Energy-Efficient Multi-Cell Massive MIMO Subject to Minimum User-Rate Constraints

Long D. Nguyen, Hoang D. Tuan, Trung Q. Duong, H. Vincent Poor, and Lajos Hanzo

***Abstract*—The capability of massive multiple-input multiple-output (mMIMO) systems supporting the throughput requirement of as many users as possible is investigated. The bottleneck of serving small numbers of users by a large number of transmit antennas in conventional mMIMO is unblocked by a new time-fraction-wise beamforming technique, which focuses signal transmission in fractions of a time slot. Based on this time-fraction-wise signal transmission, a new user service scheduling for multi-cell mMIMO, whose cell-edge users suffer not only poor channel conditions but also multi-cell interference, is proposed to support a large user-population. We demonstrate that the numbers of users served by our multi-cell mMIMO within a time-slot may be as high as twice the number of its transmit antennas.**

***Index Terms*—Multi-cell massive MIMO system, beamformer design, energy efficiency, quality-of-service, user service scheduling, nonconvex optimization**

I. Introduction

Massive multiple-input multiple-output (mMIMO) [1], [2] schemes are capable of improving the quality-of-service (QoS) in terms of the throughput of cell-edge users. As envisioned in the pioneering treatise [3], [4], mMIMO schemes have been designed for serving a number of users, which is much smaller than the number of low-power transmit antennas. Under these conditions, mMIMO schemes benefit from so-called favorable propagation characteristics such as the orthogonality of communication channels [1], [5] and the deterministic behavior of the channels' eigenvalue distribution [6], [7], which allow low-complexity zero-forcing beamforming (ZFB) to perform well [8], [9]. The performance analysis of this ZFB typically relies on having identical power allocation for all the beamformers [10]. Our previous work [11] shows that the users' QoS can be significantly improved by employing the optimal power allocation for the beamformers. It also shows that using optimal power-allocation ZFB performs much better than the optimal power-allocation aided conjugate beamformer, even though the latter was shown to perform better than the former under the equi-power allocation [3], [9]. To serve many users, mMIMO schemes have to appropriately schedule their transmissions to a small number of users served at any given time. As such, it is not known at the time of writing if mMIMO schemes are capable of maintaining a high QoS for many users simultaneously, regardless whether they are cell-center users with better channel conditions or cell-edge users with poorer channel conditions.

ZFB relies on the right-inverse of the overall mMIMO channel matrix to force the inter-user interference to zero, but the right-inverse does not exist when the number of users is higher than the number of transmit antennas. The involvement of more users results in the ill-conditioning of this right-inverse matrix, which can be overcome by the so-called regularized zero-forcing (RZFB) [8], [12]. However, by employing RZFB, the inter-user interference can no longer be forced to zero and its impact on the performance of RZFB must be addressed [13]. Another issue of mMIMO schemes is that their transmit antennas, which are closely packed in a compact space, are often assumed to be spatially uncorrelated. Under this assumption, the channel matrices are well-conditioned and ZFB is expected to perform well [14]. However, in realistic scattering environments, these antennas are inherently spatially correlated [15], [16], which reduces the rank of the channel matrices, hence reducing the capacity of mMIMO schemes.

For communication networks, the energy-efficiency (EE), which is defined as the ratio between the total information throughput and the total consumed power is an important metric [17], [18]. More explicitly, the EE has a substantial impact on the affordable number of transmit antennas, which on one hand should be high for improving the transmit power-scaling law [14] but should be low for reducing the circuit power dissipation. Reducing the circuit power requires a low number of radio frequency (RF) chains which in turn erodes the ZFB's appeals. More importantly, the diversity

L. D. Nguyen is with the Department of Engineering, Dong Nai University, Dong Nai, 810000, Vietnam (email:dinhlonghcmut@gmail.com). He would like to acknowledge the support of the Global Challenges Research Fund under the DfE-GCRF 2020-2021 scheme between Queen's University Belfast, UK and Dong Nai University, Vietnam.

H. 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). He would like to acknowledge the support of the Australian Research Council's Discovery Projects under Project DP190102501.

T. Q. Duong is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, BT7 1NN, UK (email:trung.q.duong@qub.ac.uk). He would like to acknowledge the support of the UK Royal Academy of Engineering (RAEng) under the RAEng Research Chair and Senior Research Fellowship scheme Grant RCSRF2021\11\41, RAEng Research Fellowship scheme Grant RF1415\14\22, and the Global Challenges Research Fund under the DfE-GCRF 2020-2021 scheme between Queen's University Belfast, UK and Dong Nai University, Vietnam.

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA (e-mail: poor@princeton.edu). He would like to acknowledge the support of the U.S. National Science Foundation under Grant CCF-1908308.

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK (email: lh@ecs.soton.ac.uk). He would like to acknowledge the support of the Engineering and Physical Sciences Research Council projects EP/N004558/1, EP/P034284/1, EP/P034284/1, EP/P003990/1 (COALESCE), of the Royal Society's Global Challenges Research Fund Grant as well as of the European Research Council's Advanced Fellow Grant QuantCom.

order of mMIMO is limited by the number of RF chains used. Maximizing the EE under the users' QoS constraints requires the optimization of the power/network throughput ratio. It is important to emphasize here that the capabilities of mMIMO to serve many users is considered in this paper in a different light from [19]. Explicitly, the authors of [19] consider the system sum throughput without imposing the users' QoS requirement. As a result, in [19] most of the throughput would be shared by a small fraction of the users having the best channels. By contrast, the users having a low-quality channel would fail to have an adequate QoS. The authors of [20] invoked a mMIMO scheme to serve a massive number of users by a massive number of orthogonal frequency bands, where not more than 5 are served over the same frequency band within a time slot [20, Table II]. Hence both its ZFB as well as its conjugate beamforming (termed as MR in [20]) and RZFB (termed as MMSE in [20]) serve only a few users. On the other hand, the authors of [21] and [22] designed a mMIMO scheme for massive uplink connectivity and analyzed its achievable sum rate as both the numbers of users and transmit antennas tend to infinity with their ratio kept fixed, hence making the channel conditions look more deterministic [8].

In multi-cell mMIMO arrangements, the cell-edge users suffer not only from poor channel conditions as they do in single-cell mMIMO, but also from inter-cell interference, which cannot be readily mitigated by the existing cell-wide ZFB or RZFB. As such, the problem of providing high QoS for the users of multi-cell mMIMO arrangements is much more technically challenging than that in single-cell mMIMO. The inter-cell interference management is critically important, making the multi-cell mMIMO EE more relevant than ever.

Against this backdrop, this is the first treatise exploiting mMIMO schemes for serving large numbers of users whilst guaranteeing a minimum throughput for each user. From now on, let us always be specific and refer to minimum throughput (MTp), rather than to QoS, which may mean numerous things. Explicitly, our contributions are as follows:

- To mitigate the intra-cell inter-user interference, which prevents mMIMO from serving many users benefiting from a guaranteed MTp within a time-slot, we propose mini-slot based transmit beamforming (MS-TBF). This new beamforming technique uses a fraction of the time-slot, termed as a mini-slot, for serving a subgroup of users and then exploits the remaining time-fraction to serve the complementary subgroups of other users. This allows our mMIMO scheme to provide a uniformly high QoS for many users. We then proceed by considering the problem of maximizing the single-cell EE under a minimum user-throughput guarantee.
- For efficiently suppressing the multi-cell interference, which typically prevents multi-cell mMIMO arrangements from serving many users within a time slot, we propose a new scheduling scheme for the proposed MS-TBF. This allows multi-cell mMIMO scheme to provide a guaranteed MTp for many cell-edge users, which then serves as the motivation for maximizing the multi-cell EE under the users' MTp constraints. Importantly, in contrast from the existing multi-cell beamforming schemes, the proposed scheme does not need any information on the interfering channels of all the cells for rejecting the inter-cell interference.
- However, our MS-TBF technique relies on computationally challenging optimization problems. Hence one of our contributions is to develop new algorithms for its computation, which generate a sequence of improved feasible points and converge at least to a locally optimal solution. Simulation confirms that our massive MIMO scheme is capable of maintaining the required MTp for a large number of users.

The paper is organized as follows. In Section II, the problem of maximizing the multi-cell mMIMO EE subject to the users' MTp by MS-TBF is addressed, which has not been previously considered. Furthermore, ZFB and new RZFB aided mMIMO schemes are proposed. The MS-TBF schemes are capable of serving many more users at a MTp as proposed in Section III. Accordingly the problem of maximizing the multi-cell EE subject to the users' MTp is posed and solved. Our simulations are provided in Section IV, while our conclusions are provided in Section V. Appendix I derives some of the important inequalities used in the algorithmic developments, while Appendices II and III elaborate on some of the equations.

Notation. Boldface upper and lowercase letters denote matrices and vectors, respectively. The transpose and conjugate transpose of a matrix $\mathbf{X}$ are respectively represented by $\mathbf{X}^T$ and $\mathbf{X}^H$. $\mathbf{I}$ and $\mathbf{0}$ stand for identity and zero matrices of appropriate dimensions. $\mathrm{Tr}(.)$ is the trace operator. $||\mathbf{x}||$ is the Euclidean norm of the vector $\mathbf{x}$ and $||\mathbf{X}||$ is the Frobenius norm of the matrix $\mathbf{X}$. A Gaussian random vector with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{R}_{\mathbf{x}}$ is denoted by $\mathbf{x} \sim \mathcal{CN}(\bar{\mathbf{x}}, \mathbf{R}_{\mathbf{x}})$. For matrices $\mathbf{X}_i$, $i = 1, \ldots, \mathbf{X}_k$ of appropriate dimension, $\mathsf{Col}[\mathbf{X}_i]_{i=1,\ldots,K}$ or $\mathsf{Col}[\mathbf{X}_i]_{i\in\mathcal{K}}$ for $\mathcal{K} \triangleq \{1, \ldots, k\}$ arranges $\mathbf{X}_i$ in a block-column format exemplified by

$$\mathsf{Col}[\mathbf{X}_i]_{i\in\mathcal{K}} = \begin{bmatrix} \mathbf{X}_1 \\ \ldots \\ \mathbf{X}_k \end{bmatrix}.$$

Hence it is true that $\mathsf{Col}[\mathbf{X}_i]_{i\in\mathcal{K}}\mathbf{A} = \mathsf{Col}[\mathbf{X}_i\mathbf{A}]_{i\in\mathcal{K}}$. Analogously, $\mathsf{Row}[\mathbf{X}_i]_{i=1,\ldots,K}$ or $\mathsf{Row}[\mathbf{X}_i]_{i\in\mathcal{K}}$ arranges $\mathbf{X}_i$ in a block-row format, yielding

$$\mathsf{Row}[\mathbf{X}_i]_{i\in\mathcal{K}} = \begin{bmatrix} \mathbf{X}_1 & \ldots & \mathbf{X}_k \end{bmatrix}.$$

Therefore it is true that $\mathbf{A}\mathsf{Row}[\mathbf{X}_i]_{i\in\mathcal{K}} = \mathsf{Row}[\mathbf{A}\mathbf{X}_i]_{i\in\mathcal{K}}$. Lastly, $\mathrm{diag}[a_k]_{k\in\mathcal{K}}$ is a diagonal matrix with a_k in its diagonal.

II. Zero-forcing and regularized zero-forcing beamforming

A. General formulation and solution

We consider three neighboring cells replying on a mMIMO system subject to severe inter-cell interference, as illustrated by Fig. 1.[1]

[1]A system having three neighboring cells is the most popular constellation for investigating the effects of severe mutual interference in practice because each cell is a neighbour of *all* other cells and as such its boundary users are severely interfered by *all* other cells.

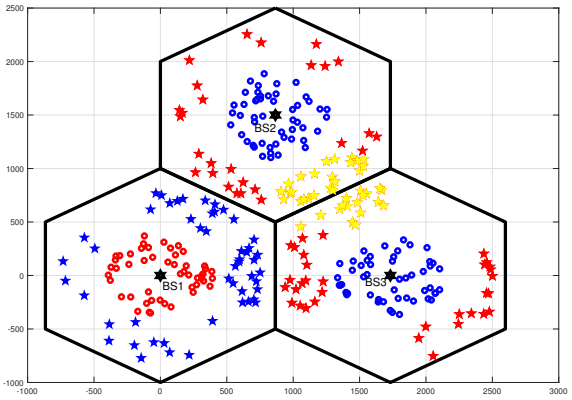


Fig. 1: Three-cell mMIMO with severe multi-cell interference. The users marked by yellow colour are served over a different time slot or frequency band. The users marked by the same colour (red or blue) are served by the same time-fraction of the time slot.

Each base station (BS) $i \in \mathcal{I} \triangleq \{1,2,3\}$ is equipped with a large-scale N-element antenna array to serve its N_{UE} single-antenna aided users (UEs)(i,k), $k \in \mathcal{K} \triangleq \{1,\ldots,N_{UE}\}$ within its cell.

Let $s_{i,k} \sim \mathcal{CN}(0,1)$ be the downlink (DL) information transmitted from BS i intended for its UE (i,k). The vector of information from BS i intended for all its UEs is defined as $\mathbf{s}_i = \mathsf{Col}[s_{i,k}]_{k\in\mathcal{K}}$. Each $s_{i,k}$ is processed by the TBF vector $\mathbf{f}_{i,k} \in \mathbb{C}^N$. The resultant TBF matrix is defined by

$$\mathbf{F}_i \triangleq \mathsf{Row}[\mathbf{f}_{i,k}]_{k\in\mathcal{K}} \in \mathbb{C}^{N\times N_{UE}}.$$

The signal transmitted from BS i is $\mathbf{x}_i = \mathbf{F}_i\mathbf{s}_i$.

The vector channel response(C/R) from BS j to UE (i,k) is modeled by $\sqrt{\beta_{j,i,k}}\mathbf{h}_{j,i,k}$, where $\sqrt{\beta_{j,i,k}}$ represents the path-loss and large-scale fading, while according to [16], [23], [24] we have

$$\mathbf{h}_{j,i,k} = \mathbf{\Theta}_{j,i,k}^{1/2}\mathbf{h}_{j,i,k}^{w}, \tag{1}$$

where $\mathbf{\Theta}_{j,i,k} \in \mathbb{C}^{N\times N}$ is a Hermitian symmetric positive semi-definite spatial correlation matrix, and $\mathbf{h}_{j,i,k}^{w} \in \mathbb{C}^N$ represents the small-scale fading having entries that are independently and identically generated by $\mathcal{CN}(0,1)$.

The channel matrix between the BS j and the UEs in i-th cell is thus given by $\boldsymbol{\beta}_{j,i}\mathbf{H}_{j,i}^H$ where $\boldsymbol{\beta}_{j,i} \triangleq \mathsf{diag}[\sqrt{\beta_{j,i,k}}]_{k\in\mathcal{K}}$ and

$$\mathbf{H}_{j,i}^H \triangleq \mathsf{Col}[\mathbf{h}_{j,i,k}^H]_{k\in\mathcal{K}} \in \mathbb{C}^{N_{UE}\times N}. \tag{2}$$

Let $y_{i,k} \in \mathbb{C}$ be the signal received at UE (i,k). Then we have $\mathbf{y}_i \triangleq \mathsf{Col}[y_{i,k}]_{k\in\mathcal{K}}$. The MIMO equation is thus

$$\mathbf{y}_i = \boldsymbol{\beta}_{i,i}\mathbf{H}_{i,i}^H\mathbf{F}_i\mathbf{s}_i + \sum_{j\in\mathcal{I}\setminus\{i\}} \boldsymbol{\beta}_{j,i}\mathbf{H}_{j,i}^H\mathbf{F}_j\mathbf{s}_j + \mathbf{n}_i,$$

where $\mathbf{n}_i = \mathsf{Col}[n_{i,k}]_{k\in\mathcal{K}}$ is the noise vector of independent entries $n_{i,k} \in \mathcal{CN}(0,\sigma^2)$. More particularly, the multi-input single-output (MISO) equation for the signal received at UE (i,k) is

$$y_{i,k} = \underbrace{\sqrt{\beta_{i,i,k}}\mathbf{h}_{i,i,k}^H\mathbf{f}_{i,k}s_{i,k}}_{\text{desired signal}} + \underbrace{\sum_{\ell\in\mathcal{K}\setminus\{k\}} \sqrt{\beta_{i,i,k}}\mathbf{h}_{i,i,k}^H\mathbf{f}_{i,\ell}s_{i,\ell}}_{\text{inter-user interference}} + \underbrace{\sum_{j\in\mathcal{I}\setminus\{i\}} \sqrt{\beta_{j,i,k}}\mathbf{h}_{j,i,k}^H\mathbf{F}_j\mathbf{s}_j}_{\text{inter-cell interference}} + n_{i,k}. \tag{3}$$

We now seek a TBF matrix $\mathbf{F}_i$ in the following class

$$\mathbf{F}_i = \bar{\mathbf{F}}_i\mathsf{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} \tag{4}$$

having a predetermined matrix

$$\bar{\mathbf{F}}_i \triangleq \mathsf{Row}[\bar{\mathbf{f}}_{i,k}]_{k\in\mathcal{K}} \in \mathbb{C}^{N\times N_{UE}}. \tag{5}$$

For $\mathbf{p}_i = \mathsf{Col}[p_{i,k}]_{k\in\mathcal{K}}$ and $\mathbf{p} = (\mathbf{p}_i)_{i\in\mathcal{I}}$, the inter-user interference and inter-cell interference functions are respectively defined from (3) as

$$\sigma_{i,k}^{\mathsf{U}}(\mathbf{p}_i) \triangleq \beta_{i,i,k}\sum_{\ell\in\mathcal{K}\setminus\{k\}} |\mathbf{h}_{i,i,k}^H\bar{\mathbf{f}}_{i,\ell}|^2 p_{i,\ell}, \tag{6}$$

and

$$\sigma_{i,k}^{\mathsf{C}}(\mathbf{p}) \triangleq \beta_{j,i,k}\sum_{j\in\mathcal{I}\setminus\{i\}}\sum_{\ell\in\mathcal{K}} ||\bar{\mathbf{f}}_{j,\ell}||^2 p_{j,\ell}. \tag{7}$$

Note that while the intra-cell channel $\mathbf{h}_{i,i,k}$ can be indeed estimated [24], the inter-cell channel $\mathbf{h}_{j,i,k}$ in (3) cannot be estimated and must be defined as in (7). Under the definitions

$$\alpha_{i,k} \triangleq \beta_{i,i,k}|\mathbf{h}_{i,i,k}^H\bar{\mathbf{f}}_{i,k}|^2 \tag{8}$$

and

$$\lambda_{i,k}(\mathbf{p}) \triangleq \sigma_{i,k}^{\mathsf{U}}(\mathbf{p}_i) + \sigma_{i,k}^{\mathsf{C}}(\mathbf{p}), \tag{9}$$

where the latter is a linear function, the information throughput of UE (i,k) is defined by

$$r_{i,k}(\mathbf{p}) = \ln\left(1 + \frac{\alpha_{i,k}p_{i,k}}{\lambda_{i,k}(\mathbf{p}) + \sigma^2}\right). \tag{10}$$

The transmit power of BS i is given by the following function, which is also linear:

$$\chi_i(\mathbf{p}_i) = \sum_{k\in\mathcal{K}} ||\bar{\mathbf{f}}_{i,k}||^2 p_{i,k}. \tag{11}$$

The entire power consumption for the DL transmission is expressed by

$$\pi(\mathbf{p}) = \sum_{i\in\mathcal{I}}(\alpha\chi_i(\mathbf{p}_i) + NP_a + P_c), \tag{12}$$

which is an affine function in $\mathbf{p}$. Here $\alpha > 1$ is the reciprocal of the drain efficiency of the BS's amplifier, while P_a and P_c are circuit power per antenna and other non-transmission power of the BSs [18].

The network's total throughput is defined as

$$\varphi(\mathbf{p}) \triangleq \sum_{(i,k)\in\mathcal{I}\times\mathcal{K}} r_{i,k}(\mathbf{p}).$$

We are interested in the following EE maximization problem under MTp and power budget constraints:

$$\max_{\mathbf{p}} \ \varphi(\mathbf{p})/\pi(\mathbf{p}) \quad \text{s.t.} \tag{13a}$$

$$\chi_i(\mathbf{p}_i) \le P_i^{\max}, i \in \mathcal{I}, \tag{13b}$$

$$r_{i,k}(\mathbf{p}) \ge \bar{r}_{i,k}, (i,k) \in \mathcal{I} \times \mathcal{K}, \tag{13c}$$

where the constraint (13c) represents the MTp thresholds at each UE, while the constraint (13b) limits the sum of transmit power to the predefined budget.

From definition (10) of $r_{i,k}(\mathbf{p})$, the constraint (13c) is equivalent to the linear constraint

$$\alpha_{i,k} p_{i,k} \ge (e^{\bar{r}_{i,k}} - 1)(\lambda_{i,k}(\mathbf{p}) + \sigma^2), (i,k) \in \mathcal{I} \times \mathcal{K}. \tag{14}$$

Hence (13) is a linear-constrained optimization problem, which is still a challenging nonconvex problem, since the objective function $\varphi(\mathbf{p})/\pi(\mathbf{p})$ is obviously not concave. Following the approach of [13], we now propose the following new computational procedure.

Let $\mathbf{p}^{(n)}$ be a feasible point of (13) found from the $(n-1)$st iteration and $t^{(n)} \triangleq \varphi(\mathbf{p}^{(n)})/\pi(\mathbf{p}^{(n)})$, so

$$\varphi(\mathbf{p}^{(n)}) - t^{(n)} \pi(\mathbf{p}^{(n)}) = 0. \tag{15}$$

Using inequality (76) in Appendix I for $x = \alpha_{i,k} p_{i,k}$, $y = \lambda_{i,k}(\mathbf{p}) + \sigma^2$, and $\bar{x} = \alpha_{i,k} p_{i,k}^{(n)}$, $\bar{y} = \lambda_{i,k}(\mathbf{p}^{(n)}) + \sigma^2$, yields the following lower bounding approximation:

$$\varphi(\mathbf{p}) \ge \varphi^{(n)}(\mathbf{p})$$

for

$$\varphi^{(n)}(\mathbf{p}) \triangleq \sum_{(i,k) \in \mathcal{I} \times \mathcal{K}} \left(\bar{a}_{i,k}^{(n)} - \frac{\bar{b}_{i,k}^{(n)}}{\alpha_{i,k} p_{i,k}} - \bar{c}_{i,k}^{(n)} (\lambda_{i,k}(\mathbf{p}) + \sigma^2) \right), \tag{16}$$

which is an affine function, where we have

$$0 < \bar{a}_{i,k}^{(n)} \triangleq r_{i,k}(\mathbf{p}^{(n)}) + \frac{2\alpha_{i,k} p_{i,k}^{(n)}}{\alpha_{i,k} p_{i,k}^{(n)} + \lambda_{i,k}(\mathbf{p}^{(n)}) + \sigma^2},$$

$$0 < \bar{b}_{i,k}^{(n)} \triangleq \frac{(\alpha_{i,k} p_{i,k}^{(n)})^2}{\alpha_{i,k} p_{i,k}^{(n)} + \lambda_{i,k}(\mathbf{p}^{(n)}) + \sigma^2},$$

$$0 < \bar{c}_{i,k}^{(n)} \triangleq \frac{\alpha_{i,k} p_{i,k}^{(n)}}{\alpha_{i,k} p_{i,k}^{(n)} + \lambda_{i,k}(\mathbf{p}^{(n)}) + \sigma^2} (\lambda_{i,k}(\mathbf{p}^{(n)}) + \sigma^2).$$

At the nth iteration, the following linear programming subproblem is solved for generating the next feasible point $\mathbf{p}^{(n+1)}$ for (13):

$$\max_{\mathbf{p}} \left[\varphi^{(n)}(\mathbf{p}) - t^{(n)} \pi(\mathbf{p}) \right] \quad \text{s.t.} \quad (13b), (14). \tag{17}$$

The computational complexity of (17) is [25]

$$\mathcal{O}(\bar{n}^2 \bar{m}^{2.5} + \bar{m}^{2.5}) \tag{18}$$

with $\bar{n} = 3N_{UE}$, which is the number of the decision variables, and $\bar{m} = 3(N_{UE} + 1)$, which is the number of constraints.

Note that $\mathbf{p}^{(n)}$ is a feasible point for (17) satisfying (15). Therefore, as long as $\mathbf{p}^{(n+1)} \ne \mathbf{p}^{(n)}$, we have

$$\begin{aligned} \varphi^{(n)}(\mathbf{p}^{(n+1)}) - t^{(n)} \pi(\mathbf{p}^{(n+1)}) &> \\ \varphi^{(n)}(\mathbf{p}^{(n)}) - t^{(n)} \pi(\mathbf{p}^{(n)}) &= \\ \varphi(\mathbf{p}^{(n)}) - t^{(n)} \pi(\mathbf{p}^{(n)}) &= 0, \end{aligned}$$

which implies that

$$t^{(n+1)} \triangleq \varphi^{(n)}(\mathbf{p}^{(n+1)})/\pi(\mathbf{p}^{(n+1)}) > t^{(n)}, \tag{19}$$

i.e. $\mathbf{p}^{(n+1)}$ is a better feasible point than $\mathbf{p}^{(n)}$ for (13). Similarly to [26, Prop.1], it can be readily shown that Algorithm 1 converges at least to a locally optimal solution of (13), satisfying the KKT conditions of optimality.

Algorithm 1 : Path-following algorithm for solving problem (13)

1: **Initialization**: Solve the convex problem $\min_{\mathbf{p}} \max_{i \in \mathcal{I}} [\chi_i(\mathbf{p}_i)/P_i^{\max}]$ s.t. (14). Exit if its optimal value is more than 1, because it means that problem (13) is infeasible. Otherwise, take its optimal solution as a feasible point $\mathbf{p}^{(0)}$ for the convex constraints (13b), (14) and set $n := 0$ and $t^{(0)} = \varphi(\mathbf{p}^{(0)})/\pi(\mathbf{p}^{(0)})$.

2: **Repeat until convergence of the objective in (13)**: Solve the problem (17) for its optimal solution $\mathbf{p}^{(n+1)}$. Set $t^{(n+1)} = \varphi(\mathbf{p}^{(n+1)})/\pi(\mathbf{p}^{(n+1)})$. Set $n := n + 1$.

Remark. It was first observed in [27] that $\varphi(\mathbf{p})$ in (13) is in fact a d.c. (difference of two concave) function [28] $\varphi_d(\mathbf{p}) - \varphi_c(\mathbf{p})$ with $\varphi_d(\mathbf{p}) = \sum_{(i,k) \in \mathcal{I} \times \mathcal{K}} \ln \left(\alpha_{i,k} p_{i,k} + \lambda_{i,k}(\mathbf{p}) + \sigma^2 \right)$ and $\varphi_c(\mathbf{p}) = \sum_{(i,k) \in \mathcal{I} \times \mathcal{K}} \ln \left(\lambda_{i,k}(\mathbf{p}) + \sigma^2 \right)$, which are concave functions. Then the problem (13) can be solved based on the DCIs [27], which invokes the following convex subproblem at the nth iteration instead of the linear programming subproblem (17) to generate the next feasible point $\mathbf{p}^{(n+1)}$ for (13), yielding:

$$\max_{\mathbf{p}} \Big[\varphi_d(\mathbf{p}) - \varphi_c(\mathbf{p}^{(n)}) - \langle \nabla \varphi_c(\mathbf{p}^{(n)}), \mathbf{p} - \mathbf{p}^{(n)} \rangle - t^{(n)} \pi(\mathbf{p}) \Big] \quad \text{s.t.} \quad (13b), (14). \tag{20}$$

However, there is no solver of polynomial complexity order for (20) since it involves logarithmic function optimization (the function $\varphi_d(\mathbf{p})$). This issue is really serious, because (20) represents a large-scale problem, so the advantage of the linear programming problem (17) over the convex problem (20) in terms of large-scale computational tractability is plausible. The algorithm proposed in [18, Alg. 1], solves the following problem at the nth iteration

$$\max_{\mathbf{p}} \frac{\varphi_d(\mathbf{p}) - \varphi_c(\mathbf{p}^{(n)}) - \langle \nabla \varphi_c(\mathbf{p}^{(n)}), \mathbf{p} - \mathbf{p}^{(n)} \rangle}{\pi(\mathbf{p})} \quad \text{s.t.} \quad (13b), (14). \tag{21}$$

to generate the next feasible point $\mathbf{p}^{(n+1)}$ for (13), which suffers from the same issue. Actually, the Dinkelbach-type algorithm used in [18, Alg. 1] is capable of solving the problem (21) by solving a sequence of the problems (20).

B. ZFB and new RZFB

In this Subsection we determine the matrices $\mathbf{F}_i$ in (5) for implementing Algorithm 1. In ZFB, $\bar{\mathbf{F}}_i$ of (4) can be viewed as the right-inverse of $\mathbf{H}_{i,i}^H$:

$$\bar{\mathbf{F}}_i = \mathsf{Row}[\bar{\mathbf{f}}_{i,k}]_{k\in\mathcal{K}} = \mathbf{H}_{i,i}(\mathbf{H}_{i,i}^H\mathbf{H}_{i,i})^{-1}, \tag{22}$$

which exists only when $\mathbf{H}_{i,i}^H\mathbf{H}_{i,i}$ is nonsingular, requiring $N > N_{UE}$. It can be seen that we have

$$\begin{aligned}\boldsymbol{\beta}_{i,i}\mathbf{H}_{i,i}^H\mathbf{F}_i &= \boldsymbol{\beta}_{i,i}\mathbf{H}_{i,i}^H\mathbf{H}_{i,i}(\mathbf{H}_{i,i}^H\mathbf{H}_{i,i})^{-1}\mathrm{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} \\ &= \mathrm{diag}[\sqrt{\beta_{i,i,k}}\sqrt{p_{i,k}}]_{k\in\mathcal{K}},\end{aligned} \tag{23}$$

and thus the inter-user interference $\sigma_{i,k}^{\mathsf{U}}(\mathbf{p}_i)$ in (3) is forced to zero. As such, $\alpha_{i,k}$ defined by (8) is $\beta_{i,i,k}$, while $\lambda_{i,k}(\mathbf{p})$ defined by (9) is

$$\lambda_{i,k}(\mathbf{p}) = \sigma_{i,k}^{\mathsf{C}}(\mathbf{p}), \tag{24}$$

with $\sigma_{i,k}^{\mathsf{C}}(\mathbf{p})$ defined from (7).

From (1) we also define $\mathbf{H}_{i,i}^w \triangleq [\mathbf{h}_{i,i,k}^w]_{k\in\mathcal{K}}$ so $\mathbf{H}_{i,i} = \boldsymbol{\Theta}_i^{1/2}\mathbf{H}_{i,i}^w$, and $\mathbf{H}_{i,i}^H\mathbf{H}_{i,i} = (\mathbf{H}_{i,i}^w)^H\boldsymbol{\Theta}_i\mathbf{H}_{i,i}^w$, which has a rank obeying $r_i < N$. This renders the matrix $(\mathbf{H}_{i,i}^w)^H\boldsymbol{\Theta}_i\mathbf{H}_{i,i}^w$ quickly ill-conditioned, as the number N_{UE} of served users increases. The conventional RZFB design of [12], [29], [30] relies on:

$$\bar{\mathbf{F}}_i \triangleq \boldsymbol{\beta}_{i,i}\mathbf{H}_{i,i}(\mathbf{H}_{i,i}^H(\boldsymbol{\beta}_{i,i})^2\mathbf{H}_{i,i} + \eta\mathbf{I}_{N_{UE}})^{-1} \tag{25}$$

for $\eta > 0$ for regularizing $\boldsymbol{\beta}_{i,i}\mathbf{H}_{i,i}^H$. However, in our recent work [13] we proposed to let

$$\bar{\mathbf{F}}_i \triangleq \mathbf{H}_{i,i}(\mathbf{H}_{i,i}^H\mathbf{H}_{i,i} + \eta\mathbf{I}_{N_{UE}})^{-1}, \tag{26}$$

for regularizing $\mathbf{H}_{i,i}^H$ only, which allows the RZFB to perform much better. The optimal value of η was also computed in [13]. For simplicity, we take

$$\eta = N_{UE}\sigma^2/P_i^{\max}. \tag{27}$$

Then, as shown by Appendix II,

$$\begin{aligned}\boldsymbol{\beta}_{i,i}\mathbf{H}_{i,i}^H\bar{\mathbf{F}}_i\mathrm{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} &= \\ \boldsymbol{\beta}_{i,i}\mathrm{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} - \eta\boldsymbol{\beta}_{i,i}\mathbf{G}_i(\eta)\mathrm{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}},\end{aligned} \tag{28}$$

where $\mathbf{G}_i(\eta)$ is a Hermitian symmetric positive definite matrix defined by

$$\begin{aligned}\mathbf{G}_i(\eta) &= \begin{bmatrix}\mathbf{g}_{i,1}\\ \ldots \\ \mathbf{g}_{i,2M}\end{bmatrix} = \begin{bmatrix} g_{i,1,1} & \ldots & g_{i,1,2M}\\ \ldots & \ldots & \ldots \\ g_{i,2M,1} & \ldots & g_{i,2M,2M}\end{bmatrix} \\ &\triangleq (\mathbf{H}_{i,i}^H\mathbf{H}_{i,i} + \eta\mathbf{I}_{N_{UE}})^{-1}.\end{aligned} \tag{29}$$

The inter-user interference $\sigma_{i,k}^{\mathsf{U}}(\mathbf{p}_i)$ defined by (6) is formulated as:

$$\sigma_{i,k}^{\mathsf{U}}(\mathbf{p}_i) = \eta^2\beta_{i,i,k}\sum_{\ell\in\mathcal{K}\setminus\{k\}}|g_{i,k,\ell}|^2p_{i,\ell}, \tag{30}$$

and the transmit power function defined by (11) is defined accordingly.

C. Cell-wide ZFB relying on closed-form computation

In cell-wide ZFB (CW-ZFB), the inter-cell interference (7) is not incorporated in the optimization formulations, hence the following function is used instead of the function $r_{i,k}(\mathbf{p})$ of the actual information throughput of UE (i,k):

$$\hat{r}_{i,k}(p_{i,k}) = \ln\left(1+\beta_{i,i,k}p_{i,k}/\sigma^2\right). \tag{31}$$

For simplicity of presentation, in this subsection only we use the notation

$$\beta_{i,i,k} \to \bar{\beta}_{i,k}. \tag{32}$$

Accordingly, CW-ZFB targets solving the following individual EE maximization problems for cells $i \in \mathcal{I}$,

$$\max_{\mathbf{p}_i} \ \frac{\sum_{k\in\mathcal{K}}\ln\left(1+\bar{\beta}_{i,k}p_{i,k}/\sigma^2\right)}{\pi_i(\mathbf{p}_i)} \quad \text{s.t.} \tag{33a}$$

$$\sum_{k\in\mathcal{K}}||\bar{\mathbf{f}}_{i,k}||^2p_{i,k} \le P_i^{\max} \tag{33b}$$

$$\ln\left(1+\bar{\beta}_{i,k}p_{i,k}/\sigma^2\right) \ge \hat{r}_{i,k},\ k\in\mathcal{K}, \tag{33c}$$

where $\hat{r}_{i,k}$ is set to $\hat{r}_{i,k} > \bar{r}_{i,k}$ for those users, who are located in the boundary areas between the cells, in order to compensate for the real performance loss expected, when taking into account the intercell-interference (7).[2] For other users we set $\hat{r}_{i,k} = \bar{r}_{i,k}$, because they do not suffer from any intercell-interference, hence their information throughput is a function of $\hat{r}_{i,k}(p_{i,k})$ as defined by (31). Note that the problem (33) is convex, since the objective function in (33a) is given by the ratio of concave and affine functions, while the constraints (33b)-(33c) are convex. Our previous work in [11] proposed the following treatment for (33). First, it follows from (33c) that

$$p_{i,k} \ge \bar{p}_{i,k} := \sigma^2(e^{\hat{r}_{i,k}}-1)/\bar{\beta}_{i,k},$$

By invoking the variable change of

$$p_{i,k} = \tilde{p}_{i,k} + \bar{p}_{i,k},$$

it becomes straightforward to solve (33) by a Dinkelbach-type algorithm, which seeks $t > 0$ such that the optimal solution of the following optimization problem is zero:

$$\max_{\tilde{\mathbf{p}}_i} \ \sum_{k\in\mathcal{K}}\ln\left(a_{i,k}+\bar{\beta}_{i,k}\tilde{p}_{i,k}/\sigma^2\right) - t\cdot\tilde{\pi}_i(\tilde{\mathbf{p}}_i) \quad \text{s.t.} \tag{34a}$$

$$\sum_{k\in\mathcal{K}}||\bar{\mathbf{f}}_{i,k}||^2\tilde{p}_{i,k} \le \bar{P}_i^{\max},\ \tilde{p}_{i,k}\ge 0,\ k\in\mathcal{K}, \tag{34b}$$

where $a_{i,k} = 1+\bar{\beta}_{i,k}\bar{p}_{i,k}/\sigma^2$, $\bar{P}_{i,\mathrm{cir}} = \alpha\sum_{k\in\mathcal{K}}||\bar{\mathbf{f}}_{i,k}||^2\bar{p}_{i,k} + P_{\mathrm{cir}}$, $P_{\mathrm{cir}} = NP_a + P_c$, $\bar{P}_i^{\max} = P_i^{\max} - \sum_{k\in\mathcal{K}}||\bar{\mathbf{f}}_{i,k}||^2\bar{p}_{i,k}$, $\tilde{\pi}_i(\tilde{\mathbf{p}}_i) \triangleq \alpha\sum_{k\in\mathcal{K}}||\bar{\mathbf{f}}_{i,k}||^2\tilde{p}_{i,k} + \bar{P}_{i,\mathrm{cir}}$.
For $t > 0$ fixed, problem (34) admits finding the optimal solution in the following closed-form:

$$\tilde{p}_{i,k}^* = \left[\frac{1}{||\bar{\mathbf{f}}_{i,k}||^2(t\alpha+\lambda)} - \frac{a_{i,k}\sigma^2}{\bar{\beta}_{i,k}}\right]^+, k\in\mathcal{K}. \tag{35}$$

[2]In our simulations, $\hat{r}_{i,k}$ is 1.4 bps/Hz for $\bar{r}_{i,k} = 1$ bps/Hz and $\hat{r}_{i,k}$ is 2.2 bps/Hz for $\bar{r}_{i,k} = 2$ bps/Hz

Here and after, $[x]^+ = \max\{0, x\}$ and $\lambda = 0$ whenever we have:

$$\sum_{k\in\mathcal{K}} \left[\frac{1}{||\bar{\mathbf{f}}_{i,k}||^2 t\alpha} - \frac{a_{i,k}\sigma^2}{\bar{\beta}_{i,k}}\right]^+ \leq \bar{P}_i^{\max}.$$

Otherwise, $\lambda > 0$ is such that

$$\sum_{k\in\mathcal{K}} \left[\frac{1}{||\bar{\mathbf{f}}_{i,k}||^2 (t\alpha+\lambda)} - \frac{a_{i,k}\sigma^2}{\bar{\beta}_{i,k}}\right]^+ = \bar{P}_i^{\max}, \qquad (36)$$

which can be readily located by the classic bisection search.

However, in contrast to [11], which indeed uses bisection for locating the optimal t, we now propose a path-following Dinkelbach-type computational procedure for (33) as follows:

Algorithm 2 : CW-ZFB path-following Dinkelbach-type algorithm

1: **Initialization**: Solve (34) for $t = 0$. Let $\tilde{\mathbf{p}}_i^{(opt)}$ be its optimal solution. Set

$$\bar{t} = \sum_{k\in\mathcal{K}} \ln\left(a_{i,k} + \bar{\beta}_{i,k}\tilde{p}_{i,k}^{(opt)}/\sigma^2\right)/\tilde{\pi}_i(\tilde{\mathbf{p}}_i^{(opt)}).$$

2: **Solve (34) for $t = \bar{t}$ until its optimal value becomes zero.** Let $\tilde{\mathbf{p}}_i^{(opt)}$ be its optimal solution. Reset $\bar{t} = \sum_{k\in\mathcal{K}} \ln\left(a_{i,k} + \bar{\beta}_{i,k}\tilde{p}_{i,k}^{(opt)}/\sigma^2\right)/\tilde{\pi}_i(\tilde{\mathbf{p}}_i^{(opt)})$.

Like Algorithm 1, this algorithm converges to a locally optimal solution of (33), satisfying the KKT conditions of optimality.

Indeed one can still reply on Algorithm 1 for computing (33). However, the above procedure is much simpler than Algorithm 1, since the solution within each iteration obeys the closed-form formula (35). One can also consider CW-RZFB, which does not incorporate the inter-cell interference (7) into the optimization formulation, but its design still has to employ Algorithm 1. Moreover, it cannot perform better than the RZFB of the previous subsection, since the former incorporates the inter-cell interference (7) in the optimization formulation (13).

III. Mini-slot zero-forcing and regularized zero-forcing beamforming

It can be seen from (7) that the distant UEs that are located in the boundary areas between the cells, suffer not only from poor channel conditions but also from severe inter-cell interference, which cannot be forced to zero or mitigated.[3] Equally importantly, mitigating the inter-cell interference defined by (7) requires the knowledge of the multi-cell channel matrices $\mathbf{H}_{j,i}^H$ for $j \neq i$ in (2), which is not easily obtained. To tackle these issues of inter-cell interference, we first arrange for user-scheduling as follows. The network serves both the nearer and farther UEs within the same time slot, where the latter are located in the boundary areas between the first, second and the third cells, all of which are marked by red or blue colour in Fig. 1. As a benefit, the inter-cell interference between the second and third cells in (3) is efficiently weak and thus can be neglected, since there are no users between these cells to be served. [4] Those distant UEs, which are located in the boundary areas between the second and the third cells are marked by yellow color in Fig. 1, which are served over a different frequency band or time slot. The reader is referred to [31], [32] and [33] for further details on the employment of similar principles in simultaneous information and energy transfer and two-way communications. It was also shown in [34] that MS-TBF is capable of achieving a higher UE throughput than non-orthogonal multi-access (NOMA) in small cells.

Among the UEs marked by red or blue colour in Fig. 1, which are served by the network within the same time slot, let us assume that the UEs (i,k), $k \in \mathcal{K}_{\mathsf{ne}} \triangleq \{1, \ldots N_{\mathsf{ne}}\}$ are located close to their BS i, while UE (i,k), $k \in \mathcal{K}_{\mathsf{fa}} \triangleq \{N_{\mathsf{ne}}+1, \ldots, N_{UE}\}$ are located farther from their BS i. Thus in each cell there are N_{ne} nearer UEs and $N_{\mathsf{fa}} \triangleq N_{UE} - N_{\mathsf{ne}}$ farther UEs. We propose the following scheme relying on a pair of separate distinct transmissions within a time slot. During time-fraction $0 \leq \tau_1 \leq 1$, BS 1 transmits its DL signal to serve its nearer UEs, while BS 2 and BS 3 transmit signals to serve their farther UEs. These users are marked by red colour. During the remaining time-fraction $\tau_2 = 1 - \tau_1$, BS 1 transmits its DL signal to serve its farther UEs, while BS 2 and BS 3 transmit signals to serve their nearer UEs. These users are marked by blue colour. Under this MS scheme, the farther UEs are almost free from inter-cell interference because they are served by their BS when the neighbouring BSs serve their nearer UEs and thus need a low transmission power that imposes no interference on other cells. Therefore, this MS scheme does not need any information on the multi-cell interfering channel matrices $\mathbf{H}_{j,i}^H$ for $j \neq i$ in (2) and it only uses the information as the individual serving cells do. More importantly, this MS scheme allows the individual BSs to serve much larger numbers of UEs within each time slot. In short, the MS scheme is capable of simultaneously mitigating both the inter-user interference and the multi-cell interference, whilst maintaining to provide the desired users' MTp target.

Let us now denote by $\mathcal{K}_{i,1}$ and $\mathcal{K}_{i,2}$ the set of those UEs in cell i, which are served during the time-fraction τ_1 and τ_2, respectively. Under the proposed scheme, we have

$$\mathcal{K}_{1,1} = \mathcal{K}_{ne}, \mathcal{K}_{1,2} = \mathcal{K}_{fa},$$
$$\mathcal{K}_{i,1} = \mathcal{K}_{fa}, \mathcal{K}_{i,2} = \mathcal{K}_{ne}, i = 2, 3.$$

Let us now stipulate the following definitions:

$$\begin{gathered}\boldsymbol{\tau} \triangleq (\tau_1, \tau_2), \mathbf{s}_i^{[q]} \triangleq \mathsf{Col}[s_{i,k}]_{k\in\mathcal{K}_{i,q}}, \mathbf{y}_i^{[q]} \triangleq \mathsf{Col}[y_{i,k}]_{k\in\mathcal{K}_{i,q}},\\ \mathbf{p}_i^{[q]} \triangleq \mathsf{Col}[p_{i,k}]_{k\in\mathcal{K}_{i,q}}, \mathbf{p}^{[q]} = [\mathbf{p}_i^{[q]}]_{i\in\mathcal{I}},\\ \mathbf{n}_i^{[q]} = \mathsf{Row}[n_{i,k}]_{k\in\mathcal{K}_{i,q}}, q = 1, 2; i \in \mathcal{I},\\ (\mathbf{H}_{j,i}^{[q]})^H \triangleq \mathsf{Col}[\mathbf{h}_{j,i,k}^H]_{k\in\mathcal{K}_{i,q}}.\end{gathered} \qquad (37)$$

As mentioned before, the inter-cell interference is weak in this MS-TBF scheme and thus can be ignored. The MIMO

[3] Those UEs, who are located far from their BS but are not in the boundary areas between the cells, suffer only poor channel condition but are free from inter-cell interference

[4] Our simulations have confirmed that the performance remains unaffected by these interference sources.

equation of signal reception in time-fraction τ_q is thus

$$\mathbf{y}_i^{[q]} = \boldsymbol{\beta}_{i,i}(\mathbf{H}_{i,i}^{[q]})^H \mathbf{F}_i^{[q]} \mathbf{s}_i^{[q]} + \mathbf{n}_i^{[q]}. \quad (38)$$

We seek $\mathbf{F}_i^{[q]}$ in the class of

$$\mathbf{F}_i^{[q]} = \bar{\mathbf{F}}_i^{[q]} \text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}_{i,q}} \quad (39)$$

with predetermined $\bar{\mathbf{F}}_i^{[q]} \in \mathbf{C}^{N\times M} = \text{Row}[\bar{\mathbf{f}}_{i,k}]_{k\in\mathcal{K}_{i,q}}$. For convenience, $M = \mathcal{K}_{i,q},\ \forall i, q$.

The inter-user interference in the time-fraction τ_q is defined by the following affine function

$$\sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) = \beta_{i,i,k} \sum_{\ell\in\mathcal{K}_{i,q}\setminus\{k\}} |\mathbf{h}_{i,i,k}^H \bar{\mathbf{f}}_{i,\ell}|^2 p_{i,\ell}, \ell \in \mathcal{K}_{i,q}. \quad (40)$$

The information throughput at UE (i,k), $k \in \mathcal{K}_{i,q}$ is $\tau_q r_{i,k}^{[q]}(\mathbf{p}_i^{[q]})$ in conjunction with

$$\begin{aligned} r_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) &\triangleq \ln\left(1 + \frac{\beta_{i,i,k}|\mathbf{h}_{i,i,k}^H \bar{\mathbf{f}}_{i,k}|^2 p_{i,k}}{\sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) + \sigma^2}\right) \\ &= \ln\left(1 + \frac{\alpha_{i,k} p_{i,k}}{\sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) + \sigma^2}\right) \end{aligned} \quad (41)$$

for

$$\alpha_{i,k} \triangleq \beta_{i,i,k}|\mathbf{h}_{i,i,k}^H \bar{\mathbf{f}}_{i,k}|^2. \quad (42)$$

The transmit beamforming power during time-fraction τ_q of each cell is $\tau_q \chi_i^{[q]}(\mathbf{p}_i^{[q]})$ with

$$\chi_i^{[q]}(\mathbf{p}_i^{[q]}) \triangleq \sum_{k\in\mathcal{K}_{i,q}} ||\bar{\mathbf{f}}_{i,k}||^2 p_{i,k}, \quad (43)$$

which must satisfy the power constraint

$$\sum_{q=1}^{2} \tau_q \chi_i^{[q]}(\mathbf{p}_i^{[q]}) \le P_i^{\max}, i \in \mathcal{I}. \quad (44)$$

We also impose additionally the following physical constraints

$$||\bar{\mathbf{f}}_{i,k}||^2 p_{i,k} \le 3P_i^{\max}, (i,k) \in \mathcal{I}\times\mathcal{K} \quad (45)$$

to substantiate the fact that it is not possible to transmit an arbitrarily high power during the time-fractions.

The entire power consumption of the DL transmission is expressed by

$$\pi(\boldsymbol{\tau}, \mathbf{p}) = \sum_{i\in\mathcal{I}} (\alpha \sum_{q=1}^{2} \tau_q \chi_i^{[q]}(\mathbf{p}_i^{[q]}) + P_{\text{cir}}). \quad (46)$$

The EE maximization problem under MTp and power budget constraints is now formulated as

$$\begin{aligned} \max_{\boldsymbol{\tau},\mathbf{p}} \quad & \frac{\sum_{q=1}^{2} \tau_q \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}_{i,q}} r_{i,k}^{[q]}(\mathbf{p}_i^{[q]})}{\pi(\boldsymbol{\tau}, \mathbf{p})} \quad \text{s.t.} && (47a) \\ & (44), (45), && (47b) \\ & \tau_q r_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) \ge \overline{r}_{i,k},\ i \in \mathcal{I}, k \in \mathcal{K}_{i,q}, q = 1,2, && (47c) \\ & \tau_1 \ge 0, \tau_2 \ge 0, \tau_1 + \tau_2 \le 1, && (47d) \end{aligned}$$

which is a challenging nonconvex problem, since its objective function is nonconcave and the constraints (44) and (47c) are nonconvex.

To address (47), let us introduce the new variable

$$\boldsymbol{\theta} = (\theta_1, \theta_2), \quad (48)$$

which satisfies the convex constraints

$$\tau_1\theta_1 \ge 1, (1-\tau_1)\theta_2 \ge 1, \theta_1 > 0, \theta_2 > 0, 0 < \tau_1 < 1. \quad (49)$$

The power constraint (44) now becomes:

$$\Pi_i(\theta_2, \mathbf{p}_i) \triangleq (1 - \frac{1}{\theta_2})\chi_i^{[1]}(\mathbf{p}_i^{[1]}) + \frac{\chi_i^{[2]}(\mathbf{p}_i^{[2]})}{\theta_2} \le P_i^{\max}, \quad (50)$$

while the problem (47) is now expressed by

$$\begin{aligned} \max_{\tau_1,\boldsymbol{\theta},\mathbf{p}} \quad & \Phi(\boldsymbol{\theta}, \mathbf{p})/\Pi(\theta_2, \mathbf{p}) \quad \text{s.t.} \quad (45), (49), (50), && (51a) \\ & r_{i,k}^{[q]}(\mathbf{p}_i^{[q]})/\theta_q \ge \bar{r}_{i,k}, q = 1,2; i \in \mathcal{I}; k \in \mathcal{K}_{i,q}, && (51b) \end{aligned}$$

where

$$\Phi(\boldsymbol{\theta}, \mathbf{p}) \triangleq \sum_{q=1}^{2} \frac{1}{\theta_q} \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}_{i,q}} r_{i,k}^{[q]}(\mathbf{p}_i^{[q]})$$

and

$$\Pi(\theta_2, \mathbf{p}) = \sum_{i\in\mathcal{I}} (\alpha \cdot \Pi_i(\theta_2, \mathbf{p}_i) + P_{\text{cir}}).$$

The maximization problem (51) is nonconvex, since its objective function $\Phi(\theta, \mathbf{p})/\Pi(\theta_2, \mathbf{p})$ is nonconcave, while the constraints (50) and (51b) are nonconvex. To obtain a path-following algorithm for its solution, we have to iteratively approximate both the objective function $\Phi(\theta, \mathbf{p})/\Pi(\theta_2, \mathbf{p})$ and the function $r_{i,k}^{[q]}(\mathbf{p}_i^{[q]})/\theta_q$ in (51b) by lower-bounding concave functions and iteratively approximating the function $\Pi_i(\theta_2, \mathbf{p}_i)$ in (50) by an upper-bounding convex function. Such approximations for $\Pi_i(\theta_2, \mathbf{p}_i)$ and $r_{i,k}^{[q]}(\mathbf{p}_i^{[q]})/\theta_q$ will provide inner convex approximations for the nonconvex constraints (50) and (51b), respectively.

Let $(\tau_1^{(n)}, \boldsymbol{\theta}^{(n)}, \mathbf{p}^{(n)})$ be a feasible point for (51) found from the $(n-1)$st iteration and

$$t^{(n)} = \Phi(\boldsymbol{\theta}^{(n)}, \mathbf{p}^{(n)})/\Pi(\theta_2^{(n)}, \mathbf{p}^{(n)}).$$

Appendix III proves that

$$\Pi_i(\theta_2, \mathbf{p}_i) \le \Pi_i^{(n)}(\theta_2, \mathbf{p}_i) \quad (52)$$

for

$$\begin{aligned} &\Pi_i^{(n)}(\theta_2, \mathbf{p}_i) \triangleq \\ &\chi_i^{[1]}(\mathbf{p}_i^{[1]}) + \sum_{k\in\mathcal{K}_{i,2}} ||\bar{\mathbf{f}}_{i,k}||^2 \frac{1}{2}\left(\frac{p_{i,k}^2}{p_{i,k}^{(n)}\theta_2^{(n)}} + p_{i,k}^{(n)}\frac{\theta_2^{(n)}}{\theta_2^2}\right) \\ &+ \sum_{k\in\mathcal{K}_{i,1}} ||\bar{\mathbf{f}}_{i,k}||^2 \left(\frac{p_{i,k}^{(n)}}{p_{i,k}} + \frac{\theta_2}{\theta_2^{(n)}} - 3\right)\frac{p_{i,k}^{(n)}}{\theta_2^{(n)}}, \end{aligned} \quad (53)$$

which is a convex function.

Therefore, the nonconvex constraint (50) is innerly approximated by the convex constraint

$$\Pi_i^{(n)}(\theta_2, \mathbf{p}_i) \le P_i^{\max}, i \in \mathcal{I}. \quad (54)$$

To innerly approximate the nonconvex constraint (51b) in (51), we apply the inequality (76) in Appendix I for $x = \alpha_{i,k}p_{i,k}$,

$y = \sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) + \sigma^2$, and $\bar{x} = \alpha_{i,k} p_{i,k}^{(n)}$, $\bar{y} = \sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2$ to obtain

$$r_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) \geq r_{i,k}^{q,(n)}(\mathbf{p}_i^{[q]}) \tag{55}$$

for

$$r_{i,k}^{q,(n)}(\mathbf{p}_i^{[q]}) = \tilde{a}_{i,k}^{(n)} - \tilde{b}_{i,k}^{(n)}/(\alpha_{i,k} p_{i,k}) - \tilde{c}_{i,k}^{(n)}(\sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) + \sigma^2), \tag{56}$$

where

$$\begin{aligned} 0 < \tilde{a}_{i,k}^{(n)} &\triangleq r_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \frac{2\alpha_{i,k} p_{i,k}^{(n)}}{\sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2 + \alpha_{i,k} p_{i,k}^{(n)}}, \\ 0 < \tilde{b}_{i,k}^{(n)} &\triangleq \frac{(\alpha_{i,k} p_{i,k}^{(n)})^2}{\sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2 + \alpha_{i,k} p_{i,k}^{(n)}}, \\ 0 < \tilde{c}_{i,k}^{(n)} &\triangleq \frac{\alpha_{i,k} p_{i,k}^{(n)}}{\sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2 + \alpha_{i,k} p_{i,k}^{(n)}} (\sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2). \end{aligned} \tag{57}$$

The nonconvex constraint (51b) is thus innerly approximated by the convex constraint:

$$r_{i,k}^{q,(n)}(\mathbf{p}_i^{[q]}) \geq \theta_q \overline{r}_{i,k}\,,\ q = 1, 2; i \in \mathcal{I}, k \in \mathcal{K}_{i,q}. \tag{58}$$

Next, we consider the terms in the numerator of the objective function in (51a). By using the inequality (73) in Appendix I for $x = \alpha_{i,k} p_{i,k}$, $y = \sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) + \sigma^2$, $t\theta_q$, and $\bar{x} = \alpha_{i,k} p_{i,k}^{(n)}$, $\bar{y} = \sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2$, $\bar{t} = \theta_q^{(n)}$, we arrive at:

$$r_{i,k}^{[q]}(\mathbf{p}_i^{[q]})/\theta_q \geq g_{i,k}^{q,(n)}(\theta_q, \mathbf{p}), \tag{59}$$

where we have

$$\begin{aligned} &g_{i,k}^{q,(n)}(\theta_q, \mathbf{p}) \triangleq \\ &a_{i,k}^{(n)} - b_{i,k}^{(n)}/(p_{i,k}\alpha_{i,k}) - c_{i,k}^{(n)}(\sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) + \sigma^2) - d_{i,k}^{(n)}\theta_q \end{aligned} \tag{60}$$

with

$$\begin{aligned} 0 < \hat{a}_{i,k}^{(n)} &\triangleq 2r_{i,k}(\mathbf{p}_i^{q,(n)})/\theta_q^{(n)} + \\ &\quad \frac{2\alpha_{i,k} p_{i,k}^{(n)}}{\left(\sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2 + \alpha_{i,k} p_{i,k}^{(n)}\right)\theta_q^{(n)}}, \\ 0 < \hat{b}_{i,k}^{(n)} &\triangleq \frac{(\alpha_{i,k} p_{i,k}^{(n)})^2}{\left(\sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2 + \alpha_{i,k} p_{i,k}^{(n)}\right)\theta_q^{(n)}}, \\ 0 < \hat{c}_{i,k}^{(n)} &\triangleq \alpha_{i,k} p_{i,k}^{(n)} / \Big((\sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2 + \alpha_{i,k} p_{i,k}^{(n)}) \\ &\quad (\sigma_{i,k}^{[q]}(\mathbf{p}_i^{q,(n)}) + \sigma^2)\theta_q^{(n)} \Big), \\ 0 < \hat{d}_{i,k}^{(n)} &\triangleq r_{i,k}(\mathbf{p}_i^{q,(n)})/(\theta_q^{(n)})^2. \end{aligned} \tag{61}$$

At the nth iteration, the following convex program is solved to generate the next feasible point $(\tau^{(n+1)}, \theta^{(n+1)}, \mathbf{p}^{(n+1)})$ for (51):

$$\begin{aligned} \max_{\tau_1, \boldsymbol{\theta}, \mathbf{p}} \ & \sum_{q=1}^{2} \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}_{i,q}} g_{i,k}^{q,(n)}(\theta_q, \mathbf{p}) \\ & -t^{(n)} \sum_{i \in \mathcal{I}} \left(\alpha \cdot \Pi_i^{(n)}(\theta_2, \mathbf{p}_i) + P_{\text{cir}} \right) \\ & \text{s.t.} \quad (45), (49), (54), (58). \end{aligned} \tag{62}$$

The computational complexity of (62) is (18) with $\bar{n} = 3(N_{UE} + 1)$ and $\bar{m} = 3(2N_{UE} + 3)$.

In Algorithm 3, we propose a path-following computational procedure for the EE maximization problem (51).

To find an initial point $[\tau_1^{(0)}, \boldsymbol{\theta}^{(0)}, \mathbf{p}^{(0)}]$ for (51) we fix $(\tau_1^{(0)}, \boldsymbol{\theta}^{(0)}) = (0.5, 2, 2)$, which obviously satisfies (49), and solve the following linear programming problem:

$$\min_{\mathbf{p}} \ \tilde{\pi}(\mathbf{p}) \quad \text{s.t.} \quad \tilde{\pi}_i(\mathbf{p}_i) \leq P_i^{\max}, i \in \mathcal{I}, \tag{63a}$$

$$\begin{aligned} \alpha_{i,k} p_{i,k} \geq (e^{\theta_q^{(0)} \overline{r}_{i,k}} - 1)(\tilde{\sigma}_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) + \sigma^2), \\ q = 1, 2; i \in \mathcal{I}, k \in \mathcal{K}_{i,q}, \end{aligned} \tag{63b}$$

where

$$\begin{aligned} \tilde{\pi}_i(\mathbf{p}_i) &\triangleq (1 - \frac{1}{\theta_2^{(0)}}) \sum_{k \in \mathcal{K}_{i,1}} ||\bar{\mathbf{f}}_{i,k}||^2 p_{i,k} \\ &\quad + \sum_{k \in \mathcal{K}_{i,2}} ||\bar{\mathbf{f}}_{i,k}||^2 p_{i,k}, i \in \mathcal{I}, \\ \tilde{\pi}(\mathbf{p}) &\triangleq \sum_{i \in \mathcal{I}} \tilde{\pi}_i(\mathbf{p}_i), \end{aligned}$$

$$\tilde{\sigma}_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) \triangleq \beta_{i,i,k} \sum_{\ell \in \mathcal{K}_{i,q} \setminus \{k\}} |\mathbf{h}_{i,i,k}^H \bar{\mathbf{f}}_{i,\ell}|^2 p_{i,k}, k \in \mathcal{K}_{i,q},$$

which are linear functions. Note that the linear constraint (63b) represents the following QoS constraints

$$\begin{aligned} \frac{1}{\theta_q^{(0)}} \ln \left(1 + \frac{\alpha_i p_{i,k}}{\tilde{\sigma}_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) + \sigma^2} \right) \geq \overline{r}_{i,k}, \\ q = 1, 2; i \in \mathcal{I}, k \in \mathcal{K}_{i,q}. \end{aligned} \tag{64}$$

Suppose $\bar{\mathbf{p}}$ is the optimal solution of (63). Then an initial point $(\boldsymbol{\theta}^{(0)}, \mathbf{p}^{(0)})$ for (51) is $p_{i,k}^{(0)} = \bar{p}_{i,k}$.

Algorithm 3 : Path-following algorithm for solving problem (51)

1: **Initialization**: Solve (63) for $(\tau_1^{(0)}, \boldsymbol{\theta}^{(0)}) = (0.5, 2, 2)$ to take its optimal solution $\mathbf{p}^{(0)}$ to create a feasible point $(\tau_1^{(0)}, \boldsymbol{\theta}^{(0)}, \mathbf{p}^{(0)})$ for (51). Set $n := 0$ and $t^{(0)} := \Phi(\boldsymbol{\theta}^{(0)}, \mathbf{p}^{(0)})/\Pi(\theta_2^{(0)}, \mathbf{p}^{(0)})$.

2: **Repeat until convergence of the objective in (51)**: Solve the problem (62) for its optimal solution $(\tau_1^{(n+1)}, \boldsymbol{\theta}^{(n+1)}, \mathbf{p}^{(n+1)})$. Set $t^{(n+1)} := \Phi(\boldsymbol{\theta}^{(n+1)}, \mathbf{p}^{(n+1)})/\Pi(\theta_2^{(n+1)}, \mathbf{p}^{(n+1)})$. Set $n := n + 1$.

Similar to Algorithm 1, Algorithm 3 converges at least to a local solution of (51), satisfying the KKT conditions of optimality.

For MS-ZFB, $\bar{\mathbf{F}}_i^{[q]}$ in (39) is the right-inverse of the matrix $(\mathbf{H}_{i,i}^{[q]})^H$:

$$\bar{\mathbf{F}}_i^{[q]} = \mathsf{Row}[\bar{\mathbf{f}}_{i,k}]_{k \in \mathcal{K}_{i,q}} = \mathbf{H}_{i,i}^{[q]}((\mathbf{H}_{i,i}^{[q]})^H \mathbf{H}_{i,i}^{[q]})^{-1}, \tag{65}$$

under which the inter-user interference $\sigma_{i,k}^q(\mathbf{p}_i^{[q]})$ in (40) is forced to zero.

On the other hand, for MS-RZFB, $\bar{\mathbf{F}}_i^{[q]}$ in (39) is

$$\bar{\mathbf{F}}_i^q = \mathbf{H}_{i,i}^{[q]}((\mathbf{H}_{i,i}^{[q]})^H \mathbf{H}_{i,i}^{[q]} + \eta \mathbf{I}_M)^{-1}. \tag{66}$$

with

$$\eta = M\sigma^2/P_i^{\max}. \tag{67}$$

Then

$$\boldsymbol{\beta}_{i,i}(\mathbf{H}_{i,i}^{[q]})^H\mathbf{F}_i^{[q]} = \boldsymbol{\beta}_{i,i}\text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}_{i,q}} - \eta\boldsymbol{\beta}_{i,i}\mathbf{G}_i^{[q]}(\eta)\text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}_{i,q}}, \tag{68}$$

where $\mathbf{G}_i^{[q]}(\eta)$ is a Hermitian symmetric positive definite matrix defined by

$$\mathbf{G}_i^{[q]}(\eta) = \begin{bmatrix}\mathbf{g}_{i,1}^{[q]}\\ \ldots \\ \mathbf{g}_{i,M}^{[q]}\end{bmatrix} = \begin{bmatrix} g_{i,1,1} & \ldots & g_{i,1,M}\\ \ldots & \ldots & \ldots \\ g_{i,M,1} & \ldots & g_{i,M,M}\end{bmatrix} \triangleq ((\mathbf{H}_{i,i}^{[q]})^H\mathbf{H}_{i,i}^{[q]} + \eta\mathbf{I}_M)^{-1}. \tag{69}$$

In this case, $\alpha_{i,k}$ defined by (42) is

$$\alpha_{i,k} = \beta_{i,k}(1-\eta g_{i,k,k})^2,$$

while the inter-user interference $\sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]})$ in (40) is

$$\sigma_{i,k}^{[q]}(\mathbf{p}_i^{[q]}) \triangleq \eta^2\beta_{i,k}\sum_{\ell\in\mathcal{K}_{i,q}\setminus\{k\}}|g_{i,k,\ell}|^2p_{i,\ell}, k\in\mathcal{K}_{i,q}. \tag{70}$$

The transmit power function $\chi_i^{[q]}(\mathbf{p}_i^{[q]})$ defined by (43) is also represented as

$$\chi_i^{[q]}(\mathbf{p}_i^{[q]}) = \text{trace}\left(\mathbf{G}_i^{[q]}(\eta)(\mathbf{H}_{i,i}^{[q]})^H\mathbf{H}_{i,i}^{[q]}\mathbf{G}_i^{[q]}(\eta)\text{diag}[p_{i,k}]_{k\in\mathcal{K}_{i,q}}\right). \tag{71}$$

IV. Numerical Simulations

This section evaluates the performance of the proposed algorithms by numerical examples for different scenarios of single-cell, twin-cell and triple-cell networks. Unless otherwise stated, it is assumed that $N_{\mathsf{ne}} = N_{\mathsf{fa}} = N_{UE}/2$. The far UEs are uniformly distributed at the cell boundaries, while the near UEs are uniformly distributed near the BSs. Each of the BSs is located at the centre of a hexagon cell having a radius of 1 km and equipped with a uniform linear array (ULA) of antennas. The total number of antennas at each BS is $N = 64$.

To investigate the impact of the spatial correlation on the number of UEs as well as on the users' MTp that a massive MIMO scheme can support, we adopt the correlated Rayleigh fading model of [35, Sec. 2.6], where the covariance matrix is modeled by the Gaussian local scattering model of:

$$[\boldsymbol{\Theta}_{j,i,k}]_{p,q} = \frac{1}{N_{clus}}\sum_{n=1}^{N_{clus}} e^{\pi(p-q)\sin(\varphi_{j,i,k}^n)}e^{-\frac{\sigma_\varphi^2}{2}(\pi(p-q)\cos(\varphi_{j,i,k}^n))^2}, \tag{72}$$

where the number of scattered clusters is $N_{clus} = 6$, the nominal angle of arrival (AoA) for the nth cluster and the angular standard deviation (ASD) are $\varphi_{j,i,k}^n \sim \mathcal{U}[\varphi_{j,i,k} - 40^\circ, \varphi_{j,i,k} + 40^\circ]$ and $\sigma_\varphi = 10^\circ$, respectively [36]. Our assumption is that the antennas are spaced at half-wavelength to result in a form-factor of 0.25 m $\times$ 0.25 m [37].

Our simulation parameters used for generating the large-scale fading in Table I are similar to those used in [38]. The MTp threshold for all the users is set as $\bar{r}_{i,k} \equiv \bar{r} = 1$ bps/Hz. In the simulations, RZFB and TF-RZFB are referred to the new class of RZFB and TF-wise RZFB proposed in (26) and (66), while Conv. TF-RZF is referred to the conventional MS-RZFB defined in (25), whose weights are computed by adjusting Algorithms 1 and 3. Also CWZF is referred to CW-ZFB proposed in Subsection II.C.

TABLE I: Large scale fading Setup

Parameter	Numerical value
Carrier frequency / Bandwidth	2GHz / 20MHz
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-cell mMIMO scenario

A typical single-cell mMIMO arrangement is depicted by Fig. 2 with $N_{UE}/2$ near users and $N_{UE}/2$ far users, which are randomly located at nearer and farther distances from the BS. Under the MS-TBF, the BS transmits its DL signal to serve the nearer UEs during the time-fraction $0 < \tau_1 < 1$, while transmitting its signal to serve the farther UEs during the remaining time-fraction $\tau_2 = 1 - \tau_1$. Algorithms 1-3 can be readily adjusted for maximizing the EE of this single-cell mMIMO.

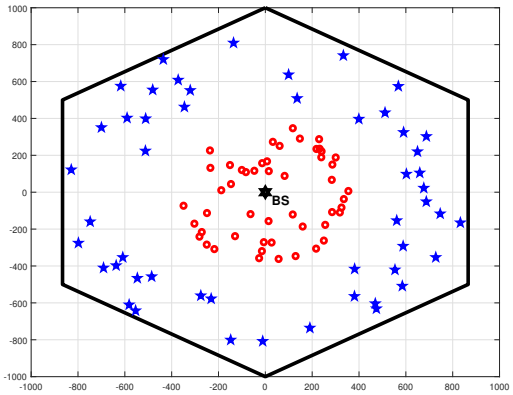


Fig. 2: Single-cell mMIMO: the users marked by the same colour are served over the same time-fraction of the time slot.

The typical convergence of the proposed Algorithm 1 used for RZFB, CW-ZFB Algorithm 2 for ZFB, and of Algorithm 3 for the MS-based ZFB and RZFB is provided in Fig. 3, where all of them are seen to converge within a few iterations. It is worth mentioning that Algorithm 2 converges much more rapidly than that proposed in [11], which is based on bisection for locating the optimal value of t in (34).

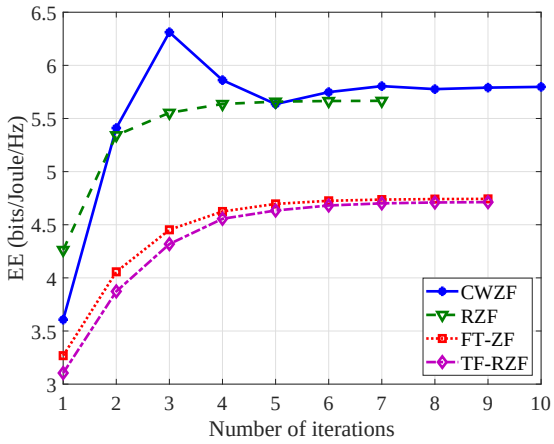
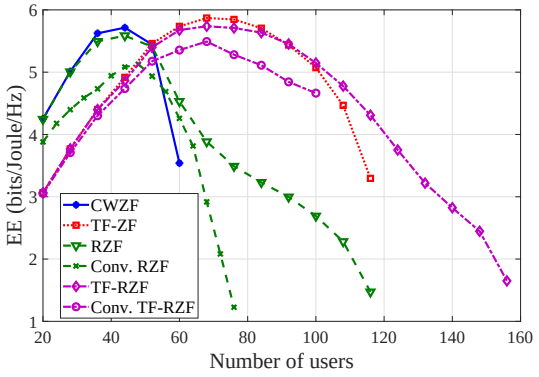
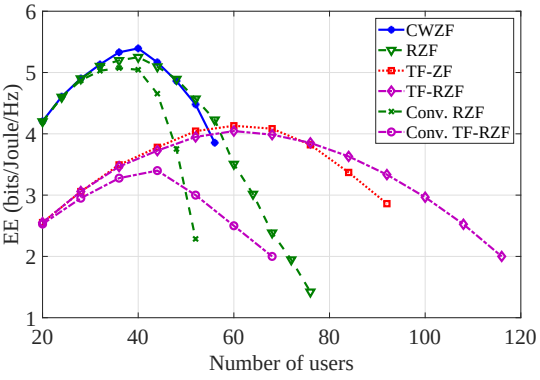


Fig. 3: The convergence of CWZF, RZF, TF-ZF and TF-RZF vs. the number of iterations under $N_{\mathsf{UE}} = 40$.



(a) $\bar{r} = 1$ bps/Hz



(b) $\bar{r} = 2$ bps/Hz

Fig. 4: The EE performance in the proposed methods vs. the number of users under different QoS requirements.

Fig. 4a and Fig. 4b portray the EE performance of the proposed beamforming approaches versus the number of users. RZFB is always capable of serving a much larger numbers of UEs than ZFB is. For the MTp threshold of $r = 1$ bps/Hz ($r = 2$ bps/Hz, resp.), CWZFB and MS-ZFB cannot serve more than 60 UEs (56 UEs, resp.) and 116 UEs (92 UEs, resp.). Meanwhile, both RZFB and MS-RZFB can serve up to 116 UEs (76 UEs, resp.) and 156 UEs (116 UEs, resp.) for $r = 1$ bps/Hz ($r = 2$ bps/Hz, resp.), but the latter clearly outperforms the former in terms of its EE. Note that when the number of served UEs is 72, this exceeds the number 64 of BS antennas. Furthermore, Conv. RZFB and Conv. TF-RZFB perform much worse than RZFB and TF-RZFB, respectively. Both the optimal MS allocation of the pair of separate transmission within the time slot and the optimal power allocation of the beamformers allow our massive MIMO scheme to serve more UEs than the number of transmit antennas.

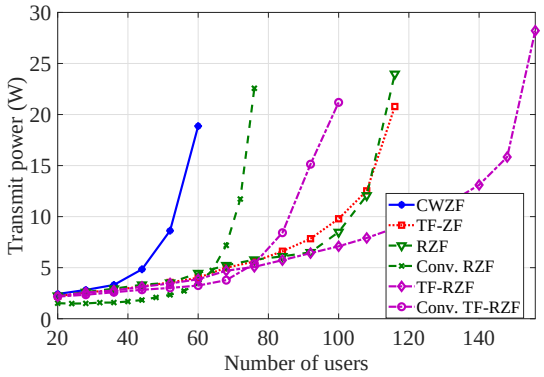
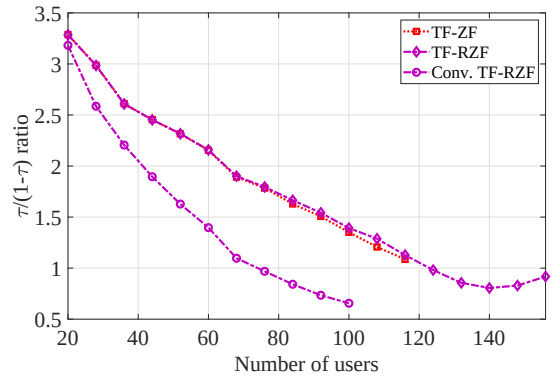


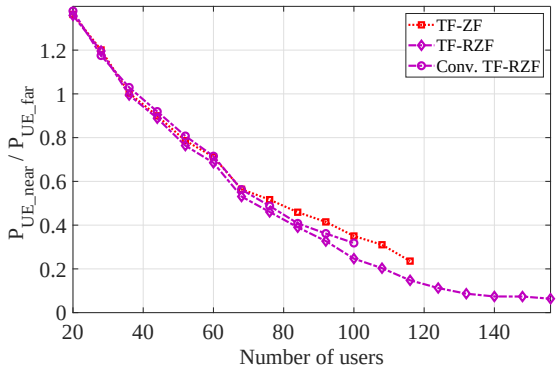
Fig. 5: The transmit power in in the proposed methods vs. the number of users under $\bar{r} = 1$ bps/Hz.

Furthermore, the EE first increases to a certain value of N_{UE} and then drops beyond that. Fig. 5 reveals that this drop is caused by the increased total transmit power. There is no 'magic' number N_{UE}, under which all the schemes attain their peak EE. Naturally, increasing the throughput threshold from 1 bps/Hz to 2 bps/Hz leads to decreasing number of UEs served and degrading the EE. Fig. 5 also shows that MS-TBF is capable of managing the power better than other beamforming schemes.

Fig. 6a and Fig. 6b plot the ratio between time-fractions dedicated to serving the near UEs and the far UEs as well as the corresponding power ratio, which are monotonically decreased with the total number N_{UE} of UEs for $r = 1$ bps/Hz. Recalling that $N_{\mathsf{ne}} = N_{\mathsf{fa}} = N_{UE}/2$ in our setting, at small N_{UE} / small N_{fa} a longer time-fraction and a higher power are allocated to the near UEs for maximizing their throughput. On the other hand, at high N_{UE} / large N_{fa}, longer time-fraction and power must be allocated to the far UEs for assuring their MTp.



(a) The value of $\tau/(1-\tau)$



(b) The total transmit power for near UEs and far UEs ratio

Fig. 6: The total transmit power for near UEs and far UEs ratio and $\tau/(1-\tau)$ value vs. the number of users under $\bar{r} = 1$ bps/Hz.

B. Twin-cell mMIMO scenario

A typical twin-cell mMIMO arrangement subject to severe multi-cell interference is depicted by Fig. 7. Under the MS-TBF, during time-fraction $0 \leq \tau \leq 1$, BS 1 serves its nearer UEs, while BS 2 serves its farther UEs. During the remaining fraction $(1-\tau)$, BS 1 serves its farther UEs, while BS 2 serves its nearer UEs. The farther UEs are thus free from inter-cell interference. Algorithms 1 and 3 can be readily adjusted for maximizing the EE of this twin-cell mMIMO scheme.

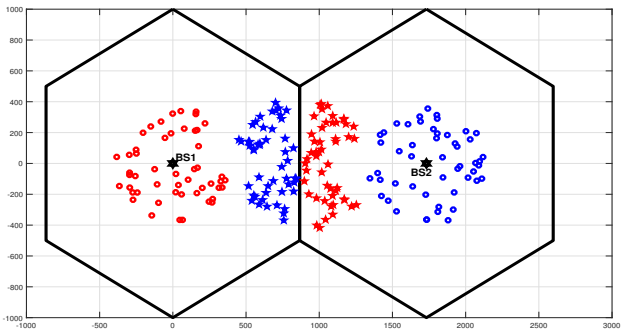
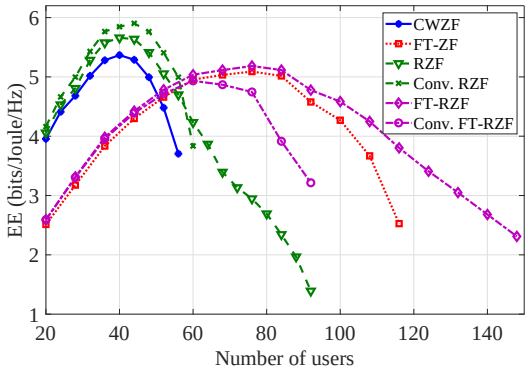
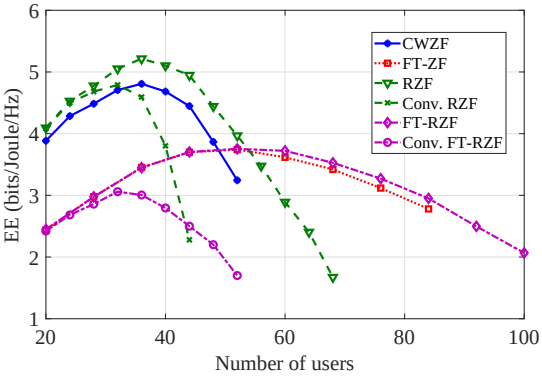


Fig. 7: Two-cell mMIMO: the users marked by the same colour are served over the same time-fraction of the time slot.

Fig. 8a and Fig. 8b show the superior performance of our MS-TBF schemes over others under different MTp thresholds. CWZFB cannot serve more than 56 UEs and 52 UEs, MS-ZFB serves up to 116 UEs and 84 UEs, respectively. Meanwhile, MS-RZFB can serve up to 148 UEs and 100 UEs, while significantly outperforming the RZFB in terms of the EE. It is observed that the EE erodes as the number N_{UE} of UEs increases. As expected, the conventional RZFB and TF-RZFB perform much worse than RZFB and TF-RZFB both in terms of the EE and the number of users served, which means that regularizing the channel matrix of small-scale fading in (26) is a much more efficient technique than regularizing the channel matrix incorporating both large-scale and small-scale fadings in (25).



(a) $\bar{r} = 1$ bps/Hz and $\hat{r} = 1.4$ bps/Hz

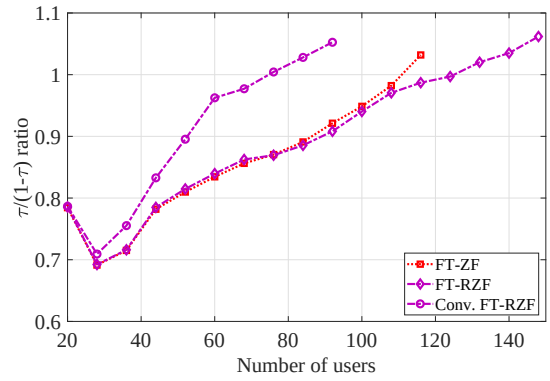


(b) $\bar{r} = 2$ bps/Hz and $\hat{r} = 2.4$ bps/Hz

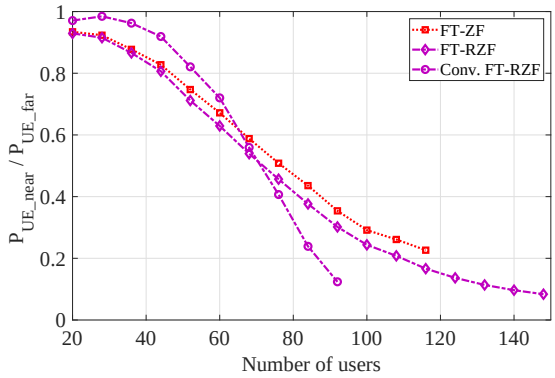
Fig. 8: The EE performance in the proposed methods vs. the number of users under different QoS requirements.

Interestingly, Fig. 9a and Fig. 9b show that the time-fraction allocation and power allocation of this twin-cell scenario are quite different from that of the single-cell scenario. The benefit of our time-fraction allocation is that the farther and nearer UEs are served in two different MSs. Meanwhile, the power allocation ratio trends are rather different for the small N_{UE} / small N_{fa} and large N_{UE} / large N_{fa} scenarios for the far UEs.

Remark. Serving those UEs, who are located at the boundary between the second and third cells which are marked with red colour in Fig. 1 substantially benefits from having no interference between the first cell of other two cells, so the MS scheme for two-cell mMIMO in the this subsection is applicable.

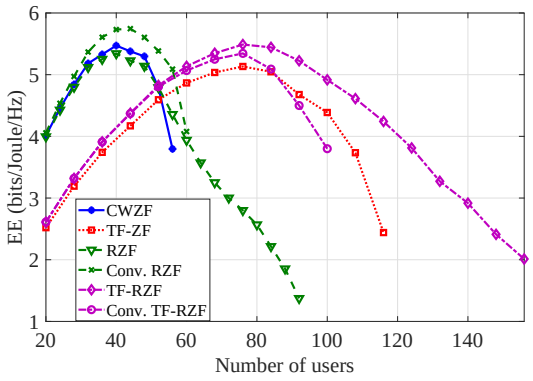


(a) The value of $\tau/(1-\tau)$

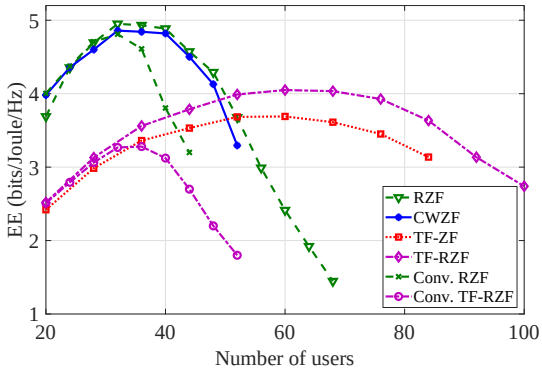


(b) The total transmit power for near UEs and far UEs ratio

Fig. 9: The total transmit power for near UEs and far UEs ratio and $\tau/(1-\tau)$ value vs. the number of users under $\bar{r} = 1$ bps/Hz.



(a) $\bar{r} = 1$ bps/Hz and $\hat{r} = 1.4$ bps/Hz



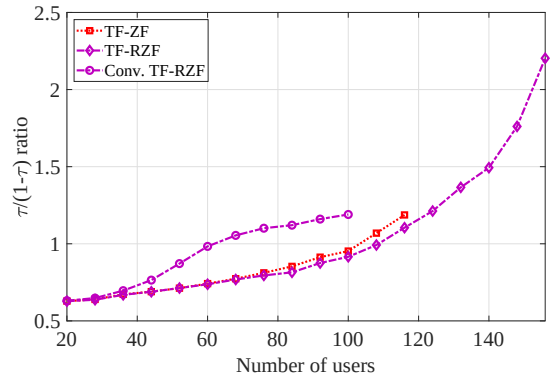
(b) $\bar{r} = 2$ bps/Hz and $\hat{r} = 2.4$ bps/Hz

Fig. 10: The EE performance in the proposed methods vs. the number of users under different QoS requirements.

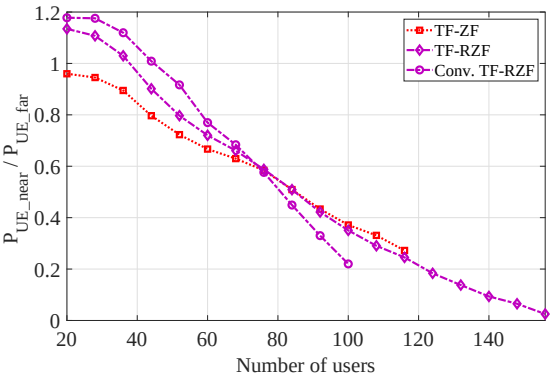
C. Triple-cell mMIMO scenario

We now return to the triple-cell mMIMO illustrated by Fig. 1. As a benefit of being free from inter-cell interference, our MS-TBF schemes can serve higher numbers of UEs at an increased EE, as Fig. 10a and Fig. 10b show. More particularly, for $r = 2$ bps/Hz, MS-ZFB and MS-RZFB are capable of serving up to 84 UEs and 100 UEs, respectively. Both RZFB and MS-RZFB can serve a larger UE population and achieve higher EE than CWZFB and MS-ZFB, since the former managed the power better than the latter in dealing with the inter-cell interference.

Fig. 11a and Fig. 11b plot the time-fraction ratio and power ratio, which are different from their counter parts in the single-cell and twin-cell scenarios. They are more or less balanced, because the far and near UEs are served in different time fractions. This fact dictates the allocation for both time-fractions and powers.



(a) The value of $\tau/(1-\tau)$



(b) The total transmit power for near UEs and far UEs ratio

Fig. 11: The total transmit power for near UEs and far UEs ratio and $\tau/(1-\tau)$ value vs. the number of users under $\bar{r} = 1$ bps/Hz.

V. Conclusions

We have considered the problem of maximizing the energy efficiency under MTp constraints for a large number of users by multi-cell mMIMO beamforming. To serve even larger numbers of users within a time slot, mini-slot-based TBF has been proposed, including new path-following computational procedures. Our simulations have demonstrated that an 8×8 antenna array aided massive MIMO is capable of serving up to 120 users at a given MTp target. Physical layer security of such multi-cell mMIMO schemes is under our current study.

Appendix I: fundamental inequalities

By noting that the function $f(x,y,t) = \frac{\ln(1+1/xy)}{t}$ is convex in $x>0, y>0, t>0$ [39], the following inequality holds true for all $x>0$, $\bar{x}>0$, $y>0$, $\bar{y}>0$, $t>0$, $\bar{t}>0$ [28]:

$$\begin{aligned}\frac{\ln(1+1/xy)}{t} &\geq f(\bar{x},\bar{y},\bar{t}) + \langle \nabla f(\bar{x},\bar{y},\bar{t}), (x,y,t)-(\bar{x},\bar{y},\bar{t})\rangle \\ &= \bar{a} - \bar{b}x - \bar{c}y - \bar{d}t, \end{aligned} \tag{73}$$

and

$$\ln(1+1/xy) \geq a - bx - cy, \tag{74}$$

where ∇ is the gradient operation and we have:

$$\bar{a} = 2\frac{\ln(1+1/\bar{x}\bar{y})}{\bar{t}} + \frac{2}{\bar{t}(\bar{x}\bar{y}+1)} > 0, \bar{b} = \frac{1}{(\bar{x}\bar{y}+1)\bar{x}\bar{t}} > 0,$$
$$\bar{c} = \frac{1}{(\bar{x}\bar{y}+1)\bar{y}\bar{t}} > 0, \bar{d} = \frac{\ln(1+1/\bar{x}\bar{y})}{\bar{t}^2} > 0,$$

and

$$a = \ln(1+1/\bar{x}\bar{y}) + 2/(\bar{x}\bar{y}+1) > 0,$$
$$b = 1/(\bar{x}\bar{y}+1)\bar{x} > 0, c = 1/(\bar{x}\bar{y}+1)\bar{y} > 0.$$

Replacing $x \to 1/x$ and $\bar{x} \to 1/\bar{x}$ in (73) and (74) leads to the following inequalities

$$\frac{\ln(1+x/y)}{t} \geq \hat{a} - \hat{b}/x - \hat{c}y - \hat{d}t, \tag{75}$$

and

$$\ln(1+x/y) \geq \tilde{a} - \tilde{b}/x - \tilde{c}y, \tag{76}$$

where we have:

$$\hat{a} = 2\frac{\ln(1+\bar{x}/\bar{y})}{\bar{t}} + \frac{2\bar{x}}{\bar{t}(\bar{y}+\bar{x})} > 0, \hat{b} = \frac{\bar{x}^2}{(\bar{y}+\bar{x})\bar{t}} > 0,$$
$$\hat{c} = \frac{\bar{x}}{(\bar{y}+\bar{x})\bar{y}\bar{t}} > 0, \hat{d} = \frac{\ln(1+\bar{x}/\bar{y})}{\bar{t}^2} > 0,$$

and

$$\tilde{a} = \ln(1+\bar{x}/\bar{y}) + 2\bar{x}/(\bar{x}+\bar{y}) > 0,$$
$$\tilde{b} = \bar{x}^2/(\bar{x}+\bar{y}) > 0, \tilde{c} = \bar{x}/(\bar{x}+\bar{y})\bar{y} > 0.$$

Observing that the function $f(z,t) = 1/zt$ is convex in $z>0, t>0$, we also have the following inequality

$$\begin{aligned}\frac{1}{zt} &\geq f(\bar{z},\bar{t}) + \langle \nabla f(\bar{z},\bar{t}), (z,t)-(\bar{z},\bar{t})\rangle \\ &= 3\frac{1}{\bar{z}\bar{t}} - \left(\frac{z/\bar{z}+t/\bar{t}}{\bar{z}\bar{t}}\right), \\ &\quad \forall\ x>0,\ \bar{x}>0, t>0, \bar{t}>0. \end{aligned} \tag{77}$$

Replacing $t \to 1/t$ and $\bar{t} \to 1/\bar{t}$ in (77) leads to the following inequality

$$\frac{t}{z} \geq 3\frac{\bar{t}}{\bar{z}} - \frac{z/\bar{z}+\bar{t}/t}{\bar{z}/\bar{t}}. \tag{78}$$

Appendix II: Derivation for (28)

From (26), we have

$$\begin{aligned}&\boldsymbol{\beta}_{i,i}\mathbf{H}_{i,i}^H\bar{\mathbf{F}}_i\text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} \\ =\ &\boldsymbol{\beta}_{i,i}\mathbf{H}_{i,i}^H\mathbf{H}_{i,i}(\mathbf{H}_{i,i}^H\mathbf{H}_{i,i}+\eta\mathbf{I}_{N_{UE}})^{-1}\text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} \\ =\ &\boldsymbol{\beta}_{i,i}\text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} \\ &-\eta\boldsymbol{\beta}_{i,i}(\mathbf{H}_{i,i}^H\mathbf{H}_{i,i}+\eta\mathbf{I}_{N_{UE}})^{-1}\text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} \\ =\ &\boldsymbol{\beta}_{i,i}\text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}} - \eta\boldsymbol{\beta}_{i,i}\mathbf{G}_i(\eta)\text{diag}[\sqrt{p_{i,k}}]_{k\in\mathcal{K}}, \end{aligned}$$

completing the proof of (28).

Appendix III: Proof for (52)

From (43) and (50), we arrive at:

$$\begin{aligned}\Pi_i(\theta_2,\mathbf{p}_i) &= \chi_i^{[1]}(\mathbf{p}_i^{[1]}) - \chi_i^{[1]}(\mathbf{p}_i^{[1]})/\theta_2 + \chi_i^{[2]}(\mathbf{p}_i^{[2]})/\theta_2 \\ &= \chi_i^{[1]}(\mathbf{p}_i^{[1]}) - \sum_{k\in\mathcal{K}_{i,1}} ||\bar{\mathbf{f}}_{i,k}||^2\frac{p_{i,k}}{\theta_2} \\ &\quad + \sum_{k\in\mathcal{K}_{i,2}} ||\bar{\mathbf{f}}_{i,k}||^2\frac{p_{i,k}}{\theta_2}. \end{aligned} \tag{79}$$

Using the inequality

$$\frac{t}{z} \leq \frac{1}{2}\left(\frac{t^2}{\bar{t}\bar{z}} + \frac{\bar{t}\bar{z}}{z^2}\right)\ \ \forall t>0, z>0 \quad \&\ \bar{t}>0, \bar{z}>0 \tag{80}$$

yields

$$\frac{p_{i,k}}{\theta_2} \leq \frac{1}{2}\left(\frac{p_{i,k}^2}{p_{i,k}^{(n)}\theta_2^{(n)}} + \frac{p_{i,k}^{(n)}\theta_2^{(n)}}{\theta_2^2}\right), \tag{81}$$

while using the inequality (75) yields

$$\frac{p_{i,k}}{\theta_2} \geq 3\frac{p_{i,k}^{(n)}}{\theta_2^{(n)}} - \frac{\theta_2/\theta_2^{(n)} - p_{i,k}^{(n)}/p_{i,k}}{\theta_2^{(n)}/p_{i,k}^{(n)}}. \tag{82}$$

Then $\Pi_i(\theta_2,\mathbf{p}_i)$ represented by (79) is bounded as

$$\begin{aligned}\Pi_i(\theta_2,\mathbf{p}_i) &\leq \chi_i^{[1]}(\mathbf{p}_i^{[1]}) \\ &\quad + \sum_{k\in\mathcal{K}_{i,1}} ||\bar{\mathbf{f}}_{i,k}||^2\left(\frac{p_{i,k}^{(n)}}{p_{i,k}} + \frac{\theta_2}{\theta_2^{(n)}} - 3\right)\frac{p_{i,k}^{(n)}}{\theta_2^{(n)}} \\ &\quad + \sum_{k\in\mathcal{K}_{i,2}} ||\bar{\mathbf{f}}_{i,k}||^2\frac{1}{2}\left(\frac{p_{i,k}^2}{p_{i,k}^{(n)}\theta_2^{(n)}} + p_{i,k}^{(n)}\frac{\theta_2^{(n)}}{\theta_2^2}\right), \end{aligned}$$

whose right hand side is the function $\Pi_i^{(n)}(\theta_2,\mathbf{p}_i)$ defined by (53). The proof of (52) is completed.

REFERENCES

[1] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan 2013.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, February 2014.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] R. Shafin and L. Liu, "Multi-cell multi-user massive FD-MIMO: Downlink precoding and throughput analysis," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 487–502, 2018.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in massive MIMO," in *22nd European Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, 2014, pp. 76–80.

[6] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications.* Delft, The Netherlands: Now Publishers Inc., 2004.

[7] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications,*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[8] 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, no. 7, pp. 4509–4537, Jul. 2012.

[9] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[10] Y. G. Lim, C. B. Chae, and G. Caire, "Performance analysis of massive MIMO for cell-boundary users," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6827–6842, Dec 2015.

[11] 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)*, Aug. 2017, p. 105.

[12] C. B. Peel, B. M. Hochwald, and A. L. Swindlerhurst, "A vector-perturbation technique for near capacity multiantenna multiuser communication- Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[13] L. D. Nguyen, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Multi-user regularized zero-forcing beamforming," *IEEE Trans. Signal Process.*, vol. 67, no. 11, pp. 2839–2853, Jun. 2019.

[14] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, April 2013.

[15] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

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

[17] 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, no. 4, pp. 697–709, Apr. 2016.

[18] A. Zappone, L. Sanguinetti, G. Bacci, E. A. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1668–1683, Apr. 2016.

[19] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 54, no. 3, pp. 528–541, Mar. 2006.

[20] B. M. Lee and H. Yang, "Massive MIMO for Industrial Internet of Things in cyber-physical systems," *IEEE Trans. Indust. Infor.*, vol. 14, no. 6, pp. 2641–2652, Jun. 2018.

[21] L. Liu and W. Yu, "Massive connectivity with massive MIMO— Part II: Achievable rate characterization," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2947–2959, Jun. 2018.

[22] L. Liu, E. G. Larsson, W. Yu, P. Popovski, C. Stefanovic, and E. de Carvalho, "Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 88–99, May 2018.

[23] A. Adhikary, H. S. Dhillon, and G. Caire, "Massive-MIMO meets HetNet: Interference coordination through spatial blanking," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1171–1186, June 2015.

[24] 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, Jul. 2017.

[25] D. Peaucelle, D. Henrion, and Y. Labit, "Users guide for SeDuMi interface 1.03," 2002. [Online]. Available: http://homepages.laas.fr/peaucell/software/sdmguide.pdf

[26] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operation Research*, vol. 26, no. 4, pp. 681–683, 1978.

[27] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local D.C. programming," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.

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

[29] V. Stankovic and M. Haardt, "Generalized design of multi-user MIMO precoding matrices," *IEEE Trans. Wirel. Commun.*, vol. 7, no. 3, pp. 953–961, Mar. 2008.

[30] H. Sung, S. Lee, and I. Lee, "Generalized channel inversion methods for multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3489–3499, Nov. 2009.

[31] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Secure and energy-efficient beamforming for simultaneous information and energy transfer," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 11, pp. 7523–7537, Nov. 2017.

[32] Z. Sheng, H. D. Tuan, T. Q. Duong, H. V. Poor, and Y. Fang, "Low-latency multiuser two-way wireless relaying for spectral and energy efficiencies," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4362–4376, Jul. 2018.

[33] A. A. Nasir, H. D. Tuan, and T. Q. Duong, "Fractional time exploitation for serving IoT users with guaranteed QoS by 5G spectrum," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 128–133, Oct. 2018.

[34] V.-D. Nguyen, H. D. Tuan, T. Q. Duong, O.-S. Shin, and H. V. Poor, "Joint fractional time allocation and beamforming for downlink multiuser MISO systems," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2650–2653, Dec. 2017.

[35] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[36] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Massive MIMO with spatially correlated Rician fading channels," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3234–3250, 2019.

[37] Y. Kim *et al.*, "Full-dimension MIMO (FD-MIMO): The next evolution of MIMO in LTE systems," *IEEE Wireless Commun.*, vol. 21, no. June, pp. 26–33, 2014.

[38] E. Bjornson, M. Kountouris, and M. Debbah, "Massive MIMO and small cells: Improving energy efficiency by optimal soft-cell coordination," in *ICT*, May 2013, pp. 1–5.

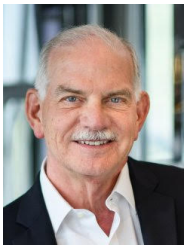
[39] Z. Sheng, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "Power allocation for energy efficiency and secrecy of interference wireless networks," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 6, pp. 3737 – 3751, Jun. 2018.

Long Dinh Nguyen 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 Univerisity 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 Department of Engineering in Dong Nai University in Vietnam as an Assistant Professor and Duy Tan University as an Adjunct Assistant Professor. His research interests include convex optimization theory and applications for modern wireless systems, resource allocation optimizations, and real-time optimization and machine learning for embedded systems, wireless communications and Internet of Things (IoTs). He is currently serving as a reviewer for IEEE Trans on Wireless Communications, IEEE Trans on Communications, IEEE Access, IEEE Communication Letter, IET Communications and server international conferences. He was awarded the Exemplary Reviewer Award in IEEE Communications Letters 2018.

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 September 2012. Currently, he is a Research Chair - Royal Academy of Engineering and a Professor with Queen's University Belfast (UK), where he was a Lecturer (Assistant Professor) (2013-2017), a Reader (Associate Professor) (2018-2020), and Full Professor from August 2020. His current research interests include wireless communications, machine learning, realtime optimisation, big data, and IoT applications to disaster management, air-quality monitoring, flood monitoring, smart agriculture, healthcare and smart cities. He is the author or co-author of over 350+ technical papers published in scientific journals (230+ articles) and presented at international conferences (140+ papers).

Dr. Duong currently serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, and an Executive 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 2019, IEEE Digital Signal Processing Conference (DSP) 2017, and International Wireless Communications & Mobile Computing Conference (IWCMC) 2019. He is the recipient of prestigious Royal Academy of Engineering Research Fellowship (2015-2020) 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, machine learning and network science, and their applications in wireless networks, energy systems and related fields. Among his publications in these areas is the recent book *Multiple Access Techniques for 5G Wireless Networks and Beyond*. (Springer, 2019).

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. Recent recognition of his work includes the 2017 IEEE Alexander Graham Bell Medal and a D.Eng. *honoris causa* from the University of Waterloo awarded in 2019.

Lajos Hanzo (http://www-mobile.ecs.soton.ac.uk, https://en.wikipedia.org/wiki/Lajos_Hanzo) (FIEEE'04, Fellow of the Royal Academy of Engineering F(REng), of the IET and of EURASIP), received his Master degree and Doctorate in 1976 and 1983, respectively from the Technical University (TU) of Budapest. He was also awarded the Doctor of Sciences (DSc) degree by the University of Southampton (2004) and Honorary Doctorates by the TU of Budapest (2009) and by the University of Edinburgh (2015). He is a Foreign Member of the Hungarian Academy of Sciences and a former Editor-in-Chief of the IEEE Press. He has served several terms as Governor of both IEEE ComSoc and of VTS. He has published 1900+ contributions at IEEE Xplore, 19 Wiley-IEEE Press books and has helped the fast-track career of 123 PhD students. Over 40 of them are Professors at various stages of their careers in academia and many of them are leading scientists in the wireless industry.