Title	Decentralized coordinated beamforming for weighted sum energy efficiency maximization in multi-cell MISO downlink
Authors(s)	Tervo, Oskari, Tran, Le-Nam, Juntti, Markku
Publication date	2015-12-16
Publication information	Tervo, Oskari, Le-Nam Tran, and Markku Juntti. "Decentralized Coordinated Beamforming for Weighted Sum Energy Efficiency Maximization in Multi-Cell MISO Downlink." IEEE, December 16, 2015. https://doi.org/10.1109/GlobalSIP.2015.7418426.
Conference details	The 3rd IEEE Conference on Signal & Information Processing, Orlando, Florida, United States of America, 14-16 December 2015
Publisher	IEEE
Item record/more information	http://hdl.handle.net/10197/12188
Publisher's statement	© 2015 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.
Publisher's version (DOI)	10.1109/GlobalSIP.2015.7418426

Decentralized Coordinated Beamforming for Weighted Sum Energy Efficiency Maximization in Multi-Cell MISO Downlink

Oskari Tervo*, Le-Nam Tran†, and Markku Juntti*
*Centre for Wireless Communications, University of Oulu, P.O.Box 4500, FI-90014 University of Oulu, Finland
Email: {oskarite,markku.juntti}@ee.oulu.fi
†Department of Electronic Engineering, Maynooth University, Maynooth, Co Kildare, Ireland
Email: ltran@eeng.nuim.ie

***Abstract*—We study energy-efficient decentralized coordinated beamforming in multi-cell multiuser multiple-input single-output system. The problem of interest is to maximize the weighted sum energy efficiency subject to user-specific quality of service constraints. The original problem is iteratively approximated as a convex program according to successive convex approximation (SCA) principle. The convex problem at each iteration is then formulated as a general global consensus problem, which is solved via alternating direction method of multipliers (ADMM). This enables base stations to independently and in parallel optimize their beamformers relying only on local channel state information and limited backhaul information exchange. In addition to waiting for the ADMM to converge as conventionally when solving the approximate convex program, we propose a method where only one ADMM iteration is performed after each SCA update step. Numerical results illustrate the fast convergence of the proposed methods and show that performing only one ADMM iteration per each convex problem can significantly improve the convergence speed.**

I. Introduction

To address high volume of wireless traffic and high data rate demand, current wireless communications systems have been built upon multi-antenna technology. Although the multi-antenna techniques have been shown to provide huge spectral efficiency gains, they are achieved at the expense of increased energy consumption. To aim at energy savings for the future networks, energy efficiency has become an increasingly important optimization criterion [1]–[4]. By optimizing the ratio of the data rate and the total power consumption (including the data independent circuit power consumption), significant energy efficiency gains can be achieved over the traditional design criteria, e.g., transmit power minimization and sum rate maximization.

The 5G visions include network densification as an important upcoming trend [5]. In small-cell networks, inter-cell interference becomes a limiting performance factor which has to be accounted for in the network optimization. One efficient concept to coordinate interference is *coordinated beamforming*, where base stations (BSs) can jointly design their data transmissions by exchanging limited amount of control information via reliable backhaul links [6].

The energy efficiency maximization (EEmax) problems belong to the class of fractional programs, which are highly difficult to solve even in the case of single-cell beamforming [7]–[9]. Energy-efficient multi-cell beamforming has been studied in some recent works. Nguyen *et al.* [10] considered the problem of maximizing the minimum energy efficiency among the BSs to provide EE fairness in a multi-cell multiuser multiple-input single-output (MISO) system. Coordinated beamforming for network EEmax in multi-cell multi-antenna systems was studied in [11], [12], where the latter incorporated the data rate constraints of the users in the optimization. To satisfy the heterogeneous energy efficiency requirements of different cells, He *et al.* [13] used the weighted sum energy efficiency as the optimization criterion. The above mentioned works require a centralized controller which carries out all the optimization. Distributed methods for network EEmax problem was studied in [14], [15].

In this paper, we propose a decentralized energy-efficient coordinated beamforming method for multi-cell multiuser MISO system. On contrary to the network EEmax problem studied in [14], [15], we aim at maximizing the weighted sum of the energy efficiencies of the cells subject to user-specific signal-to-interference-plus-noise ratio (SINR) constraints. This optimization criterion can potentially satisfy heterogeneous energy efficiency requirements of different BS types. The original problem is iteratively approximated as a convex problem according to successive convex approximation (SCA) principle. The convex problem at each iteration is then formulated as a global consensus problem [16], which is solved via alternating direction method of multipliers (ADMM). Different to prior works, we propose a method, which can be realized in a decentralized manner using only local channel state information and limited scalar information exchange between the BSs, and guarantees certain SINR level for each user.

II. System Model and Problem Formulation

A. System Model

We consider a B-cell multiuser multiple-input single-output downlink channel. Each base station $b \in \mathcal{B} = \{1, \dots, B\}$ equipped with N antennas transmits data to a group of K_b single-antenna users in its cell, represented by the set $\mathcal{K}_b$. Each user $k \in \mathcal{K} \triangleq \cup_{b\in\mathcal{B}}\mathcal{K}_b$ in the network is served only by a single BS which is denoted by $b_k \in \mathcal{B}$, i.e., $\mathcal{K}_b \cap \mathcal{K}_m = \emptyset \ \forall b \neq m$. The channel (row) vector from BS b to user k is represented by $\mathbf{h}_{b,k} \in \mathbb{C}^{1\times N}$. We adopt linear beamforming, where data symbol s_k intended for user k is multiplied with the beamforming vector $\mathbf{w}_k \in \mathbb{C}^{N\times 1}$ before being transmitted. Accordingly, the received signal for user k is given by

$$y_k = \mathbf{h}_{b_k,k}\mathbf{w}_k s_k + \sum\nolimits_{j\in\mathcal{K}_{b_k}\setminus\{k\}} \mathbf{h}_{b_k,k}\mathbf{w}_j s_j + \sum\nolimits_{m\in\bar{\mathcal{B}}_{b_k}} \sum\nolimits_{j\in\mathcal{K}_m} \mathbf{h}_{m,k}\mathbf{w}_j s_j + n_k \quad (1)$$

where $\bar{\mathcal{B}}_b = \mathcal{B}\setminus\{b\}$ is the set of all BSs, excluding b, and $n_k \sim \mathcal{CN}(0, N_0)$ is background noise with noise power spectral density N_0. The data streams are independent and have zero mean and unit power. Considering a time-division duplex system, we assume that BSs can perfectly estimate the channels of all the users from the uplink pilots. Accordingly, the data rate of user k is given by

$$R_k(\mathbf{w}) = W\log(1+\Gamma_k(\mathbf{w})) \quad (2)$$

where W is the bandwidth and

$$\Gamma_k(\mathbf{w}) \triangleq \frac{|\mathbf{h}_{b_k,k}\mathbf{w}_k|^2}{\mathcal{I}_k(\tilde{\mathbf{w}}_{b_k}) + \sum_{m\in\bar{\mathcal{B}}_{b_k}} \mathcal{I}_{m,k}(\tilde{\mathbf{w}}_m)} \quad (3)$$

is the SINR of user k. In (2) and (3), $\tilde{\mathbf{w}}_b$ is defined as the vector stacking all the beamformers of users served by BS b, i.e., $\tilde{\mathbf{w}}_b = \{\mathbf{w}_k\}_{k\in\mathcal{K}_b}$, and $\mathbf{w}$ as the vector including the beamformers of all users in the network, i.e., $\mathbf{w} = \{\tilde{\mathbf{w}}_b\}_{b\in\mathcal{B}}$. Moreover, $\mathcal{I}_k(\tilde{\mathbf{w}}_{b_k}) \triangleq WN_0 + \sum_{j\in\mathcal{K}_{b_k}\setminus\{k\}} |\mathbf{h}_{b_k,k}\mathbf{w}_j|^2$ is the intra-cell interference-plus-noise and $\mathcal{I}_{m,k}(\tilde{\mathbf{w}}_m) \triangleq \sum_{j\in\mathcal{K}_m} |\mathbf{h}_{m,k}\mathbf{w}_j|^2$ is the inter-cell interference (ICI) from BS $m \in \bar{\mathcal{B}}_{b_k}$ to user k.

B. Power Consumption Model

We adopt the power model from [17], where the power consumption of each cell is modeled as

$$P_{\text{tot},b} = \tfrac{1}{\eta}P_{\text{data,b}} + P_{\text{FIX}} + NP_{\text{BS}} + P_{\text{SYN}} + K_bP_{\text{UE}} \quad (4)$$

where $P_{\text{data},b}$ is the data transmit power of BS b in downlink, $\eta \in [0,1]$ is the power amplifier efficiency at the BS, P_{FIX} is fixed power consumption required for site-cooling, control signaling, backhaul infrastructure, and baseband processors, P_{BS} is the power per radio frequency chain at each antenna, P_{SYN} is the power consumed by local oscillator and P_{UE} is the circuit power of each user. We denote the total circuit power as $P_0 \triangleq P_{\text{FIX}} + NP_{\text{BS}} + P_{\text{SYN}} + K_bP_{\text{UE}}$.

C. Weighted Sum Energy Efficiency Maximization

The problem of weighted sum energy efficiency maximization with per-user SINR constraints can be written as

$$\begin{aligned}
\max_{\{\mathbf{w}_k\}_{k\in\mathcal{K}}} \quad & \sum\nolimits_{b\in\mathcal{B}} \omega_b \frac{f_b(\mathbf{w})}{g_b(\tilde{\mathbf{w}}_b)} && (5a)\\
\text{s.t.} \quad & \Gamma_k(\mathbf{w}) \geq \bar{\Gamma}_k, k\in\mathcal{K} && (5b)\\
& \sum\nolimits_{k\in\mathcal{K}_b} ||\mathbf{w}_k||_2^2 \leq P_b, \forall b\in\mathcal{B} && (5c)
\end{aligned}$$

where ω_b is the energy efficiency priority weight for BS b, $f_b(\mathbf{w}) \triangleq \sum_{k\in\mathcal{K}_b} R_k(\mathbf{w})$, $g_b(\tilde{\mathbf{w}}_b) \triangleq \frac{1}{\eta}\sum_{k\in\mathcal{K}_b} ||\mathbf{w}_k||_2^2 + P_0$, and $\bar{\Gamma}_k$ is the target SINR. We note that the constraint in (5b) admits a second-order cone representation as

$$\begin{cases}
\frac{\mathbf{h}_{b_k,k}\mathbf{w}_k}{\sqrt{\bar{\Gamma}_k}} \geq \left(\mathcal{I}_k(\tilde{\mathbf{w}}_{b_k}) + \sum_{m\in\bar{\mathcal{B}}_{b_k}} \mathcal{I}_{m,k}(\tilde{\mathbf{w}}_m)\right)^{\frac{1}{2}} & (6a)\\
\text{Im}(\mathbf{h}_{b_k,k}\mathbf{w}_k) = 0. & (6b)
\end{cases}$$

The equivalence between (5b) and (6) is due to the fact that a phase rotation on $\mathbf{w}_k$ will offer the same objective in (5a) and still satisfies the power constraint in (5c). This observation has been exploited in [18] to transform the SPmin problem into the second-order cone program. The difficulty of the problem is now in the nonconvex objective function (5a).

III. Successive Convex Approximation

As a first step, we can write the following equivalent transformation

$$\begin{aligned}
\max_{\{t_b\}_{b\in\mathcal{B}},\{\mathbf{w}_k\}_{k\in\mathcal{K}}} \quad & W\sum\nolimits_{b\in\mathcal{B}} \omega_b t_b && (7a)\\
\text{s.t.} \quad & t_b \leq \frac{f_b(\mathbf{w})}{g_b(\tilde{\mathbf{w}}_b)}, \forall b\in\mathcal{B} && (7b)\\
& (5c), (6) && (7c)
\end{aligned}$$

which is in fact an epigraph form of (5). To lighten notations we will drop the constant W in (7a) onwards. Next, we can further equivalently reformulate (7) as

$$\begin{aligned}
\max_{\{t_b,\alpha_b\}_{b\in\mathcal{B}},\{\mathbf{w}_k,\gamma_k\}_{k\in\mathcal{K}}} \quad & \sum\nolimits_{b\in\mathcal{B}} \omega_b t_b && (8a)\\
\text{s.t.} \quad & g_b(\tilde{\mathbf{w}}_b) \leq \frac{\alpha_b^2}{t_b}, \forall b\in\mathcal{B} && (8b)\\
& \gamma_k \leq \Gamma_k(\mathbf{w}), \forall k\in\mathcal{K} && (8c)\\
& \sum\nolimits_{k\in\mathcal{K}_b} \log(1+\gamma_k) \geq \alpha_b^2, \forall b\in\mathcal{B} && (8d)\\
& (5c), (6). && (8e)
\end{aligned}$$

The equivalence between (8) and (7) is guaranteed since all the constraints (8b)-(8d) are active at optimality. Intuitively, the newly introduced variables $\{\gamma_k\}_{k\in\mathcal{K}}, \{\alpha_b\}_{b\in\mathcal{B}}$ represent the SINR of user k and the square root of the sum data rate of cell b, respectively. It is easy to see that the nonconvexity of (8) is due to (8b) and (8c), and the decomposability of (8) depends on (8c) and (6a).

To address the nonconvexity of (8) we first equivalently replace (8c) by the following inequality constraints

$$\gamma_k \leq \frac{|\mathbf{h}_{b_k,k}\mathbf{w}_k|^2}{\beta_k}, \forall k\in\mathcal{K} \quad (9a)$$

$$\beta_k \geq \mathcal{I}_k(\tilde{\mathbf{w}}_{b_k}) + \sum\nolimits_{m\in\bar{\mathcal{B}}_{b_k}} \mathcal{I}_{m,k}(\tilde{\mathbf{w}}_m), \forall k\in\mathcal{K}, \quad (9b)$$

where $\{\beta_k\}_{k\in\mathcal{K}}$ are newly introduced variables which can be interpreted as total interference plus noise experienced by user k. We note that (9b) are convex constraints. At this point, we can express the original problem (5) as

$$\begin{aligned}
\max_{\{t_b,\alpha_b\}_{b\in\mathcal{B}},\{\mathbf{w}_k,\gamma_k,\beta_k\}_{k\in\mathcal{K}}} \quad & \sum\nolimits_{b\in\mathcal{B}} \omega_b t_b && (10a)\\
\text{s.t.} \quad & g_b(\tilde{\mathbf{w}}_b) \leq \alpha_b^2/t_b, \forall b\in\mathcal{B} && (10b)\\
& \gamma_k \leq |\mathbf{h}_{b_k,k}\mathbf{w}_k|^2/\beta_k, \forall k\in\mathcal{K} && (10c)\\
& (5c), (6), (8d), (9b). && (10d)
\end{aligned}$$

We can see that all the constraint except (10b) and (10c) are convex. We note that $\frac{\alpha_b^2}{t_b}$ and $\frac{|\mathbf{h}_{b_k,k}\mathbf{w}_k|^2}{\beta_k}$ are quadratic-over-linear, i.e., jointly convex functions of the same form and admit linear lower approximations at point $(\alpha_b^{(n)}, t_b^{(n)})$ and $(\mathbf{w}_k^{(n)}, \beta_k^{(n)})$ as

$$\frac{\alpha_b^2}{t_b} \geq \frac{2\alpha_b^{(n)}}{t_b^{(n)}}\alpha_b - \left(\frac{\alpha_b^{(n)}}{t_b^{(n)}}\right)^2 t_b \triangleq \phi^{(n)}(\alpha_b, t_b) \quad (11)$$

and

$$\frac{|\mathbf{h}_{b_k,k}\mathbf{w}_k|^2}{\beta_k} \geq \frac{2\mathrm{Re}((\mathbf{w}_k^{(n)})^H\mathbf{h}_{b_k,k}^H\mathbf{h}_{b_k,k}\mathbf{w}_k)}{\beta_k^{(n)}} - (\frac{|\mathbf{h}_{b_k,k}\mathbf{w}_k^{(n)}|}{\beta_k^{(n)}})^2\beta_k \triangleq \psi^{(n)}(\mathbf{w}_k,\beta_k), \quad (12)$$

respectively. Thus, we iteratively approximate the original problem (5) as a convex program

$$\max_{\{t_b,\alpha_b\}_{b\in\mathcal{B}},\{\mathbf{w}_k,\gamma_k,\beta_k\}_{k\in\mathcal{K}}} \quad \sum\nolimits_{b\in\mathcal{B}}\omega_b t_b \quad (13a)$$
$$\text{s.t.} \quad g_b(\tilde{\mathbf{w}}_b) \leq \phi^{(n)}(\alpha_b,t_b), \forall b\in\mathcal{B} \quad (13b)$$
$$\gamma_k \leq \psi^{(n)}(\mathbf{w}_k,\beta_k), \forall k\in\mathcal{K} \quad (13c)$$
$$\text{(5c), (6), (8d), (9b)} \quad (13d)$$

at iteration n of the SCA method until convergence. The convergence of SCA is studied in [19] and is omitted here for brevity.

IV. Decentralized Algorithm

After a careful check we can see that only the constraints (6a) and (9b) couple all the beamformers across the BSs. If all BSs are connected with a central processing node, then (13) can be solved in a centralized manner. However, this is not often the case in practice and decentralized algorithms are thus always appreciated. The focus of this section is to develop a consensus based decentralized approach to solve (13) [16]. Toward this end we first rewrite (13) equivalently as

$$\max \quad \sum\nolimits_{b\in\mathcal{B}}\omega_b t_b \quad (14a)$$
$$\text{s.t.} \quad \frac{\mathbf{h}_{b_k,k}\mathbf{w}_k}{\sqrt{\Gamma_k}} \geq (\mathcal{I}_k(\tilde{\mathbf{w}}_{b_k}) + \sum\nolimits_{m\in\bar{\mathcal{B}}_{b_k}} z_{m,k}^2)^{\frac{1}{2}}, \forall k\in\mathcal{K} \quad (14b)$$
$$\beta_k \geq \mathcal{I}_k(\tilde{\mathbf{w}}_{b_k}) + \sum\nolimits_{m\in\bar{\mathcal{B}}_{b_k}} z_{m,k}^2, \forall k\in\mathcal{K} \quad (14c)$$
$$z_{b,j}^2 \geq \sum\nolimits_{k\in\mathcal{K}_b}|\mathbf{h}_{b_k,j}\mathbf{w}_k|^2, \forall b\in\mathcal{B}, j\in\bar{\mathcal{K}}_b \quad (14d)$$
$$\text{(5c), (6b), (8d), (13b), (13c)} \quad (14e)$$

where the optimization variables are $\{t_b,\alpha_b\}_{b\in\mathcal{B}}, \{\mathbf{w}_k,\gamma_k,\beta_k\}_{k\in\mathcal{K}}, \{z_{b,j}\}_{b\in\mathcal{B},j\in\bar{\mathcal{K}}_b}$, and $\bar{\mathcal{K}}_b \triangleq \mathcal{K}\setminus\mathcal{K}_b$ is the set of users who are not served by cell b. The newly added variables $\{z_{b,j}\}_{b\in\mathcal{B},j\in\bar{\mathcal{K}}_b}$ can be interpreted as the square root of the ICI introduced by cell b to user $j\in\bar{\mathcal{K}}_b$. For cell b we define a local feasible set $\mathcal{S}_b$ as

$$\begin{aligned}\mathcal{S}_b = \Big\{&(\tilde{\mathbf{w}}_b,\boldsymbol{\gamma}_b,\alpha_b,t_b,\boldsymbol{\beta}_b,\tilde{\mathbf{z}}_b)\big|\\ &\sum\nolimits_{k\in\mathcal{K}_b}||\mathbf{w}_k||_2^2 \leq P_b\\ &\mathrm{Im}(\mathbf{h}_{b_k,k}\mathbf{w}_k) = 0, \forall k\in\mathcal{K}_b\\ &\sum\nolimits_{k\in\mathcal{K}_b}\log(1+\gamma_k) \geq \alpha_b^2\\ &g_b(\tilde{\mathbf{w}}_b) \leq \phi^{(n)}(\alpha_b,t_b)\\ &\gamma_k \leq \psi^{(n)}(\mathbf{w}_k,\beta_k), \forall k\in\mathcal{K}_b\\ &\frac{\mathbf{h}_{b_k,k}\mathbf{w}_k}{\sqrt{\Gamma_k}} \geq (\mathcal{I}_k(\tilde{\mathbf{w}}_b) + \sum\nolimits_{m\in\bar{\mathcal{B}}_{b_k}}(\tilde{z}_{m,k}^b)^2)^{\frac{1}{2}}, \forall k\in\mathcal{K}_b\\ &\beta_k \geq \mathcal{I}_k(\tilde{\mathbf{w}}_b) + \sum\nolimits_{m\in\bar{\mathcal{B}}_b}(\tilde{z}_{m,k}^b)^2, \forall k\in\mathcal{K}_b\\ &(\tilde{z}_{b,j}^b)^2 \geq \sum\nolimits_{k\in\mathcal{K}_b}|\mathbf{h}_{b,j}\mathbf{w}_k|^2, \forall j\in\bar{\mathcal{K}}_b\Big\} \quad (15)\end{aligned}$$

where $\boldsymbol{\gamma}_b = \{\gamma_k\}_{k\in\mathcal{K}_b}$, $\boldsymbol{\beta}_b = \{\beta_k\}_{k\in\mathcal{K}_b}$ and we have replaced each $z_{m,k}$ by a new variable $\tilde{z}_{m,k}^b$. In fact, $\tilde{z}_{m,k}^b$ can be viewed as the local copy of $z_{m,k}$ handled by cell b and $\tilde{\mathbf{z}}_b \in \mathcal{R}^{|\bar{\mathcal{K}}_b|+K_b(B-1)}$ includes all the variables that are relevant to BS b, i.e., the interference that BS b causes to the users of other cells and the interference that other cells $m\in\bar{\mathcal{B}}_b$ cause to the users of cell b. With the introduction of $\mathcal{S}_b$ we can equivalently write (14) in a more compact form as

$$\max \quad \sum\nolimits_{b\in\mathcal{B}}\omega_b t_b \quad (16a)$$
$$\text{s.t.} \quad (\tilde{\mathbf{w}}_b,\boldsymbol{\gamma}_b,\alpha_b,t_b,\boldsymbol{\beta}_b,\tilde{\mathbf{z}}_b) \in \mathcal{S}_b, \forall b\in\mathcal{B} \quad (16b)$$
$$\tilde{\mathbf{z}}_b = \mathbf{z}_b, \forall b\in\mathcal{B} \quad (16c)$$

where the optimization variables are $\{\tilde{\mathbf{w}}_b,\boldsymbol{\gamma}_b,\alpha_b,t_b,\boldsymbol{\beta}_b,\tilde{\mathbf{z}}_b,\mathbf{z}_b\}_{b\in\mathcal{B}}$ and a vector $\mathbf{z}_b \in \mathcal{R}^{|\bar{\mathcal{K}}_b|+K_b(B-1)}$ includes the corresponding global variables that are relevant to BS b, i.e., all the elements from $\{z_{b,j}^b\}_{j\in\bar{\mathcal{K}}_b}$ and $\{\tilde{z}_{m,k}^b\}_{m\in\bar{\mathcal{B}}_b,k\in\mathcal{K}_b}$ with the same ordering as in $\tilde{\mathbf{z}}_b$. The equality constraints in (16c) force the local copies to be equal to the global variables. Now we note that (16) is in a form of general global consensus problem which can be solved using ADMM [16]. Let us denote $\boldsymbol{\Psi}_b = \{\tilde{\mathbf{w}}_b,\boldsymbol{\gamma}_b,\alpha_b,t_b,\boldsymbol{\beta}_b,\tilde{\mathbf{z}}_b\}$, including the local variables associated with BS b. Furthermore, we denote

$$h_b(\boldsymbol{\Psi}_b) = \begin{cases}\omega_b t_b, \boldsymbol{\Psi}_b\in\mathcal{S}_b\\ \infty, \quad \text{otherwise.}\end{cases} \quad (17)$$

The basic idea of the ADMM is to alternately update local, global and dual variables. As a first step of the ADMM, we write the partial augmented Lagrangian of (16) as

$$\begin{aligned}&L(\{\boldsymbol{\Psi}_b,\mathbf{z}_b,\boldsymbol{\nu}_b\}_{\forall b\in\mathcal{B}})\\ &= \sum\nolimits_{b\in\mathcal{B}}\left(h_b(\boldsymbol{\Psi}_b) - \boldsymbol{\nu}_b^T(\tilde{\mathbf{z}}_b-\mathbf{z}_b) - \tfrac{\rho}{2}||\tilde{\mathbf{z}}_b-\mathbf{z}_b||_2^2\right)\end{aligned} \quad (18)$$

where $\{\boldsymbol{\nu}_b\}_{b\in\mathcal{B}}$ are the dual variables associated with the interference equality constraints (16c). The last term of (18) is a quadratic penalty term with penalty parameter $\rho > 0$, which penalizes for the violation of the equality constraints (16c). Due to the added penalty terms, the ADMM is able to converge without the need of strict convexity or finiteness of the original objective function (16). At iteration $l+1$ of the ADMM, we can write the update steps as

$$\boldsymbol{\Psi}_b^{(l+1)} = \arg\max_{\boldsymbol{\Psi}_b} \quad L(\boldsymbol{\Psi}_b,\mathbf{z}_b^{(l)},\boldsymbol{\nu}_b^{(l)}), \forall b\in\mathcal{B} \quad (19a)$$
$$\{\mathbf{z}_b^{(l+1)}\}_{\forall b\in\mathcal{B}} = \arg\max_{\{\mathbf{z}_b\}_{\forall b\in\mathcal{B}}} \quad L(\{\boldsymbol{\Psi}_b^{(l+1)},\mathbf{z}_b,\boldsymbol{\nu}_b^{(l)}\}_{\forall b\in\mathcal{B}}) \quad (19b)$$
$$\boldsymbol{\nu}_b^{(l+1)} = \boldsymbol{\nu}_b^{(l)} + \rho(\tilde{\mathbf{z}}_b^{(l+1)} - \mathbf{z}_b^{(l+1)}), \forall b\in\mathcal{B}. \quad (19c)$$

The steps (19a) and (19c) are completely decentralized and can be solved independently in parallel at each BS. Updating the global variables in (19b) requires sharing local interference terms $\tilde{\mathbf{z}}_b$ to the coupled BSs. Assuming that there is a dedicated backhaul link between each BS and $K_b = K, \forall b\in\mathcal{B}$, each BS needs to share only $2K$ real scalars per link at each iteration.[1] The local variables in (19a) are updated as $\boldsymbol{\Psi}_b^{(l+1)} = \boldsymbol{\Psi}_b^*$, where $\boldsymbol{\Psi}_b^*$ is the optimal solution of the generalized convex program

$$\max_{\boldsymbol{\Psi}_b} \quad h_b(\boldsymbol{\Psi}_b) - (\boldsymbol{\nu}_b^{(l)})^T(\tilde{\mathbf{z}}_b - \mathbf{z}_b^{(l)}) - \tfrac{\rho}{2}||\tilde{\mathbf{z}}_b - \mathbf{z}_b^{(l)}||_2^2. \quad (20)$$

[1] Note that a centralized solution would require each BS to share all the local channel state information to a centralized controller (or to other BSs in the cooperation set). In this case, each BS has to send BKN complex values per link.

The global variables in (19b) can be found as $\mathbf{z}_b^{(l+1)} = \mathbf{z}_b^*$, where $\mathbf{z}_b^*$ is the optimal solution of the convex program

$$\max_{\{\mathbf{z}_b\}_{b\in\mathcal{B}}} -\sum\nolimits_{b\in\mathcal{B}}((\boldsymbol{\nu}_b^{(l)})^T(\tilde{\mathbf{z}}_b^{(l+1)}-\mathbf{z}_b)-\tfrac{\rho}{2}||\tilde{\mathbf{z}}_b^{(l+1)}-\mathbf{z}_b||_2^2) \quad (21)$$

which is separable in $\mathbf{z}_b$ and can be solved independently after the local copies have been exchanged. Since the problem is quadratic and unconstrained, setting the gradient of (21) to zero gives

$$z_{b,j}^* = \frac{1}{2}\big(\tilde{z}_{b,j}^{b,(l+1)} + \tilde{z}_{b,j}^{b_j,(l+1)} + \frac{1}{\rho}(\nu_{b,j}^{b,(l)} + \nu_{b,j}^{b_j,(l)})\big). \qquad (22)$$

By substituting $\mathbf{z}_b^{(l+1)}$ to (19c), we obtain $\nu_{b,j}^{b_j,(l)} + \nu_{b,j}^{b,(l)} = 0$, which yields $z_{b,j}^{(l+1)} = z_{b,j}^* = 1/2(\tilde{z}_{b,j}^{b,(l+1)} + \tilde{z}_{b,j}^{b_j,(l+1)})$. The algorithm is summarized in Algorithm 1, where $\mathbf{r}_b^{(l)} \triangleq \tilde{\mathbf{z}}_b^{(l)} - \mathbf{z}_b^{(l)}$ and ϵ is a small threshold. The convergence proof is omitted here due to the lack of space. However, for each fixed SCA step, we solve a convex problem (13) via ADMM. For convex problems, the ADMM, i.e., steps 3-7 of Algorithm 1 is guaranteed to converge to the optimal solution [16]. For the SCA method, the convergence is studied in [19]. By combining the results of [16] and [19], the convergence is guaranteed. Note that during the intermediate iterations of the ADMM, the feasibility of the original problem (14) is not guaranteed, because local copies of the interference terms can differ from the global ones, and, thus, the SINR constraints may be violated.

Algorithm 1 Proposed decentralized solution.

Initialization: Set $n = 0$, and generate initial points $\mathbf{\Phi}_b^{(0)}$.
1: **repeat**
2: $n := n + 1$.
3: **repeat**
4: BS $b, \forall b \in \mathcal{B}$: Update local variables using (19a) and signal local copies $\tilde{\mathbf{z}}_b$ to the coupled BSs.
5: BS $b, \forall b \in \mathcal{B}$: Update global variables $\mathbf{z}_b$ using (19b).
6: BS $b, \forall b \in \mathcal{B}$: Update local dual variables using (19c).
7: **until** $||\mathbf{r}_b^{(l)}||_2 \leq \epsilon, \forall b \in \mathcal{B}$
8: BS $b, \forall b \in \mathcal{B}$: Perform SCA step, i.e., update functions $\psi^{(n+1)}(\mathbf{w}_k, \beta_k)$ and $\phi^{(n+1)}(\alpha_b, t_b)$ using $\mathbf{w}_k^*, \beta_k^*$ and α_b^*, t_b^* obtained from step 4, respectively.
9: **until** convergence

A simplified version of Algorithm 1

In Algorithm 1, the stopping criterion of ADMM is satisfied when $||\mathbf{r}_b^{(l)}||_2 \leq \epsilon, \forall b \in \mathcal{B}$, where $\mathbf{r}_b^{(l)} \triangleq \tilde{\mathbf{z}}_b^{(l)} - \mathbf{z}_b^{(l)}$ is the residual at BS b after the lth update and ϵ is a small positive constant. This requires all the BS to repeat steps 4-6 several times until reaching a predetermined accuracy. In an effort to combine SCA method and ADMM, we discover an interesting result. That is, Algorithm 1 will also converge even if steps 4-6 are performed an arbitrary number of times at each BS. In this paper, we illustrate this result by performing steps 4-6 only once after each SCA update (i.e., step 8). We refer to this variant of Algorithm 1 as 'simplified Algorithm 1'. Numerical results show that simplified Algorithm 1 can remarkably improve the convergence speed. However, the convergence of simplified Algorithm 1 is not theoretically guaranteed and is very challenging to prove but we have observed its convergence in all numerical experiments.

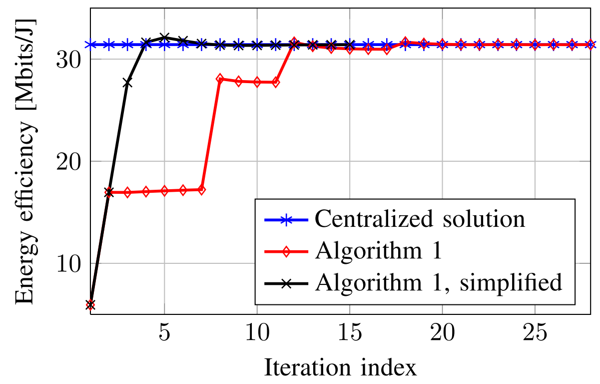

Fig. 1. Convergence of the proposed algorithm.

V. Numerical Results

We evaluate the performance by assuming quasistatic frequency flat Rayleigh fading channels and consider a cluster of $B = 3$ circular cells, with inter-BS distance d_{B} of 400 m. The radius of each cell is assumed to be $\frac{d_{\text{B}}}{2}$, i.e., the cell edges are overlapping and the users are randomly dropped to the cell edges on the center area of the three BSs. The following simulation parameters adopted from [20] are used: $W = 20$ MHz, $P_{\text{FIX}} = 18$ Watts, $P_{\text{BS}} = 1$ Watt, $P_{\text{SYN}} = 2$ Watts, $\rho = 0.04, \bar{\Gamma}_k = \bar{\Gamma} = 0$ dB, $P_b = 46$ dBm, $\eta = 0.39, WN_0 =$ -96 dBm, $N = 9, K_b = K = 3$. The threshold for ADMM to stop is set to $\epsilon = 10^{-2}$.

Fig. 1 illustrates the convergence of Algorithm 1. We can see that Algorithm 1 converges to the value produced by the centralized method and the speed of convergence is fast even with quite large number of antennas/users. Note that the during ADMM iterations, the objective value of Algorithm 1 can be higher than that of the centralized method, because the feasibility is not guaranteed.

VI. Conclusion

We have proposed the decentralized coordinated beamforming method for weighted sum energy efficiency maximization in multi-cell multiuser MISO system. The method is a combination of successive convex approximation and alternating direction method of multipliers, where each base station can independently optimize its beamformers relying only on local channel state information and limited backhaul information exchange. Numerical results illustrate the fast convergence of the proposed method.

Acknowledgment

This project has been co-funded by the Irish Government and the European Union under Ireland's EU Structural and Investment Funds Programmes 2014-2020 through the SFI Research Centres Programme under Grant 13/RC/2077. The work has also been supported by the Finnish Funding Agency for Technology and Innovation, Infotech Oulu Doctoral Program, HPY Research Foundation, Walter Ahlström Foundation, Tauno Tönning Foundation, and KAUTE Foundation.

References

[1] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.

[2] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3304, Sep. 2012.

[3] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 1–14, Jan. 2015.

[4] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell ofdma systems with limited backhaul capacity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3618–3631, Oct. 2012.

[5] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.

[6] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[7] C. Jiang and L. Cimini, "Downlink energy-efficient multiuser beamforming with individual SINR constraints," in *IEEE MILCOM*, 2011, pp. 495–500.

[8] O. Tervo, L.-N. Tran, and M. Juntti, "Joint transmit beamforming and antenna selection for energy efficiency maximization in MISO downlink," in *Proc. IEEE Global Conf. Signal Inform. Process.*, 2014, pp. 118–122.

[9] O. Tervo, L.-N. Tran, and M. Juntti, "Energy-efficient transmit beamforming for MISO downlink via sequential convex approximation," in *Proc. IEEE Int. Workshop on Signal Process. Adv. Wireless Commun.*, 2015, pp. 415–419.

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

[11] S. He, Y. Huang, S. Jin, and L. Yang, "Coordinated beamforming for energy efficient transmission in multicell multiuser systems," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4961–4971, 2013.

[12] Y. Li, Y. Tian, and C. Yang, "Energy-efficient coordinated beamforming under minimal data rate constraint of each user," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2387–2397, Jun. 2015.

[13] S. He, Y. Huang, L. Yang, and B. Ottersten, "Coordinated multicell multiuser precoding for maximizing weighted sum energy efficiency," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 741–751, Feb 2014.

[14] Y. Huang, S. He, S. Jin, and W. Chen, "Decentralized energy-efficient coordinated beamforming for multicell systems," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4302–4314, Nov 2014.

[15] S. He, W. Chen, Y. Huang, S. Jin, L. Jiang, and G. Wang, "Distributed energy-efficient design for coordinated multicell downlink transmission," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2015, pp. 539–544.

[16] C. S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," 2011.

[17] R. Kumar and J. Gurugubelli, "How green the LTE technology can be?" in *Proc. Wireless VITAE*, 2011, pp. 1–5.

[18] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[19] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *Journal of Global Optimization*, vol. 47, no. 1, pp. 29–51, 2010.

[20] 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.