

RIS-Aided Multiple-Input Multiple-Output Broadcast Channel Capacity

H. D. Tuan¹, A. A. Nasir², E. Dutkiewicz¹, H. V. Poor³, and L. Hanzo⁴

Abstract—Scalable algorithms are conceived for obtaining the sum-rate capacity of the reconfigurable intelligent surface (RIS)-aided multiuser (MU) multiple-input multiple-output (MIMO) broadcast channel (BC), where a multi-antenna base station (BS) transmits signals to multi-antenna users with the help of an RIS equipped with a massive number of finite-resolution programmable reflecting elements (PREs). As a byproduct, scalable path-following algorithms emerge for determining the sum-rate capacity of the conventional MIMO BCs, closing a long-standing open problem of information theory. The paper also develops scalable algorithms for maximizing the minimum rate (max-min rate optimization) of the users achieved by the joint design of RIS's PRE and transmit beamforming for such an RIS-aided BC. The simulations provided confirm the high performance achieved by the algorithms developed, despite their low computational complexity.

Index Terms—Reconfigurable intelligent surface (RIS)-aided communication, sum capacity, mixed discrete continuous optimization, scalable algorithms.

I. INTRODUCTION

Reconfigurable intelligent surfaces (RISs) have emerged as a potential technology for low-power and energy-efficient information delivery in communication and sensing [1], [2]. Tuning of the RISs' reconfigurable programmable elements (PREs) has been shown to be effective in maximizing the spectral efficiency of the Internet-of-Things (IoT) (see e.g. [3]–[9] and references therein). However, there is a paucity of literature on the capacity of RIS-aided multiple-input multiple-output (MIMO) broadcast channels (BCs). This is not a surprise, because since the publication of [10] it has been widely acknowledged that the capacity of conventional (RIS-less) BCs can be obtained by solving a semi-definite convex optimization problem, but no scalable algorithm has been discovered for its computation. Hence its investigation remains limited to simple

The work was supported in part by the Australian Research Council's Discovery Projects under Grant DP190102501, in part by the Deanship of Research Oversight and Coordination (DROC) at KFUPM for funding under the Interdisciplinary Research Center for Communication Systems and Sensing through project No. INCS2203, in part by the U.S National Science Foundation under Grants CNS-2128448 and ECCS-2335876, in part by the Engineering and Physical Sciences Research Council projects EP/W016605/1, EP/X01228X/1 and EP/Y026721/1 as well as by the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028)

¹School of Electrical and Data Engineering, University of Technology Sydney, Broadway, NSW 2007, Australia (email:Tuan.Hoang@uts.edu.au, Eryk.Dutkiewicz@uts.edu.au); ²Department of Electrical Engineering and Center for Communication Systems and Sensing King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia (email: anasir@kfupm.edu.sa); ³Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA (email: poor@princeton.edu); ⁴School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK (email: lh@ecs.soton.ac.uk)

low-dimensional scenarios. Meanwhile, as RISs have to use large numbers of low resolution PREs to exploit their full potential [11]–[13], the capacity of RIS-aided BCs is obtained by solving a large-scale mixed discrete-continuous problem, which is computationally intractable. Another computationally challenging problem is that of maximizing the users' minimum rate achieved by RIS-aided BCs, which leads to the fairest possible user rate distribution. This problem of non-smooth optimization has been addressed in our previous work [4] for the case of infinite-resolution PREs. In fact, the non-smoothness of the minimum rate based objective function does not impose any additional computational difficulty on convex-solver based algorithms, which exploit the fact that the pointwise minimum of concave functions is a non-smooth but a concave function [14]. However, this non-smoothness constitutes a challenge for the development of scalable algorithms, which iterate by evaluating closed-form expressions instead of solving convex subproblems. Thus, to avoid this non-smooth minimum rate objective function, our recent contributions [5], [6], [15], [16] unveiled the benefits of having smooth objective functions relying on the user rates' geometric mean (GM-rate), whose maximization conveniently leads to improving the UEs' fairness. By exploiting the smoothness of the latter, scalable algorithms have been developed in [5], [15], [16] for its maximization.

Against the above background, the paper offers the following new contributions:

- It develops scalable algorithms for determining the capacity of RIS-aided MIMO BCs. Their byproducts are scalable path-following algorithms characterizing the capacity of the conventional (RIS-less) MIMO BCs. Apart from their communications theoretical insights, these algorithms are also novel from an optimization perspective.
- It develops scalable algorithms for addressing the problem of maximizing the users' minimum rate achieved by the joint design of the RIS's PREs and transmit beamforming. More explicitly, we propose a surrogate problem of max-min optimization, for which the non-smooth max-min objective function is closely approximated by a smooth function termed as a soft-min function. Accordingly, a scalable algorithm is developed for its computation. The simulations provided show that the algorithm achieved a near-unity minimum to maximum rate ratio, demonstrating an excellent rate-fairness.

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

The remainder of the paper is structured as follows. Sec-

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

Contents \ Literature	This work	[10]	[4]	[5]	[6]	[15], [16]
BC capacity	✓	✓				
RIS-aided BC capacity	✓					
RIS-aided max-min rate	✓		✓			
GM-rate maximization				✓	✓	✓
RIS-aided GM-rate maximization				✓	✓	
Quantized RISs	✓			✓	✓	
Scalable algorithms	✓			✓	✓	✓

tion II and Section III are devoted to the development of scalable algorithms conceived for determining the RIS-aided BC capacity and to maximizing the users' minimum rates, respectively. Section IV provides simulations and reports on the associated computational aspects, while our conclusions are given by Section V. The mathematical ingredients of the paper are matrices inequalities, which are provided in the Appendix.

Notation. Only optimization variables are boldfaced in the paper; $X \succeq 0$ ($X \succ 0$, resp.) indicates that X is a positive semi-definite (definite, resp.) matrix; Whenever $X \succeq 0$, $\lambda_{\max}(X)$ stands for the maximal eigenvalue of X ; Whenever $X = (X_1, \dots, X_K)$, $X \succeq 0$ means $X_k \succeq 0$, $k = 1, \dots, K$. The reader is referred to [17]–[19] for fundamental results on matrix inequalities that are useful for computational solutions of communication and signal processing systems. Furthermore, we also use notations such as $\langle X, Y \rangle \triangleq \text{trace}(X^H Y)$, $\langle X \rangle = \text{trace}(X)$, $[X]^2 = X X^H$ for the matrices X and Y . Thus, the Frobenius squared norm of the matrix X can be defined by $\|X\|^2 \triangleq \langle [X]^2 \rangle$.

Basic definition. According to [14, p. 366], a function \bar{f} is said to be a tight minorant (majorant, resp.) of a function f over the domain $\text{dom}(f)$ at a point $\bar{x} \in \text{dom}(f)$ if $f(\bar{x}) = \bar{f}(\bar{x})$ and $f(x) \geq \bar{f}(x) \forall x \in \text{dom}(f)$ ($f(x) \leq \bar{f}(x) \forall x \in \text{dom}(f)$, resp.). Then $f(x^{opt}) \geq \bar{f}(\bar{x})$ for $x^{opt} = \arg \max_{x \in \text{dom}(f)} f(x)$ ($f(x^{opt}) \leq \bar{f}(\bar{x})$ for $x^{opt} = \arg \min_{x \in \text{dom}(f)} f(x)$, resp.).

II. RIS-AIDED BC CAPACITY

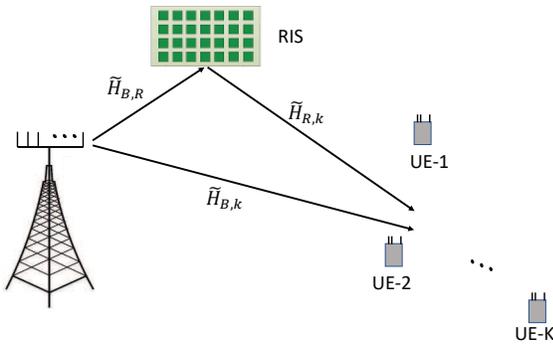


Fig. 1: RIS-aided BC

We consider the RIS-aided MIMO BC illustrated by Fig. 1, where a RIS of N reflecting elements supports the broadcast by an N_t -antenna BS to K N_r -antenna users (UEs) $k \in \mathcal{K} \triangleq \{1, \dots, K\}$. It is conventional to assume line-of-sight (LoS) links between the BS and RIS, and the RIS and

UEs, and but non-LoS (NLoS) link between the BS and UEs because the RIS is typically in a prominent high-rise position [20]. Accordingly, the quasi-static and flat-fading channels spanning from the BS and the RIS to UE k and from the BS to the RIS are modelled by $\tilde{H}_{B,k} = \sqrt{\beta_{B,k}} H_{B,k} \in \mathbb{C}^{N_r \times N_t}$, $\tilde{H}_{R,k} = \sqrt{\beta_{R,k}} H_{R,k} \in \mathbb{C}^{N_r \times N}$, and $\tilde{H}_{B,R} = \sqrt{\beta_{B,R}} H_{B,R} \in \mathbb{C}^{N \times N_t}$, where $\sqrt{\beta_{B,k}}$, $\sqrt{\beta_{R,k}}$, and $\sqrt{\beta_{B,R}}$ model the path-loss and large-scale fading of the BS-to-UE k link, the RIS-to-UE k link, and the BS-to-RIS link, respectively [11], [21]. Furthermore, $H_{R,k}$ is modelled by Rician fading for representing the LoS channels between the RIS and the UEs [22]. By contrast, $H_{B,k}$ is modelled by Rayleigh fading in the face of NLoS channels between the BS and the UEs. Similarly to the seminal papers on network capacity, we assume having perfect channel state information (CSI), which can be obtained by channel estimation [23]–[27]. The channel matrix of the RIS-aided connection between the BS and UE $k \in \mathcal{K}$ is given by

$$\begin{aligned} \mathbb{C}^{N_r \times N_t} \ni \mathcal{H}_k(\boldsymbol{\theta}) &\triangleq \tilde{H}_{R,k} R_{R,k}^{1/2} \text{diag}(e^{j\boldsymbol{\theta}}) \tilde{H}_{B,R} + \tilde{H}_{B,k} \quad (1) \\ &= \tilde{H}_{B,R,k} \text{diag}(e^{j\boldsymbol{\theta}}) H_{B,R} + \tilde{H}_{B,k}, \quad (2) \end{aligned}$$

with $\tilde{H}_{B,R,k} \triangleq \sqrt{\beta_{B,R}} \sqrt{\beta_{R,k}} H_{R,k} R_{R,k}^{1/2} \in \mathbb{C}^{N_r \times N}$, where $R_{R,k} \in \mathbb{C}^{N \times N}$ represents the spatial correlation matrix of the RIS elements with respect to user k [21], and $\text{diag}(e^{j\boldsymbol{\theta}})$ in (5) for $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N)^T \in [0, 2\pi)^N$ represents the matrix of PREs. We are interested in quantized PREs having b -bit resolution, formulated as:

$$\boldsymbol{\theta}_n \in \mathcal{B} \triangleq \left\{ \nu \frac{2\pi}{2^b}, \nu = 0, 1, \dots, 2^b - 1 \right\}, \quad (3)$$

for $n \in \mathcal{N} \triangleq \{1, \dots, N\}$, i.e.

$$\boldsymbol{\theta} \in \mathcal{B}^N. \quad (4)$$

Let $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ be the signal to be broadcast by the BS. The signal received at UE $k \in \mathcal{K}$ is

$$y_k = \mathcal{H}_k(\boldsymbol{\theta}) \mathbf{x} + n_k, \quad (5)$$

where $n_k \in \mathcal{C}(0, \sigma I_{N_r})$ is the background noise at UE k . Let Π be the set of all bijections from \mathcal{K} to itself. Given the power budget P , the sum-rate capacity of this RIS-aided MIMO BC with $\boldsymbol{\theta}$ held fixed at $\bar{\boldsymbol{\theta}}$ is achievable by dirty paper coding (DPC) [28] given by

$$\max_{\substack{\mathbf{W}_k \succeq 0, \\ k \in \mathcal{K}, \pi \in \Pi}} \sum_{k=1}^K \ln \left| I_{N_r} + \mathcal{H}_{\pi(k)}(\bar{\boldsymbol{\theta}}) \mathbf{W}_{\pi(k)} \mathcal{H}_{\pi(k)}^H(\bar{\boldsymbol{\theta}}) \right|$$

$$\max_{\mathbb{C}^{N_r \times N_r} \ni \mathbf{X}_k \succeq 0} \sum_{k=1}^K \ln \left| \sigma I_{N_t} + \left(\sigma I_{N_t} + \sum_{i \neq k}^K \mathcal{H}_i^H(\bar{\theta}) H X_i^{(\kappa)} \mathcal{H}_i(\bar{\theta}) \right)^{-1/2} \mathcal{H}_k^H(\bar{\theta}) \mathbf{X}_k \mathcal{H}_k(\bar{\theta}) \right. \\ \left. \left(\sigma I + \sum_{i \neq k}^K \mathcal{H}_i^H(\bar{\theta}) X_i^{(\kappa)} \mathcal{H}_i(\bar{\theta}) \right)^{-1/2} \right| \quad \text{s.t.} \quad (7b), \quad (8)$$

$$\times \left(\sigma I + \sum_{i=1}^{k-1} \mathcal{H}_{\pi(i)}(\bar{\theta}) \mathbf{W}_{\pi(i)} \mathcal{H}_{\pi(i)}^H(\bar{\theta}) \right)^{-1} \Big| \\ \text{s.t.} \quad \sum_{k=1}^K \langle \mathbf{W}_k \rangle \leq P, \quad (6)$$

where \mathbf{W}_k represents the covariance matrix to be designed for the UE $k \in \mathcal{K}$. However, this problem is computationally intractable, so the main result of [10] is that it can be obtained via its duality

$$\max_{\substack{\mathbb{C}^{N_r \times N_r} \ni \mathbf{X}_k \succeq 0, \\ k \in \mathcal{K}}} \ln \left| \sigma I_{N_t} + \sum_{k=1}^K \mathcal{H}_k^H(\bar{\theta}) \mathbf{X}_k \mathcal{H}_k(\bar{\theta}) \right| \\ - \ln |\sigma I_{N_t}| \quad (7a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \langle \mathbf{X}_k \rangle \leq P, \quad (7b)$$

which is a semi-definite problem of convex optimization, but there is no scalable algorithm for its solution (see e.g. [29]), where \mathbf{X}_k represents the covariance matrix to be designed for the UE $k \in \mathcal{K}$. More explicitly, there is no convex solver of polynomial complexity for its computation. One can see the computational difficulty of this problem by briefly recalling how it was addressed in [29]. To elaborate at the κ -th iteration, the algorithm of [29] solves the problem shown at the bottom of this page to generate $X^{(\kappa+1)}$: which admits a closed-form solution. However, since this next iterative point $X^{(\kappa+1)}$ is not necessarily better $X^{(\kappa)}$, it is not easy to show the convergence of this iterative procedure [30].

The authors of [31] did not address (7) but solved the problem of individual power constraints formulated as

$$\max_{\mathbb{C}^{N_r \times N_r} \ni \mathbf{X}_k \succeq 0} \ln \left| \sigma I_{N_t} + \sum_{k=1}^K \mathcal{H}_k^H(\bar{\theta}) \mathbf{X}_k \mathcal{H}_k(\bar{\theta}) \right| \quad (9a)$$

$$\text{s.t.} \quad \langle \mathbf{X}_k \rangle \leq P_k, k = 1, \dots, K, \quad (9b)$$

by alternating optimization for all \mathbf{X}_k : fixing all $\mathbf{X}_i = \bar{X}_i$, $i \neq k$ to solve

$$\max_{\mathbb{C}^{N_r \times N_r} \ni \mathbf{X}_k \succeq 0} \ln \left| \sigma I_{N_t} + \sum_{i \neq k}^K \mathcal{H}_i^H(\bar{\theta}) \bar{X}_i \mathcal{H}_i(\bar{\theta}) \right. \\ \left. + \mathcal{H}_k^H(\bar{\theta}) \mathbf{X}_k \mathcal{H}_k(\bar{\theta}) \right| \quad \text{s.t.} \quad \langle \mathbf{X}_k \rangle \leq P_k. \quad (10)$$

Apart from its slow convergence due to the optimization relying on a single variable per iteration, the main drawback of (10) is that it does not work for the sum power constraint in (7) (because it is clear that $\langle \mathbf{X}_k \rangle \equiv P_k$ during the whole process!).

Indeed, the construction of a scalable algorithm for obtaining the global solution of (7) has remained an open problem for twenty years at the time of writing.

Thus, in this section we aim for determining the sum-rate capacity of the above RIS-aided BC. According to (7), this may be formulated by the following problem of mixed discrete continuous optimization:

$$\max_{\boldsymbol{\theta}, \mathbf{X} \succeq 0} f(\boldsymbol{\theta}, \mathbf{X}) \triangleq \ln \left| \sigma I_{N_t} + \sum_{k=1}^K \mathcal{H}_k^H(\boldsymbol{\theta}) \mathbf{X}_k \mathcal{H}_k(\boldsymbol{\theta}) \right| \\ - \ln |\sigma I_{N_t}| \quad \text{s.t.} \quad (4), (7b). \quad (11)$$

On one hand, alternating optimization in the continuous variable \mathbf{X} with $\boldsymbol{\theta}$ held fixed, is still challenging as discussed above. More particularly, one cannot use (8) because it does not necessarily result in a better feasible point. On the other hand, alternating optimization in the discrete variable $\boldsymbol{\theta}$ with \mathbf{X} held fixed poses a challenging combinatoric problem. We now develop scalable alternating iterations for both these problems.

Initialized by a $(\theta^{(0)}, X^{(0)})$ feasible for (11), let $(\theta^{(\kappa)}, X^{(\kappa)})$ be a feasible point of (11) that is found from the $(\kappa - 1)$ th iteration.

A. Covariance iteration

1) *SVD-based iteration*: We generate $X^{(\kappa+1)}$ so that

$$f(\theta^{(\kappa)}, X^{(\kappa+1)}) > f(\theta^{(\kappa)}, X^{(\kappa)}). \quad (12)$$

Upon defining $\mathcal{H}_k^{(\kappa)} \triangleq \mathcal{H}(\theta^{(\kappa)})$, and then $f^{(\kappa)}(\mathbf{X}) \triangleq f(\theta^{(\kappa)}, \mathbf{X}) = \ln |\sigma I_{N_t} + \sum_{k=1}^K (\mathcal{H}_k^{(\kappa)})^H \mathbf{X}_k \mathcal{H}_k^{(\kappa)}| - \ln |\sigma I_{N_t}|$, we generate $X^{(\kappa+1)}$ verifying (12) by addressing the following convex problem

$$\max_{\mathbb{C}^{N_r \times N_r} \ni \mathbf{X} \succeq 0} f^{(\kappa)}(\mathbf{X}) \triangleq \ln \left| \sigma I_{N_t} + \sum_{k=1}^K (\mathcal{H}_k^{(\kappa)})^H \mathbf{X}_k \mathcal{H}_k^{(\kappa)} \right| \\ - \ln |\sigma I_{N_t}| \quad \text{s.t.} \quad (7b). \quad (13)$$

Using the inequality (100) yields the following tight minorant of $f^{(\kappa)}(\mathbf{X})$ at $X^{(\kappa)}$:

$$\tilde{f}^{(\kappa)}(\mathbf{X}) \triangleq f^{(\kappa)}(X^{(\kappa)}) - \sum_{k=1}^K \left\langle A_k^{(\kappa)}, (\mathbf{X}_k + \epsilon I_{N_r})^{-1} - \right. \\ \left. (X_k^{(\kappa)} + \epsilon I_{N_r})^{-1} \right\rangle - \ln |\sigma I_{N_t}|, \quad (14)$$

where

$$\mathbb{C}^{N_r \times N_r} \ni A_k^{(\kappa)} = (X_k^{(\kappa)} \epsilon I_{N_r}) \mathcal{H}_k^{(\kappa)} \\ \times \left[\sigma I_{N_t} + \sum_{k=1}^K (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa)} \mathcal{H}_k^{(\kappa)} \right]^{-1} \\ \times (\mathcal{H}_k^{(\kappa)})^H (X_k^{(\kappa)} + \epsilon I_{N_r}) \succeq 0, \quad (15)$$

and

$$\epsilon = \frac{\sigma}{2\lambda_{\max} \left(\sum_{k=1}^K (\mathcal{H}_k^{(\kappa)})^H \mathcal{H}_k^{(\kappa)} \right)}, \quad (16)$$

resulting in:

$$\sigma I_{N_t} - \epsilon \sum_{k=1}^K (\mathcal{H}_k^{(\kappa)})^H \mathcal{H}_k^{(\kappa)} \succeq (\sigma/2) I_{N_t} \succ 0. \quad (17)$$

We solve the following problem to generate $X^{(\kappa+1)}$:

$$\max_{\mathbf{X} \succeq 0} \tilde{f}^{(\kappa)}(\mathbf{X}) \quad \text{s.t.} \quad (7b), \quad (18)$$

or, equivalently

$$\min_{\mathbf{X} \succ 0} \sum_{k=1}^K \langle A_k^{(\kappa)}, (\mathbf{X}_k + \epsilon I_{N_r})^{-1} \rangle \quad \text{s.t.} \quad (7b). \quad (19)$$

Upon harnessing the singular value decomposition (SVD),

$$A_k^{(\kappa)} = U_k^{(\kappa)} \Sigma_k^{(\kappa)} (U_k^{(\kappa)})^H \quad (20)$$

in conjunction with a unitary $U_k^{(\kappa)}$ and a positive diagonal $\Sigma_k^{(\kappa)} = \text{diag}\{\gamma_{k,1}, \dots, \gamma_{k,N_r}\}$, (19) becomes equivalent to

$$\min_{\bar{\mathbf{X}} \succ 0} \sum_{k=1}^K \langle \Sigma_k^{(\kappa)}, \bar{\mathbf{X}}_k^{-1} \rangle \quad \text{s.t.} \quad (7b) \quad (21)$$

for

$$\bar{\mathbf{X}}_k = (U_k^{(\kappa)})^H (\mathbf{X}_k + \epsilon I_{N_r}) U_k^{(\kappa)}, k = 1, \dots, K. \quad (22)$$

It can be inferred that the optimal solution $\bar{\mathbf{X}}_k$ in (21) must be diagonal, i.e.

$$\bar{\mathbf{X}}_k = \text{diag}\{\mathbf{x}_{k,1} + \epsilon, \dots, \mathbf{x}_{k,N_r} + \epsilon\} \quad (23)$$

and (22) leads to

$$\min_{\mathbf{x}_{k,j} > 0} \sum_{k=1}^K \sum_{j=1}^{N_r} \frac{\gamma_{k,j}}{\mathbf{x}_{k,j} + \epsilon} \quad \text{s.t.} \quad \sum_{k=1}^K \sum_{j=1}^{N_r} \mathbf{x}_{k,j} = P. \quad (24)$$

The optimal solution of (24) is given by

$$x_{k,j}^{opt} = \max\{\lambda \sqrt{\gamma_{k,j}} - \epsilon, 0\}, \quad (25)$$

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

$$\sum_{k=1}^K \sum_{j=1}^{N_r} \max\{\lambda \sqrt{\gamma_{k,j}} - \epsilon, 0\} = P. \quad (26)$$

Thus, it follows from (22) and (25) that a $X^{(\kappa+1)}$ verifying (12) obeys:

$$X_k^{(\kappa+1)} \triangleq U_k^{(\kappa)} \text{diag}\{x_{k,1}^{opt}, \dots, x_{k,N_r}^{opt}\} (U_k^{(\kappa)})^H, \quad (27)$$

with $U_k^{(\kappa)}$ and $x_{k,j}^{opt}$ are defined in (20) and (25).

Remark 1. When we replace the sum power constraint (7b) by the individual power constraints (9b) for the problem (19), the solution is given by (27) in conjunction with

$$x_{k,j}^{opt} = \max\{\lambda_k \sqrt{\gamma_{k,j}} - \epsilon, 0\}, \quad (28)$$

where $\lambda_k > 0$ is found by bisection satisfying that

$$\sum_{j=1}^{N_r} \max\{\lambda_k \sqrt{\gamma_{k,j}} - \epsilon, 0\} = P_k. \quad (29)$$

Remark 2. Note that the problem (13) is convex and thus (27) provides a scalable path-following iteration for computing its globally optimal solution, which is new even from algorithmic optimization perspective.

2) *SVD-free based iterations:* To avoid the SVD of (20), we make the following variable changes

$$\mathbf{X}_k = [\mathbf{P}_k]^2, \mathbf{P}_k \in \mathbb{C}^{N_r \times N_r}, k \in \mathcal{K}. \quad (30)$$

Accordingly, we also set $X_k^{(\kappa)} = [P_k^{(\kappa)}]^2$, $k = 1, \dots, K$. For $\mathbf{P} \triangleq (\mathbf{P}_1, \dots, \mathbf{P}_K)$, the problem (13) used for generating $X^{(\kappa+1)}$ verifying (12) is expressed by the following problem

$$\begin{aligned} \max_{\mathbf{P}} g^{(\kappa)}(\mathbf{P}) &\triangleq \ln |I_{N_t}| + \sum_{k=1}^K [(\mathcal{H}_k^{(\kappa)})^H \mathbf{P}_k]^2 / \sigma \\ \text{s.t.} \quad &\sum_{k=1}^K \langle [\mathbf{P}_k]^2 \rangle \leq P. \end{aligned} \quad (31)$$

Using the inequality (101) yields the following minorant of $g^{(\kappa)}(\mathbf{P})$ at $P^{(\kappa)}$:

$$\begin{aligned} \tilde{g}^{(\kappa)}(\mathbf{P}) &\triangleq a^{(\kappa)} + \frac{2}{\sigma} \sum_{k=1}^K \Re\{\langle (P_k^{(\kappa)})^H [\mathcal{H}_k^{(\kappa)}]^2 \mathbf{P}_k \rangle\} \\ &\quad - \sum_{k=1}^K \langle C^{(\kappa)}, [(\mathcal{H}_k^{(\kappa)})^H \mathbf{P}_k]^2 \rangle \\ &= a^{(\kappa)} + \frac{2}{\sigma} \sum_{k=1}^K \Re\{\langle (P_k^{(\kappa)})^H [\mathcal{H}_k^{(\kappa)}]^2 \mathbf{P}_k \rangle\} \\ &\quad - \sum_{k=1}^K \langle \mathbf{P}_k^H \mathcal{H}_k^{(\kappa)} C^{(\kappa)} (\mathcal{H}_k^{(\kappa)})^H \mathbf{P}_k \rangle, \end{aligned} \quad (32)$$

in conjunction with

$$\begin{aligned} a^{(\kappa)} &\triangleq g^{(\kappa)}(P^{(\kappa)}) - \frac{1}{\sigma} \sum_{k=1}^K \langle [(\mathcal{H}_k^{(\kappa)})^H P_k^{(\kappa)}]^2 \rangle - \sigma \langle C^{(\kappa)} \rangle, \\ 0 \leq C^{(\kappa)} &\triangleq \frac{1}{\sigma} I_{N_t} - \left(\sigma I_{N_t} + \sum_{k=1}^K [(\mathcal{H}_k^{(\kappa)})^H P_k^{(\kappa)}]^2 \right)^{-1}. \end{aligned} \quad (33)$$

For $C_\sigma^{(\kappa)} \triangleq \sigma C^{(\kappa)} = I_{N_t} - \left(I_{N_t} + \frac{1}{\sigma} \sum_{k=1}^K [(\mathcal{H}_k^{(\kappa)})^H P_k^{(\kappa)}]^2 \right)^{-1}$, we thus solve the following problem of tight minorant maximization of (30) to generate $P^{(\kappa+1)}$:

$$\max_{\mathbf{P}} \tilde{g}^{(\kappa)}(\mathbf{P}) \quad \text{s.t.} \quad \sum_{k=1}^K \langle [\mathbf{P}_k]^2 \rangle \leq P, \quad (34)$$

which admits the following closed-form solution given at the top of the next page, where $\mu > 0$ is found by bisection, so that

$$\sum_{k=1}^K \langle [(\mathcal{H}_k^{(\kappa)} C_\sigma^{(\kappa)} (\mathcal{H}_k^{(\kappa)})^H + \mu I_{N_r})^{-1} [\mathcal{H}_k^{(\kappa)}]^2 P_k^{(\kappa)}]^2 \rangle = P. \quad (36)$$

Thus, $X_k^{(\kappa+1)} \triangleq [P_k^{(\kappa+1)}]^2$ verifies (12).

Remark 3. The advantage of harnessing the variable change in (30) is that it still works for $P_k \in \mathbb{C}^{N_r \times q}$ with $q < N_r$, where P_k plays the role of MIMO-aided uplink beamforming in multi-access channels (MAC).

$$P_k^{(\kappa+1)} = \begin{cases} (\mathcal{H}_k^{(\kappa)} C_\sigma^{(\kappa)} (\mathcal{H}_k^{(\kappa)})^H)^{-1} [\mathcal{H}_k^{(\kappa)}]^2 [P_k^{(\kappa)}]^2 \\ \quad \text{if } \sum_{k=1}^K \langle (\mathcal{H}_k^{(\kappa)} C_\sigma^{(\kappa)} (\mathcal{H}_k^{(\kappa)})^H)^{-1} [\mathcal{H}_k^{(\kappa)}]^2 P_k^{(\kappa)} \rangle \leq P, \\ (\mathcal{H}_k^{(\kappa)} C_\sigma^{(\kappa)} (\mathcal{H}_k^{(\kappa)})^H + \mu I_{N_r})^{-1} [\mathcal{H}_k^{(\kappa)}]^2 P_k^{(\kappa)} \quad \text{otherwise,} \end{cases} \quad (35)$$

B. PREs iteration

With $P_k^{(\kappa+1)} \triangleq \sqrt{X_k^{(\kappa+1)}}$, and $\varphi^{(\kappa)}(\boldsymbol{\theta}) \triangleq f(\boldsymbol{\theta}, X^{(\kappa+1)}) = \ln |I_{N_t} + \sum_{k=1}^K [\mathcal{H}_k^H(\boldsymbol{\theta}) P_k^{(\kappa+1)}]^2 / \sigma|$, to generate $\theta^{(\kappa+1)}$ so that

$$\begin{aligned} f(\theta^{(\kappa+1)}, X^{(\kappa+1)}) &= \varphi^{(\kappa)}(\theta^{(\kappa+1)}) \\ &> \varphi^{(\kappa)}(\theta^{(\kappa)}) \\ &= f(\theta^{(\kappa)}, X^{(\kappa+1)}), \end{aligned} \quad (37)$$

we address the following problem

$$\max_{\boldsymbol{\theta}} \varphi^{(\kappa)}(\boldsymbol{\theta}) \quad \text{s.t.} \quad (4). \quad (38)$$

Using the inequality (101) yields the following minorant of $\varphi^{(\kappa)}(\boldsymbol{\theta})$ at $\theta^{(\kappa)}$:

$$\begin{aligned} \tilde{\varphi}^{(\kappa)}(\boldsymbol{\theta}) &\triangleq \tilde{a}^{(\kappa)} + \frac{2}{\sigma} \sum_{k=1}^K \Re\{ \langle X_k^{(\kappa+1)} \mathcal{H}_k^{(\kappa)} \mathcal{H}_k^H(\boldsymbol{\theta}) \rangle \} \\ &\quad - \sum_{k=1}^K \langle \tilde{C}^{(\kappa)}, \mathcal{H}_k^H(\boldsymbol{\theta}) X_k^{(\kappa+1)} \mathcal{H}_k(\boldsymbol{\theta}) \rangle \\ &= \tilde{a}^{(\kappa)} + \frac{2}{\sigma} \sum_{k=1}^K \Re\{ \langle (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa+1)} \mathcal{H}_k(\boldsymbol{\theta}) \rangle \} \\ &\quad - \sum_{k=1}^K \langle \tilde{C}^{(\kappa)}, \mathcal{H}_k^H(\boldsymbol{\theta}) X_k^{(\kappa+1)} \mathcal{H}_k(\boldsymbol{\theta}) \rangle, \end{aligned} \quad (39)$$

with

$$\begin{aligned} \tilde{a}^{(\kappa)} &\triangleq \varphi^{(\kappa)}(\theta^{(\kappa)}) - \frac{1}{\sigma} \sum_{k=1}^K \langle (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa+1)} \mathcal{H}_k^{(\kappa)} \rangle - \sigma \langle \tilde{C}^{(\kappa)} \rangle, \\ 0 &\preceq \tilde{C}^{(\kappa)} \triangleq \frac{1}{\sigma} I_{N_t} - \left(\sigma I_{N_t} + \sum_{k=1}^K (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa+1)} \mathcal{H}_k^{(\kappa)} \right)^{-1}. \end{aligned} \quad (40)$$

Then we have

$$\begin{aligned} \sigma \tilde{\varphi}^{(\kappa)}(\boldsymbol{\theta}) &= \sigma \tilde{a}^{(\kappa)} + 2 \sum_{k=1}^K \Re\{ \langle (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa+1)} \mathcal{H}_k(\boldsymbol{\theta}) \rangle \} \\ &\quad - \sum_{k=1}^K \langle \tilde{C}^{(\kappa)}, \mathcal{H}_k^H(\boldsymbol{\theta}) X_k^{(\kappa+1)} \mathcal{H}_k(\boldsymbol{\theta}) \rangle, \end{aligned} \quad (41)$$

along with $\tilde{C}_\sigma^{(\kappa)} \triangleq I_{N_t} - \left(I_{N_t} + \frac{1}{\sigma} \sum_{k=1}^K (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa+1)} \mathcal{H}_k^{(\kappa)} \right)^{-1}$.

We may write $\text{diag}(e^{j\boldsymbol{\theta}}) = \sum_{n=1}^N e^{j\boldsymbol{\theta}_n} \Upsilon_n$, where Υ_n is the matrix of size $N \times N$ having only zero entries, except for its (n, n) -entry, which is 1. Then the matrix $\mathcal{H}_k(\boldsymbol{\theta})$ defined by (2) is represented by

$$\mathcal{H}_k(\boldsymbol{\theta}) = \sum_{n=1}^N e^{j\boldsymbol{\theta}_n} \mathcal{H}_{k,n} + \tilde{H}_{B,k}, \quad (42)$$

with $\mathcal{H}_{k,n} \triangleq \tilde{H}_{B,R,k} \Upsilon_n H_{B,R}$. Therefore, we have

$$\langle (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa+1)} \mathcal{H}_k(\boldsymbol{\theta}) \rangle = \alpha_{k,1}^{(\kappa)} + \sum_{n=1}^N \tilde{b}_{k,1}^{(\kappa)}(n) e^{j\boldsymbol{\theta}_n}, \quad (43)$$

in conjunction with¹

$$\begin{aligned} \alpha_{k,1}^{(\kappa)} &\triangleq \langle (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa+1)} \tilde{H}_{B,k} \rangle, \\ \tilde{b}_{k,1}^{(\kappa)}(n) &= \langle (\mathcal{H}_k^{(\kappa)})^H X_k^{(\kappa+1)} \mathcal{H}_{k,n} \rangle, n \in \mathcal{N}. \end{aligned} \quad (44)$$

Furthermore,

$$\begin{aligned} \langle \tilde{C}_\sigma^{(\kappa)}, \mathcal{H}_k^H(\boldsymbol{\theta}) X_k^{(\kappa+1)} \mathcal{H}_k(\boldsymbol{\theta}) \rangle &= \alpha_{k,2}^{(\kappa)} + 2 \sum_{n=1}^N \Re\{ \tilde{b}_{k,2}^{(\kappa)}(n) e^{j\boldsymbol{\theta}_n} \} \\ &\quad + (e^{j\boldsymbol{\theta}})^H \Phi_k^{(\kappa+1)} e^{j\boldsymbol{\theta}}, \end{aligned} \quad (45)$$

with

$$\begin{aligned} \alpha_{k,2}^{(\kappa)} &\triangleq \langle \tilde{C}_\sigma^{(\kappa)} (\tilde{H}_{B,k})^H X_k^{(\kappa+1)} \tilde{H}_{B,k} \rangle, \\ \tilde{b}_{k,2}^{(\kappa)}(n) &= \langle \tilde{C}_\sigma^{(\kappa)} (\tilde{H}_{B,k})^H X_k^{(\kappa+1)} \mathcal{H}_{k,n} \rangle, n = 1, \dots, N, \\ \Phi_k^{(\kappa+1)}(n', n) &\triangleq \langle \tilde{C}_\sigma^{(\kappa)} \mathcal{H}_{k,n'}^H X_k^{(\kappa+1)} \mathcal{H}_{k,n} \rangle, (n', n) \in \mathcal{N} \times \mathcal{N}. \end{aligned} \quad (46)$$

Based on (43), and (45), we obtain

$$\begin{aligned} \sigma \tilde{\varphi}^{(\kappa)}(\boldsymbol{\theta}) &= \tilde{a}^{(\kappa+1)} + 2 \Re\left\{ \sum_{n=1}^N \tilde{b}^{(\kappa+1)}(n) e^{j\boldsymbol{\theta}_n} \right\} \\ &\quad - (e^{j\boldsymbol{\theta}})^H \Phi^{(\kappa+1)} e^{j\boldsymbol{\theta}}, \end{aligned} \quad (47)$$

with

$$\begin{aligned} \tilde{a}^{(\kappa+1)} &\triangleq \sigma \tilde{a}^{(\kappa)} + \sum_{k=1}^K \left(2 \Re\{ \alpha_{k,1}^{(\kappa)} \} - \alpha_{k,2}^{(\kappa)} \right), \\ \tilde{b}^{(\kappa+1)}(n) &\triangleq \sum_{k=1}^K \left(\tilde{b}_{k,1}^{(\kappa)}(n) - \tilde{b}_{k,2}^{(\kappa)}(n) \right), n \in \mathcal{N}, \\ \Phi^{(\kappa+1)} &\triangleq \sum_{k=1}^K \Phi_k^{(\kappa+1)}. \end{aligned} \quad (48)$$

Furthermore, a tight minorant of $\sigma \tilde{\varphi}^{(\kappa)}(\boldsymbol{\theta})$ and $\sigma \varphi^{(\kappa)}(\boldsymbol{\theta})$ is

$$\begin{aligned} \tilde{\tilde{\varphi}}_\sigma^{(\kappa)}(\boldsymbol{\theta}) &\triangleq \tilde{a}^{(\kappa+1)} + 2 \Re\left\{ \sum_{n=1}^N \left(\tilde{b}^{(\kappa+1)}(n) \right. \right. \\ &\quad \left. \left. - \sum_{n'=1}^N e^{-j\boldsymbol{\theta}_{n'}} \Phi^{(\kappa+1)}(n', n) \right. \right. \\ &\quad \left. \left. + \lambda_{\max}(\Phi^{(\kappa+1)}) e^{-j\boldsymbol{\theta}_n} \right) e^{j\boldsymbol{\theta}_n} \right\} \\ &\quad - (e^{j\boldsymbol{\theta}^{(\kappa)}})^H \Phi^{(\kappa+1)} e^{j\boldsymbol{\theta}^{(\kappa)}} - 2 \lambda_{\max}(\Phi^{(\kappa+1)}) N. \end{aligned} \quad (49)$$

We thus solve the following discrete problem of tight minorant maximization for (38) to generate $\theta^{(\kappa+1)}$:

$$\max_{\boldsymbol{\theta} \in \mathcal{B}^N} \tilde{\tilde{\varphi}}_\sigma^{(\kappa)}(\boldsymbol{\theta}), \quad (50)$$

which admits the closed-form solution of

$$\theta_n^{(\kappa+1)} = 2\pi - \left\lfloor \angle \left(\tilde{b}^{(\kappa+1)}(n) - \sum_{m=1}^N e^{-j\boldsymbol{\theta}_m^{(\kappa)}} \Phi^{(\kappa+1)}(m, n) \right) \right\rfloor$$

¹In what follows $b(i)$ is the i -th entry of b and $[A](i, i)$ is the i -th diagonal entry of A , and $[A]^*(i, i)$ is the complex conjugate of $[A](i, i)$

$$+ \lambda_{\max}(\Phi^{(\kappa+1)})e^{-j\theta_n^{(\kappa)}} \Big]_b, \quad n \in \mathcal{N}, \quad (51)$$

where $[\alpha]_b$ represents the projection of $\alpha \in [0, 2\pi]$ into \mathcal{B} defined by

$$[\alpha]_b = \nu_\alpha \frac{2\pi}{2^b}, \quad (52)$$

with

$$\nu_\alpha \triangleq \arg \min_{\nu=0,1,\dots,2^b} \left| \nu \frac{2\pi}{2^b} - \alpha \right|, \quad (53)$$

which can be readily found, because we have $\nu_\alpha \in \{\nu, \nu+1\}$ for $\alpha \in [\nu \frac{2\pi}{2^b}, (\nu+1) \frac{2\pi}{2^b}]$. We also reset $\nu_\alpha = 0$, when the optimal solution of (53) is 2^b .

C. Scalable path-following algorithms and their convergence

The pseudo code of the SVD-based iteration (27) (SVD-free iteration (35), resp.) for generating $X^{(\kappa+1)}$ and the iteration (51) for generating $\theta^{(\kappa+1)}$ is provided by Algorithm 1 (Algorithm 2, resp.). It follows from (27) and (37) that

$$f(\theta^{(\kappa+1)}, X^{(\kappa+1)}) > f(\theta^{(\kappa)}, X^{(\kappa)}), \quad (54)$$

i.e. both of them are path-following computational procedures, which iteratively improve the feasible points. As such, the sequence $\{(\theta^{(\kappa)}, X^{(\kappa)})\}$ converges to a limit point $(\bar{\theta}, \bar{X})$ according to the Cauchy theorem.

Algorithm 1 SVD-based sum-rate capacity (SVD-based SRC) path-following algorithm to solve the problem (11).

- 1: **Initialization.** Randomly generate $P_k^{(0)}$ satisfying $\sum_{k=1}^K \langle [P_k^{(0)}]^2 \rangle \leq P$. Set $X^{(0)} \triangleq ([P_1^{(0)}]^2, \dots, [P_K^{(0)}]^2)$. Generate a feasible point $\theta^{(0)}$. Set $\kappa = 0$.
 - 2: **Repeat until convergence of the objective in (11):** Generate $X^{(\kappa+1)}$ by (23)-(26), and $\theta^{(\kappa+1)}$ by (51). Reset $\kappa := \kappa + 1$.
 - 3: **Output** $(X^{(\kappa)}, \theta^{(\kappa)})$ and the achieved sum-rate capacity $f(\theta^{(\kappa)}, X^{(\kappa)})$.
-

Algorithm 2 SVD-free sum-rate capacity (SVD-free SRC) path-following algorithm to solve the problem (11).

- 1: **Initialization.** Randomly generate $P_k^{(0)}$ satisfying $\sum_{k=1}^K \langle [P_k^{(0)}]^2 \rangle \leq P$. Set $X^{(0)} \triangleq ([P_1^{(0)}]^2, \dots, [P_K^{(0)}]^2)$. Generate a feasible point $\theta^{(0)}$. Set $\kappa = 0$.
 - 2: **Repeat until convergence of the objective in (11):** Generate $P^{(\kappa+1)}$ by (35), $X^{(\kappa+1)}$ by (30), and $\theta^{(\kappa+1)}$ by (51). Reset $\kappa := \kappa + 1$.
 - 3: **Output** $(X^{(\kappa)}, \theta^{(\kappa)})$ and the achieved sum-rate capacity $f(\theta^{(\kappa)}, X^{(\kappa)})$.
-

III. MAX-MIN RATE OPTIMIZATION

In this section, the transmit signal \mathbf{x} in (5) is

$$\mathbf{x} = \sum_{k=1}^K \mathbf{P}_k s_k, \quad (55)$$

where $s_k \in \mathbb{C}^d$ with $d \leq N_r$ and $\mathbb{E}(\|s_k\|^2) = 1$ encodes the information stream intended for UE k , which is beamformed by $\mathbf{P}_k \in \mathbb{C}^{N_t \times d}$. For $\mathbf{P} \triangleq (\mathbf{P}_1, \dots, \mathbf{P}_K)$, the rate at UE k is

$$r_k(\boldsymbol{\theta}, \mathbf{P}) \triangleq \ln \left| I_{N_r} + [\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_k]^2 (\sigma I_{N_r} + \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_i]^2)^{-1} \right|. \quad (56)$$

Given the power budget P , we are interested in the following problem of max-min rate optimization

$$\max_{\boldsymbol{\theta}, \mathbf{P}} \min_{k=1, \dots, K} r_k(\boldsymbol{\theta}, \mathbf{P}) \quad \text{s.t. (4),} \quad (57a)$$

$$\sum_{k=1}^K \langle [\mathbf{P}_k]^2 \rangle \leq P. \quad (57b)$$

Our previous work [4] provides a convex-solver based algorithm for solving (57) for $d = N_r = 1$ and $b = \infty$, under which the variable $\boldsymbol{\theta} \triangleq e^{j\theta}$ in (57) is also continuous with its entries satisfying the unit modulus constraint, so they can be handled by using an exact penalty based method. The computational complexity of alternating iterations in \mathbf{P} is $\mathcal{O}((KN_t)^3)$ while that of alternating iterations in $\boldsymbol{\theta}$ is $\mathcal{O}(N^3)$. As mentioned in the Introduction, the non-smoothness of the objective function in (57) does not add an additional computational difficulty for convex-solver based algorithms, but it is still a hurdle for the development of scalable algorithms. Instead of the non-smooth objective function $\min_{k=1, \dots, K} r_k$ in (57), our recent treatises [5], [15], [16] used the smooth objective function $(\prod_{k=1}^K r_k)^{1/K}$ of the rate functions' geometric mean (GM-rate), whose maximization leads to improved rate-fairness for the users. By exploiting the smoothness of the latter, the scalable algorithms have been developed for its maximization in [5], [15], [16].

Given the discrete constraint (4) for the PREs' b -bit resolution, (57) represents a computationally challenging problem of large scale mixed discrete-continuous optimization, for which an exact penalty based method for its solution unlikely exists. Moreover, the convex-solver based approach is hardly applicable to discrete problems.

To develop scalable algorithms for addressing (57), we conceive the following surrogate max-min problem

$$\max_{\boldsymbol{\theta}, \mathbf{P}} \min_{k \in \mathcal{K}} \ln \left| I_{N_r} + [\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_k]^2 \left(c\sigma I_{N_r} + c \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_i]^2 \right)^{-1} \right| \quad \text{s.t. (4), (57b)}. \quad (58)$$

for $c > 0$. In fact, for $N_r = 1$,

$$r_k(\boldsymbol{\theta}, \mathbf{P}) = \ln \left(1 + \frac{|\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_k|^2}{\sigma + \sum_{i \neq k}^K |\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_i|^2} \right),$$

so

$$(57) \Leftrightarrow \max_{\boldsymbol{\theta}, \mathbf{P}} \min_{k \in \mathcal{K}} \frac{|\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_k|^2}{\sigma + \sum_{i \neq k}^K |\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_i|^2} \quad \text{s.t. (4), (57b)} \quad (59)$$

$$\Leftrightarrow \max_{\boldsymbol{\theta}, \mathbf{P}} \min_{k \in \mathcal{K}} \frac{|\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_k|^2}{c \left(\sigma + \sum_{i \neq k}^K |\mathcal{H}_k(\boldsymbol{\theta})\mathbf{P}_i|^2 \right)}$$

$$\begin{aligned} & \text{s.t. (4), (57b)} \quad (60) \\ \Leftrightarrow & \max_{\boldsymbol{\theta}, \mathbf{P}} \min_{k \in \mathcal{K}} \ln \left(1 + \frac{|\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k|^2}{c \left(\sigma + \sum_{i \neq k}^K |\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_i|^2 \right)} \right) \\ & \text{s.t. (4), (57b),} \quad (61) \end{aligned}$$

i.e. (57) is equivalent to (58).

Let us define

$$\begin{aligned} \Gamma_c(\boldsymbol{\theta}, \mathbf{P}) \triangleq & \sum_{k=1}^K \left[I_{N_r} - (\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k)^H \left([\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k]^2 + c\sigma I_{N_r} \right. \right. \\ & \left. \left. + c \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_i]^2 \right)^{-1} \mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k \right]. \quad (62) \end{aligned}$$

Then one has

$$\begin{aligned} f_{MM}(\boldsymbol{\theta}, \mathbf{P}) \triangleq & \max_{k \in \mathcal{K}} \ln \left(\left| I_{N_r} + [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k]^2 \left(c\sigma I_{N_r} \right. \right. \right. \\ & \left. \left. \left. + c \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_i]^2 \right)^{-1} \right| \right) \quad (63) \end{aligned}$$

$$\begin{aligned} = & \max_{k \in \mathcal{K}} \ln \left| I_{N_r} - (\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k)^H \left([\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k]^2 \right. \right. \\ & \left. \left. + c\sigma I_{N_r} + c \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_i]^2 \right)^{-1} \mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k \right| \quad (64) \end{aligned}$$

$$\leq \ln \left| \Gamma_c(\boldsymbol{\theta}, \mathbf{P}) \right| \quad (65)$$

$$\begin{aligned} \leq & \frac{1}{K} \sum_{k \in \mathcal{K}} \ln \left| I_{N_r} - (\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k)^H \left([\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k]^2 \right. \right. \\ & \left. \left. + c\sigma I_{N_r} + c \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_i]^2 \right)^{-1} \mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k \right| \\ & + \ln K \quad (66) \end{aligned}$$

$$\begin{aligned} = & \frac{1}{K} \sum_{k \in \mathcal{K}} \ln \left| \left(I_{N_r} + [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k]^2 \left(c\sigma I_{N_r} \right. \right. \right. \\ & \left. \left. \left. + c \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_i]^2 \right)^{-1} \right) \right| + \ln K \quad (67) \end{aligned}$$

$$\leq f_{MM}(\boldsymbol{\theta}, \mathbf{P}) + \ln K. \quad (68)$$

Note that $\ln K$ in (68) becomes small compared to $f_{MM}(\boldsymbol{\theta}, \mathbf{P})$, as c becomes smaller. Thus, with c smaller, the smooth function $\ln \left| \Gamma_c(\boldsymbol{\theta}, \mathbf{P}) \right|$ closely approximates the non-smooth function $f_{MM}(\boldsymbol{\theta}, \mathbf{P})$. Furthermore, we have

$$\begin{aligned} -f_{MM}(\boldsymbol{\theta}, \mathbf{P}) = & -\max_{k \in \mathcal{K}} \left[-\ln \left| I_{N_r} + [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k]^2 \right. \right. \\ & \left. \left. \times \left(c\sigma I_{N_r} + c \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_i]^2 \right)^{-1} \right| \right] \quad (69) \end{aligned}$$

$$\begin{aligned} = & \min_{k \in \mathcal{K}} \left[\ln \left| I_{N_r} + [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_k]^2 \left(c\sigma I_{N_r} \right. \right. \right. \\ & \left. \left. \left. + c \sum_{i \neq k}^K [\mathcal{H}_k(\boldsymbol{\theta}) \mathbf{P}_i]^2 \right)^{-1} \right| \right], \quad (70) \end{aligned}$$

so the non-smooth objective function in (58) is closely approximated by the function $-\ln \left| \Gamma_c(\boldsymbol{\theta}, \mathbf{P}) \right| = \ln \left| \Gamma_c(\boldsymbol{\theta}, \mathbf{P}) \right|^{-1}$. We thus propose to consider the following problem having a smooth objective function, which is called a soft-min objective function, to address (61)/(57):

$$\max_{\boldsymbol{\theta}, \mathbf{P}} f_c(\boldsymbol{\theta}, \mathbf{P}) \triangleq \ln \left| \Gamma_c(\boldsymbol{\theta}, \mathbf{P}) \right|^{-1} \quad \text{s.t. (4), (57b).} \quad (71)$$

Initialized by $(\boldsymbol{\theta}^{(0)}, P^{(0)})$ feasible for (71), let $(\boldsymbol{\theta}^{(\kappa)}, P^{(\kappa)})$ be a feasible point for (71) that is found from the $(\kappa - 1)$ -th iteration.

A. Beamforming iteration

To generate a feasible $P^{(\kappa+1)}$ so that

$$f_c(\boldsymbol{\theta}^{(\kappa)}, P^{(\kappa+1)}) > f_c(\boldsymbol{\theta}^{(\kappa)}, P^{(\kappa)}), \quad (72)$$

we address the following problem

$$\max_{\mathbf{P}} f_b^{(\kappa)}(\mathbf{P}) \triangleq \ln \left| \Gamma_b^{(\kappa)}(\mathbf{P}) \right|^{-1} \quad \text{s.t. (57b),} \quad (73)$$

in conjunction with

$$\begin{aligned} \Gamma_b^{(\kappa)}(\mathbf{P}) & \triangleq \Gamma_c(\boldsymbol{\theta}^{(\kappa)}, \mathbf{P}) \\ & = \sum_{k=1}^K \left[I_{N_r} - (\mathcal{H}_k^{(\kappa)} \mathbf{P}_k)^H \left([\mathcal{H}_k^{(\kappa)} \mathbf{P}_k]^2 \right. \right. \\ & \left. \left. + c\sigma I_{N_r} + c \sum_{i \neq k}^K [\mathcal{H}_k^{(\kappa)} \mathbf{P}_i]^2 \right)^{-1} \mathcal{H}_k^{(\kappa)} \mathbf{P}_k \right] \quad (74) \end{aligned}$$

where $\mathcal{H}_k^{(\kappa)} \triangleq \mathcal{H}_k(\boldsymbol{\theta}^{(\kappa)})$, $k = 1, \dots, K$. Applying the inequality (108) for

$$\begin{aligned} \bar{X}_k & \triangleq \mathcal{H}_k^{(\kappa)} P_k^{(\kappa)}, \quad k \in \mathcal{K} \\ \bar{Y}_k & \triangleq [\mathcal{H}_k^{(\kappa)} P_k^{(\kappa)}]^2 + c\sigma I_{N_r} + c \sum_{i \neq k}^K [\mathcal{H}_k^{(\kappa)} P_i^{(\kappa)}]^2, \quad k \in \mathcal{K} \quad (75) \end{aligned}$$

yields the following tight minorant of $f_b^{(\kappa)}(\mathbf{P})$:

$$\begin{aligned} \tilde{f}_b^{(\kappa)}(\mathbf{P}) \triangleq & a^{(\kappa)} + 2 \sum_{k=1}^K \Re\{\langle \mathcal{A}_k \mathbf{P}_k \rangle\} - \sum_{k=1}^K \left\langle \mathcal{B}_k, \left([\mathcal{H}_k^{(\kappa)} \mathbf{P}_k]^2 \right. \right. \\ & \left. \left. + c \sum_{i \neq k}^K [\mathcal{H}_k^{(\kappa)} \mathbf{P}_i]^2 \right) \right\rangle \quad (76) \end{aligned}$$

$$= a^{(\kappa)} + 2 \sum_{k=1}^K \Re\{\langle \mathcal{A}_k \mathbf{P}_k \rangle\} - \sum_{k=1}^K \langle \mathcal{C}_k, [\mathbf{P}_k]^2 \rangle, \quad (77)$$

where we have

$$\begin{aligned} a^{(\kappa)} \triangleq & \ln \left| \Gamma_b^{(\kappa)}(P^{(\kappa)}) \right|^{-1} - \sum_{k=1}^K \left\langle (\Gamma_b^{(\kappa)}(\bar{X}, \bar{Y}))^{-1} \right. \\ & \left. \bar{X}_k^H \bar{Y}_k^{-1} \bar{X}_k \right\rangle - c\sigma \sum_{k=1}^K \langle \mathcal{B}_k \rangle, \quad (78) \end{aligned}$$

$$\begin{aligned} \mathcal{A}_k & \triangleq (\Gamma_b^{(\kappa)}(\bar{X}, \bar{Y}))^{-1} \bar{X}_k^H \bar{Y}_k^{-1} \mathcal{H}_k^{(\kappa)}, \\ \mathcal{B}_k & \triangleq \bar{Y}_k^{-1} \bar{X}_k (\Gamma_b^{(\kappa)}(\bar{X}, \bar{Y}))^{-1} \bar{X}_k^H \bar{Y}_k^{-1} \end{aligned}$$

and

$$\begin{aligned} \mathbb{C}^{N_t \times N_t} \ni \mathcal{C}_k \triangleq & (\mathcal{H}_k^{(\kappa)})^H \mathcal{B}_k \mathcal{H}_k^{(\kappa)} + c \sum_{i \neq k}^K (\mathcal{H}_i^{(\kappa)})^H \mathcal{B}_i \mathcal{H}_i^{(\kappa)}, \\ & k \in \mathcal{K}. \quad (79) \end{aligned}$$

We thus solve the following problem of tight minorant maximization of (73) to generate $P^{(\kappa+1)}$ verifying (72)

$$\max_{\mathbf{P}_k} \tilde{f}_b^{(\kappa)}(\mathbf{P}) \quad \text{s.t.} \quad (57b), \quad (80)$$

which admits the closed-form solution of

$$P_k^{(\kappa+1)} = \begin{cases} \mathcal{C}_k^{-1} \mathcal{A}_k^H & \text{if } \sum_{k \in \mathcal{K}} \langle [\mathcal{C}_k^{-1} \mathcal{A}_k^H]^2 \rangle \leq P, \\ (\mathcal{C}_k + \mu I_{N_t})^{-1} \mathcal{A}_k^H & \text{otherwise,} \end{cases} \quad (81)$$

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

$$\sum_{k=1}^K \langle [(\mathcal{C}_k + \mu I_{N_t})^{-1} \mathcal{A}_k^H]^2 \rangle = P.$$

B. PREs iteration

To generate a feasible $\theta^{(\kappa+1)}$ sothat

$$f_c(\theta^{(\kappa+1)}, P^{(\kappa+1)}) > f_c(\theta^{(\kappa)}, P^{(\kappa+1)}), \quad (82)$$

we address the following problem

$$\max_{\boldsymbol{\theta}} f_p^{(\kappa)}(\boldsymbol{\theta}) \triangleq \ln \left| \Gamma_p^{(\kappa)}(\boldsymbol{\theta}) \right|^{-1} \quad \text{s.t.} \quad (4), \quad (83)$$

with

$$\begin{aligned} \Gamma_p^{(\kappa)}(\boldsymbol{\theta}) &\triangleq \Gamma_c(\boldsymbol{\theta}, P^{(\kappa+1)}) \\ &= \sum_{k=1}^K \left[I_{N_r} - (\mathcal{H}_k(\boldsymbol{\theta}) P_k^{(\kappa+1)})^H \left([\mathcal{H}_k(\boldsymbol{\theta}) P_k^{(\kappa+1)}]^2 \right. \right. \\ &\quad \left. \left. + c \sigma I_{N_r} + c \sum_{i \neq k} [\mathcal{H}_k(\boldsymbol{\theta}) P_i^{(\kappa+1)}]^2 \right)^{-1} \mathcal{H}_k(\boldsymbol{\theta}) P_k^{(\kappa+1)} \right]. \end{aligned} \quad (84)$$

Applying the inequality (108) for

$$\begin{aligned} \bar{X}_k &\triangleq \mathcal{H}_k^{(\kappa)} P_k^{(\kappa+1)}, k \in \mathcal{K} \\ \bar{Y}_k &\triangleq [\mathcal{H}_k^{(\kappa)} P_k^{(\kappa+1)}]^2 + c \sigma I_{N_r} + c \sum_{i \neq k} [\mathcal{H}_k^{(\kappa)} P_i^{(\kappa+1)}]^2, \end{aligned} \quad (85)$$

yields the following tight minorant of $f_p^{(\kappa)}(\boldsymbol{\theta})$:

$$\begin{aligned} \tilde{f}_p^{(\kappa)}(\boldsymbol{\theta}) &\triangleq \tilde{a}^{(\kappa)} + 2 \sum_{k=1}^K \Re \{ \langle \tilde{\mathcal{A}}_k \mathcal{H}_k(\boldsymbol{\theta}) \rangle \} \\ &\quad - \sum_{k=1}^K \langle \tilde{\mathcal{B}}_k, \mathcal{H}_k(\boldsymbol{\theta}) [P_k^{(\kappa+1)}]^2 (\mathcal{H}_k(\boldsymbol{\theta}))^H \rangle \\ &\quad + c \sum_{i \neq k} \langle \mathcal{H}_k(\boldsymbol{\theta}) [P_i^{(\kappa+1)}]^2 (\mathcal{H}_k(\boldsymbol{\theta}))^H \rangle \end{aligned} \quad (86)$$

$$\begin{aligned} &= \tilde{a}^{(\kappa)} + 2 \sum_{k=1}^K \Re \{ \langle \tilde{\mathcal{A}}_k \mathcal{H}_k(\boldsymbol{\theta}) \rangle \} \\ &\quad - \sum_{k=1}^K \langle \tilde{\mathcal{B}}_k, \mathcal{H}_k(\boldsymbol{\theta}) \Psi_k^{(\kappa+1)} (\mathcal{H}_k(\boldsymbol{\theta}))^H \rangle, \end{aligned} \quad (87)$$

where

$$\begin{aligned} \tilde{a}^{(\kappa)} &\triangleq \ln \left| \Gamma_p^{(\kappa)}(\theta^{(\kappa)}) \right|^{-1} - \sum_{k=1}^K \langle (\Gamma_p^{(\kappa)}(\bar{X}, \bar{Y}))^{-1} \bar{X}_k^H \bar{Y}_k^{-1} \bar{X}_k \rangle \\ &\quad - c \sigma \sum_{k=1}^K \langle \tilde{\mathcal{B}}_k \rangle, \\ \tilde{\mathcal{A}}_k &\triangleq P_k^{(\kappa+1)} (\Gamma_p^{(\kappa)}(\bar{X}, \bar{Y}))^{-1} \bar{X}_k^H \bar{Y}_k^{-1}, \\ \tilde{\mathcal{B}}_k &\triangleq \bar{Y}_k^{-1} \bar{X}_k (\Gamma_p^{(\kappa)}(\bar{X}, \bar{Y}))^{-1} \bar{X}_k^H \bar{Y}_k^{-1}, \\ \Psi_k^{(\kappa+1)} &\triangleq [P_k^{(\kappa+1)}]^2 + c \sum_{i \neq k} [P_i^{(\kappa+1)}]^2. \end{aligned} \quad (88)$$

Upon recalling (42), we can arrive at

$$\langle \tilde{\mathcal{A}}_k \mathcal{H}_k(\boldsymbol{\theta}) \rangle = \alpha_{k,1}^{(\kappa+1)} + \sum_{n=1}^N \tilde{b}_{k,1}^{(\kappa+1)}(n) e^{j\theta_n} \quad (89)$$

in conjunction with

$$\alpha_{k,1}^{(\kappa+1)} \triangleq \langle \tilde{\mathcal{A}}_k \tilde{H}_{B,k} \rangle, k \in \mathcal{K}; \tilde{b}_{k,1}^{(\kappa+1)}(n) \triangleq \langle \tilde{\mathcal{A}}_k \mathcal{H}_{k,n} \rangle, n \in \mathcal{N}, \quad (90)$$

and

$$\begin{aligned} \langle \tilde{\mathcal{B}}_k, \mathcal{H}_k(\boldsymbol{\theta}) \Psi_k^{(\kappa+1)} (\mathcal{H}_k(\boldsymbol{\theta}))^H \rangle &= \alpha_{k,2}^{(\kappa+1)} + (e^{j\theta})^H \Phi_k^{(\kappa+1)} e^{j\theta} \\ &\quad + 2 \sum_{n=1}^N \Re \{ \tilde{b}_{k,2}^{(\kappa+1)}(n) e^{j\theta_n} \} \end{aligned} \quad (91)$$

with

$$\begin{aligned} \alpha_{k,2}^{(\kappa+1)} &\triangleq \langle \tilde{\mathcal{B}}_k (\tilde{H}_{B,k}) \Psi_k^{(\kappa+1)} \tilde{H}_{B,k}^H \rangle, \\ \tilde{b}_{k,2}^{(\kappa+1)}(n) &= \langle \tilde{\mathcal{B}}_k (\tilde{H}_{B,k}) \Psi_k^{(\kappa+1)} \mathcal{H}_{k,n}^H \rangle, n = 1, \dots, N, \\ \Phi_k^{(\kappa+1)}(n', n) &\triangleq \langle \tilde{\mathcal{B}}_k \mathcal{H}_{k,n} \Psi_k^{(\kappa+1)} \mathcal{H}_{k,n'}^H \rangle, (n', n) \in \mathcal{N} \times \mathcal{N}, \end{aligned} \quad (92)$$

Therefore,

$$\begin{aligned} \tilde{f}_p^{(\kappa)}(\boldsymbol{\theta}) &= \tilde{a}^{(\kappa+1)} + 2 \sum_{n=1}^N \Re \{ \tilde{b}^{(\kappa+1)}(n) e^{j\theta_n} \} \\ &\quad - (e^{j\theta})^H \Phi^{(\kappa+1)} e^{j\theta} \end{aligned} \quad (93)$$

with

$$\begin{aligned} \tilde{a}^{(\kappa+1)} &\triangleq \tilde{a}^{(\kappa)} + \sum_{k=1}^K \left[2 \Re \{ \alpha_{k,1}^{(\kappa+1)} \} - \alpha_{k,2}^{(\kappa+1)} \right] \\ \tilde{b}^{(\kappa+1)}(n) &\triangleq \sum_{k=1}^K \left[\tilde{b}_{k,1}^{(\kappa+1)}(n) - \tilde{b}_{k,2}^{(\kappa+1)}(n) \right], n \in \mathcal{N}, \\ \Phi^{(\kappa+1)} &\triangleq \sum_{k=1}^K \Phi_k^{(\kappa+1)}. \end{aligned} \quad (94)$$

Furthermore, a tight minorant of $\tilde{f}_p^{(\kappa)}(\boldsymbol{\theta})$ and $f_p^{(\kappa)}(\boldsymbol{\theta})$ is

$$\begin{aligned} \tilde{\tilde{f}}_p^{(\kappa)}(\boldsymbol{\theta}) &\triangleq \tilde{a}^{(\kappa+1)} + 2 \Re \left\{ \sum_{n=1}^N \left(\tilde{b}^{(\kappa+1)}(n) - \sum_{n'=1}^N e^{-j\theta_{n'}} \right) \right. \\ &\quad \left. \times \Phi^{(\kappa+1)}(n', n) + \lambda_{\max}(\Phi^{(\kappa+1)}) e^{-j\theta_n^{(\kappa)}} \right\} e^{j\theta_n} \\ &\quad - (e^{j\theta^{(\kappa)}})^H \Phi^{(\kappa+1)} e^{j\theta^{(\kappa)}} - 2 \lambda_{\max}(\Phi^{(\kappa+1)}) N \end{aligned} \quad (95)$$

We thus solve the following discrete problem of tight minorant maximization for (83) to generate $\theta^{(\kappa+1)}$ verifying (82):

$$\max_{\boldsymbol{\theta} \in \mathcal{B}^N} \tilde{\tilde{\varphi}}_{\sigma}^{(\kappa)}(\boldsymbol{\theta}), \quad (96)$$

which admits the closed-form solution of

$$\begin{aligned} \theta_n^{(\kappa+1)} &= 2\pi - \left[\angle \left(\tilde{b}^{(\kappa+1)}(n) - \sum_{m=1}^N e^{-j\theta_m^{(\kappa)}} \Phi^{(\kappa+1)}(m, n) \right. \right. \\ &\quad \left. \left. + \lambda_{\max}(\Phi^{(\kappa+1)}) e^{-j\theta_n^{(\kappa)}} \right) \right]_b, n \in \mathcal{N}. \end{aligned} \quad (97)$$

C. Soft max-min algorithm and its convergence

Algorithm 3 provides the pseudo-code for solving the problem (71) based on iterating (81) and (97). It follows from (72) and (82) that

$$f_c\left(\theta^{(\kappa+1)}, P^{(\kappa+1)}\right) > f_c\left(\theta^{(\kappa)}, P^{(\kappa)}\right),$$

i.e. Algorithm 3 provides a path-following procedure, which iteratively improves feasible points. As such, the sequence $\{(\theta^{(\kappa)}, P^{(\kappa)})\}$ converges to a limit point $(\bar{\theta}, \bar{P})$ by the Cauchy theorem.

Algorithm 3 Soft max-min algorithm to solve problem (71).

- 1: **Initialization.** Generate a feasible point $(P^{(0)}, \theta^{(0)})$ for problem (71). Set $\kappa = 0$.
 - 2: **Repeat until convergence of the objective in (71):** Generate $P^{(\kappa+1)}$ by (81), and $\theta^{(\kappa+1)}$ by (97). Reset $\kappa := \kappa + 1$.
 - 3: **Output** $(P^{(\kappa)}, \theta^{(\kappa)})$ and the achieved individual UE rates $r_k(\theta^{(\kappa)}, P^{(\kappa)})$, $k = 1, \dots, K$.
-

IV. SIMULATION RESULTS

Unless otherwise specified, we assume having $N_t = 10$ downlink transmit-antennas at the BS, $d = N_r = 2$, $N = 100$ PREs at the RIS, $P = 10$ dBm transmit power budget, and $b = 3$ bit-resolution of quantized PREs. The noise power is set to $\sigma = -114$ dBm, i.e. the noise power spectral density is -174 dBm/Hz over the transmission bandwidth of 1 MHz. The convergence tolerance of the proposed algorithms is set to 10^{-3} . The antenna gains of the AP (G_{AP}) and the RIS's elements (G_{RIS}) are set to 5 dBi. The AP and RIS are deployed at the coordinates of (20, 0, 25) and (0, 30, 40) in the three-dimensional (3D) space, respectively. We set $K = 10$ UEs, which are randomly distributed in a (120×120) area right of the AP and RIS. The spatial correlation matrix, which models the correlation between the RIS elements with respect to the users, is given by $[R_{R,k}]_{n,n'} = e^{j\pi(n-n') \sin \tilde{\phi}_k \sin \tilde{\theta}_k}$, where $\tilde{\phi}_k$ and $\tilde{\theta}_k$ represent the azimuth and elevation angle of UE k , respectively [21], [32].

We simulate a pair of practical scenarios. The first scenario assumes the availability of a direct link between the BS and UEs, while the second scenario assumes that the direct transmission path between the BS and UEs is blocked by some obstruction. We refer to them as **Scenario 1** and **Scenario 2**, respectively. Due to the absence of the direct path in Scenario 2, the distances between the nodes have to be kept lower, i.e., the UEs are randomly placed in a (60×60) area to the right of the BS and RIS.

A. Results for sum-rate capacity optimization

The SVD-based and SVD-free sum-rate capacity optimization Algorithms 1 and 2 conceived for solving the problem (11) are represented as “SVD-based SRC” and “SVD-free SRC”.

Fig. 2 shows the convergence of SVD-based SRC and SVD-free SRC at the transmit power budget $P = 0$ dBm.

Observe that both algorithms converge monotonically. In addition to observing the convergence performance with RIS ($b = \infty$), Fig. 2 also shows the convergence performance for the RIS-less case, i.e., without RIS, which implies that communication only takes place via the direct path between the BS and UEs. The sum-rate capacity slightly improves with the RIS's assistance and it takes more iterations to converge. In addition, we can observe that the SVD-based SRC converges more promptly than SVD-free SRC, which does not rely on SVD.

Table II shows that, on average, SVD-based SRC and SVD-free SRC require 37.8 and 72.9 iterations for convergence, respectively. Furthermore, the covariance iterations in SVD-based SRC are computationally more efficient than their counterparts in SVD-free SRC. This shows that SVD-based SRC may have a computational advantage over SVD-free SRC. Note that overall, both of them are computationally efficient, thanks to iterating by evaluating *closed-form expressions*.

Fig. 3 plots the sum-rate capacity versus the transmit power budget P of both SVD-based SRC and SVD-free SRC, while using $b = 3$ bit-resolution of quantized PREs. Fig. 3 also plots the sum-rate capacity of the RIS-less case, which shows that the sum-rate capacity of the RIS-less case and RIS-aided case ($b = 3$) is close to each other, especially at higher power budgets. This is because under **Scenario 1**, the RIS only offers a modest advantage due to the availability of the direct path between the BS and UEs. Fig. 3 shows that both algorithms improve the sum-rate capacity vs. the power budget P . Both algorithms perform similarly at lower power budgets, but SVD-based SRC outperforms SVD-free SRC at higher power budgets. We observe that by adjusting the convergence tolerance to 10^{-4} , SVD-free SRC starts performing closer to SVD-based SRC. However, we decide to use the convergence tolerance of 10^{-3} in our simulations because 10^{-4} requires a very large number of iterations for convergence and the resultant improvement in the sum-rate capacity is negligible.

Figures 4-7 consider **Scenario 2**, which assumes that the direct transmission path between the BS and UEs is blocked by some obstruction. Fig. 4 plots the sum-rate capacity versus the transmit power budget P by SVD-based SRC and SVD-free SRC. For comparison, Fig. 4 also plots the sum-rate capacity without PRE optimization, which assumes having some random phase-shifts θ for PREs. We can observe the advantage of the proposed SVD-based SRC and SVD-free SRC, which employ joint optimization of \mathbf{X} and θ over the “without PRE optimization” scenario.

Fig. 4 shows that under **Scenario 2**, SVD-free SRC slightly outperforms SVD-based SRC at higher power budgets. This observation is different from the one that we had under **Scenario 1** and it can be explained as follows. Under **Scenario 1**, the RIS only offers a minor advantage because there is a direct path between the BS and UEs, as demonstrated in Fig. 3. Therefore, we may conclude that under **Scenario 1**, the problem (11) is effectively about optimizing the covariance \mathbf{X} . The optimization of \mathbf{X} alone in (13) constitutes a convex problem and SVD-based SRC offers the globally optimal solution. That is why SVD-based SRC was seen to outperform

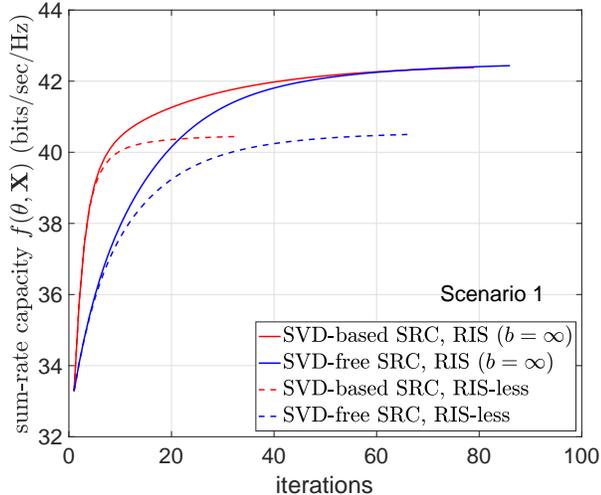


Fig. 2: Convergence of SVD-based SRC and SVD-free SRC.

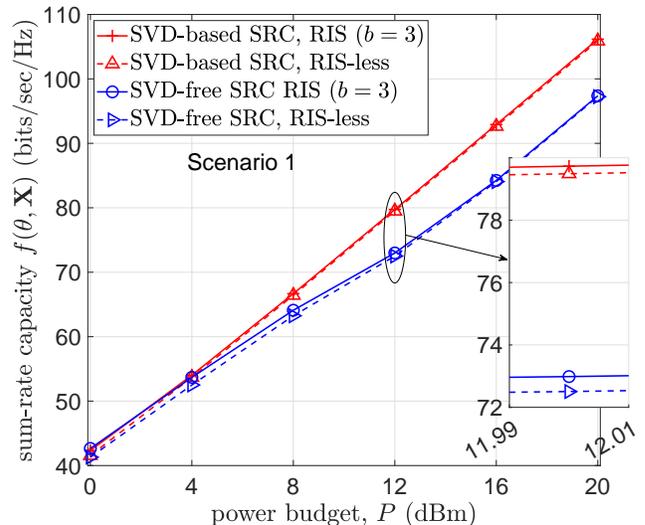


Fig. 3: Sum-rate capacity versus transmit power budget P under Scenario 1.

TABLE II: Complexity of SVD-based SRC and SVD-free SRC for N PREs, N_r receive antenna elements and K UEs.

	SVD-based Alg. 1		SVD-free Alg. 2	
	Covariance iteration	PREs iteration	Covariance iteration	PREs iteration
Computational Complexity	$\mathcal{O}(N_r K)$	$\mathcal{O}(N)$	$\mathcal{O}(N_r \log_2(N_r) K)$	$\mathcal{O}(N)$
Average # of iterations for convergence	37.8		72.9	

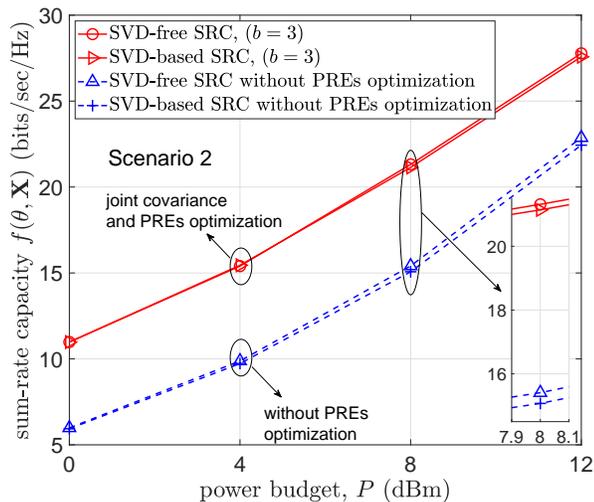


Fig. 4: Sum-rate capacity versus transmit power budget P under Scenario 2.

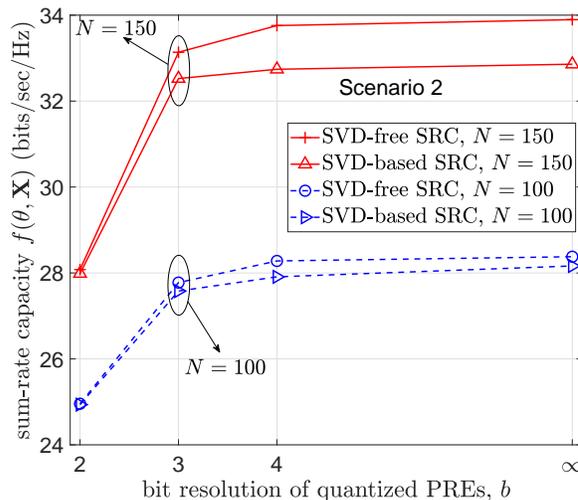


Fig. 5: Sum-rate capacity versus quantized PREs' resolution under Scenario 2.

SVD-free SRC under Scenario 1. However, under Scenario 2, where we totally rely on the RIS-assisted path, RIS promises a clear advantage. Thus, jointly optimizing the covariance \mathbf{X} and phase-shifts θ is important, as demonstrated by Fig. 4. Since, the joint optimization of \mathbf{X} and θ in (11) is a non-convex problem, both SVD-based SRC and SVD-free SRC can only promise to find a locally optimal solution at best and achieve similar performance, as shown in Fig. 4.

Fig. 5 plots the sum-rate capacity versus the bit-resolution b of PREs. Two different values of the number of RIS elements $N = \{100, 150\}$ are considered. Fig. 5 shows that the sum-rate capacity increases upon increasing N due to the increase in

the number of resources. Observe that the performance gap between the sum-rate capacity of 3-bit quantized PREs and ∞ -resolution PREs is quite small.

Fig. 6 plots the sum-rate capacity versus the number of transmit antennas N_t at the BS for both SVD-based SRC and SVD-free SRC under Scenario 2, where $N_r = \{1, 2\}$ are considered at the UE. Fig. 6 shows that the sum-rate capacity increases with N_t and N_r due to the increase in the number of resources. In addition, we can observe that the performance advantage of employing $N_r = 2$ antennas over single-antenna UEs increases with the increase in N_t . Fig. 7 plots the sum-rate capacity versus the number of UEs K . Observe that the

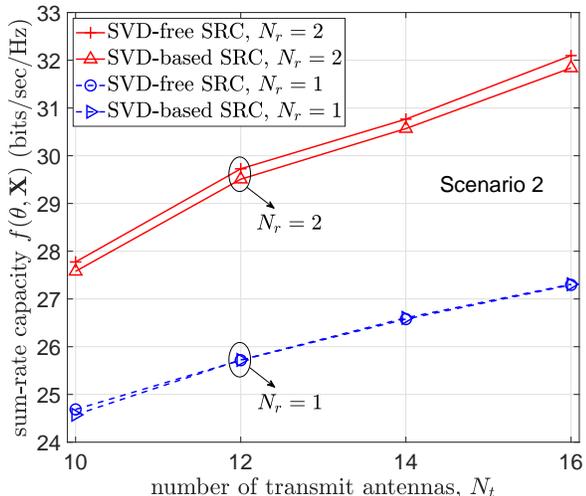


Fig. 6: Sum-rate capacity versus number of BS antennas N_t under Scenario 2.

proposed algorithms accommodate an increased number of UEs and the sum-rate capacity increases with the increase in K .

B. Results for max-min rate optimization

In this subsection, we characterize the performance of the max-min rate optimization Alg. 3, which solves the problem (71) and it is represented by the label “Soft max-min”. The problem (71) is based on function Γ_c , which is defined in (62) and depends on the value of c . In order to choose the best value of c , we observe the minimum UE rate by solving problem (71) for varying values of simulation parameters. We select $c = 0.1$, because it overall achieves better minimum UE rate and quicker convergence than the other values over a wide range of simulation parameters.

Fig. 8 shows the minimum UE rate versus the Soft max-min iterations for $c = \{0.05, 0.1\}$ and $P = 0$ dBm. We only consider $c = 0.05$ and $c = 0.1$ in Fig. 8, because we find them better than other values of c both in terms of convergence and throughput. The results of Fig. 8 consider (i) Scenario 1 with RIS ($b = \infty$), (ii) Scenario 1 with RIS-less case, and (iii) Scenario 2 with RIS ($b = \infty$). Fig. 8 shows that the minimum UE-rate converges monotonically in all the cases. We observe that although the minimum UE rate of $c = 0.05$ is slightly better than that of $c = 0.1$ under Scenario 1, we observe the clear supremacy of the latter under Scenario 2.

Fig. 9 shows the ratio of the maximum UE rate to the minimum UE rate, which is a measure of rate-fairness among the UEs, versus the number of iterations in Alg. 3 for $c = 0.1$ and $P = 0$ dBm. Fig. 9 shows that the ratio is smaller for Scenario 1 than that for Scenario 2, because the former has the direct link between the BS and UEs and this is helpful for achieving a higher minimum UE rate and lower deviation among the UEs rates. Fig. 9 also shows that the Soft max-min achieves quite a small deviation among the UE rates, where the ratio of the maximum UE rate to the minimum UE rate is in the range of (1.5, 2).

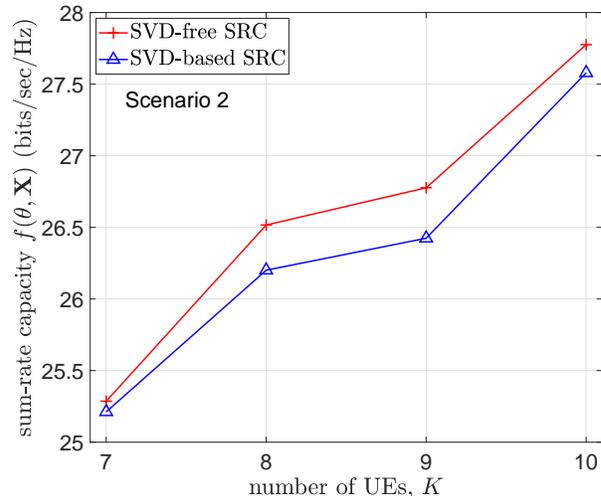


Fig. 7: Sum-rate capacity versus quantized PREs’ resolution under Scenario 2.

Fig. 10 considers the RIS-less case of Scenario 1 and compares the minimum UE rate of the proposed Soft max-min and that of the convex solver based approach. More particularly, by using off-the-shelf convex solvers, we solve the (i) max-min rate optimization problem (57) and (ii) problem (71) under the RIS-less case. These problems are non-convex and we use the inequalities (101) and (108) to develop iterative convex approximation procedures for solving (57) and (71), respectively. Fig. 10 shows that Soft max-min and the solution of problem (71) using a convex-solver based approach yield similar minimum user-rates. However, solving the max-min rate optimization problem of (57) using a convex-solver based approach yields a higher minimum user-rate. This is expected, because our proposed Soft max-min does not explicitly solve the max-min rate optimization problem. Therefore, we cannot expect it to yield the same minimum UE rate as that achieved by explicitly solving the max-min rate optimization problem. Nonetheless, the proposed Soft max-min optimization is beneficial because it is computationally efficient compared to the convex solver based approaches, as shown in Table III. This is because our proposed Soft max-min iterates by evaluating *closed-form expressions* to generate feasible points. Secondly, the achievable minimum UE rate of Soft max-min is not far from that of the max-min rate optimization of (57), just 16.3% lower at $P = 12$ dBm, for example. Further features of our proposed Soft max-min can be seen from Fig. 11, which plots the individual UE rates at $P = 12$ dBm. We can observe that only three out of $K = 10$ UEs are slightly behind in terms of achievable rate, when compared to the max-min rate optimization of (57).

Fig. 12 plots the ratio of the maximum UE rate to the minimum UE rate versus P under Scenario 2. Three different quantized PREs associated with $b = \{2, 3, \infty\}$, are considered. We can observe that this rate-ratio, which reflects the spread of the UE rates, decreases upon increasing b . A benefit of our proposed Soft max-min is that the performance gap between 2-bit quantized PREs and ∞ -resolution PREs is very small. The rate-ratio also decreases with the increase in P due to the

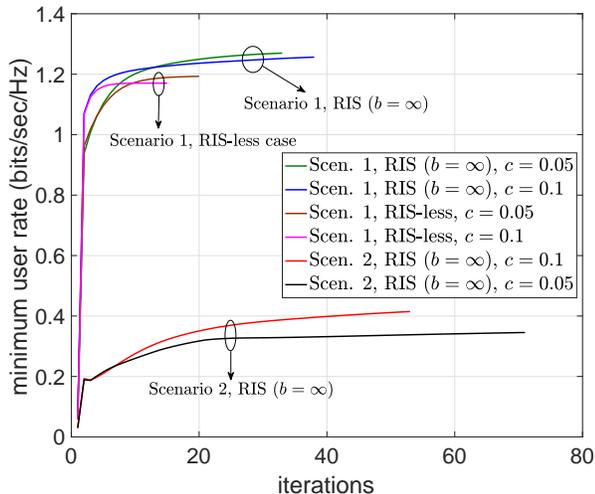


Fig. 8: The achievable minimum UE rate versus the number of Soft max-min iterations.

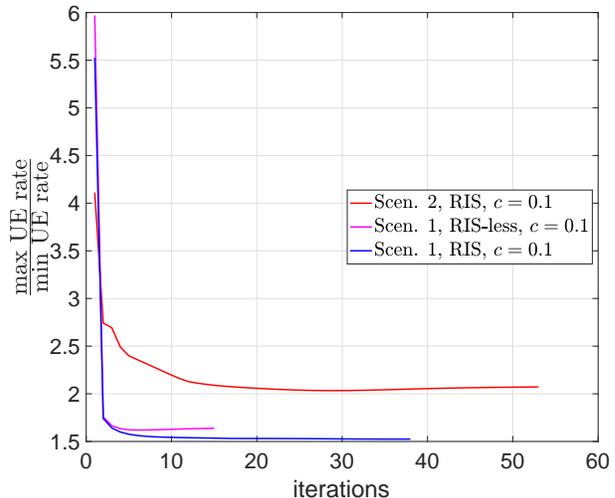


Fig. 9: Ratio of maximum UE rate to minimum UE rate versus the number of Soft max-min iterations.

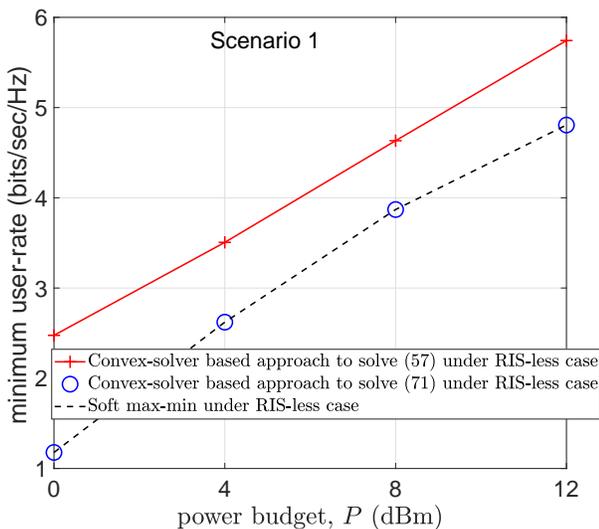


Fig. 10: Comparison of the proposed Soft max-min and of the convex solver based approach.

TABLE III: Computational complexity comparison of Soft max-min and of the convex solver based approaches for a RIS-less case.

	Soft max-min under RIS-less case	Convex solver approach to solve (57) under RIS-less case	Convex solver approach to solve (71) under RIS-less case
Computational Complexity	$\mathcal{O}(N_t \log_2(N_t)K)$	$\mathcal{O}((N_t N_r K)^3)$	$\mathcal{O}((N_t N_r K)^3)$
Average # of iterations for convergence	43.8	18.2	43.8
Average computation time per iteration	0.017 sec	1.104 sec	1 sec

increase in power budget. It is noteworthy that the rate-ratio becomes even lower than 1.3 at $P = 20$ dBm.

Fig. 13 plots the minimum UE rate versus the number of transmit antennas N_t at the BS for $N_r = \{1, 2\}$ in Scenario 2. Fig. 13 shows that the minimum UE rate increases with the increase in N_t , N_r , or b , due to the increase in the number of resources. We can observe that the performance gap between the single-antenna and two-antenna UEs reduces with the increase of N_t for both $b = 3$ bit-resolution or ∞ -resolution PREs.

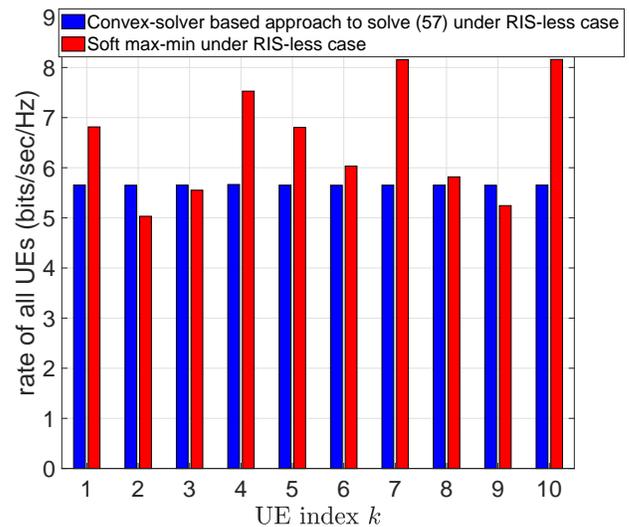


Fig. 11: UEs rate distribution of the Soft max-min and of the convex solver based approach.

TABLE III: Computational complexity comparison of Soft max-min and of the convex solver based approaches for a RIS-less case.

Fig. 14 plots the achievable minimum rate under the bit-resolution b of PREs in Scenario 2. Three different values of the number of RIS elements $N = \{100, 150, 200\}$ are considered. Fig. 14 shows that the minimum rate increases with N . However, the improvement in the achievable minimum rate becomes marginal upon increasing N from 150 to 200 compared to that when we increase N from 100 to 150. Fig. 14 shows that the quantized PREs designed by the proposed Soft max-min is very efficient because the minimum UE rate of $b = 4$ bit-resolution is not far from that of ∞ -resolution

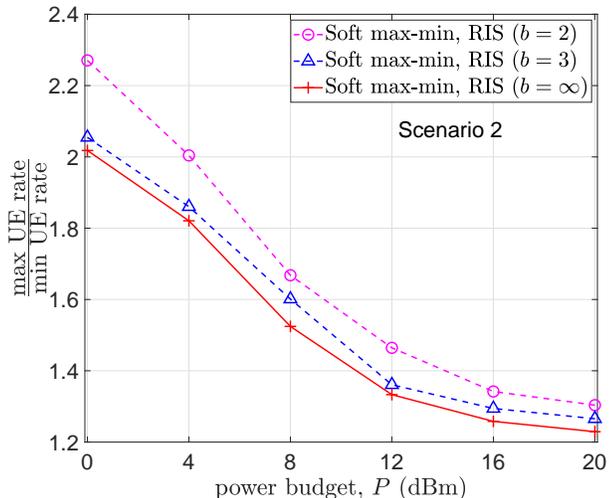


Fig. 12: Ratio of maximum to minimum UE rate versus P under Scenario 2.

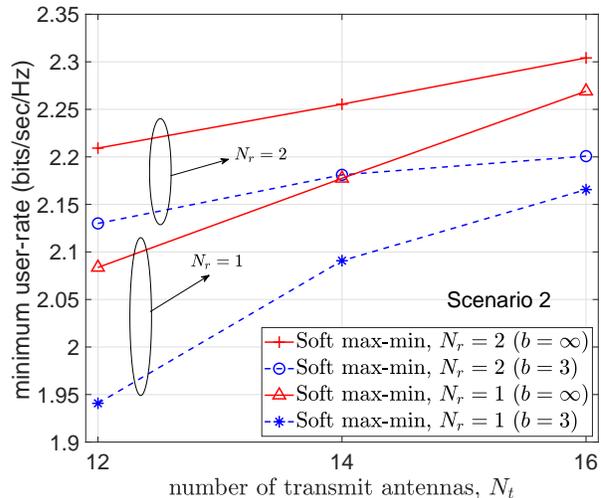


Fig. 13: Minimum UE rate by Alg. 3 versus N_t under Scenario 2.

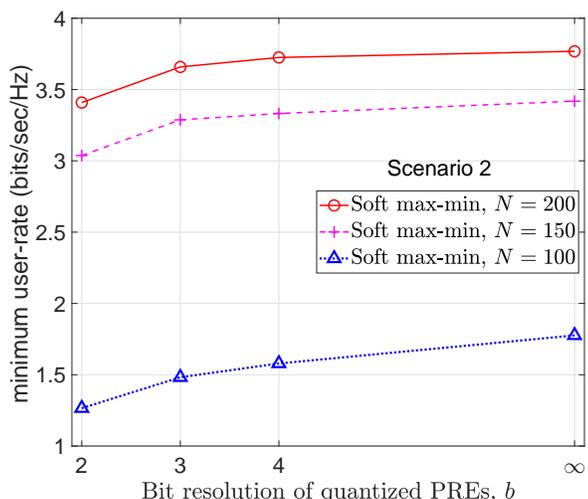


Fig. 14: Minimum UE rate by Soft max-min versus PREs' resolution under Scenario 2.

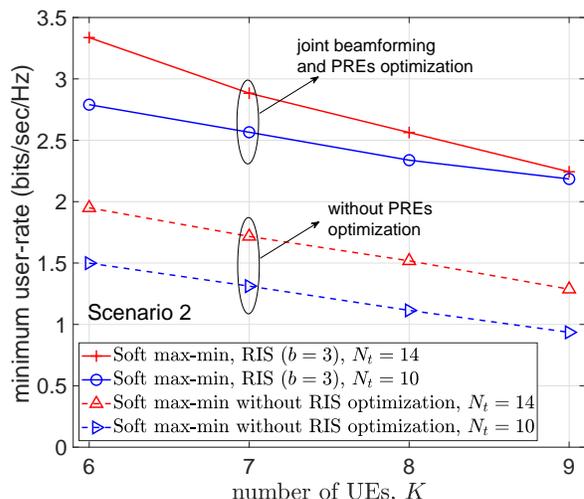


Fig. 15: Minimum UE rate by Soft max-min versus K under Scenario 2.

PREs.

Fig. 15 plots the minimum UE rate versus the number of UEs K . We can observe that the minimum UE rate decreases with the increase in K . This is because as we increase K , the competition among the UEs increases for the limited resources, i.e., under fixed values of P and N_t . For comparison, Fig. 15 also plots the minimum UE rate without PRE optimization. We can clearly observe the advantage of the proposed Soft max-min, which employs joint beamforming and PREs optimization over that “without PRE optimization” case.

V. CONCLUSIONS

Determining the capacity and maximizing the users' minimum rate in RIS-aided BCs have been regarded as computationally intractable problems due to their large-scale mixed discrete-continuous nature. This paper has demonstrated that both of these challenges can be efficiently addressed using our

scalable path-following algorithms, opening up new avenues for exploring large-scale RIS-aided BCs. Their extensions to scenarios based on channel statistics only for the sake of avoiding the potential bottleneck caused by the channel overhead required for channel estimation in large scale networks are of great interest and are under current study.

APPENDIX: MATRIX INEQUALITIES

Theorem 1: [33, Th.2] For the $A > 0$ the function $\ln|A + H\mathbf{X}^{-1}H^H|$ is convex in $\mathbf{X} > 0$.

Theorem 2: [18] The matrix-valued function $\mathbf{X}^H\mathbf{Y}^{-1}\mathbf{X}$ is convex over the domain $\{(\mathbf{X}, \mathbf{Y}) : \mathbf{Y} > 0\}$, i.e. the following matrix inequality holds for all $\alpha \geq 0$ and $\beta \geq 0$ with $\alpha + \beta = 1$, and $\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_1 > 0$ and $\mathbf{Y}_2 > 0$

$$(\alpha\mathbf{X}_1 + \beta\mathbf{X}_2)^H (\alpha\mathbf{Y}_1 + \beta\mathbf{Y}_2)^{-1} (\alpha\mathbf{X}_1 + \beta\mathbf{X}_2) \preceq \alpha\mathbf{X}_1^H \mathbf{Y}_1^{-1} \mathbf{X}_1 + \beta\mathbf{X}_2^H \mathbf{Y}_2^{-1} \mathbf{X}_2. \quad (98)$$

Theorem 3: For the $A \succ 0$ function $\ln |A + \sum_{i=1}^n H_i \mathbf{X}_i^{-1} H_i^H|$ is convex in $\mathbf{X}_i \succ 0$, $i = 1, 2, \dots, n$, so the following inequalities hold for all $\mathbf{X}_i \succ 0$ and $\bar{X}_i \succ 0$, $i = 1, \dots, n$:

$$\begin{aligned} & \ln \left| A + \sum_{i=1}^n H_i \mathbf{X}_i^{-1} H_i^H \right| \geq \\ & \ln \left| A + \sum_{i=1}^n H_i \bar{X}_i^{-1} H_i^H \right| \\ & + \sum_{i=1}^n \left\langle \bar{X}_i^{-1} H_i^H \left(A + \sum_{i=1}^n H_i \bar{X}_i^{-1} H_i^H \right)^{-1} \right. \\ & \quad \left. \times H_i \bar{X}_i^{-1}, \mathbf{X}_i - \bar{X}_i \right\rangle. \end{aligned} \quad (99)$$

Proof: Rewrite

$$\begin{aligned} \ln |A + \sum_{i=1}^n H_i \mathbf{X}_i^{-1} H_i^H| &= \ln |A + H \mathbf{X}^{-1} H^H|, \\ H &= [H_1 \ H_2 \ \dots \ H_n], \\ \mathbf{X} &= \text{diag}[\mathbf{X}_i]_{i=1,2,\dots,n} \end{aligned}$$

which is convex in \mathbf{X} according to Theorem 1. \square

By resetting $\mathbf{X}_i^{-1} \leftarrow \mathbf{X}_i$ and $\bar{X}_i \leftarrow \bar{X}_i^{-1}$ in (99), we obtain the following inequality for all $\mathbf{X}_i \succ 0$ and $\bar{X}_i \succ 0$, $i = 1, \dots, n$:

$$\begin{aligned} & \ln \left| A + \sum_{i=1}^n H_i \mathbf{X}_i H_i^H \right| \geq \\ & \ln \left| A + \sum_{i=1}^n H_i \bar{X}_i H_i^H \right| \\ & + \sum_{i=1}^n \left\langle \bar{X}_i H_i^H \left(A + \sum_{i=1}^n H_i \bar{X}_i H_i^H \right)^{-1} \right. \\ & \quad \left. \times H_i \bar{X}_i, \mathbf{X}_i^{-1} - \bar{X}_i^{-1} \right\rangle. \end{aligned} \quad (100)$$

Another inequality is for all \mathbf{V} and \bar{V} of size $n \times m$ and $\mathbf{Y} \succ 0$ and $\bar{Y} \succ 0$ of size $n \times n$ [34], [35]:

$$\begin{aligned} \ln |I_n + [\mathbf{V}]^2 (\mathbf{Y})^{-1}| &\geq \ln |I_n + [\bar{V}]^2 (\bar{Y})^{-1}| - \langle [\bar{V}]^2 (\bar{Y})^{-1} \\ & \quad + 2\Re\{\langle \bar{V}^H (\bar{Y})^{-1} \mathbf{V} \rangle\} \\ & \quad - \langle (\bar{Y})^{-1} - (\bar{Y} + [\bar{V}]^2)^{-1}, [\mathbf{V}]^2 + \mathbf{Y} \rangle. \end{aligned} \quad (101)$$

Considering both sides of (101) as functions of the variables (\mathbf{V}, \mathbf{Y}) , they match at (\bar{V}, \bar{Y}) , i.e. the function defined by the right-hand-side (RHS), which is concave quadratic because $(\bar{Y})^{-1} - (\bar{Y} + [\bar{V}]^2)^{-1} \succeq 0$, provides a tight minorant of the log-determinant function defined by the left-hand-side (LHS) at (\bar{V}, \bar{Y}) .

A particular case of (101) is the following inequality for all $\mathbf{V} \in \mathbb{C}^{n \times m}$ and $\bar{V} \in \mathbb{C}^{n \times m}$ and $\sigma > 0$:

$$\begin{aligned} \ln |I_n + [\mathbf{V}]^2 / \sigma| &\geq \ln |I_n + [\bar{V}]^2 / \sigma| - \frac{1}{\sigma} \langle [\bar{V}]^2 \rangle \\ & \quad + \frac{2}{\sigma} \Re\{\langle \bar{V}^H \mathbf{V} \rangle\} - \langle \sigma^{-1} I_n \\ & \quad - (\sigma I_n + [\bar{V}]^2)^{-1}, [\mathbf{V}]^2 + \sigma I_n \rangle. \end{aligned} \quad (102)$$

For $\mathbf{X} \triangleq (\mathbf{X}_1, \dots, \mathbf{X}_K)$ and $\mathbf{Y} \triangleq (\mathbf{Y}_1, \dots, \mathbf{Y}_K)$ with $\mathbf{X}_k \in \mathbb{C}^{n \times m}$ and $0 \prec \mathbf{Y}_k \in \mathbb{C}^{n \times n}$, $k = 1, \dots, K$, consider the function

$$\psi(\mathbf{X}, \mathbf{Y}) \triangleq \ln |\Gamma(\mathbf{X}, \mathbf{Y})| \quad (103)$$

in conjunction

$$\Gamma(\mathbf{X}, \mathbf{Y}) \triangleq \sum_{k=1}^K (I_m - \mathbf{X}_k^H \mathbf{Y}_k^{-1} \mathbf{X}_k) \quad (104)$$

in the domain

$$\{[\mathbf{X}_k]^2 \prec \mathbf{Y}_k, k = 1, \dots, K\}. \quad (105)$$

Theorem 4: In the domain constrained by (105), the function $\psi(\mathbf{X}, \mathbf{Y})$ is concave with its differential at a point (\bar{X}, \bar{Y}) given by

$$\begin{aligned} & \left\langle \psi(\bar{X}, \bar{Y}), (\mathbf{X}, \mathbf{Y}) - (\bar{X}, \bar{Y}) \right\rangle = \\ & \left\langle \Gamma^{-1}(\bar{X}, \bar{Y}) \sum_{k=1}^K \left(-\bar{X}_k^H \bar{Y}_k^{-1} \mathbf{X}_k - \mathbf{X}_k^H \bar{Y}_k^{-1} \bar{X}_k \right. \right. \\ & \quad \left. \left. + \bar{X}_k^H \bar{Y}_k^{-1} \bar{X}_k + \bar{X}_k^H \bar{Y}_k^{-1} \mathbf{Y}_k \bar{Y}_k^{-1} \bar{X}_k \right) \right\rangle. \end{aligned} \quad (106)$$

Thus, the following inequality holds for all (\mathbf{X}, \mathbf{Y}) and (\bar{X}, \bar{Y}) in the domain constrained by (105):

$$\begin{aligned} \ln |\Gamma(\mathbf{X}, \mathbf{Y})| &\leq \ln |\Gamma(\bar{X}, \bar{Y})| + \sum_{k=1}^K \langle \Gamma^{-1}(\bar{X}, \bar{Y}) \bar{X}_k^H \bar{Y}_k^{-1} \bar{X}_k \rangle \\ & \quad - 2 \sum_{k=1}^K \Re\{\langle \Gamma^{-1}(\bar{X}, \bar{Y}) \bar{X}_k^H \bar{Y}_k^{-1} \mathbf{X}_k \rangle\} \\ & \quad + \sum_{k=1}^K \langle \bar{Y}_k^{-1} \bar{X}_k \Gamma^{-1}(\bar{X}, \bar{Y}) \bar{X}_k^H \bar{Y}_k^{-1} \mathbf{Y}_k \rangle, \end{aligned} \quad (107)$$

or equivalently,

$$\begin{aligned} \ln |\Gamma(\mathbf{X}, \mathbf{Y})|^{-1} &\geq \ln |\Gamma(\bar{X}, \bar{Y})|^{-1} - \sum_{k=1}^K \langle \Gamma^{-1}(\bar{X}, \bar{Y}) \bar{X}_k^H \bar{Y}_k^{-1} \\ & \quad \times \bar{X}_k \rangle + 2 \sum_{k=1}^K \Re\{\langle \Gamma^{-1}(\bar{X}, \bar{Y}) \bar{X}_k^H \bar{Y}_k^{-1} \mathbf{X}_k \rangle\} \\ & \quad - \sum_{k=1}^K \langle \bar{Y}_k^{-1} \bar{X}_k \Gamma^{-1}(\bar{X}, \bar{Y}) \bar{X}_k^H \bar{Y}_k^{-1} \mathbf{Y}_k \rangle. \end{aligned} \quad (108)$$

Proof: By Theorem 2, we have:

$$\Gamma(\alpha \mathbf{X} + \beta \bar{X}, \alpha \mathbf{Y} + \beta \bar{Y}) \preceq \alpha \Gamma(\mathbf{X}, \mathbf{Y}) + \beta \Gamma(\bar{X}, \bar{Y})$$

for all (\mathbf{X}, \mathbf{Y}) and (\bar{X}, \bar{Y}) in the domain constrained by (105) and $\alpha \geq 0$ and $\beta \geq 0$ with $\alpha + \beta = 1$. Therefore,

$$\begin{aligned} \ln |\Gamma(\alpha \mathbf{X} + \beta \bar{X}, \alpha \mathbf{Y} + \beta \bar{Y})| &\leq \\ \ln |\alpha \Gamma(\mathbf{X}, \mathbf{Y}) + \beta \Gamma(\bar{X}, \bar{Y})| &\leq \end{aligned} \quad (109)$$

$$\alpha \ln |\Gamma(\mathbf{X}, \mathbf{Y})| + \beta \ln |\Gamma(\bar{X}, \bar{Y})|, \quad (110)$$

where the second inequality follows from the concavity of the function $\ln \det$. This shows the concavity of the log-determinant function. Then (107) is a fundamental property of concave functions [14]. \square

REFERENCES

- [1] M. D. Renzo, F. H. Danufane, and S. Tret'yakov, "Communication models for reconfigurable intelligent surfaces: From surface electromagnetics to wireless networks optimization," *Proc. IEEE*, vol. 110, pp. 1164–1208, Sept. 2022.
- [2] H. Zhang, B. Di, K. Bian, Z. Han, H. V. Poor, and L. Song, "Toward ubiquitous sensing and localization with reconfigurable intelligent surfaces," *Proc. IEEE*, vol. 110, pp. 1401–1422, Sept. 2022.
- [3] C. Pan *et al.*, "Multicell MIMO communications relying on intelligent reflecting surface," *IEEE Trans. Wirel. Commun.*, vol. 19, pp. 5218–5233, Aug. 2020.
- [4] 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.
- [5] 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, pp. 295–309, Jan. 2022.
- [6] A. A. Nasir, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, "Low-resolution RIS-aided multiuser MIMO signaling," *IEEE Trans. Commun.*, vol. 70, pp. 6517–6531, Oct. 2022.
- [7] L. Wei *et al.*, "Joint channel estimation and signal recovery for RIS-empowered multiuser communications," *IEEE Trans. Commun.*, vol. 70, pp. 4640–4655, July 2022.
- [8] M. Abughalwa, H. D. Tuan, D. N. Nguyen, H. V. Poor, and L. Hanzo, "Finite-blocklength RIS-aided transmit beamforming," *IEEE Trans. Vehic. Techn.*, vol. 71, pp. 12374–12379, Nov. 2022.
- [9] X. Zhai, G. Han, Y. Cai, and L. Hanzo, "Beamforming design based on two-stage stochastic optimization for RIS-assisted over-the-air computation systems," *IEEE Internet Things J.*, vol. 9, pp. 5474–5488, Apr. 2022.
- [10] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Info. Theory*, vol. 50, no. 4, pp. 768–783–1580, 2004.
- [11] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wirel. Commun. Lett.*, vol. 9, pp. 581–585, May 2020.
- [12] W. Tang *et al.*, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Trans. Wirel. Commun.*, vol. 20, pp. 421–439, Jan. 2021.
- [13] M. Najafi, V. Jamali, R. Schober, and H. V. Poor, "Physics-based modeling and scalable optimization of large intelligent reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, pp. 2673–2691, Apr. 2021.
- [14] H. Tuy, *Convex Analysis and Global Optimization (second edition)*. Springer International, 2017.
- [15] 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, pp. 6050–6065, Sept. 2022.
- [16] W. Zhu, H. D. Tuan, E. Dutkiewicz, and L. Hanzo, "Collaborative beamforming aided fog radio access networks," *IEEE Trans. Veh. Techn.*, vol. 71, pp. 7805–7820, Jul. 2022.
- [17] H. D. Tuan, H. H. Kha, H. H. Nguyen, and V. J. Luong, "Optimized training sequences for spatially correlated MIMO-OFDM," *IEEE Trans. Wirel. Commun.*, vol. 9, pp. 2768–2778, Sept. 2010.
- [18] U. Rashid, H. D. Tuan, and H. H. Nguyen, "Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks," *IEEE Trans. Comm.*, vol. 62, pp. 488–499, Feb. 2014.
- [19] Z. Sheng, H. D. Tuan, H. H. Nguyen, and M. Debbah, "Optimal training sequences for large-scale MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3329–3343, 2017.
- [20] M. Di Renzo, M. Debbah, and *et al.*, "Smart radio environments empowered by AI reconfigurable meta-surfaces: An idea whose time has come," *EURASIP J. Wirel. Commun. Network.*, no. 1, p. 129, 2019.
- [21] Q. U. A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M. S. Alouini, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 12, pp. 7748–7764, 2020.
- [22] E. Bjornson, O. Ozdogan, and E. G. Larsson, "Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?," *IEEE Wirel. Commun. Lett.*, vol. 9, no. 2, pp. 244–248, 2020.
- [23] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Tran. Commun.*, vol. 69, pp. 4144–4157, June 2021.
- [24] A. L. Swindlehurst, G. Zhou, R. Liu, C. Pan, and M. Li, "Channel estimation with reconfigurable intelligent surfaces: A general framework," *Proc. IEEE*, vol. 110, pp. 1312–1338, Sept. 2022.
- [25] J. An *et al.*, "Joint training of the superimposed direct and reflected links in reconfigurable intelligent surface assisted multiuser communications," *IEEE Trans. Green Commun. Network.*, vol. 6, pp. 739–754, June 2022.
- [26] J. An, C. Xu, L. Gan, and L. Hanzo, "Low-complexity channel estimation and passive beamforming for RIS-assisted MIMO systems relying on discrete phase shifts," *IEEE Trans. Commun.*, vol. 70, pp. 1245–1260, Feb. 2022.
- [27] C. Xu *et al.*, "Channel estimation for reconfigurable intelligent surface assisted high-mobility wireless systems," *IEEE Trans. Veh. Techn.*, vol. 72, pp. 718–734, Jan. 2023.
- [28] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," *IEEE Trans. Info. Theory*, vol. 49, pp. 2658–2668, Oct. 2003.
- [29] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Info. Theory*, vol. 51, no. 4, pp. 1570–1580, 2005.
- [30] P. He and L. Zhao, "Correction of convergence proof for iterative water-filling in Gaussian MIMO broadcast channels," *IEEE Trans. Info. Theory*, vol. 57, pp. 2539–2543, Apr. 2011.
- [31] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Infor. Theory*, vol. 50, pp. 145–152, Jan 2004.
- [32] Q.-U.-A. Nadeem, A. Kammoun, M. Debbah, and M.-S. Alouini, "A generalized spatial correlation model for 3D MIMO channels based on the Fourier coefficients of power spectrums," *IEEE Trans. Signal Process.*, vol. 63, pp. 3671–3686, Jul. 2015.
- [33] 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, Nov. 2019.
- [34] 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.
- [35] H. D. Tuan, H. H. M. Tam, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Superposition signaling in broadcast interference networks," *IEEE Trans. Commun.*, vol. 65, pp. 4646–4656, Nov. 2017.