Long-Term Rate-Fairness-Aware Beamforming Based Massive MIMO Systems

W. Zhu[1,2], H. D. Tuan[2], E. Dutkiewicz[2], Y. Fang[1], H. V. Poor[3], and L. Hanzo[4]

Abstract—This is the first treatise on multi-user (MU) beamforming designed for achieving long-term rate-fairness in full-dimensional MU massive multi-input multi-output (m-MIMO) systems. Explicitly, based on the channel covariances, which can be assumed to be known beforehand, we address this problem by optimizing the following objective functions: the users' signal-to-leakage-noise ratios (SLNRs) using SLNR max-min optimization, geometric mean of SLNRs (GM-SLNR) based optimization, and SLNR soft max-min optimization. We develop a convex-solver based algorithm, which invokes a convex subproblem of cubic time-complexity at each iteration for solving the SLNR max-min problem. We then develop closed-form expression based algorithms of scalable complexity for the solution of the GM-SLNR and of the SLNR soft max-min problem. The simulations provided confirm the users' improved-fairness ergodic rate distributions.

Index Terms—Full-dimensional massive MIMO, statistical beamforming, ergodic rate, signal to leakage plus noise ration (SLNR) optimization.

I. INTRODUCTION

Massive multi-input multi-output (m-MIMO) [1] schemes have been shown to constitute a disruptive technology capable of providing reliable quality of service for multiple users at a low transmission power. With the explosive proliferation of transmit antennas, one has to rely on a large two-dimensional (2D) uniformly spaced rectangular antenna array (URA) to implement m-MIMO schemes [2]–[5]. Transmit beamforming (TBF) constitutes a prominent member of the m-MIMO family. The most popular m-MIMO beamformers such as zero-forcing/regularized zero-forcing and conjugate beamformers [1], [6], [7] rely on accurate channel state information (CSI), which is a challenging task. Fortunately, the second-order statistics of full-dimensional m-MIMO channels can be known beforehand [8], [9] or efficiently signaled back to the transmitter through feedback channels. In this context,

it is of great practical interest to consider the problem of designing statistical beamformers that are capable of maintaining fairness in terms of the users' ergodic rates, including their fair long-term evolution. Unfortunately, this problem is computationally intractable, because the ergodic rate function is not analytical in beamformers [10]–[15]. To resolve this issue, statistical beamforming tends to rely on the so-called signal-to-leakage-noise-ratio (SLNR) optimization [16]. Due to the statistical independence of the individual signal and its leakage plus noise, the mathematical expectation of an individual SLNR admits an analytical lower bound as the ratio of two convex quadratic functions of the individual beamformer, thanks to the Mullen theorem [17]. With respect to the individual beamformer power, obtaining the optimal beamformer that maximizes this ratio can be readily achieved, thanks to the application of the Rayleigh-Ritz quotient theorem. This approach has been explored in previous studies using equal or fixed power allocation techniques [11], [16], [18]–[22]. However, the equal-power allocation falls short of achieving rate-fairness, even in the case of zero-forcing or regularized zero-forcing beamforming [6], [7]. Therefore, the problem of SLNR max-min optimization under a specific power constraint, which represents the most intuitive approach to SLNR optimization, still lacks a computational solution.

Against the above background, this paper is the first one optimizing the SLNRs for improving all users' ergodic rates. Explicitly, the paper offers the following contributions:

- We develop an algorithm to tackle the aforementioned max-min SLNR optimization problem. This algorithm iterates by solving convex problems (CPs);
- Since there is no precise proportionality relationship between the SLNRs and the individual users' ergodic rates, we have to find a work-around. To elaborate, given that the SLNR max-min and ergodic rate max-min problems are not equivalent, we propose other SLNR optimization problems, such as maximizing the geometric mean of the SLNRs (GM-SLNR maximization) and a soft min SLNR (SLNR soft max-min) problem. We demonstrate that these objective functions achieve an improved minimum ergodic rate. Importantly, we devise algorithms that iterate by evaluating closed-form (CF) expressions of scalable complexity, rendering them practical solutions for m-MIMO systems;
- The simulations provided reveal a surprising fact: the SLNR max-min optimization having the most balanced SLNR distributions associated with the highest minimum SLNR actually tends to result in the most unbalanced

[1]School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China (email: wenbozhu@shu.edu.cn, yfang@staff.shu.edu.cn); [2]School of Electrical and Data Engineering, University of Technology Sydney, Broadway, NSW 2007, Australia (email: wenbo.zhu@student.uts.edu.au, tuan.hoang@uts.edu.au, eryk.dutkiewicz@uts.edu.au); [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, U.K (email: lh@ecs.soton.ac.uk) (Corresponding Author: Y. Fang)
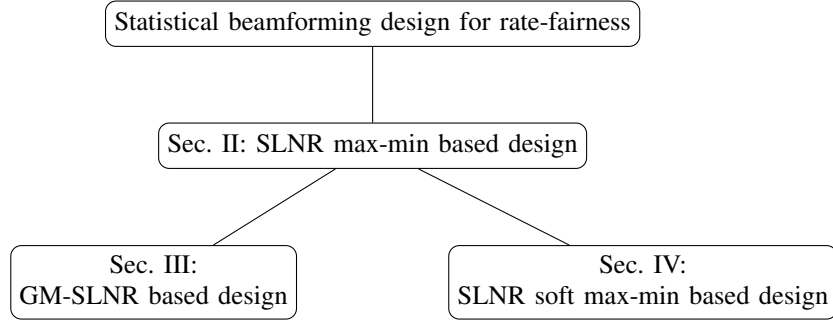
Fig. 1: The outline of the paper highlighting the proposed beamforming designs.

ergodic rate distributions. Unexpectedly, it also has the lowest minimum ergodic rate. By contrast, the soft max-min optimization exhibits the most balanced ergodic rate distributions and has the highest minimum ergodic rate. *Thus, the soft max-min optimization has the best performance vs. complexity trade-off*;

- Statistical beamforming is based on channel statistics alone, hence it is free from channel-estimation overhead. As such, a byproduct of our results is a new beamforming design, which requires very limited feedback overhead. Hence it is eminently suitable for long-term multicast beamforming [23] or MIMO-OFDM [24].

The paper is organized as follows. Section II is devoted to our beamforming design relying on maximizing the minimum SLNR. By contrast, the beamforming designs based on maximizing the GM-SLNR and the soft min SLNR is considered in Section III and Section IV, respectively. Our simulations are provided in Section V, while our conclusions are discussed in Section VI. The Appendix provides mathematical ingredients for the paper, which includes a new quadratic minorant of a ratio of two quadratic functions. To help the reader to smoothly follow the technical content of the paper, we provide the flowchart in Fig. 1.

Notation. Boldface fonts are exclusively used to represent decision variables; The $N \times N$ identity matrix is represented by I_N; $\mathcal{C}(0, a)$ represents the set of circular Gaussian random variables having zero means and a variance of a; The dot product of matrices X and Y is given by $\langle X, Y \rangle = \text{trace}(X^H Y)$, but we also use $\langle X \rangle$ to represent $\text{trace}(X)$; For the matrix X, the outer product XX^H is denoted by $[X]^2$; The expression $X \succeq 0$ indicates that the matrix X is positive semi-definite; $\otimes$ represents the Kronecker product operator; Lastly,

$$\text{vec} \begin{bmatrix} x_1 & \dots & x_L \end{bmatrix} \triangleq \begin{bmatrix} x_1 \\ \dots \\ x_L \end{bmatrix}.$$

Table I presents a summary of the used notations.

II. SLNR MAX-MIN BASED BEAMFORMING

We consider a network supported by a base station (BS) equipped with a $Q \times Q$-uniform rectangular array (URA) to serve L single-antenna downlink users (DUs), which are indexed by $\ell \in \mathcal{L} \triangleq \{1, \dots, L\}$. The matrix $H_\ell \in \mathbb{C}^{Q \times Q}$

of entries $H_\ell(q, q')$ represents the channel spanning from the (q, q')-th antenna to DU ℓ is modeled by [25], [26]

$$H_\ell = \sqrt{\beta_\ell} \sqrt{R_{v,\ell}} H_{s,\ell} \sqrt{R_{h,\ell}}, \tag{1}$$

where β_ℓ is a deterministic positive number representing the path-loss, $R_{v,\ell}$ and $R_{h,\ell}$ are deterministic positive definite matrices representing the vertical and horizontal correlation matrices, while $H_{s,\ell} \in \mathbb{C}^{Q \times Q}$ is a random matrix having zero-mean and unit-variance identically distributed (iid) complex Gaussian elements, representing the small-scale fading. Thus, we have

$$h_\ell \triangleq \text{vec}(H_\ell) = \sqrt{\beta_\ell}(\sqrt{R_{h,\ell}}^T \otimes \sqrt{R_{v,\ell}})\text{vec}(H_{s,\ell}), \tag{2}$$

so

$$\mathbb{E}(h_\ell h_\ell^H) = \beta_\ell R_{h,\ell}^T \otimes R_{v,\ell}. \tag{3}$$

Furthermore, $h_\ell^T = \sqrt{\beta_\ell}\text{vec}^T(H_{s,\ell})(\sqrt{R_{h,\ell}} \otimes \sqrt{R_{v,\ell}}^T)$, hence we have

$$\mathcal{R}_\ell \triangleq \mathbb{E}[(h_\ell^T)^H h_\ell^T] = \beta_\ell R_{h,\ell} \otimes R_{v,\ell}^T. \tag{4}$$

Additionally,

$$H_\ell^T = \sqrt{\beta_\ell} \sqrt{R_{h,\ell}}^T H_{s,\ell}^T \sqrt{R_{v,\ell}}^T, \tag{5}$$

therefore

$$\tilde{h}_\ell \triangleq \text{vec}(H_\ell^T) = \sqrt{\beta_\ell}(\sqrt{R_{v,\ell}} \otimes \sqrt{R_{h,\ell}}^T)\text{vec}(H_{s,\ell}), \tag{6}$$

and $\tilde{h}_\ell^T = \sqrt{\beta_\ell}\text{vec}^T(H_{s,\ell})(\sqrt{R_{v,\ell}}^T \otimes \sqrt{R_{h,\ell}})$, for which we have

$$\widetilde{\mathcal{R}}_\ell \triangleq \mathbb{E}[(\tilde{h}_\ell^T)^H \tilde{h}_\ell^T] = \beta_\ell R_{v,\ell}^T \otimes R_{h,\ell}. \tag{7}$$

Let $s_\ell \in \mathcal{C}(0, 1)$ be the information symbol intended for DU $\ell \in \mathcal{L}$, which is beamformed by the matrix $\mathbf{V}_\ell \in \mathbb{C}^{Q \times Q}$ for the BS's downlink (DL) transmission, i.e. the (q, q')-th antenna transmits the signal $\sum_{\ell \in \mathcal{L}} \mathbf{V}_\ell(q, q')s_\ell$. The DL signal received at DU ℓ is given by

$$y_\ell = \sum_{\ell' \in \mathcal{L}} \langle H_\ell^T \mathbf{V}_{\ell'} \rangle s_{\ell'} + n_\ell, \tag{8}$$

where $n_\ell \in \mathcal{C}(0, \sigma)$ is the background noise at DU ℓ. In the previous paper [27], we proposed the following new class of $\mathbf{V}_\ell \in \mathbb{C}^{Q \times Q}$

$$\mathbf{V}_\ell = \sum_{m \in \mathcal{M}} \mathbf{v}_{m,\ell}^e (\mathbf{v}_{m,\ell}^a)^T, \tag{9}$$

TABLE I: Notations used in the paper

Notation	Description
Q	number of vertical and horizontal antennas at the base station
M	number of outer products of the azimuth and elevation beamformers
L	number of downlink users (DUs)
$\mathbf{v}_\ell^a, \mathbf{v}_\ell^e$	azimuth and elevation beamformers for user ℓ
$\mathbf{v}^a, \mathbf{v}^e$	sets of azimuth and elevation beamformers
$\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)$	SLNR for user ℓ
$\rho_\ell(\mathbf{v}^a, \mathbf{v}^e)$	achievable ergodic rate for user ℓ

with $\mathcal{M} \triangleq \{1, \ldots, M\}$. In (9), $\mathbf{v}_{m,\ell}^e \in \mathbb{C}^Q$ plays the role of elevation beamforming, while $\mathbf{v}_{m,\ell}^a \in \mathbb{C}^Q$ plays the role of azimuth beamforming. The $2D$-structured beamformer constructed as the outer product of the elevation and azimuth beamformers [22], [25], [28]–[30] corresponds to $M = 1$, while the unstructured FD $\mathbf{V}_\ell \in \mathbb{C}^{M \times M}$ corresponds to $M = Q$. Importantly, it has been shown in [27] that the beamformer (9) constructed for $M = 2$ already achieves instantaneous rates (under the full availability of channel state information) close to those achieved by relying on $M = Q$. By contrast, the distinct objective of the present paper is to investigate its performance in terms of the ergodic rate achieved.

With $\mathbf{V}_\ell$ defined by (9), equation (8) of the DL signal received at DU ℓ becomes

$$y_\ell = \sum_{\ell' \in \mathcal{L}} \left(\sum_{m \in \mathcal{M}} (\mathbf{v}_{m,\ell'}^e)^T H_\ell \mathbf{v}_{m,\ell'}^a \right) s_{\ell'} + n_\ell. \quad (10)$$

In what follows, we will use the notations

$$\mathbf{v}_\ell^z \triangleq \text{vec} \begin{bmatrix} \mathbf{v}_{1,\ell}^z & \cdots & \mathbf{v}_{M,\ell}^z \end{bmatrix} \in \mathbb{C}^{QM}, z \in \{a, e\}, \quad (11)$$

and then $\mathbf{v}^a \triangleq \{\mathbf{v}_\ell^a, \ell \in \mathcal{L}\}$, and $\mathbf{v}^e \triangleq \{\mathbf{v}_\ell^e, \ell \in \mathcal{L}\}$, and $\mathbf{v} \triangleq \{\mathbf{v}^a, \mathbf{v}^e\}$. The achievable ergodic rate at DU ℓ is defined by

$$\rho_\ell(\mathbf{v}^a, \mathbf{v}^e) \triangleq \mathbb{E}_{h_\ell} \left[\ln \left(1 + \text{SINR}_\ell(\mathbf{v}^a, \mathbf{v}^e) \right) \right], \quad (12)$$

where we have

$$\text{SINR}_\ell(\mathbf{v}^a, \mathbf{v}^e) \triangleq \frac{|\sum_{m \in \mathcal{M}} (\mathbf{v}_{m,\ell}^e)^T H_\ell \mathbf{v}_{m,\ell}^a|^2}{\sum_{\ell' \in \mathcal{L} \setminus \{\ell\}} |\sum_{m \in \mathcal{M}} (\mathbf{v}_{m,\ell}^e)^T H_\ell \mathbf{v}_{m,\ell'}^a|^2 + \sigma}. \quad (13)$$

Given $(\mathbf{v}^a, \mathbf{v}^e)$, the statistical expectation in (12) is calculated over all realization of H_ℓ with the covariance defined from (4). This is contrasted to the instantaneous rate of $\ln(1 + \text{SINR}_\ell(\mathbf{v}^a, \mathbf{v}^e))$, which is calculated at the specific instantaneous value of H_ℓ.

Thus, the problem of maximizing the minimum ergodic rate is formulated as

$$\max_{\mathbf{v} = (\mathbf{v}^a, \mathbf{v}^e)} \min_{\ell \in \mathcal{L}} \rho_\ell(\mathbf{v}^a, \mathbf{v}^e) \quad (14a)$$

$$\text{s.t.} \quad \sum_{\ell \in \mathcal{L}} || \sum_{m \in \mathcal{M}} \mathbf{v}_{m,\ell}^a (\mathbf{v}_{m,\ell}^e)^T ||^2 \leq P. \quad (14b)$$

However, the ergodic rate function ρ_ℓ is not computationally tractable and one has to use the so-called SLNR defined by

$$\mathbb{E}_h \left[\frac{|\sum_{m \in \mathcal{M}} (\mathbf{v}_{m,\ell}^e)^T H_\ell \mathbf{v}_{m,\ell}^a|^2}{\sum_{\ell' \in \mathcal{L} \setminus \{\ell\}} |\sum_{m \in \mathcal{M}} (\mathbf{v}_{m,\ell}^e)^T H_{\ell'} \mathbf{v}_{m,\ell}^a|^2 + \sigma} \right], \quad (15)$$

and use the Mullen theorem [17] of $\mathbb{E}[x/y] \geq \mathbb{E}[x]/\mathbb{E}[y]$ for independent random variables to obtain its lower bound

$$\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e) \triangleq \frac{\mathbb{E}_{h_\ell}[|\sum_{m \in \mathcal{M}} (\mathbf{v}_{m,\ell}^e)^T H_\ell \mathbf{v}_{m,\ell}^a|^2]}{\sum_{\ell' \in \mathcal{L} \setminus \{\ell\}} \mathbb{E}_{h_{\ell'}}[|\sum_{m \in \mathcal{M}} (\mathbf{v}_{m,\ell}^e)^T H_{\ell'} \mathbf{v}_{m,\ell}^a|^2] + \sigma}. \quad (16)$$

Thus, the most natural computationally tractable surrogate of problem (14) representing a computationally intractable max-min ergodic rate optimization is the following problem

$$\max_{\mathbf{v}^a, \mathbf{v}^e} \gamma(\mathbf{v}^a, \mathbf{v}^e) \triangleq \min_{\ell \in \mathcal{L}} \gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e) \quad \text{s.t.} \quad (14b). \quad (17)$$

For convenience of terminology, we will still refer to this problem as SLNR max-min optimization. This section aims to solve this problem.

Using the initialization of $(v^{a,(0)}, v^{e,(0)})$ that is feasible for (14b), for $\nu = 1, \ldots$, and letting $(v^{a,(\nu)}, v^{e,(\nu)})$ be a feasible point of (14b) that is found from the $(\nu - 1)$-st iteration, we process the ν-th iteration to generate $(v^{a,(\nu+1)}, v^{e,(\nu+1)})$.

A. Azimuth beamforming optimization

For $\mathbf{v}^e$ fixed at $v^{e,(\nu)}$, the power constraint (14b) is expressed by

$$\sum_{\ell \in \mathcal{L}} \langle \mathcal{A}_\ell^{e,(\nu)}, [\mathbf{v}_\ell^a]^2 \rangle \leq P, \quad (18)$$

with

$$0 \preceq \mathcal{A}_\ell^{e,(\nu)} \triangleq \left[\langle v_{m,\ell}^{e,(\nu)}, v_{m',\ell}^{e,(\nu)} \rangle I_M \right]_{(m,m') \in \mathcal{M} \times \mathcal{M}}$$
$$\in \mathbb{C}^{(QM) \times (QM)}. \quad (19)$$

Let us now define the random variables

$$h_{\ell',\ell}^{e,(\nu)} \triangleq \left[(v_{1,\ell'}^{e,(\nu)})^T H_\ell \ldots \quad (v_{M,\ell'}^{e,(\nu)})^T H_\ell \right]$$
$$= \begin{bmatrix} H_\ell^T v_{1,\ell'}^{e,(\nu)} \\ \ldots \\ H_\ell^T v_{M,\ell'}^{e,(\nu)} \end{bmatrix}^T$$
$$= \begin{bmatrix} \left((v_{1,\ell'}^{e,(\nu)})^T \otimes I_Q \right) \text{vec}(H_\ell^T) \\ \ldots \\ \left((v_{M,\ell'}^{e,(\nu)})^T \otimes I_Q \right) \text{vec}(H_\ell^T) \end{bmatrix}^T$$
$$= \text{vec}^T(H_\ell^T) \left[v_{1,\ell'}^{e,(\nu)} \otimes I_Q \quad \ldots \quad v_{M,\ell'}^{e,(\nu)} \otimes I_Q \right]$$
$$= \tilde{h}_\ell^T \Psi_{\ell'}^{e,(\nu)} \in \mathbb{C}^{1 \times (QM)}, (\ell', \ell) \in \mathcal{L} \times \mathcal{L}, \quad (20)$$

with $\tilde{h}_\ell$ defined from (6), and

$$\Psi_{\ell'}^{e,(\nu)} \triangleq \left[v_{1,\ell'}^{e,(\nu)} \quad \ldots \quad v_{M,\ell'}^{e,(\nu)} \right] \otimes I_Q. \quad (21)$$

We then express $\gamma_\ell(\mathbf{v}_\ell^a, v_\ell^{e,(\nu)})$ by

$$\gamma_\ell(\mathbf{v}_\ell^a, v_\ell^{e,(\nu)}) = \gamma_\ell^a(\mathbf{v}_\ell^a)$$
$$\triangleq \frac{\mathbb{E}[|h_{\ell,\ell}^{e,(\nu)}\mathbf{v}_\ell^a|^2]}{\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}}\mathbb{E}[|h_{\ell,\ell'}^{e,(\nu)}\mathbf{v}_\ell^a|^2]+\sigma}$$
$$= \frac{||\sqrt{\widetilde{\mathcal{R}}_\ell}\Psi_\ell^{e,(\nu)}\mathbf{v}_\ell^a||^2}{\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}}||\sqrt{\widetilde{\mathcal{R}}_{\ell'}}\Psi_\ell^{e,(\nu)}\mathbf{v}_\ell^a||^2+\sigma}, \quad (22)$$

where $\widetilde{\mathcal{R}}_\ell$ is defined from (7). Thus, alternating optimization in $\mathbf{v}^a$ is formulated by the following problem:

$$\max_{\mathbf{v}^a}\min_{\ell\in\mathcal{L}}\gamma_\ell^a(\mathbf{v}_\ell^a) \quad \text{s.t.} \quad (18). \quad (23)$$

The following tight minorant of $\gamma_\ell^a(\mathbf{v}_\ell^a)$ at $v^{a,(\nu)}$ is obtained by applying the inequality (84) for $(\bar\chi,\bar x)=(\bar\chi_\ell,\bar x_\ell)\triangleq (\sqrt{\widetilde{\mathcal{R}}_\ell}\Psi_\ell^{e,(\nu)}v_\ell^{a,(\nu)},\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}}||\sqrt{\widetilde{\mathcal{R}}_{\ell'}}\Psi_\ell^{e,(\nu)}v_\ell^{a,(\nu)}||^2+\sigma)$,

$$\gamma_\ell^{a,(\nu)}(\mathbf{v}_\ell^a)\triangleq a_\ell^{a,(\nu)}+2\Re\{b_\ell^{a,(\nu)}\mathbf{v}_\ell^a\}$$
$$-c_\ell^{a,(\nu)}\sum_{\ell'\in\mathcal{L}}||\sqrt{\widetilde{\mathcal{R}}_{\ell'}}\Psi_\ell^{e,(\nu)}\mathbf{v}_\ell^a||^2, \quad (24)$$

for

$$a_\ell^{a,(\nu)}\triangleq -\frac{||\bar\chi_\ell||^4}{\bar x_\ell^2}-\sigma c_\ell^{a,(\nu)},$$
$$b_\ell^{a,(\nu)}\triangleq \frac{\bar x_\ell+||\bar\chi_\ell||^2}{\bar x_\ell^2}(v_\ell^{a,(\nu)})^H(\Psi_\ell^{e,(\nu)})^H\widetilde{\mathcal{R}}_\ell\Psi_\ell^{e,(\nu)}, \quad (25)$$
$$c_\ell^{a,(\nu)}\triangleq \frac{||\bar\chi_\ell||^2}{\bar x_\ell^2}.$$

Thus, we solve the following convex problem

$$\max_{\mathbf{v}^a}\min_{\ell\in\mathcal{L}}\gamma_\ell^{a,(\nu)}(\mathbf{v}_\ell^a) \quad \text{s.t.} \quad (18), \quad (26)$$

to generate $v^{a,(\nu+1)}$ satisfying

$$\gamma(v^{a,(\nu+1)},v^{e,(\nu)})>\gamma(v^{a,(\nu)},v^{e,(\nu)}) \quad (27)$$

whenever $\gamma(v^{a,(\nu)},v^{e,(\nu)})\neq\gamma(v^{a,(\nu+1)},v^{e,(\nu)})$.

B. Elevation beamforming optimization

For $\mathbf{v}^a$ fixed at $v^{a,(\nu+1)}$, the power constraint (14b) is formulated as

$$\sum_{\ell\in\mathcal{L}}\langle\mathcal{A}_\ell^{a,(\nu)},[\mathbf{v}_\ell^e]^2\rangle\leq P, \quad (28)$$

with

$$\mathcal{A}_\ell^{a,(\nu)}\triangleq\left[\langle v_{m,\ell}^{a,(\nu+1)},v_{m',\ell}^{a,(\nu+1)}\rangle I_M\right]_{(m,m')\in\mathcal{M}\times\mathcal{M}}$$
$$\in\mathbb{C}^{(QM)\times(QM)}. \quad (29)$$

Let us now define the random variables

$$h_{\ell',\ell}^{a,(\nu)}\triangleq\left[(v_{1,\ell'}^{a,(\nu+1)})^TH_\ell^T \quad \cdots \quad (v_{M,\ell'}^{a,(\nu+1)})^TH_\ell^T\right]$$
$$=\begin{bmatrix}H_\ell v_{1,\ell'}^{a,(\nu+1)}\\ \cdots \\ H_\ell v_{M,\ell'}^{a,(\nu+1)}\end{bmatrix}^T$$

$$=\begin{bmatrix}\left((v_{1,\ell'}^{a,(\nu+1)})^T\otimes I_Q\right)\mathsf{vec}(H_\ell)\\ \cdots \\ \left((v_{M,\ell'}^{a,(\nu+1)})^T\otimes I_Q\right)\mathsf{vec}(H_\ell)\end{bmatrix}^T$$
$$=\mathsf{vec}^T(H_\ell)\left[v_{1,\ell'}^{a,(\nu+1)}\otimes I_Q \quad \cdots \quad v_{M,\ell'}^{a,(\nu+1)}\otimes I_Q\right]$$
$$=h_\ell^T\Psi_{\ell'}^{a,(\nu)}\in\mathbb{C}^{1\times(QM)}, \quad (30)$$

with h_ℓ defined from (2), and

$$\Psi_{\ell'}^{a,(\nu)}\triangleq\left[v_{1,\ell'}^{a,(\nu+1)} \quad \cdots \quad v_{M,\ell'}^{a,(\nu+1)}\right]\otimes I_Q. \quad (31)$$

We then express $\gamma_\ell(v^{a,(\nu+1)},\mathbf{v}_\ell^e)$ by

$$\gamma_\ell(v^{a,(\nu+1)},\mathbf{v}_\ell^e) = \gamma_\ell^e(\mathbf{v}_\ell^e)$$
$$\triangleq \frac{\mathbb{E}[|h_{\ell,\ell}^{a,(\nu)}\mathbf{v}_\ell^e|^2]}{\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}}\mathbb{E}[|h_{\ell,\ell'}^{a,(\nu)}\mathbf{v}_\ell^e|^2]+\sigma}$$
$$= \frac{||\sqrt{\mathcal{R}_\ell}\Psi_\ell^{a,(\nu)}\mathbf{v}_\ell^e||^2}{\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}}||\sqrt{\mathcal{R}_{\ell'}}\Psi_\ell^{a,(\nu)}\mathbf{v}_\ell^e||^2+\sigma}, \quad (32)$$

with $\mathcal{R}_\ell$ defined from (4).

The alternating optimization in $\mathbf{v}^e$ is formulated by the problem:

$$\max_{\mathbf{v}^e}\min_{\ell\in\mathcal{L}}\gamma_\ell^e(\mathbf{v}_\ell^e) \quad \text{s.t.} \quad (28). \quad (33)$$

The following tight minorant of $f_\ell^e(\mathbf{v}_\ell^e)$ at $v_\ell^{e,(\nu)}$ is obtained by applying the inequality (84) for $(\bar\chi,\bar x)=(\bar\chi_\ell,\bar x_\ell)\triangleq (\sqrt{\mathcal{R}_\ell}\Psi_\ell^{a,(\nu)}v_\ell^{e,(\nu)},\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}}||\sqrt{\mathcal{R}_{\ell'}}\Psi_\ell^{a,(\nu)}v_\ell^{e,(\nu)}||^2+\sigma)$,

$$\gamma_\ell^{e,(\nu)}(\mathbf{v}_\ell^e)\triangleq a_\ell^{e,(\nu)}+2\Re\{b_\ell^{e,(\nu)}\mathbf{v}_\ell^e\}$$
$$-c_\ell^{e,(\nu)}\sum_{\ell'\in\mathcal{L}}||\sqrt{\mathcal{R}_{\ell'}}\Psi_\ell^{a,(\nu)}\mathbf{v}_\ell^e||^2, \quad (34)$$

for

$$a_\ell^{e,(\nu)}\triangleq -\frac{||\bar\chi_\ell||^4}{\bar x_\ell^2}-\sigma c_\ell^{e,(\nu)},$$
$$b_\ell^{e,(\nu)}\triangleq \frac{\bar x_\ell+||\bar\chi||^2}{\bar x_\ell^2}(v_\ell^{e,(\nu)})^H(\Psi_\ell^{a,(\nu)})^H\mathcal{R}_\ell\Psi_\ell^{a,(\nu)}, \quad (35)$$
$$c_\ell^{e,(\nu)}\triangleq \frac{||\bar\chi_\ell||^2}{\bar x_\ell^2}.$$

We now solve the following convex problem

$$\max_{\mathbf{v}^e}\min_{\ell\in\mathcal{L}}\gamma_\ell^{e,(\nu)}(\mathbf{v}_\ell^e) \quad \text{s.t.} \quad (28) \quad (36)$$

to generate $v^{e,(\nu+1)}$ satisfying

$$\gamma(v^{a,(\nu+1)},v^{e,(\nu+1)})>\gamma(v^{a,(\nu+1)},v^{e,(\nu)}), \quad (37)$$

whenever $\gamma(v^{a,(\nu+1)},v^{e,(\nu+1)})\neq\gamma(v^{a,(\nu+1)},v^{e,(\nu)})$.

C. Algorithm and its computational complexity

Algorithm 1 provides the pseudo code for computing (17) by solving the convex problems (26) and (36) having the complexity order of $\mathcal{O}(Q^3M^3)$ to generate the next iterative feasible point $(v^{a,(\nu+1)},v^{e,(\nu+1)})$, which according to (27) and (37) satisfies $\gamma(v^{a,(\nu+1)},v^{e,(\nu+1)})>\gamma(v^{a,(\nu)},v^{e,(\nu)})$. The sequence $\{(v^{a,(\nu)},v^{e,(\nu)})\}$ thus cannot be cyclic and thus it is convergent. Since having a higher minimum SLNR does not necessarily imply having a higher minimum ergodic rate

---

Algorithm 1 SLNR max-min based statistical beamforming algorithm

---

1: **Initialization:** Initialize a feasible $(v^{a,(0)}, v^{e,(0)})$ for (14b). Set $(v^{a,opt}, v^{e,opt}) = (v^{a,(0)}, v^{e,(0)})$ and $\text{rate}_{\min} = \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,opt}, v^{e,opt})$.

2: **Repeat until convergence:** Generate $v^{a,(\nu+1)}$ by solving the convex problem (26). If $\min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu)}) > \text{rate}_{\min}$, update $(v^{a,opt}, v^{e,opt}) \leftarrow (v^{a,(\nu+1)}, v^{e,(\nu)})$ and $\text{rate}_{\min} \leftarrow \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu)})$. Generate $v^{e,(\nu+1)}$ by solving the convex problem (36). If $\min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu+1)}) > \text{rate}_{\min}$, update $(v^{a,opt}, v^{e,opt}) \leftarrow (v^{a,(\nu+1)}, v^{e,(\nu+1)})$ and $\text{rate}_{\min} \leftarrow \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu+1)})$. Reset $\nu \leftarrow \nu + 1$.

3: **Output** $(v^{a,opt}, v^{e,opt})$ and $\text{rate}_{\min}$.

---

defined from (12)[1], whenever generating a new feasible point for (17), Algorithm 1 also updates the incumbent beamformers achieving the best minimum user rate.

III. GM-SLNR MAXIMIZATION BASED BEAMFORMING

Note that the SLNR max min problem (17) is equivalent to the problem

$$\max_{\mathbf{v}^a, \mathbf{v}^e} \min_{\ell \in \mathcal{L}} \ln \left[1 + \gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e) \right] \quad \text{s.t.} \quad (14b), \tag{38}$$

which is seen computationally as difficult as problem (17). However, following our previous developments [31]–[33], it can be efficiently handled via the following surrogate GM-SLNR problem:

$$\max_{\mathbf{v}^a, \mathbf{v}^e} \left[\prod_{\ell \in \mathcal{L}} \ln \left(1 + \gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e) \right) \right]^{1/L} \quad \text{s.t.} \quad (14b), \tag{39}$$

i.e. maximizing the smooth objective function in (39) enhances the nonsmooth objective in (38). The aim of this section is to provide a scalable algorithm for its computation. Initializing $(v^{a,(0)}, v^{e,(0)})$ that is feasible for (14b), for $\nu = 1, \ldots,$ we let $(v^{a,(\nu)}, v^{e,(\nu)})$ be a feasible point of (14b) that is found from the $(\nu - 1)$-st iteration. Similarly to [27], [33], the ν-th iteration to generate $(v^{a,(\nu+1)}, v^{e,(\nu+1)})$ is based on the following problem:

$$\max_{\mathbf{v}^a, \mathbf{v}^e} \Gamma^{(\kappa)}(\mathbf{v}^a, \mathbf{v}^e) \triangleq \sum_{\ell \in \mathcal{L}} \lambda_\ell^{(\nu)} \ln \left[1 + \gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e) \right] \text{ s.t. } (14b), \tag{40}$$

for[2]

$$\lambda_\ell^{(\nu)} \triangleq \frac{\max_{\ell' \in \mathcal{L}} \ln \left[1 + \gamma_{\ell'}(v^{a,(\nu)}, v^{e,(\nu)}) \right]}{\ln \left[1 + \gamma_\ell(v^{a,(\nu)}, v^{e,(\nu)}) \right]}, \ell \in \mathcal{L}. \tag{41}$$

---

[1]Given the beamformers, the ergodic rate function defined by (12) can be calculated based on [10].

[2]$\left[\prod_{\ell \in \mathcal{L}} \ln \left(1 + \gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e) \right) \right]^{1/L} = \min_{\lambda_\ell > 0, \prod_{\ell=1}^{L} \lambda_\ell = 1}$
$\left[\sum_{\ell=1}^{L} \lambda_\ell \ln \left[1 + \gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e) \right] \right]$, which is achieved at
$\lambda_\ell = \frac{\max_{\ell' \in \mathcal{L}} \ln \left[1 + \gamma_{\ell'}(v^{a,(\nu)}, v^{e,(\nu)}) \right]}{\ln \left[1 + \gamma_\ell(v^{a,(\nu)}, v^{e,(\nu)}) \right]}, \ell \in \mathcal{L}.$

A. Azimuth beamforming optimization

The alternating optimization on $\mathbf{v}^a$ is based on optimizing $\mathbf{v}^a$ with $\mathbf{v}^e$ fixed at $v^{e,(\nu)}$, which is based on the following problem

$$\max_{\mathbf{v}^a} \Gamma^{(\kappa)}(\mathbf{v}^a, v^{e,(\nu)}) = \sum_{\ell \in \mathcal{L}} \lambda_\ell^{(\nu)} \ln \left[1 + \gamma_\ell^a(\mathbf{v}_\ell^a) \right] \text{ s.t. } (18), \tag{42}$$

with $\gamma_\ell^a(\mathbf{v}_\ell^a)$ defined from (22).

The following minorant of $\Gamma^{(\kappa)}(\mathbf{v}^a, v^{e,(\nu)})$ at $v^{a,(\nu)}$ is obtained by applying the inequality (82) for $(\bar{\chi}, \bar{x}) = (\bar{\chi}_\ell^a, \bar{x}_\ell^a) \triangleq \left(\sqrt{\widetilde{\mathcal{R}}_\ell} \Psi_\ell^{e,(\nu)} v_\ell^{a,(\nu)}, \sum_{\ell' \in \mathcal{L} \setminus \{\ell\}} || \sqrt{\widetilde{\mathcal{R}}_{\ell'}} \Psi_\ell^{e,(\nu)} v_\ell^{a,(\nu)} ||^2 + \sigma \right)$:

$$\Gamma^{a,(\kappa)}(\mathbf{v}^a) \triangleq \sum_{\ell \in \mathcal{L}} \lambda_\ell^{(\nu)} \left(a_\ell^{a,(\nu)} + 2\Re\{b_\ell^{a,(\nu)} \mathbf{v}_\ell^a\} - c_\ell^{a,(\nu)} \sum_{\ell' \in \mathcal{L}} || \sqrt{\widetilde{\mathcal{R}}_{\ell'}} \Psi_\ell^{e,(\nu)} \mathbf{v}_\ell^a ||^2 \right), \tag{43}$$

where

$$a_\ell^{a,(\nu)} \triangleq \alpha(\bar{\chi}_\ell^a, \bar{x}_\ell^a), \ b_\ell^{a,(\nu)} \triangleq \frac{(\bar{\chi}_\ell^a)^H \sqrt{\widetilde{\mathcal{R}}_\ell} \Psi_\ell^{e,(\nu)}}{\bar{x}_\ell^a},$$
$$c_\ell^{a,(\nu)} \triangleq \psi(\bar{\chi}_\ell^a, \bar{x}_\ell^a), \tag{44}$$
$$\alpha(\bar{\chi}_\ell^a, \bar{x}_\ell^a) \triangleq \ln \left(1 + \frac{||\bar{\chi}_\ell^a||^2}{\bar{x}_\ell^a} \right) - \frac{||\bar{\chi}_\ell^a||^2}{\bar{x}_\ell^a}.$$

We thus solve the following problem to generate $v^{a,(\nu+1)}$:

$$\max_{\mathbf{v}^a} \Gamma^{a,(\kappa)}(\mathbf{v}^a) \quad \text{s.t.} \quad (18). \tag{45}$$

The optimal solution of this problem is given by the following closed-form under the definition of $\widetilde{\Psi}_\ell^{a,(\nu)} \triangleq c_\ell^{a,(\nu)} (\Psi_\ell^{e,(\nu)})^H \left(\sum_{\ell' \in \mathcal{L}} \widetilde{\mathcal{R}}_{\ell'} \right) \Psi_\ell^{e,(\nu)}$,

$$v_\ell^{a,(\nu+1)} = \begin{cases} (\widetilde{\Psi}_\ell^{a,(\nu)})^{-1} (b_\ell^{a,(\nu)})^H \\ \text{if } \sum_{\ell \in \mathcal{L}} \langle \mathcal{A}_\ell^{e,(\nu)}, [(\widetilde{\Psi}_\ell^{a,(\nu)})^{-1} (b_\ell^{a,(\nu)})^H]^2 \rangle \leq P, \\ \lambda_\ell^{(\nu)} \left(\lambda_\ell^{(\nu)} \widetilde{\Psi}_\ell^{a,(\nu)} + \mu I_{QM} \right)^{-1} (b_\ell^{a,(\nu)})^H \\ \text{otherwise,} \end{cases} \tag{46}$$

where $\mu > 0$ is found by bisection so that

$$\sum_{\ell \in \mathcal{L}} \langle \mathcal{A}_\ell^{e,(\nu)}, [\lambda_\ell^{(\nu)} \left(\lambda_\ell^{(\nu)} \widetilde{\Psi}_\ell^{a,(\nu)} + \mu I_{QM} \right)^{-1} (b_\ell^{a,(\nu)})^H]^2 \rangle = P.$$

B. Elevation beamforming optimization

The alternating optimization on $\mathbf{v}^e$ is based on optimizing $\mathbf{v}^e$ with $\mathbf{v}^a$ fixed at $v^{a,(\nu+1)}$, which is based on the following problem

$$\max_{\mathbf{v}^a} \Gamma^{(\kappa)}(v^{a,(\nu+1)}, \mathbf{v}^e) = \sum_{\ell \in \mathcal{L}} \lambda_\ell^{(\nu)} \ln \left[1 + \gamma_\ell^e(\mathbf{v}_\ell^e) \right] \text{ s.t. } (28), \tag{47}$$

with $\gamma_\ell^e(\mathbf{v}_\ell^e)$ defined by (32).

The following minorant of $\Gamma^{(\kappa)}(v^{a,(\nu+1)}, \mathbf{v}^e)$ at $v^{e,(\nu)}$ is obtained by applying the inequality (82) for $(\bar{\chi}, \bar{x}) = (\bar{\chi}_\ell^e, \bar{x}_\ell^e) \triangleq \left(\sqrt{\mathcal{R}_\ell} \Psi_\ell^{a,(\nu)} v_\ell^{e,(\nu)}, \sum_{\ell' \in \mathcal{L} \setminus \{\ell\}} || \sqrt{\mathcal{R}_{\ell'}} \Psi_\ell^{a,(\nu)} v_\ell^{e,(\nu)} ||^2 + \sigma \right)$,

$$\Gamma^{e,(\kappa)}(\mathbf{v}^e) \triangleq \sum_{\ell \in \mathcal{L}} \lambda_\ell^{(\nu)} \left(a_\ell^{e,(\nu)} + 2\Re\{b_\ell^{e,(\nu)} \mathbf{v}_\ell^e\} \right.$$

Algorithm 2 GM-SLNR based statistical beamforming algorithm

1: **Initialization:** Initialize a feasible $(v^{a,(0)}, v^{e,(0)})$ for (14b). Set $(v^{a,opt}, v^{e,opt}) = (v^{a,(0)}, v^{e,(0)})$ and $\text{rate}_{\min} = \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,opt}, v^{e,opt})$.
2: **Repeat until convergence:** Generate $v^{a,(\nu+1)}$ by the closed form (68). If $\min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu)}) > \text{rate}_{\min}$, update $(v^{a,opt}, v^{e,opt}) \leftarrow (v^{a,(\nu+1)}, v^{e,(\nu)})$ and $\text{rate}_{\min} \leftarrow \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu)})$. Generate $v^{e,(\nu+1)}$ by the closed form (51). If $\min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu+1)}) > \text{rate}_{\min}$, update $(v^{a,opt}, v^{e,opt}) \leftarrow (v^{a,(\nu+1)}, v^{e,(\nu+1)})$ and $\text{rate}_{\min} \leftarrow \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu+1)})$. Reset $\nu \leftarrow \nu + 1$.
3: **Output** $(v^{a,opt}, v^{e,opt})$ and $\text{rate}_{\min}$.

$$-c_\ell^{e,(\nu)} \sum_{\ell' \in \mathcal{L}} ||\sqrt{\mathcal{R}_{\ell'}} \Psi_\ell^{a,(\nu)} \mathbf{v}_\ell^e||^2 \Bigg), \qquad (48)$$

where

$$a_\ell^{e,(\nu)} \triangleq \alpha(\bar{\chi}_\ell^e, \bar{x}_\ell^e), \ b_\ell^{e,(\nu)} \triangleq \frac{(\bar{\chi}_\ell^e)^H \sqrt{\mathcal{R}_\ell} \Psi_\ell^{a,(\nu)}}{\bar{x}_\ell^e}, \\ c_\ell^{e,(\nu)} \triangleq \psi(\bar{\chi}_\ell^e, \bar{x}_\ell^e). \qquad (49)$$

We thus solve the following problem to generate $v^{a,(\nu+1)}$:

$$\max_{\mathbf{v}^e} \Gamma^{e,(\kappa)}(\mathbf{v}^e) \quad \text{s.t.} \quad (28). \qquad (50)$$

The optimal solution of this problem is given by the following closed-form under the definition of $\widetilde{\Psi}_\ell^{e,(\nu)} \triangleq c_\ell^{e,(\nu)}(\Psi_\ell^{a,(\nu)})^H \left(\sum_{\ell' \in \mathcal{L}} \mathcal{R}_{\ell'} \right) \Psi_\ell^{a,(\nu)}$,

$$v_\ell^{e,(\nu+1)} = \begin{cases} (\widetilde{\Psi}_\ell^{e,(\nu)})^{-1}(b_\ell^{e,(\nu)})^H \\ \quad \text{if } \sum_{\ell \in \mathcal{L}} \langle \mathcal{A}_\ell^{a,(\nu)}, [(\widetilde{\Psi}_\ell^{e,(\nu)})^{-1}(b_\ell^{e,(\nu)})^H]^2 \rangle \leq P, \\ \lambda_\ell^{(\nu)} \left(\lambda_\ell^{(\nu)} \widetilde{\Psi}_\ell^{e,(\nu)} + \mu I_{QM} \right)^{-1} (b_\ell^{e,(\nu)})^H \\ \quad \text{otherwise}, \end{cases} \qquad (51)$$

where $\mu > 0$ is found by bisection so that

$$\sum_{\ell \in \mathcal{L}} \langle \mathcal{A}_\ell^{a,(\nu)}, [\lambda_\ell^{(\nu)} \left(\lambda_\ell^{(\nu)} \widetilde{\Psi}_\ell^{e,(\nu)} + \mu I_{QM} \right)^{-1} (b_\ell^{e,(\nu)})^H]^2 \rangle = P.$$

C. Scalable algorithm and its computational complexity

Algorithm 2 provides the pseudo-code of computing (39) and also for updating the beamformers achieving the best minimum ergodic rate. The computational complexity of each iteration is on the order of $\mathcal{O}(QM)$. Its convergence can be shown by following the proof provided in [27], [33].

IV. SLNR SOFT MAX-MIN OPTIMIZATION BASED BEAMFORMING

The SLNR max min problem (17) is also equivalent to the problem

$$\max_{\mathbf{v}^a, \mathbf{v}^e} \min_{\ell \in \mathcal{L}} \ln \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right) \quad \text{s.t.} \quad (14b), \qquad (52)$$

for any $c > 0$. Then we have

$$\min_{\ell \in \mathcal{L}} \ln \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right) = -\max_{\ell \in \mathcal{L}} \ln \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1} \qquad (53)$$

while

$$\max_{\ell \in \mathcal{L}} \ln \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1}$$
$$\leq \ln \frac{\sum_{\ell \in \mathcal{L}} \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1}}{L} + \ln L \qquad (54)$$
$$\leq \max_{\ell \in \mathcal{L}} \ln \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1} + \ln L. \qquad (55)$$

For small c, $\ln L$ is small compared to the nonsmooth function in the LHS of (54) and the RHS of (55), so we have

$$\max_{\ell \in \mathcal{L}} \ln \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1}$$
$$\approx \ln \frac{\sum_{\ell \in \mathcal{L}} \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1}}{L} \qquad (56)$$
$$= \ln \left(\sum_{\ell \in \mathcal{L}} \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1} \right) - \ln L. \qquad (57)$$

Therefore, we can approximate the problem (52) of nonsmooth function optimization by the following problem of smooth function optimization:

$$\max_{\mathbf{v}^a, \mathbf{v}^e} \left[-\ln \left(\sum_{\ell \in \mathcal{L}} \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1} \right) + \ln L \right] \text{ s.t. } (14b), \qquad (58)$$

which is equivalent to the problem

$$\min_{\mathbf{v}^a, \mathbf{v}^e} \tilde{\Gamma}(\mathbf{v}^a, \mathbf{v}^e) \triangleq \ln \left(\sum_{\ell \in \mathcal{L}} \left(1 + \frac{\gamma_\ell(\mathbf{v}_\ell^a, \mathbf{v}_\ell^e)}{c} \right)^{-1} \right) \text{ s.t. } (14b). \qquad (59)$$

In this paper, we refer to (58)/(59) as SLNR soft max-min optimization. This section aims for developing an algorithm of scalable complexity for solving (59) and optimizing the users' ergodic rates.

Use the initialization $(v^{a,(0)}, v^{e,(0)})$ that is feasible for (14b), for $\nu = 1, \ldots$, and let $(v^{a,(\nu)}, v^{e,(\nu)})$ be a feasible point of (14b) that is found from the $(\nu - 1)$-st iteration. Then we can carry out the ν-th iteration to generate $(v^{a,(\nu+1)}, v^{e,(\nu+1)})$.

A. Azimuth beamforming optimization

With $\gamma_\ell^a(\mathbf{v}_\ell^a)$ defined from (22), the alternating optimization in $\mathbf{v}^a$ is based on the problem

$$\min_{\mathbf{v}^a} \tilde{\Gamma}^a(\mathbf{v}^a) \quad \text{s.t.} \quad (18), \qquad (60)$$

where we have $\tilde{\Gamma}^a(\mathbf{v}^a)$ given by (61) shown at the top of next page.

The following tight majorant of $\tilde{\Gamma}^a(\mathbf{v}^a)$ at $v^{a,(\nu)}$ is obtained by applying the inequality (91) for $(\bar{\chi}_\ell, \bar{x}_\ell) = \left(\sqrt{\widetilde{\mathcal{R}}_\ell} \Psi_\ell^{e,(\nu)} v_\ell^{a,(\nu)}, \sum_{\ell' \in \mathcal{L} \setminus \{\ell\}} ||\sqrt{\widetilde{\mathcal{R}}_{\ell'}} \Psi_\ell^{e,(\nu)} v_\ell^{a,(\nu)}||^2 + \sigma \right)$:

$$\tilde{\Gamma}^{a,(\kappa)}(\mathbf{v}^a) \triangleq a_1^{(\nu)} - 2 \sum_{\ell \in \mathcal{L}} \Re\{b_{1,\ell}^{(\nu)} \mathbf{v}_\ell^a\}$$

$$\tilde{\Gamma}^a(\mathbf{v}^a) \triangleq \tilde{\Gamma}(\mathbf{v}^a, v^{e,(\nu)})$$

$$= \ln\left(\sum_{\ell \in \mathcal{L}} \left(1 + \frac{\gamma_\ell^a(\mathbf{v}_\ell^a)}{c}\right)^{-1}\right)$$

$$= \ln\left(\sum_{\ell \in \mathcal{L}}\left(1 - \frac{||\sqrt{\widetilde{\mathcal{R}}_\ell}\Psi_\ell^{e,(\nu)}\mathbf{v}_\ell^a||^2}{||\sqrt{\widetilde{\mathcal{R}}_\ell}\Psi_\ell^{e,(\nu)}\mathbf{v}_\ell^a||^2 + c\left(\sum_{\ell' \in \mathcal{L}\setminus\{\ell\}} ||\sqrt{\widetilde{\mathcal{R}}_{\ell'}}\Psi_\ell^{e,(\nu)}\mathbf{v}_\ell^a||^2 + \sigma\right)}\right)\right). \tag{61}$$

$$+\sum_{\ell \in \mathcal{L}} c_{1,\ell}^{(\nu)}\left(||\sqrt{\widetilde{\mathcal{R}}_\ell}\Psi_\ell^{e,(\nu)}\mathbf{v}_\ell^a||^2\right.$$

$$\left. +c\sum_{\ell' \in \mathcal{L}\setminus\{\ell\}} ||\sqrt{\widetilde{\mathcal{R}}_{\ell'}}\Psi_\ell^{e,(\nu)}\mathbf{v}_\ell^a||^2\right) \tag{62}$$

$$=a_1^{(\nu)} - 2\sum_{\ell \in \mathcal{L}} \Re\{b_{1,\ell}^{(\nu)}\mathbf{v}_\ell^a\} + \sum_{\ell \in \mathcal{L}}(\mathbf{v}_\ell^a)^H\widetilde{\Psi}_\ell^{a,(\nu)}\mathbf{v}_\ell^a, \tag{63}$$

where we have

$$a_1^{(\nu)} \triangleq \tilde{\Gamma}^a(v^{a,(\nu)})$$

$$+ \frac{1}{\chi_1^{(\nu)}}\sum_{\ell \in \mathcal{L}}\left(\frac{||\bar{\chi}_\ell||^2}{c\bar{x}_\ell + ||\bar{\chi}_\ell||^2} + \frac{||\bar{\chi}_\ell||^2}{(c\bar{x}_\ell + ||\bar{\chi}_\ell||^2)^2}c\sigma\right), \tag{64a}$$

$$b_{1,\ell}^{(\nu)} \triangleq \frac{1}{\chi_1^{(\nu)}}\frac{(\sqrt{\widetilde{\mathcal{R}}_\ell}\Psi_\ell^{e,(\nu)}v_\ell^{a,(\nu)})^H\sqrt{\widetilde{\mathcal{R}}_\ell}\Psi_\ell^{e,(\nu)}}{c\bar{x}_\ell + ||\bar{\chi}_\ell||^2}, \ell \in \mathcal{L}, \tag{64b}$$

$$c_{1,\ell}^{(\nu)} \triangleq \frac{1}{\chi_1^{(\nu)}}\frac{||\bar{\chi}_\ell||^2}{(c\bar{x}_\ell + ||\bar{\chi}_\ell||^2)^2}, \ell \in \mathcal{L}, \tag{64c}$$

$$\chi_1^{(\nu)} \triangleq \sum_{\ell \in \mathcal{L}}\left(1 + \frac{\gamma_\ell^a(v_\ell^{a,(\nu)})}{c}\right)^{-1}, \tag{64d}$$

and

$$\widetilde{\Psi}_\ell^{a,(\nu)} \triangleq c_{1,\ell}^{(\nu)}\left((\Psi_\ell^{e,(\nu)})^H\widetilde{\mathcal{R}}_\ell\Psi_\ell^{e,(\nu)}\right.$$

$$\left. +c\sum_{\ell' \in \mathcal{L}\setminus\{\ell\}}(\Psi_\ell^{e,(\nu)})^H\widetilde{\mathcal{R}}_{\ell'}\Psi_\ell^{e,(\nu)}\right). \tag{65}$$

We thus solve the following problem

$$\min_{\mathbf{v}^a}\tilde{\Gamma}^{a,(\kappa)}(\mathbf{v}^a) \quad \text{s.t.} \quad (18). \tag{66}$$

to generate $v^{a,(\nu+1)}$ so that

$$\tilde{\Gamma}(v^{a,(\nu+1)}, v^{e,(\nu)}) < \tilde{\Gamma}(v^{a,(\nu)}, v^{e,(\nu)}). \tag{67}$$

The optimal solution of this problem is given by the following closed-form:

$$v_\ell^{a,(\nu+1)} = \begin{cases} (\widetilde{\Psi}_\ell^{a,(\nu)})^{-1}(b_\ell^{a,(\nu)})^H \\ \quad \text{if } \sum_{\ell \in \mathcal{L}}\langle\mathcal{A}_\ell^{e,(\nu)}, [(\widetilde{\Psi}_\ell^{a,(\nu)})^{-1}(b_\ell^{a,(\nu)})^H]^2\rangle \leq P, \\ \left(\widetilde{\Psi}_\ell^{a,(\nu)} + \mu I_{QM}\right)^{-1}(b_\ell^{a,(\nu)})^H \\ \quad \text{otherwise,} \end{cases} \tag{68}$$

where $\mu > 0$ is found by bisection so that $\sum_{\ell \in \mathcal{L}}\langle\mathcal{A}_\ell^{e,(\nu)}, [\left(\widetilde{\Psi}_\ell^{a,(\nu)} + \mu I_{QM}\right)^{-1}(b_\ell^{a,(\nu)})^H]^2\rangle = P$.

B. Elevation beamforming optimization

With $\gamma_\ell^e(\mathbf{v}_\ell^e)$ defined by (32), the alternating optimization in $\mathbf{v}^e$ is based on the problem

$$\min_{\mathbf{v}^e}\tilde{\Gamma}^e(\mathbf{v}^e) \quad \text{s.t.} \quad (28), \tag{69}$$

where $\tilde{\Gamma}^e(\mathbf{v}^e)$ is given by (70) shown at the top of next page.

The following tight majorant of $\tilde{\Gamma}^e(\mathbf{v}^e)$ at $v^{a,(\nu)}$ is obtained by applying the inequality (91) for $(\bar{\chi}_\ell, \bar{x}_\ell) = \left(\sqrt{\mathcal{R}_\ell}\Psi_\ell^{a,(\nu)}v_\ell^{e,(\nu)}, \sum_{\ell' \in \mathcal{L}\setminus\{\ell\}}||\sqrt{\mathcal{R}_\ell}\Psi_\ell^{e,(\nu)}v_\ell^{e,(\nu)}||^2 + \sigma\right)$:

$$\tilde{\Gamma}^{e,(\kappa)}(\mathbf{v}^e) \triangleq a_2^{(\nu)} - 2\sum_{\ell \in \mathcal{L}}\Re\{b_{2,\ell}^{(\nu)}\mathbf{v}_\ell^e\}$$

$$+ \sum_{\ell \in \mathcal{L}} c_{2,\ell}^{(\nu)}\left(||\sqrt{\mathcal{R}_\ell}\Psi_\ell^{a,(\nu)}\mathbf{v}_\ell^e||^2\right.$$

$$\left. +c\sum_{\ell' \in \mathcal{L}\setminus\{\ell\}}||\sqrt{\mathcal{R}_{\ell'}}\Psi_\ell^{a,(\nu)}\mathbf{v}_\ell^e||^2\right) \tag{71}$$

$$=a_2^{(\nu)} - 2\sum_{\ell \in \mathcal{L}}\Re\{b_{2,\ell}^{(\nu)}\mathbf{v}_\ell^e\} + \sum_{\ell \in \mathcal{L}}(\mathbf{v}_\ell^e)^H\widetilde{\Psi}_\ell^{e,(\nu)}\mathbf{v}_\ell^e, \tag{72}$$

where we have

$$a_2^{(\nu)} \triangleq \tilde{\Gamma}^e(v^{e,(\nu)})$$

$$+ \frac{1}{\chi_2^{(\nu)}}\sum_{\ell \in \mathcal{L}}\left(\frac{||\bar{\chi}_\ell||^2}{c\bar{x}_\ell + ||\bar{\chi}_\ell||^2} + \frac{||\bar{\chi}_\ell||^2}{(c\bar{x}_\ell + ||\bar{\chi}_\ell||^2)^2}c\sigma\right), \tag{73a}$$

$$b_{2,\ell}^{(\nu)} \triangleq \frac{1}{\chi_2^{(\nu)}}\frac{(\sqrt{\mathcal{R}_\ell}\Psi_\ell^{a,(\nu)}v_\ell^{e,(\nu)})^H\sqrt{\mathcal{R}_\ell}\Psi_\ell^{a,(\nu)}}{c\bar{x}_\ell + ||\bar{\chi}_\ell||^2}, \ell \in \mathcal{L}, \tag{73b}$$

$$c_{2,\ell}^{(\nu)} \triangleq \frac{1}{\chi_2^{(\nu)}}\frac{||\bar{\chi}_\ell||^2}{(c\bar{x}_\ell + ||\bar{\chi}_\ell||^2)^2}, \ell \in \mathcal{L}, \tag{73c}$$

$$\chi_2^{(\nu)} \triangleq \sum_{\ell \in \mathcal{L}}\left(1 + \frac{\gamma_\ell^e(v_\ell^{e,(\nu)})}{c}\right)^{-1}, \tag{73d}$$

and

$$\widetilde{\Psi}_\ell^{e,(\nu)} \triangleq c_{2,\ell}^{(\nu)}\left((\Psi_\ell^{a,(\nu)})^H\mathcal{R}_\ell\Psi_\ell^{a,(\nu)}\right.$$

$$\left. +c\sum_{\ell' \in \mathcal{L}\setminus\{\ell\}}(\Psi_\ell^{a,(\nu)})^H\mathcal{R}_{\ell'}\Psi_\ell^{a,(\nu)}\right). \tag{74}$$

We thus solve the following problem

$$\min_{\mathbf{v}^e}\tilde{\Gamma}^{e,(\kappa)}(\mathbf{v}^e) \quad \text{s.t.} \quad (28), \tag{75}$$

$$\tilde{\Gamma}^e(\mathbf{v}^e) \triangleq \tilde{\Gamma}(v^{a,(\nu+1)}, \mathbf{v}^e)$$

$$= \ln\left(\sum_{\ell \in \mathcal{L}} \left(1 + \frac{\gamma_\ell^e(\mathbf{v}_\ell^e)}{c}\right)^{-1}\right)$$

$$= \ln\left(\sum_{\ell \in \mathcal{L}} \left(1 - \frac{||\sqrt{\mathcal{R}_\ell}\Psi_\ell^{a,(\nu)}\mathbf{v}_\ell^e||^2}{||\sqrt{\mathcal{R}_\ell}\Psi_\ell^{a,(\nu)}\mathbf{v}_\ell^e||^2 + c\left(\sum_{\ell' \in \mathcal{L}\setminus\{\ell\}} ||\sqrt{\mathcal{R}_{\ell'}}\Psi_\ell^{a,(\nu)}\mathbf{v}_\ell^e||^2 + \sigma\right)}\right)\right). \tag{70}$$

Algorithm 3 SLNR Soft max-min based statistical beamforming algorithm

1: **Initialization:** Initialize a feasible $(v^{a,(0)}, w^{e,(0)})$ for (28). Set $(v^{a,opt}, v^{e,opt}) = (v^{a,(0)}, v^{e,(0)})$ and $\text{rate}_{\min} = \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,opt}, v^{e,opt})$.

2: **Repeat until convergence of the objective function in (59):** Generate $v^{a,(\nu+1)}$ by (68). If $\min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu)}) > \text{rate}_{\min}$, update $(v^{a,opt}, v^{e,opt}) \leftarrow (v^{a,(\nu+1)}, v^{e,(\nu)})$ and $\text{rate}_{\min} \leftarrow \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu)})$. Generate $v^{e,(\nu+1)}$ by (77). If $\min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu+1)}) > \text{rate}_{\min}$, update $(v^{a,opt}, v^{e,opt}) \leftarrow (v^{a,(\nu+1)}, v^{e,(\nu+1)})$ and $\text{rate}_{\min} \leftarrow \min_{\ell \in \mathcal{L}} \rho_\ell(v^{a,(\nu+1)}, v^{e,(\nu+1)})$. Reset $\nu \leftarrow \nu + 1$.

3: **Output** $(v^{a,opt}, v^{e,opt})$ and $\text{rate}_{\min}$.

---

to generate $v^{e,(\nu+1)}$ so that

$$\tilde{\Gamma}(v^{a,(\nu+1)}, v^{e,(\nu+1)}) < \tilde{\Gamma}(v^{a,(\nu+1)}, v^{e,(\nu)}). \tag{76}$$

The optimal solution of this problem is given by the following closed-form:

$$v_\ell^{e,(\nu+1)} = \begin{cases} (\widetilde{\Psi}_\ell^{e,(\nu)})^{-1}(b_\ell^{e,(\nu)})^H \\ \quad \text{if } \sum_{\ell \in \mathcal{L}} \langle \mathcal{A}_\ell^{a,(\nu)}, [(\widetilde{\Psi}_\ell^{e,(\nu)})^{-1}(b_\ell^{e,(\nu)})^H]^2 \rangle \leq P, \\ \left(\widetilde{\Psi}_\ell^{e,(\nu)} + \mu I_{QM}\right)^{-1}(b_\ell^{e,(\nu)})^H \\ \quad \text{otherwise}, \end{cases} \tag{77}$$

where $\mu > 0$ is found by bisection so that $\sum_{\ell \in \mathcal{L}} \langle \mathcal{A}_\ell^{a,(\nu)}, [\left(\widetilde{\Psi}_\ell^{e,(\nu)} + \mu I_{QM}\right)^{-1}(b_\ell^{e,(\nu)})^H]^2 \rangle = P$.

C. Scalable algorithm and its computational complexity

Algorithm 3 provides the pseudo code for solving problem (59) by generating the next iterative feasible point $(v^{a,(\nu+1)}, v^{e,(\nu+1)})$ by the closed forms (68) and (77), which according to (67) and (76) satisfies $\tilde{\Gamma}(v^{a,(\nu+1)}, v^{e,(\nu+1)}) < \tilde{\Gamma}(v^{a,(\nu)}, v^{e,(\nu)})$. The sequence $\{(v^{a,(\nu)}, v^{e,(\nu)})\}$ thus cannot be cyclic and thus it is convergent. Whenever generating a new feasible point for (59), Algorithm 3 also updates the incumbent beamformers achieving the best minimum user rate. The computational complexity of each iteration is on the order of $\mathcal{O}(QM)$. Given the cubic-time computational complexity of the convex-solver based Algorithm 1, both Algorithm 2 and Algorithm 3 have scalable complexities, rendering them computationally efficient.

V. SIMULATION RESULTS

A $Q \times Q$-URA BS is positioned at the center of a cell having a 200-meter radius. Within the cell, there are $L = 10$ DUs that are randomly placed. The BS antennas have a height of 10 meters, while the DUs have a height of 1.5 meters. The path-loss and shadow-fading of DU ℓ experienced at a distance d_ℓ from the BS is given by [34] $\rho_\ell = 28.8 + 35.3 \log 10(d_\ell) + \xi_\ell$ (in dB), where $\xi_\ell \sim \mathcal{N}(0, \sigma_{\text{sf}}^2)$ represents the shadow-fading coefficient with $\sigma_{\text{sf}} = 7.8$. The carrier frequency is set to 2 GHz, while the bandwidth is set to 10 MHz. The power density of the background noise is set to -174 dBm/Hz. A convergence tolerance of 10^{-3} is specified for the algorithms with respect to the SLNR performance. We refer to the m-th vertical and n-th horizontal direction of the URA as the (m, n)-th antenna element. In accordance with [25], the vertical correlation matrix $R_{v,\ell}$ is represented by

$$[R_{v,\ell}]_{m,p} = e^{j\pi(p-m)\cos\beta_\ell} e^{-\frac{1}{2}[\sigma_\beta \pi(p-m)\sin\beta_\ell]^2}, \tag{78}$$

while the horizontal correlation matrix is formulated as

$$[R_{h,\ell}]_{n,q} = \frac{1}{\sqrt{\gamma_3}} e^{-\frac{\gamma_2^2 \cos^2\alpha_\ell}{2\gamma_3}} e^{j\frac{\gamma_1 \cos\alpha}{\gamma_3}} e^{-\frac{(\gamma_1 \sigma_\alpha \sin\alpha_\ell)^2}{2\gamma_3}}, \tag{79}$$

where $\gamma_1 = \pi(q - n)\sin\beta_\ell$, $\gamma_2 = \sigma_\beta \pi(q - n)\cos\beta_\ell$, $\gamma_3 = \gamma_2^2 \sigma_\alpha^2 \sin^2\alpha_\ell + 1$. The azimuth and elevation angles of DU ℓ, denoted as α_ℓ and β_ℓ, respectively, are determined based on the relative location of the DU with respect to the BS. Furthermore, σ_α and σ_β denote the angular spreads in the azimuth and elevation domains, respectively, and both are set to $5°$.

We will compare the minimum ergodic rate achieved by our optimized beamforming with the following baseline uniform power beamformers given by [16]

$$\sqrt{P/K} w_{u,\ell}, \tag{80}$$

where $w_{u,\ell} \in \mathbb{C}^{Q^2}$ is the eigenvalue vector corresponding to the largest eigenvalue of the matrix

$$\left(\sum_{\ell' \in \mathcal{L}\setminus\{\ell\}} \mathcal{R}_{\ell'} + (K\sigma/P)I_{Q^2}\right)^{-1} \mathcal{R}_\ell$$

with $\mathcal{R}_\ell$ defined from (4).

Once a better feasible point $(v^{a,(\nu+1)}, v^{e,(\nu+1)})$ is obtained in the ν-th iteration, which maximizes the minimum SLNR, or in its GM form or soft max-min form, we define $V_\ell^{(\nu+1)}$ by (9) and then $v_\ell^{(\nu+1)} \triangleq \text{vec}(V_\ell^{(\nu+1)})$, and $W^{(\nu+1)} \triangleq \begin{bmatrix} v_1^{(\nu+1)} & \cdots & v_L^{(\nu+1)} \end{bmatrix}$. Furthermore, we define $S_\ell = (W^{(\nu+1)})^H \mathcal{R}_\ell W^{(\nu+1)}$ and then $\tilde{S}_\ell$ by removing the ℓ-th
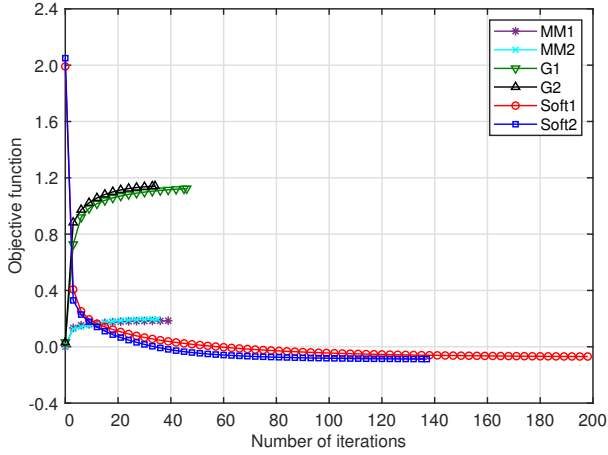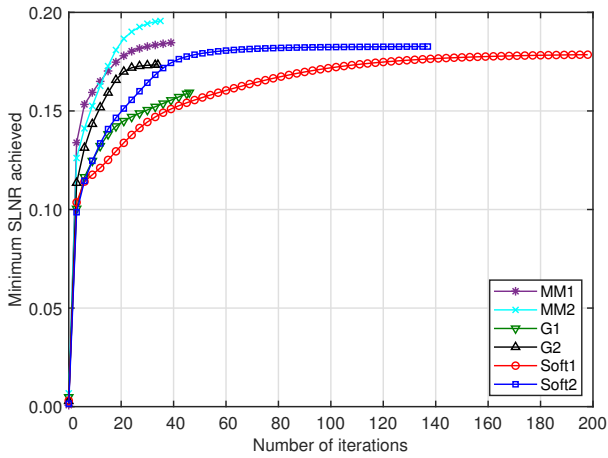
Fig. 2: The convergence of the algorithms.

Fig. 3: The convergence of the minimum SLNR.

Fig. 4: The convergence of the minimum ergodic rate.

TABLE II: Minimum ergodic rate (Mbps) vs. coefficient c for the soft max-min algorithm.

	$c = 1$	$c = 0.1$	$c = 0.01$
Soft1	2.97	4.95	4.27
Soft2	2.82	5.34	5.26

column and the ℓ-th row of S_ℓ. The ergodic rate is calculated according to [10, Th.1] as follows:

$$\mathbb{E}\{\ln(1 + \text{SINR}_\ell)\} = \sum_{m=1}^{\text{rank}(S_\ell)} \frac{e^{\frac{\sigma}{\mu_m}} E_1\left(\frac{\sigma}{\mu_m}\right)}{\prod_{l \neq m}\left(1 - \frac{\mu_l}{\mu_m}\right)}$$
$$- \sum_{m=1}^{\text{rank}(\tilde{S}_\ell)} \frac{e^{\frac{\sigma}{\lambda_m}} E_1\left(\frac{\sigma}{\lambda_m}\right)}{\prod_{l \neq m}\left(1 - \frac{\lambda_l}{\lambda_m}\right)}, \quad (81)$$

where $\mu_1, \ldots, \mu_{\text{rank}(S_\ell)}$ and $\lambda_1, \ldots, \lambda_{\text{rank}(\tilde{S}_\ell)}$ are the distinct non-zero eigenvalues of S_ℓ and $\tilde{S}_\ell$, respectively, while $E_1(x) = \int_1^\infty e^{-xu}/u \, du$ is the exponential integral.

We use the following legends to specify the proposed implementations: "MM1/MM2" refer to the CP-based algorithms with $M = 1/M = 2$; "G1/G2" refer to the GM-based algorithms with $M = 1/M = 2$; "Soft1/Soft2" refer to the soft max-min based algorithms with $M = 1/M = 2$; "EQ" refers to the baseline algorithm that uses uniform power beamformers.

Fig. 2 depicts the convergence of the algorithms, while Fig. 3 and Fig. 4 portray the convergence of the minimum SLNR and the minimum ergodic rate achieved, respectively. The transmit power budget P is maintained at 24 dBm, and the num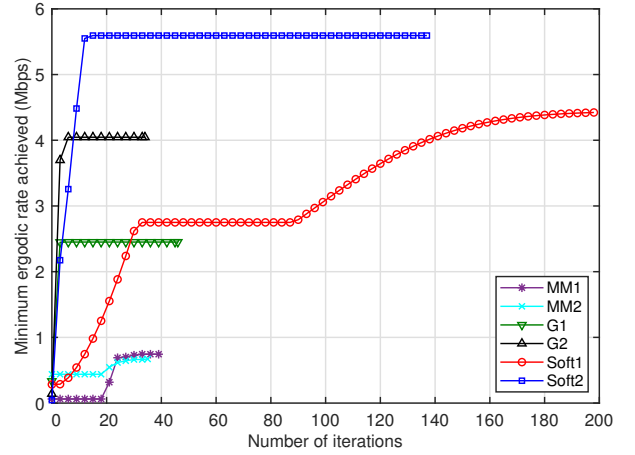ber of BS antennas is kept fixed at 64. The coefficient c used in soft max-min based algorithms is set to 0.1. It is worth noting that we use the minimization form (59) for the objective function of the soft max-min based algorithm. Therefore, in Fig. 2, the objective functions of Soft1/Soft2 show a gradually decreasing trend until convergence. A noteworthy observation from Fig. 3 and Fig. 4 is that the minimum SLNR exhibits a gradual increase with each iteration, while the minimum ergodic rate does not follow this pattern. This implies that the beamformers that yield the highest SLNR values may not necessarily lead to the highest minimum ergodic rate.

Importantly, the choice of the coefficient c in the soft max-min algorithm should be determined based on specific factors, such as the dimensions of the beamformers and the transmit power budget P. We present Table II to illustrate the influence of different c values on the minimum ergodic rate achieved (in Mbps). The results are obtained at a transmit power P of 24 dBm and 64 BS antennas. We can observe that setting c to 0.1 yields the highest minimum ergodic rate.

We evaluate the performance of the proposed algorithms upon varying the numbers of BS antennas. Fig. 5 plots the minimum SLNR achieved versus the number of BS antennas at a power of $P = 24$ dBm. The convex-solver based algorithms achieve the highest minimum SLNR, followed by the soft max-min based algorithms, and then the GM-based algorithms. The minimum SLNR obtained by the baseline algorithm, which calculates the minimum ergodic rate using the uniform power beamformers, is significantly lower than that of the other algorithms. Moreover, using the sum of two outer products yields better results than using the sum of just one outer product.

Fig. 6 shows the minimum ergodic rate achieved versus the number of BS antennas at the same power of $P = 24$ dBm. It is worth noting that the minimum ergodic rate is determined by selecting the maximum value obtained during
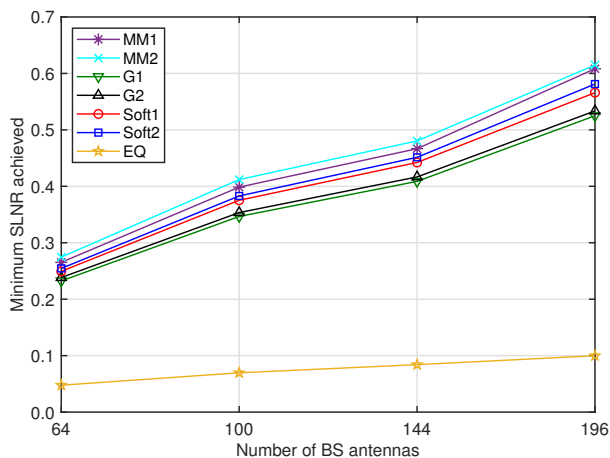
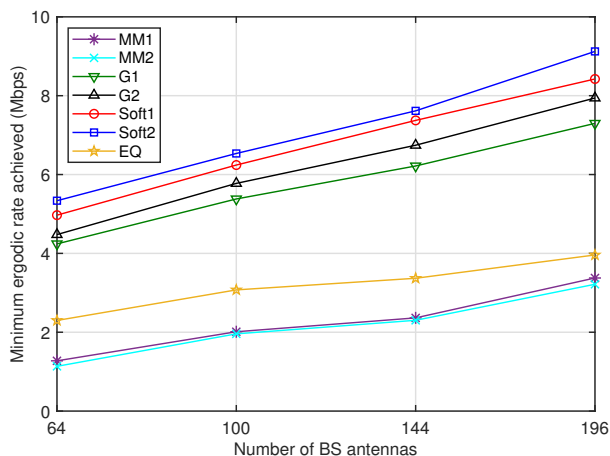Fig. 5: The minimum SLNR achieved vs. the number of BS antennas.

Fig. 7: Ergodic rate distribution for three algorithms associated with $M = 1$.

Fig. 6: The minimum ergodic rate achieved vs. the number of BS antennas.

Fig. 8: Ergodic rate distribution for three algorithms associated with $M = 2$ and the baseline algorithm.

TABLE III: The minimum-to-maximum ergodic rate ratio vs. the number of BS antennas

	64 antennas	100 antennas	144 antennas	196 antennas
MM1	0.0519	0.0760	0.0810	0.1014
MM2	0.0536	0.0780	0.0838	0.1039
G1	0.1106	0.1421	0.1510	0.1589
G2	0.1303	0.1494	0.1713	0.1759
Soft1	0.1410	0.1675	0.1859	0.1915
Soft2	0.1794	0.1861	0.2007	0.2264
EQ	0.0281	0.0345	0.0355	0.0386

the iterative optimization process, rather than the one calculated by the beamformers that yields the highest SLNR. The results demonstrate that both the GM-based and soft max-min based algorithms outperform the baseline algorithm, which calculates the minimum ergodic rate using the uniform power beamformers. Additionally, the soft max-min based algorithms achieve a higher minimum ergodic rate than the GM-based algorithms. We can also find that the sum of two outer products performs better than the sum of one outer product in terms of the ergodic rate. However, the minimum ergodic rates obtained by the convex-solver based algorithms are significantly lower than those obtained by the other algorithms, including the baseline algorithm. The baseline algorithm performs better than expected, probably due to its connection with regularized zero forcing beamforming under the equal power allocation [35], [36].

Fig. 7 portrays the users' ergodic rate distribution pattern for algorithms associated with $M = 1$, while Fig. 8 portrays this pattern for algorithms associated with $M = 2$ and the baseline algorithm. This simulation uses 196 antennas at the BS and a power of $P = 24$ dBm. Similar to Fig. 6, these
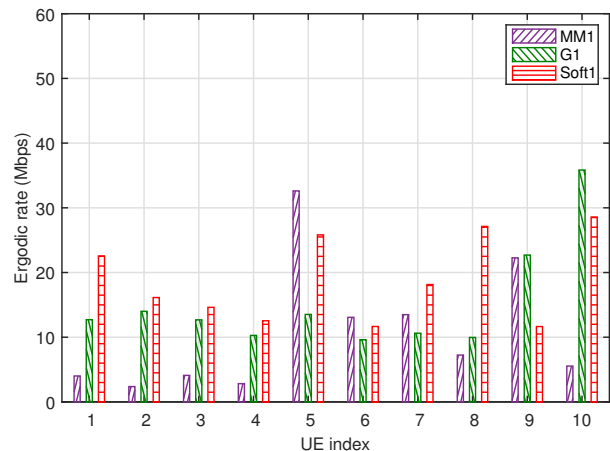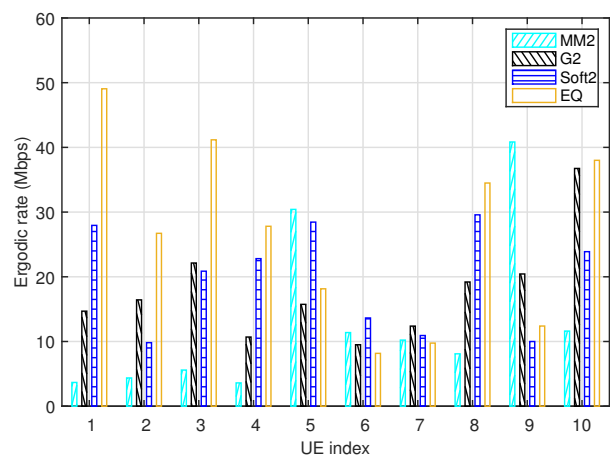
figures illustrate that both the GM-based and the soft max-min based algorithms G2 and Soft2 succeed in enhancing the ergodic rates of all DUs, while also ensuring their minimum value. Furthermore, the baseline algorithm EQ performs better than the SLNR max-min based M2, probably owing to its association with regularized zero forcing beamforming under equal power allocation [35], [36].

For more qualitative analysis, Table III and Table IV present the minimum-to-maximum ergodic rate ratio and Jain's fairness index [37] of the ergodic rate, respectively, at $P = 24$ dBm. These two metrics are used for evaluating the distribution pattern of the ergodic rate obtained by the proposed

TABLE IV: Jain's fairness index of the ergodic rate vs. the number of BS antennas

	64 antennas	100 antennas	144 antennas	196 antennas
MM1	0.4901	0.5528	0.5690	0.5849
MM2	0.4807	0.5613	0.5729	0.5833
G1	0.6303	0.6875	0.7087	0.7186
G2	0.6706	0.7007	0.7250	0.7261
Soft1	0.6942	0.7211	0.7371	0.7399
Soft2	0.7138	0.7286	0.7456	0.7635
EQ	0.4930	0.5236	0.5478	0.5573
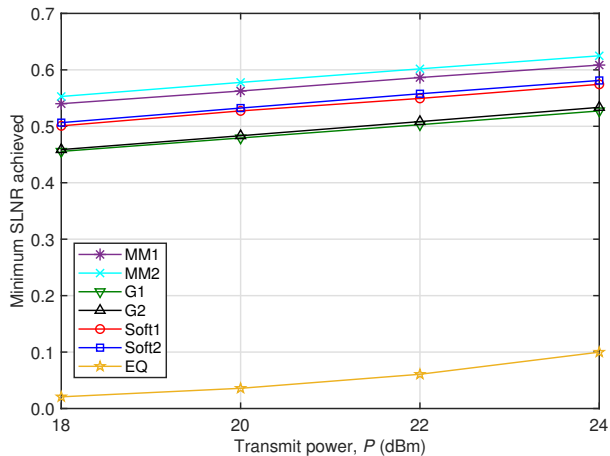
Fig. 9: The minimum SLNR achieved vs. P using 196 antennas at the BS.

Fig. 10: The minimum ergodic rate achieved vs. P using 196 antennas at the BS.

algorithms. As the number of BS antennas increases from 64 to 196, both the minimum-to-maximum ergodic rate ratio and Jain's fairness index increase for our proposed algorithms, which illustrates that employing more antennas for beamforming flexibility helps achieve more uniformly distributed ergodic user rates. Moreover, the soft max-min based algorithms yield the highest minimum-to-maximum ergodic rate ratio and Jain's fairness index, demonstrating their ability to achieve a more fair ergodic rate distribution. Additionally, both the GM-based and soft max-min based algorithms outperform the convex-solver based algorithms and the baseline algorithm in terms of rate fairness.

In addition, we investigate the effect of the transmit power budget P on the minimum SLNR and minimum ergodic rate achieved, as depicted in Fig. 9 and Fig. 10, for a fixed number of 196 BS antennas. As previous noted in Fig. 4, the highest SLNR value does not necessarily guarantee the highest minimum ergodic rate. Observe that the minimum SLNR increases with P, but the minimum ergodic rate does not exhibit the same trend. However, similar to the observations gleaned from Fig. 6, the soft max-min based and GM-based algorithms perform better than the baseline algorithm, while the convex-solver based algorithms yield the lowest minimum ergodic rate.

VI. CONCLUSIONS

In order to investigate the long-term ergodic rate distributions provided by full-dimensional m-MIMO systems, we have proposed different statistic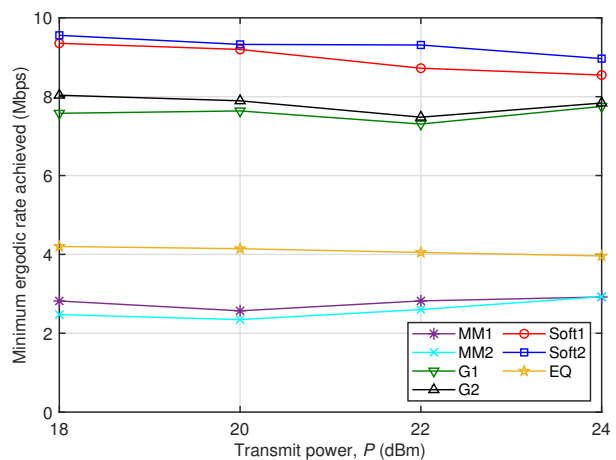al beamforming schemes, which are based on the optimization of the users' SLNRs. Accordingly, the paper has developed computationally efficient algorithms for their solutions, which are also used for generating feasible beamforming in updating the incumbent ergodic minimum rate. Furthermore, we have provided simulations to show the impact of SLNR metric on the ergodic rates. An extension to statistical hybrid beamforming for mmWave communications is currently under our further investigation.

APPENDIX: MATHEMATICAL INGREDIENTS

The following inequalities hold true for all $\boldsymbol{\chi} \in \mathbb{C}^N$, $\boldsymbol{x} > 0$ and $\bar{\chi} \in \mathbb{C}^N$, $\bar{x} > 0$:

$$\ln\left(1 + \frac{||\boldsymbol{\chi}||^2}{\boldsymbol{x}}\right) \geq \alpha(\bar{\chi}, \bar{x}) + 2\frac{\Re\{\bar{\chi}^H \boldsymbol{\chi}\}}{\bar{x}} - \psi(\bar{\chi}, \bar{x})(||\boldsymbol{\chi}||^2 + \boldsymbol{x}), \tag{82}$$

with

$$\begin{aligned} \alpha(\bar{\chi}, \bar{x}) &\triangleq \ln\left(1 + \frac{||\bar{\chi}||^2}{\bar{x}}\right) - \frac{||\bar{\chi}||^2}{\bar{x}}, \\ 0 &< \psi(\bar{\chi}, \bar{x}) \triangleq \frac{||\bar{\chi}||^2}{\bar{x}(||\bar{\chi}||^2 + \bar{x})} \end{aligned} \tag{83}$$

and

$$\frac{||\boldsymbol{\chi}||^2}{\boldsymbol{x}} \geq -\frac{||\bar{\chi}||^4}{\bar{x}^2} + 2\frac{\bar{x} + ||\bar{\chi}||^2}{\bar{x}^2}\Re\{\bar{\chi}^H \boldsymbol{\chi}\} - \frac{||\bar{\chi}||^2}{\bar{x}^2}(||\boldsymbol{\chi}||^2 + \boldsymbol{x}). \tag{84}$$

The inequality (82) has been proved in [38]. The following steps assist in proving (84):

$$\frac{||\boldsymbol{\chi}||^2}{\boldsymbol{x}} = \frac{1}{1 - \frac{||\boldsymbol{\chi}||^2}{\boldsymbol{x} + ||\boldsymbol{\chi}||^2}} - 1 \tag{85}$$

$$\geq \frac{2}{1 - \frac{||\bar{\chi}||^2}{\bar{x} + ||\bar{\chi}||^2}} - \frac{1}{\left(1 - \frac{||\bar{\chi}||^2}{\bar{x} + ||\bar{\chi}||^2}\right)^2}\left(1 - \frac{||\boldsymbol{\chi}||^2}{\boldsymbol{x} + ||\boldsymbol{\chi}||^2}\right)$$

$$- 1 \tag{86}$$

$$= 2\left(1 + \frac{||\bar{\chi}||^2}{\bar{x}}\right) - \left(1 + \frac{||\bar{\chi}||^2}{\bar{x}}\right)^2$$

$$+ \left(1 + \frac{||\bar{\chi}||^2}{\bar{x}}\right)^2 \frac{||\boldsymbol{\chi}||^2}{\boldsymbol{x} + ||\boldsymbol{\chi}||^2} - 1 \tag{87}$$

$$= -\frac{||\bar{\chi}||^4}{\bar{x}^2} + \frac{\left(|\bar{x}| + ||\bar{\chi}||^2\right)^2}{\bar{x}^2}\frac{||\boldsymbol{\chi}||^2}{\boldsymbol{x} + ||\boldsymbol{\chi}||^2} \tag{88}$$

$$\geq -\frac{||\bar{\chi}||^4}{\bar{x}^2} + \frac{(|\bar{x}| + ||\bar{\chi}||^2)^2}{\bar{x}^2}\left(2\frac{\Re\{\bar{\chi}^H\chi\}}{\bar{x} + ||\bar{\chi}||^2}\right.$$
$$\left. -\frac{||\bar{\chi}||^2}{(\bar{x} + ||\bar{\chi}||^2)^2}(||\chi||^2 + x)\right) \quad (89)$$

$$= -\frac{||\bar{\chi}||^4}{\bar{x}^2} + 2\frac{\bar{x} + ||\bar{\chi}||^2}{\bar{x}^2}\Re\{\bar{\chi}^H\chi\} - \frac{||\bar{\chi}||^2}{\bar{x}^2}(||\chi||^2 + x), \quad (90)$$

which is the same as (84). Both expressions (86) and (89) are derived through the linearization of the convex function $1/(1 - t)$ for $t = ||\chi||^2/(x + ||\chi||^2) \in [0, 1)$.

The following inequality holds true for all $\chi_\ell \in \mathbb{C}^N$, $\bar{\chi}_\ell \in \mathbb{C}^N$, and $x_\ell > 0$, $\bar{x}_\ell > 0$, $\ell \in \mathcal{L}$, and $c > 0$:

$$\ln\left(\sum_{\ell \in \mathcal{L}}(1 - \frac{||\chi_\ell||^2}{||\chi_\ell||^2 + cx_\ell})\right)$$
$$\leq \ln\left(\sum_{\ell \in \mathcal{L}}(1 - \frac{||\bar{\chi}_\ell||^2}{||\bar{\chi}_\ell||^2 + c\bar{x}_\ell})\right)$$
$$+ \left(\sum_{\ell \in \mathcal{L}}(1 - \frac{||\bar{\chi}_\ell||^2}{c\bar{x}_\ell + ||\bar{\chi}_\ell||^2})\right)^{-1}\sum_{\ell \in \mathcal{L}}\frac{||\bar{\chi}_\ell||^2}{c\bar{x}_\ell + ||\bar{\chi}_\ell||^2}$$
$$- \left(\sum_{\ell \in \mathcal{L}}(1 - \frac{||\bar{\chi}_\ell||^2}{c\bar{x}_\ell + ||\bar{\chi}_\ell||^2})\right)^{-1}\sum_{\ell \in \mathcal{L}}\left(2\frac{\Re\{\bar{\chi}_\ell^H\chi_\ell\}}{c\bar{x}_\ell + ||\bar{\chi}_\ell||^2}\right.$$
$$\left. -\frac{||\bar{\chi}_\ell||^2}{(c\bar{x}_\ell + ||\bar{\chi}_\ell||^2)^2}(||\chi_\ell||^2 + cx_\ell)\right). \quad (91)$$

Define $\pi_\ell(\chi_\ell, x_\ell) \triangleq ||\chi_\ell||^2 + cx_\ell$ to write

$$\ln\left(\sum_{\ell \in \mathcal{L}}(1 - \frac{||\chi_\ell||^2}{||\chi_\ell||^2 + cx_\ell})\right)$$
$$= \ln\left(\sum_{\ell \in \mathcal{L}}(1 - \frac{||\chi_\ell||^2}{\pi_\ell(\chi_\ell, x_\ell)})\right)$$
$$\leq \ln\left(\sum_{\ell \in \mathcal{L}}(1 - \frac{||\bar{\chi}_\ell||^2}{\pi_\ell(\bar{\chi}_\ell, \bar{x}_\ell)})\right) + \left(\sum_{\ell \in \mathcal{L}}(1 - \frac{||\bar{\chi}_\ell||^2}{\pi_\ell(\bar{\chi}_\ell, \bar{x}_\ell)})\right)^{-1}$$
$$\times \sum_{\ell \in \mathcal{L}}\left(\frac{||\bar{\chi}_\ell||^2}{\pi_\ell(\bar{\chi}_\ell, \bar{x}_\ell)} - \frac{||\chi_\ell||^2}{\pi_\ell(\chi_\ell, x_\ell)}\right), \quad (92)$$

while

$$\frac{||\chi_\ell||^2}{\pi_\ell(\chi_\ell, x_\ell)} \geq 2\frac{\Re\{\bar{\chi}_\ell^H\chi_\ell\} - ||\bar{\chi}_\ell||^2}{\pi_\ell(\bar{\chi}_\ell, \bar{x}_\ell)} - \frac{|\bar{\chi}_\ell|^2}{\pi_\ell^2(\bar{\chi}_\ell, \bar{x}_\ell)}\pi_\ell(\chi_\ell, x_\ell). \quad (93)$$

Thus (91) follows from (92) and (93).

Considering the both sides of (82) and (84) ((91), resp.) as the functions of (χ, x) ((χ_ℓ, x_ℓ), $\ell \in \mathcal{L}$, resp.), the RHS of (82) and (84) ((91), resp.) match with their LHS counterpart at $(\bar{\chi}, \bar{x})$ (($\bar{\chi}_\ell, \bar{x}_\ell$), $\ell \in \mathcal{L}$, resp.), so the latter is a tight minorant (majorant, resp.) of the former at $(\bar{\chi}, \bar{x})$ (($\bar{\chi}_\ell, \bar{x}_\ell$), $\ell \in \mathcal{L}$, resp.) [39].

REFERENCES

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

[2] Y.-H. Nam *et al.*, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 172–179, Jun. 2013.

[3] Y. Kim *et al.*, "Full dimension MIMO (FD-MIMO): the next evolution of MIMO in LTE systems," *IEEE Wirel. Commun.*, vol. 21, no. 2, pp. 26–33, Feb. 2014.

[4] H. Ji *et al.*, "Overview of full-dimension MIMO in LTE-advanced pro," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 176–184, Feb. 2017.

[5] B. Mondal *et al.*, "3D channel model in 3GPP," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 16–23, Mar. 2015.

[6] 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, no. 11, pp. 2839–2853, Jun. 2019.

[7] 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, no. 2, pp. 914–928, Feb. 2021.

[8] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[9] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing: The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[10] E. Bjornson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4298–4310, Aug. 2010.

[11] C. Zhang, Y. Huang, Y. Jing, S. Jin, and L. Yang, "Sum-rate analysis for massive MIMO downlink with joint statistical beamforming and user scheduling," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 4, pp. 2181–2194, Apr. 2017.

[12] X. Li, C. Li, S. Jin, and X. Gao, "Interference coordination for 3-D beamforming-based hetnet exploiting statistical channel-state information," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 10, pp. 6887–6900, Oct. 2018.

[13] S. Qiu, D. Gesbert, D. Chen, and T. Jiang, "A covariance-based hybrid channel feedback in FDD massive MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8365–8377, Dec. 2019.

[14] H. Liu, X. Yuan, and Y. J. Zhang, "Statistical beamforming for FDD downlink massive MIMO via spatial information extraction and beam selection," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 7, pp. 4617–4631, Jul. 2020.

[15] X. Li, Z. Liu, N. Qin, and S. Jin, "FFR based joint 3D beamforming interference coordination for multi-cell FD-MIMO downlink transmission systems," *IEEE Trans. Veh. Techn.*, vol. 69, no. 3, pp. 3015–3118, Mar. 2020.

[16] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Trans. Wirel. Commun.*, vol. 7, no. 5, pp. 1711–1721, May 2007.

[17] K. Mullen, "A note on the ratio of two independent random variables," *Amer. Stat.*, vol. 21, no. 3, pp. 30–31, Jun. 1967.

[18] R. Zhang and L. Hanzo, "Cooperative downlink multicell preprocessing relying on reduced-rate back-haul data exchange," *IEEE Trans. Veh. Techn.*, vol. 60, no. 2, pp. 539–545, Feb 2011.

[19] J. Park, G. Lee, Y. Sung, and M. Yukawa, "Coordinated beamforming with relaxed zero forcing: The sequential orthogonal projection combining method and rate control," *IEEE Trans. Signal Process*, vol. 61, no. 12, pp. 3100–3112, Jun. 2013.

[20] D. Kim, G. Lee, and Y. Sung, "Two-stage beamformer design for massive MIMO downlink by trace quotient formulation," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2200–2211, Jun. 2015.

[21] X. Li, S. Jin, X. Gao, and R. W. Heath, "Three-dimensional beamforming for large-scale FD-MIMO systems exploiting statistical channel state information," *IEEE Trans. Veh. Techn.*, vol. 65, no. 11, pp. 8992–9005, Nov. 2016.

[22] Y. Song, C. Liu, W. Wang, N. Cheng, M. Wang, W. Zhuang, and X. Shen, "Domain selective precoding in 3-D massive MIMO systems," *IEEE J. Select. Topics Signal Process.*, vol. 13, no. 5, pp. 1103–1117, May 2019.

[23] A. Lozano, "Long-term transmit beamforming for wireless multicasting," in *Proc. 2007 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2007, pp. III–417–III–420.

[24] Y.-P. Lin, "Hybrid MIMO-OFDM beamforming for wideband mmWave channels without instantaneous feedback," *IEEE Trans. Signal Process.*, vol. 66, no. 19, pp. 5142–5151, Oct. 2018.

[25] D. Ying, F. W. Vook, T. A. Thomas, D. J. Love, and A. Ghosh, "Kronecker product correlation model and limited feedback codebook design in a 3D channel model," in *Proc. 2014 IEEE Inter. Conf. Commun. (ICC)*, 2014, pp. 5865–5870.

[26] Q.-U.-A. Nadeem, A. Kammoun, and M.-S. Alouini, "Elevation beamforming with full dimension MIMO architectures in 5G systems: A tutorial," *IEEE Commun. Surv. Tut.*, vol. 21, no. 4, pp. 3238–3273, 2019.

[27] W. Zhu, H. D. Tuan, E. Dutkiewicz, Y. Fang, and L. Hanzo, "Low-complexity Pareto-optimal 3D beamforming for the full-dimensional multi-user massive MIMO downlink," *IEEE Trans. Veh. Tech.*, vol. 72, no. 7, pp. 8869–8885, Jul. 2023.

[28] A. Alkhateeb, G. Leus, and R. W. Heath, "Multi-layer precoding: A potential solution for full-dimensional massive MIMO systems," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 3, pp. 5810–5824, Mar. 2017.

[29] J. Kang, O. Simeone, J. Kang, and S. Shamai, "Layered downlink precoding for C-RAN systems with full dimensional MIMO," *IEEE Trans. Veh. Techn.*, vol. 66, no. 3, pp. 2170–2182, Mar. 2017.

[30] Z. Wang, W. Liu, C. Qian, S. Chen, and L. Hanzo, "Two-dimensional precoding for 3-D massive MIMO," *IEEE Trans. Veh. Techn.*, vol. 66, no. 6, pp. 5488–5493, Jun. 2017.

[31] 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, Jan. 2022.

[32] 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.*, vol. 70, no. 9, pp. 6050–6065, Sept. 2022.

[33] W. Zhu, H. D. Tuan, E. Dutkiewicz, and L. Hanzo, "Collaborative beamforming aided fog radio access networks," *IEEE Trans. Veh. Techn.*, vol. 71, no. 7, pp. 7805–7820, Jul. 2022.

[34] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Sophia Antipolis Cedex, France, Tech. Rep. TR 38.901 (V14.2.0), Sep. 2017. [Online]. Available: http:// www.3gpp.org/.

[35] P. Patcharamaneepakorn, S. Armour, and A. Doufex, "On the equivalence between SLNR and MMSE precoding schemes with single-antenna receivers," *IEEE Commun. Lett.*, vol. 16, no. 3, pp. 1034–1037, Mar. 2012.

[36] E. Bjornson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Found. Trends Commun. Inf. Theory*, vol. 9, no. 23, pp. 113–381, 2013.

[37] R. Jain, D.-M. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *Digital Equipment, Tech. Rep. DEC-TR-301*, Sept. 1984.

[38] H. H. M. Tam, H. D. Tuan, and D. T. Ngo, "Successive convex quadratic programming for quality-of-service management in full-duplex MU-MIMO multicell networks," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2340–2353, Jun. 2016.

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