Title	Globally Optimal Energy Efficiency Maximization for Capacity-Limited Fronthaul Crans with þÿD y n a m i c P o w e r A m p l i f i e r s E f f i c i e n c y
Authors(s)	Nguyen, Kien-Giang; Vu, Quang-Doanh; Tran, Le-Nam; Juntti, Markku
Publication date	2018-09-13
Publication information	2018 IEEE International Conference on Acoustics, Speech and Signal Processing: Proceedings
Conference details	The 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, Canada, 15-20 April 2018
Publisher	IEEE
Link to online version	https://2018.ieeeicassp.org/
Item record/more information	http://hdl.handle.net/10197/10384
Publisher's statement	© 2018 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/icassp.2018.8461308

GLOBALLY OPTIMAL ENERGY EFFICIENCY MAXIMIZATION FOR CAPACITY-LIMITED FRONTHAUL CRANS WITH DYNAMIC POWER AMPLIFIERS' EFFICIENCY

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

* Centre for Wireless Communications, University of Oulu, P.O.Box 4500, FI-90014, Finland;
Email: {giang.nguyen, doanh.vu, markku.juntti}@oulu.fi.
† School of Electrical and Electronic Engineering, University College Dublin, Ireland. Email: nam.tran@ucd.ie.

ABSTRACT

A joint beamforming and remote radio head (RRH)-user association design for downlink of cloud radio access networks (CRANs) is considered. The aim is to maximize the system energy efficiency subject to constraints on users' quality-of-service, capacity of fronthaul links and transmit power. Different to the conventional power consumption models, we embrace the dependence of baseband signal processing power on the data rate, and the dynamics of the power amplifiers' efficiency. The considered problem is a mixed Boolean nonconvex program whose optimal solution is difficult to find. As our main contribution, we provide a discrete branch-reduce-and-bound (DBRnB) approach to solve the problem globally. We also make some modifications to the standard DBRnB procedure. Those remarkably improve the convergence performance. Numerical results are provided to confirm the validity of the proposed method.

***Index Terms*—** Energy efficiency, CRAN, limited fronthaul, nonlinear power amplifier, discrete branch-reduce-and-bound.

1. INTRODUCTION

Cloud radio access network (CRAN) is a novel network architecture which effectively supports the low-latency deployments such as joint transmission coordinated multipoint [1, 2]. In CRANs, the functionality of the conventional base stations is divided into two entities called baseband units (BBUs) and remote radio heads (RRHs). The BBUs including the signal processing functionalities are located at a central cloud computing platform, while the RRHs placed close to the antennas are responsible for wireless interface of the network. This architecture allows CRANs to alleviate the strict synchronization requirements among RRHs, and also to leverage powerful computing capabilities for full cooperation. One of the main challenges in CRAN design is to deal with the limited capacity of the fronthaul links, the means of transporting baseband signals between BBU and RRHs [3].

An effective and widely used method overcoming the problem of the limited fronthaul is to select a set of users that can be served by a RRH [4–7]. The approach gives rise to the RRH-user association problems modeled by a set of binary preference variables. As a result, the problems are cast as mixed binary integer programs whose optimal solutions are difficult to derive.

This work was supported in part by the projects "Wireless Connectivity for Internet of Everything-Energy Efficient Transceiver and System Design (WiConIE)" funded by Academy of Finland under Grant 297803, and "Flexible Uplink-Downlink Resource Management for Energy and Spectral Efficiency Enhancing in Future Wireless Networks (FURMESFuN)" funded by Academy of Finland under Grant 31089. It was also supported by Tauno Tönning Foundation, Nokia Foundation, and Walter Ahlström Foundation.

Due to a growing concern over the power consumption in mobile networks, recent research in wireless communications has shifted its focus on energy efficiency (EE). The problem of energy efficiency maximization (EEmax) has been studied in many prior publications [8–13] for different contexts. The mentioned works and other related studies assume that signal processing power is independent of the data rate and the efficiency of the power amplifiers (PAs) can be modeled as a constant factor. However, these assumptions are not true in reality. In fact, signal processing power increases proportionally with the data rates [14–17], and PA's efficiency is dynamic depending on the output power [18–20].

In this paper, we focus on EE downlink transmission in CRANs. Specifically, we jointly design transmit beamforming and RRH-user association with the objective of maximizing the network EE under the constraints of per-RRH fronthaul capacity, transmit power budget and user's quality-of-service (QoS). Our contribution is fourfold: i) we include the rate-dependent signal processing power and the dynamics of PA's efficiency in the total power consumption model; ii) we develop a globally optimal solution to the problem of interest, which is a mixed Boolean nonconvex program, based on the general discrete branch-reduce-and-bound (DBRnB) framework introduced in [21]; iii) special modifications are made to the standard DBRnB procedure to improve the convergence performance; and iv) the numerical results assessing the proposed method are provided.

2. SYSTEM MODEL AND PROBLEM FORMULATION

We consider downlink transmission of a multiuser multiple-input single-output (MISO) wireless system consisting a set of B RRHs, denoted by $\mathcal{B} \triangleq \{1, \dots, B\}$, each equipped with I antennas, and a set of K single-antenna users, denoted by $\mathcal{K} \triangleq \{1, \dots, K\}$. The BBU pool is assumed to have perfect channel state information of the network. A specific user can simultaneously receive data from multiple RRHs. Let $\mathbf{h}_{b,k} \in \mathbb{C}^{1\times I}$ be the channel between RRH b and user k, d_k denote the normalized data symbol intended for user k, and $\mathbf{w}_{b,k} \in \mathbb{C}^{I\times 1}$ be the beamforming vector from RRH b to user k. Assuming channels are flat, the received signal at user k is

$$y_k = \Big(\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k}\mathbf{w}_{b,k}\Big) d_k + \sum_{j\in\mathcal{K}\setminus k} \Big(\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k}\mathbf{w}_{b,j}\Big) d_j + n_k \quad (1)$$

where $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the additive white Gaussian noise. For notational convenience, let $\mathbf{h}_k \triangleq [\mathbf{h}_{1,k}, \mathbf{h}_{2,k}, \dots, \mathbf{h}_{B,k}] \in \mathbb{C}^{1\times IB}$ and $\mathbf{w}_k \triangleq [\mathbf{w}_{1,k}^T, \mathbf{w}_{2,k}^T, \dots, \mathbf{w}_{B,k}^T]^T \in \mathbb{C}^{IB\times 1}, \forall k$. Also, let $\mathbf{w}$ denote the beamforming vector stacking all $\mathbf{w}_k$. Assuming single-user decoding, the SINR at user k is written as

$$\gamma_k(\mathbf{w}) \triangleq \frac{|\mathbf{h}_k\mathbf{w}_k|^2}{\sum_{j\in\mathcal{K}\setminus k} |\mathbf{h}_k\mathbf{w}_j|^2 + \sigma_k^2}. \quad (2)$$

Let r_k be the achievable data rate transmitted to user k. Then by Shannon's coding theory, we have

$$r_k \leq \log(1+\gamma_k(\mathbf{w})), \forall k \in \mathcal{K}. \quad (3)$$

In reality, the fronthaul link from the BBU pool to RRH b has a finite capacity, denoted by $\bar{C}_b$. For the problem formulation purpose let us introduce Boolean variables $\{x_{b,k}\}_{b,k}$ such that $x_{b,k}=1$ indicates that user k receives data from RRH b and $x_{b,k}=0$ otherwise. Then the total data rate transmitted by the wireless interface of RRH b is $\sum_{k\in\mathcal{K}} x_{b,k} r_k$, and thus, for feasible transmission we should have

$$\textstyle\sum_{k\in\mathcal{K}} x_{b,k} r_k \leq \bar{C}_b, \forall b \in \mathcal{B}. \quad (4)$$

Power Consumption Model: The power consumption in the network can be divided into three parts: rate-independent power consumption, power for signal processing, and power consumed in PAs [20, 22–24]. Details of these parts are presented below.

The rate-independent power consumption is modeled as [22,24]

$$P_{\text{I}} \triangleq KP_{\text{ms}} + P_{\text{OLT}} + \textstyle\sum_{b\in\mathcal{B}} (s_b(P_{\text{RRH}}^{\text{active}} + P_{\text{FH}}^{\text{active}}) + (1-s_b)(P_{\text{RRH}}^{\text{sleep}} + P_{\text{FH}}^{\text{sleep}})). \quad (5)$$

In (5), P_{ms} is the circuit power consumed by a user device, $P_{\text{RRH}}^{\text{active}}$ and $P_{\text{RRH}}^{\text{sleep}}$ are the power consumption at a RRH corresponding to the active and sleep modes, respectively [22]. It is assumed that all RRHs connect to the BBU through a passive optical network which consists of an optical line terminal (OLT) and a set of network units (NUs) [23]. The OLT is always active and consumes a fixed power, denoted by P_{OLT}; NUs are switchable between the active or sleep mode, each consuming an amount of power denoted by $P_{\text{FH}}^{\text{active}}$ and $P_{\text{FH}}^{\text{sleep}}$, respectively. We introduce Boolean variables $\{s_b\}_b$ such that $s_b=1$ indicates RRH and NU b are active and $s_b=0$ otherwise. The relationship between s_b and $x_{b,k}$ can be represented as

$$s_b = \max_{k\in\mathcal{K}}\{x_{b,k}\} \Leftrightarrow \begin{cases} s_b \geq x_{b,k}, \forall k \in \mathcal{K} \\ s_b \leq \sum_{k\in\mathcal{K}} x_{b,k} \end{cases}, \forall b \in \mathcal{B}. \quad (6)$$

For RRH b, the power consumed by the signal processing functions is measured by a continuous function of the fronthaul rate $\tilde{r}_b$, denoted by $\psi_b(\tilde{r}_b)$ where $\tilde{r}_b \triangleq \sum_{k\in\mathcal{K}} x_{b,k} r_k$ [14–16]. According to [14], $\psi_b(\tilde{r}_b)$ is linearly scaled w.r.t. $\tilde{r}_b$, i.e.

$$\psi(\tilde{r}_b) = p_{\text{SP}}\tilde{r}_b \quad (7)$$

where p_{SP} is a constant coefficient in (W/(Gnats/s)).

In reality, the efficiencies of the PAs depend on the output power modeled as $\epsilon_{b,i}(\{\mathbf{w}_{b,k}\}_k) \triangleq \frac{1}{\tilde{\epsilon}}\sqrt{\sum_{k\in\mathcal{K}} |[\mathbf{w}_{b,k}]_i|^2}, \forall k \in \mathcal{K}, b \in \mathcal{B}, i=1,...,I$, where $\tilde{\epsilon} \triangleq \sqrt{P_{\text{a}}}/\epsilon_{\max}$, and P_{a} and $\epsilon_{\max} \in [0,1]$ are the maximum power of the PA and the maximum PA's efficiency, respectively [20]. Let $\phi_b(\{\mathbf{w}_{b,k}\}_k)$ measure the amount of power consumed by the PAs at RRH b. Then we have

$$\phi_b(\{\mathbf{w}_{b,k}\}_k) = \sum_{i=1}^{I} \frac{\sum_{k\in\mathcal{K}} |[\mathbf{w}_{b,k}]_i|^2}{\epsilon_{b,i}(\{\mathbf{w}_{b,k}\}_k)} = \tilde{\epsilon}\textstyle\sum_{i=1}^{I} ||\tilde{\mathbf{w}}_{b,i}||_2 \quad (8)$$

where $\tilde{\mathbf{w}}_{b,i} \triangleq [[\mathbf{w}_{b,1}]_i; [\mathbf{w}_{b,2}]_i; ...; [\mathbf{w}_{b,K}]_i] \in \mathbb{C}^{K\times 1}$.

Let us define $\mathbf{x} \triangleq \{x_{b,k}\}_{b\in\mathcal{B}k\in\mathcal{K}}$, $\mathbf{s} \triangleq \{s_b\}_{b\in\mathcal{B}}$, and $\mathbf{r} \triangleq \{r_k\}_k$. Based on the above discussions, the total consumed power in the considered system, denoted by $f_{\text{P}}(\mathbf{w},\mathbf{x},\mathbf{r},\mathbf{s})$, is expressed as

$$\begin{aligned} f_{\text{P}}(\mathbf{w},\mathbf{x},\mathbf{r},\mathbf{s}) &\triangleq P_{\text{I}} + \textstyle\sum_{b\in\mathcal{B}}(\psi(\tilde{r}_b) + \phi_b(\{\mathbf{w}_{b,k}\}_k)) \\ &= \textstyle\sum_{b\in\mathcal{B}}(\tilde{\epsilon}\sum_{i=1}^{I}||\tilde{\mathbf{w}}_{b,i}||_2 + p_{\text{SP}}\sum_{k\in\mathcal{K}} x_{b,k}r_k + s_b\Delta P) + P_{\text{const}} \end{aligned} \quad (9)$$

in which $P_{\text{const}} \triangleq B(P_{\text{RRH}}^{\text{sleep}} + P_{\text{FH}}^{\text{sleep}}) + KP_{\text{ms}} + P_{\text{OLT}}$ and $\Delta P \triangleq P_{\text{RRH}}^{\text{active}} + P_{\text{FH}}^{\text{active}} - P_{\text{RRH}}^{\text{sleep}} - P_{\text{FH}}^{\text{sleep}}$ are constants.

Problem Formulation: We jointly design beamforming and RRH-user association such that the overall network EE is maximized. The problem of interest reads

$$\begin{aligned} &\underset{\mathbf{w},\mathbf{x},\mathbf{s},\mathbf{r}}{\text{maximize}}\ \frac{\sum_{k\in\mathcal{K}} r_k}{f_{\text{P}}(\mathbf{w},\mathbf{x},\mathbf{r},\mathbf{s})} && (10a)\\ &\text{subject to } r_k \geq r_0,\ \forall k \in \mathcal{K} && (10b)\\ &\textstyle\sum_{k\in\mathcal{K}} ||\mathbf{w}_{b,k}||_2^2 \leq \bar{P}_b,\ \forall b \in \mathcal{B} && (10c)\\ &||\tilde{\mathbf{w}}_{b,i}||_2^2 \leq P_{\text{a}},\ \forall b \in \mathcal{B}, i=1,...,I && (10d)\\ &||\mathbf{w}_{b,k}||_2^2 \leq x_{b,k}\bar{P}_b,\ \forall k \in \mathcal{K}, b \in \mathcal{B} && (10e)\\ &\textstyle\sum_{b\in\mathcal{B}} x_{b,k} \geq 1, \forall k \in \mathcal{K} && (10f)\\ &\mathbf{x} \in \{0,1\}^{BK},\ \mathbf{s} \in \{0,1\}^{B} && (10g)\\ &(3), (4), (6). && (10h) \end{aligned}$$

The constraints in (10b), (10c) and (10d) represent users' QoS, the total transmit power and per antenna power constraints at each individual RRH, respectively. Constraint (10e) guarantees that $||\mathbf{w}_{b,k}||_2^2 = 0$ if RRH b does not serve user k. Constraint (10f) implies that each user is served by at least one RRH (due to QoS).

3. GLOBALLY OPTIMAL SOLUTION BY DISCRETE MONOTONIC OPTIMIZATION

Problem (10) is a mixed Boolean nonconvex program (MBNP) generally known to be NP-hard. Although numerous wireless communications nonconvex problems can be solved using *general monotonic optimization* (GMO) [12,13,25,26], the GMO principle is inapplicable to MBNP since it outputs only approximate solutions of discrete variables [27]. Here we globally solve (10) based on the so-called *discrete monotonic optimization* (DMO) [21]. We follow the definitions of *box*, *increasing function*, and *normal cone* in [21]; $[\mathbf{a};\mathbf{b}]$ denotes the box with lower and upper vertices $\mathbf{a}$ and $\mathbf{b}$.

The standard procedure solving a DMO problem is DBRnB [21]. It is an iterative procedure performing three basic operations at each iteration: *branching*, *reduction,* and *bounding*. Starting from original box $[\mathbf{a};\mathbf{b}]$, we iteratively divide it into smaller and smaller ones, remove boxes that do not contain an optimal solution, search over remaining boxes for an improved solution until an error tolerance is met. During the branching and reduction steps, elements corresponding to discrete constraints are adjusted to stay in the discrete set. Details of using DBRnB to solve (10) are presented next.

The current form of (10) is not a (standard) DMO problem since the objective in (10a) is not an increasing function w.r.t. the involved variables. Thus we need to reformulate (10) as

$$\begin{aligned} &\underset{\eta,\mathbf{w},\mathbf{x},\mathbf{s},\mathbf{r},\mathbf{t}}{\text{maximize}}\ \eta && (11a)\\ &\text{subject to } \eta\tilde{f}_{\text{P}}(\mathbf{x},\mathbf{s},\mathbf{r},\mathbf{t}) - \textstyle\sum_{k\in\mathcal{K}} r_k \leq 0 && (11b)\\ &\textstyle\sum_{i=1}^{I} ||\tilde{\mathbf{w}}_{b,i}||_2 \leq t_b,\ \forall b \in \mathcal{B} && (11c)\\ &(10b)-(10h) && (11d) \end{aligned}$$

where η and $\mathbf{t} \triangleq \{t_b\}_b$ are the newly introduced variables and $\tilde{f}_{\text{P}}(\mathbf{x},\mathbf{s},\mathbf{r},\mathbf{t}) \triangleq \sum_{b\in\mathcal{B}}(\tilde{\epsilon}t_b + s_b\Delta P + p_{\text{SP}}\sum_{k=1}^{K} x_{bk}r_k) + P_{\text{const}}$. The equivalence between (10) and (11) in terms of optimal solution set can be easily proved since (11) is indeed the epigraph of (10), i.e. at the optimum constraints (11b), (11c) are active. We now have a useful observation to solve (11). Let $(\mathbf{x}^*,\mathbf{s}^*,\mathbf{r}^*,\mathbf{t}^*)$ be an optimal solution to (11). Then we have $\eta^* = \frac{\sum_{k=1}^{K} r_k^*}{f_{\mathrm{P}}(\mathbf{x}^*,\mathbf{s}^*,\mathbf{r}^*,\mathbf{t}^*)}$, and an optimal beamforming vector $\mathbf{w}^*$ can be computed as

$$\mathbf{w}^* = \mathrm{find}\{\mathbf{w}|(3),(10\mathrm{c})-(10\mathrm{e}),(11\mathrm{c})\} \tag{12}$$

in which we replace $(\mathbf{s},\mathbf{x},\mathbf{r},\mathbf{t})$ by $(\mathbf{s}^*,\mathbf{x}^*,\mathbf{r}^*,\mathbf{t}^*)$. This implies that $\mathbf{w}^*$ and η^* can be determined easily if $(\mathbf{x}^*,\mathbf{s}^*,\mathbf{r}^*,\mathbf{t}^*)$ are known. Moreover, we observe that constraints (4), (6), (10f), (11b), and (11c) are monotone w.r.t. $(\mathbf{s},\mathbf{x},\mathbf{r},\mathbf{t})$, and thus, the branching step can be done over $(\mathbf{s},\mathbf{x},\mathbf{r},\mathbf{t})$. The proposed DBRnB-based algorithm is outlined in Alg. 1. Details of the main steps are as follows.

Let $\mathcal{S}$ be the feasible set of problem (11), i.e.,

$$\begin{aligned}\mathcal{S} \triangleq \big\{&\{\mathbf{s},\mathbf{x},\mathbf{r},\mathbf{t}\}|(3),(6),(10\mathrm{b}),(10\mathrm{d})-(10\mathrm{g}),(11\mathrm{b}),\\ &\textstyle\sum_{k\in\mathcal{K}} x_{b,k} r_k \le s_b \bar{C}_b, \sum_{k\in\mathcal{K}}\|\mathbf{w}_{b,k}\|_2^2 \le s_b \bar{P}_b,\\ &\textstyle\sum_{i=1}^{I}||\tilde{\mathbf{w}}_{b,i}||_2^2 \le s_b t_b, \forall b\in\mathcal{B}\big\}.\end{aligned}$$

Remark that we have equivalently rewritten (4), (10c) and (11c) by multiplying their right hand sides by s_b to improve the proposed algorithm's efficiency. In other words, if $s_b = 0$, we can skip examining the constraints involving s_b. Since $\mathcal{S}$ is upper bounded by the power and backhaul constraints, it satisfies the normal and finite properties. At the initialization stage, we need to find a box $[\mathbf{a};\mathbf{b}] \in \mathbb{R}_+^{B(K+2)+K}$ such that $\mathcal{S} \subseteq [\mathbf{a};\mathbf{b}]$, where $\mathbf{a} \triangleq [\underline{\mathbf{s}},\underline{\mathbf{x}},\underline{\mathbf{r}},\underline{\mathbf{t}}]$ and $\mathbf{b} \triangleq [\overline{\mathbf{s}},\overline{\mathbf{x}},\overline{\mathbf{r}},\overline{\mathbf{t}}]$. Obviously, $\underline{s}_b = 0, \overline{s}_b = 1, \underline{x}_{b,k} = 0, \overline{x}_{b,k} = 1, \underline{t}_b = 0$ and $\overline{t}_b = I\sqrt{P_{\mathrm{a}}}$. We also have $\underline{r}_k = r_0$ due to (10b), and $r_k \le \min\{\bar{C}_b, \log(1+|\mathbf{h}_k\mathbf{w}_k|^2/WN_0)\} \le \min\{\bar{C}_b, \log(1+B\bar{P}_b\|\mathbf{h}_k\|_2^2/WN_0)\} = \overline{r}_k$, since $|\mathbf{h}_k\mathbf{w}_k|^2 \le \|\mathbf{h}_k\|_2^2\|\mathbf{w}_k\|_2^2$ (the Cauchy-Schwarz inequality) and $\|\mathbf{w}_k\|_2^2 \le B\bar{P}_b$. Notation η^{best} denotes the best current objective value, and set $\mathcal{R}$ contains the candidate boxes which includes only red($[\mathbf{a};\mathbf{b}]$) at starting stage; here red($[\mathbf{a};\mathbf{b}]$) denotes the box returned by the reduction operation (over $[\mathbf{a};\mathbf{b}]$) presented following.

Reduction: A box $[\mathbf{a}';\mathbf{b}']$ possibly contains segments either infeasible to (11) or resulting in an objective smaller than η^{best}. Reduction refers to removing those portions, thus reducing the search space in the next iterations, i.e. we create $[\mathbf{a}'';\mathbf{b}''] \subseteq [\mathbf{a}';\mathbf{b}']$ such that an optimal solution (if exists in $[\mathbf{a}';\mathbf{b}']$) must be contained in $[\mathbf{a}'';\mathbf{b}'']$. Mathematically, we can replace $\mathbf{a}'$ by $\mathbf{a}'' \ge \mathbf{a}'$ where $\mathbf{a}'' = \mathbf{b}' - \sum_{j=1}^{B(K+2)+K} \mathbf{e}_j \alpha_j (b'_j - a'_j)$ and $\alpha_j = \sup\{\alpha\ |0 \le \alpha \le 1,\ \mathbf{b}' - \alpha(b'_j - a'_j)\mathbf{e}_j \in [\mathbf{a};\mathbf{b}]\backslash\mathcal{S}, \eta(\mathbf{b}' - \alpha(b'_j - a'_j)\mathbf{e}_j) \ge \eta^{\text{best}}\}$, for each j; here $\eta(\mathbf{a})$ denotes the value of η at vertex $\mathbf{a}$, and $\mathbf{e}_j$ is the jth unit vector. Similarly, vertex $\mathbf{b}'$ is replaced by $\mathbf{b}'' \le \mathbf{b}'$ where $\mathbf{b}'' = \mathbf{a}'' + \sum_{j=1}^{B(K+2)+K} \beta_j (b'_j - a''_j)\mathbf{e}_j$ and $\beta_j = \sup\{\beta\ |0 \le \beta \le 1,\ \mathbf{a}'' + \beta(b'_j - a''_j)\mathbf{e}_j \in \mathcal{S}\}$, for each j. The values of $\{\alpha_j\}$ and $\{\beta_j\}$ can be found by the bisection method. Note that for the case of Boolean variables, the reduction can be simplified as follows. If $a'_j, b'_j \in \{0,1\}$ and $b'_j - a'_j = 1$, we can quickly set that $a''_j = 1$ if $\mathbf{b}' - \mathbf{e}_j \in [\mathbf{a};\mathbf{b}]\backslash\mathcal{S}$ and $a''_j = 0$ otherwise. If $a''_j = 0$, we obtain $b''_j = 1$ if $\mathbf{a}'' + \mathbf{e}_j \in \mathcal{S}$ and $b''_j = 0$ otherwise.

Branching: At each iteration, a box in $\mathcal{R}$ is selected and split into two new equal size boxes. To be bound improving, the box with the largest upper bound of η is picked denoted by $[\tilde{\mathbf{a}};\tilde{\mathbf{b}}]$ for convenience. Let $k \triangleq \arg\max_j(\tilde{b}_j - \tilde{a}_j)$. Then the two new smaller boxes $[\tilde{\mathbf{a}};\tilde{\mathbf{b}}']$ and $[\tilde{\mathbf{a}}';\tilde{\mathbf{b}}]$ are created where $\tilde{b}'_j = \tilde{b}_j$ and $\tilde{a}'_j = \tilde{a}_j$ for all $j \ne k$, $\tilde{b}'_k = \lfloor \tilde{b}_k - (\tilde{b}_k - \tilde{a}_k)/2 \rfloor_{\mathbb{Z}}$ and $\tilde{a}'_k = \lceil \tilde{a}_k + (\tilde{b}_k - \tilde{a}_k)/2 \rceil_{\mathbb{Z}}$ for the case of Boolean variables, and $\tilde{b}'_k = \tilde{b}_k - (\tilde{b}_k - \tilde{a}_k)/2$ and $\tilde{a}'_k = \tilde{a}_k + (\tilde{b}_k - \tilde{a}_k)/2$ for the case of continuous variables; notations $\lceil p \rceil_{\mathcal{Q}}$ and $\lfloor p \rfloor_{\mathcal{Q}}$ denote the upper and lower nearest neighbor elements of p in set $\mathcal{Q}$, respectively. Note that for the case $\tilde{a}_k, \tilde{b}_k \in \{0,1\}$, if $\tilde{b}_k - \tilde{a}_k = 1$, $\lfloor \tilde{b}_k - (\tilde{b}_k - \tilde{a}_k)/2 \rfloor_{\mathbb{Z}} = 0$ and $\lceil \tilde{a}_k + (\tilde{b}_k - \tilde{a}_k)/2 \rceil_{\mathbb{Z}} = 1$.

Bounding: Bounding is applied to ensure the convergence for DBRnB via improving the upper and lower bounds of η. Due to the monotonicity, the upper and lower bounds of η corresponding to box $[\mathbf{a}';\mathbf{b}'] \triangleq [\underline{\mathbf{s}},\underline{\mathbf{x}},\underline{\mathbf{r}},\underline{\mathbf{t}};\overline{\mathbf{s}},\overline{\mathbf{x}},\overline{\mathbf{r}},\overline{\mathbf{t}}]$ can be easily found as $\eta_{\mathrm{up}}([\mathbf{a}';\mathbf{b}']) \triangleq \frac{\sum_{k\in\mathcal{K}} \overline{r}_k}{f_{\mathrm{P}}(\underline{\mathbf{s}},\underline{\mathbf{x}},\underline{\mathbf{r}},\underline{\mathbf{t}})}$ and $\eta_{\mathrm{low}}([\mathbf{a}';\mathbf{b}']) \triangleq \frac{\sum_{k\in\mathcal{K}} \underline{r}_k}{f_{\mathrm{P}}(\overline{\mathbf{s}},\overline{\mathbf{x}},\overline{\mathbf{r}},\overline{\mathbf{t}})}$, respectively. These bounds are then used to update η^{best} and remove the boxes whose upper bound is smaller than η^{best} [21].

A DBRnB algorithm basically stores sequences of boxes until an optimal solution is found which requires some memory capacity. To reduce the memory requirement we can remove boxes which contain no feasible point. For (11), we make a simple feasibility condition based on the fronthaul constraints given as

$$\textstyle\sum_{k\in\mathcal{K}} \underline{r}_k \le \sum_{b\in\mathcal{B}} \overline{s}_b \bar{C}_b, \tag{13}$$

i.e. if a box does not satisfy (13), it contains no feasible point. Here (13) is due to $\sum_{b\in\mathcal{B}} \overline{s}_b \bar{C}_b \ge \sum_{b\in\mathcal{B}}\sum_{k\in\mathcal{K}} x_{b,k} r_k \ge \sum_{k\in\mathcal{K}} r_k \sum_{b\in\mathcal{B}} x_{b,k} \ge \sum_{k\in\mathcal{K}} \underline{r}_k$ where the last inequality follows (10f).

The standard DBRnB procedure would require a large number of iterations for convergence. In the following, we provide some modifications derived based on the specific structure of (11), which significantly improve the convergence performance.

Algorithm 1 The proposed DBRnB-based algorithm solving (11)

Initialization: Determine $\mathbf{a}, \mathbf{b}$ such that $\mathcal{S} \subseteq [\mathbf{a};\mathbf{b}]$. Set $\eta^{\text{best}} = 0$ and $\mathcal{R} = \mathrm{red}([\mathbf{a};\mathbf{b}])$.
repeat
  Branching: select $[\tilde{\mathbf{a}};\tilde{\mathbf{b}}]$, then create $[\tilde{\mathbf{a}};\tilde{\mathbf{b}}']$ and $[\tilde{\mathbf{a}}';\tilde{\mathbf{b}}]$. $\mathcal{R} := \mathcal{R} \setminus \{[\tilde{\mathbf{a}};\tilde{\mathbf{b}}]\}$.
  Reduction: determine $V_1 = \mathrm{red}([\tilde{\mathbf{a}};\tilde{\mathbf{b}}'])$, $V_2 = \mathrm{red}([\tilde{\mathbf{a}}';\tilde{\mathbf{b}}])$
  Bounding: For each box V_n, $n = 1,2$, not violating (13)
  if (14) is feasible **then**
    Calculate $\mathbf{t}^*$ and extract $\mathbf{x}^*$from the optimal solution of (14) $(\mathbf{w}^*,\mathbf{u}^*)$.
    Update $\underline{\mathbf{t}} := \mathbf{t}^*$ then calculate $\eta_{\mathrm{up}}(V_n)$ using (16).
    Check $\mathbf{x}^*$ with (18), if true, calculate $\eta_{\mathrm{low}}(V_n)$ using (17) then update $\eta^{\text{best}} := \max\{\eta_{\mathrm{low}}(V_n), \eta^{\text{best}}\}$, otherwise $\eta_{\mathrm{low}}(V_n) = \frac{\sum_{k\in\mathcal{K}} \underline{r}_k}{f_{\mathrm{P}}(\overline{\mathbf{s}},\overline{\mathbf{x}},\overline{\mathbf{r}},\overline{\mathbf{t}})}$.
    Find the set of boxes which do not contain optimal solution $\tilde{\mathcal{R}} = \{V_i \in \mathcal{R} | \eta_{\mathrm{up}}(V_i) < \eta^{\text{best}}\}$.
    Update $\mathcal{R} := \{V_n | \eta_{\mathrm{up}}(V_n) \ge \eta^{\text{best}}\} \cup \mathcal{R} \setminus \tilde{\mathcal{R}}$.
  end if
until Convergence

Improved Convergence Modifications

Improved Branching: For (11), we can skip branching on $\mathbf{t}$. Specifically, let us consider the following second-order-cone program

$$\underset{\mathbf{w},\mathbf{u}}{\text{minimize}} \quad \textstyle\sum_{b\in\mathcal{B}}\sum_{i=1}^{I} u_{b,i} \tag{14a}$$

$$\text{subject to } \Re(\mathbf{h}_k\mathbf{w}_k) \ge \sqrt{(e^{\underline{r}_k}-1)(\textstyle\sum_{j\ne k}^{K}|\mathbf{h}_k\mathbf{w}_j|^2+\sigma^2)} \tag{14b}$$

$$||\tilde{\mathbf{w}}_{b,i}||_2 \le u_{b,i},\ \underline{s}_b\underline{t}_b \le \textstyle\sum_{i=1}^{I} u_{b,i} \le \overline{s}_b\overline{t}_b \tag{14c}$$

$$||\tilde{\mathbf{w}}_{b,i}||_2^2 \le \overline{s}_b P_{\mathrm{a}},\ \|\mathbf{w}_{b,k}\|_2^2 \le \overline{x}_{b,k}\bar{P}_b, b\in\mathcal{B} \tag{14d}$$

$$\textstyle\sum_{k\in\mathcal{K}}\|\mathbf{w}_{bk}\|_2^2 \le \overline{s}_b\bar{P}_b\ \forall b\in\mathcal{B} \tag{14e}$$

which can be viewed as minimizing the power consumption subject to minimum users' rate requirement $\underline{\mathbf{r}}$. Let $\mathbf{u}^*$ denote the optimal solution (if (14) is feasible), and $\mathbf{t}^* \triangleq \{\sum_{i=1}^{I} u_{b,i}^*\}_b$. Obviously $\mathbf{1}^T\mathbf{t}^*$ (where $\mathbf{1}$ is the all-one vector) is the minimum power required to achieve $\underline{\mathbf{r}}$, and it holds $\underline{\mathbf{t}} \leq \mathbf{t}^*$. In addition, $\mathbf{t}^*$ is unique since the objective is the epigraph of $\sum_{b\in\mathcal{B}}\sum_{i=1}^{I}||\tilde{\mathbf{w}}_{b,i}||_2$ [28, Chap. 3]. Hence we can replace $\underline{\mathbf{t}}$ by $\mathbf{t}^*$ to obtain a tighter lower bound of $\mathbf{t}$. Therefore, we can only branch $(\mathbf{s}, \mathbf{x}, \mathbf{r})$, since $\mathbf{t}^*$ is always improved with $\underline{\mathbf{r}}$.

Improved Branching Order: We can potentially reduce the computational complexity if we opt to branch $\mathbf{s}$ first due to its dependency on other factors. Intuitively, the number of active RRHs provides the degree-of-freedoms that can make the desired data rate $\underline{\mathbf{r}}$ achievable. Moreover, we immediately obtain $x_{b,k} = 0, \forall k \in \mathcal{K}$ whenever $s_b = 0$. Therefore by first keeping branching on $\mathbf{s}$ until $\underline{\mathbf{s}} = \overline{\mathbf{s}}$, we can quickly remove combinations of $\{s_b\}_b$ infeasible to (11). This is done by solving (14) with given $\overline{\mathbf{s}}$ and target rate r_0.

Improved Bounds: We now present a way to obtain tighter upper and lower bounds of η. Let us consider box $[\mathbf{a}; \mathbf{b}] = [\underline{\mathbf{s}}, \underline{\mathbf{x}}, \underline{\mathbf{r}}, \underline{\mathbf{t}}; \overline{\mathbf{s}}, \overline{\mathbf{x}}, \overline{\mathbf{r}}, \overline{\mathbf{t}}]$. An upper bound of $\tilde{f}_{\mathrm{P}}(\underline{\mathbf{x}}, \underline{\mathbf{s}}, \underline{\mathbf{r}}, \underline{\mathbf{t}})$ is

$$\begin{aligned} \underline{f}_{\mathrm{P}}(\underline{\mathbf{x}}, \underline{\mathbf{s}}, \underline{\mathbf{r}}, \mathbf{t}^*) &\triangleq \tilde{\epsilon}\mathbf{1}^T\mathbf{t}^* + \Delta P \max\{1, \textstyle\sum_{b\in\mathcal{B}}\underline{s}_b\} \\ &+ p_{\mathrm{SP}} \max\{\textstyle\sum_{k\in\mathcal{K}}\underline{r}_k, \sum_{b\in\mathcal{B}}\sum_{k\in\mathcal{K}}\underline{x}_{b,k}\underline{r}_k\} + P_{\mathrm{const}} \end{aligned} \tag{15}$$

where $\mathbf{t}^*$ is determined via (14) (if feasible); the second term is due to the fact that at least one RRH is active; and the third term is achieved by the following inequality $\sum_{b\in\mathcal{B}}(\sum_{k\in\mathcal{K}} x_{b,k}r_k) \geq \sum_{k\in\mathcal{K}} r_k(\sum_{b\in\mathcal{B}} x_{b,k}) \geq \sum_{k\in\mathcal{K}} r_k$. Obviously, $\tilde{f}_{\mathrm{P}}(\underline{\mathbf{x}}, \underline{\mathbf{s}}, \underline{\mathbf{r}}, \underline{\mathbf{t}}) \leq \underline{f}_{\mathrm{P}}(\underline{\mathbf{x}}, \underline{\mathbf{s}}, \underline{\mathbf{r}}, \mathbf{t}^*)$, and replacing $\tilde{f}_{\mathrm{P}}(\underline{\mathbf{s}}, \underline{\mathbf{x}}, \underline{\mathbf{r}}, \underline{\mathbf{t}})$ by $\underline{f}_{\mathrm{P}}(\underline{\mathbf{x}}, \underline{\mathbf{s}}, \underline{\mathbf{r}}, \mathbf{t}^*)$ does not remove any feasible solution. Thus, a tighter upper bound of η over $[\mathbf{a}; \mathbf{b}]$ can be calculated as

$$\eta_{\mathrm{up}}([\mathbf{a}; \mathbf{b}]) = \frac{\sum_{k\in\mathcal{K}} \overline{r}_k}{\underline{f}_{\mathrm{P}}(\underline{\mathbf{x}}, \underline{\mathbf{s}}, \underline{\mathbf{r}}, \mathbf{t}^*)}. \tag{16}$$

Similarly, suppose $(\hat{\mathbf{s}}, \hat{\mathbf{x}}, \underline{\mathbf{r}}, \hat{\mathbf{t}})$ to be some feasible point within $[\mathbf{a}; \mathbf{b}]$. We can easily check that $f_{\mathrm{P}}(\hat{\mathbf{s}}, \hat{\mathbf{x}}, \underline{\mathbf{r}}, \hat{\mathbf{t}}) \leq f_{\mathrm{P}}(\overline{\mathbf{s}}, \overline{\mathbf{x}}, \overline{\mathbf{r}}, \overline{\mathbf{t}})$ due to the monotonicity property. Then an improved lower bound of η over $[\mathbf{a}; \mathbf{b}]$ can be obtained as

$$\eta_{\mathrm{low}}([\mathbf{a}; \mathbf{b}]) = \frac{\sum_{k\in\mathcal{K}} \underline{r}_k}{\tilde{f}_{\mathrm{P}}(\hat{\mathbf{s}}, \hat{\mathbf{x}}, \underline{\mathbf{r}}, \hat{\mathbf{t}})}. \tag{17}$$

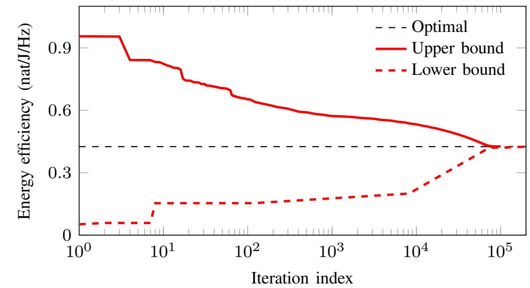
Note that if $\eta_{\mathrm{low}}([\mathbf{a}; \mathbf{b}]) > \eta^{\mathrm{best}}$, we update $\eta^{\mathrm{best}} := \eta_{\mathrm{low}}([\mathbf{a}; \mathbf{b}])$ and remove boxes whose upper bound are smaller than η^{best} (i.e. Step 11 in Alg. 1). Thus, obtaining a feasible point is vital to improving the algorithm's efficiency. For this purpose we propose a simple heuristic trick which may quickly find a feasible point in $[\mathbf{a}; \mathbf{b}]$. First, we note that a feasible point $(\hat{\mathbf{s}}, \hat{\mathbf{x}}, \underline{\mathbf{r}}, \hat{\mathbf{t}})$ must satisfy two conditions: $\underline{\mathbf{r}}$ is achievable by $(\hat{\mathbf{s}}, \hat{\mathbf{x}}, \hat{\mathbf{t}})$; and

$$\hat{\mathbf{x}} \in \{\mathbf{x} \mid \textstyle\sum_{b\in\mathcal{B}} x_{b,k} \geq 1, k \in \mathcal{K}, \sum_{k\in\mathcal{K}} x_{b,k}\underline{r}_k \leq \bar{C}_b, b \in \mathcal{B}\} \tag{18}$$

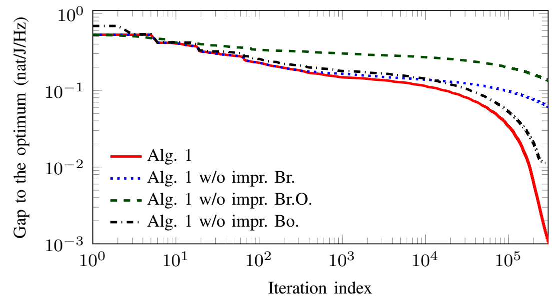
Second, it is easily seen that the point returned by solving (14) always satisfies the first condition. Consequently, we can extract $\hat{\mathbf{x}}$ from the optimal solution of (14), and then verify (18).

4. NUMERICAL RESULTS

We consider a simulation model as follows. The distance between the RRHs is 200 m; the channel $\mathbf{h}_{b,k}$ between RRH b and user k is generated as $\mathbf{h}_{b,k} \sim \mathcal{CN}(0, \rho_{b,k}\mathbf{I}_I)$, where $\rho_{b,k}$ represents the large-scale fading and is calculated as $\rho_{b,k}[\mathrm{dB}] = 30\log_{10}(d_{b,i_j}) + 38 + \mathcal{N}(0, 8)$ (d_{b,i_j} is the distance in meters); the system bandwidth is 10 MHz; the noise power is -143 dBW; we take $B = 3$, $K = 4$, $\bar{P} = \bar{P}_b = IP_{\mathrm{a}} = 30$ dBm and $\bar{C}_b = \bar{C} = 10$ nats/s/Hz,$\forall b$; other parameters are set as follows: $P_{\mathrm{RRH}}^{\mathrm{active}} + P_{\mathrm{FH}}^{\mathrm{active}} = 10.65$ W [22, 23], $P_{\mathrm{RRH}}^{\mathrm{sleep}} + P_{\mathrm{FH}}^{\mathrm{sleep}} = 5.05$ W [22, 23], P_{ms}=0.1 W, $\epsilon_{\max} = 0.55$ [20], $p_{\mathrm{SP}} = 10$ W/(Gnats/Hz), $I = 2$, and $r_0 = 1$.



(a) Convergence of the upper and lower bounds.



(b) Convergence speed with and without the proposed mofications.

Fig. 1. Convergence behavior of Algorithm 1 for one channel realization.

Fig. 1 plots convergence performance of Alg. 1 for a random channel realization. Particularly, Fig. 1(a) depicts the upper and lower bounds over iterations. It is seen that the bounds monotonically converge to the optimal value. Fig. 1(b) shows the convergence speed via the gap between the upper bound and the optimal value. In this figure we also provide the performance of other schemes to confirm the effectiveness of the proposed modifications. Specifically, the schemes labelled '*w/o. impr. Br.*', '*w/o. impr. Br-O.*', '*w/o. impr. Bo.*' represent for Alg. 1 without applying improved branching, improved branching order and improved bounding, respectively. The results clearly demonstrate that applying the proposed modifications remarkably improves the convergence performance.

5. CONCLUSION

We have studied the joint design of beamforming and RRH-user association in CRANs which maximizes the system EE subject to per-RRH fronthaul capacity, transmit power budget and per-user QoS. We have adopted the power consumption model wherein the impacts of rate-dependent signal processing power and the dynamics of the PA's efficiency are included. To investigate the optimal performance we have developed the new globally optimal method based on the DBRnB framework. We have also provided the useful modifications which improve the algorithm's efficiency. Numerical evaluations have been provided to confirm the effectiveness of the proposed algorithm. The algorithm can be implemented in small-scale networks or serve as benchmark in evaluating low-complexity suboptimal solutions, which are considered in our future work.

6. REFERENCES

[1] C. Mobile, "C-RAN: the road towards green RAN," *White Paper*, vol. 2, 2011.

[2] J. Wu, "Green wireless communications: from concept to reality [industry perspectives]," *IEEE Wireless Commun.*, vol. 19, no. 4, pp. 4–5, Aug. 2012.

[3] M. Peng, C. Wang, V. Lau, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: insights and challenges," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 152–160, Apr. 2015.

[4] B. Dai and W. Yu, "Sparse beamforming and user-centric clustering for downlink cloud radio access network," *IEEE Access*, vol. 2, no. 6, pp. 1326–1339, Oct. 2014.

[5] ——, "Energy efficiency of downlink transmission strategies for cloud radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 1037–1050, April 2016.

[6] T. X. Tran and D. Pompili, "Dynamic radio cooperation for user-centric cloud-RAN with computing resource sharing," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2379–2393, Apr. 2017.

[7] F. Zhuang and V. K. N. Lau, "Backhaul limited asymmetric cooperation for MIMO cellular networks via semidefinite relaxation," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 684–693, Feb. 2014.

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

[9] D. W. K. Ng, E. S. 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.

[10] 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, Dec. 2013.

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

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

[13] A. Zappone, E. Björnson, L. Sanguinetti, and E. Jorswieck, "Globally optimal energy-efficient power control and receiver design in wireless networks," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2844–2859, Jun. 2017.

[14] C. Isheden and G. P. Fettweis, "Energy-efficient multi-carrier link adaptation with sum rate-dependent circuit power," in *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, Dec. 2010, pp. 1–6.

[15] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient resource allocation in OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 60, no. 12, pp. 3767–3778, Dec. 2012.

[16] T. Wang and L. Vandendorpe, "On the optimum energy efficiency for flat-fading channels with rate-dependent circuit power," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4910–4921, Dec. 2013.

[17] 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.

[18] S. Mikami, T. Takeuchi, H. Kawaguchi, C. Ohta, and M. Yoshimoto, "An efficiency degradation model of power amplifier and the impact against transmission power control for wireless sensor networks," in *2007 IEEE Radio and Wireless Symposium,*, 2007, pp. 447–450.

[19] E. Björnemo, "Energy constrained wireless sensor networks: Communication principles and sensing aspects," Ph.D. dissertation, Institutionen för teknikvetenskaper, 2009.

[20] D. Persson, T. Eriksson, and E. G. Larsson, "Amplifier-aware multiple-input multiple-output power allocation," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1112–1115, Jun. 2013.

[21] H. Tuy, M. Minoux, and N. Hoai-Phuong, "Discrete monotonic optimization with application to a discrete location problem," *SIAM Journal on Optimization*, vol. 17, no. 1, pp. 78–97, 2006.

[22] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.

[23] A. R. Dhaini, P. H. Ho, G. Shen, and B. Shihada, "Energy efficiency in TDMA-Based next-generation passive optical access networks," *IEEE/ACM Trans. Netw.*, vol. 22, no. 3, pp. 850–863, Jun. 2014.

[24] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *Proc.19th Future Network & Mobile Summit (ICT Summit'10)*, Florence, Italy, Jun. 2010, pp. 1–8.

[25] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2508–2523, May 2012.

[26] E. A. Jorswieck and E. G. Larsson, "Monotonic optimization framework for the two-user MISO interference channel," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2159–2168, Jul. 2010.

[27] H. Tuy, F. Al-Khayyal, and P. T. Thach, "Monotonic optimization: Branch and cut methods," in *Essays and Surveys in Global Optimization*. Springer, 2005, pp. 39–78.

[28] J. Dattorro, *Convex optimization & Euclidean distance geometry*. Lulu. com, 2010.