

“© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

# Beamforming and Power Allocation for Energy-Efficient Massive MIMO

Long D. Nguyen\*, Hoang D. Tuan<sup>†</sup>, Trung Q. Duong\* and H. Vincent Poor<sup>‡</sup>

\*Queen's University, Belfast, UK (e-mail: {Inguyen04, trung.q.duong}@qub.ac.uk)

<sup>†</sup>University of Technology Sydney, Sydney, Australia (e-mail: tuan.hoang@uts.edu.au)

<sup>‡</sup>Princeton University, Princeton, NJ, USA (e-mail: poor@princeton.edu)

**Abstract**—Massive multiple input multiple output (MIMO) has emerged as a promising technology, which utilizes a large number of antennas at base stations (BSs) to significantly improve the spectral efficiency in terms of bits/s/Hz while reducing the radiated signal power. A critical issue with massive MIMO is the costly circuit power consumption, which is proportional to the number of antennas. This paper develops low-complexity power allocation techniques to apply beamforming and to maximize the energy efficiency of massive MIMO while meeting users' quality-of-service requirements. Algorithms of low computational complexity with rapid convergence are proposed to solve for the optimal beamformer in this sense. Numerical examples are provided to show the merit of the proposed computational approach.

## I. INTRODUCTION

Massive multiple input multiple output (MIMO) [1], in which a large scale array of antennas is deployed at base stations (BSs) to enhance both spectral efficiency and radiated power efficiency, has emerged as a solution to help enable the dramatic increase in network capacity required to support emerging wireless networks. However, the increased amount of infrastructure needed for deploying such large numbers of antennas also results in a substantial growth in the circuit power consumption at BSs. Since the energy-efficiency (EE) performance in terms of bits/Joule/Hz is a very important figure-of-merit in 5G systems [2], [3], it is of great interest to employ signal processing techniques to improve the EE performance of massive MIMO. The objective of such improvement is to maximize the ratio of the sum throughput and consumed power. The consumed power in the denominator contains not only the radiated power, which can be well controlled, but also the circuit power consumption, which is proportional to the number of transmit antennas and thus constitutes a significant fraction of the total. Maximizing the system EE is thus different from minimizing the transmit power in guaranteeing users' quality-of-service (QoS) requirements, which was addressed in [4] using semi-definite programming (SDP). Moreover, the dimensionality of SDP increases dramatically for massive MIMO, making such an approach computationally very inefficient.

One of the most appealing aspects of massive MIMO is its favorable propagation characteristics [1], [5], which allow low-complexity beamforming such as zero forcing or conjugate beamforming to perform well [6]. However, this performance analysis has been considered under equi-power

allocated beamforming only [7]. Most works in this area (see e.g. [7] and references therein) have focused on the QoS, which equi-power beamforming can offer to the cell edge users, i.e., those suffering from poor channel conditions. On the other hand, optimizing power allocation (PA) in massive MIMO beamforming can be useful in offering better QoS. The main objective of this paper is to consider optimal PA for both zero forcing and conjugate beamforming to maximize the EE of massive MIMO while meeting users' QoS requirements. As a result, both the service quality for every user and EE for massive MIMO are simultaneously optimized in the proposed approach. Our main contributions are as follows:

- The PA problem of zero-forcing beamforming to maximize the EE is shown to be quasi-concave and solved very effectively, where every Dinkelbach's iteration of fractional programming [8] admits a closed-form solution;
- The PA problem of conjugate beamforming to maximize the EE is no longer concave/quasi-concave but is still efficiently solved by a path-following computational procedure of rapid convergence, which invokes a simple quadratic program of moderate dimension at each iteration;
- An interesting insight is that optimal power-allocated zero-forcing beamforming performs much better than its conjugate counterpart although according to [9], equi-power conjugate beamforming performs better in terms of EE compared with equi-power zero-forcing beamforming.

*Notation:* Boldface upper and lowercase letters denote matrices and vectors, respectively. The transposition and conjugate transposition of a matrix  $\mathbf{X}$  are respectively represented by  $\mathbf{X}^T$  and  $\mathbf{X}^H$ .  $\mathbf{I}_M$  denotes the identity matrix of size  $M \times M$ .  $\langle \mathbf{x}, \mathbf{y} \rangle$  for complex vectors  $\mathbf{x}$  and  $\mathbf{y}$  is their dot product  $\mathbf{x}^H \mathbf{y}$  and  $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$  is the norm of  $\mathbf{x}$ . A Gaussian random vector with mean  $\bar{\mathbf{x}}$  and covariance  $\mathbf{R}_{\mathbf{x}}$  is denoted by  $\mathbf{x} \sim \mathcal{CN}(\bar{\mathbf{x}}, \mathbf{R}_{\mathbf{x}})$ .  $[\mathbf{X}_1; \dots; \mathbf{X}_k]$  is a matrix created by stacking vertically  $\mathbf{X}_1, \dots, \mathbf{X}_k$ .

## II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider downlink communication from a base station (BS) equipped with a large-scale  $M$ -element antenna array with  $M$  up to several hundred to serve  $K$  ( $K \ll M$ ) users (UEs), each of which is equipped with a single antenna. Each

symbol  $x_k$  intended for UE  $k$  is beamformed by a vector  $\mathbf{f}_k \in \mathbb{C}^M$  before being transmitted by the BS. The transmitted signal at the BS is thus  $\sum_{k=1}^K \mathbf{f}_k x_k$ . Let  $\mathbf{g}_k \in \mathbb{C}^M$  be the channel vector between the BS and UE  $k$ , which is modelled by  $\mathbf{g}_k = \sqrt{\beta_k} \mathbf{h}_k$  [1], [10], where  $\sqrt{\beta_k}$  models the path loss and large-scale fading while  $\mathbf{h}_k = (h_{1k}, \dots, h_{Mk})^T$  with  $h_{mk} \in \mathcal{CN}(0, 1)$  represents the small-scale fading.

For notational convenience, denote  $\{1, \dots, K\}$  by  $\mathcal{K}$ . The received signal at UE  $k$  is given by

$$y_k = \sqrt{\beta_k} \mathbf{h}_k^T \mathbf{f}_k x_k + \sum_{i \in \mathcal{K} \setminus \{k\}} \sqrt{\beta_k} \mathbf{h}_k^T \mathbf{f}_i x_i + n_k, \quad (1)$$

where  $n_k \in \mathcal{CN}(0, \sigma_k^2)$ . The channel state information (CSI) is assumed known at the BS, which can be obtained from the UE's uplink pilots via time-division-duplexing (TDD) and uplink/downlink radio channel reciprocity [1], [11].

For  $\mathbf{f} = [\mathbf{f}_k]_{k \in \mathcal{K}}$ , the information throughput at UE  $k$  is

$$r_k(\mathbf{f}) = \log_2(1 + \text{SINR}_k(\mathbf{f})) \quad (2)$$

with the signal-to-interference-plus-noise ratio (SINR) given by

$$\text{SINR}_k(\mathbf{f}) = \frac{\bar{\beta}_k |\mathbf{h}_k^T \mathbf{f}_k|^2}{\bar{\beta}_k \sum_{i \in \mathcal{K} \setminus \{k\}} |\mathbf{h}_k^T \mathbf{f}_i|^2 + \sigma_k^2}.$$

The consumed power for the instantaneous downlink transmission can be expressed as

$$P_{\text{total}}(\mathbf{f}) = \alpha \sum_{k=1}^K \|\mathbf{f}_k\|^2 + MP_A + P_C, \quad (3)$$

where  $\alpha > 1$  is the reciprocal of the drain efficiency of the amplifier of the BS,  $P_A$  denotes the per-antenna circuit power and  $P_C$  represents the non-transmission power of the BS.

The EE maximization problem subject to users' QoS constraints and transmission power budget can be formulated as

$$\max_{\mathbf{f}=[\mathbf{f}_k]_{k \in \mathcal{K}}} \sum_{k=1}^K r_k(\mathbf{f}) / P_{\text{total}}(\mathbf{f}) \quad (4a)$$

$$\text{s.t. } r_k(\mathbf{f}) \geq \bar{r}_k, \forall k \in \mathcal{K}, \quad (4b)$$

$$\sum_{k=1}^K \|\mathbf{f}_k\|^2 \leq P^{\max}, \quad (4c)$$

where the constraint (4b) represents the QoS data rate requirement for each UE. The constraint (4c) enforces the power budget at the BS.

The objective in (4a), which is the ratio of the sum throughput and consumed power, expresses the energy-efficiency in terms of bits/Joule/Hz. The circuit power consumption  $MP_A$  in (3) is proportional to the number  $M$  of antennas and cannot be controlled. In general, (4) is a very difficult large-scale nonconvex optimization problem because the numerator in the objective function in (4a) is not a concave function of the beamforming vector  $\mathbf{f} \in \mathbb{C}^{MK}$  and  $M$  is large.

### III. PA FOR ZERO FORCING BEAMFORMING

Define a matrix  $\mathbf{H} = [\mathbf{h}_1^T, \dots, \mathbf{h}_K^T] \in \mathbb{C}^{K \times M}$ , which is very fat due to the fact that  $K \ll M$  and  $M$  is large. Accordingly, the square matrix  $\mathbf{H}\mathbf{H}^H \in \mathbb{C}^{K \times K}$  of much smaller size is very well-conditioned, whose eigenvalue distribution becomes more deterministic as  $M$  increases [12]. To exploit these favorable propagation characteristics in the massive MIMO transmission, we first seek a beamforming vector  $\mathbf{f}_k$  in the class of zero-forcing (ZF) beams as follows. Let

$$\bar{\mathbf{F}} = [\bar{\mathbf{f}}_1, \dots, \bar{\mathbf{f}}_K] = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1}, \quad (5)$$

so

$$\mathbf{I}_M = \mathbf{H}\bar{\mathbf{F}} = [\mathbf{h}_1^T \bar{\mathbf{F}}; \dots; \mathbf{h}_K^T \bar{\mathbf{F}}] = [\mathbf{h}_i^T \bar{\mathbf{f}}_j]_{(i,j) \in \mathcal{K} \times \mathcal{K}}. \quad (6)$$

We normalize  $\tilde{\mathbf{f}}_k = \bar{\mathbf{f}}_k / \|\bar{\mathbf{f}}_k\|$ ,  $k = 1, \dots, K$ , and seek  $\mathbf{f}_k$  in the set

$$\mathbf{f}_k = \sqrt{p_k} \tilde{\mathbf{f}}_k, k \in \mathcal{K}. \quad (7)$$

The equation (1) becomes

$$y_k = \sqrt{\beta_k} \sqrt{p_k} \mathbf{h}_k^T \tilde{\mathbf{f}}_k x_k + n_k, \quad (8)$$

where the multiple user (MU) interference in (1) has been cancelled thanks to the zero-forcing condition (6).

For  $\mathbf{p} = (p_1, \dots, p_K)^T$ , the information throughput for UE  $k$  defined by (2) becomes the following concave function of  $p_k$ :

$$r_k(p_k) = \log_2(1 + \bar{\beta}_k |\mathbf{h}_k^T \tilde{\mathbf{f}}_k|^2 p_k / \sigma_k^2), \quad (9)$$

while the consumed power for the BS transmission defined by (3) is now an affine function of  $\mathbf{p}$ :

$$P_{\text{total}}(\mathbf{p}) = \alpha \sum_{k=1}^K p_k + MP_A + P_C. \quad (10)$$

Therefore, the EE maximization problem (4) for zero-forcing beamforming can now be formulated in terms of PA optimization as

$$\max_{\mathbf{p}=(p_1, \dots, p_K)^T} \sum_{k=1}^K \log_2(1 + \beta_k p_k / \sigma_k^2) / P_{\text{total}}(\mathbf{p}) \quad (11a)$$

$$\text{s.t. } \log_2(1 + \beta_k p_k / \sigma_k^2) \geq \bar{r}_k, k = 1, \dots, K \quad (11b)$$

$$\sum_{k=1}^K p_k \leq P^{\max}, p_k \geq 0, k = 1, \dots, K, \quad (11c)$$

where  $\beta_k = \bar{\beta}_k |\mathbf{h}_k^T \tilde{\mathbf{f}}_k|^2$ .

As the objective function in (11a) is the ratio of concave and affine functions while the constraints (11b)-(11c) are convex, the problem (11) can be solved by Dinkelbach's algorithm of fractional programming [8], which finds the optimal value for  $\tau > 0$  such that zero is the optimal value of the convex program

$$\max_{\mathbf{p}} \sum_{k=1}^K \log_2(1 + \beta_k p_k / \sigma_k^2) - \tau P_{\text{total}}(\mathbf{p}) \text{ s.t. } (11b) - (11c). \quad (12)$$

However, being convex, the program (12) is still computationally difficult as it involves optimization of logarithmic functions. Our first contribution is to provide a closed-form solution for (12).

It follows from (11b) that  $p_k \geq \bar{p}_k := \sigma_k^2(2^{\bar{r}_k} - 1)/\beta_k$ . Therefore, by making the variable change  $p_k = \tilde{p}_k + \bar{p}_k$ , (12) is equivalent to

$$\max_{\tilde{\mathbf{p}}=(\tilde{p}_1, \dots, \tilde{p}_K)^T} \sum_{k=1}^K \log_2(a_k + \beta_k \tilde{p}_k / \sigma_k^2) - \tau \tilde{P}_{\text{total}}(\tilde{\mathbf{p}}) \quad (13a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \tilde{p}_k \leq \bar{P}^{\max}, \quad \tilde{p}_k \geq 0, \quad k = 1, \dots, K, \quad (13b)$$

where  $a_k = 1 + \bar{p}_k \beta_k / \sigma_k^2$ ,  $\bar{p}_{\text{cir}} = \alpha \sum_{k=1}^K \bar{p}_k + p_{\text{cir}}$ ,  $p_{\text{cir}} = MP_A + P_C$ ,  $\bar{P}^{\max} = P^{\max} - \sum_{k=1}^K \bar{p}_k$  and  $\tilde{P}_{\text{total}}(\tilde{\mathbf{p}}) \triangleq \alpha \sum_{k=1}^K \tilde{p}_k + \bar{p}_{\text{cir}}$ .

Problem (13) admits a closed-form solution

$$\tilde{p}_k^* = \left[ \frac{1}{\ln 2 (\tau \alpha + \lambda)} - \frac{a_k \sigma_k^2}{\beta_k} \right]^+, \quad (14)$$

where  $[x]^+ = \max\{0, x\}$ . If

$$\sum_{k=1}^K \left[ \frac{1}{\ln 2 \tau \alpha} - \frac{a_k \sigma_k^2}{\beta_k} \right]^+ \leq \bar{P}^{\max},$$

then  $\lambda = 0$ . Otherwise,  $\lambda > 0$  is such that

$$\sum_{k=1}^K \left[ \frac{1}{\ln 2 (\tau \alpha + \lambda)} - \frac{a_k \sigma_k^2}{\beta_k} \right]^+ = \bar{P}^{\max}, \quad (15)$$

which can be easily located using bisection search.

In summary, (11) is solved by the following Dinkelbach's type algorithm.

- *Initialization.* Solve (13) for initial  $\tau > 0$ . If its optimal value is more than zero set  $\underline{\tau} = \tau$  and reset  $\tau \leftarrow 2\tau$  and solve (13) again. Otherwise (its optimal value is less than zero) set  $\bar{\tau} = \tau$ . We end up by having  $\underline{\tau}$  and  $\bar{\tau}$  such that the optimal value of (13) is positive for  $\tau = \underline{\tau}$  and is negative for  $\tau = \bar{\tau}$ . The optimal  $\tau$  for zero optimal value of (13) lies on  $[\underline{\tau}, \bar{\tau}]$  so from now we locate it by bisection in the next stage;
- *Bisection.* Solve (13) for  $\tau = (\underline{\tau} + \bar{\tau})/2$ . If its optimal value is positive reset  $\underline{\tau} \leftarrow \tau$ , otherwise (its optimal value is negative) reset  $\bar{\tau} \leftarrow \tau$ . Process until  $\bar{\tau} - \underline{\tau} \leq \epsilon$  (tolerance) to have the optimal value of (13) is zero.

#### IV. PA FOR CONJUGATE BEAMFORMING

Another class of low-complexity beamforming is conjugate beamforming

$$\mathbf{f}_k = \sqrt{p_k} \mathbf{h}_k^* / \|\mathbf{h}_k\|, \quad (16)$$

where  $\mathbf{h}_k^* = (h_{1k}^*, \dots, h_{Mk}^*)^T$ . The received equation (1) becomes

$$y_k = \sqrt{\beta_k} \sqrt{p_k} \|\mathbf{h}_k\| x_k + \sqrt{\beta_k} \sum_{i \in \mathcal{K} \setminus \{k\}} \sqrt{p_i} \frac{\langle \mathbf{h}_k, \mathbf{h}_i \rangle}{\|\mathbf{h}_i\|} x_i + n_k, \quad (17)$$

where unlike (8) with the MU interference completely cancelled due to the zero-forcing condition (6), the second term in the right hand side of (17) cannot be cancelled. Beside the inherent low-complexity, conjugated beamforming is justified by the fact that  $\|\mathbf{h}_k\|^2 = O(M)$  while  $|\langle \mathbf{h}_k, \mathbf{h}_i \rangle| = O(\sqrt{M})$  as  $M \rightarrow +\infty$  [1], [11], making the SINR in (17) scaled up by  $\sqrt{M}$ .

For  $\beta_{kk} = \bar{\beta}_k \|\mathbf{h}_k\|^2$  and  $\beta_{ki} = \bar{\beta}_k |\langle \mathbf{h}_k, \mathbf{h}_i \rangle|^2 / \|\mathbf{h}_i\|^2$  for  $i \neq k$ , the information throughput at UE  $k$  is given by

$$\tilde{r}_k(\mathbf{p}) = \log_2 \left( 1 + \frac{\beta_{kk} p_k}{\sum_{i \in \mathcal{K} \setminus \{k\}} \beta_{ki} p_i + \sigma_k^2} \right). \quad (18)$$

The EE maximization problem is formulated by

$$\max_{\mathbf{p}=(p_1, \dots, p_K)^T} \frac{1}{\ln 2} F(\mathbf{p}) \quad \text{s.t.} \quad (11c), \quad (19a)$$

$$\log_2 \left( 1 + \frac{\beta_{kk} p_k}{\sum_{i \in \mathcal{K} \setminus \{k\}} \beta_{ki} p_i + \sigma_k^2} \right) \geq \bar{r}_k, \quad k \in \mathcal{K}, \quad (19b)$$

where

$$F(\mathbf{p}) \triangleq \sum_{k=1}^K \tilde{r}_k(\mathbf{p}) / P_{\text{total}}(\mathbf{p}),$$

and  $P_{\text{total}}(\mathbf{p})$  is defined from (10).

Note that (19b) consists of the linear constraints

$$\beta_{kk} p_k \geq (2^{\bar{r}_k} - 1) \left( \sum_{i \in \mathcal{K} \setminus \{k\}} \beta_{ki} p_i + \sigma_k^2 \right), \quad k = 1, \dots, K. \quad (20)$$

As a result, (19) is the following linearly constrained optimization problem:

$$\max_{\mathbf{p}=(p_1, \dots, p_K)^T} \frac{1}{\ln 2} F(\mathbf{p}) \quad \text{s.t.} \quad (11c), (20). \quad (21)$$

Unlike (11), the numerator of the objective function in (21) is no longer concave but is a d.c. (difference of two convex functions) function [13] so the aforementioned Dinkelbach's algorithm is not applicable. The authors in [2] suggested to approximate this objective at each outer iteration by a concave function to facilitate Dinkelbach's algorithm. As mentioned above, Dinkelbach's algorithm requires a few inner iterations, each of which involves optimization of a logarithmic function and thus is still computationally difficult.

We now propose an efficient computational procedure for (21), which needs to solve only a few quadratic convex programs (QPs) of moderate dimension.

Let  $\mathbf{p}^{(n)}$  be a feasible point for (11c) and (20). Using the inequality (31) in the Appendix for  $x = \beta_{kk} p_k (\sum_{i \in \mathcal{K} \setminus \{k\}} \beta_{ki} p_i + \sigma_k^2)$ ,  $t = P_{\text{total}}(\mathbf{p})$ , and  $\bar{x} = \beta_{kk} p_k^{(n)} (\sum_{i \neq k} \beta_{ki} p_i^{(n)} + \sigma_k^2)$ ,  $\bar{t} = P_{\text{total}}(\mathbf{p}^{(n)})$  yields

$$F(\mathbf{p}) \geq F^{(n)}(\mathbf{p}), \quad (22)$$

for

TABLE I: Simulation Setup

| Parameter                     | Assumption                                 |
|-------------------------------|--------------------------------------------|
| Carrier frequency / Bandwidth | 2GHz / 10MHz                               |
| BS transmission power         | 46 dBm                                     |
| Path loss from BS to user     | $148.1 + 37.6 \log_{10} R$ [dB], $R$ in km |
| Shadowing standard deviation  | 8dB                                        |
| Noise power density           | -174 dBm/Hz                                |
| Drain efficiency of amplifier | $\lambda = 0.388$                          |
| The circuit power per antenna | $P_A = 189$ mW                             |
| The non-transmission power    | $P_C = 40$ dBm                             |

$$F^{(n)}(\mathbf{p}) \triangleq \sum_{k=1}^K \left[ a_k^{(n)} - \frac{b_k^{(n)} \sigma_k^2}{\beta_{kk} p_k} - b_k^{(n)} \sum_{i \in \mathcal{K} \setminus \{k\}} \left[ \frac{2\beta_{ki} \sqrt{p_i^{(n)}}}{\beta_{kk} p_k^{(n)}} \sqrt{p_i} \right. \right. \\ \left. \left. - \frac{\beta_{ki} p_i^{(n)}}{(\beta_{kk} p_k^{(n)})^2} \beta_{kk} p_k \right] - c_k^{(n)} P_{\text{total}}(\mathbf{p}) \right], \quad (23)$$

$$\text{and } 0 < a_k^{(n)} \triangleq 2 \frac{\ln(1 + x_k^{(n)})}{\bar{t}} + \frac{x_k^{(n)}}{\bar{t}(x_k^{(n)} + 1)}, \quad 0 < b_k^{(n)} \triangleq \frac{(x_k^{(n)})^2}{\bar{t}(x_k^{(n)} + 1)}, \quad \text{and } 0 < c_k^{(n)} \triangleq \frac{\ln(1 + x_k^{(n)})}{\bar{t}^2}, \quad k = 1, \dots, K.$$

It should be noted that the inequality (32) in the Appendix has been used in deriving (22). Beside relation (22), it is straightforward to see that

$$F(\mathbf{p}^{(n)}) = F^{(n)}(\mathbf{p}^{(n)}). \quad (24)$$

Suppose that  $\mathbf{p}^{(n+1)}$  is the optimal solution of the QP

$$\max_{\mathbf{p}=(p_1, \dots, p_K)^T} \frac{1}{\ln 2} F^{(n)}(\mathbf{p}) \quad \text{s.t.} \quad (11c), (20). \quad (25)$$

It follows from (22) and (24) that

$$F(\mathbf{p}^{(n+1)}) \geq F^{(n)}(\mathbf{p}^{(n+1)}) > F^{(n)}(\mathbf{p}^{(n)}) = F(\mathbf{p}^{(n)})$$

as far as  $\mathbf{p}^{(n+1)} \neq \mathbf{p}^{(n)}$ , i.e.  $\mathbf{p}^{(n+1)}$  is a better point than  $\mathbf{p}^{(n)}$  for problem (21). It can be easily shown that the sequence  $\{\mathbf{p}^{(n)}\}$  converges at least to a locally optimal solution of (21). Algorithm 1 summarizes the proposed QP-based path-following computational procedure for solving (21). The initial point  $\mathbf{p}^{(0)}$  for (21) can be easily located because the constraints in (21) are linear.

**Algorithm 1** : Path-following algorithm for the EE maximization (21)

- 1: **Initialization**: Choose a feasible point  $\mathbf{p}^{(0)}$  for (21). Set  $n := 0$ .
- 2: **Repeat**
- 3:   Solve the QP (25) for the optimal solution  $\mathbf{p}^{(n+1)}$ .
- 4:   Set  $n := n + 1$ .
- 5: **Until** convergence of the objective in (21).

## V. NUMERICAL RESULTS

In this section, the performance of the proposed algorithms will be evaluated using numerical examples. The BS is located in the center of a circular cell with radius 1 km. There are  $K = 16$  UEs, which are uniformly distributed at random within the cell. The other simulation parameters are provided in Table I, which are similar to those used in [4]. Rate QoS of 2 Mbps is set for all users. The proposed computational procedures for solving (11) and (19) typically converge within 10 iterations, yielding their optimal solutions.

Analyzing the objective functions in (11) and (19) we can see that when the circuit power  $MP_A$  dominates their denominators, they are maximized by maximizing the sum throughput in the numerator. On the other hand, by setting the thresholds  $\bar{r}_k$  in (11b) and (19b) higher for better QoS, one needs more transmit power, so the objective functions in (11) and (19) are maximized by minimizing the total transmission power needed to meet the users' QoS. This analysis motivates us to compare the EE performance of (11) and (19) with that obtained by the following throughput maximization:

$$\text{Rmax ZF: } \max_{\mathbf{p}=(p_1, \dots, p_K)} \sum_{k=1}^K r_k(p_k) \quad \text{s.t.} \quad (11c), (11b), \quad (26)$$

for zero-forcing beamforming and

$$\text{Rmax Conj: } \max_{\mathbf{p}=(p_1, \dots, p_K)} \sum_{k=1}^K \tilde{r}_k(\mathbf{p}) \quad \text{s.t.} \quad (11c), (19b) \quad (27)$$

for conjugate beamforming, which can be solved by Algorithm 1, and by the following transmit power minimization:

$$\text{Pmin ZF: } \min_{\mathbf{p}=(p_1, \dots, p_K)} \sum_{k=1}^K p_k \quad \text{s.t.} \quad (11c), (11b), \quad (28)$$

for zero-forcing beamforming and

$$\text{Pmin Conj: } \min_{\mathbf{p}=(p_1, \dots, p_K)} \sum_{k=1}^K p_k \quad \text{s.t.} \quad (11c), (19b), \quad (29)$$

for conjugate beamforming, which are linear programs.

Fig. 1a shows the EE performance vs. the number of antennas  $M$  in all considered optimization problems. The optimal power-allocated ZF beamforming significantly outperforms the optimal power-allocated conjugate beamforming. Although the sum throughput is still increasing in  $M$ , as Fig. 1b shows, the EE performance decreases for  $M > 80$ . This shows that increasing the number of antennas may not result in a gain in the EE because the saved transmit power cannot always compensate for the circuit power increase.

Observe that the EE performance resulting from the transmit power minimizations (28) and (29) is very bad because the circuit power  $MP_A$  actually dominates the consumed power in these optimization problems. Even though the minimal transmit powers in (28) and (29) are different, they do not lead to much change in the consumed power so the EE performances of (28) and (29) are not visually distinguishable.

The EE performance attained by the sum throughput maximizations (26) and (27) is much better than those attained by

the transmit power minimizations (28) and (29). Both Fig. 1a and Fig. 1b also reveal that zero-forcing beamforming also significantly outperforms its conjugate counterpart in these maximization problems. Since the sum throughput is scaled up with the number of antennas, the EE performance of the sum throughput maximizations (26) and (27) also increases. Interestingly, the EE performance of the sum throughput maximization problem (26) for zero-forcing beamforming catches up with that of the EE maximization problem (19) for conjugate beamforming at  $M = 200$ .

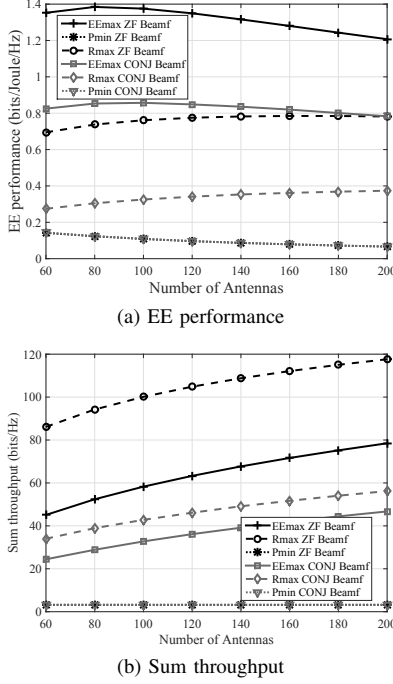


Fig. 1: The EE performance and sum throughput versus the number of antennas  $M$ .

## VI. CONCLUSIONS

We have proposed numerical algorithms of low computational complexity and rapid convergence for maximizing the energy efficiency of zero-forcing and conjugate beamforming in the downlink transmission of a massive MIMO BS while meeting the users' QoS constraints. The effectiveness of these algorithms was confirmed via numerical examples. Their extensions to massive MIMO heterogeneous networks are currently under our consideration.

## APPENDIX: FUNDAMENTAL INEQUALITIES

We exploit the fact that the function  $f(x, t) = \frac{\ln(1+1/x)}{t}$  is convex in  $x > 0, t > 0$  which can be proved by examining its Hessian. The following inequality for all  $x > 0, \bar{x} > 0, t > 0$  and  $\bar{t} > 0$  then holds true [13]:

$$\begin{aligned} \frac{\ln(1+1/x)}{t} &\geq f(\bar{x}, \bar{t}) + \langle \nabla f(\bar{x}, \bar{t}), (x, t) - (\bar{x}, \bar{t}) \rangle \\ &= 2 \frac{\ln(1+1/\bar{x})}{\bar{t}} + \frac{1}{\bar{t}(\bar{x}+1)} - \frac{x}{(\bar{x}+1)\bar{x}\bar{t}} - \frac{\ln(1+1/\bar{x})}{\bar{t}^2} t. \end{aligned} \quad (30)$$

By substituting  $1/x \rightarrow x$  and  $1/\bar{x} \rightarrow \bar{x}$  in (30), we have

$$\frac{\ln(1+x)}{t} \geq a - \frac{b}{x} - ct, \quad (31)$$

where  $a = 2 \frac{\ln(1+\bar{x})}{\bar{t}} + \frac{\bar{x}}{\bar{t}(\bar{x}+1)} > 0$ ,  $b = \frac{\bar{x}^2}{\bar{t}(\bar{x}+1)} > 0$ , and  $c = \frac{\ln(1+\bar{x})}{\bar{t}^2} > 0$ . Finally, by exploiting the fact that the function  $x^2/t$  is convex in  $x > 0$  and  $t > 0$  which leads to the inequality  $\frac{x^2}{t} \geq 2 \frac{\bar{x}x}{\bar{t}} - \frac{\bar{x}^2}{\bar{t}^2} t \forall x > 0, \bar{x} > 0, t > 0, \bar{t} > 0$ , we also have the following inequality by substituting  $x \rightarrow \sqrt{x}$  and  $\bar{x} \rightarrow \sqrt{\bar{x}}$ :

$$\frac{x}{t} \geq 2 \frac{\sqrt{\bar{x}}\sqrt{x}}{\bar{t}} - \frac{\bar{x}}{\bar{t}^2} t \quad \forall x > 0, \bar{x} > 0, t > 0, \bar{t} > 0. \quad (32)$$

## ACKNOWLEDGEMENTS

This work was supported in part by a U.K. Royal Academy of Engineering Research Fellowship under Grant RF1415\14\22 and by the U.K. Engineering and Physical Sciences Research Council under Grant EP/P019374/1, by the Australian Research Councils Discovery Projects under Project DP130104617, and by the U.S. National Science Foundation under Grants CNS-1456793 and ECCS-1343210.

## REFERENCES

- [1] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [2] A. Zappone, L. Sanguinetti, G. Bacci, E. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 1668–1683, Apr. 2016.
- [3] 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. Sel. Areas Commun.*, vol. 34, no. 4, pp. 697–709, Apr. 2016.
- [4] E. Bjornson, M. Kountouris, and M. Debbah, "Massive MIMO and small cells: Improving energy efficiency by optimal soft-cell coordination," in *Proc. 20th Int. Conf. on Telecom. (ICT)*, May. 2013, pp. 1–5.
- [5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in massive MIMO," in *Proc. 22nd European Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, 2014, pp. 76–80.
- [6] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.
- [7] 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.
- [8] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, 1967.
- [9] H. Yang and T. L. Marzetta, "Energy efficient design of massive MIMO: How many antennas?" in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Glasgow, UK, 2015, pp. 1–5.
- [10] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots," in *Proc. 51st Annual Allerton Conf. on Commun., Control, and Computing*, Monticello, IL, USA, Oct. 2013, pp. 293–298.
- [11] 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, Feb. 2014.
- [12] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Now Publishers, 2004.
- [13] H. Tuy, *Convex Analysis and Global Optimization (second edition)*. Springer, 2016.