Joint Energy Efficiency Optimization with Nonlinear Precoding in Multi-cell Broadcast Systems

Xin Gui, Kyoung-Jae Lee, Jaehoon Jung, and Inkyu Lee

Abstract: In this paper, we focus on maximizing weighted sum energy efficiency (EE) for a multi-cell multi-user channel. In order to solve this non-convex problem, we first decompose the original problem into a sequence of parallel subproblems which can optimized separately. For each subproblem, a base station employs dirty paper coding to maximize the EE for users within a cell while regulating interference induced to other cells. Since each subproblem can be transformed to a convex multiple-access channel problem, the proposed method provides a closed-form solution for power allocation. Then, based on the derived optimal covariance matrix for each subproblem, a local optimal solution is obtained to maximize the sum EE. Finally, simulation results show that our algorithm based on non-linear precoding achieves about 20 percent performance gains over the conventional linear precoding method.

Index Terms: Dirty paper coding, energy efficiency, multicell broadcast channels

I. INTRODUCTION

W IRELESS communication system designs have faced great challenges due to increased demand on higher transmission capacity, lower bit error rate, and better cell coverage. Multiple-input multiple-output (MIMO) technologies exploit multi-antennas to provide extra spatial degrees of freedom and brings diversity and multiplexing gains [1], [2], [5]. As a result, the MIMO has become an essential technique for next generation cellular networks due to its significant performance improvement without requiring additional time or frequency resources.

For multi-user MIMO broadcast channels (BC), the channel capacity has been studied in [6]–[9]. It was shown in [6] that dirty paper coding (DPC) achieves the capacity region of the MIMO-BC. In [7], an efficient iterative water-filling method was proposed to compute the sum-rate of the MIMO-BC with sum

Digital object identifier: 10.1109/JCN.2016.000122

transmit power constraint. Also, several linear precoding methods were proposed in [8] and [9] to reduce the complexity of the DPC.

In addition, to mitigate inter-cell interference due to frequency reuse, base station (BS) cooperation, also known as network MIMO or coordinated beamforming, has drawn much attentions [10]-[16]. A coordinated beamforming scheme was introduced in [10] to achieve the Pareto boundary of user rate tuples for a multiple-input single-output (MISO) interference channel. A distributed beamforming technique in [11] for a MISO channel is sub-optimal but has much lower complexity, while a gradient ascent scheme in [12] obtains a locally optimal linear precoder for MIMO interference channels. In [13], a successive convex approximation linear precoder was employed to decompose the original non-convex problem into a sequence of convex subproblems, which can be solved independently. Besides, the sum-rate maximization problem for interfering broadcast channels (IFBC) was addressed in [14] by exploiting the relationship between the rate and the minimum mean squared error (MMSE). Also, a distributed beamforming based on a high signal-to-interference-plus-noise (SINR) approximation was presented for the MISO-IFBC [15], and an asymptotic approach based on random matrix theory was introduced in [16].

Aforementioned references mainly focus on rate or spectral efficiency optimization. Recently, energy efficient wireless communications, which are often called green communications, have attracted increasing attentions [17]–[22]. Since enormous energy consumption in wireless communication equipments makes a negative impact on the environment, pursuing high energy efficiency (EE) has become a key issue for future wireless communications. Generally, the EE is defined as the sum-rate divided by the total power consumption.

The EE optimization problem for a single cell MIMO-BC has been solved in [17] by the fractional programming (FP) theory and the multiple access channel (MAC)-BC duality. The optimal precoding design for the EE maximization in a cognitive radio MIMO-BC was introduced in [18]. The adaptive transmission methods in [19] maximize the EE performance. Also, the EE optimization was studied in orthogonal frequency division multiplexing systems [20] and distributed antenna systems [21]. Recently, the sum-EE optimization for multi-cell coordinated joint transmission systems was investigated in [22] with linear precoding designs in IFBC, where a local optimal solution is iteratively found based on the weighted MMSE minimization problem.

In this paper, we examine the weighted sum EE optimization

1229-2370/16/\$10.00 © 2016 KICS

This paper is specially handled by EICs and Division Editors with the help of three anonymous reviewers in a fast manner.

This work was supported in by the National Research Foundation (NRF) funded by the Ministry of Science, ICT and Future Planning, Korean Government, under Grant 2014R1A2A1A10049769. The work of K.-J. Lee was supported through the NRF funded by the Ministry of Science, ICT & Future Planning, Korean Government, under Grant 2013R1A1A1060503 and grant 2014K1A3A1A09063284.

X. Gui was with the School of Electrical Engineering, Korea University, Seoul 136-701, Korea. Now he is with Samsung R&D Center-Beijing, Beijing, China, email: guixinjxn1234@163.com.

K.-J. Lee is with the Department of Electronics and Control Engineering, Hanbat National University, Daejeon, Korea, email: kyoungjae@hanbat.ac.kr.

J. Jung and I. Lee are with the the School of Electrical Engineering, Korea University, Seoul 136-701, Korea, email: {jhnjung, inkyu}@korea.ac.kr.

for a coordinated multi-cell IFBC using non-linear precoding schemes. Since DPC achieves the optimal performance in single cell systems [17], [18], we extend the study of DPC onto a multicell system with interference coordination. Although DPC has been considered as a theoretical benchmark due to its high complexity implementation, our study on DPC for multicell network enables to realize extra performance gains over the linear precoding schemes. The achievable EE performance by DPC allows us to reveal a theoretical limit.

Moreover, by using the weighted sum energy efficiency as the optimization criterion, we can satisfy heterogeneous requirements from different cells, which is more difficult to solve due to its sum-of-ratio form. To handle the non-convex weighted sum EE problem, we first apply an iterative linear approximation method and the FP algorithm to transform the original problem into a sequence of simpler subproblems. By employing the linear approximation method, each cell aims to maximize its own EE function while minimizing interference to other cells. As a result, the original problem can be solved by sequentially addressing each subproblem. After that, convex MAC problems are obtained by adopting the BC-MAC duality, and we propose a gradient descent method to solve the MAC problem.

Also, for the case of equal weights for user within a cell, we introduce a coordinated iterative water-filling method by applying the Karush-Kuhn-Tucker (KKT) optimality conditions. The coordinated water-filling solution based on the MAC problem achieves the globally optimal EE for each original BC subproblem, and the computational complexity is greatly reduced compared to an interior-point method. Also, each BS can iteratively update its precoding matrices with a small amount of information exchanges among the BSs. Finally, simulation results demonstrate the convergence and the performance advantage of our proposed algorithm. We show that the performance gain is about 20 percent compared to the system with linear precoding.

The rest of this paper is organized as follows: Section II presents the system model of a multi-cell IFBC. In Section III, the proposed EE optimization algorithm is analyzed. Section IV illustrates the simulation results. Finally, Section V concludes our paper.

II. SYSTEM MODEL

Since SNMPv2 is not as widespread as SNMPv1, which does not scale to large complex networks, we use two different solutions for gathering MIB-II variables on managed elements: A mobile agent-based solution and an SNMP based one.

Consider a downlink multi-user MISO system with L cells operating on the same frequency, as shown in Fig. 1. At each cell, a BS equipped with N_T antennas sends signals simultaneously to K mobile users with a single antenna. Denoting the kth user in the lth cell as user l_k , the received signal at user l_k is expressed as

$$y_{l_k} = \mathbf{h}_{l,l_k}^{\mathrm{H}} \mathbf{x}_{l_k} + \mathbf{h}_{l,l_k}^{\mathrm{H}} \sum_{j=1, \ j \neq k}^{K} \mathbf{x}_{l_j} + \sum_{i=1, \ i \neq l}^{L} \mathbf{h}_{i,l_k}^{\mathrm{H}} \sum_{j=1}^{K} \mathbf{x}_{i_j} + n_{l_k},$$
(1)

where $\mathbf{h}_{i,l_k} \in \mathbf{C}^{N_T \times 1}$ is the frequency-flat channel vector between the *i*th BS to user l_k including both large-scale fading



Fig. 1. System model of a multi-cell IFBC.

and small-scale fading, \mathbf{x}_{l_j} indicates the transmitted signal from the *l*th BS to its connected *j*th user, and n_{l_k} represents the additive white Gaussian noise with zero mean and unit variance. We can see from (1) that y_{l_k} contains both intra-cell interference $\mathbf{h}_{l,l_k}^{\mathrm{H}} \sum_{j=1, j \neq k}^{K} \mathbf{x}_{l_j}$ and inter-cell interference $\sum_{i=1, i \neq l}^{L} \mathbf{h}_{i,l_k}^{\mathrm{H}} \sum_{j=1}^{K} \mathbf{x}_{i_j}$. In addition, the transmitted signal \mathbf{x}_{l_j} for the *j*th user in the *l*th cell is given as $\mathbf{x}_{l_j} = \mathbf{v}_{l_j} s_{l_j}$, where $\mathbf{v}_{l_j} \in \mathbf{C}^{N_T \times 1}$ denotes the beamforming vector and s_{l_j} is the transmit complex signal. In this paper, we assume that a BS applies DPC to achieve the Shannon capacity region at each cell. Since the maximum throughput for the BC can be attainted by arbitrary encoding order [6] in single cell case, without loss of generality, it is assumed that \mathbf{x}_{l_k} in the *l*th BS is encoded in the order of user indices. Thus, the *k*th user does not suffer interference from user 1 to user k - 1.

Accordingly, the achievable rate is given by

$$r_{l_k}^{BC} = \log_2 \left(\frac{R_{l_k} + \mathbf{h}_{l,l_k}^{\mathrm{H}} \left(\sum_{j=k}^{K} \mathbf{Q}_{l_j} \right) \mathbf{h}_{l,l_k}}{R_{l_k} + \mathbf{h}_{l,l_k}^{\mathrm{H}} \left(\sum_{j=k+1}^{K} \mathbf{Q}_{l_j} \right) \mathbf{h}_{l,l_k}} \right), \quad (2)$$

where $\mathbf{Q}_{l_j} = \mathbf{v}_{l_j} \mathbf{v}_{l_j}^{\mathrm{H}}$ is the transmit covariance matrix, and R_{l_k} defines the variance of the total inter-cell interference plus additive noise as

$$R_{l_k} = \sum_{i=1, i \neq l}^{L} \mathbf{h}_{i, l_k}^{\mathrm{H}} \left(\sum_{j=1}^{K} \mathbf{Q}_{i_j} \right) \mathbf{h}_{i, l_k} + N_{l_k},$$

where N_{l_k} is the covariance of addictive noise term.

Our goal is to design an energy efficient transmission scheme for a multi-cell MISO system. The EE formula for the lth cell is defined as the ratio of the weighted sum-rate to the per-cell power consumption which is written by [22]

$$e_l\left(\mathbf{Q}_l\right) = \frac{\sum\limits_{k=1}^{K} \alpha_{l_k} r_{l_k}^{BC}}{\eta_l \sum\limits_{k=1}^{K} \operatorname{tr}\left(\mathbf{Q}_{l_k}\right) + N_T P_{\mathrm{d}} + P_{\mathrm{s}}},\tag{3}$$

where α_{l_k} represents the weight of user l_k , η_l is the inefficiency of the power amplifier, $N_T P_d$ denotes the dynamic power consumption proportional to the number of active transmit antennas, P_s accounts for the static power independent of N_T , such as circuit power consumption of radio frequency chains and the baseband processing. Note that since the power consumption at the user is very small compared with BS, similar to many related references, we ignore the user consumed power in this paper.

For notational simplicity, we define $\mathbf{Q}_l = {\{\mathbf{Q}_{l_k}\}}_{k=1}^{K}$ as the transmit covariance matrix set at the *l*th cell, and $\mathbf{Q} = {\{\mathbf{Q}_l\}}_{l=1}^{L}$ as the set of the transmit covariance of all cells. Then, the weighted sum EE can be obtained as

$$\max_{\mathbf{Q}} \sum_{l=1}^{L} \omega_l \ e_l \left(\mathbf{Q}_l \right)$$

s.t.
$$\sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{Q}_{l_k} \right) \le P_l, \ \forall l,$$
 (4)

where ω_l is a nonnegative weight coefficient, P_l indicates sum power constraint at the *l*th BS. Note that the formulated problem (4) is non-convex since the optimization variables in (2) and (3) are jointly coupled. In addition, the weighted sum of fractions makes the problem even more intractable. In the next section, we focus on solving (4) by using some transformations and optimization methods.

III. MULTI-CELL EE BEAMFORMING DESIGN

In this section, we aim to solve the fractional problem (4). We first employ an iterative linear approximation method and the FP algorithm to transform the original problem into a sequence of simpler subproblems. For each problem, convex MAC problems are obtained by applying the BC-MAC duality, and we propose a gradient descent method to solve the MAC problem. Also, as a special case of equal weights for users within a cell, a closed form solution is derived based on KKT optimality conditions. After that, an iterative EE optimization algorithm is described. Moveover, we will show that the proposed algorithm allows distributed implementation and also provide algorithm complexity analysis.

A. Problem Formulation

Since the objective function in (4) is in fractional form, by adopting the fractional programming method in [22] and [23] and introducing a set of auxiliary variables β_l , we can rewrite (4) into an equivalent form as

$$\max_{\mathbf{Q}, \beta} \sum_{l=1}^{L} \omega_l \beta_l$$

s.t.
$$\frac{h_l(\mathbf{Q}_l)}{g_l(\mathbf{Q}_l)} \ge \beta_l, \forall l$$

$$\sum_{k=1}^{K} \operatorname{tr}(\mathbf{Q}_{l_k}) \le P_l, \forall l,$$
 (5)

where we define $\beta = \{\beta_l\}_{l=1}^L$, $h_l(\mathbf{Q}_l) = \sum_{k=1}^K \alpha_{l_k} r_{l_k}^{BC}$ and $g_l(\mathbf{Q}_l) = \eta_l \sum_{k=1}^K \operatorname{tr}(\mathbf{Q}_{l_k}) + N_T P_d + P_s$, and the first inequality set represents the EE constraints. We then introduce the follow-

set represents the EE constraints. We then introduce the following lemma to decompose (5) into L parallel subproblems.

Lemma 1: Denoting (\mathbf{Q}^*, β^*) as a solution of problem (5), $(\mathbf{Q}_1^*, \dots, \mathbf{Q}_L^*)$ satisfies the KKT condition of the following L subproblems

$$\max_{\mathbf{Q}_{l}} \lambda_{l} \left(h_{l} \left(\mathbf{Q}_{l} \right) - \beta_{l} g_{l} \left(\mathbf{Q}_{l} \right) \right) - \sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{A}_{l} \mathbf{Q}_{l_{k}} \right)$$

s.t.
$$\sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{Q}_{l_{k}} \right) \leq P_{l}, \ \forall j,$$
 (6)

where A_l is given by

$$\mathbf{A}_{l} = \sum_{i=1, i \neq l}^{L} \lambda_{i} \sum_{j=1}^{K} \mathbf{h}_{l,i_{j}} \left[\frac{1}{\ln 2 \left(R_{i_{j}} + \sum_{k=j+1}^{K} \mathbf{h}_{i,i_{j}}^{\mathrm{H}} \mathbf{Q}_{i_{k}} \mathbf{h}_{i,i_{j}} \right)} - \frac{1}{\ln 2 \left(R_{i_{j}} + \sum_{k=j}^{K} \mathbf{h}_{i,i_{j}}^{\mathrm{H}} \mathbf{Q}_{i_{k}} \mathbf{h}_{i,i_{j}} \right)} \right] \mathbf{h}_{l,i_{j}}^{\mathrm{H}}.$$
(7)

Also, there exists $\lambda = \{\lambda_l\}_{l=1}^{L}$ such that $(\mathbf{Q}_1^*, \dots, \mathbf{Q}_L^*)$ also satisfies the following equations

$$\lambda_{l}^{*} = rac{\omega_{l}}{g_{l}\left(\mathbf{Q}_{l}^{*}
ight)} \text{ and } \beta_{l}^{*} = rac{h_{l}\left(\mathbf{Q}_{l}^{*}
ight)}{g_{l}\left(\mathbf{Q}_{l}^{*}
ight)}, \forall l$$

On the contrary, if $(\mathbf{Q}_1^*, \dots, \mathbf{Q}_L^*)$ are the solutions of the *L* subproblems in (6) and satisfy (7), then (\mathbf{Q}^*, β^*) also fulfills the KKT condition of problem (5).

Proof: See Appendix A. \Box

Problem (6) essentially maximizes the objective function in Problem (5) with respect to \mathbf{Q}_l , except that the weighted sum EE of other cells is approximated to the first order at the point $\overline{\mathbf{Q}}_l$. Since the KKT conditions of the above *L* subproblems are exactly the same as those of problem (5), we can obtain \mathbf{Q}^* that satisfies (23)–(27) by iteratively solving the above subproblems until arriving at a stationary point. Note that the last term of the objective function in (6) can be interpreted as the interference cost that the *l*th BS pays for inter-cell interference induced to other cells as in [13]. This discourages selfish behavior of cell *l*, which would otherwise maximize its own EE measurement.

Next, it remains to determine the optimal solution for each subproblem in (6), which can be transformed by the Lagrange dual decomposition method [24]. The Lagrangian of problem (6) is given by

$$L_{l} \left(\mathbf{Q}_{l}, \ \mu_{l} \right) = \lambda_{l} h_{l} \left(\mathbf{Q}_{l} \right)$$
$$- \sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{Q}_{l_{k}} \left(\left(\lambda_{l} \beta_{l} \eta_{l} + \mu_{l} \right) \mathbf{I}_{N_{T}} + \mathbf{A}_{l} \right) \right) + \mu_{l} P_{l} - N_{T} P_{s} - P_{d},$$

$$\tag{8}$$

where μ_l is the Lagrange multiplier corresponding to the power constraint. The Lagrange dual function is then written by

$$D_{l}(\mu_{l}) = \max_{\mathbf{Q}_{l} \succeq \mathbf{0}} \hat{L}_{l}(\mathbf{Q}_{l}, \mu_{l}).$$
(9)

The optimal solution can be found by solving the dual problem

$$\min_{\mu_l \ge 0} D_l\left(\mu_l\right).$$

To compute the dual problem, we first focus on the maximization of the dual function for a given μ_l . By deleting unrelated terms, we get

$$\max_{\mathbf{Q}_l \succeq \mathbf{0}} \lambda_l h_l \left(\mathbf{Q}_l \right) - \sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{Q}_{l_k} \left(\left(\lambda_l \beta_l \eta_l + \mu_l \right) \mathbf{I}_{N_T} + \mathbf{A}_l \right) \right).$$
(10)

In the following, we will introduce a BC-MAC duality to solve problem (10).

B. BC-MAC Duality

Problem (10) is still non-convex on \mathbf{Q}_l , and thus it cannot be solved by applying the Hadamard's inequality as in [13]. Instead, we adopt some transformations and the BC-MAC duality property to address problem (10). First, by setting new variables

$$\begin{split} \tilde{\mathbf{Q}}_{l_k} = & \left((\lambda_l \beta_l \eta_l + \mu_l) \, \mathbf{I}_{N_T} + \mathbf{A}_l \right)^{\frac{1}{2}} \mathbf{Q}_{l_k} ((\lambda_l \beta_l \eta_l + \mu_l) \, \mathbf{I}_{N_T} + \mathbf{A}_l)^{\frac{1}{2}}, \\ \tilde{\mathbf{h}}_{l,l_k} = & \left(R_{l_k} \right)^{-\frac{1}{2}} \mathbf{h}_{l,l_k} ((\lambda_l \beta_l \eta_l + \mu_l) \, \mathbf{I}_{N_T} + \mathbf{A}_l)^{-\frac{1}{2}}, \end{split}$$

the achievable rate in (2) can be rewritten as

$$r_{l_k}^{BC} = \log_2 \left(\frac{1 + \tilde{\mathbf{h}}_{l,l_k}^{\mathrm{H}} \left(\sum_{j=k}^{K} \tilde{\mathbf{Q}}_{l_j} \right) \tilde{\mathbf{h}}_{l,l_k}}{1 + \tilde{\mathbf{h}}_{l,l_k}^{\mathrm{H}} \left(\sum_{j=k+1}^{K} \tilde{\mathbf{Q}}_{l_j} \right) \tilde{\mathbf{h}}_{l,l_k}} \right),$$

and the problem in (10) is equivalent to the following problem as

$$\max_{\tilde{\mathbf{Q}}_l \succeq \mathbf{0}} \lambda_l h_l \left(\tilde{\mathbf{Q}}_l \right) - \sum_{k=1}^{K} \operatorname{tr} \left(\tilde{\mathbf{Q}}_{l_k} \right).$$
(11)

We then introduce the following proposition.

Proposition 1: According to the property of the BC-MAC duality [6], [17], problem (11) is equivalent to its dual MAC problem as

$$\max_{\left\{P_{1}^{M},\dots,P_{K}^{M}\right\}} \tilde{f}_{l}\left(P_{1}^{M},\dots,P_{K}^{M}\right) = \\
\max_{P_{k}^{M}} \sum_{i=1}^{K} \Delta_{l_{i}} \log_{2} \left| \mathbf{I}_{N_{T}} + \sum_{j=1}^{i} P_{j}^{M} \tilde{\mathbf{h}}_{l,l_{j}} \tilde{\mathbf{h}}_{l,l_{j}}^{H} \right| - \sum_{k=1}^{K} P_{k}^{M} \tag{12}$$

where P_k^{M} denotes the transmit power at the *k*th user for the MAC uplink channel, and we have $\Delta_{l_i} = \alpha_{l_i} - \alpha_{l_{i+1}}$ and $\alpha_{l_{K+1}} = 0$. If there exists the optimal solution $\tilde{\mathbf{Q}}_l^*$ for problem (11), we can also find a set of $P_k^{\mathrm{M}*}$ satisfying $\sum_{k=1}^{K} \operatorname{tr} \left(\tilde{\mathbf{Q}}_{l_k}^* \right) = \sum_{k=1}^{K} P_k^{\mathrm{M}*}$ by solving (12), which obtains the maximum objective

value in (11), and vice versa.

According to Proposition 1, instead of solving the non-convex problem (11), we can address problem (12). Since problem (12) is convex, we can apply conventional convex optimization methods such as the interior point scheme. However, convex optimization becomes computationally expensive due to the special structure in (12). Thus, we propose a gradient descent method [12] for this problem. The gradient of $\tilde{f}_l(P_1^M, \dots, P_K^M)$ with respect to P_k^M equals

$$\nabla_{k} = \frac{\partial \tilde{f}_{l}\left(P_{1}^{\mathrm{M}}, \cdots, P_{K}^{M}\right)}{\partial P_{k}^{\mathrm{M}}} = \sum_{j=k}^{K} \frac{\Delta_{j}\left(\tilde{\mathbf{h}}_{l,l_{j}}^{\mathrm{H}} \tilde{\mathbf{h}}_{l,l_{j}}\right)}{\ln 2\left(1 + \sum_{i=1}^{j} P_{i}^{\mathrm{M}} \tilde{\mathbf{h}}_{l,l_{i}}^{\mathrm{H}} \tilde{\mathbf{h}}_{l,l_{i}}\right)} - 1.$$
(13)

Let $\nabla_k(n)$ and $P_k^M(n)$ be ∇_k and P_k^M at the *n* th iteration step, respectively. Then $P_k^M(n+1)$ can be updated according to

$$P_{k}^{M}(n+1) = \left[P_{k}^{M}(n) + t\nabla_{k}(n)\right]^{+},$$
(14)

where $[x]^+ = \max(0, x)$ and t is the step size. As long as the step size t is small enough, the MAC problem at the (n + 1) th iteration $\tilde{f}_l(P_1^M(n+1), \dots, P_K^M(n+1))$ is greater than (or equal to which indicates the optimality) $\tilde{f}_l(P_1^M(n), \dots, P_K^M(n))$. To efficiently determine the step size t, we employ Armijo's Rule which provides provable convergence [12].

C. Equal Weight Case

In this subsection, we consider a special case of equal weights for users within a cell, which results in a closed-form power allocation solution. Setting $\alpha_{l_1} = \cdots = \alpha_{l_K} = \alpha_l$, (12) equals

$$\max_{\left\{P_{1}^{M},\dots,P_{K}^{M}\right\}} \lambda_{l}\alpha_{l}\log_{2} \left| \mathbf{I}_{N_{T}} + \sum_{k=1}^{K} P_{k}^{M}\tilde{\mathbf{h}}_{l,l_{k}}\tilde{\mathbf{h}}_{l,l_{k}}^{H} \right| - \sum_{k=1}^{K} P_{k}^{M}.$$
(15)

By using the transformation in [7], we can redefine for given $(P_1^M, \dots, P_{k-1}^M, P_{k+1}^M, \dots, P_K^M)$ as

$$\max_{P_k^M} \lambda_l \alpha_l \log_2 \left| \mathbf{I}_{N_T} + \tilde{\mathbf{g}}_{l,l_k} P_k^M \tilde{\mathbf{g}}_{l,l_k}^{\mathrm{H}} \right| + \lambda_l \alpha_l \log_2 |\mathbf{Z}_l| - \sum_{k=1}^K P_k^M$$
(16)

where $\tilde{\mathbf{g}}_{l,l_k} = \left(\mathbf{I}_{N_T} + \sum_{j \neq k}^{K} \tilde{\mathbf{h}}_{l,l_j} P_j^M \tilde{\mathbf{h}}_{l,l_j}^{\mathrm{H}}\right)^{-1/2} \tilde{\mathbf{h}}_{l,l_j}$ and $\mathbf{Z}_l = \mathbf{I}_{N_T} + \sum_{j \neq k} \tilde{\mathbf{h}}_{l,l_j} P_j^M \tilde{\mathbf{h}}_{l,l_j}^{\mathrm{H}}$. We can solve (16) by checking the first-order KKT condition as

$$P_k^M = \left(\frac{\lambda_l \alpha_l}{\ln 2} - \frac{1}{\tilde{\mathbf{g}}_{l,l_k} \tilde{\mathbf{g}}_{l,l_k}^{\mathrm{H}}}\right)^+.$$
 (17)

Since problem (16) is convex, the power allocation solution can be iteratively found by applying (17) and the block-coordinate ascent algorithm. Moreover, we can see that the water-filling solution (17) is different from [7] such that (17) contains both the EE auxiliary variable β_l and the interference pricing matrix \mathbf{A}_l in $\tilde{\mathbf{g}}_{l,l_k}$. Thus, both parameters affect the power allocation results at each update step. We will call this method for the MAC problem (15) as a coordinated EE iterative water-filling method. Note that the closed-form solution (17) has much lower complexity than the gradient descent method for problem (12) with unequal weights.

After computing the optimal $P_k^{M^*}$ for the MAC problem (15), the optimal solution $\tilde{\mathbf{Q}}_l^*$ for the BC problem (11) can also be obtained by using the MAC-BC mapping [6]. Thus, the optimal solution for problem (10) can be expressed as

$$\mathbf{Q}_{l_{k}}^{*} = \\ \left(\left(\lambda_{l} \beta_{l} \eta_{l} + \mu_{l} \right) \mathbf{I}_{N_{T}} + \mathbf{A}_{l} \right)^{-\frac{1}{2}} \tilde{\mathbf{Q}}_{l_{k}}^{*} \left(\left(\lambda_{l} \beta_{l} \eta_{l} + \mu_{l} \right) \mathbf{I}_{N_{T}} + \mathbf{A}_{l} \right)^{-\frac{1}{2}}$$
(18)

due to the equivalence between (10) and (11).

Next, it remains to solve the dual problem (9) with a given \mathbf{Q}^* . Since the Lagrangian function $D_l(\mu_l)$ is not necessarily differentiable, a sub-gradient based method is employed for this problem. Denote $\bar{\mathbf{Q}}_l$ and $\hat{\mathbf{Q}}_l$ as the optimal covariance matrices in (9) with $\mu_l = \bar{\mu}_l$, and $\mu_l = \hat{\mu}_l$, respectively, $D_l(\bar{\mu}_l)$ can be expressed as

$$\begin{split} D_{l}\left(\bar{\mu}_{l}\right) &= \max_{\mathbf{Q}_{l}\succ\mathbf{0}}\lambda_{l}h_{l}\left(\mathbf{Q}_{l}\right) \\ &- \sum_{k=1}^{K}\operatorname{tr}\left(\mathbf{Q}_{l_{k}}\left(\left(\lambda_{l}\beta_{l}\eta_{l}+\bar{\mu}_{l}\right)\mathbf{I}_{N_{T}}+\mathbf{A}_{l}\right)\right)+\bar{\mu}_{l}P_{l} \\ &= \lambda_{l}h_{l}\left(\bar{\mathbf{Q}}_{l}\right) - \sum_{k=1}^{K}\operatorname{tr}\left(\bar{\mathbf{Q}}_{l_{k}}\left(\left(\lambda_{l}\beta_{l}\eta_{l}+\bar{\mu}_{l}\right)\mathbf{I}_{N_{T}}+\mathbf{A}_{l}\right)\right)+\bar{\mu}_{l}P_{l} \\ &\geq \lambda_{l}h_{l}\left(\hat{\mathbf{Q}}_{l}\right) - \sum_{k=1}^{K}\operatorname{tr}\left(\hat{\mathbf{Q}}_{l_{k}}\left(\left(\lambda_{l}\beta_{l}\eta_{l}+\bar{\mu}_{l}\right)\mathbf{I}_{N_{T}}+\mathbf{A}_{l}\right)\right)+\bar{\mu}_{l}F \\ &= \lambda_{l}h_{l}\left(\hat{\mathbf{Q}}_{l}\right) - \sum_{k=1}^{K}\operatorname{tr}\left(\hat{\mathbf{Q}}_{l_{k}}\left(\left(\lambda_{l}\beta_{l}\eta_{l}+\hat{\mu}_{l}\right)\mathbf{I}_{N_{T}}+\mathbf{A}_{l}\right)\right) \\ &+ \sum_{k=1}^{K}\operatorname{tr}\left(\hat{\mathbf{Q}}_{l_{k}}\left(\left(\lambda_{l}\beta_{l}\eta_{l}+\hat{\mu}_{l}\right)\mathbf{I}_{N_{T}}+\mathbf{A}_{l}\right)\right) \\ &- \sum_{k=1}^{K}\operatorname{tr}\left(\hat{\mathbf{Q}}_{l_{k}}\left(\left(\lambda_{l}\beta_{l}\eta_{l}+\bar{\mu}_{l}\right)\mathbf{I}_{N_{T}}+\mathbf{A}_{l}\right)\right)+\bar{\mu}_{l}P_{l} \\ &= D_{l}\left(\hat{\mu}_{l}\right) + \left(P_{l} - \sum_{k=1}^{K}\operatorname{tr}\left(\hat{\mathbf{Q}}_{l_{k}}\right)\right)\left(\bar{\mu}_{l} - \hat{\mu}_{l}\right). \end{split}$$

Thus, $P_l - \sum_{k=1}^{K} \operatorname{tr} (\mathbf{Q}_{l_k}^*)$ can be chosen as the sub-gradient of $D_l(\mu_l)$, where $\mathbf{Q}_{l_k}^*$ for $k = 1, \dots, K$ is the optimal covariance matrix for a fixed μ_l in (9).

In each iterative step, μ_l is updated according to the sub-gradient direction. The value of μ_l should increase if $\sum_{k=1}^{K} \operatorname{tr}(\mathbf{Q}_{l_k}) < P_l$, and vice versa. Moreover, a bisection method can be efficiently applied to find the optimal μ_l^* . If such μ_l^* does not exist, we set $\mu_l^* = 0$. We summarize the proposed

Algorithm 1 Iterative power allocation

algorithm for problem (6) in Algorithm 1 below,

Initialize $\mu_{l,\min}$ and $\mu_{l,\max}$. **Repeat** Set $\mu_l = (\mu_{l,\min} + \mu_{l,\max})/2$. **Repeat** for $k = 1, 2, \dots, K$ Update P_k^M according to (17). end

Until convergence

Calculate $\tilde{\mathbf{Q}}_{l}^{*}$ according to the MAC-BC mapping, and obtain \mathbf{Q}_{l}^{*} according to (18).

If
$$\sum_{k=1}^{K} \operatorname{tr}(\mathbf{Q}_{l_k}) > P_l$$
, then $\mu_{l,\min} = \mu_l$, else $\mu_{l,\max} = \mu_l$.
Until $|\mu_{l,\max} - \mu_{l,\min}|$ is small enough

which adopts the Lagrange dual decomposition method for problem (6). Since there is generally a duality gap between the dual problem and the primal problem for a non-convex problem, a question remains whether the proposed algorithm can converge to a global optimal solution. We then introduce the following proposition to show that it is indeed the case.

Proposition 2: The optimal transmit covariance matrices \mathbf{Q}_l^* with the optimal μ_l^* obtained by Algorithm1 achieves the global optimal solution of problem (6).

Proof: See Appendix B.

In Algorithm 1, the MAC problem is addressed by employing the EE iterative water-filling method, which has much lower complexity than the interior-point method. Then, \mathbf{Q}_l^* is identified by the bisection method which converges fast. Thus, we can see that Algorithm 1 is efficient to optimize each subproblem in (6). It is interesting to note that our proposed algorithm can easily be extended to multi-cell multi-user MIMO systems by finding the optimal MAC covariance matrix in problem (16) for each user with the EE iterative method similarly as in [17]. Then we can get the optimal covariance for the BC by the MAC-BC P_l mapping for the MIMO case as that in [6]. Next, the following lemma proves the convergence of Algorithm 1.

Lemma 2: The optimization algorithm for problem (6) carried out at each cell always improves the network sum EE measurements $\sum_{l=1}^{L} \tilde{e}(\mathbf{Q}_l, \mathbf{Q}_{-l})$, where $\tilde{e}(\mathbf{Q}_l, \mathbf{Q}_{-l}) = \lambda_l (h_l(\mathbf{Q}_l) - \beta_l g_l(\mathbf{Q}_l))$ and $\mathbf{Q}_{-l} = {\mathbf{Q}_1, \dots, \mathbf{Q}_{l-1}, \mathbf{Q}_{l+1}, \dots, \mathbf{Q}_L}$.

Proof: See Appendix C.

As indicated in Lemma 2, updating the BC precoding covariance matrix independently at each BS always increases the multi-cell sum EE. Thus, this iterative algorithm can be performed across all BSs until the whole system reaches a static state. Moreover, this Gauss-Seidel iteration [13] can converge to the KKT point of problem (6). The auxiliary variables λ and β are updated by a Newton-like method. We consider an update of the BS precoding for the BC problem in Algorithm 1 as an inner loop, while the update of λ , β , and \mathbf{A}_l is referred to as an outer loop. The overall algorithm is summarized in Algorithm2 as below.

Algorithm 2 Iterative EE optimization
Initialize $\left\{ \mathbf{Q}_{l}^{(0)} \right\}_{l=1}^{L}$ satisfying transmit power constraint.
Set $\lambda_l = \frac{\omega_l}{g_l(\mathbf{Q}_l^{(0)})}$ and $\beta_l = \frac{h_l(\mathbf{Q}_l^{(0)})}{g_l(\mathbf{Q}_l^{(0)})}$, $\forall l$.
Repeat
for $l = 1, 2, \dots, L$
Update A_l according to (7).
Compute $\left\{\mathbf{Q}_{i_j}\right\}_{j=1}^K$ according to Algorithm 1.
Obtain λ_l and β_l as $\lambda_l = \frac{\omega_l}{g_l(\mathbf{Q}_l^*)}$ and $\beta_l = \frac{h_l(\mathbf{Q}_l^*)}{g_l(\mathbf{Q}_l^*)}$.
end
Until convergence

Remark 1: In some scenarios, we may need to satisfy additional rate constraints while maximizing the energy efficiency. It is worthwhile to mention that our algorithm is also able to maximize the weighted sum energy efficiency with the rate requirements of individual cells. Let the required minimum rate at the *l*th cell be C_l . By adding the minimum rate constraint condition to the original problem (5), it can be decomposed into L parallel subproblems by applying Lemma 1 similarly, where each subproblem can be written as

$$\max_{\mathbf{Q}_{l}} \lambda_{l} \left(h_{l} \left(\mathbf{Q}_{l} \right) - \beta_{l} g_{l} \left(\mathbf{Q}_{l} \right) \right) - \sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{A}_{l} \mathbf{Q}_{l_{k}} \right)$$
(19)

s.t.
$$\sum_{k=1}^{K} \operatorname{tr}\left(\mathbf{Q}_{l_{k}}\right) \leq P_{l}, \ \forall j,$$

$$(20)$$

$$h_l\left(\mathbf{Q}_l\right) \ge C_l \;\forall j. \tag{21}$$

Considering that problem (19) is feasible, the constraints of (20) and (21) correspond to an upper bound and a lower bound of the feasible sum transmit power given by P_l and \bar{P}_l , respectively. Setting the optimal sum power to be $P_l^* = \sum_{k=1}^{K} \operatorname{tr} (\mathbf{Q}_{l_k}^*)$, if the condition $P_l^* \leq \bar{P}_l \leq P_l$ is satisfied, i.e., the constraint (21) is active, problem (19) is equivalent to a sum transmit power minimization problem under sum-rate constraint [Section II, 17].

By employing the transformation in (11) and the BC-MAC duality similarly in Proposition 1, the equivalent MAC problem can be written as

$$\begin{array}{l} \min_{\left\{P_{1}^{\mathrm{M}}, \cdots, P_{K}^{\mathrm{M}}\right\}} \sum_{k=1}^{K} P_{k}^{\mathrm{M}} \\ \text{s.t.} \sum_{i=1}^{K} \Delta_{l_{i}} \log_{2} \left| \mathbf{I}_{N_{T}} + \sum_{j=1}^{i} P_{j}^{\mathrm{M}} \tilde{\mathbf{h}}_{l, l_{j}} \tilde{\mathbf{h}}_{l, l_{j}}^{\mathrm{H}} \right| = C_{l}. \end{array}$$

$$(22)$$

Problem (22) can be effectively solved by applying the iterative algorithm in [25] (a detailed algorithm is omitted due to page limit). After computing the optimal $P_k^{M^*}$ for the above MAC problem, the optimal solution $\tilde{\mathbf{Q}}_l^*$ for the BC problem (19) can also be obtained by using the MAC-BC mapping.

D. Sum Rate Maximization Algorithm

For comparison purpose, we introduce the problem of sumrate maximization with transmit power constraint at each BS as

$$\max_{\mathbf{Q}} \sum_{l=1}^{L} \omega_{l} \bar{R}_{l}$$
s.t.
$$\sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{Q}_{l_{k}} \right) \leq P_{l}, \; \forall l,$$
(23)

where $\bar{R}_l = \sum_{k=1}^{K} r_{l_k}^{BC}$. Similar to the method in the previous sub-section, we can address a BC problem at each cell by determining the dual MAC problem. Let us define $\hat{\mathbf{h}}_{l,l_k}$ as $\hat{\mathbf{h}}_{l,l_k} = (R_{l_k})^{-\frac{1}{2}} \mathbf{h}_{l,l_k} (\tilde{\tau} \mathbf{I}_{N_T} + \mathbf{A}_l)^{-\frac{1}{2}}$, where $\tilde{\tau}$ denotes a Lagrange variable associated with the total transmit power constraint.

The dual MAC problem can be solved similarly as the coordinated iterative EE water-filling algorithm to obtain the transmit power of each user as

$$P_k^{\mathrm{M}}(\tilde{\tau}) = \left(\frac{\omega_l}{\ln 2} - \frac{1}{\hat{\mathbf{g}}_{l,l_k} \hat{\mathbf{g}}_{l,l_k}^{\mathrm{H}}}\right)^+, \tag{24}$$

where
$$\hat{\mathbf{g}}_{l,l_k} = \left(\mathbf{I}_{N_T} + \sum_{j \neq k}^{K} P_j^{\mathrm{M}} \hat{\mathbf{h}}_{l,l_j} \hat{\mathbf{h}}_{l,l_j}^{\mathrm{H}} \right)^{-1/2} \hat{\mathbf{h}}_{l,l_k}$$
. Note that

 $\tilde{\tau}$ is also calculated by the bisection method. The whole algorithm is summarized in Algorithm 3.

Note that the convergence of the WSR algorithm can be proved similarly as in Lemma 2, since an update of the transmit BS covariance matrix Q_l always improves the network sum rate. Thus, the convergence is guaranteed due to the bounded transmit power.

Algorithm 3 Weighted sum-rate (WSR) algorithm
Initialize $\left\{ \mathbf{Q}_{l}^{(0)} \right\}_{l=1}^{L}$ satisfying transmit power constraint.
Repeat
for $l = 1, 2,, L$
Update \mathbf{A}_l according to (7).
Update $\left\{\mathbf{Q}_{r_j}\right\}_{j=1}^{K}$ by Algorithm 1 with (20).
end
Until convergence

E. Distributed Implementation

The proposed algorithm can be implemented in a distributed manner at each cell. We first assume that each BS has local CSI, i.e., BS *l* knows only the channel vectors $\mathbf{h}_{l,l_k}^{\mathrm{H}}$ to all the connected users. This is reasonable especially in time-division-duplex (TDD) systems where local CSI can be effectively estimated at the BS by exploiting downlink-uplink reciprocity. Also, each BS needs to obtain \mathbf{A}_l at each iteration in Algorithm 2. Although the calculation of \mathbf{A}_l depends on the channels \mathbf{h}_{i,i_j} is not necessarily required at the *l*th BS. Note that \mathbf{A}_l includes B_{i_j} given as

$$B_{i_j} = \frac{1}{R_{i_j} + \sum_{k=j+1}^{K} \mathbf{h}_{i,i_j}^{\mathrm{H}} \mathbf{Q}_{i_k} \mathbf{h}_{i,i_j}} - \frac{1}{R_{i_j} + \sum_{k=j}^{K} \mathbf{h}_{i,i_j}^{\mathrm{H}} \mathbf{Q}_{i_k} \mathbf{h}_{i,i_j}},$$
(25)

where $R_{ij} + \sum_{k=j+1}^{K} \mathbf{h}_{i,i_j}^{\mathrm{H}} \mathbf{Q}_{i_k} \mathbf{h}_{i,i_j}$ represents the total interference (i.e., intra-cell and inter-cell interference) plus noise, and $R_{i_j} + \sum_{k=j}^{K} \mathbf{h}_{i,i_j}^{\mathrm{H}} \mathbf{Q}_{i_k} \mathbf{h}_{i,i_j}$ denotes the desired signal power and the interference plus noise.

In (25), both two terms can be estimated locally at user i_j . Therefore, user i_j can send the scalar value B_{r_j} to its serving BS through a feedback link. Then, BSs can exchange the value of B_{i_j} between each other. Since B_{i_j} is a scalar number, the information exchange overhead may be low. After obtaining the matrix \mathbf{A}_l , each cell operates independently to update the precoding matrices \mathbf{Q}_l and the EE parameters λ_l and β_l .

F. Complexity Analysis

In what follows, the computational complexity is measured by counting the number of flops as real floating point operations as in [9]. That is, each function such as addition, multiplication, division and square root is counted as one flop. A multiplication of an $m \times n$ matrix and $n \times p$ matrix involves 2mnp flops, and an inversion of an $m \times m$ matrix using Gauss-Jordan elimination requires $\frac{4}{3}m^3$ flops.

As for Algorithm 1, an iterative EE water-filling method needs $2\theta K^2 N_T^2 + \frac{4}{3}\theta K N_T^3 + 2\theta K N_T^2$ flops, where θ is the iteration number of the EE water-filling algorithm. The MAC-BC mapping takes $4K^2N_T^2 + \frac{4}{3}K^2N_T^3$ flops. Thus, Algorithm 1 uses $\rho_1 \left[\frac{4}{3}K^2N_T^3 + (2\theta + 4)K^2N_T^2 + 2\theta K N_T^2\right]$ flops, where ρ_1 is the iteration numbers of the bisection method in Algorithm 1.

For Algorithm 2, the calculation of \mathbf{A}_l requires $8LK^2N_T^2 + 2LKN_T^2$ floating flops, and the updating of the EE parameters involves $8KN_T^2 + KN_T$ floating flops. Thus, the total complexity of Algorithm 2 is $\rho_2 L \left[\frac{4}{3}\rho_1 KN_T^3 (K+1) + (2\theta\rho_1 + 4\rho_1 + 8L)K^2N_T^2\right] = o\left(\frac{28}{3}L\tau^5\right)$, where ρ_2 is the iteration numbers of Algorithm 2, and $\tau = \max(N_T, K)$.

IV. SIMULATION RESULTS

In this section, we investigate the performance of the proposed algorithm via numerical simulations. We consider that the number of cooperative cells equals L = 3, unless specified otherwise. The cell radius is set to be d = 500 m and each user is at least $0.5 \times d$ m away from a serving BS. The user location is randomly generated. The circuit power per antenna is $P_{\rm d} = 30$ dBm, and the basic power consumed at the BS is $P_{\rm s} = 40$ dBm. In addition, we assume that all BSs have the same transmit power constraint. The noise figure is set to be 9 dB, and the EE weight ω_l and the inefficiency PA factor η_l are fixed to be one. The channel vector \mathbf{h}_{i,l_k} from the *i*th cell to user l_k is generated as $\mathbf{h}_{i,l_k} \stackrel{\Delta}{=} \sqrt{\kappa_{i,l_k}} \mathbf{h}_{i,l_k}^{\omega}$, where $\mathbf{h}_{i,l_k}^{\omega}$ is Gaussian distributed with zero mean and unit variance, and κ_{r,l_k} denotes the large scale fading given as $\kappa_{i,l_k} = -38\log_{10} \left(d_{i,l_k} \right) - 34.5 + \xi_{i,l_k}$ in decibels. Here ξ_{i,l_k} is the log-normal shadow fading with zero mean and standard deviation of 8 dB [20], [22].

In Fig. 2, we exhibit the convergence property of the gradient descent algorithm under different signal- to-noise-ratios (SNR) and step size t. In this example, we assume $[\alpha_{l_1}, \alpha_{l_2}, \alpha_{l_3}, \alpha_{l_4}] = [0.8, 0.6, 0.4, 0.2] \quad \forall l, L = 2, \text{ and } N_T = K = 4$. It can be observed from this figure that with an appropriately chosen step size t, the algorithm always approaches to the optimal solution, and the algorithm converges faster with a larger t. In addition, the required iterations grow with SNR.

The EE performance with a different number of users at each cell with $N_T = 5$ is illustrated in Fig. 3. We also plot the linear precoding method for EE optimization in [22] for comparison. It is observed from Fig. 3 that the proposed algorithm always outperforms the linear precoding method. At a high SNR region, there exists a performance gain of nearly 20%. Moreover, Fig. 3 shows that more users allow better EE performance. It can be explained that more users can provide higher multiplexing gain.

Fig. 4 presents the EE performance with a different number



Fig. 2. Convergence behavior of the gradient descent method ($L = 2, N_T = K = 4$).



Fig. 3. Sum EE performance as a function of transmit power ($N_T = 5$).



Fig. 4. Sum EE performance as a function of transmit power (K = 3).

of transmit antennas and K = 3. The figure demonstrates that our proposed algorithms have better performance for all configurations. We also find that Algorithm 2 and WSR Algorithm achieve almost the same EE at the low SNR region. However,



Fig. 5. Sum EE performance as a function of transmit power ($N_T = 5$, $N_R = 2$).



Fig. 6. Sum EE performance as increasing of number of transmit antenna $(SNR=5~{\rm dB}).$

as SNR increases, Algorithm 2 generates higher EE, while the curves of the WSR algorithm become deteriorated. In addition, Fig. 4 confirms that more transmit antennas improves the EE performance.

In Fig. 5, the MIMO antenna case is considered. The EE performance is shown as a function of a different number of users, with $N_T = 5$ and $N_R = 2$. We can observe a similar phenomenon as in Fig. 3 that the proposed algorithm always exhibits a significant performance gain over the linear precoding method. Moreover, numerical results also show that the average EE increases as the number of users grows. However, a performance gain shrinks with the number of users.

Fig. 6 illustrates that the average EE performance with respect to the number of transmit antennas, where the number of served users at each BS increases with the number of transmit antennas according to $K = N_T - 2$. The plot shows that the EE of two algorithms both increases with the number of transmit antennas. Moreover, the slope of the curves decreases as P_d grows. It is also observed that a gain of the WSEE algorithm over the WSR tends to shrink with the number of transmit antennas.



Fig. 7. Sum EE performance as a function of transmit power ($N_T = K = 4$).



Fig. 8. Convergence behavior of Algorithm 2 ($N_T = K = 4$).

Fig. 7 compares the average EE for different circuit power $P_{\rm d}$ and $P_{\rm s}$. We can see that the performance of both algorithms decreases as $P_{\rm d}$ increases. If $P_{\rm d}$ is reduced from 40 dBm to 30 dBm, the performance of the proposed algorithm improves about 200%. Moreover, we also observe that the maximum Tx power at which the two algorithms achieve the same EE performance becomes larger as $P_{\rm d}$ increases. This implies that if the circuit power is reduced, the proposed algorithm obtains more gain than the sum rate maximization algorithm. In addition, we can observe a similar phenomenon as $P_{\rm s}$ varies.

Finally, the convergence property of the proposed EE optimization algorithm with various SNR is illustrated in Fig. 8 with $N_T = K = 4$. We can check that all the curves of Algorithm 2 converge within 10 iterations, and the convergence speed grows in the low SNR regime. In addition, we can see that at SNR of 10 dB and 20 dB, the curves converge to the same optimal point, which coincides with the results shown in Fig. 3. GUI et al.: JOINT ENERGY EFFICIENCY OPTIMIZATION WITH NONLINEAR PRECODING

V. CONCLUSION

This paper has proposed transmit covariance matrix designs for the weighted EE maximization problem in coordinated multi-cell multi-user systems considering non-linear precoding techniques. Since the objective problem is non-convex with a fractional form, we have decomposed it into a sequence of subproblems. Then the BC-MAC duality has been exploited to change each subproblem into a convex MAC problem. We have proposed a coordinated EE iterative water-filling method to solve the MAC problem, and the corresponding transmit covariance matrices for the BC problem have been obtained from the MAC-BC mapping. Finally, the proposed EE algorithm which can be implemented in a distributed manner has been introduced to iteratively find a local optimal solution for the multi-cell sum EE maximization problem. The simulation results have demonstrated the performance advantage and the convergence of the proposed DPC based algorithm. Among interesting open issues for future work, we point to the development of effective global optimization algorithms to tackle the non-convex joint precoding design problems formulated in this work.

APPENDIX A Proof of Lemma 1

By introducing Lagrange variables $\lambda = \{\lambda_l\}_{l=1}^{L}$ for EE constraint and $\mu = \{\mu_l\}_{l=1}^{L}$ for the transmit power constraint, respectively, the Lagrange function of problem (5) can be written as

$$L\left(\mathbf{Q},\beta,\lambda,\mu\right) = \sum_{l=1}^{L} \omega_{l}\beta_{l} + \sum_{l=1}^{L} \lambda_{l}\left(h_{l}\left(\mathbf{Q}_{l}\right) - \beta_{l}g_{l}\left(\mathbf{Q}_{l}\right)\right) - \sum_{l=1}^{L} \mu_{l}\left(\sum_{k=1}^{K} \operatorname{tr}\left(\mathbf{Q}_{l_{k}}\right) - P_{l}\right).$$
(26)

Then, according to (26), the KKT conditions are given as

$$\frac{\partial L}{\partial \mathbf{Q}_{l}} = \lambda_{l} \frac{\partial h_{l} \left(\mathbf{Q}_{l}\right)}{\partial \mathbf{Q}_{l}} - \lambda_{l} \beta_{l} \frac{\partial g_{l} \left(\mathbf{Q}_{l}\right)}{\partial \mathbf{Q}_{l}} + \sum_{r=1, r\neq l}^{L} \lambda_{r} \frac{\partial h_{r} \left(\mathbf{Q}_{r}\right)}{\partial \mathbf{Q}_{l}} - \mu_{l} \frac{\partial}{\partial \mathbf{Q}_{l}} \sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{Q}_{l_{k}}\right) = 0,$$
(27)

$$\frac{\partial L}{\partial \beta_l} = \omega_l - \lambda_l g_l \left(\mathbf{Q}_l \right) = 0, \tag{28}$$

$$\lambda_l \left(h_l \left(\mathbf{Q}_l \right) - \beta_l g_l \left(\mathbf{Q}_l \right) \right) = 0, \tag{29}$$

$$\mu_l \left(\sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{Q}_{l_k} \right) - P_l \right) = 0, \tag{30}$$

$$\lambda_l \ge 0, \ \mu_l \ge 0, \ \forall l. \tag{31}$$

Note that the third term of the right hand side of (27) is obtained as

$$\sum_{r=1,r\neq l}^{L} \lambda_r \frac{\partial h_r(\mathbf{Q}_r)}{\partial \mathbf{Q}_l} = \sum_{r=1,r\neq l}^{L} \lambda_r \sum_{k=1}^{K} \frac{\partial r_{l_k}^{BC}}{\partial \mathbf{Q}_{l_k}}$$
$$= -\sum_{r=1,r\neq l}^{L} \lambda_i \sum_{j=1}^{K} \mathbf{h}_{l,r_j} \left[\frac{1}{\ln 2 \left(R_{r_j} + \sum_{k=j+1}^{K} \mathbf{h}_{r,r_j}^{H} \mathbf{Q}_{r_k} \mathbf{h}_{r,r_j} \right)} - \frac{1}{\ln 2 \left(R_{r_j} + \sum_{k=j}^{K} \mathbf{h}_{r,r_j}^{H} \mathbf{Q}_{r_k} \mathbf{h}_{r,r_j} \right)} \right] \mathbf{h}_{l,r_j}^{H}.$$
(32)

We can see that (27), (30) and (31) are just the KKT conditions of problem (6). Moreover, since $g_l(\mathbf{Q}_l)$ is positive, (28) and (29) are equivalent to

$$\lambda_l = \frac{\omega_l}{g_l(\mathbf{Q}_l)} \text{ and } \beta_l = \frac{h_l(\mathbf{Q}_l)}{g_l(\mathbf{Q}_l)}, \tag{33}$$

respectively. Thus, the first conclusion of Lemma 1 is proved. The proof of the contrary case can be easily done by using a similar method.

APPENDIX B Proof of Proposition 2

First, we assume $\mu_l^* = 0$. In this case, the transmit power constraint in (6) is relaxed due to the complementary slack condition. Then, we can see that problem (6) is equivalent to problem (10) by setting $\mu_l = 0$. Thus, the global optimality follows since the BC problem in (11) becomes equivalent to the convex MAC problem in (12).

Next, supposing $\mu_l^* > 0$, we have $\sum_{k=1}^K \operatorname{tr} (\mathbf{Q}_{l_k}^*) = P_l$. We then prove the global optimality by contradiction. Assume that transmit covariance matrices \mathbf{Q}_l^* obtained by Algorithm 1 are not globally optimal to problem (6), and there is another set $\tilde{\mathbf{Q}}_l$ which is the globally optimal solution. Then it follows

$$\lambda_{l}\left(h_{l}\left(\tilde{\mathbf{Q}}_{l}\right)-\beta_{l}g_{l}\left(\tilde{\mathbf{Q}}_{l}\right)\right)-\sum_{k=1}^{K}\operatorname{tr}\left(\mathbf{A}_{l}\tilde{\mathbf{Q}}_{l_{k}}\right)$$

> $\lambda_{l}\left(h_{l}\left(\mathbf{Q}_{l}^{*}\right)-\beta_{l}g_{l}\left(\mathbf{Q}_{l}^{*}\right)\right)-\sum_{k=1}^{K}\operatorname{tr}\left(\mathbf{A}_{l}\mathbf{Q}_{l_{k}}^{*}\right).$ (34)

Since \mathbf{Q}_l^* is the optimal solution of the Lagrange dual function of (10), we have

$$\lambda_{l}h_{l}\left(\mathbf{Q}_{l}^{*}\right) - \sum_{k=1}^{K} \operatorname{tr}\left(\mathbf{Q}_{l_{k}}^{*}\left(\left(\lambda_{l}\beta_{l}\eta_{l}+\mu_{l}^{*}\right)\mathbf{I}_{N_{T}}+\mathbf{A}_{l}\right)\right)\right)$$

$$\geq \lambda_{l}h_{l}\left(\tilde{\mathbf{Q}}_{l}\right) - \sum_{k=1}^{K} \operatorname{tr}\left(\tilde{\mathbf{Q}}_{l_{k}}\left(\left(\lambda_{l}\beta_{l}\eta_{l}+\mu_{l}^{*}\right)\mathbf{I}_{N_{T}}+\mathbf{A}_{l}\right)\right).$$
(35)

By adding (35) to (34), we can get

$$\mu_l^* \sum_{k=1}^K \operatorname{tr}\left(\tilde{\mathbf{Q}}_{l_k}\right) > \mu_l^* \sum_{k=1}^K \operatorname{tr}\left(\mathbf{Q}_{l_k}^*\right).$$
(36)

Since $\sum_{k=1}^{K} \operatorname{tr} \left(\mathbf{Q}_{l_k}^* \right) = P_l$ due to $\mu_l^* > 0$, (36) leads to

$$\sum_{k=1}^{K} \operatorname{tr}\left(\tilde{\mathbf{Q}}_{l_{k}}\right) > \sum_{k=1}^{K} \operatorname{tr}\left(\mathbf{Q}_{l_{k}}^{*}\right) = P_{l}, \quad (37)$$

which contradicts that $\tilde{\mathbf{Q}}_l$ is the global optimal solution for problem (6). Thus, the assumption cannot be true, and \mathbf{Q}_l^* must be the global optimal solution.

APPENDIX C Proof of Lemma 2

Let us define
$$f(\mathbf{Q}_l, \mathbf{Q}_{-l}) = \sum_{r \neq l}^{L} \lambda_r (h_r(\mathbf{Q}_r) - \beta_r g_r(\mathbf{Q}_r)).$$

We first check the convexity of $f(\mathbf{Q}_l, \mathbf{Q}_{-l})$ with respect to $\mathbf{Q}_l \in Q_l \stackrel{\Delta}{=} \left\{ \mathbf{Q}_{l_k} | \mathbf{Q}_{l_k} \succeq \mathbf{0}, \sum_{k=1}^K \operatorname{tr}(\mathbf{Q}_{l_k}) \le P_l \right\}$. Denoting $\mathbf{Q}_l = \mathbf{X}_l + t_1 \mathbf{Y}_l$, where $t_1 \in [0, 1]$, and $\mathbf{X}_l, \mathbf{Y}_l \in Q_l$, the second derivative of $f(\mathbf{X}_l + t_1 \mathbf{Y}_l, \mathbf{Q}_{-l})$ equals

$$\frac{\partial^{2}}{\partial t^{2}}f\left(\mathbf{X}_{l}+t_{1}\mathbf{Y}_{l}, \mathbf{Q}_{-l}\right)$$

$$=\sum_{r=1, r\neq l}^{L} \lambda_{r} \operatorname{tr}\left[\mathbf{h}_{r,l}\mathbf{Y}_{l}\left(\frac{\mathbf{h}_{r,l}^{\mathrm{H}}\mathbf{Y}_{l}\mathbf{h}_{r,l}}{\ln 2(R_{r})^{2}}-\frac{\mathbf{h}_{r,l}^{\mathrm{H}}\mathbf{Y}_{l}\mathbf{h}_{r,l}}{\ln 2\left(R_{r}+\mathbf{h}_{r,r}^{\mathrm{H}}\mathbf{Y}_{r}\mathbf{h}_{r,r}\right)^{2}}\right)\mathbf{h}_{r,l}^{\mathrm{H}}\right].$$
(38)

Since $\mathbf{h}_{r,r}^{\mathrm{H}} \mathbf{Y}_r \mathbf{h}_{r,r} \geq 0$, we can easily see that $\frac{\partial^2}{\partial t_1^2} f(\mathbf{X}_l + t_1 \mathbf{Y}_l, \mathbf{Q}_{-l}) \geq 0$. Therefore, the convexity of $f(\mathbf{Q}_l, \mathbf{Q}_{-l})$ is proved.

Assume that $\mathbf{Q}_l = \mathbf{Q}_l$ for all $l = 1, 2, \dots, L$ from the previous iteration. Also we denote \mathbf{Q}_l^* as the optimal solution of problem (6) for the *l*th cell and define $\tilde{e}(\mathbf{Q}_l, \mathbf{Q}_{-l}) = \lambda_l (h_l(\mathbf{Q}_l) - \beta_l g_l(\mathbf{Q}_l))$. Then we have

$$\sum_{l=1}^{L} \tilde{e} \left(\mathbf{Q}_{l}^{*}, \ \tilde{\mathbf{Q}}_{-l} \right) = \tilde{e} \left(\mathbf{Q}_{l}^{*}, \ \tilde{\mathbf{Q}}_{-l} \right) + f \left(\mathbf{Q}_{l}^{*}, \ \tilde{\mathbf{Q}}_{-l} \right)$$

$$\geq \tilde{e} \left(\mathbf{Q}_{l}^{*}, \ \tilde{\mathbf{Q}}_{-l} \right) + f \left(\tilde{\mathbf{Q}}_{l}, \ \tilde{\mathbf{Q}}_{-l} \right) - \sum_{i=1}^{K} \operatorname{tr} \left(\mathbf{A}_{l} \left(\mathbf{Q}_{l_{k}}^{*} - \tilde{\mathbf{Q}}_{l_{k}} \right) \right)$$
(39)

$$\geq \tilde{e}\left(\tilde{\mathbf{Q}}_{l}, \ \tilde{\mathbf{Q}}_{-l}\right) + f\left(\tilde{\mathbf{Q}}_{l}, \ \tilde{\mathbf{Q}}_{-l}\right) - \sum_{i=1}^{K} \operatorname{tr}\left(\mathbf{A}_{l}\left(\tilde{\mathbf{Q}}_{l_{k}} - \tilde{\mathbf{Q}}_{l_{k}}\right)\right)$$

$$\tag{40}$$

$$=\sum_{l=1}^{L}\tilde{e}\left(\tilde{\mathbf{Q}}_{l},\;\tilde{\mathbf{Q}}_{-l}\right)$$

where (39) follows from the convexity of $f(\mathbf{Q}_l, \mathbf{Q}_{-l})$, and (40) comes from the fact that \mathbf{Q}_l^* is the optimal solution of Problem (6). Thus, the objective function $\sum_{l=1}^{L} \tilde{e}(\mathbf{Q}_l, \mathbf{Q}_{-l})$ is non-decreasing after each updating at each cell. Since the objective function is bounded from the above due to transmit power transmit, Lemma 2 is proved.

REFERENCES

 D. Gesbert, M. Kountouris, R. Heath, Jr., C. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sept. 2007.

- [2] H. Lee, B. Lee, and I. Lee, "Iterative detection and decoding with an improved V-BLAST for MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 504–513, Mar. 2006.
- [3] C. Jiang and Leonard J. Cimini, "Energy-efficient transmission for MIMO interference channels," *in Proc. IEEE WCNC*, Apr. 2012. pp. 522–527,
- [4] D. Feng et al., "A survey of energy-efficient wireless communications," IEEE Commun. Surveys Tuts., vol. 15, no. 1, pp. 167–178, Jan. 2013.
- [5] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.
- [6] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and H.-V. Poor, "On Gaussian MIMO BC-MAC duality with multiple transmit covariance constraints," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, Apr. 2012.
- [7] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.
- [8] Q. H. Spencer, A. L. Swindelhurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 461–471, Feb. 2004.
- [9] H. Sung, S.-R. Lee, and I. Lee, "Generalized channel inversion methods for multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3489–3499, Nov. 2009.
- [10] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450–5458, Oct. 2010.
- [11] S.-H. Park, H. Park, and I. Lee, "Distributed beamforming techniques for weighted sum-rate maximization in MISO interference channels," *IEEE Commun. Lett.*, vol. 14, no. 12, pp. 1131–1133, Dec. 2010.
- [12] H. Sung, S.-H. Park, K.-J. Lee, and I. Lee, "Linear precoder designs for Kuser interference channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 291–301, Jan. 2010.
- [13] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Inf. Theory* vol. 57, no. 5, pp. 3117–3131, May 2011.
- [14] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
- [15] H.-J. Choi, S.-H. Park, S.-R. Lee, and I. Lee, "Distributed beamforming techniques for weighted sum-rate maximization in MISO interfering broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1314–1320, Apr. 2012.
- [16] S.-R. Lee, H.-B. Kong, H. Park, and I. Lee, "Beamforming designs based on an asymptotic approach in MISO interference channels," *IEEE Trans. Wireless Commun.*, vol.12, no. 12, pp. 6430–6438, Dec. 2013.
- [17] J. Xu and L. Qiu, "Energy efficiency optimazition for MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 12 no. 2, pp. 690–701, Feb. 2013.
- [18] J. Mao, G. Xie, J. Gao, and Y. Liu, "Energy efficiency optimization for cognitive radio MIMO broadcast channels," *IEEE Commun. Lett.*, vol. 17, no. 2, pp. 337–340, Feb. 2013.
- [19] G. Miao, N. Himayat, and G. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 545–554, Feb. 2010.
- [20] 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.
- [21] H. Kim, S.-R. Lee, C. Song, and I. Lee, "Optimal power allocation for energy efficiency maximization in distributed antenna systems," *in Proc. IEEE ICC*, May 2013, pp. 1–5.
- [22] 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.
- [23] W. Dinkelbach, "On nonlinear fractional programming," Bulletin of the Australian Mathematical Society, vol. 13, pp. 492–498, Mar. 1967.
- [24] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [25] C. Fung, W. Yu, and T. Lim, "Multi-antenna downlink precoding with individual rate constraints: power minimization and user ordering," *in Proc. ICCS*, pp. 45–49.



Xin Gui received the Ph.D. degree degrees in Communication and Information System from Beijing University of Posts and Telecommunications, Beijing, China, July 2013. From 2014 to 2015, he was a Postdoctoral Fellow at Korea University. Since Jan. 2015, he works at Samsung institution of communication, Beijing.



Jaehoon Jung received the B.S. and M.S. degrees in Electrical Engineering from Korea University, Seoul, Korea, in 2011 and 2013, respectively. He is currently working toward the Ph.D. degree in the School of Electrical Engineering, Korea University. His research interests include communication theory and signal processing techniques for multicell MIMO wireless network systems. He was awarded the Student Travel Grant at the IEEE International Conference on Communications in 2013.



Kyoung-Jae Lee received the B.S., M.S., and Ph.D. degrees from the School of Electrical Engineering, Korea University, Seoul, Korea, in 2005, 2007, and 2011, respectively. During the winter of 2006, he interned at Beceem Communications, Inc., Santa Clara, CA, USA, and during the summer of 2009, he visited the University of Southern California, Los Angeles, CA, as a Visiting Student. He worked as a Research Professor at Korea University in 2011. From 2011 to 2012, he was a Postdoctoral Fellow at the Wireless Networking and Communications Group, University

of Texas at Austin, Austin, TX, USA. Since September 2012, he has been with the Department of Electronics and Control Engineering, Hanbat National University, Daejeon, Korea. His research interests are in communication theory, signal processing, and information theory applied to the next-generation wireless communications. He was a recipient of the Best Paper Award at IEEE VTC Fall in 2009, the IEEE ComSoc APB Outstanding Paper Award in 2013, and the IEEE ComSoc APB Outstanding Young Researcher in 2013.



Inkyu Lee received the B.S. (Hons.) degree in Control and Instrumentation Engineering from Seoul National University, Seoul, South Korea, in 1990, and the M.S. and Ph.D. degrees in Electrical Engineering from Stanford University, Stanford, CA, USA, in 1992 and 1995, respectively. From 1995 to 2001, he was a Member of Technical Staff with Bell Laboratories, Lucent Technologies, where he studied highspeed wireless system designs. From 2001 to 2002, he was with Agere Systems, Murray Hill, NJ, USA, as a Distinguished Member of Technical Staff. Since

2002, he has been with Korea University, Seoul, where he is currently a Professor with the School of Electrical Engineering. In 2009, he visited the University of Southern California, Los Angeles, CA, USA, as a Visiting Professor. He has authored over 130 journal papers in the IEEE. He has 30 U.S. patents granted or pending. His research interests include digital communications, signal processing, and coding techniques applied for next generation wireless systems. He has served as an Associate Editor of the IEEE TRANSACTIONS ON COMMU-NICATIONS from 2001 to 2011, and the IEEE TRANSACTIONS ON WIRE-LESS COMMUNICATIONS from2007 to 2011. He has been a Chief Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICA-TIONS (Special Issue on 4G Wireless Systems) in 2006. He currently serves as an Editor of the IEEE ACCESS. He was a recipient of the IT Young Engineer Award at the IEEE/IEEK Joint Award in 2006, and of the Best Paper Award at APCC in 2006, the IEEE VTC in 2009, and ISPACS in 2013. He was also a recipient of the Best Research Award from the Korea Information and Communications Society in 2011, and the Best Young Engineer Award from the National Academy of Engineering in Korea (NAEK) in 2013. He has been elected as a Member of NAEK in 2015. He is an IEEE Fellow and an IEEE Distinguished Lecturer