21) of max-min rate optimization. It is plausible that the objective function in (21) is maximized at $\mathbf{P}_k \equiv \mathbf{p}_0 I_{N_R}$, $k \in \mathcal{K}$ so all users' rates are $N_R \ln(1+\mathbf{p}_0^2/\sigma)$. The approximated power constraint (24) becomes $\mathbf{p}_0^2\langle([H\mathbf{\Theta}]^2)^{-1}\rangle \leq P/N$. The users' rates are all equal, which are explicitly expressed as

$$N_R \ln(1+\frac{P}{N\sigma\langle([H\mathbf{\Theta}]^2)^{-1}\rangle}). \tag{47}$$

Maximizing (47) is losslessly reduced to

$$\min_{\boldsymbol{\theta}} \ \langle([H\mathbf{\Theta}]^2)^{-1}\rangle \quad \text{s.t.} \quad (7), \tag{48}$$

which is a particular case of (36) for $P^{(\kappa+1)} \equiv I_{KN_R}$, and it is equivalent to

$$\max_{\boldsymbol{\theta}} \ \varphi_{MM}(\boldsymbol{\theta}) \triangleq \alpha\langle[\mathbf{\Theta}]^2\rangle - \langle([H\mathbf{\Theta}]^2)^{-1}\rangle \quad \text{s.t.} \quad (7), \tag{49}$$

where $\alpha > 0$ is chosen such that $\varphi_{MM}(\boldsymbol{\theta})$ is convex. Similar to (39), we have:

$$\begin{aligned}\varphi_{MM}(\boldsymbol{\theta}) &\geq \varphi_{MM}^{(\kappa)}(\boldsymbol{\theta}) \\ &\triangleq 2\Re\{\langle\mathbf{\Theta}(\Theta^{(\kappa)})^H\rangle\} - a^{(\kappa)} + 2\Re\{\langle B^{(\kappa)}\mathbf{\Theta}\rangle\}\end{aligned} \tag{50}$$

for

$$\begin{aligned} a^{(\kappa)} &\triangleq \alpha N N_{RF} + 3\langle([H\Theta^{(\kappa)}]^2)^{-1}\rangle, \\ B^{(\kappa)} &\triangleq (\Theta^{(\kappa)})^H H^H([H\Theta^{(\kappa)}]^2)^{-2}H \in \mathbb{C}^{N_{RF}\times N}. \end{aligned} \tag{51}$$

Initialized by $\theta^{(0)}$ feasible for (7), the FWI at the $\kappa$-th iteration for $\kappa = 0, 1, \ldots$, generates $\theta^{(\kappa+1)}$ by solving the following problem of discrete optimization:

$$\max_{\boldsymbol{\theta}} \ \varphi_{MM}^{(\kappa)}(\boldsymbol{\theta}) \quad \text{s.t.} \quad (7), \tag{52}$$

which like (41) admits the closed-form solution

$$\theta^{(\kappa+1)} = [2\pi - \lfloor\angle\left(\alpha e^{-\jmath\theta_{n,j}^{(\kappa)}} + B^{(\kappa)}(j,n)\right)\rfloor_b]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}. \tag{53}$$

Similarly to (44), it can be shown that $\varphi_{MM}(\theta^{(\kappa+1)}) > \varphi_{MM}(\theta^{(\kappa)})$ as far as $\varphi_{MM}(\theta^{(\kappa+1)}) \neq \varphi_{MM}(\theta^{(\kappa)})$, so Algorithm 3 provides scalable-complex ascent iterations for maximizing the users' rate defined by (47).

**Algorithm 3** Max-min rate MIMO ZFB scalable algorithm

1: **Initialization:** Initialize $\theta^{(0)}$. Set $\kappa = 0$.
2: **Repeat until convergence of $\theta^{(\kappa)}$:** Generate $\theta^{(\kappa+1)}$ by FWI (53). Reset $\kappa := \kappa + 1$.
3: **Output** $\theta^{opt} = \theta^{(\kappa)}$ and $p_k^{opt} \equiv \sqrt{P/\langle[H\Theta^{(\kappa)}(\Theta^{opt})^H]^2\,[([H\Theta^{opt}]^2)^{-1}]^2\rangle}$, $k \in \mathcal{K}$, verifying the power constraint (19).

## III. The Brunn-Minkowski geometry for the Joint design of ABF and MIMO RZFB

When $KN_R \geq N_{RF}$ the matrix $[H\mathbf{\Theta}]^2$ is singular, so the ZFB defined by (14) does not exist. We thus use the RZFB [5] formulated as

$$\mathbf{V}^B = \mathbf{\Theta}^H H^H\left([H\mathbf{\Theta}]^2 + \alpha I_{KN_R}\right)^{-1}\mathsf{diag}[\mathbf{P}_k]_{k\in\mathcal{K}},$$

$$\mathbf{P}_k \in \mathbb{C}^{N_R\times N_R}, \tag{54}$$

in (12), for $\alpha = N_{RF}\sigma/P$, leading the MIMO equation (13) to

$$\begin{aligned} y &= H\mathbf{\Theta}\mathbf{\Theta}^H H^H\left([H\mathbf{\Theta}]^2 + \alpha I_{KN_R}\right)^{-1}\sum_{k\in\mathcal{K}}\mathbf{P}_k s_k + n \qquad (55)\\ &= \Psi(\boldsymbol{\theta})\sum_{k\in\mathcal{K}}\mathbf{P}_k s_k + n, \qquad (56)\end{aligned}$$

for

$$\Psi(\boldsymbol{\theta}) \triangleq I_{KN_R} - \alpha\left([H\mathbf{\Theta}]^2 + \alpha I_{KN_R}\right)^{-1}. \tag{57}$$

It is important to observe that $[H\mathbf{\Theta}]^2 + \alpha I_{KN_R} \succeq \alpha I_{KN_R} \succ 0$, so $\left([H\mathbf{\Theta}]^2 + \alpha I_{KN_R}\right)^{-1} \succ 0$ [44], and consequently

$$\Psi(\boldsymbol{\theta}) \prec I_{KN_R}. \tag{58}$$

The performance of RZFB is thus critically dependent on the matrix $\Psi(\boldsymbol{\theta})$ in (56): the closer and sparser $\Psi(\boldsymbol{\theta})$ approaches to the identity matrix $I_{KN_R}$, the more efficiently RZFB regularizes the multi-user interference. Moreover, $\Psi(\boldsymbol{\theta}) \succeq 0$ and $\Phi(\boldsymbol{\theta}) \triangleq I_{KN_R} - \Psi(\boldsymbol{\theta}) \succ 0$, so we can explore the Brunn-Minkowski geometry [50] of positive definite matrices for gauging the closeness of $\Psi(\boldsymbol{\theta})$ to $I_{KN_R}$ and its sparseness via the closeness of $\Phi(\boldsymbol{\theta})$ to the zero matrix and its sparseness. Thanks to the matrix inequality (58), both the $\ell_1$-distance and the Bures-Wasserstein distance [51] between $\Psi(\boldsymbol{\theta})$ and $I_{KN_R}$ is the $\ell_1$-norm of $\Phi(\boldsymbol{\theta})$, which is simply denoted by $\langle\Phi(\boldsymbol{\theta})\rangle$. Accordingly, we consider either the problem of minimizing the $\ell_1$-distance/Bures-Wasserstein distance between $\Psi(\boldsymbol{\theta})$ and $I_{KN_R}$:

$$\min_{\boldsymbol{\theta}}\langle\Phi(\boldsymbol{\theta})\rangle \Leftrightarrow \min_{\boldsymbol{\theta}} f(\boldsymbol{\theta}) \triangleq \langle\left([H\mathbf{\Theta}]^2 + \alpha I_{KN_R}\right)^{-1}\rangle, \tag{59}$$

or the problem of minimizing the volume of $\Phi(\boldsymbol{\theta})$ [50]:

$$\min_{\boldsymbol{\theta}} |\Phi(\boldsymbol{\theta})| \Leftrightarrow \max_{\boldsymbol{\theta}} f(\boldsymbol{\theta}) \triangleq \ln\left|I_{KN_R} + \frac{1}{\alpha}[H\mathbf{\Theta}]^2\right|, \tag{60}$$

subject to the constraint (7) of b-bit resolution, both of which also automatically promote the sparseness of $\Phi(\boldsymbol{\theta})$ (and $\Psi(\boldsymbol{\theta})$).

The next two subsections are devoted to the computation of (59) and (60).

### A. Inverse matrix trace minimization based ABF design

Let us now aim for computing (59) subject to (7), which is equivalent to

$$\max_{\boldsymbol{\theta}} \ f_t(\boldsymbol{\theta}) \triangleq t\langle[\mathbf{\Theta}]^2\rangle - \langle([H\mathbf{\Theta}]^2 + \alpha I_{KN_R})^{-1}\rangle \quad \text{s.t.} \quad (7), \tag{61}$$

where $t > 0$ is chosen such that $f_t(\boldsymbol{\theta})$ is convex. Similar to (39):

$$\begin{aligned} f_t(\boldsymbol{\theta}) &\geq f_t^{(\kappa)}(\boldsymbol{\theta}) \\ &\triangleq 2t\Re\{\langle\mathbf{\Theta}(\Theta^{(\kappa)})^H\rangle\} - a^{(\kappa)} + 2\Re\{\langle B^{(\kappa)}\mathbf{\Theta}\rangle\}\end{aligned} \tag{62}$$

for

$$\begin{aligned} A^{(\kappa)} &\triangleq [H\Theta^{(\kappa)}]^2 + \alpha I_{KN_R})^{-1}, \\ a^{(\kappa)} &\triangleq tNN_{RF} + 3\langle A^{(\kappa)}\rangle - 2\alpha\langle[A^{(\kappa)}]^2\rangle, \\ B^{(\kappa)} &\triangleq (\Theta^{(\kappa)})^H[H^H A^{(\kappa)}]^2 \in \mathbb{C}^{N_{RF}\times N}. \end{aligned} \tag{63}$$

[5] In many existing contributions such as [49] ZFB was used for $KN_R = N_{RF}$, but that is not correct because in fact $[H\mathbf{\Theta}]^2$ is often singular.

Initialized by $\theta^{(0)}$ feasible for (7), the FWI generates $\theta^{(\kappa+1)}$ by solving the following problem of discrete optimization

$$\max_{\boldsymbol{\theta}} \ f_t^{(\kappa)}(\boldsymbol{\theta}) \quad \text{s.t.} \quad (7), \tag{64}$$

which like (41) admits the closed form solution

$$\theta^{(\kappa+1)} = [2\pi - \lfloor \angle \left( te^{-\jmath\theta^{(\kappa)}_{n,j}} + B^{(\kappa)}(j,n) \right) \rfloor_b]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}. \tag{65}$$

The computational complexity of (65) is on the order of $\mathcal{O}(NN_{RF})$, i.e. it is linearly scalable in $NN_{RF}$. Like (44), it can be shown that

$$f(\theta^{(\kappa+1)}) < f(\theta^{(\kappa)}), \tag{66}$$

as far as $f(\theta^{(\kappa+1)}) \neq f(\theta^{(\kappa)})$. Algorithm 4 provides scalable-complex iterations for computing (59).

**Algorithm 4** Inverse matrix trace minimization scalable-complex FWI

1: **Initialization:** Initialize a feasible $\theta^{(0)}$ for (7). Set $\kappa = 0$.
2: **Repeat until convergence of** $\theta^{(\kappa)}$**:** Generate $\theta^{(\kappa+1)}$ by FWI (65). Reset $\kappa := \kappa + 1$.
3: **Output** $\theta^{opt} = \theta^{(\kappa)}$.

### *B. Log determinant maximization based ABF design*

Let us now aim for computing (60) subject to (7). In Appendix B, we show that

$$\begin{aligned} f(\boldsymbol{\theta}) &\geq a^{(\kappa)} + \frac{1}{\alpha}\left[2\Re\{\langle A^{(\kappa)}\boldsymbol{\Theta}\rangle\} + 2\Re\{\langle C^{(\kappa)}\boldsymbol{\Theta}\rangle\}\right] &(67)\\ &\triangleq f^{(\kappa)}(\boldsymbol{\theta}), &(68)\end{aligned}$$

for

$$\begin{array}{c}\tilde{a}^{(\kappa)} \triangleq f(\theta^{(\kappa)}) - \frac{1}{\alpha}\langle [H\Theta^{(\kappa)}]^2\rangle - \langle \left(\frac{1}{\alpha}[H\Theta^{(\kappa)}]^2 + I_{N_RK}\right)^{-1}\rangle,\\ A^{(\kappa)} \triangleq (\Theta^{(\kappa)})^H H^H H \in \mathbb{C}^{N_{RF}\times N},\\ B^{(\kappa)} \triangleq H^H\left[I_{N_RK} - \left(\frac{1}{\alpha}[H\Theta^{(\kappa)}]^2 + I_{N_RK}\right)^{-1}\right]H,\end{array} \tag{69}$$

and

$$\begin{array}{c}a^{(\kappa)} \triangleq \tilde{a}^{(\kappa)} - \frac{1}{\alpha}(2\lambda_{\max}(B^{(\kappa)})NN_{RF} - \langle B^{(\kappa)}, [\Theta^{(\kappa)}]^2\rangle,\\ C^{(\kappa)} \triangleq (\Theta^{(\kappa)})^H\left(\lambda_{\max}(B^{(\kappa)})I_N - B^{(\kappa)}\right).\end{array} \tag{70}$$

The FWI generates $\theta^{(\kappa+1)}$ by solving the following problem of discrete optimization

$$\max_{\boldsymbol{\theta}} \ f^{(\kappa)}(\boldsymbol{\theta}) \quad \text{s.t.} \quad (7), \tag{71}$$

which like (41) admits the following closed-form solution

$$\theta^{(\kappa+1)} = [2\pi - \lfloor \angle \left( A^{(\kappa)}(j,n) + C^{(\kappa)}(j,n) \right) \rfloor_b]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}. \tag{72}$$

The computational complexity of (72) is $\mathcal{O}(NN_{RF})$, i.e. it is also linearly scalable in $NN_{RF}$. Similarly to (44), we can show (66) as far as $f(\theta^{(\kappa+1)}) \neq f(\theta^{(\kappa)})$. Algorithm 5 provides scalable-complex iterations for computing (60) subject to (7).

**Algorithm 5** Log determinant maximization scalable-complex FWI

1: **Initialization:** Initialize a feasible $\theta^{(0)}$ for (7). Set $\kappa = 0$.
2: **Repeat until convergence of** $\theta^{(\kappa)}$**:** Generate $\theta^{(\kappa+1)}$ by FWI (72). Reset $\kappa := \kappa + 1$.
3: **Output** $\theta^{opt} = \theta^{(\kappa)}$.

Initial $\theta^{(0)}$ for Algorithms 4 and 5 when $N_R = 1$ and $K = N_{RF}$ is chosen according to [31], [52] extended to $b$-bit as

$$\theta^0 = [2\pi - \lfloor \angle H_j(n) \rfloor_b]_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}}, \tag{73}$$

where $H_j$ is the channel defined from (4) and $H_j(n)$ is its $n$-th entry. The rationale of this choice is to maximize $|H_j\boldsymbol{\Theta}_{:,j}|$.

### *C. MIMO RZFB design*

Based on the ABF designed in the previous subsections, this subsection addresses the design of RZFB. Having obtained $\theta^{opt}$ by Algorithm 4 or Algorithm 5, for $\beta_k \triangleq ||H_k\Theta^{opt}||^2$, and then $H_k^B \triangleq H_k\Theta^{opt}/\sqrt{\beta_k} \in \mathbb{C}^{N_R\times N_{RF}}$, $k \in \mathcal{K}$, and

$$H_B = \begin{bmatrix} H_B^1 \\ \dots \\ H_B^K \end{bmatrix} \in \mathbb{C}^{(N_RK)\times N_{RF}},$$

we re-write (13) as

$$y = \mathsf{diag}[\sqrt{\beta_k}I_{N_R}]_{k\in\mathcal{K}}H_B\mathbf{V}^Bs + \nu. \tag{74}$$

We use the equality $H_B^H\left([H_B]^2 + \alpha I_{KN_R}\right)^{-1} = \left([H_B^H]^2 + \alpha I_{N_{RF}}\right)^{-1}H_B^H$, and employ the RZF beamformers formulated as:

$$\mathbf{V}_k^B \triangleq \left([H_B^H]^2 + \alpha I_{N_{RF}}\right)^{-1}(H_B^i)^H\mathbf{P}_k, k \in \mathcal{K}, \tag{75}$$

to write equation (55) of the signal received at UE $k$ in the form of:

$$y_k = \sqrt{\beta_k}\sum_{\ell\in\mathcal{K}}\Xi_{k,\ell}\mathbf{P}_\ell s_\ell + \nu_k, \tag{76}$$

for

$$\Xi_{k,\ell} \triangleq H_B^i\left([H_B^H]^2 + \alpha I_{N_{RF}}\right)^{-1}(H_B^\ell)^H \in \mathbb{C}^{N_R\times N_R}. \tag{77}$$

The rate of UE $k$ is formulated as:

$$r_k(\mathbf{P}) = \ln\left|I_{N_R} + [\Xi_{k,k}\mathbf{P}_k]^2\Gamma_k^{-1}(\mathbf{P})\right|, \tag{78}$$

where we have

$$\Gamma_k(\mathbf{P}) \triangleq \sum_{\ell\in\mathcal{K}\setminus\{k\}}[\Xi_{k,\ell}\mathbf{P}_\ell]^2 + (\sigma/\beta_k)I_{N_R}. \tag{79}$$

The transmit power $\mathbb{E}(||\sum_{k\in\mathcal{K}}\Theta^{opt}\mathbf{V}_k^Bs_k||^2)$ is expressed as

$$\sum_{k\in\mathcal{K}}||\Theta^{opt}\mathbf{V}_k^B||^2 = \sum_{k\in\mathcal{K}}\langle\mathcal{R}_k, [\mathbf{P}_k]^2\rangle, \tag{80}$$

with

$$\mathcal{R}_k \triangleq [H_B^i\left([H_B^H]^2 + \alpha I_{N_{RF}}\right)^{-1}V_{RF}^H(\theta^{opt})]^2 \succeq 0, k \in \mathcal{K}. \tag{81}$$

The problem of designing RZFB to maximize the GM-rate subject to transmit power constraints is formulated as

$$\max_{\mathbf{P}}\ f_{GM}(\mathbf{P}) \triangleq \left(\prod_{k\in\mathcal{K}} r_k(\mathbf{P})\right)^{1/K} \quad (82a)$$

$$\text{s.t.}\quad \sum_{k\in\mathcal{K}}\langle \mathcal{R}_k, [\mathbf{P}_k]^2\rangle \le P. \quad (82b)$$

Similarly to (28), a specific initialized by $P^{(0)}$ which is feasible for (82), for $\kappa = 0,1,\dots,$ we iterate $P^{(\kappa+1)}$ by solving the optimization problem

$$\max_{\mathbf{P}}\ f^{(\kappa)}(\mathbf{P}) \triangleq \sum_{k\in\mathcal{K}} \gamma_k^{(\kappa)} r_k(\mathbf{P}) \quad \text{s.t.} \quad (82b), \quad (83)$$

associated with

$$\gamma_k^{(\kappa)} = \frac{\max_{k'\in\mathcal{K}} r_{k'}(P^{(\kappa)})}{r_k(P^{(\kappa)})}, k\in\mathcal{K}. \quad (84)$$

Upon exploiting using the inequality (2), we arrive at:

$$\begin{aligned} r_k(\mathbf{P}) &\ge \tilde{a}_k^{(\kappa)} + 2\Re\{\langle A_k^{(\kappa)}\mathbf{P}_k\rangle\} \\ &\quad -\langle B_k^{(\kappa)}\left(\sum_{\ell\in\mathcal{K}}[\Xi_{k,\ell}\mathbf{P}_\ell]^2 + (\sigma/\beta_k)I_{N_R}\right)\rangle \quad (85)\\ &= a^{(\kappa)} + 2\Re\{\langle A_k^{(\kappa)}\mathbf{P}_k\rangle\} \\ &\quad -\sum_{\ell\in\mathcal{K}}\langle \Xi_{k,\ell}^H B_k^{(\kappa)}\Xi_{k,\ell}, [\mathbf{P}_\ell]^2\rangle \\ &\triangleq r_k^{(\kappa)}(\mathbf{P}), \quad (86) \end{aligned}$$

with

$$\begin{aligned} \tilde{a}_k^{(\kappa)} &\triangleq r_k(P^{(\kappa)}) - \langle [\Xi_{k,k}P_k^{(\kappa)}]^2\Gamma_k^{-1}(P^{(\kappa)})\rangle, \\ A_k^{(\kappa)} &\triangleq (P_k^{(\kappa)})^H\Xi_{k,k}^H\Gamma_k^{-1}(P^{(\kappa)})\Xi_{k,k}, \\ B_k^{(\kappa)} &\triangleq \Gamma_k^{-1}(P^{(\kappa)}) - \left(\Gamma_k(P^{(\kappa)}) + [\Xi_{k,k}P_k^{(\kappa)}]^2\right)^{-1}, \\ a_k^{(\kappa)} &\triangleq \tilde{a}_k^{(\kappa)} - (\sigma/\beta_k)\langle B_k^{(\kappa)}\rangle. \end{aligned} \quad (87)$$

Therefore,

$$\begin{aligned} f^{(\kappa)}(\mathbf{P}) &\ge \tilde{f}^{(\kappa)}(\mathbf{P}) \\ &\triangleq \sum_{k\in\mathcal{K}}\gamma_k^{(\kappa)} r_k^{(\kappa)}(\mathbf{P}) \\ &= a^{(\kappa)} + 2\sum_{k\in\mathcal{K}}\Re\{\langle\gamma_k^{(\kappa)}A_k^{(\kappa)}\mathbf{P}_k\rangle\} \\ &\quad -\sum_{k\in\mathcal{K}}\gamma_k^{(\kappa)}\sum_{\ell\in\mathcal{K}}\langle\Xi_{k,\ell}^H B_k^{(\kappa)}\Xi_{k,\ell}, [\mathbf{P}_\ell]^2\rangle \\ &= a^{(\kappa)} + 2\sum_{k\in\mathcal{K}}\Re\{\langle\gamma_k^{(\kappa)}A_k^{(\kappa)}\mathbf{P}_k\rangle\} \\ &\quad - \sum_{k\in\mathcal{K}}\langle\mathcal{Q}_k^{(\kappa)}, [\mathbf{P}_k]^2\rangle \quad (88) \end{aligned}$$

for

$$a^{(\kappa)} \triangleq \sum_{k\in\mathcal{K}} a_k^{(\kappa)}, \mathcal{Q}_k^{(\kappa)} \triangleq \sum_{\ell\in\mathcal{K}}\gamma_\ell^{(\kappa)}\Xi_{\ell,k}^H B_\ell^{(\kappa)}\Xi_{\ell,k}. \quad (89)$$

We solve the following problem of minorant maximization to generate $P^{(\kappa+1)}$

$$\max_{\mathbf{P}}\ \tilde{f}^{(\kappa)}(\mathbf{P}) \triangleq \quad \text{s.t.} \quad (82b), \quad (90)$$

which admits the following closed-form solution

$$P_k^{(\kappa+1)} = \begin{cases} \gamma_k^{(\kappa)}(\mathcal{Q}_k^{(\kappa)})^{-1}(A_k^{(\kappa)})^H \\ \text{if} \quad \sum_{k\in\mathcal{K}}||\mathcal{R}_k^{1/2}(\mathcal{Q}_k^{(\kappa)})^{-1}\gamma_k^{(\kappa)}(A_k^{(\kappa)})^H||^2 \le P \\ \left(\mathcal{Q}_k^{(\kappa)} + \mu\mathcal{R}_k\right)^{-1}\gamma_k^{(\kappa)}(A_k^{(\kappa)})^H \text{otherwise,} \end{cases} \quad (91)$$

where $\mu > 0$ is chosen for ensuring that $\sum_{k\in\mathcal{K}}||\mathcal{R}_k^{1/2}(\mathcal{Q}_k^{(\kappa)} + \mu\mathcal{R}_k)^{-1}\gamma_k^{(\kappa)}(A_k^{(\kappa)})^H||^2 = P$.

Similarly to (29), we can show that $f^{(\kappa)}(P^{(\kappa+1)}) > f^{(\kappa)}(P^{(\kappa)})$, provided that $f^{(\kappa)}(P^{(\kappa+1)}) \ne f^{(\kappa)}(P^{(\kappa)})$. Algorithm 6 provides scalable-complex iterations for solving the optimization the problem (82).

**Algorithm 6** GM-rate maximization based MIMO RZFB scalable algorithm

1: **Initialization:** Initialize a feasible $P_k^{(\kappa)}$ for the constraint (82b). Set $\kappa = 0$.
2: **Repeat until convergence of the objective function given by (82a):** Define $\gamma_k^{(\kappa)}$ according to (84) and then generate $P_k^{(\kappa+1)}$ by (91). Reset $\kappa := \kappa + 1$.
3: **Output** $P_k^{opt} = P_k^{(\kappa+1)}$.

## IV. NEW STRUCTURED MIMO RZFB

Under using the RZFB solution of (75), the transmit signal $x = \sum_{k\in\mathcal{K}}\Theta^{opt}\mathbf{V}_k^B s_k$ is proper Gaussian, since we have $\mathbb{E}(xx^T) = 0$. In this section, we propose the following new RZFB solution:

$$\left([H_B^H]^2 + \alpha I_{N_{RF}}\right)^{-1}(H_B^i)^H\left(\mathbf{P}_{k,1}s_k + \mathbf{P}_{k,2}s_k^*\right), \quad (92)$$

with $\mathbf{P}_{k,1}\in\mathbb{C}^{N_R\times N_R}$ and $\mathbf{P}_{k,2}\in\mathbb{C}^{N_R\times N_R}$, $k\in\mathcal{K}$. As a result, the transmit signal

$$x = \sum_{k\in\mathcal{K}}\Theta^{opt}\left([H_B^H]^2 + \alpha I_{N_{RF}}\right)^{-1}(H_B^i)^H\left(\mathbf{P}_{k,1}s_k + \mathbf{P}_{k,2}s_k^*\right) \quad (93)$$

is improper Gaussian, because we have $\mathbb{E}(xx^T) \ne 0$ [53]. The reader is referred e.g. to [43], [54]–[57] and references therein for characterizing the efficiency of improper Gaussian signaling in interference-limited networks.

Instead of (76), the signal received at UE $k$ now becomes:

$$y_k = \sqrt{\beta_k}\sum_{\ell\in\mathcal{K}}\Xi_{k,\ell}\left(\mathbf{P}_{\ell,1}s_\ell + \mathbf{P}_{\ell,2}s_\ell^*\right) + \nu_k, \quad (94)$$

where $\Xi_{k,\ell}$ is defined in (77).

For

$$\bar{y}_k \triangleq \begin{bmatrix}\Re\{y_k\}\\ \Im\{y_k\}\end{bmatrix}, \bar{s}_k \triangleq \begin{bmatrix}\Re\{s_k\}\\ \Im\{s_k\}\end{bmatrix}, \bar{\nu}_k \triangleq \begin{bmatrix}\Re\{\nu_k\}\\ \Im\{\nu_k\}\end{bmatrix},$$

$$\bar{\Xi}_{k,\ell} \triangleq \begin{bmatrix}\Re\{\Xi_{k,\ell}\} & -\Im\{\Xi_{k,\ell}\}\\ \Im\{\Xi_{k,\ell}\} & \Re\{\Xi_{k,\ell}\}\end{bmatrix},$$

the equivalent real composite form of (94) becomes (95). By making the variable change (96), we can represent (95) by

$$\bar{y}_k = \sqrt{\beta_k}\sum_{\ell\in\mathcal{K}}\bar{\Xi}_{k,\ell}\boldsymbol{X}_\ell\bar{s}_\ell + \bar{\nu}_k. \quad (97)$$

$$\bar{y}_k = \bar{\nu}_k + \sqrt{\beta_k}\sum_{\ell\in\mathcal{K}}\bar{\Xi}_{k,\ell}\begin{bmatrix}\Re\{\mathbf{P}_{\ell,1}\}+\Re\{\mathbf{P}_{\ell,2}\} & -\Im\{\mathbf{P}_{\ell,1}\}+\Im\{\mathbf{P}_{\ell,2}\}\\ \Im\{\mathbf{P}_{\ell,1}\}+\Im\{\mathbf{P}_{\ell,2}\} & \Re\{\mathbf{P}_{\ell,1}\}-\Re\{\mathbf{P}_{\ell,2}\}\end{bmatrix}\bar{s}_\ell. \quad (95)$$

$$\boldsymbol{X}_k = \begin{bmatrix}\boldsymbol{X}_k^{11} & \boldsymbol{X}_k^{12}\\ \boldsymbol{X}_k^{21} & \boldsymbol{X}_k^{22}\end{bmatrix} = \begin{bmatrix}\Re\{\mathbf{P}_{k,1}\}+\Re\{\mathbf{P}_{k,2}\} & -\Im\{\mathbf{P}_{k,1}\}+\Im\{\mathbf{P}_{k,2}\}\\ \Im\{\mathbf{P}_{k,1}\}+\Im\{\mathbf{P}_{k,2}\} & \Re\{\mathbf{P}_{k,1}\}-\Re\{\mathbf{P}_{k,2}\}\end{bmatrix}\in\mathbb{R}^{(2N_R)\times(2N_R)}, k\in\mathcal{K}, \quad (96)$$

For $\boldsymbol{X}\triangleq\{\boldsymbol{X}_k, k\in\mathcal{K}\}$, the UE $k$'s rate is $0.5\rho_k(\boldsymbol{X})$ [58] in conjunction with

$$\rho_k(\boldsymbol{X})\triangleq\ln\left|I_{2N_R}+[\bar{\Xi}_{k,k}\boldsymbol{X}_k]^2\bar{\Gamma}_k^{-1}(\boldsymbol{X})\right|, \quad (98)$$

where $\bar{\Gamma}_k\triangleq\sum_{\ell\in\mathcal{K}\setminus\{k\}}[\bar{\Xi}_{k,\ell}\boldsymbol{X}_\ell]^2+(\sigma/\beta_k)I_{2N_R}$.

Under the variable change (96), the real composite form of the transmit signal $x$ defined in (93) becomes (99). By noting that $E(\bar{s}_k\bar{s}_k^T)=0.5I_{2N_R}$, we have $E(||\bar{s}_k||^2)=0.5\langle(R_k^TR_k),[\boldsymbol{X}_k]^2\rangle$ and so the power constraint is formulated as:

$$\sum_{k\in\mathcal{K}}E(||x_k||^2)\le P \quad\Leftrightarrow\quad \sum_{k\in\mathcal{K}}\langle(R_k^TR_k),[\boldsymbol{X}_k]^2\rangle\le 2P. \quad (101)$$

The problem of GM-rate maximization under RZFB (92) can be formulated by[6]

$$\max_{\boldsymbol{X}}\ \bar{f}_{GM}(\boldsymbol{X})\triangleq\left(\prod_{k\in\mathcal{K}}\rho_k(\boldsymbol{X})\right)^{1/K}\quad \text{s.t.}\quad (101). \quad (102)$$

Similarly to (28), starting from a specific $X^{(0)}$ which is feasible for (101), for $\kappa=0,1,\dots$, we iterate $X^{(\kappa+1)}$ by addressing the problem

$$\max_{\boldsymbol{X}}\ f^{(\kappa)}(\boldsymbol{X})\triangleq\sum_{k\in\mathcal{K}}\gamma_k^{(\kappa)}\rho_k(\boldsymbol{X})\quad\text{s.t.}\quad(101), \quad (103)$$

for

$$\gamma_k^{(\kappa)}\triangleq\frac{\max_{k'\in\mathcal{K}}\rho_{k'}(X^{(\kappa)})}{\rho_k(X^{(\kappa)})}, k\in\mathcal{K}. \quad (104)$$

Upon exploiting the inequality (2), we arrive at:

$$\begin{aligned}\rho_k(\boldsymbol{X}) &\ge \tilde{a}_k^{(\kappa)}+2\Re\{\langle A_k^{(\kappa)}\boldsymbol{X}_k\rangle\}\\ &\quad -\langle B_k^{(\kappa)}\left(\sum_{\ell\in\mathcal{K}}[\bar{\Xi}_{k,\ell}\boldsymbol{X}_\ell]^2+(\sigma/\beta_k)I_{2N_R}\right)\rangle\\ &= a^{(\kappa)}+2\Re\{\langle A_k^{(\kappa)}\boldsymbol{X}_k\rangle\}\\ &\quad -\sum_{\ell\in\mathcal{K}}\langle\bar{\Xi}_{k,\ell}^TB_k^{(\kappa)}\bar{\Xi}_{k,\ell},[\boldsymbol{X}_\ell]^2\rangle\\ &\triangleq \rho_k^{(\kappa)}(\boldsymbol{X}),\end{aligned} \quad (105)$$

with

$$\begin{aligned}\tilde{a}_k^{(\kappa)}&\triangleq\rho_k(X^{(\kappa)})-\langle[\bar{\Xi}_{k,k}X_k^{(\kappa)}]^2\bar{\Gamma}_k^{-1}(X^{(\kappa)})\rangle,\\ A_k^{(\kappa)}&\triangleq(X_k^{(\kappa)})^T\bar{\Xi}_{k,k}^T\bar{\Gamma}_k^{-1}\ (X^{(\kappa)})\bar{\Xi}_{k,k},\\ B_k^{(\kappa)}&\triangleq\bar{\Gamma}_k^{-1}(X^{(\kappa)})-\left(\bar{\Gamma}_k(X^{(\kappa)})+[\bar{\Xi}_{k,k}X_k^{(\kappa)}]^2\right)^{-1},\\ a_k^{(\kappa)}&\triangleq\tilde{a}_k^{(\kappa)}-(\sigma/\beta_k)\langle B_k^{(\kappa)}\rangle.\end{aligned} \quad (106)$$

[6]The result must be divided by 2.

Therefore, we have:

$$\begin{aligned}f^{(\kappa)}(\boldsymbol{X}) &\ge \tilde{f}^{(\kappa)}(\boldsymbol{X})\\ &\triangleq \sum_{k\in\mathcal{K}}\gamma_k^{(\kappa)}\rho_k^{(\kappa)}(\boldsymbol{X})\\ &= a^{(\kappa)}+2\sum_{k\in\mathcal{K}}\Re\{\langle\gamma_k^{(\kappa)}A_k^{(\kappa)}\boldsymbol{X}_k\rangle\}\\ &\quad -\sum_{k\in\mathcal{K}}\gamma_k^{(\kappa)}\sum_{\ell\in\mathcal{K}}\langle\bar{\Xi}_{k,\ell}^TB_k^{(\kappa)}\bar{\Xi}_{k,\ell},[\boldsymbol{X}_\ell]^2\rangle\\ &= a^{(\kappa)}+2\sum_{k\in\mathcal{K}}\Re\{\langle\gamma_k^{(\kappa)}A_k^{(\kappa)}\boldsymbol{X}_k\rangle\}\\ &\quad -\sum_{k\in\mathcal{K}}\langle\bar{\mathcal{Q}}_k^{(\kappa)},[\boldsymbol{X}_k]^2\rangle,\end{aligned} \quad (107)$$

for

$$a^{(\kappa)}\triangleq\sum_{k\in\mathcal{K}}a_k^{(\kappa)}, \bar{\mathcal{Q}}_k^{(\kappa)}\triangleq\sum_{\ell\in\mathcal{K}}\gamma_\ell^{(\kappa)}\bar{\Xi}_{\ell,k}^TB_\ell^{(\kappa)}\bar{\Xi}_{\ell,k}. \quad (108)$$

We thus solve the following problem of minorant maximization for (103) to generate $X^{(\kappa+1)}$

$$\max_{\boldsymbol{X}}\ \tilde{f}^{(\kappa)}(\boldsymbol{X})\quad\text{s.t.}\quad(101), \quad (109)$$

which admits the following closed-form solution

$$X_k^{(\kappa+1)}=\begin{cases}\gamma_k^{(\kappa)}(\bar{\mathcal{Q}}_k^{(\kappa)})^{-1}(A_k^{(\kappa)})^T\\ \text{if}\quad\sum_{k\in\mathcal{K}}||R_k(\bar{\mathcal{Q}}_k^{(\kappa)})^{-1}\gamma_k^{(\kappa)}(A_k^{(\kappa)})^T||^2\le 2P\\ \left(\bar{\mathcal{Q}}_k^{(\kappa)}+\mu R_k^TR_k\right)^{-1}\gamma_k^{(\kappa)}(A_k^{(\kappa)})^T\text{otherwise},\end{cases} \quad (110)$$

where $\mu>0$ is chosen for ensuring that $\sum_{k\in\mathcal{K}}||R_k(\bar{\mathcal{Q}}_k^{(\kappa)}+\mu R_k^TR_k)^{-1}\gamma_k^{(\kappa)}(A_k^{(\kappa)})^T||^2=2P$.

Similarly to (29), we can show that $f^{(\kappa)}(X^{(\kappa+1)})>f^{(\kappa)}(X^{(\kappa)})$ provided that $f^{(\kappa)}(X^{(\kappa+1)})\ne f^{(\kappa)}(X^{(\kappa)})$. Algorithm 7 provides scalable-complex iterations for computing the problem (102). Then the optimal matrices $P_{k,1}^{opt}$ and $P_{k,2}^{opt}$ used for determing the RZFB (92) are recovered from the optimal solution $X_k^{opt}=\begin{bmatrix}X_k^{11,opt} & X_k^{12,opt}\\ X_k^{21,opt} & X_k^{22,opt}\end{bmatrix}$ of (102) as (111).

## V. Numerical Results

This section evaluates the numerical efficiency of the proposed algorithms. With the $K_N$ users randomly located within the cell radius of 50 meters, where the links between the BS and UEs are assumed to be LOS channels, the path-loss of UE $k_N$, $k_N\in\mathcal{K}_N\triangleq\{1,\dots,K_N\}$ experienced at a distance $d_{k_N}$ from the BS is set to $\rho_{k_N}=44.84+21\log 10(d_{K_N})$ dB, which takes into account a 16.5 dB gain due to beamforming-aided mmWave transmission [3], [6], [59], and the complex

$$\bar{x} \triangleq \begin{bmatrix} \Re\{x\} \\ \Im\{x\} \end{bmatrix} = \sum_{k\in\mathcal{K}} R_k \boldsymbol{X}_k \bar{s}_k, \tag{99}$$

for

$$R_k \triangleq \begin{bmatrix} \Re\{\Theta^{opt}\left([H_B^H]^2+\alpha I_{N_{RF}}\right)^{-1}(H_B^k)^H\} & -\Im\{\Theta^{opt}\left([H_B^H]^2+\alpha I_{N_{RF}}\right)^{-1}(H_B^k)^H\} \\ \Im\{\Theta^{opt}\left([H_B^H]^2+\alpha I_{N_{RF}}\right)^{-1}(H_B^k)^H\} & \Re\{\Theta^{opt}\left([H_B^H]^2+\alpha I_{N_{RF}}\right)^{-1}(H_B^k)^H\} \end{bmatrix}. \tag{100}$$

$$\begin{bmatrix} \Re\{P_{k,1}^{opt}\} & \Im\{P_{k,1}^{opt}\} \\ \Re\{P_{k,2}^{opt}\} & \Im\{P_{k,2}^{opt}\} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} X_k^{11,opt}+X_k^{22,opt} & X_k^{21,opt}-X_k^{12,opt} \\ X_k^{11,opt}-X_k^{22,opt} & X_k^{21,opt}+X_k^{12,opt} \end{bmatrix}, k\in\mathcal{K}. \tag{111}$$

**Algorithm 7** New structured MIMO RZFB optimization algorithm

1: **Initialization:** Initialize a feasible $X_k^{(\kappa)}$ for the constraint (101). Set $\kappa = 0$.
2: **Repeat until convergence of the objective function given by (102):** Define $\gamma_k^{(\kappa)}$ according to (84) and then generate $X_k^{(\kappa+1)}$ by (110). Reset $\kappa := \kappa + 1$.
3: **Output** $X_k^{opt} = X_k^{(\kappa+1)}$, $k \in \mathcal{K}$.

gain $\alpha_{k_N,c,\ell}$ follows the Ricean distribution with a K-factor of 10dB [60], [61]. Similarity, with $K_F$ users randomly located between 50 and 200 meters radius having NLOS environment, the path-loss of UE $k_F$, $k_F \in \mathcal{K}_F \triangleq \{K_N+1,\dots,K\}$ is set to $\rho_{K_F} = 36.72 + 35.3\log 10(d_{K_F})$ dB, and the complex gain $\alpha_{k_F,c,\ell}$ follows Rayleigh fading. The azimuth angle of departure (arrival, resp.) $\phi^t_{k,c,\ell}$ ($\phi^r_{k,c,\ell}$, resp.) and the elevation angle of departure $\theta^t_{k,c,\ell}$ are generated according to the Laplacian distribution in conjunction with random mean cluster angles in the interval $[0, 2\pi)$ and with spreads of 10 degrees within each cluster, while $N_c = 2$ and $N_{sc} = 3$ [59]. 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.

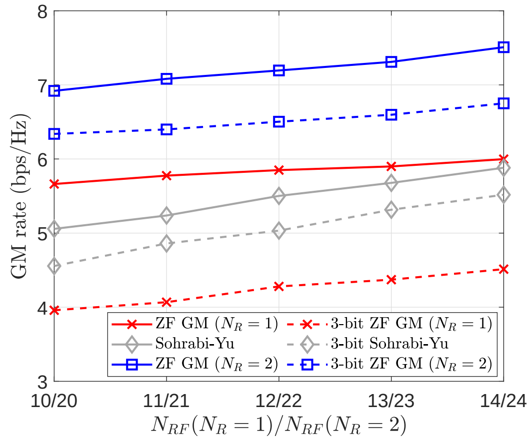
Unless otherwise stated, we have $K = 8$, $K_N = 3$, $N_{RF} = 8$, $N = 64$ ($N_1 = 8$, $N_2 = 8$), $P = 20$ dBm, and $b = 3$. The results are multiplied by $\log_2 e$ to convert the unit nats/sec into the unit bps/Hz. The convergence threshold of the proposed algorithms is set to $10^{-3}$.

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

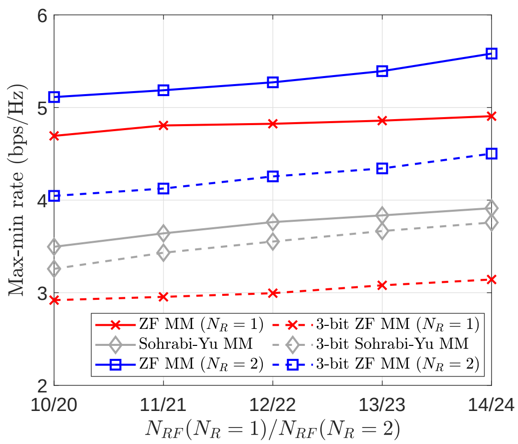
- ”ZF GM” and ”3-bit ZF GM” refer to the performance of the alternating optimization Algorithm 1 and Algorithm 2 for $b = \infty$ and $b = 3$, respectively;
- ”ZF MM” and ”3-bit ZF MM” refer to the performance of Algorithm 3 employed for for $b = \infty$ and $b = 3$, respectively;
- ”Sohrabi-Yu”/”Sohrabi-Yu MM” and ”3-bit Sohrabi-Yu”/”3-bit Sohrabi-Yu MM” refer to the performance of generating $P^{(\kappa+1)}$ by (36)/$P^{(\kappa+1)} \equiv I_K$, but generating $\theta^{(\kappa+1)}$ according to [28] by addressing (36) by alternating optimization in each $\boldsymbol{\theta}_{n,j}$ with other $\boldsymbol{\theta}_{n',j}$ held fixed, which needs the exhaustive search over $\mathcal{B}$ for $b = 3$. The final $P^{(\kappa)}$ must be scaled to satisfy the power constraint in (45);
- ”Trace-Max” and ”3-bit Trace-Max” refer to the results of the trace maximization based Algorithm 4 for for $b = \infty$ and $b = 3$, respectively, and then implementing the RZFB Algorithm 6;
- ”Log-det-Max” and ”3-bit Log-det-Max” refer to the performance of the log-det maximization based Algorithm 5 for $b = \infty$ and $b = 3$, respectively, and then harnessing the RZFB Algorithm 6;
- ”Nasir et al.” and ”3-bit Nasir et al.” refer to the results based on (73) for $b = \infty$ and $b = 3$, respectively, and then implementing the RZFB Algorithm 6;
- ”IGS Trace-Max” and ”3-bit IGS Trace-Max” refer to the results of the trace maximization based Algorithm 4 for $b = \infty$ and $b = 3$, respectively, and then implementing the new structured RZFB optimization Algorithm 7;
- ”IGS Log-det-Max” and ”3-bit IGS Log-det-Max” refer to the results of the trace maximization based Algorithm 5 for $b = \infty$ and $b = 3$, respectively, and then implementing the new structured RZFB optimization Algorithm 7.

Fig. 1 plots the achievable GM rate and max-min rate versus the number $N_{RF}$ in the ZFB based maximization, which shows that all the ZFB based algorithms outperform their 3-bit resolution counterpart. As expected, the ZFB maximization based algorithms achieve better GM rates, while ZFB max-min rate based algorithms achieve better max-min rates. Increasing the number of RF chains does not lead to a significant improvement, which is not unexpected for multi-user communications. Furthermore, all the algorithms benefit from the improved spatial diversity due to increasing the number of receive antennas.

Fig. 2 plots GM versus the number of RF chains at the UE, attained by RZFB based maximization. For $N_R = 1$, ”IGS Trace-Max” is the best performer and ”Trace-Max” performs better than ”Log-det-Max” does. Furthermore, ”IGS Log-det-Max” has better performance than ”Log-det-max”, but the gap becomes smaller with the increasing number of RF chains. For $N_R = 2$, ”IGS Trace-Max”, ”IGS Log-det-Max”, ”3 bit IGS Trace-Max” and ”3 bit IGS Log-det-Max” perform similarly, and their PGS based algorithms also have similar performance. Moreover, 3-bit resolution algorithms benefit a greater extend from the increasing number of UEs' receive antennas than their infinite-bit resolution counterparts. Fig. 2 also shows that the performance of ”Trace-Max” and ”IGS Trace-Max” degrade upon increasing the number of the

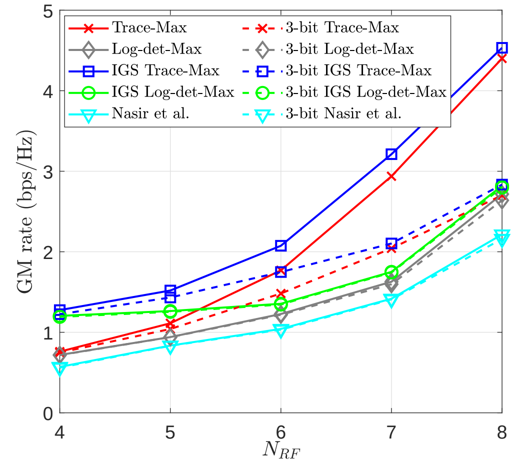



(a)

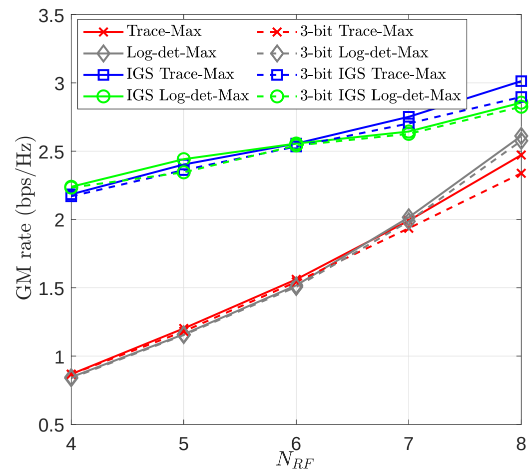



(b)

Fig. 1: (a) Achievable GM rate vs the number $N_{RF}$ of RF chains; (b) Achievable max-min rate vs the number $N_{RF}$ of RF chains;




(a)




(b)

Fig. 2: Achievable GM rate vs the number $N_{RF}$ of RF chains: (a) $N_R = 1$; (b) $N_R = 2$

receiver antennas at the UEs for $N_{RF} \geq 7$, which underlines that "Trace-Max" and "IGS Trace-Max" are poor at processing multiple information streams. As expected, all the algorithms benefit from increasing the number of RF chains. However, in contrast to the ZF based algorithms, RZF trace maximization based algorithms benefit to a greater extent than their 3-bit resolution counterparts for $N_R = 1$. Furthermore, all the IGS based algorithms outperform their proper Gaussian counterparts, confirming the advantage of employing IGS.

Fig. 3 and Fig. 4 portray the min-rate/max-rate ratio (MMR) and the rate variance/mean rate (RV) parameterized by $N_{RF}$. Fig. 3 shows that IGS trace based maximization algorithms have best MMR, and that "IGS Trace-Max" and "IGS Log-det-Max" have better MMR than that of their "Trace-Max" and "Log-det-Max" counterparts under $N_R = 2$, respectively. Fig. 4 shows that "3-bit Trace-Max" has the best RV for $N_R = 1$, while "IGS Trace-Max" and "3-bit IGS Trace-Max" have the best RV for $N_R = 2$. Furthermore, "IGS Log-det-Max" and "3-bit IGS Log-det-Max" have better rate distribution with the increasing number of UEs' receive antennas.

Fig. 5 plots the sum rates (SRs) achieved by the proposed algorithms. Observe that the SR achieved follows the GM rate trend of Fig. 2. Explicitly, "IGS Trace-Max" achieves the overall best SR for $N_R = 1$, "3-bit IGS Log-det-Max" has the best SR among all the 3-bit resolution algorithms, and all the algorithms benefit from increasing the number of RF chains. Fig. 5 also confirms the advantage of employing IGS.

Fig. 6 plots the achievable GM rate versus the number of BS transmit antennas $N$. "IGS Trace-Max" is the overall best performer, while the 3-bit resolution log-det based maximization algorithms have better performance than that of trace based maximization algorithms. Upon increasing the number of BS transmit antennas, all the proposed algorithms achieve better GM rates. Furthermore, trace based maximization algorithms benefit a greater extent from the spatial diversity attained by increasing the number of BS transmit antennas.

We also examine the achievable GM rate under varying power budgets $P$ in Fig. 7. As expected, the achievable GM rate is monotonically increasing. Fig. 7 also shows that all the trace maximization based algorithms degrade upon increasing the number of the receiver antennas at the UEs, confirming that trace maximization based algorithms are poor at processing multiple information streams.

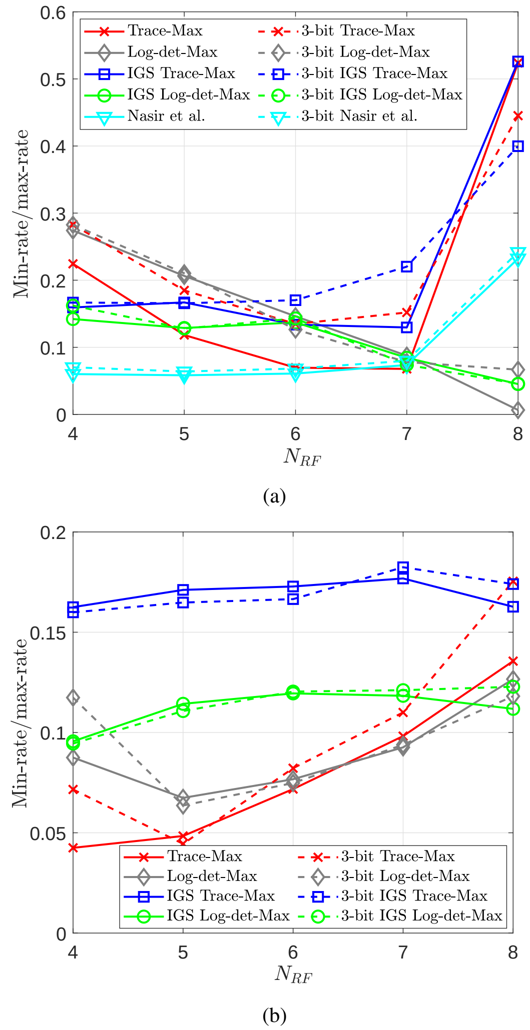



Fig. 3: MMR vs the number $N_{RF}$ of RF chains: (a) $N_R = 1$; (b) $N_R = 2$

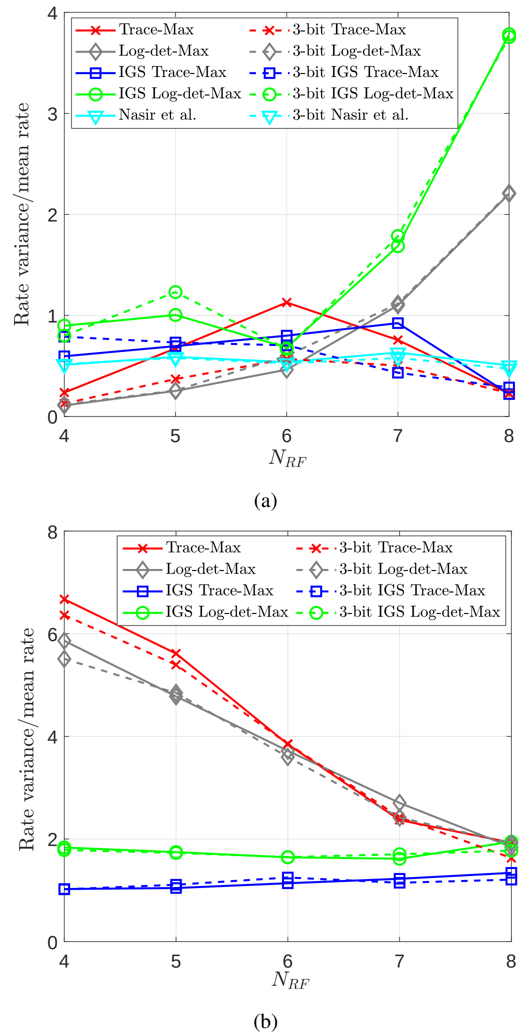



Fig. 4: RV vs the number $N_{RF}$ of RF chains: (a) $N_R = 1$; (b) $N_R = 2$

Furthermore, Fig. 8 allows us to compare the performance achieved by the $b$-bit solution for different values of $b$. Observe that unlike trace maximization base $b$-bit solution algorithms can benefit the increasing $b$, the performance of $b$-bit solutions for log-det maximization is similar with their infinite-solution algorithms.

## VI. Conclusions

A multi-user mmWave network was designed, where a base station uses hybrid beamforming consisting of finite-resolution analog beamforming coupled with digital zero-forcing or regularized zero forcing beamforming. We have proposed several novel algorithms, which iterate by relying on scalable-complex expressions for determining the hybrid beamformer weights maximizing the geometric means of the users' rates. Our extensive simulations have showed that our hybrid beamformers achieve fair user-rate distributions without unduly eroding the overall sum rates.

## Appendices

### *Appendix A: the proof for (38)*

One has (112) with $a^{(\kappa)}$ and $B^{(\kappa)}$ defined from (40), which is the RHS of (39). Note that the RHS of (112) is the linearized function at $\Theta^{(\kappa)}$ of the convex function in the LHS so the later is lower bounded by the former [36].

### *Appendix B: the proof for (67)*

By using the inequality (2), we obtain (113) with $\tilde{a}^{(\kappa)}$, $A^{(\kappa)}$, and $B^{(\kappa)}$ defined from (69). Then the RHS of (113) is the same as the RHS of (67) with $a^{(\kappa)}$ and $C^{(\kappa)}$ defined from (69).

## References

[1] 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.

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

$$\begin{aligned}
\varphi(\boldsymbol{\theta}) &\geq \alpha\left(2\Re\{\langle\boldsymbol{\Theta}(\Theta^{(\kappa)})^H\rangle\} - NN_{RF}\right) - 3\langle[P^{(\kappa+1)}]^2, ([H\Theta^{(\kappa)}]^2)^{-1}\rangle \\
&\quad +2\Re\{\langle[H\Theta^{(\kappa)}]^H([H\Theta^{(\kappa)}]^2)^{-1}[P^{(\kappa+1)}]^2([H\Theta^{(\kappa)}]^2)^{-1}[H\boldsymbol{\Theta}]\rangle\} \\
&= 2\alpha\Re\{\langle\boldsymbol{\Theta}(\Theta^{(\kappa)})^H\rangle\} - a^{(\kappa)} + 2\Re\{\langle B^{(\kappa)}\boldsymbol{\Theta}\rangle\} \\
&= 2\sum_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}} \Re\{\left(\alpha e^{-\jmath\theta^{(\kappa)}_{n,j}} + B^{(\kappa)}(j,n)\right)e^{\jmath\boldsymbol{\theta}_{n,j}}\} - a^{(\kappa)} \\
&= 2\sum_{(n,j)\in\mathcal{N}\times\mathcal{N}_{RF}} \left|\alpha e^{-\jmath\theta^{(\kappa)}_{n,j}} + B^{(\kappa)}(j,n)\right|\cos\left(\angle\left(\alpha e^{-\jmath\theta^{(\kappa)}_{n,j}} + B^{(\kappa)}(j,n)\right) + \boldsymbol{\theta}_{n,j}\right) - a^{(\kappa)}
\end{aligned} \tag{112}$$

$$\begin{aligned}
f(\boldsymbol{\theta}) &\geq f(\theta^{(\kappa)}) - \frac{1}{\alpha}\langle[H\Theta^{(\kappa)}]^2\rangle + \frac{2}{\alpha}\Re\{\langle(\Theta^{(\kappa)})^H H^H H\boldsymbol{\Theta}\rangle\} - \langle\frac{1}{\alpha}I_{N_RK} - \left(\alpha I_{N_RK} + [H\Theta^{(\kappa)}]^2\right)^{-1}, [H\boldsymbol{\Theta}]^2 + \alpha I_{N_RK}\rangle \\
&= f(\theta^{(\kappa)}) - \frac{1}{\alpha}\langle[H\Theta^{(\kappa)}]^2\rangle + \frac{2}{\alpha}\Re\{\langle(\Theta^{(\kappa)})^H H^H H\boldsymbol{\Theta}\rangle\} - \frac{1}{\alpha}\langle I_{N_RK} - \left(\frac{1}{\alpha}[H\Theta^{(\kappa)}]^2 + I_{N_RK}\right)^{-1}, [H\boldsymbol{\Theta}]^2 + \alpha I_{N_RK}\rangle \\
&= \tilde{a}^{(\kappa)} + \frac{1}{\alpha}\left[2\Re\{\langle A^{(\kappa)}\boldsymbol{\Theta}\rangle\} - \langle B^{(\kappa)}, [\boldsymbol{\Theta}]^2\rangle\right] \\
&= \tilde{a}^{(\kappa)} + \frac{1}{\alpha}\left[2\Re\{\langle A^{(\kappa)}\boldsymbol{\Theta}\rangle\} - \lambda_{\max}(B^{(\kappa)})\langle[\boldsymbol{\Theta}]^2\rangle + \langle\lambda_{\max}(B^{(\kappa)})I_N - B^{(\kappa)}, [\boldsymbol{\Theta}]^2\rangle\right]
\end{aligned} \tag{113}$$


[3] 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.

[4] 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. Surv. Tut.*, vol. 20, pp. 870–913, Secondquarter 2018.

[5] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, pp. 3029–3056, Sept. 2015.

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

[7] 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, pp. 106–113, Feb. 2014.

[8] T. S. Rappaport, R. W. Heath, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Pearson/Prentice Hall, 2015.

[9] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: An inclusive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, pp. 949–973, 2nd Quart. 2016.

[10] P. Sudarshan, N. B. Mehta, A. F. Molisch, and J. Zhang, "Channel statistics-based RF pre-processing with antenna selection," *IEEE Trans. Wirel. Commun.*, vol. 5, no. 12, pp. 3501–3511, 2006.

[11] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, 2005.

[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 beamforming for millimeter wave systems using the MMSE criterion," *IEEE Trans. Commun.*, vol. 67, pp. 3693–3708, May 2019.

[16] S. Lyu, Z. Wang, Z. Gao, H. He, and L. Hanzo, "Lattice-based mmWave hybrid beamforming," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4907–4920, 2021.

[17] V. Raghavan, S. Subramanian, J. Cezanne, A. Sampath, O. H. Koymen, and J. Li, "Single-user versus multi-user precoding for millimeter wave MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, p. 13871401, Jun. 2017.

[18] S. He, J. Wang, Y. Huang, B. Ottersten, and W. Hong, "Codebook-based hybrid precoding for millimeter wave multiuser systems," *IEEE Trans. Signal Process.*, vol. 65, pp. 5289–5303, 2017.

[19] Z. Li, S. Han, S. Sangodoyin, R. Wang, and A. F. Molisch, "Joint optimization of hybrid beamforming for multi-user massive MIMO downlink," *IEEE Trans. Wirel. Commun.*, vol. 17, pp. 3600–3614, Jun. 2018.

[20] X. Sun, C. Qi, and G. Y. Li, "Beam training and allocation for multiuser millimeter wave massive MIMO systems," *IEEE Trans. Wirel. Commun.*, vol. 18, pp. 1041–1053, Feb. 2019.

[21] M. Zeng, W. Hao, O. A. Dobre, and H. V. Poor, "Energy-efficient power allocation in uplink mmwave massive MIMO with NOMA," *IEEE Trans. Veh. Techn.*, vol. 68, pp. 3000–3004, Mar. 2019.

[22] M. A. Almasi, M. Vaezi, and H. Mehrpouyan, "Impact of beam misalignment on hybrid beamforming NOMA for mmwave communications," *IEEE Trans. Commun.*, vol. 67, pp. 4505–4518, Jun. 2019.

[23] S. Gong, C. Xing, V. K. N. Lau, S. Chen, and L. Hanzo, "Majorization-minimization aided hybrid transceivers for MIMO interference channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 4903–4918, 2020.

[24] H. Ruan, P. Xiao, L. Xiao, and J. R. Kelly, "Joint iterative optimization-based low-complexity adaptive hybrid beamforming for massive MU-MIMO systems," *IEEE Trans. Commun.*, vol. 69, pp. 1707–1722, Mar. 2021.

[25] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. UK: Cambridge Univ. Press, 2016.

[26] 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.

[27] L. D. Nguyen, H. D. Tuan, T. Q. Duong, H. V. Poor, and L. Hanzo, "Energy-efficient multi-cell massive MIMO subject to minimum user-rate constraints," *IEEE Trans. Commun.*, vol. 69, pp. 914–928, Feb. 2021.

[28] 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.

[29] 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.

[30] S. Sun, T. S. Rappaport, M. Shafi, and H. Tataria, "Analytical framework of hybrid beamforming in multi-cell millimeter-wave systems," *IEEE Trans. Wirel. Commun.*, vol. 17, pp. 7528–7543, Nov. 2018.

[31] A. A. Nasir, H. D. Tuan, T. Q. Duong, H. V. Poor, and L. Hanzo,

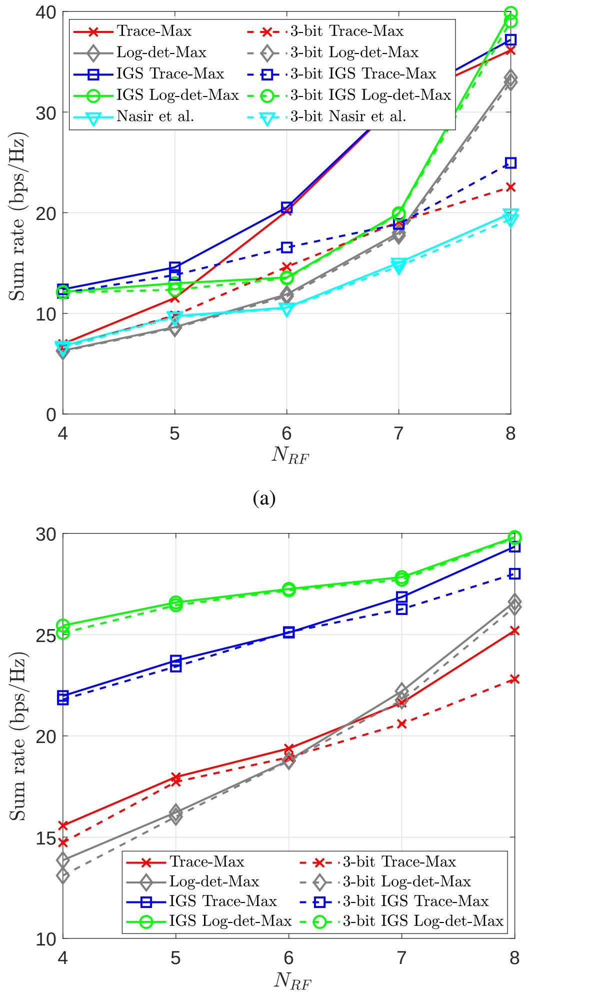



Fig. 5: Achieved SR vs the number $N_{RF}$ of RF chains: (a) $N_R = 1$; (b) $N_R = 2$

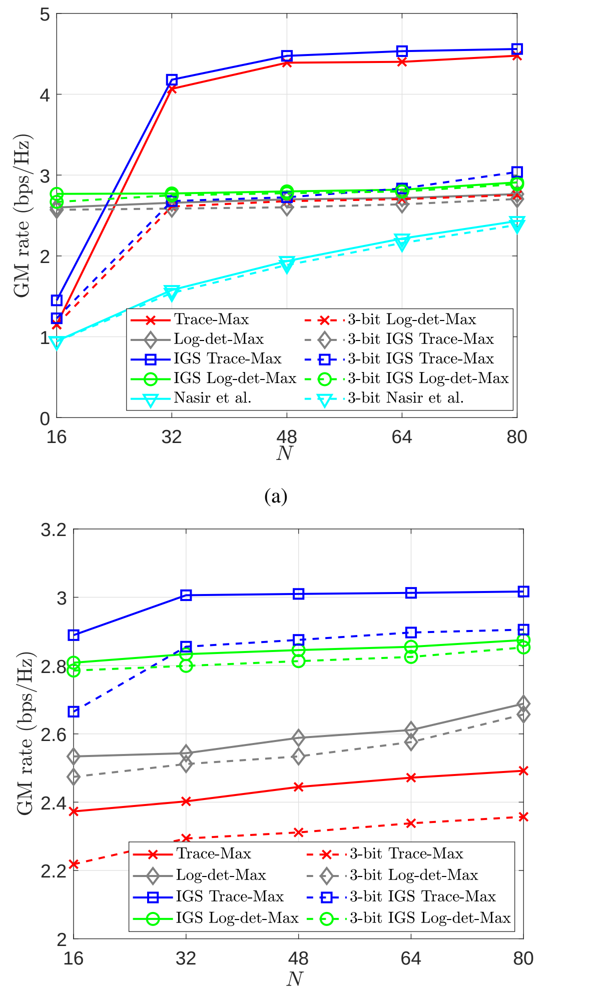



Fig. 6: Achievable GM vs the number $N$ of BS antennas: (a) $N_R = 1$ ; (b) $N_R = 2$


"Hybrid beamforming for multi-user millimeter-wave networks," *IEEE Trans. Veh. Techn.*, vol. 69, pp. 2943–2956, Mar. 2020.

[32] N. Prasad, K. Li, and X. Wang, "Fair-rate allocation in multiuser ofdm-sdma networks," *IEEE Trans. Signal Process.*, vol. 57, pp. 2797–2808, Jul. 2009.

[33] H. Ma, J. Cheng, and X. Wang, "Proportional fair secrecy beamforming for MISO heterogeneous cellular networks with wireless information and power transfer," *IEEE Trans. Commun. vol. 67, no. 8, pp. 5659-5673, Aug. 2019.*, vol. 67, pp. 5659–5673, Aug. 2019.

[34] 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, pp. 488–499, Feb. 2014.

[35] 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.

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

[37] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, pp. 1164–1179, June 2014.

[38] Y.-P. Lin, "On the quantization of phase shifters for hybrid precoding systems," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2237–2246, 2017.

[39] 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.

[40] 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.

[41] 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.

[42] J. D. Krieger, C.-P. Yeang, and G. W. Wornell, "Dense delta-sigma phased arrays," *IEEE Trans. Antenn. Propag.*, vol. 61, p. 18251837, Apr. 2013.

[43] H. Yu, H. D. Tuan, 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, no. 1, pp. 295–309, 2021.

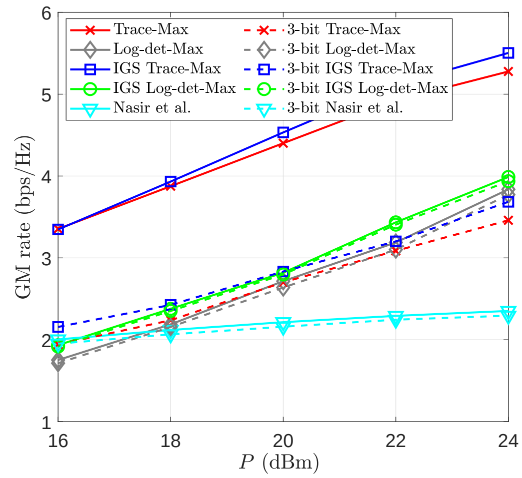
[44] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 1985.

[45] P. Apkarian and H. D. Tuan, "Robust control via concave minimization local and global algorithms," in *Proc. of the 37th IEEE Conf. Decision Control*, pp. 3855–3860, 1998.

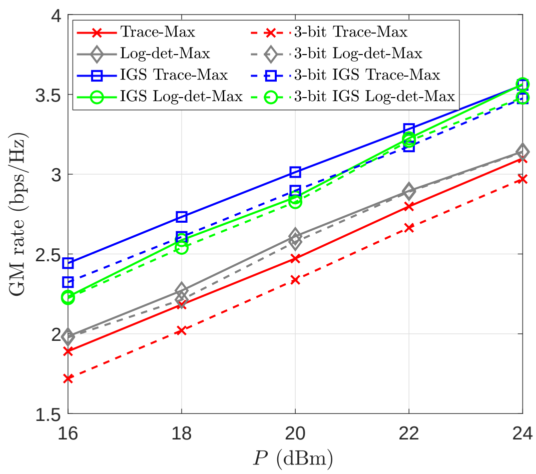
[46] P. Apkarian and H. D. Tuan, "Concave programming in control theory," *J. Glob. Opt.*, vol. 15, pp. 243–270, Apr. 1999.

[47] P. Apkarian and H. D. Tuan, "Robust control via concave optimization: local and global algorithms," *IEEE Trans. Autom. Control*, vol. 45, pp. 299–305, Feb. 2000.

[48] E. Che, H. D. Tuan, and H. H. Nguyen, "Joint optimization of cooperative beamforming and relay assignment in multi-user wireless relay networks," *IEEE Trans. Wirel. Commun.*, vol. 13, pp. 5481–5495, Oct.

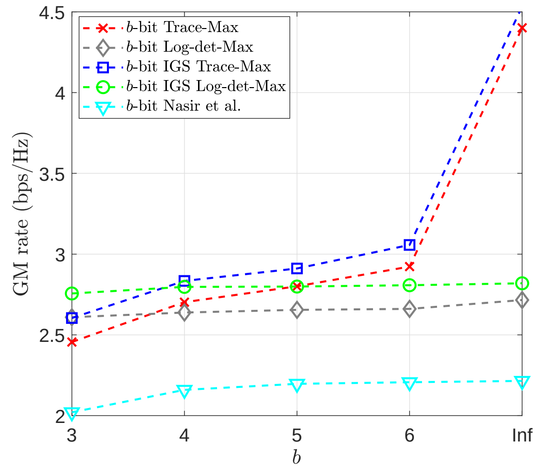



(a)

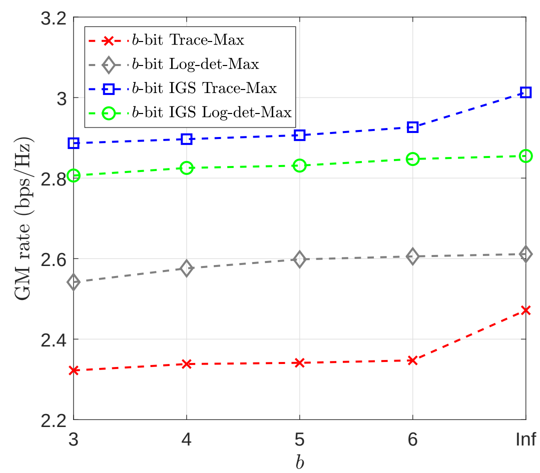



(b)

Fig. 7: Achievable GM vs power budget $P$: (a) $N_R = 1$ ; (b) $N_R = 2$




(a)




(b)

Fig. 8: Achievable GM vs the analog beamforming resolution $b$: (a) $N_R = 1$ ; (b) $N_R = 2$

2014.

[49] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wirel. Commun.*, vol. 14, pp. 6481–6494, Nov. 2015.

[50] R. Schneider, *Convex bodies: The Brunn-Minkowski theory,*. Cambridge University Press, Cambridge, 1993.

[51] C. Villani, *Optimal Transport: Old and New*. Springer Verlag, 2009.

[52] 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.

[53] P. J. Schrerier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge University Press, 2010.

[54] S. Lagen, A. Agustin, and J. Vidal, "Coexisting linear and widely linear transceivers in the MIMO interference channel," *IEEE Trans. Signal Process.*, vol. 64, pp. 652–664, Feb 2016.

[55] H. D. Tuan, A. A. Nasir, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Non-orthogonal multiple access with improper Gaussian signaling," *IEEE J. Selec. Topics Signal Process.*, vol. 13, pp. 496–507, Mar. 2019.

[56] 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.

[57] 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. Sel. Areas Commun.*, vol. 38, pp. 2589–2603, Nov. 2020.

[58] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[59] "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.

[60] A. K. Gupta, J. G. Andrews, and R. W. Heath, "On the feasibility of sharing spectrum licenses in mmWave cellular systems," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3981–3995, 2016.

[61] T. S. Rappaport, G. R. MacCartney, S. Sun, H. Yan, and S. Deng, "Small-scale, local area, and transitional millimeter wave propagation for 5G communications," *IEEE Trans. Antenn. Propag.*, vol. 65, no. 12, pp. 6474–6490, 2017.