**QUEEN'S
UNIVERSITY
BELFAST**

Multi-User Regularized Zero-Forcing Beamforming

Tuan, H. D., Duong, T. Q., & Poor, H. V. (2019). Multi-User Regularized Zero-Forcing Beamforming. *IEEE Transactions on Signal Processing*, 67(11), 2839-2853. <https://doi.org/10.1109/TSP.2019.2905833>

Published in:
IEEE Transactions on Signal Processing

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights
© 2019 IEEE.

This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Multi-user regularized zero-forcing beamforming

L. D. Nguyen, H. D. Tuan, T. Q. Duong, and H. V. Poor

Abstract—Regularized zero-forcing beamforming (RZFB) is an interesting class of linear signal processing problems, which is very attractive for use in large-scale communication networks due to its simple visualization as a straightforward extension of the well-accepted zero-forcing beamforming (ZFB). However, unlike ZFB, which is multi-user interference free, RZFB must manage multi-user interference to achieve its high throughput performance. Most of existing works focus on the performance analysis of particular RZBF schemes such as the equip-power allocated RZBF under a fixed regularization parameter. This paper is the first work to consider the joint design of power allocation and regularization parameter for RZFB to maximize the worst users' throughput or the quality-of-service awarded energy efficiency under a fixed transmit power constraint. Such designs pose very computationally challenging optimization problems, for which the paper proposes two-stage optimization algorithms of low computational complexity. Their computational and performance efficiencies are substantiated through numerical examples.

Index Terms—Multi-antenna communications, regularized zero-forcing beamforming, nonconvex optimization algorithms

I. INTRODUCTION

Linear beamforming is considered as the main driver of 5G signal processing techniques [1]. Zero-forcing beamforming (ZFB) [2] is a very popular class of linear beamforming due to its simplicity in design [3], [4]. Based on the channel matrix right inversion, ZFB is most attractive when the number of transmit antennas is much larger than the number of users (UEs) as the case of massive multiple input multiple output (MIMO) serving a few UEs [5]. Furthermore, ZFB for heterogeneous networks has been considered in [6] and [7]. Reference [8] presented a multi-cell cell-edge-aware ZFB to suppress interference at cell-edge UEs, where the main results are derivations for the number of UEs and antennas going to infinity. Reference [9] considered the two-stage beamformer design for massive MIMO to maximize the signal-to-leakage-plus-noise-ratio, where ZFB is used to eliminate the intra-group user interference. ZFB is often based on equal-power allocation to all UEs for computational tractability, which achieves the performance that is far from its capacity [10].

This work was supported in part by Institute for Computational Science and Technology, Hochiminh city, Vietnam, in part by the Australian Research Council's Discovery Projects under Project DP190102501, in part by the U.K. Royal Academy of Engineering Research Fellowship under Grant RF1415\14\22, and in part by U.S. National Science Foundation under Grants CCF-093970 and CCF-1513915

Long D. Nguyen was with Queen's University, Belfast BT7 1NN, UK. He is now with Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam (email: dinhlonghcmut@gmail.com)

Hoang D. Tuan is with the School of Electrical and Data Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia (email: Tuan.Hoang@uts.edu.au)

Trung Q. Duong is with Queen's University, Belfast BT7 1NN, UK (email: trung.q.duong@qub.ac.uk)

H. Vincent Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA (email: poor@princeton.edu)

When the number of transmit antennas is not much larger than the number of served UEs, the channel right inversion may not exist or is ill-posed, making ZFB either useless or poorly perform. A practical remedy for the channel matrix inversion is regularized zero-forcing beamforming (RZFB) [11]–[13], which uses a regularization parameter to make the channel matrix well-posed for the right inversion. Unlike ZFB, multi-user interference cannot be cancelled by RZFB and thus needs to be appropriately managed by optimizing the regularization parameter. So far, most of the existing works considered the problem of optimizing the regularization parameter based on the equal-power allocation to UEs, for which the regularization parameter is analytically found for the case of sum rate capacity under sum power constraint, providing that both number of transmit antennas and number of served UEs go to infinity with fixed ratio [11]. This value of the regularization parameter is used for analyzing the spectral efficiency of RZFB [14] or comparing the performance of RZFB with that of other schemes [15]. An important contribution is [16], which derived the deterministic equivalence of signal-to-interference-plus-noise (SINR) in the sense of almost sure probability when both the numbers of antennas and UEs go to infinity with a bounded ratio. The analytical formulas in [16, Corollaries 1 and 2] indeed rule out the opportunity of power allocation optimization as the individual rates are the same function of their allocated power. A large scale analysis has been made in [17], [18] under setting the regularization parameter to one. Ref [19] used a polynomial expansion for reducing the complexity of matrix inversion calculation, which works for SINR large scale analysis. An optimal ratio of the numbers of UEs and antennas for maximizing the asymptotic signal-to-leakage-plus-noise when both of these numbers go to infinity has been derived in [20]. Again, a heuristic method of low complexity for MIMO RZFB at high signal-to-noise ratio (SNR) was proposed in [21] with achievable sum rate analysis. There is no constraint for quality-of-service (QoS) for UEs in these works, making the multi-user services less meaningful.

Motivated by the above aforementioned existing issues in RZFB design, in this paper considers the joint design of power allocation and regularization parameter to maximize the worst UEs' throughput or the energy-efficiency subject to QoS constraints for UEs in terms of their throughput. Under a fixed regularization parameter, such problems for single-antenna UEs have been particularly considered in our previous work [22]. However, the joint optimization in the power allocation and regularization parameter causes much more computational challenges that make the computational method [22] no longer applicable, even for the case of single-antenna UEs considered there. Our contribution is thus two-fold:

- Development of computational procedures for RZFB for

single-antenna UEs, which constitute of two stages. As the optimization problems are very complicated, in the first stage path-following computational procedures are developed to solve their nonconvex relaxed problems, which are simpler than the original problem but still not convex. In the second stage, exact penalty optimization algorithms, which compensate the relaxation done in the first stage, are employed to locate a stationary point of the original problems.

- Development of two-stage computational procedures for RZFB for multi-antenna UEs, which are very different from that for single-antenna UEs in term of mathematical structures and thus require quite different mathematic tools for their computation.

The paper is organized as follows. After this Introduction, Section II is devoted to RZFB optimization for single-antenna UEs while Section III is devoted to RZFB optimization for multi-antenna UEs. Numerical examples to demonstrate the viability of the Algorithms developed in the previous section are provided in Section IV. Conclusions are drawn in Section V. The two appendices serve as the mathematic foundations for the developments in Section II and Section III. To the authors' best knowledge, these mathematical results are new.

Notation. Bold-faced lower-case and upper-case letters, e.g., $\mathbf{x}$ and $\mathbf{X}$, are respectively used for vectors and matrices, while lower-case letters, e.g., x , are used for scalars. The inner product between the vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^H \mathbf{y}$. Analogously, $\langle \mathbf{X}, \mathbf{Y} \rangle = \text{trace}(\mathbf{X}^H \mathbf{Y})$ for the matrices $\mathbf{X}$ and $\mathbf{Y}$. The shorthand notation $\mathbf{X}^2$ for the matrix $\mathbf{X}$ denotes the Hermitian symmetric positive definite matrix $\mathbf{X} \mathbf{X}^H$. $\mathbf{I}_N$ stands for the identity matrix of N dimensions. The notation $\mathbf{A} \succ 0$ means that $\mathbf{A}$ is a Hermitian symmetric strictly positive definite. $\text{diag}[a_i]_{i=1}^N$ ($\text{diag}[\mathbf{A}_i]_{i=1}^N$, resp.) is a diagonal (block-diagonal, resp.) matrix with scalars a_i ($i = 1, \dots, N$, (matrices $\mathbf{A}_i$, $i = 1, \dots, N$, resp.) on its diagonal (diagonal block, resp.). Also $\mathbb{R}_+^N \triangleq \{(x_1, \dots, x_N)^T \in \mathbb{R}^N : x_i > 0, i = 1, \dots, N\}$.

II. MISO REGULARIZED ZERO-FORCING BEAMFORMING

Consider a communication system with one base station (BS) serving N UEs. The BS is equipped with a large-scale M -antenna array while the UEs are equipped with a single antenna. The multi-input single-output (MISO) equation of such system is

$$\mathbf{y} = \text{diag}[\sqrt{\beta_n}]_{n=1}^N \mathbf{H}^H \tilde{\mathbf{s}} + \mathbf{w}, \quad (1)$$

where

$$\mathbf{H}^H = [h_{ij}^*]_{i=1, \dots, N; j=1, \dots, M} = \begin{bmatrix} \mathbf{h}_1^H \\ \vdots \\ \mathbf{h}_N^H \end{bmatrix} \in \mathbb{C}^{N \times M},$$

and $\mathbf{y} \in \mathbb{C}^N$ is the signal received at N UEs, $\tilde{\mathbf{s}} \in \mathbb{C}^M$ is the signal sent from BS and $\mathbf{w} \in \mathcal{CN}(0, \sigma^2 \mathbf{I}_N)$ is the background noise. The channel from the BS to user n is thus

$$\sqrt{\beta_n} \mathbf{h}_n, \mathbf{h}_n = [h_{n1} \dots h_{nM}]^T, \quad (2)$$

where $\sqrt{\beta_n}$ expresses the large-scale fading, while $\mathbf{h}_n = \mathbf{R}_n^{1/2} [\tilde{h}_{n1} \dots \tilde{h}_{nM}]^T$ with $h_{ij} \in \mathcal{CN}(0, 1)$ expressing the small

scale fading and $\mathbf{R}_n$ expressing the spatial correlation due to netted spacing of large numbers of antennas. Like the existing works on beamforming such as [11]–[13], [23]–[26], we assume the availability of the full channel state information and refer the reader to [27] and references therein for the mechanism of large-scale MIMO channel estimation.

In the above system, unlike the conventional massive MIMO, we do not assume that $N \ll M$, under which the right inversion for $\tilde{\mathbf{H}} = \text{diag}[\sqrt{\beta_n}]_{n=1}^N \mathbf{H}$ may not exist or be ill-conditioned; thus ZFB is not applicable. We are interested in the following class of RZFB, which admits a large-scale tractable computation

$$\tilde{\mathbf{s}} = \mathbf{F} \mathbf{s} = \mathbf{H}(\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_N)^{-1} \text{diag}[\sqrt{p_n}]_{n=1}^N \mathbf{s}, \quad (3)$$

where $\mathbf{s} = (s_1, \dots, s_N)^T$ is the information intended for the UEs, $\alpha > 0$ is the scalar to regularize $\mathbf{H}^H \mathbf{H}$, and $p_n, n = 1, \dots, N$ are the power control coefficients. Instead of designing the complex beamforming matrix $\mathbf{F}$ in (3) of the large size $M \times N$, which obviously leads to a large scale intractable computation (see e.g. [23], [28]), by (3) we aim to design the real N -dimensional vector $\mathbf{p} \triangleq (p_1, \dots, p_N)^T \in \mathbb{R}_+^N$ of power allocation for computational tractability. By (3) we propose a new RZFB compared to the existing RZFB defined by [11]–[13]

$$\mathbf{F} = \tilde{\mathbf{H}}(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}} + \alpha \mathbf{I}_N)^{-1} \text{diag}[\sqrt{p_n}]_{n=1}^N. \quad (4)$$

Then (1) is written by

$$\mathbf{y} = \text{diag}[\sqrt{\beta_n}]_{n=1}^N \mathbf{H}^H \mathbf{H}(\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_N)^{-1} \text{diag}[\sqrt{p_n}]_{n=1}^N \mathbf{s} + \mathbf{w} \quad (5)$$

$$= \text{diag}[\sqrt{\beta_n}]_{n=1}^N \mathbf{H}^H (\mathbf{H} \mathbf{H}^H + \alpha \mathbf{I}_M)^{-1} \mathbf{H} \text{diag}[\sqrt{p_n}]_{n=1}^N \mathbf{s} + \mathbf{w}, \quad (6)$$

where (6) is derived from (5) by using the following identity

$$\mathbf{H}(\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_N)^{-1} = (\mathbf{H} \mathbf{H}^H + \alpha \mathbf{I}_M)^{-1} \mathbf{H} \quad \forall \alpha > 0. \quad (7)$$

Make singular value decomposition (SVD)

$$\mathbf{H} \mathbf{H}^H = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^H \quad (8)$$

with a unitary $\mathbf{U}$ and diagonal $\mathbf{\Sigma} = \text{diag}[\lambda_m]_{m=1}^M$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M \geq 0$, and then define the matrix

$$\tilde{\mathbf{H}} = \mathbf{U}^H \mathbf{H}$$

of N columns $\tilde{\mathbf{h}}_n \in \mathbb{C}^M$, $n = 1, \dots, N$, to express (6) by

$$\mathbf{y} = \text{diag}[\sqrt{\beta_n}]_{n=1}^N \mathcal{H}_\alpha \text{diag}[\sqrt{p_n}]_{n=1}^N \mathbf{s} + \mathbf{w}, \quad (9)$$

where

$$\begin{aligned} \mathcal{H}_\alpha &\triangleq \tilde{\mathbf{H}}^H \mathcal{D}(\alpha) \tilde{\mathbf{H}} \\ &= [\tilde{\mathbf{h}}_n^H \mathcal{D}(\alpha) \tilde{\mathbf{h}}_j]_{n=1, \dots, N; j=1, \dots, N} \in \mathbb{C}^{N \times N}, \end{aligned}$$

which is a Hermitian symmetric matrix, and

$$\mathcal{D}(\alpha) \triangleq \text{diag}[(\lambda_m + \alpha)^{-1}]_{m=1}^M. \quad (10)$$

The throughput at user (UE) n is

$$r_n(p, \alpha) = \ln \left(1 + \frac{\beta_n p_n (\tilde{\mathbf{h}}_n^H \mathcal{D}(\alpha) \tilde{\mathbf{h}}_n)^2}{\beta_n \sum_{j \neq n} p_j |\tilde{\mathbf{h}}_j^H \mathcal{D}(\alpha) \tilde{\mathbf{h}}_j|^2 + \sigma^2} \right), \quad (11)$$

while by using identity (7), the transmit power at the BS is expressed by

$$\pi(p, \alpha) = \text{trace}(\mathbf{F}^H \mathbf{F}) = \sum_{n=1}^N p_n (\bar{\mathbf{h}}_n^H (\mathcal{D}(\alpha))^2 \bar{\mathbf{h}}_n) \quad (12)$$

as shown by Appendix III.

We consider the following max-min throughput optimization problem

$$\max_{\mathbf{p} \in \mathbb{R}_+^N, \alpha > 0} \min_{n=1, \dots, N} r_n(\mathbf{p}, \alpha) \quad (13a)$$

$$\text{s.t. } \pi(\mathbf{p}, \alpha) \leq P, \quad (13b)$$

where P is a given power threshold, i.e. we seek to maximize the worst UEs' throughput under the total transmit power constraint.

We also consider the following quality-of-service (QoS)-aware energy efficiency problem [29], [30]

$$\max_{\mathbf{p} \in \mathbb{R}_+^N, \alpha} \phi(\mathbf{p}, \alpha) \triangleq \frac{\sum_{n=1}^N r_n(\mathbf{p}, \alpha)}{\eta \pi(\mathbf{p}, \alpha) + P_{\text{non}}} \quad (14a)$$

$$\text{s.t. } (13b), \quad (14b)$$

$$r_n(\mathbf{p}, \alpha) \geq \bar{r}, n = 1, \dots, N, \quad (14c)$$

with a throughput threshold $\bar{r}$ to express the UEs' QoS. Here $\eta > 1$ is reciprocal of the drain efficiency of the amplifier of BS. $P_{\text{non}} = MP_a + P_c$ is the circuit non-transmit power consumed at BS where P_a and P_c are circuit power per antenna and non-transmission power. The denominator of the objective function in (14) represents the total power consumption in broadcasting signals to UEs.

One can see from the definition (11) and (12) that the throughput function $r_n(p, \alpha)$ is a complex nonconcave function, while the sum power function $\pi(p, \alpha)$ is nonconvex. Therefore, both the max-min throughput optimization problem (13) and the energy-efficiency maximization problem (14) are maximization of a nonconcave objective function over nonconvex constraints, which are very computationally difficult. Our next subsections are devoted to their computation.

A. Max-min throughput optimization

To address the max-min throughput optimization problem (13), make the following variable changes

$$x_n = 1/p_n > 0, n = 1, \dots, N \quad (15)$$

to express the rate function r_n and sum power function π as the following functions of $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}_+^N$ and $\alpha > 0$:

$$\begin{aligned} r_n(\mathbf{p}, \alpha) &\rightarrow \tilde{r}_n(\mathbf{x}, \alpha) \\ &\triangleq \ln \left(1 + \frac{(\bar{\mathbf{h}}_n^H \mathcal{D}(\alpha) \bar{\mathbf{h}}_n)^2}{x_n \left(\sum_{j \neq n} |\bar{\mathbf{h}}_n^H \mathcal{D}(\alpha) \bar{\mathbf{h}}_j|^2 / x_j + \sigma^2 / \beta_n \right)} \right) \\ &= \ln \left(1 + \frac{\left(\sum_{m=1}^M |\bar{h}_{n,m}|^2 / (\lambda_m + \alpha) \right)^2}{x_n \left(\sum_{j \neq n} \left| \frac{\sum_{m=1}^M \bar{h}_{n,m}^* \bar{h}_{j,m}}{(\lambda_m + \alpha)} \right|^2 / x_j + \frac{\sigma^2}{\beta_n} \right)} \right), \end{aligned}$$

and

$$\begin{aligned} \pi(\mathbf{p}, \alpha) &\rightarrow \tilde{\pi}(\mathbf{x}, \alpha) \triangleq \sum_{n=1}^N \frac{\bar{\mathbf{h}}_n (\mathcal{D}(\alpha))^2 \bar{\mathbf{h}}_n}{x_n} \\ &= \sum_{n=1}^N \sum_{m=1}^M \frac{|\bar{h}_{n,m}|^2}{(\lambda_m + \alpha)^2 x_n} \end{aligned}$$

which is seen as a convex function by checking its Hessian.

Then (13) is written by the following convex constrained optimization problem:

$$\max_{\mathbf{x} \in \mathbb{R}_+^N, \alpha > 0} \min_{n=1, \dots, N} \tilde{r}_n(\mathbf{x}, \alpha) \quad (16a)$$

$$\text{s.t. } \tilde{\pi}(\mathbf{x}, \alpha) \leq P. \quad (16b)$$

The difficulty of computing (16) is now concentrated at its objective function, which is still a very complex nonconcave function. To treat it, introduce the new slack variable $\mathbf{z} = (z_1, \dots, z_M)^T \in \mathbb{R}_+^M$ constrained by

$$z_m = \frac{1}{\lambda_m + \alpha}, m = 1, \dots, M, \quad (17)$$

which is equivalent to the nonconvex constraint

$$z_m \leq \frac{1}{\lambda_m + \alpha}, m = 1, \dots, M \quad (18)$$

plus the convex constraint

$$z_m(\lambda_m + \alpha) \geq 1, m = 1, \dots, M. \quad (19)$$

As the first stage, we treat the following nonconvex relaxed problem by dropping the nonconvex constraints (18):

$$\max_{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}_+^{N+M}, \alpha > 0} \Phi(\mathbf{x}, \alpha, \mathbf{z}) \triangleq \min_{n=1, \dots, N} g_n(\mathbf{x}, \alpha, \mathbf{z}) \quad (20a)$$

$$\text{s.t. } (19), (16b), (20b)$$

where

$$g_n(\mathbf{x}, \alpha, \mathbf{z}) \triangleq \ln \left(1 + \frac{\left(\sum_{m=1}^M |\bar{h}_{n,m}|^2 / (\lambda_m + \alpha) \right)^2}{x_n \left(\sum_{j \neq n} \frac{|\sum_{m=1}^M \bar{h}_{n,m}^* \bar{h}_{j,m} z_m|^2}{x_j} + \frac{\sigma^2}{\beta_n} \right)} \right). \quad (21)$$

Suppose $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ is a feasible point for (20) found from the $(\kappa - 1)$ th iteration. Use inequality (78) in the Appendix for

$$\begin{aligned} x &= x_n, y = 1 / \left(\sum_{m=1}^M |\bar{h}_{n,m}|^2 / (\lambda_m + \alpha) \right)^2, \\ t &= \left(\sum_{j \neq n} \left| \sum_{m=1}^M \bar{h}_{n,m}^* \bar{h}_{j,m} z_m \right|^2 / x_j + \sigma^2 / \beta_n \right), \end{aligned}$$

and

$$\begin{aligned} \bar{x} &= x_n^{(\kappa)}, \bar{y} = y_n^{(\kappa)} \triangleq 1 / \left(\sum_{m=1}^M |\bar{h}_{n,m}|^2 / (\lambda_m + \alpha^{(\kappa)}) \right)^2, \\ \bar{t} &= t_n^{(\kappa)} \triangleq \left(\sum_{j \neq n} \left| \sum_{m=1}^M \bar{h}_{n,m}^* \bar{h}_{j,m} z_m^{(\kappa)} \right|^2 / x_j^{(\kappa)} + \sigma^2 / \beta_n \right), \end{aligned}$$

to obtain

$$g_n(\mathbf{x}, \alpha, \mathbf{z}) \geq g_n^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) \quad (22)$$

over the trust region

$$\lambda_m + 2\alpha^{(\kappa)} - \alpha > 0, m = 1, \dots, M, \quad (23)$$

for

$$\begin{aligned} & g_n^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) = \\ & \tilde{a}_n^{(\kappa)} - \tilde{b}_n^{(\kappa)} x_n - \frac{\tilde{c}_n^{(\kappa)}}{\left(\sum_{m=1}^M |\bar{h}_{n,m}|^2 \ell_m^{(\kappa)}(\alpha)\right)^2} \\ & - \tilde{d}_n^{(\kappa)} \left(\sum_{j \neq n} \frac{\left| \sum_{m=1}^M \bar{h}_{n,m}^* \bar{h}_{j,m} z_m \right|^2}{x_j} + \frac{\sigma^2}{\beta_n} \right), \end{aligned} \quad (24)$$

which is a concave function with

$$\begin{aligned} 0 < \tilde{a}_n^{(\kappa)} & \triangleq \ln \left(1 + 1/x_n^{(\kappa)} y_n^{(\kappa)} t_n^{(\kappa)} \right) \\ & \quad + 3/(x_n^{(\kappa)} y_n^{(\kappa)} t_n^{(\kappa)} + 1), \\ 0 < \tilde{b}_n^{(\kappa)} & \triangleq 1/(x_n^{(\kappa)} y_n^{(\kappa)} t_n^{(\kappa)} + 1) x_n^{(\kappa)}, \\ 0 < \tilde{c}_n^{(\kappa)} & \triangleq 1/(x_n^{(\kappa)} y_n^{(\kappa)} t_n^{(\kappa)} + 1) y_n^{(\kappa)}, \\ 0 < \tilde{d}_n^{(\kappa)} & \triangleq 1/(x_n^{(\kappa)} y_n^{(\kappa)} t_n^{(\kappa)} + 1) t_n^{(\kappa)}, \end{aligned} \quad (25)$$

and

$$\ell_m^{(\kappa)}(\alpha) = (\lambda_m + 2\alpha^{(\kappa)} - \alpha)/(\lambda_m + \alpha^{(\kappa)})^2. \quad (26)$$

At the κ th iteration, the following convex program is solved to generate the next feasible point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$ for (20):

$$\begin{aligned} \max_{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}_+^{N+M}, \alpha > 0} \quad & \Phi^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) \triangleq \min_{n=1, \dots, N} g_n^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) \quad (27a) \\ \text{s.t.} \quad & (16b), (19), (23). \quad (27b) \end{aligned}$$

The computational complexity of (27) is

$$\mathcal{O}(\bar{n}^2 \bar{m}^{2.5} + \bar{m}^{3.5}), \quad (28)$$

with $\bar{n} = N + M + 1$, which the number of scalar variables, and $\bar{m} = 2 + 3M + N$, which is the number of constraints.

Note that

$$\Phi^{(\kappa)}(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)}) > \Phi^{(\kappa)}(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$$

because $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$ and $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ respective are the optimal solution and a feasible point for the convex optimization problem (27). This together with

$$\Phi(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)}) \geq \Phi^{(\kappa)}(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)}),$$

which follows from (22) yield

$$\Phi(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)}) > \Phi^{(\kappa)}(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)}).$$

But one can check immediately that $\Phi(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)}) = \Phi^{(\kappa)}(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ because

$$g_n(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)}) = g_n^{(\kappa)}(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)}),$$

so

$$\Phi(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)}) > \Phi(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)}), \quad (29)$$

showing that $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$ is a better feasible point for (20) than $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$. The objective $\Phi(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ thus converges. The sequence $\{(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})\}$ of improved feasible points for (20) can be proved converged at least to a stationary point $(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$ of (20) [31].

If $(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$ satisfies the nonconvex constraint (18) with some tolerance then it is also a stationary point of (16). Otherwise, we go to the second stage, where we follow [24], [32], [33] by introducing the penalty function

$$\theta_{\min}(\alpha, \mathbf{z}) = \min_{m=1, \dots, M} [1/(\lambda_m + \alpha) - z_m], \quad (30)$$

for which $\theta_{\min}(\bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) < 0$, and then the penalty factor

$$\mu \triangleq - \frac{\min_{n=1, \dots, N} g_n(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})}{\theta_{\min}(\bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})} \quad (31)$$

to consider the following penalized optimization problem

$$\begin{aligned} \max_{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}_+^{N+M}, \alpha > 0} \quad & \left[\min_{n=1, \dots, N} g_n(\mathbf{x}, \alpha, \mathbf{z}) + \mu \theta_{\min}(\alpha, \mathbf{z}) \right] \\ \text{s.t.} \quad & (16b), (19). \end{aligned} \quad (32)$$

Note that the convex constraint (19) implies

$$z_m \geq 1/(\lambda_m + \alpha), m = 1, \dots, M,$$

so $-\theta_{\min}(\alpha, \mathbf{z}) = \max_{m=1, \dots, M} (z_m - 1/(\lambda_m + \alpha)) \geq 0$ and $-\theta_{\min}(\alpha, \mathbf{z}) = 0$ if and only if $(\alpha, \mathbf{z})$ satisfies the nonconvex constraint (18). Therefore $-\theta_{\min}(\alpha, \mathbf{z})$ can be used to measure the degree of satisfaction of the nonconvex constraint (18). Instead of handling the nonconvex constraint (18) we incorporate this degree of its satisfaction into the objective function in (32), which is an exact penalty optimization formulation for (16), i.e. they share the same optimal solution with a finite value of μ [34, Chapter 16]. Initialized from $(\mathbf{x}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)}) = (\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$, at the κ th iteration, the following convex optimization problem is solved to generate the next feasible point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$ for (32):

$$\begin{aligned} \max_{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}_+^{N+M}, \alpha > 0} \quad & \left[\min_{n=1, \dots, N} g_n^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) + \mu \theta_{\min}^{(\kappa)}(\alpha, \mathbf{z}) \right] \quad (33a) \\ \text{s.t.} \quad & (16b), (19), (23), (33b) \end{aligned}$$

where

$$\theta_{\min}^{(\kappa)}(\alpha, \mathbf{z}) \triangleq \min_{m=1, \dots, M} [\ell_m^{(\kappa)}(\alpha) - z_m], \quad (34)$$

which a lower bounding concave approximation of $\theta_{\min}(\alpha, \mathbf{z})$. Its computational complexity is the same as that of (27). Analogously to (29), $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$ is seen as a better feasible point for (32) than $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$. The sequence $\{(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})\}$ at least converges to a stationary point $(\mathbf{x}_{opt}, \alpha_{opt}, \mathbf{z}_{opt})$ of (32) with

$$z_m^{opt} = 1/(\lambda_m + \alpha^{opt}), m = 1, \dots, M.$$

This computational procedure in the second stage terminates when¹

$$\min_{m=1, \dots, M} [1/(\lambda_m + \alpha^{(\kappa)}) - z_m^{(\kappa)}] \geq -\epsilon_{tot} \quad (35)$$

¹For instance $\epsilon_{tot} = 0.1$

so

$$z_m^{(\kappa)} = 1/(\lambda_m + \alpha^{(\kappa)})$$

for tolerance ϵ_{tot} . Algorithm 1 sums up our computational procedure.

Algorithm 1 Two-stage Max-min Optimization Algorithm

- 1: **Stage I:** Set $\kappa = 0$. Take a feasible point $(\mathbf{x}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)})$ for the convex constraints (16b) and (19).
 - 2: **Repeat until convergence:** Solve the convex optimization problem (27) to generate the next feasible point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$. Set $\kappa := \kappa + 1$.
 - 3: **Set** $(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) = (\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$. If it satisfies (18) declare that it is a stationary point of (16) and stop the Algorithm. Otherwise go to Stage II.
 - 4: **Stage II:** Set μ by (31). Reset $\kappa \rightarrow 0$, and $(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) \rightarrow (\mathbf{x}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)})$.
 - 5: **Repeat Until (35):** Solve the convex optimization problem (33) to generate the next feasible point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$; Set $\kappa := \kappa + 1$.
 - 6: **Output** $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ as the stationary point of (16).
-

B. EE maximization

Similarly, at the first stage, we address the EE maximization problem (14) via its following nonconvex relaxation

$$\max_{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}_+^{N+M}, \alpha > 0} \frac{\sum_{n=1}^N g_n(\mathbf{x}, \alpha, \mathbf{z})}{\Pi(\mathbf{x}, \alpha)} \quad (36a)$$

$$\text{s.t. } (16b), (19), g_n(\mathbf{x}, \alpha, \mathbf{z}) \geq \bar{r}, n = 1, \dots, N, \quad (36b)$$

where $g_n(\mathbf{x}, \alpha, \mathbf{z})$ is defined from (21) and $\Pi(\mathbf{x}, \alpha) = \eta\tilde{\pi}(\mathbf{x}, \alpha) + P_{non}$.

Suppose that $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ is a feasible point for (36) found from the $(\kappa - 1)$ th iteration and

$$\delta^{(\kappa)} \triangleq \frac{\sum_{n=1}^N g_n(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})}{\Pi(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)})}. \quad (37)$$

At the κ th iteration we solve the following convex optimization problem to generate the next iterative point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$:

$$\max_{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}_+^{N+M}, \alpha > 0} \sum_{n=1}^N g_n^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) - \delta^{(\kappa)} \Pi(\mathbf{x}, \alpha) \quad (38a)$$

$$\text{s.t. } (16b), (19), (23), g_n^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) \geq \bar{r}, n = 1, \dots, N, \quad (38b)$$

where functions $g_n^{(\kappa)}$ are defined from (24). Its computational complexity is (28) for $\bar{n} = M + N + 1$ and $\bar{m} = 2 + 3M + 2N$. Since $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ is a feasible point for (38), it follows that

$$\begin{aligned} \sum_{n=1}^N g_n^{(\kappa)}(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)}) - \delta^{(\kappa)} \Pi(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}) &> \\ \sum_{n=1}^N g_n^{(\kappa)}(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)}) - \delta^{(\kappa)} \Pi(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}) &= \\ 0. \end{aligned}$$

Hence

$$\frac{\sum_{n=1}^N g_n^{(\kappa)}(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})}{\Pi(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)})} > \delta^{(\kappa)},$$

so

$$\begin{aligned} \delta^{(\kappa+1)} &= \frac{\sum_{n=1}^N g_n(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})}{\Pi(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)})} \\ &\geq \frac{\sum_{n=1}^N g_n^{(\kappa)}(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})}{\Pi(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)})} \\ &> \delta^{(\kappa)}, \end{aligned} \quad (39)$$

i.e. the next iterative point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$ is better than $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ in optimizing (36). Suppose $(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$ is the stationary point of (36) found by this procedure. Again, if $(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$ satisfies the convex constraint (18) then it is also a stationary point for the EE maximization problem (14). Otherwise, we initialize the second stage with setting

$$\mu = - \sum_{n=1}^N g_n(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) / \theta_{\text{sum}}(\bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}), \quad (40)$$

where

$$\theta_{\text{sum}}(\alpha, \mathbf{z}) = \sum_{m=1}^M \left(\frac{1}{\lambda_m + \alpha} - z_m \right). \quad (41)$$

Initialized from $(\mathbf{x}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)}) = (\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$, at the κ th iteration we solve the following convex optimization problem to generate the next iterative point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$:

$$\begin{aligned} \max_{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}_+^{N+M}, \alpha > 0} \sum_{n=1}^N g_n^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) + \mu \theta_{\text{sum}}^{(\kappa)}(\alpha, \mathbf{z}) \\ - \delta^{(\kappa)} \Pi(\mathbf{x}, \alpha) \end{aligned} \quad (42a)$$

$$\text{s.t. } (16b), (19), (23), g_n^{(\kappa)}(\mathbf{x}, \alpha, \mathbf{z}) \geq \bar{r}, n = 1, \dots, N, \quad (42b)$$

where

$$\theta_{\text{sum}}^{(\kappa)} = \sum_{m=1}^M (\ell_m^{(\kappa)}(\alpha) - z_m), \quad (43)$$

with $\ell_m^{(\kappa)}$ defined from (26) and

$$\delta^{(\kappa)} \triangleq \frac{\sum_{n=1}^N g_n(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)}) + \mu \theta(\alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})}{\Pi(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)})}. \quad (44)$$

Its computational complexity is the same as that of (38). Algorithm 2 sums up our computational procedure.

III. MIMO REGULARIZED ZERO-FORCING

Now, consider the case when each UE is equipped with N_r antennas and the number of information streams for each use is N_r so each information $\mathbf{s}_i \in \mathbb{C}^{N_r}$, $E(\|\mathbf{s}_i\|^2) = 1$ for UE i is precoded by $\tilde{\mathbf{P}}_n \in \mathbb{C}^{M \times N_r}$.

Instead of MISO equation (1), we have the MIMO equation

$$\mathbf{y} = \text{diag}[\sqrt{\beta_n} \mathbf{I}_{N_r}]_{n=1}^N \mathbf{H}^H \tilde{\mathbf{s}} + \mathbf{w}, \quad (45)$$

Algorithm 2 Two-stage EE Maximization Algorithm

- 1: **State I:** Use Algorithm 1 to find a feasible point $(\mathbf{x}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)})$ for constraints (16b), (19) and (36b). Set $\kappa = 0$.
 - 2: **Repeat until convergence:** Solve the convex optimization problem (38) to generate the next iterative point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$; Set $\kappa := \kappa + 1$.
 - 3: **Set** $(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) = (\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$. If it satisfies (18) declare that it is a stationary point of (14) and stop the Algorithm. Otherwise go to the next stage
 - 4: **Stage II:** Set μ by (40). Reset $\kappa \rightarrow 0$, and $(\bar{\mathbf{x}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) \rightarrow (\mathbf{x}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)})$.
 - 5: **Repeat Until (35):** Solve the convex optimization problem (42) to generate the next feasible point $(\mathbf{x}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$; Set $\kappa := \kappa + 1$.
 - 6: **Output** $(\mathbf{x}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ as the stationary point of (14).
-

where

$$\mathbf{H}^H = [h_{ij}^*]_{i=1, \dots, NN_r; j=1, \dots, M} = \begin{bmatrix} \mathbf{H}_1^H \\ \vdots \\ \mathbf{H}_N^H \end{bmatrix} \in \mathbb{C}^{(NN_r) \times M},$$

$$\mathbf{H}_n^H \in \mathbb{C}^{N_r \times M}$$

and $\mathbf{y} = (\mathbf{y}_1^T, \dots, \mathbf{y}_N^T)^T \in \mathbb{C}^{NN_r}$, $\mathbf{y}_i \in \mathbb{C}^{N_r}$, is the signal received at N UEs, $\tilde{\mathbf{s}} = \sum_{i=1}^N \tilde{\mathbf{P}}_n \mathbf{s}_n \in \mathbb{C}^M$ is the signal sent from the BS, and $\mathbf{w} \in \mathcal{CN}(0, \sigma^2 \mathbf{I}_{NN_r})$ is the background noise. Compared to the channel (2), which is a vector, the channel from the BS to user n is a matrix

$$\sqrt{\beta_n} \mathbf{H}_n, \quad (46)$$

where $\sqrt{\beta_n}$ expresses the large-scale fading as in (2), while entries of $\mathbf{H}_n = \mathbf{R}_{T,n}^{1/2} \tilde{\mathbf{H}}_n \mathbf{R}_{R,n}^{1/2}$ with $h_{ij} \in \mathcal{CN}(0, 1)$ expressing the small scale fading, $\tilde{\mathbf{H}}_n = [\tilde{\mathbf{h}}_{11} \dots \tilde{\mathbf{h}}_{1M}; \dots; \tilde{\mathbf{h}}_{N_r,1} \dots \tilde{\mathbf{h}}_{N_r,M}]^T$ and $\mathbf{R}_{T,n}$ and $\mathbf{R}_{R,n}$ expressing the spatial correlation at BS and UE n , respectively.

We consider the following new class of RZFB

$$\tilde{\mathbf{P}} = \mathbf{H}(\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_{NN_r})^{-1} \text{diag}[\mathbf{P}_n]_{n=1}^N,$$

$$\mathbf{P}_n = \text{diag}[\sqrt{p_{n,i}}]_{i=1}^{N_r} \in \mathbb{C}^{N_r \times N_r}, \quad (47)$$

where $\alpha > 0$ is the scalar to regularize $\mathbf{H}^H \mathbf{H}$, i.e. instead of designing N beamforming matrices $\tilde{\mathbf{P}}_n$ of size $M \times N_r$ (see e.g. [23], [26]) we design N_r -dimensional vectors $\mathbf{p}_n \triangleq (p_{n,1}, \dots, p_{n,N_r})^T \in \mathbb{R}_+^{N_r}$, $n = 1, \dots, N$ of power allocation.

Then (45) is written by

$$\begin{aligned} \mathbf{y} &= \text{diag}[\sqrt{\beta_n} \mathbf{I}_{N_r}]_{n=1}^N \mathbf{H}^H \mathbf{H} (\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_{NN_r})^{-1} \\ &\quad \times \text{diag}[\mathbf{P}_n]_{n=1}^N \mathbf{s} + \mathbf{w} \\ &= \text{diag}[\sqrt{\beta_n} \mathbf{I}_{N_r}]_{n=1}^N \mathbf{H}^H (\mathbf{H} \mathbf{H}^H + \alpha \mathbf{I}_M)^{-1} \mathbf{H} \\ &\quad \times \text{diag}[\mathbf{P}_n]_{n=1}^N \mathbf{s} + \mathbf{w} \\ &= \text{diag}[\sqrt{\beta_n} \mathbf{I}_{N_r}]_{n=1}^N \mathcal{H}_\alpha \text{diag}[\mathbf{P}_n]_{n=1}^N \mathbf{s} + \mathbf{w}, \end{aligned} \quad (48)$$

where $\mathbf{s} = (\mathbf{s}_1, \dots, \mathbf{s}_N)^T$, and

$$\mathcal{H}_\alpha \triangleq [\mathbf{H}_n^H (\mathbf{H} \mathbf{H}^H + \alpha \mathbf{I}_M)^{-1} \mathbf{H}_j]_{n=1, \dots, N; j=1, \dots, N}$$

$$\in \mathbb{C}^{(NN_r) \times (NN_r)}$$

which is a Hermitian symmetric matrix. Each its block entry is

$$\mathbf{H}_n^H (\mathbf{H} \mathbf{H}^H + \alpha \mathbf{I}_M)^{-1} \mathbf{H}_j \in \mathbb{C}^{N_r \times N_r}.$$

Upon the SVD (8), define

$$\bar{\mathbf{H}}_n = \mathbf{U}^H \mathbf{H}_n, n = 1, \dots, N.$$

Then,

$$\mathcal{H}_\alpha \triangleq [\bar{\mathbf{H}}_n^H \mathcal{D}(\alpha) \bar{\mathbf{H}}_j]_{n=1, \dots, N; j=1, \dots, N} \in \mathbb{C}^{(NN_r) \times (NN_r)},$$

where $\mathcal{D}(\alpha)$ is defined from (10).

Then, for $\mathbf{P} \triangleq \text{diag}[\mathbf{P}_n]_{n=1}^N$, Appendix IV shows that the throughput at UE n is

$$\begin{aligned} r_n(\mathbf{P}, \alpha) &= \ln |\mathbf{P}_n^2| + 2 \ln |\bar{\mathbf{H}}_n^H \mathcal{D}(\alpha) \bar{\mathbf{H}}_n| \\ &\quad + \ln \left| (\bar{\mathbf{H}}_n^H \mathcal{D}(\alpha) \bar{\mathbf{H}}_n \mathbf{P}_n)^{-2} \right. \\ &\quad \left. + \left(\sum_{j \neq n} (\bar{\mathbf{H}}_n^H \mathcal{D}(\alpha) \bar{\mathbf{H}}_j \mathbf{P}_j)^2 + \frac{\sigma^2}{\beta_n} \mathbf{I}_{N_r} \right)^{-1} \right| \end{aligned} \quad (49)$$

while the power constraint at the BS is

$$\pi(\mathbf{P}, \alpha) = \sum_{n=1}^N \text{trace}(\bar{\mathbf{H}}_n^H (\mathcal{D}(\alpha))^2 \bar{\mathbf{H}}_n \mathbf{P}_n^2) \quad (50)$$

$$\leq P. \quad (51)$$

The max-min rate optimization is now formulated as

$$\max_{\mathbf{P} \succ 0, \alpha > 0} \min_{n=1, \dots, N} r_n(\mathbf{P}, \alpha) \quad (52a)$$

$$\text{s.t. (51)}, \quad (52b)$$

while the EE maximization problem is

$$\max_{\mathbf{P} \succ 0, \alpha > 0} \frac{\sum_{n=1}^N r_n(\mathbf{P}, \alpha)}{\eta \pi(\mathbf{P}, \alpha) + P_{non}} \quad (53a)$$

$$\text{s.t. (51)}, \quad (53b)$$

$$r_n(\mathbf{P}, \alpha) \geq \bar{r}, n = 1, \dots, N. \quad (53c)$$

Compared to their counterparts for single-antenna UEs defined by (11) and (12), the throughput and sum power functions for multi-antenna UEs defined by (49) and (50) are even more complex. Obviously, the tools for solving (14) and (13) in the previous section are hardly usable for solving (52) and (53). The next subsections provide a new development for their computation.

A. MIMO max-min throughput optimization

There are lots of ways to handle the optimization problem (52). Perhaps, a natural way is to introduce the variables

$$\mathbf{X}_n \triangleq \mathbf{P}_n^{-2} = \text{diag}[1/p_{n,i}]_{i=1, \dots, N_r},$$

$$\mathbf{X} \triangleq \text{diag}[\mathbf{X}_n]_{n=1, \dots, N}. \quad (54)$$

Then

$$\begin{aligned} r_n(\mathbf{P}, \alpha) &\rightarrow \tilde{r}_n(\mathbf{X}, \alpha) \\ &= \ln |\mathbf{X}_n^{-1}| + 2 \ln |\bar{\mathbf{H}}_n^H \mathcal{D}(\alpha) \bar{\mathbf{H}}_n| \\ &\quad + \ln \left| \Phi_{nn}^{-1}(\mathbf{X}_n, \alpha) + \left(\Psi_n(\mathbf{X}, \alpha) + \frac{\sigma^2}{\beta_n} \mathbf{I}_{N_r} \right)^{-1} \right| \end{aligned} \quad (55)$$

and

$$\begin{aligned}\pi(\mathbf{P}, \alpha) \rightarrow \tilde{\pi}(\mathbf{X}, \alpha) &= \sum_{n=1}^N \text{trace}((\bar{\mathbf{H}}_n^H(\mathcal{D}(\alpha))^2 \bar{\mathbf{H}}_n X_n^{-1})) \\ &= \sum_{n=1}^N \sum_{i=1}^{N_r} \sum_{m=1}^M \frac{|\bar{H}_n(m, i)|^2}{(\lambda_m + \alpha)^2 x_{ni}},\end{aligned}\quad (56)$$

which is a convex function for

$$\Phi_{nj}(\mathbf{X}_j, \alpha) \triangleq (\bar{\mathbf{H}}_n^H \mathcal{D}(\alpha) \mathbf{H}_j) \mathbf{X}_j^{-1} (\bar{\mathbf{H}}_n^H \mathcal{D}(\alpha) \mathbf{H}_j)^H, \quad n = 1, \dots, N; j = 1, \dots, N, \quad (57)$$

and

$$\Psi_n(\mathbf{X}, \alpha) \triangleq \sum_{j \neq n} \Phi_{nj}(\mathbf{X}_j, \alpha). \quad (58)$$

The max-min rate optimization (52) is now equivalent to

$$\max_{\mathbf{X} \succ 0, \alpha > 0} \min_{n=1, \dots, N} \tilde{r}_n(\mathbf{X}, \alpha) \quad (59a)$$

$$\text{s.t. } \tilde{\pi}(\mathbf{X}, \alpha) \leq P, \quad (59b)$$

where the sum power constraint (59b) is convex.

Again, by introducing the slack variable $\mathbf{z} = (z_1, \dots, z_m)^T \in \mathbb{R}_+^M$ satisfying the convex constraint (19) we consider the following nonconvex relaxation of (59)

$$\max_{\mathbf{X} \succ 0, \mathbf{z} \in \mathbb{R}_+^M, \alpha > 0} \min_{n=1, \dots, N} g_n(\mathbf{X}, \alpha, \mathbf{z}) \quad (60a)$$

$$\text{s.t. } (19), (59b), \quad (60b)$$

where

$$\begin{aligned}g_n(\mathbf{X}, \alpha, \mathbf{z}) &= \ln |\mathbf{X}_n^{-1}| + 2 \ln |\bar{\mathbf{H}}_n^H (\Sigma + \alpha \mathbf{I}_M)^{-1} \bar{\mathbf{H}}_n| \\ &\quad + \ln \left| \left[(\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_n) \mathbf{X}_n^{-1} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_n)^H \right]^{-1} \right. \\ &\quad \left. + \left(\sum_{j \neq n} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \mathbf{H}_j) \mathbf{X}_j^{-1} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \mathbf{H}_j)^H + \frac{\sigma^2}{\beta_n} \mathbf{I}_{N_r} \right) \right|^{-1}\end{aligned}\quad (61)$$

for $\Delta(\mathbf{z}) \triangleq \text{diag}[z_m]_{m=1}^M$.

Suppose that $(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ is a feasible point for (60) found from the $(\kappa - 1)$ th iteration.

We address the first term in the objective (61). By applying inequality (82) in the Appendix for $\mathbf{X}_1 = \mathbf{X}_n$ and $\bar{\mathbf{X}}_1 = \mathbf{X}_n^{(\kappa)}$, we obtain

$$\ln |\mathbf{X}_n^{-1}| \geq g_{1,n}^{(\kappa)}(\mathbf{X}) \quad (62)$$

for

$$\begin{aligned}g_{1,n}^{(\kappa)}(\mathbf{X}) &\triangleq \\ \log |(\mathbf{X}_n^{(\kappa)})^{-1}| + \text{trace}(\mathbf{I}) - \text{trace}((\mathbf{X}_n^{(\kappa)})^{-1} \mathbf{X}_n).\end{aligned}\quad (63)$$

We address the second term in the objective (61). By applying inequality (83) in the Appendix for $\alpha > 0$ and $\bar{\alpha} = \alpha^{(\kappa)} > 0$, we obtain

$$\ln |\bar{\mathbf{H}}_n^H (\Sigma + \alpha \mathbf{I}_M)^{-1} \bar{\mathbf{H}}_n| \geq g_{2,n}^{(\kappa)}(\alpha) \quad (64)$$

for

$$\begin{aligned}g_{2,n}^{(\kappa)}(\alpha) &= \ln |\bar{\mathbf{H}}_n^H (\Sigma + \alpha^{(\kappa)} \mathbf{I}_M)^{-1} \bar{\mathbf{H}}_n| \\ &\quad - \text{trace} \left([\bar{\mathbf{H}}_n^H (\Sigma + \alpha^{(\kappa)} \mathbf{I}_M)^{-1} \bar{\mathbf{H}}_n]^{-1} \times \right. \\ &\quad \left. \bar{\mathbf{H}}_n^H (\Sigma + \alpha^{(\kappa)} \mathbf{I}_M)^{-2} \bar{\mathbf{H}}_n \right) (\alpha - \alpha^{(\kappa)}).\end{aligned}\quad (65)$$

We address the third term in the objective (61). By applying inequality (81) in the Appendix for

$$\mathbf{X}_1 = (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_n) \mathbf{X}_n^{-1} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_n)^H,$$

$$\mathbf{X}_2 = \sum_{j \neq n} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_j) \mathbf{X}_j^{-1} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_j)^H + (\sigma^2 / \beta_n) \mathbf{I}_{N_r},$$

and

$$\begin{aligned}\bar{\mathbf{X}}_1 &= \mathbf{X}_1^{(\kappa)} \\ &\triangleq (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}^{(\kappa)}) \bar{\mathbf{H}}_n) (\mathbf{X}_n^{(\kappa)})^{-1} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}^{(\kappa)}) \bar{\mathbf{H}}_n)^H,\end{aligned}$$

$$\begin{aligned}\bar{\mathbf{X}}_2 &= \mathbf{X}_2^{(\kappa)} \\ &\triangleq \sum_{j \neq n} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}^{(\kappa)}) \bar{\mathbf{H}}_j) (\mathbf{X}_j^{(\kappa)})^{-1} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}^{(\kappa)}) \bar{\mathbf{H}}_j)^H \\ &\quad + (\sigma^2 / \beta_n) \mathbf{I}_{N_r},\end{aligned}$$

we obtain

$$\ln |\mathbf{X}_1^{-1} + \mathbf{X}_2^{-1}| \geq g_{3,n}^{(\kappa)}(\mathbf{X}, \mathbf{z}) \quad (66)$$

for

$$\begin{aligned}g_{3,n}^{(\kappa)}(\mathbf{X}, \mathbf{z}) &= \\ \log |(\mathbf{X}_1^{(\kappa)})^{-1} + (\mathbf{X}_2^{(\kappa)})^{-1}| + \text{trace}(\mathbf{I}) \\ &\quad + \text{trace}([(X_1^{(\kappa)} + X_2^{(\kappa)})^{-1} - (X_1^{(\kappa)})^{-1}] \mathbf{X}_1) \\ &\quad + \text{trace}([(X_1^{(\kappa)} + X_2^{(\kappa)})^{-1} - (X_2^{(\kappa)})^{-1}] \mathbf{X}_2).\end{aligned}\quad (67)$$

For $\mathcal{A} \succeq 0$, it can be derived that

$$\begin{aligned}\text{trace}(\mathcal{A}(\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_j) \mathbf{X}_j^{-1} (\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_j)^H) &= \\ \text{trace}((\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_j)^H \mathcal{A}(\bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_j) \mathbf{X}_j^{-1}) &= \\ \text{trace}(\bar{\mathbf{H}}_j^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_n \mathcal{A} \bar{\mathbf{H}}_n^H \Delta(\mathbf{z}) \bar{\mathbf{H}}_j \text{diag}[1/p_{ji}]_{i=1}^{N_r}) &= \\ \sum_{i=1}^{N_r} (\bar{\mathbf{H}}_{ji}^H \Delta(\mathbf{z}) (\bar{\mathbf{H}}_n \mathcal{A} \bar{\mathbf{H}}_n^H) \Delta(\mathbf{z}) \bar{\mathbf{H}}_{ji}) / p_{ji} &= \\ \sum_{i=1}^{N_r} \|(\bar{\mathbf{H}}_n \mathcal{A} \bar{\mathbf{H}}_n^H)^{1/2} \Delta(\mathbf{z}) \bar{\mathbf{H}}_{ji}\|^2 / p_{ji},\end{aligned}$$

where $\bar{\mathbf{H}}_{ji} = \bar{\mathbf{H}}_j(:, i) \in \mathbb{C}^M$ is the i th column of $\bar{\mathbf{H}}_j$. Therefore, for $0 \preceq \mathcal{A}_i \triangleq (\mathbf{X}_i^{(\kappa)})^{-1} - (\mathbf{X}_1^{(\kappa)} + \mathbf{X}_2^{(\kappa)})^{-1}$, $i = 1, 2$ one has

$$\begin{aligned}g_{3,n}^{(\kappa)}(\mathbf{X}, \mathbf{z}) &= \log |(\mathbf{X}_1^{(\kappa)})^{-1} + (\mathbf{X}_2^{(\kappa)})^{-1}| \\ &\quad + \text{trace}(\mathbf{I}) - \text{trace}(\mathcal{A}_2) \sigma^2 / \beta_n \\ &\quad - \sum_{i=1}^{N_r} \|(\bar{\mathbf{H}}_n \mathcal{A}_1^{(\kappa)} \bar{\mathbf{H}}_n^H)^{1/2} \Delta(\mathbf{z}) \bar{\mathbf{H}}_{ni}\|^2 / p_{ni} \\ &\quad - \sum_{j \neq n} \sum_{i=1}^{N_r} \|(\bar{\mathbf{H}}_n \mathcal{A}_2^{(\kappa)} \bar{\mathbf{H}}_n^H)^{1/2} \Delta(\mathbf{z}) \bar{\mathbf{H}}_{ji}\|^2 / p_{ji},\end{aligned}$$

which is a concave function.

At the κ th iteration, the following convex program is solved to generate the next feasible point $(\mathbf{X}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$ for (59):

$$\max_{\mathbf{X} \succ 0, \mathbf{z} \in \mathbb{R}_+^M, \alpha} \min_{n=1, \dots, N} g_n^{(\kappa)}(\mathbf{X}, \alpha, \mathbf{z}) \text{ s.t. (19), (59b).} \quad (68)$$

where $g_n^{(\kappa)}(\mathbf{X}, \alpha, \mathbf{z}) = g_{1,n}^{(\kappa)}(\mathbf{X}) + 2g_{2,n}^{(\kappa)}(\alpha) + g_{3,n}^{(\kappa)}(\mathbf{X}, \mathbf{z})$. Its computational complexity is (28) for $\bar{n} = NN_r + M + 1$ and $\bar{m} = NN_r + 2M + 2$.

Similar to the first stage of Algorithm 1, we can show that the sequence $\{(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})\}$ generated by a such procedure will at least converge to a stationary point $(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$ of the nonconvex optimization problem (60). If $(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$ satisfies the nonconvex constraint (18) with some tolerance then it is also a stationary point of (59). Otherwise, recalling the definition (30) of the penalty function $\theta_{\min}(\alpha, \mathbf{z})$, we define the penalty factor

$$\mu \triangleq -\frac{\min_{n=1, \dots, N} g_n(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})}{\theta_{\min}(\bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})} \quad (69)$$

to consider the following penalized optimization problem

$$\max_{\mathbf{X} \succ 0, \mathbf{z} \in \mathbb{R}_+^M, \alpha > 0} \left[\min_{n=1, \dots, N} g_n(\mathbf{X}, \alpha, \mathbf{z}) + \mu \theta_{\min}(\alpha, \mathbf{z}) \right] \quad (70a)$$

s.t. (19), (59b). (70b)

Initialized from $(\mathbf{X}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)}) = (\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$, at the κ th iteration, the following convex program is solved to generate the next feasible point $(\mathbf{X}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$ for (70):

$$\max_{\mathbf{X} \succ 0, \mathbf{z} \in \mathbb{R}_+^M, \alpha > 0} \left[\min_{n=1, \dots, N} g_n^{(\kappa)}(\mathbf{X}, \alpha, \mathbf{z}) + \mu \theta_{\min}^{(\kappa)}(\alpha, \mathbf{z}) \right] \quad (71a)$$

s.t. (19), (59b), (71b)

where $\theta_{\min}^{(\kappa)}(\alpha, \mathbf{z})$ is defined from (34). Its computational complexity is the same as that of (59). The procedure terminates once (35) is verified.

Algorithm 3 sums up our computational procedure.

Algorithm 3 Two-stage MIMO Max-min Optimization Algorithm

- 1: **Stage I:** Set $\kappa = 0$. Take a feasible point $(\mathbf{X}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)})$ for the convex constraints (19) and (59b),.
 - 2: **Repeat until convergence** Solve the convex optimization problem (68) to generate the next feasible point $(\mathbf{X}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$; Set $\kappa := \kappa + 1$.
 - 3: **Set** $(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) = (\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$. If it satisfies (18) declare that it is a stationary point of (59) and stop the Algorithm. Otherwise go to the next stage.
 - 4: **Stage II:** Set μ by (69). Reset $\kappa \rightarrow 0$, and $(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) \rightarrow (\mathbf{X}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)})$.
 - 5: **Repeat until (35):**
 - 6: Solve the convex optimization problem (71) to generate the next feasible point $(\mathbf{X}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$;
 - 7: Set $\kappa := \kappa + 1$.
 - 8: **Output** $(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ as a stationary point of (59).
-

B. EE maximization

At the first stage we address the EE maximization problem (53) via its following nonconvex relaxation

$$\max_{\mathbf{X} \succ 0, \mathbf{z} \in \mathbb{R}_+^M, \alpha > 0} \frac{\sum_{n=1}^N g_n(\mathbf{X}, \alpha, \mathbf{z})}{\Pi(\mathbf{X}, \alpha)} \quad (72a)$$

$$\text{s.t. (19), (59b), } g_n(\mathbf{X}, \alpha, \mathbf{z}) \geq \bar{r}, n = 1, \dots, N, \quad (72b)$$

where $g_n(\mathbf{X}, \alpha, \mathbf{z})$ is defined from (61) and $\Pi(\mathbf{X}, \alpha) = \eta \tilde{\pi}(\mathbf{X}, \alpha) + P_{non}$.

Suppose that $(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ is a feasible point for (72) found from the $(\kappa - 1)$ th iteration and

$$\delta^{(\kappa)} \triangleq \frac{\sum_{n=1}^N g_n(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})}{\Pi(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)})}. \quad (73)$$

At the κ th iteration we solve the following convex optimization problem to generate the next iterative point $(\mathbf{X}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$:

$$\max_{\mathbf{X} \succ 0, \mathbf{z} \in \mathbb{R}_+^M, \alpha > 0} \sum_{n=1}^N g_n^{(\kappa)}(\mathbf{X}, \alpha, \mathbf{z}) - \delta^{(\kappa)} \Pi(\mathbf{X}, \alpha) \quad (74a)$$

$$\text{s.t. (19), (59b), } g_n^{(\kappa)}(\mathbf{X}, \alpha, \mathbf{z}) \geq \bar{r}, n = 1, \dots, N. \quad (74b)$$

Its computational complexity is (28) for $\bar{n} = NN_r + M + 1$ and $\bar{m} = 2 + N(N_r + 1) + 2M$. Similar to (39), we can show that $\delta^{(\kappa+1)} > \delta^{(\kappa)}$ and so the generated sequence $\{(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})\}$ at least converges to a stationary point $(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$ of (72). Again, if $(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$ satisfies the convex constraint (18) then it is also a stationary point of the EE maximization problem (53). Otherwise, we need to process one more stage to compute a stationary point of (53). To this end, with θ_{sum} defined from (41), set

$$\mu = -\sum_{n=1}^N g_n(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) / \theta_{\text{sum}}(\bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}). \quad (75)$$

Initialized from $(\mathbf{X}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)}) = (\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt})$, at the κ th iteration we solve the following convex optimization problem to generate the next iterative point $(\mathbf{X}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$:

$$\max_{\mathbf{X} \succ 0, \mathbf{z} \in \mathbb{R}_+^M, \alpha > 0} \sum_{n=1}^N g_n^{(\kappa)}(\mathbf{X}, \alpha, \mathbf{z}) + \mu \theta_{\text{sum}}^{(\kappa)}(\alpha, \mathbf{z}) - \delta^{(\kappa)} \Pi(\mathbf{X}, \alpha) \quad (76a)$$

$$\text{s.t. (19), (59b), } g_n^{(\kappa)}(\mathbf{X}, \alpha, \mathbf{z}) \geq \bar{r}, n = 1, \dots, N, \quad (76b)$$

where $\theta_{\text{sum}}^{(\kappa)}$ is defined from (43) and

$$\delta^{(\kappa)} \triangleq \frac{\sum_{n=1}^N g_n(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)}) + \mu \theta_{\text{sum}}^{(\kappa)}(\alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})}{\Pi(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)})}. \quad (77)$$

Its computational complexity is the same as that of (74). Algorithm 4 sums up our computational procedure.

Algorithm 4 MIMO EE Maximization Algorithm

- 1: **Initialization:** Use Algorithm 3 to find a feasible point $(\mathbf{X}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)})$ for constraints (19), (59b), and $g_n(\mathbf{X}, \alpha, \mathbf{z}) \geq \bar{r}_n$, $n = 1, \dots, N$. Set $\kappa = 0$.
 - 2: **Repeat until convergence:** Solve the convex optimization problem (74) to generate the next iterative point $(\mathbf{X}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$. Set $\kappa := \kappa + 1$.
 - 3: **Set** $(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) = (\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$. If it satisfies (18) declare that it is a stationary point of (53) and stop the Algorithm. Otherwise set μ by (75). Reset $\kappa \rightarrow 0$, and $(\bar{\mathbf{X}}^{opt}, \bar{\alpha}^{opt}, \bar{\mathbf{z}}^{opt}) \rightarrow (\mathbf{X}^{(0)}, \alpha^{(0)}, \mathbf{z}^{(0)})$.
 - 4: **Repeat until 35):** Solve the convex optimization problem (76) to generate the next feasible point $(\mathbf{X}^{(\kappa+1)}, \alpha^{(\kappa+1)}, \mathbf{z}^{(\kappa+1)})$. Set $\kappa := \kappa + 1$.
 - 5: **Output** $(\mathbf{X}^{(\kappa)}, \alpha^{(\kappa)}, \mathbf{z}^{(\kappa)})$ as a stationary point of (53).
-

IV. NUMERICAL SIMULATIONS

In this section, we evaluate the performance of the proposed algorithms by numerical examples. The BS is located at the centre of a hexagon cell with radius 1 km. Unless otherwise stated, it is assumed that half of number of UEs called by far UEs are equally distributed at farther distances with another half of UEs called by nearer UEs equally distributed at near distances to the BS.

We adopt the standard exponential correlation model [16], where the medium spatial correlation between antenna (p, q) and antenna (m, n) , which constitute an entry of matrix $\mathbf{R}$ in (2) is expressed by

$$[\mathbf{R}]_{(p,q),(m,n)} = (0.5)^{|p-m|+|q-n|}$$

Table I provides other simulation parameters for generating large-scale fading channel and power consumption, which are similar to those used in [35]. The throughput threshold for all UEs is $\bar{r} \in \{0.4, 0.8\}$ bps/Hz as in [36, Table I]. In the simulations, Opt. RZF is referred to the RZFB found by the proposed Algorithms 1-4, while Sub. RZF is referred to the RZFB found by the proposed Algorithms 1-4 for the regularization parameter α fixed at $N\sigma^2/P$ as proposed in [11], [37]. Also, Equip-power RZF is referred to the RZFB under equal-power allocation, while Conv. RZF is referred to the conventional RZFB defined from (4), which is computed by adjusting Algorithm 1-4. Each point of the numerical results is the average of 1,000 random channel realizations.

TABLE I: Large scale fading Setup

Parameter	Numerical value
Carrier frequency / Bandwidth	2GHz / 10MHz
BS transmission power	46 dBm
Path loss from BS to UE	$128.1 + 37.6 \log_{10} R$ [dB], R in km
Shadowing standard deviation	8 dB
Noise power density	-174 dBm/Hz
Noise figure	9 dB
Drain efficiency of amplifier	$\alpha = 1/0.388$
Circuit power per antenna	$P_A = 189$ mW
Non-transmission power	$P_C = 40$ dBm

A. Single-antenna UEs

In this scenario, the BS is equipped with an 4×4 (4 rows in the horizontal dimension and 4 columns in the vertical dimension) or 8×8 (8 rows in the horizontal dimension and 8 columns in the vertical dimension) uniform planar array (UPA) of antennas. Thus, the total number of antennas at BS is $M = 16$ or $M = 64$, respectively.

1) *Max-min throughput performance:* Table II provides the average numbers of iterations of Algorithm 1 for computing Opt. RZF and Sub. RZF under different numbers of UEs to plot Fig. 1a. Algorithm 1 needs a few iterations, where the second stage for penalized optimization is either not necessary or takes only one iteration. Fig. 1a and Fig. 1b plot the achievable UEs'

TABLE II: The average number of iterations for $M = 16$.

# of UEs (N)	8	12	16	20	24	28	32
Sub. RZF	4.5	4.6	4.7	4.8	5.0	5.0	5.1
Opt. RZF	5.9	6.9	5.8	5.0	4.5	4.6	4.6

max-min throughput by different RZFB schemes versus the number of UEs for $M = 16$ and $M = 64$, respectively. In both figures, Opt. RZF obviously outperforms Sub. RZF, showing the importance of optimizing the regularization parameter α . With the number of BS antennas increased to $M = 64$, the number of served UEs also increases. For example, 16-antenna array BS can offer the QoS of 0.8 bps/Hz to 12 and 16 UEs according to Fig. 1a but 64-antenna array BS offer the same QoS to 54 and 70 UEs according to Fig. 1b. As expected, Equip-power RZF achieves the worst UEs' max-min throughput. Conv. RZF performs much worse than Sub. RZF and Opt. RZF, which means that regularizing the channel $\mathbf{H}^H \mathbf{H}$ of small-scale fading in (3) is a much more efficient way than regularizing the channel $\hat{\mathbf{H}}^H \hat{\mathbf{H}}$ incorporating both large-scale and small-scale fadings in (4).

(a) $M = 16$ (b) $M = 64$

Fig. 1: The achievable users' max-min throughput for $M = \{16, 64\}$.

2) *EE performance*: Fig. 2a plots the achievable EE by Opt. RZF and Sub. RZF at QoS threshold $\bar{r} = 0.4$ bps/Hz. As expected, Opt. RZF achieves much better EE compared with Sub. RZF. The gap is very wide for the number of UEs exceeding 12.

TABLE III: The average iterations of Algorithm 2 to plot Fig. 2b for $M = 16$ and $\bar{r} = 0.8$ bps/Hz.

# of UEs (N)	8	10	12	14	16	18
Sub. RZF	9.7	8.9	8.6	18.8	17	15.4
Opt. RZF	16.4	17.4	17.6	18.8	17	15.4

Table III shows that the average numbers of iterations of Algorithm 2 for computing Opt. RZF are higher compared to that for computing Sub. RZF. Fig. 2b reveals that Opt. RZF achieves a much better EE than Sub. RZF for a higher QoS threshold $\bar{r} = 0.8$ bps/Hz. Especially, Opt. RZF guarantees this QoS for a higher number of UEs than Sub. RZF does. Opt. RZF is able to serve up to 18 UEs but Sub. RZF cannot serve more than 12 UEs. Fig. 3 justifies this intuition. Sub. RZF consumes almost all the allowable transmit power to serve 12 UEs, while Opt. RZF controls the power much better in serving more UEs.

(a) $\bar{r} = 0.4$ bps/Hz (b) $\bar{r} = 0.8$ bps/Hz

Fig. 2: EE performance for $M = 16$ and $\bar{r} = \{0.4, 0.8\}$ bps/Hz.

Fig. 3: Transmit power for $M = 16$ and $\bar{r} = 0.8$ bps/Hz.

TABLE IV: The average iterations of Algorithm 3 to plot Fig. 4 for $M = 64$ and $\bar{r} = 0.8$ bps/Hz.

# of UEs (N)	30	36	42	48	54	60	66
Sub. RZF	10	10	10.5	10.5	22	22	19.5
Opt. RZF	20	20.5	21	21.5	22	22	19.5

(a) EE performance

(b) Transmit power

Fig. 4: EE performance and transmit power for $M = 64$ and $\bar{r} = 0.8$ bps/Hz.

For a higher number of BS antennas such as $M = 64$, Fig. 4a demonstrates the superior performance of Opt. RZF compared with Sub. RZF. The average iterations of Algorithm 2 for their computation are given in Table IV. As expected, Opt. RZF serves a much higher number of UEs (up to 66 UEs) than Sub. RZF does at the QoS threshold of $\bar{r} = 0.8$ bps/Hz. This can be explained by the result of transmit power as shown in Fig. 4b, where Opt. RZF requires much less transmission power to optimize EE than Sub. RZF.

B. Multi-antenna users

In this scenario, the BS is equipped with an 4×4 uniform planar array (UPA) of antennas corresponding to $M = 16$ and each UE is equipped with 2 antennas. To see the trade-off between the computational complexity and performance, we also include the simulation for unstructured MIMO beamforming, which uses the algorithm of [25], [26] for the max-min throughput optimization and the algorithm of [38], [39] for the EE maximization. The computational complexity of each iteration is $\mathcal{O}(\bar{n}^2 \bar{m}^{2.5} + \bar{m}^{3.5})$, where $\bar{n} = MN N_r + 1$ and $\bar{m} = N + 1$ for max-min throughput optimization and $\bar{n} = N(M^2 + N_r^2)$ and $\bar{m} = 2N + M^2 N + 1$ for the EE optimization, which is much more complex than that of each

iteration of Algorithms 3 and 4.

Table V provides the average numbers of iterations for plotting Fig. 5. Although Opt. RZF requires a larger number of iterations than Sub. RZF for convergence, the worst UEs' throughput of the former is much higher compared to the latter as shown in Fig. 5. The performance of the unstructured beamforming is slightly higher than that by Opt. RZF.

Fig. 5 plots the achievable UEs' max-min throughput by different RZFB schemes vs the number of UEs. Opt. RZF is seen achieving the UEs' max-min throughput much better than Sub. RZF, especially for $N < 20$. Like Fig. 1a and Fig. 1b, the achievable UE's max-min throughput by Equip-power RZF is the worst and Conv. RZF performs much worse than Sub. RZF and Opt. RZF.

TABLE V: The average iterations of Algorithm 3 and Unstruct. Max-Min for plotting Fig. 5.

# of UEs (N)	8	12	16	20	24	28	32
Sub. RZF	25.3	36.5	45.2	38.6	36	21.8	15.3
Opt. RZF	38.8	47.3	48.8	45.4	45.9	35.4	26.3
Unstruct. Max-Min	25	26	35	41	39	35	22

Fig. 5: The achievable UEs' max-min throughput.

In terms of EE performance, Fig. 6a and Fig. 6b also show the superior EE of Opt. RZF over Sub. RZF particularly because according to Fig. 7. Opt. RZF requires a much lower transmit power than Sub RZF in guaranteeing QoS threshold $\bar{r} = 0.8$ bps/Hz. The average iterations of Algorithm 4 for their computation are provided in Table VI.

(a) $\bar{r} = 0.4$ bps/Hz (b) $\bar{r} = 0.8$ bps/HzFig. 6: EE performance for $\bar{r} = \{0.4, 0.8\}$ bps/Hz. Fig. 7: Transmit power for $\bar{r} = 0.8$ bps/Hz.TABLE VI: The average iterations of Algorithm 4 and Unstruct. GEE for plotting Fig. 6b for $\bar{r} = 0.8$ bps/Hz.

# of UEs (N)	4	6	8	10	12
Sub. RZF	21.3	26.7	33	36	
Opt. RZF	34	44.7	43.7	49.3	51
Unstruct. EE	8	11	12	18	18

V. CONCLUSIONS

In this paper, we have considered the joint design of power allocation and regularization parameter to maximize the UEs' minimal throughput or the energy-efficiency in guaranteeing the UEs' QoS in term of their throughput requirement for a new class of RZFB. We have developed the two-stage algorithms to solve the posed problems and provided numerical examples to support their efficiency. We have also compared the performance achieved by the propose joint design with that achieved by the existing RZFB schemes to show the importance of the joint optimization in power allocation and regularization parameter. An extension of the joint design to multi-cell RZFB is under our current study.

APPENDIX I: FUNDAMENTAL INEQUALITIES

Theorem 1: Function $f(x, y, t) = \ln(1 + 1/xyt)$ is convex in $(x, y, t)^T \in \mathbb{R}_+^3$.

Proof: Function $\ln(1 + 1/t^3)$ is convex in $t > 0$ as it is immediate to see that $d^2 \ln(1 + 1/t^3)/dt^2 > 0$, while function $g(x, y, z) = (xyt)^{1/3}$ is concave in $(x, y, t)^T \in \mathbb{R}_+^3$ [40]. Then for $\alpha \geq 0, \beta \geq 0$ with $\alpha + \beta = 1$, it is true that

$$\begin{aligned}
 f(\alpha(x, y, t) + \beta(x_1, y_1, t_1)) &= \\
 \ln(1 + 1/(g(\alpha(x, y, t) + \beta(x_1, y_1, t_1)))^3) &\leq \\
 \ln(1 + 1/(\alpha g(x, y, t) + \beta g(x_1, y_1, t_1))^3) &\leq \\
 \alpha \ln(1 + 1/(g(x, y, t))^3) + \beta \ln(1 + 1/(g(x_1, y_1, t_1))^3) &= \\
 \alpha f(x, y, t) + \beta f(x_1, y_1, t_1) & \\
 \forall (x, y, t)^T \in \mathbb{R}_+^3, (x_1, y_1, t_1)^T \in \mathbb{R}_+^3, &
 \end{aligned}$$

showing the convexity of $f(x, y, t)$. $\square$

The following linear approximation then holds true for all $x > 0, \bar{x} > 0, y > 0, \bar{y} > 0, t > 0, \bar{t} > 0$ [40]:

$$\begin{aligned}
 \ln(1 + 1/xyt) &\geq f(\bar{x}, \bar{y}, \bar{t}) + \langle \nabla f(\bar{x}, \bar{y}, \bar{t}), (x, y, t) - (\bar{x}, \bar{y}, \bar{t}) \rangle \\
 &= \tilde{a} - \tilde{b}x - \tilde{c}y - \tilde{d}t,
 \end{aligned} \tag{78}$$

where ∇ is the gradient operation and

$$\begin{aligned}
 \tilde{a} &= \ln(1 + 1/\bar{x}\bar{y}\bar{t}) + \frac{3}{(\bar{x}\bar{y}\bar{t} + 1)} > 0, \tilde{b} = \frac{1}{(\bar{x}\bar{y}\bar{t} + 1)\bar{x}} > 0, \\
 \tilde{c} &= \frac{1}{(\bar{x}\bar{y}\bar{t} + 1)\bar{y}} > 0, \tilde{d} = \frac{1}{(\bar{x}\bar{y}\bar{t} + 1)\bar{t}} > 0.
 \end{aligned}$$

Furthermore, as function x^2/t is convex on x and $t > 0$, the following inequality holds true

$$\frac{x^2}{t} \geq \frac{2\bar{x}}{\bar{t}}x - \frac{\bar{x}^2}{\bar{t}^2}t. \tag{79}$$

Finally, noting that function $f(\alpha, t) = \mathbf{h}^H(\mathbf{\Sigma} + \alpha\mathbf{I}_M)^{-2}\mathbf{h}/t$ is convex in $\alpha > 0$ and $t > 0$, we have the following inequality

$$\begin{aligned}
 \frac{\mathbf{h}^H(\mathbf{\Sigma} + \alpha\mathbf{I}_M)^{-2}\mathbf{h}}{t} &\geq \\
 \frac{\mathbf{h}^H(\mathbf{\Sigma} + \bar{\alpha}\mathbf{I}_M)^{-2}\mathbf{h}}{\bar{t}} + \langle \nabla f(\bar{\alpha}, \bar{t}), (\alpha, t) - (\bar{\alpha}, \bar{t}) \rangle &= \\
 4 \frac{\mathbf{h}^H(\mathbf{\Sigma} + \bar{\alpha}\mathbf{I}_M)^{-2}\mathbf{h}}{\bar{t}} & \\
 - 2 \frac{\mathbf{h}^H(\mathbf{\Sigma} + \bar{\alpha}\mathbf{I}_M)^{-2}\mathbf{\Sigma}(\mathbf{\Sigma} + \bar{\alpha}\mathbf{I}_M)^{-1}\mathbf{h}}{\bar{t}} & \\
 - 2 \frac{\mathbf{h}^H(\mathbf{\Sigma} + \bar{\alpha}\mathbf{I}_M)^{-3}\mathbf{h}}{\bar{t}}\alpha - \frac{\mathbf{h}^H(\mathbf{\Sigma} + \bar{\alpha}\mathbf{I}_M)^{-2}\mathbf{h}}{\bar{t}^2}t & \tag{80}
 \end{aligned}$$

for all $\alpha > 0, t > 0, \bar{\alpha} > 0$ and $\bar{t} > 0$.

APPENDIX II: CONVEXITY OF NONSPECTRAL FUNCTIONS AND LOG-DET INEQUALITIES

In the previous appendix we obtained some results on convexity of function of real vector variables. There are lots of interesting results on convexity of the so-called spectral functions of matrix variables (see e.g. [41]) such as $f(\text{trace}(\mathbf{X}))$ or $f(\lambda_{\min}(\mathbf{X}))$ or $f(\lambda_{\max}(\mathbf{X}))$, where $\lambda_{\min}(\mathbf{X})$ and $\lambda_{\max}(\mathbf{X})$ are respectively the minimal and maximal eigenvalue of the matrix variable $\mathbf{X}$, which are dependent on the set of eigenvalues of the matrix variable but not dependent on the order of these eigenvalues. As expected, it is much more difficult to analyse convexity of nonspectral functions and in this Appendix we provide some first results about it.

Theorem 2: For $\mathbf{A} \succ 0$ function $\log|\mathbf{A} + \mathbf{H}\mathbf{X}^{-1}\mathbf{H}^H|$ is convex in $\mathbf{X} \succ 0$.

Proof: One has

$$(\mathbf{A} + \mathbf{H}\mathbf{X}^{-1}\mathbf{H}^H)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{H}^H\mathbf{A}^{-1}\mathbf{H} + \mathbf{X})^{-1}\mathbf{A}^{-1}$$

By [42, Appendix B]

$$f(\mathbf{X}) = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{H}^H\mathbf{A}^{-1}\mathbf{H} + \mathbf{X})^{-1}\mathbf{A}^{-1}$$

is concave

$$f(\alpha\mathbf{X} + \beta\mathbf{Y}) \succ \alpha f(\mathbf{X}) + \beta f(\mathbf{Y})$$

for all $\mathbf{X} \succ 0, \mathbf{Y} \succ 0$ and $\alpha \geq 0, \beta \geq 0, \alpha + \beta = 1$. Therefore

$$\begin{aligned} \log|f(\alpha\mathbf{X} + \beta\mathbf{Y})| &\geq \\ \log|\alpha f(\mathbf{X}) + \beta f(\mathbf{Y})| &\geq \\ \alpha \log|f(\mathbf{X})| + \beta \log|f(\mathbf{Y})|, \end{aligned}$$

showing that $\log|\mathbf{A} + \mathbf{H}\mathbf{X}^{-1}\mathbf{H}^H|^{-1} = -\log|\mathbf{A} + \mathbf{H}\mathbf{X}^{-1}\mathbf{H}^H|$ is concave in $\mathbf{X}$. $\square$

Theorem 3: Function $\log|\mathbf{X}_1^{-1} + \mathbf{X}_2^{-1}|$ is convex in $\mathbf{X}_1 \succ 0$ and $\mathbf{X}_2 \succ 0$.

Proof: Define

$$\begin{aligned} f_1(\mathbf{X}_1, \mathbf{X}_2) &\triangleq (\mathbf{X}_1^{-1} + \mathbf{X}_2^{-1})^{-1} \\ &= \mathbf{X}_1 - \mathbf{X}_1(\mathbf{X}_1 + \mathbf{X}_2)^{-1}\mathbf{X}_1. \end{aligned}$$

By [42, Appendix B] f_1 is concave:

$$\begin{aligned} f_1(\alpha(\mathbf{X}_1, \mathbf{X}_2) + \beta(\mathbf{Y}_1, \mathbf{Y}_2)) &\succeq \\ \alpha f_1(\mathbf{X}_1, \mathbf{X}_2) + \beta f_1(\mathbf{Y}_1, \mathbf{Y}_2) \end{aligned}$$

for all $\mathbf{X}_i \succ 0, \mathbf{Y}_i \succ 0$ and $\alpha \geq 0, \beta \geq 0, \alpha + \beta = 1$. Therefore

$$\begin{aligned} \log|f_1(\alpha(\mathbf{X}_1, \mathbf{X}_2) + \beta(\mathbf{Y}_1, \mathbf{Y}_2))| &\geq \\ \log|\alpha f_1(\mathbf{X}_1, \mathbf{X}_2) + \beta f_1(\mathbf{Y}_1, \mathbf{Y}_2)| &\geq \\ \alpha \log|f_1(\mathbf{X}_1, \mathbf{X}_2)| + \beta \log|f_1(\mathbf{Y}_1, \mathbf{Y}_2)| \end{aligned}$$

showing that $\log|f_1(\mathbf{X}_1, \mathbf{X}_2)| = -\log|\mathbf{X}_1^{-1} + \mathbf{X}_2^{-1}|$ is concave. $\square$

Applying Theorem 3 yields the following inequality

$$\begin{aligned} \log|\mathbf{X}_1^{-1} + \mathbf{X}_2^{-1}| &\geq \log|\bar{\mathbf{X}}_1^{-1} + \bar{\mathbf{X}}_2^{-1}| + \text{trace}(\mathbf{I}) \\ &\quad + \text{trace}([\bar{\mathbf{X}}_1 + \bar{\mathbf{X}}_2]^{-1} - \bar{\mathbf{X}}_1^{-1})\mathbf{X}_1 \\ &\quad + \text{trace}([\bar{\mathbf{X}}_1 + \bar{\mathbf{X}}_2]^{-1} - \bar{\mathbf{X}}_2^{-1})\mathbf{X}_2 \\ &\quad \forall \mathbf{X}_i \succ 0, \bar{\mathbf{X}}_i \succ 0, i = 1, 2. \end{aligned} \quad (81)$$

Using the convexity of function $\log|\mathbf{X}^{-1}|$ yields the following inequality

$$\begin{aligned} \log|\mathbf{X}_1^{-1}| &\geq \log|\bar{\mathbf{X}}_1^{-1}| + \text{trace}(\mathbf{I}) - \text{trace}(\bar{\mathbf{X}}_1^{-1}\mathbf{X}_1) \\ &\quad \forall \mathbf{X}_1 \succ 0, \bar{\mathbf{X}}_1 \succ 0. \end{aligned} \quad (82)$$

Noting that function $f(\alpha) = \ln|\bar{\mathbf{H}}_n^H(\Sigma + \alpha\mathbf{I}_M)^{-1}\bar{\mathbf{H}}_n|$ is convex in $\alpha > 0$ by Theorem 2, we then have the following inequality

$$\begin{aligned} \ln|\bar{\mathbf{H}}_n^H(\Sigma + \alpha\mathbf{I}_M)^{-1}\bar{\mathbf{H}}_n| &\geq \\ f(\bar{\alpha}) + \langle \nabla f(\bar{\alpha}), \alpha - \bar{\alpha} \rangle &= \\ \ln|\bar{\mathbf{H}}_n^H(\Sigma + \bar{\alpha}\mathbf{I}_M)^{-1}\bar{\mathbf{H}}_n| & \\ -\text{trace}([\bar{\mathbf{H}}_n^H(\Sigma + \bar{\alpha}\mathbf{I}_M)^{-1}\bar{\mathbf{H}}_n]^{-1} & \\ \times \bar{\mathbf{H}}_n^H(\Sigma + \bar{\alpha}\mathbf{I}_M)^{-2}\bar{\mathbf{H}}_n) (\alpha - \bar{\alpha}) & \end{aligned} \quad (83)$$

for all $\alpha > 0$ and $\bar{\alpha} > 0$.

APPENDIX III: DERIVATION FOR (12)

$$\begin{aligned} \pi(p, \alpha) &= \text{trace}(\mathbf{F}^H\mathbf{F}) \\ &= \text{trace}(\text{diag}[\sqrt{p_n}]_{n=1}^N (\mathbf{H}^H\mathbf{H} + \alpha\mathbf{I}_N)^{-1}\mathbf{H}^H \\ &\quad \times \mathbf{H}(\mathbf{H}^H\mathbf{H} + \alpha\mathbf{I}_N)^{-1}\text{diag}[\sqrt{p_n}]_{n=1}^N) \\ &= \text{trace}((\mathbf{H}^H\mathbf{H} + \alpha\mathbf{I}_N)^{-1}\mathbf{H}^H \\ &\quad \times \mathbf{H}(\mathbf{H}^H\mathbf{H} + \alpha\mathbf{I}_N)^{-1}\text{diag}[p_n]_{n=1}^N) \\ &= \text{trace}(\mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \alpha\mathbf{I}_M)^{-1}(\mathbf{H}\mathbf{H}^H + \alpha\mathbf{I}_M)^{-1} \\ &\quad \times \mathbf{H}\text{diag}[p_n]_{n=1}^N) \\ &= \sum_{n=1}^N p_n (\mathbf{h}_n^H(\mathbf{H}\mathbf{H}^H + \alpha\mathbf{I}_M)^{-2}\mathbf{h}_n) \\ &= (12). \end{aligned}$$

APPENDIX IV: DERIVATION FOR (49) AND (50)

$$\begin{aligned} r_n(\mathbf{P}, \alpha) &= \\ \ln|\mathbf{I}_{N_r} + (\bar{\mathbf{H}}_n^H\mathcal{D}(\alpha)\bar{\mathbf{H}}_n\mathbf{P}_n)^2 & \\ \times \left(\sum_{j \neq n} (\bar{\mathbf{H}}_n^H\mathcal{D}(\alpha)\bar{\mathbf{H}}_j\mathbf{P}_j)^2 + (\sigma^2/\beta_n)\mathbf{I}_{N_r} \right)^{-1}| &= \\ \ln|(\bar{\mathbf{H}}_n^H\mathcal{D}(\alpha)\bar{\mathbf{H}}_n\mathbf{P}_n)^2| & \\ + \ln|(\bar{\mathbf{H}}_n^H\mathcal{D}(\alpha)\bar{\mathbf{H}}_n\mathbf{P}_n)^{-2} + & \\ \left(\sum_{j \neq n} (\bar{\mathbf{H}}_n^H\mathcal{D}(\alpha)\bar{\mathbf{H}}_j\mathbf{P}_j)^2 + (\sigma^2/\beta_n)\mathbf{I}_{N_r} \right)^{-1}| &= (49), \end{aligned}$$

and

$$\begin{aligned} \pi(\mathbf{P}, \alpha) &= \text{trace}(\tilde{\mathbf{P}}^H\tilde{\mathbf{P}}) \\ &= \text{trace}(\text{diag}[\mathbf{P}_n]_{n=1}^N (\mathbf{H}^H\mathbf{H} + \alpha\mathbf{I}_{NN_r})^{-1}\mathbf{H}^H \\ &\quad \times \mathbf{H}(\mathbf{H}^H\mathbf{H} + \alpha\mathbf{I}_{NN_r})^{-1}\text{diag}[\mathbf{P}_n]_{n=1}^N) \\ &= \text{trace}(\mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \alpha\mathbf{I}_M)^{-1} \\ &\quad \times (\mathbf{H}\mathbf{H}^H + \alpha\mathbf{I}_M)^{-1}\mathbf{H}\text{diag}[\mathbf{P}_n]_{n=1}^N) \\ &= (50). \end{aligned}$$

REFERENCES

- [1] M. Shafi et al, "5G: A tutorial overview of standards trials challenges deployment and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, pp. 1201–1221, Jun. 2017.
- [2] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, pp. 461–471, Feb. 2004.
- [3] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, pp. 4409–4418, Sep. 2008.

- [4] S.-R. Lee, J.-S. Kim, S.-H. Moon, H.-B. Kong, and I. Lee, "Zero-forcing beamforming in multiuser MISO downlink systems under per-antenna power constraint and equal-rate metric," *IEEE Trans. Wirel. Commun.*, vol. 12, pp. 228–236, Jan. 2013.
- [5] F. Rusek et al., "Scaling up MIMO: opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, pp. 40–60, Jan 2013.
- [6] Z. Xu, C. Yang, G. Y. Li, Y. Liu, and S. Xu, "Energy-efficient CoMP precoding in heterogeneous networks," *IEEE Trans. Signal Process.*, vol. 62, pp. 1005–1017, Feb. 2014.
- [7] L. D. Nguyen, H. D. Tuan, T. Q. Duong, O. A. Dobre, and H. V. Poor, "Downlink beamforming for energy-efficient heterogeneous networks with massive MIMO and small cells," *IEEE Trans. Wirel. Commun.*, vol. 17, pp. 3386–3400, May 2018.
- [8] H. H. Yang, G. Geraci, T. Q. S. Quek, and J. G. Andrews, "Cell-edge-aware precoding for downlink massive MIMO cellular networks," *IEEE Trans. Signal Process.*, vol. 65, pp. 3344–3358, Jul. 2017.
- [9] D. Kim, G. Lee, and Y. Sung, "Two-stage beamformer design for massive MIMO downlink by trace quotient formulation," *IEEE Trans. Commun.*, vol. 63, pp. 2200–2211, Jun. 2015.
- [10] L. D. Nguyen, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Beamforming and power allocation for energy-efficient massive MIMO," in *Proc. 22nd Inter. Conf. Digital Signal Process. (DSP2017)*, p. 105, Aug. 2017.
- [11] C. B. Peel, B. M. Hochwald, and A. L. Swindlerhurst, "A vector-perturbation technique for near capacity multi-antenna multiuser communication-part I: channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, pp. 195–202, Jan. 2005.
- [12] V. Stankovic and M. Haardt, "Generalized design of multi-user MIMO precoding matrices," *IEEE Trans. Wirel. Commun.*, vol. 7, pp. 953–961, Mar. 2008.
- [13] H. Sung, S. Lee, and I. Lee, "Generalized channel inversion methods for multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 57, pp. 3489–3499, Nov. 2009.
- [14] H. Tataria, P. J. Smith, and P. A. Dmochowski, "On the general analysis of coordinated regularized zero-forcing precoding: An application to two-tier small-cell networks," *IEEE Trans. Commun.*, vol. 65, pp. 3133–3150, Jul. 2017.
- [15] R. de Francisco and D. T. Slock, "An optimized unitary beamforming technique for MIMO broadcast channels," *IEEE Trans. Wirel. Commun.*, vol. 9, pp. 990–1000, Mar. 2010.
- [16] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, pp. 4509–4537, Jul. 2012.
- [17] A. Muller, R. Couillet, E. Björnson, S. Wagner, and M. Debbah, "Interference-aware RZF precoding for multicell downlink systems," *IEEE Trans. Signal Process.*, vol. 63, pp. 3959–3973, Aug. 2015.
- [18] H. Lin, F. Gao, S. Jin, and G. Y. Li, "A new view of multi-user hybrid massive MIMO: Non-orthogonal angle division multiple access," *IEEE J. Sel. Areas Commun.*, vol. 35, pp. 2268–2280, Oct. 2017.
- [19] A. Kammoun, A. Muller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, pp. 861–875, Oct. 2014.
- [20] S. Park, J. Park, A. Yazdan, and R. W. Heath, "Optimal user loading in massive MIMO systems with regularized zero forcing precoding," *IEEE Wireless Commun. Lett.*, vol. 6, pp. 118–121, Feb. 2017.
- [21] K. Zu, R. C. de Lamare, and M. Haardt, "Generalized design of low-complexity block diagonalization type precoding algorithms for multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, pp. 4232–4242, Oct. 2013.
- [22] L. D. Nguyen, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Multi-cell massive MIMO beamforming in assuring QoS for large numbers of users," <https://arxiv.org/abs/1712.03548>, 2017.
- [23] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffio, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wirel. Commun.*, vol. 7, pp. 4792–4799, Dec. 2008.
- [24] A. H. Phan, H. D. Tuan, H. H. Kha, and D. T. Ngo, "Nonsmooth optimization for efficient beamforming in cognitive radio multicast transmission," *IEEE Trans. Signal Process.*, vol. 60, pp. 2941–2951, Jun. 2012.
- [25] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "Linear transceiver design for a MIMO interfering broadcast channel achieving max-min fairness," *Signal Process.*, vol. 93, no. 12, pp. 3327–3340, 2013.
- [26] 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, Jun. 2016.
- [27] J. Chen, H. Yin, L. Cottatellucci, and D. Gesbert, "Feedback mechanisms for FDD massive MIMO with D2D-based limited CSI sharing," *IEEE Trans. Wirel. Commun.*, vol. 16, pp. 5162–5175, Aug. 2017.
- [28] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Secrecy rate beamforming for multicell networks with information and energy harvesting," *IEEE Trans. Signal Process.*, vol. 65, pp. 677–689, Feb. 2017.
- [29] S. Buzzi, C.-L. I, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5G networks and challenges ahead," *IEEE J. Select. Areas Commun.*, vol. 34, pp. 697–709, Apr. 2016.
- [30] A. Zappone, L. Sanguinetti, G. Bacci, E. A. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wirel. technologies," *IEEE Trans. Signal Process.*, vol. 64, pp. 1668–1683, Apr. 2016.
- [31] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.
- [32] Y. Shi, H. D. Tuan, S. Su, and H. Tuy, "Global optimization for optimal power flow over transmission networks," *J. of Global Optimization*, vol. 69, pp. 745–760, 2017.
- [33] Y. Shi, H. D. Tuan, A. V. Savkin, T. Q. Duong, and H. V. Poor, "Model predictive control for smart grids with multiple electric-vehicle charging stations," *IEEE Trans. Smart Grid*, vol. 10, pp. 2127–2136, Feb. 2019.
- [34] J. F. Bonnans, J. C. Gilbert, C. Lemarechal, and C. Sabastigabal, *Numerical Optimization Theoretical and Practical Aspects (second edition)*. Springer, 2006.
- [35] E. Björnson, M. Kountouris, and M. Debbah, "Massive MIMO and small cells: improving energy efficiency by optimal soft-cell coordination," in *ICT*, pp. 1–5, May 2013.
- [36] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, pp. 1065–1082, June 2014.
- [37] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: a large system analysis," in *Proc. of 2008 Conf. Global Commun. (GLOBECOM)*, pp. 1–4, 2008.
- [38] Z. Sheng, H. D. Tuan, H. H. M. Tam, H. H. Nguyen, and Y. Fang, "Energy-efficient precoding in multicell networks with full-duplex base stations," *EURASIP J. Wirel. Commun. Network.*, vol. 2017:48, Dec. 2017. <https://doi.org/10.1186/s13638-017-0831-5>.
- [39] H. H. M. Tam, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "MIMO energy harvesting in full-duplex multi-user networks," *IEEE Trans. Wirel. Commun.*, vol. 16, pp. 3282–3297, May 2017.
- [40] H. Tuy, *Convex Analysis and Global Optimization (second edition)*. Springer International, 2017.
- [41] A. Lewis, "Derivatives of spectral functions," *Mathematics of Operations Research*, vol. 21, pp. 576–588, Aug. 1996.
- [42] U. Rashid, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks," *IEEE Trans. Commun.*, vol. 62, pp. 488–499, Feb. 2014.

Long D. Nguyen was born in Dong Nai, Vietnam. He received his B.S. degree in Electrical and Electronics Engineering and M.S. degree in Telecommunication Engineering from Ho Chi Minh City University of Technology (HCMUT), Vietnam, in 2013 and 2015, respectively. He received his Ph.D. degree in Electronics and Electrical Engineering from Queen's University Belfast (QUB), UK, in 2018. He was a Research Fellow at Queen's University Belfast, UK for a part of Newton project (2018-2019). He is currently with Dong Nai University in Vietnam as an Assistant Professor and Duy Tan University as an Adjunct Assistant Professor. His research interests include convex optimization techniques for resource management in wireless communications, energy efficiency approaches (heterogeneous networks, relay networks, cell-free networks and massive MIMO), and real-time embedded optimization for wireless networks and Internet of Things (IoT).

Hoang Duong Tuan received the Diploma (Hons.) and Ph.D. degrees in applied mathematics from Odessa State University, Ukraine, in 1987 and 1991, respectively. He spent nine academic years in Japan as an Assistant Professor in the Department of Electronic-Mechanical Engineering, Nagoya University, from 1994 to 1999, and then as an Associate Professor in the Department of Electrical and Computer Engineering, Toyota Technological Institute, Nagoya, from 1999 to 2003. He was a Professor with the School of Electrical Engineering and Telecommunications, University of New South Wales, from 2003 to 2011. He is currently a Professor with the School of Electrical and Data Engineering, University of Technology Sydney. He has been involved in research with the areas of optimization, control, signal processing, wireless communication, and biomedical engineering for more than 20 years.

Trung Q. Duong (S'05, M'12, SM'13) received his Ph.D. degree in Telecommunications Systems from Blekinge Institute of Technology (BTH), Sweden in 2012. Currently, he is with Queen's University Belfast (UK), where he was a Lecturer (Assistant Professor) from 2013 to 2017 and a Reader (Associate Professor) from 2018. His current research interests include Internet of Things (IoT), wireless communications, molecular communications, and signal processing. He is the author or co-author of 290 technical papers published in scientific journals

(165 articles) and presented at international conferences (125 papers).

Dr. Duong currently serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, IET COMMUNICATIONS, and a Lead Senior Editor for IEEE COMMUNICATIONS LETTERS. He was awarded the Best Paper Award at the IEEE Vehicular Technology Conference (VTC-Spring) in 2013, IEEE International Conference on Communications (ICC) 2014, IEEE Global Communications Conference (GLOBECOM) 2016, and IEEE Digital Signal Processing Conference (DSP) 2017. He is the recipient of prestigious Royal Academy of Engineering Research Fellowship (2016-2021) and has won a prestigious Newton Prize 2017.

H. Vincent Poor (S'72, M'77, SM'82, F'87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor of Electrical Engineering. During 2006 to 2016, he served as Dean of Princeton's School of Engineering and Applied Science. He has also held visiting appointments at several other universities, including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory and signal processing, and their applications in wireless networks, energy systems and related fields. Among his publications in these areas is the recent book *Information Theoretic Security and Privacy of Information Systems* (Cambridge University Press, 2017).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and is a foreign member of the Chinese Academy of Sciences, the Royal Society and other national and international academies. He received the Technical Achievement and Society Awards of the IEEE Signal Processing Society in 2007 and 2011, respectively. Recent recognition of his work includes the 2017 IEEE Alexander Graham Bell Medal, Honorary Professorships from Peking University and Tsinghua University, both conferred in 2017, and a D.Sc. *honoris causa* from Syracuse University, awarded in 2017.