Distributed Energy Efficiency Fairness Optimization by ADMM in Multicell MISO Downlink

Kien-Giang Nguyen*, Le-Nam Tran†, Quang-Doanh Vu*, and Markku Juntti*

*Centre for Wireless Communications, University of Oulu, P.O.Box 4500, FI-90014 University of Oulu, Finland
Email: {nguyen.k.giang,quvu,markku.juntti}@ee.oulu.fi

†Department of Electronic Engineering, Maynooth University, Maynooth, Co Kildare, Ireland
Email: ltran@eeng.nuim.ie

Abstract—This paper studies the fairness of achievable energy efficiency (EE) in a multicell multiuser multiple-input single-output downlink. The objective is to maximize the minimum EE among all base stations (BSs) subject to per-BS power constraints. The resulting optimization problem is a max-min fractional program, and, thus, difficult to solve in general. Our goal is to develop a decentralized algorithm for the max-min EE problem which solves the problem locally. The idea behind the proposed method is to combine the framework of successive convex approximation (SCA) and alternative direction method of multipliers (ADMM). We transform the convex program obtained at each step of the SCA procedure into a form that lends itself to the ADMM. The resulting formulation is solved optimally by allowing the BSs to exchange the required information until the ADMM converges. In addition to further reduce backhaul overhead, the proposed algorithm is modified to enhance the convergence speed. Numerical results are provided to demonstrate the effectiveness of the proposed algorithms.

I. INTRODUCTION

Multiple-antenna techniques can offer impressive improvements on the achievable capacity of wireless communications systems. Currently, a dominant objective in network design is to maximize the spectral efficiency (SE) to satisfy the increasing demand on data traffic of cellular networks. While SE maximization still remains important, we also need to consider the total energy spent by wireless systems, especially due to recent explosive growth of wireless devices and traffic. Therefore, energy-efficient transmission has got increasing attention recently [1].

Energy efficiency (EE), also known as bit-per-Joule capacity, is defined as the ratio of throughput and the total power consumption of the network. The problem of EE maximization under individual quality-of-service constraints and/or transmit power constraints has been of particular interest [2]–[4]. To maintain fairness of the individual users, maximizing the minimum of EEs of users has been considered [5]–[7] This is basically a max-min fractional program. This optimization problem is nonconvex and generally difficult to solve globally optimally.

For nonconvex problems in general and the max-min EE problem in particular, a classical goal is to find a stationary solution, i.e., a solution that satisfies the Karush-Kuhn-Tucker (KKT) conditions. This was done in [5] by combining the Dinkelbach's approach and alternating optimization. More explicitly, the parameterized problem attained from the Dinkelbach's method is solved by alternately optimizing the beamformers, receivers and other auxiliary variables. We also proposed a more computationally efficient beamforming design for the max-min EE problem, using an inner approximation framework which is now better known as successive convex approximation (SCA) [8]. Briefly, novel transformations were introduced to expose the hidden convexity of the considered problem, and the remaining nonconvexity is iteratively replaced by convex approximants. For single-input single-output (SISO) systems, the SCA principle was also used in [7] to derive power control policies for the max-min EE problem.

The existing beamforming designs or power control policies for EE fairness are centralized algorithms, which need a central node to collect all the system information (i.e., channel state information (CSI) and power parameters) and then solving the EE problem. This may not be appealing from a practical implementation perspective, especially when the amount of CSI needs to be shared is large and/or a central node is difficult to build. In [6], a distributed energy-efficient power optimization for SISO interference channels was studied. Although the main computational efforts can be done distributedly at each BS, some information still needs to be collected and processed at a central node.

In this paper, we consider multicell multiuser multiple-input single-output (MISO) downlink channels and propose a decentralized solution for the problem of achieving fairness EE among BSs, each of which is limited by a total power budget. The proposed algorithm are based on applying the SCA method to obtain convex approximations of the nonconvex max-min EE problem and applying the alternative direction method of multipliers (ADMM) to solve the arrived convex programs. However, this cannot be accomplished by a direct application of the ADMM. To make it possible we introduce the local and global versions of the so-called interference temperature, and decompose the SCA problem into subproblems that can be solved independently at each BS. The proposed distributed method follows strictly the ADMM where we allow the BSs to exchange the relevant information to the others until the ADMM declares convergence. In other words, the ADMM is run until a stationary solution to the SCA problem is found. Once this has been done, all the BSs update the relevant parameters to create a new SCA problem and enter the

next loop of the ADMM. In this way the proposed algorithm can yield the same performance as the centralized method. We will numerically observe that the amount of exchanged information is significant in some cases. Thus, the proposed algorithm is modified to reduce the exchanged information by enhancing the convergence speed. In particular, we limit the maximum number iterations of the ADMM so that the SCA parameters can be updated earlier, and, hence, the total number of iterations is significant reduced. Numerical results are provided to demonstrate the effectiveness of the proposed decentralized algorithms.

Notation: The following notations are used throughout the paper. Vectors are denoted by bold lowercase letters. $\mathbb{C}^{a \times b}$ represents the space of complex matrices of dimension given as superscripts. $(.)^T$ and $(.)^H$ represent the transpose and Hermitian transpose operator, respectively. $|.|$ and $\Re(.)$ represent the absolute value and real part of a complex number, respectively. $\|.\|_2$ represents the ℓ_2 norm. The calligraphic letters denote a set. $\{\mathbf{x}_b\}_{b \in \mathcal{B}}$ refers to a vector $\mathbf{x}_b$ whose indices belong to the set $\mathcal{B}$.

II. System Model and Problem Formulation

We consider a MISO downlink system consisting of B base stations (BS), each is equipped with N antennas. The set of BSs is denoted by $\mathcal{B} = \{1, \ldots, B\}$. To lighten the notation we assume each BS serves a different group of K users and refer to the kth user served by BS b as b_k for $k = 1, 2, \ldots, K$. Note that the total number of users in the considered system model is BK. Assuming a flat fading channel model, the received signal at user b_k is written as

$$
\begin{aligned}
y_{b_k} = \mathbf{h}_{b,b_k} \mathbf{w}_{b_k} s_{b_k} &+ \sum_{i=1, j \neq k}^{K} \mathbf{h}_{b,b_k} \mathbf{w}_{b_j} s_{b_j} \\
&+ \sum_{i=1, i \neq b}^{B} \sum_{j=1}^{K} \mathbf{h}_{i,b_k} \mathbf{w}_{i_j} s_{i_j} + n_{b_k},
\end{aligned} \tag{1}
$$

where $\mathbf{h}_{i,b_k} \in \mathbb{C}^{1 \times N}$ is the channel vector from BS i to user b_k; $\mathbf{w}_{b_k} \in \mathbb{C}^{N \times 1}$ and s_{b_k} are the beamforming vector and the transmit data symbol from BS b to user b_k, respectively; and $n_{b_k} \sim \mathcal{CN}(0, \sigma_{b_k}^2)$ is the additive white Gaussian noise with variance $\sigma_{b_k}^2$. We treat the inter-user interference as Gaussian noise and thus the SINR for user b_k is given by

$$
\gamma_{b_k}(\mathbf{w}) = \frac{|\mathbf{h}_{b,b_k} \mathbf{w}_{b_k}|^2}{I_{b_k}(\mathbf{w}) + W \sigma_{b_k}^2} \tag{2}
$$

where W is the system bandwidth, $\mathbf{w}$ is the vector encompassing the beamformers of all BK users, and $I_{b_k}(\mathbf{w})$ is defined as

$$
I_{b_k}(\mathbf{w}) \triangleq \sum_{j=1, j \neq k}^{K} |\mathbf{h}_{b,b_k} \mathbf{w}_{b_j}|^2 + \sum_{i=1, i \neq b}^{B} \sum_{j=1}^{K} |\mathbf{h}_{i,b_k} \mathbf{w}_{i_j}|^2. \tag{3}
$$

Accordingly, the EE of BS b is expressed as

$$
f_b(\mathbf{w}) = \frac{\sum_{k=1}^{K} W \log_2\left(1 + \gamma_{b_k}(\mathbf{w})\right)}{\frac{1}{\epsilon} \sum_{k=1}^{K} \|\mathbf{w}_{b_k}\|_2^2 + N P_{\mathrm{dp}} + P_{\mathrm{sp}}} \tag{4}
$$

where $\epsilon \in (0, 1)$ is the power amplifier efficiency, P_{dp} is the dynamic power consumption corresponding to the power radiation of all circuit blocks in each active radio frequency chain, and P_{sp} is the static power spent by the cooling system, power supply, etc. We denote by $P_0 \triangleq N P_{\mathrm{dp}} + P_{\mathrm{sp}}$ the total circuit power. The problem of max-min EE fairness among all the BSs is mathematically stated as

$$
\max_{\mathbf{w}} \quad \min_{1 \leq b \leq B} \quad f_b(\mathbf{w}) \tag{5a}
$$

$$
\text{subject to} \quad \sum_{k=1}^{K} \|\mathbf{w}_{b_k}\|_2^2 \leq P_b, \quad \forall b \in \mathcal{B} \tag{5b}
$$

where P_b is the transmit power budget of BS b.

We remark that existing solutions aim at finding a stationary solution to (5), which is a classical goal for general nonconvex problems [5], [7], [8]. However, these previous studies ended up with centralized solutions, requiring a central node in their implementation to collect the required information and carry out all the computations. Among them, the one-layer iterative method proposed in [8] based on the SCA framework, has been shown to achieve lowest computational complexity. This method will form a basis for the development of decentralized algorithms to solve the max-min EE problem in this paper.

III. Beamforming Design for Max-Min EE Fairness

A. Centralized Approach

In this subsection we briefly review the low-complexity beamforming design for the max-min EE problem introduced in our earlier work [8]. We start by introducing an equivalent reformulation of (5) which more exposes the hidden convexity. Specifically, (5) can be rewritten as

$$
\max_{\mathbf{w}, \eta, \mathbf{z}, \mathbf{t}} \quad \eta \tag{6a}
$$

$$
\text{subject to} \quad z_b^2 / t_b \geq \eta, \quad \forall b \in \mathcal{B} \tag{6b}
$$

$$
\sum_{k=1}^{K} \log\left(1 + \frac{|\mathbf{h}_{b,b_k} \mathbf{w}_{b_k}|^2}{I_{b_k}(\mathbf{w}) + W \sigma_{b_k}^2}\right) \geq z_b^2, \tag{6c}
$$

$$
\frac{1}{\epsilon} \sum_{k=1}^{K} \|\mathbf{w}_{b_k}\|_2^2 + P_0 \leq t_b, \quad \forall b \in \mathcal{B} \tag{6d}
$$

$$
\sum_{k=1}^{K} \|\mathbf{w}_{b_k}\|_2^2 \leq P_b, \quad \forall b \in \mathcal{B}, \tag{6e}
$$

where η, $\mathbf{z} \triangleq [z_1, \ldots, z_B]^T$, $\mathbf{t} \triangleq [t_1, \ldots, t_B]^T$ are newly introduced slack variables. To deal with the nonconvexity of (6c) we further introduce new auxillirary variables $g_{b_k} \in \mathbf{g} \triangleq \{g_{b_1}, \ldots, g_{b_K}\}_{b \in \mathcal{B}}$ and $q_{b_k} \in \mathbf{q} \triangleq \{q_{b_1}, \ldots, q_{b_K}\}_{b \in \mathcal{B}}$, and

rewrite (6) as

$$\max_{\mathbf{w},\eta,\mathbf{z},\mathbf{t},\mathbf{g},\mathbf{q}} \quad \eta \tag{7a}$$

$$\text{subject to} \quad z_b^2/t_b \geq \eta, \ \forall b \in \mathcal{B} \tag{7b}$$

$$\sum_{k=1}^{K} \log(1 + g_{b_k}) \geq z_b^2, \ \forall b \in \mathcal{B} \tag{7c}$$

$$\frac{|\mathbf{h}_{b,b_k}\mathbf{w}_{b_k}|^2}{q_{b_k}} \geq g_{b_k}, \ \forall b \in \mathcal{B}, k \in \mathcal{K}_b \tag{7d}$$

$$q_{b_k} \geq I_{b_k}(\mathbf{w}) + W\sigma_{b_k}^2, \ \forall b \in \mathcal{B}, k \in \mathcal{K}_b \tag{7e}$$

$$(6d), (6e). \tag{7f}$$

It is easy to see that the constraints from (7b) to (7f) must hold with equality at the optimality. Thus, the equivalence between (5) and (7) is guaranteed. We further note that all the constraints in (7) are convex, excluding (7b) and (7d). This implies that we need to find ways to deal with the nonconvexity in (7b) and (7d). Based on the observation that that (7b) and (7d) admit the same form, i.e., the left side is a quadratic-over-affine function (which is convex) and the right side is an affine function, we apply the application of the inner approximation algorithm [9] to approximate the nonconvex part of (7b) and (7d). More explicitly, we iteratively replace (7b) and (7d) by

$$\phi_b^{(n)}(z_b, t_b) \geq \eta \tag{8}$$

and

$$\psi_{b_k}^{(n)}(\mathbf{w}_{b_k}, q_{b_k}) \geq g_{b_k} \tag{9}$$

respectively, where

$$\phi_b^{(n)}(z_b, t_b) \triangleq \frac{2z_b^{(n)}}{t_b^{(n)}} z_b - \frac{(z_b^{(n)})^2}{(t_b^{(n)})^2} t_b \tag{10}$$

and

$$\psi_{b_k}^{(n)}(\mathbf{w}_{b_k}, q_{b_k}) \triangleq \frac{2\Re(\mathbf{h}_{b,b_k}^{(n)}\mathbf{w}_{b_k})}{q_{b_k}^{(n)}} - \frac{|\mathbf{h}_{b,b_k}\mathbf{w}_{b_k}^{(n)}|^2 q_{b_k}}{(q_{b_k}^{(n)})^2} \tag{11}$$

are the first order of z_b^2/t_b around the point $(z_b^{(n)}, t_b^{(n)})$ and $|\mathbf{h}_{b,b_k}\mathbf{w}_{b_k}|^2/q_{b_k}$ around the point $(\mathbf{w}_{b_k}^{(n)}, q_{b_k}^{(n)})$, respectively, and $\mathbf{h}_{b,b_k}^{(n)} = (\mathbf{w}_{b_k}^{(n)})^H \mathbf{h}_{b,b_k}^H \mathbf{h}_{b,b_k}$. The superscript n in above equations denotes the nth iteration of the iterative procedure. In summary, the approximate convex program at iteration $n+1$ of the proposed iterative algorithm proposed in [8] is given by

$$\max_{\mathbf{w},\eta,\mathbf{z},\mathbf{t},\mathbf{g},\mathbf{q}} \quad \eta \tag{12a}$$

$$\text{subject to} \quad (6d), (6e), (7c), (7e), (8), (9). \tag{12b}$$

B. Decentralized Approaches for the max-min EE

Our goal in this subsection is to propose a decentralized solution to solve (7). Such solutions would be of practical interest when a central processing station is not available, or when sending the channel state information of all BSs to a central node is overwhelming. To the best of our knowledge, decentralized solutions for the max-min EE problem have not been investigated previously in the related literature. The idea is to propose a distributed approach for solving the convex

program obtained in each step of the iterative procedure. Thus we base our decentralized algorithms on the formulation in (12) to solve it optimally in the sequel.

A simple and more popular approach in the context of distributed optimization is the dual decomposition method. However, this approach cannot be applied to solve the problem (12) due to the lack of strict convexity of the objective (see [10] for further details). A similar problem was also mentioned in [11], but in a different context. A remedy for this issue is to combine the dual decomposition and augmented Lagrangian method which results in the so-called ADMM [10], [11]. In addition to the lack of strict convexity, the formulation in (12) is not amendable to a direct application of ADMM since (7e) is not decomposable due to the intercell interference (ICI) terms contained in $I_{b_k}(\mathbf{w})$. To overcome these issues, we rewrite (12) as

$$\min_{\mathbf{w},\eta,\mathbf{z},\mathbf{t},\mathbf{g},\mathbf{q},\boldsymbol{\tau},\boldsymbol{\tau}'} \quad -\sum_{b=1}^{B} \eta_b \tag{13a}$$

$$\text{subject to} \quad \eta_b = \eta \tag{13b}$$

$$\tau_{b,i_j} \geq \sum_{k=1}^{K} |\mathbf{h}_{b,i_j}\mathbf{w}_{b_k}|^2 \ \forall i \neq b \tag{13c}$$

$$q_{b_k} \geq \sum_{j=1, j\neq k}^{K} |\mathbf{h}_{b,b_k}\mathbf{w}_{b_j}|^2 + \sum_{i=1, i\neq b}^{B} \tau'_{b_k, i} + W\sigma_{b_k}^2 \tag{13d}$$

$$\tau_{b,i_j} = \mu_{b,i_j} \ \forall i \neq b \tag{13e}$$

$$\tau'_{b_k, i} = \mu_{i, b_k} \ \forall i \neq b \tag{13f}$$

$$(6d), (6e), (7c), (8), (9), \tag{13g}$$

where $\eta_b, \tau_{b,i_j} \in \boldsymbol{\tau}, \tau'_{b_k, i} \in \boldsymbol{\tau}'$ are newly introduced slack variables, representing the energy efficiency of BS b, the ICI producing by BS b to user i_j, and the ICI at user b_k generated by neighboring BS i ($i \neq b$), respectively. We remark that $\eta_b, \tau_{b,i_j}, \tau'_{b_k, i}$ locally stored at BS b. On the other hand, η, μ_{b,i_j} and μ_{i, b_k} are respectively the global copies of the interference variables $\tau_{b,i_j}, \tau'_{b_k, i}$ to satisfy the equivalence between (13) and (12). We will refer the interference (local or global) variables as interference temperatures since their value varies after every iteration. Next we include all the constraints that can be handled locally at BS b in the set $\mathcal{S}_b$ which is defined as

$$\mathcal{S}_b = \{\mathbf{s}_b \mid \phi_b^{(n)}(z_b, t_b) \geq \eta_b, \tau_{b,i_j} \geq \sum_{k=1}^{K} |\mathbf{h}_{b,i_j}\mathbf{w}_{b_k}|^2,$$

$$\psi_{b_k}^{(n)}(\mathbf{w}_{b_k}, q_{b_k}) \geq g_{b_k}, \sum_{k=1}^{K} \log(1 + g_{b_k}) \geq z_b^2,$$

$$\frac{1}{\epsilon}\sum_{k=1}^{K} \|\mathbf{w}_{b_k}\|_2^2 + P_0 \leq t_b, \sum_{k=1}^{K} \|\mathbf{w}_{b_k}\|_2^2 \leq P_b,$$

$$q_{b_k} \geq \sum_{j=1, j\neq k}^{K} |\mathbf{h}_{b,b_k}\mathbf{w}_{b_j}|^2 + \sum_{i\neq b, i=1}^{B} \tau'_{b_k, i} + W\sigma_{b_k}^2\}$$

$$\tag{14}$$

where $\mathbf{s}_b \triangleq \{\eta_b, z_b, t_b, \{g_{b_k}\}_{k=1}^{K}, \{\mathbf{w}_{b_k}\}_{k=1}^{K}, \{q_{b_k}\}_{k=1}^{K}, \{\{\tau_{b,i_j}\}_{j=1}^{K}\}_{i\in\mathcal{I}_b}, \{\{\tau'_{b_k, i}\}_{k=1}^{K}\}_{i\in\mathcal{I}_b}\}$ represents the set of the local variables at BS b, and the set $\mathcal{I}_b$ is denoted as $\mathcal{I}_b \triangleq \mathcal{B}\setminus\{b\}$, i.e, $\mathcal{I}_b$ is the set including the indices of all other BSs. For notation convenience, we denote two sets of variables

for each BS, i.e., $\boldsymbol{\theta}_b \triangleq \left\{ \{\{\tau_{b,i_j}\}_{j=1}^K\}_{i\in\mathcal{I}_b}, \{\{\tau'_{b_k,i}\}_{k=1}^K\}_{i\in\mathcal{I}_b} \right\}$ and $\boldsymbol{\nu}_b \triangleq \left\{ \{\{\mu_{b,i_j}\}_{j=1}^K\}_{i\in\mathcal{I}_b}, \{\{\mu_{i,b_k}\}_{k=1}^K\}_{i\in\mathcal{I}_b} \right\}$ rearranging the local interference temperatures of BS b and the corresponding global versions, respectively. We remark that $\boldsymbol{\theta}_b \in \mathbf{s}_b$ and further denote by $\mathbf{s} \triangleq \{\mathbf{s}_b\}_{b\in\mathcal{B}}$ and $\boldsymbol{\vartheta} = \{\eta, \{\boldsymbol{\nu}_b\}_{b\in\mathcal{B}}\}$ the set of local and global variables, respectively. From the steps above, (13) can be equivalently rewritten as

$$\min_{\mathbf{s},\boldsymbol{\vartheta}} \quad -\sum_{b=1}^B \eta_b \tag{15a}$$

$$\text{subject to} \quad \mathbf{s}_b \in \mathcal{S}_b, \ \forall b \in \mathcal{B} \tag{15b}$$

$$\eta_b = \eta, \ \forall b \in \mathcal{B} \tag{15c}$$

$$\boldsymbol{\theta}_b = \boldsymbol{\nu}_b, \ \forall b \in \mathcal{B}. \tag{15d}$$

The reformulation (15) of (12) is now amendable for applying ADMM to find a decentralized solution. In particular, we derive the augmented Lagrangian function of (15) as

$$\mathcal{L}^{(n+1)}(\mathbf{s}, \boldsymbol{\vartheta}, \boldsymbol{\zeta}, \boldsymbol{\xi}) = \sum_{b=1}^B [-\eta_b + \xi_b(\eta_b - \eta) + \boldsymbol{\zeta}_b^T(\boldsymbol{\theta}_b - \boldsymbol{\nu}_b)$$
$$+ \frac{c}{2}((\eta_b - \eta)^2 + \|\boldsymbol{\theta}_b - \boldsymbol{\nu}_b\|_2^2)] \tag{16}$$

where $c > 0$ is the penalty parameter and $\boldsymbol{\xi} \triangleq \{\xi_b\}_{b\in\mathcal{B}}$, $\boldsymbol{\zeta} \triangleq \{\boldsymbol{\zeta}_b\}_{b\in\mathcal{B}}$ (in which $\boldsymbol{\zeta}_b = \left\{ \{\{\zeta_{b,i_j}\}_{j=1}^K\}_{i\in\mathcal{I}_b}, \{\{\zeta_{b_k,i}\}_{k=1}^K\}_{i\in\mathcal{I}_b} \right\}$ corresponding to variables τ_{b,i_j} and $\tau'_{b_k,i}$, respectively) are the Lagrange multipliers. The general idea of the ADMM is to use the Gauss-Seidel method to update the local variables (i.e. $\mathbf{s}$), the global variables (i.e. $\boldsymbol{\vartheta}$), and the Lagrange multipliers (i.e. $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$). In other words, we fix the other variables when updating the local variables, and the same manner is alternatively applied for the global variables and Lagrange multipliers, respectively. Particularly, the variables updating procedure at the iteration $l+1$ of the ADMM to solve (15) is detailed as follows.

First, the set of local variables $\mathbf{s}$ is updated by solving the convex problem

$$\mathbf{s}^{(l+1)} = \arg\min_{\mathbf{s}} \mathcal{L}^{(n+1)}(\mathbf{s}, \boldsymbol{\vartheta}^{(l)}, \boldsymbol{\xi}^{(l)}, \boldsymbol{\zeta}^{(l)}). \tag{17}$$

Interestingly, it is easily seen that the augmented Lagrangian function in (16) is decomposable, thus, the optimization problem (17) can be separately implemented at each BS. Accordingly, B subproblems are solved in the parallel at each BS, i.e., the subproblem at the BS b is given by

$$\mathbf{s}_b^{(l+1)} = \arg\min_{\mathbf{s}_b \in \mathcal{S}_b} [-\eta_b + \xi_b^{(l)}(\eta_b - \eta^{(l)}) + \boldsymbol{\zeta}_b^{(l)T}(\boldsymbol{\theta}_b - \boldsymbol{\nu}_b^{(l)})$$
$$+ \frac{c}{2}((\eta_b - \eta^{(l)})^2 + \|\boldsymbol{\theta}_b - \boldsymbol{\nu}_b^{(l)}\|_2^2)]. \tag{18}$$

We remark that since all the constraints listed in (14) are convex and SOC representable (when the conic approximation for (6c) is applied [8]), then (18) can be efficiently solved by the modern SOCP solvers.

Next, the update of global variables is carried out by using the exchanged information η_b and interference temperature $\boldsymbol{\theta}_b$ of each BS. Similarly, the separable form of the augmented Lagrangian function (16) turns this step to the distributed fashion where η and $\boldsymbol{\nu}_b$ are independently updated by

$$\eta^{(l+1)} = \arg\min_{\eta} \sum_{b=1}^B \left[-\xi_b^{(l)}\eta + \frac{c}{2}(\eta_b^{(l+1)} - \eta)^2 \right] \tag{19}$$

and

$$\boldsymbol{\nu}_b^{(l+1)} = \arg\min_{\boldsymbol{\nu}_b} -\boldsymbol{\zeta}_b^{(l)T}\boldsymbol{\nu}_b + \frac{c}{2}\|\boldsymbol{\theta}_b^{(l+1)} - \boldsymbol{\nu}_b\|_2^2. \tag{20}$$

As can be seen, the equations (19) and (20) are unconstrained quadratic programming w. r. t. the involved variables. Thus, the solutions admit closed-forms, i.e,

$$\eta^{(l+1)} = \frac{1}{B}\sum_{b=1}^B \left[\eta_b^{(l+1)} + \frac{1}{c}\xi_b^{(l)} \right] \tag{21}$$

and

$$\mu_{b,i_j}^{(l+1)} = \frac{\tau_{b,i_j}^{(l+1)} + \tau'^{(l+1)}_{i_j,b}}{2} + \frac{\zeta_{b,i_j}^{(l)} + \zeta_{i_j,b}^{(l)}}{2c} \tag{22}$$

where ζ_{b,i_j} and $\zeta_{i_j,b}$ are the dual variables associated to the primal variables τ_{b,i_j} of BS b and $\tau'_{i_j,b}$ of BS i ($i \neq b$), respectively. To compute $\eta^{(l+1)}$ in (21), all BSs can run an average consensus algorithm [10], [12]. We note that $\mu_{b,i_j}^{(l+1)}$ is computed after gathering $\tau'^{(l+1)}_{i_j,b}$ and $\zeta_{i_j,b}^{(l)}$ from BS $i \neq b$

The last step of the ADMM is to update the Lagrangian multipliers as

$$\xi_b^{(l+1)} = \xi_b^{(l)} + c\left(\eta_b^{(l+1)} - \eta^{(l+1)}\right) \tag{23}$$

$$\boldsymbol{\zeta}_b^{(l+1)} = \boldsymbol{\zeta}_b^{(l)} + c\left(\boldsymbol{\theta}_b^{(l+1)} - \boldsymbol{\nu}_b^{(l+1)}\right). \tag{24}$$

Because the obligated information for calculating (23) and (24) is available at all BSs after updating global variables, the multipliers update can be implemented at each BS without extra gathered information.

After the ADMM procedure converges i.e, the residual $\varepsilon_{\text{ADMM}}$ below a threshold, we update the involved SCA parameter $(\mathbf{w}^{(n+1)}, \mathbf{q}^{(n+1)}, \mathbf{z}^{(n)}, \mathbf{t}^{(n)})$ as in the centralized algorithm until satisfying the stopping criteria. To summarize, Algorithm 1 outlines the proposed decentralized algorithm for max-min EE beamforming design problem based on ADMM leveraging SCA.

We remark that the global variables update can be carried out without the information of Lagrangian multipliers, i.e,

$$\eta^{(l+1)} = \frac{1}{B}\sum_{b=1}^B \eta_b^{(l+1)} \tag{25}$$

$$\mu_{b,i_j}^{(l+1)} = (\tau_{b,i_j}^{(l+1)} + \tau'^{(l+1)}_{i_j,b})/2. \tag{26}$$

Indeed, we have $\sum_{b=1}^B \xi_b^{(l+1)} = \sum_{b=1}^B [\xi_b^{(l)} + c(\eta_b^{(l+1)} - \eta^{(l+1)})] = 0$. Following the same manner, we can easily see that $\sum_{b=1}^B \boldsymbol{\zeta}_b = 0$ and thereby obtaining $\zeta_{b,i_j}^{(l)} + \zeta_{i_j,b}^{(l)} = 0$.

Algorithm 1 Proposed decentralized beamformer ADMM-based design for max-min EE in multicell MISO downlink

---

1: **Initialization:** Set $n := 0$, choose initial value for $(\mathbf{w}^{(n)}, \mathbf{q}^{(n)}, \mathbf{z}^{(n)}, \mathbf{t}^{(n)})$ such that these satisfy the feasible set of each BS $\mathcal{S}_b$, choose the initial value for $\boldsymbol{\vartheta}^{(0)}, \boldsymbol{\xi}^{(0)}, \boldsymbol{\zeta}^{(0)}$.

2: **repeat**

3:    $l := 0$.

4:    **while** $\varepsilon_{\mathrm{ADMM}} \geq 10^{-5}$ **do**

5:      **for** $b \in \mathcal{B}$ **do**

6:        BS b updates $\mathbf{s}_b^{(l+1)}$ by (18).

7:        BS b updates $\eta^{(l+1)}$ through an average consensus algorithm [12].

8:        BS b receives $\tau'^{(l+1)}_{i_j,b}$, then updates $\mu_{b,i_j}^{(l+1)}$ by (26).

9:        Update Lagrange multipliers $\xi_b^{(l+1)}$ and $\zeta_b^{(l+1)}$ by (23) and (24), respectively.

10:      **end for**

11:    $l := l + 1$.

12:    **end while**

13:    Obtain the optimal value $(\mathbf{w}^*, \mathbf{q}^*, \mathbf{z}^*, \mathbf{t}^*)$.

14:    Update the SCA parameters $(\mathbf{w}^{(n)}, \mathbf{q}^{(n)}, \mathbf{z}^{(n)}, \mathbf{t}^{(n)}) = (\mathbf{w}^*, \mathbf{q}^*, \mathbf{z}^*, \mathbf{t}^*)$, and $(\boldsymbol{\vartheta}^{(0)}, \boldsymbol{\xi}^{(0)}, \boldsymbol{\zeta}^{(0)}) = (\boldsymbol{\vartheta}^*, \boldsymbol{\xi}^*, \boldsymbol{\zeta}^*)$.

15:    $n := n + 1$.

16: **until** the SCA converges

Thus, the actual obligated information required to exchange among the BSs is only the local variables.

The convergence of Algorithm 1 includes the convergence analysis of the ADMM and the SCA convergence. It is worth noting that the convergence behavior both of the algorithms has been well investigated in literature, and we skip it due to the spaces. The interested reader can refer to the results in [10] and [9].

Per-Base station Complexity Analysis: We remark that the per-iteration complexity of Algorithm 1 is dominated by the complexity of solving the subproblem (18) at each BS. To reduce the complexity when solving (18) which is in fact the general convex program due to the constraint $\sum_{k=1}^{K} \log(1 + g_{b_k}) \geq z_b^2$, the SOC approximation of exponential cone as in [8, (13)] is applied. Thus, the cost of solving the subproblem (18) at each BS is $\mathcal{O}(N^3 K^3 + (m + 7)^3 K^3)$, where m is the conic approximation parameter of accuracy [13, Sect. 6.6].

Amount of Exchanged Information: The information of the local ICI $\boldsymbol{\theta}_b$ and EE η_b at each BS is required to broadcast to the other BSs. The amount of exchanged information mainly depends on the steps 7 and 8 of Algorithm 1, i.e., the average consensus algorithm. This implies that BS b needs to send out total numbers of $(B-1)(K+1)$ real values to the other $(B-1)$ BSs. In cellular networks, e.g LTE Radio Access Network, this could be done by deploying the X2 interface which connects neighboring eNodeBs in point-to-point fashion.

Implementation issues: Waiting for the ADMM-based loop to completely converge at each SCA iteration ensures the global convergence of Algorithm 1 but generally slows down its convergence rate. We now provide some modifications that

can be made to Algorithm 1 to improve its convergence rate in practice. The idea is to allow the ADMM part to terminate early in first SCA iterations. In particular, the ADMM loop terminates after I_{ADMM} updates and this number can be varied as the SCA outer loop evolves. In some first SCA iterations, we numerically obverse that the increase in the cost function between two consecutive iterations is large, and thus it usually requires a large number of updates for the ADMM to converge. However, the solutions return by the ADMM part are still a rough estimate of the solution of the considered problem in the first SCA steps. Thus, it is unnecessary to let the ADMM complete converge in these cases. Our rationale is that the ADMM part can produce a good estimate after a number of updates. On the other hand, when the SCA is nearly convergent, we can set I_{ADMM} to be smaller. By adapting I_{ADMM} we can enhance the convergence rate of Algorithm 1 significantly as shown in Section IV.

IV. NUMERICAL RESULTS

Table I
SIMULATION PARAMETERS

PARAMETERS	VALUE
Pathloss and shadowing	$38\log(d) + 34.5 + \mathcal{N}(0, 8)$ [dB]
Cell radius	500 [m]
Power budget P_b	35 [dBm]
Static power consumption P_{sp}	33 [dBm]
Dynamic power consumption P_{dp}	40 [dBm]
Power amplifier efficiency ϵ	0.35
Number of BSs B	3
Number of per-BS users	2
Number of BS antennas	4
Signal bandwidth W	10 [kHz]
Power spectral density of noise	-174 [dBm/Hz]

In this section, we demonstrate the effectiveness of the proposed method by numerical experiments. The simulation parameters are listed in Table I. All the convex problems considered in this paper are solved using the MOSEK solver in MATLAB environment. Specially, we apply the conic approximation presented in [8] for the nonlinear logarithmic constraints to improve the processing speed. The initial values of Algorithm 1 are generated as follows. First, a set of beamformers that satisfy (6e) is created at each BS. Then the initial values of other variables $(z_b^{(0)}, t_b^{(0)}, q_{b_k}^{(0)})$ are obtained by setting all the constrains in (14) to be equality. The initial global values $\boldsymbol{\mu}^{(0)}$ and $\eta^{(0)}$ and the Lagrangian multipliers $\boldsymbol{\xi}^{(0)}$ and $\boldsymbol{\zeta}^{(0)}$ are set to zero. The iterations of the proposed decentralized algorithms are terminated if the increase of the value of η in two consecutive SCA iterations is less than 10^{-5}. The centralized algorithm compared to proposed approaches uses the same setting as in [8]. We note that the proposed algorithm using limited number of the ADMM iterations to enhance the convergence rate is referred as 'Algorithm 1 modified'.

Fig. 1 plots the convergence results of Algorithm 1 and its modified version. We can see that the proposed algorithms are able to obtain the same stationary point as the centralized
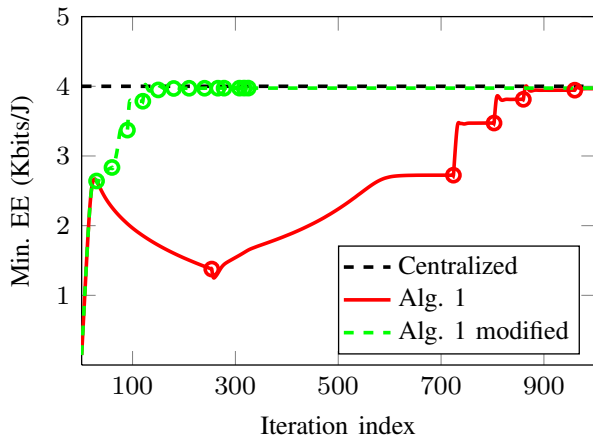
(a) Example 1
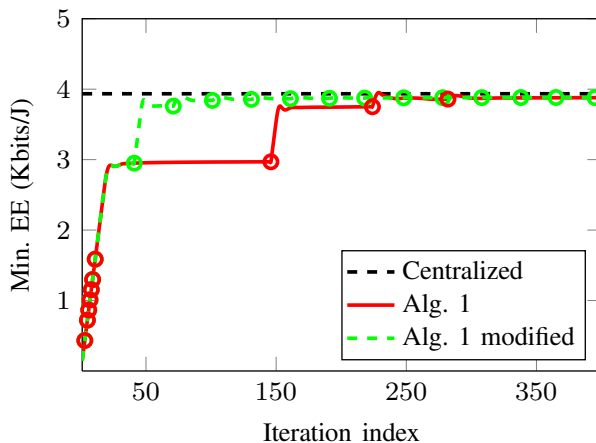
(b) Example 2

Fig. 1. Convergence of the Algorithm 1 and its modified version with $I_{ADMM} = 40$. The circles represent the iterations where the SCA parameters are updated

approach. Since Algorithm 1 updates the SCA parameters after the ADMM completely converging, it requires hundreds of iterations to achieve the stationary solution. On the other hand, the modified approach succeeds to speed up the convergence rate. Fig. 1(a) shows an example of the bad choice of initial set. As can be seen, the very first ADMM procedure in Algorithm 1 takes a large number of iterations to obtain the convergence point at the first update of the SCA. The other example in Fig. 1(b) shows that lots of iterations are spent to obtain a very slight gain in the achieved EE. Both scenarios result in a slow convergence rate. By limiting the maximum iterations of the ADMM, the SCA parameters are updated earlier and the total number of required iterations are significantly reduced. This task is equivalent to re-pick a better initial point for iterative procedure (as shown in Fig. 1(a)), and eliminate the inefficient iterations (as shown in Fig. 1(b)).

V. CONCLUSION

This paper has studied the problem of EE optimization for multicell multiuser MISO downlink, which aims at maximiz-

ing the minimum EE among all BSs. We have proposed the decentralized algorithm to distributedly solve the EE fairness problem at each BS. The idea of combining the framework of the SCA and the ADMM was carried out. In particular, the proposed algorithm resorts the convex approximation achieved at each iteration of the SCA method to apply the ADMM to solve the problem in the distributed manner. The proposed approach follows strictly the ADMM where we allow the BSs to exchange the relevant information to the others until the ADMM declares a convergence. The simulation results demonstrate that the decentralized algorithm is able to achieve the stationary solution as the centralized algorithm. We have further numerically shown that the convergence rate of the proposed method can be improved by limiting the maximum number iterations of the ADMM.

REFERENCES

[1] D. Feng, C. Jiang, G. Lim, J. Cimini, L. J., G. Feng, and G. Li, "A survey of energy-efficient wireless communication," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 167–178, Feb. 2013.

[2] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user mimo systems: Is massive mimo the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059 – 3075, Jun. 2015.

[3] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser mimo systems: Spectral and energy efficiency maximization," *IEEE Signal Process. Lett.*, vol. 61, no. 16, pp. 4038–3050, Aug. 2013.

[4] O. Tervo, L.-N. Tran, and M. Juntti, "Optimal energy-efficient transmit beamforming for multi-user miso downlink," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5574 – 5588, Oct. 2015.

[5] S. He, Y. Huang, S. Jin, F. Yu, and L. Yang, "Max-min energy efficient beamforming for multicell multiuser joint transmission systems," *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1956–1959, Oct. 2013.

[6] B. Du, C. Pan, W. Zhang, and M. Chen, "Distributed energy-efficient power optimization for CoMP systems with max-min fairness," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 999–1002, Jun. 2014.

[7] A. Zappone, L. Sanguinetti, G. Bacci, E. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," arXiv:1503.04609v1.

[8] K.-G. Nguyen, L.-N. Tran, O. Tervo, Q.-D. Vu, and M. Juntti, "Achieving energy efficiency fairness in multicell multiuser miso downlink," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1426 – 1429, Aug. 2015.

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

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statisticallearning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1.

[11] C. Shen, T.-H. Chang, K.-Y. Wang, and Z. Qiu, "Distributed robust multicell coordinated beamforming with imperfect CSI: An ADMM approach," *IEEE Signal Process. Lett.*, vol. 60, no. 6, pp. 2988–3003, Jun. 2012.

[12] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, pp. 65–78, 2004.

[13] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization*. Philadelphia: MPS-SIAM Series on Optimization, SIAM, 2001.