Efficient Algorithms for Sum Rate Maximization in Fronthaul-Constrained C-RANs

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 University of Oulu, 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*—We consider downlink transmission of a fronthaul-constrained cloud radio access network. Our aim is to maximize the system sum data rate via jointly designing beamforming and user association. The problem is basically a mixed integer non-convex programs for which a global solution requires a prohibitively high computational effort. The focus is thus on efficient solutions capable of achieving the near optimal performance with low complexity. To this end, we transform the design problem into continuous programs by two approaches: penalty and sparse approximation methods. The resulting continuous nonconvex problems are then solved by the successive convex approximation framework. Numerical results indicate that the proposed methods are near-optimal, and outperform existing suboptimal methods in terms of achieved performances and computational complexity.**

I. Introduction

In recent years, *cloud radio access networks (C-RANs)* has received significant attention due to its potential of implementing the fifth-generation (5G) standard [1]. Therein, the baseband (BB) signal processing units are no longer equipped at base stations (BSs) but migrated at a central cloud computing platform, called BB unit (BBU) pool. As a result, BSs in C-RANs mainly account for the wireless interface of the network, and now referred to as remote radio heads (RRHs). However, the BB signals from the BBU pool need to be delivered to the RRHs through the fronthaul links of finite capacity. This forms one of the main practical challenges of the C-RAN designs [2]–[4].

Sum rate maximization for capacity-limited fronthaul C-RANs has been investigated in [5]–[7]. Therein, to cope with the constraint of fronthaul capacity, the strategies of selecting a proper set of users that can be served by an RRH were considered. This leads to the approaches of joint designs transmit beamforming/power and RRH-user association. However, such design problems are basically mixed integer non-convex programs (MINPs) which are difficult to solve globally [7]. Thus most of the related works focused on suboptimal yet low complexity solutions. In particular, in [5], a combination of reweighted ℓ_1-norm and alternating optimization was proposed. Nevertheless, its convergence is not analytically guaranteed. To overcome the convergence issue, the authors in [6] devised a method based on a combination of sparse optimization and successive convex approximation (SCA). However, it is numerically observed that these mentioned methods are far from optimal (see Fig. 1). In addition, they are sort of two-layer iterative procedure which might require high computational cost.

Motivated by the above discussions, this paper aims at proposing efficient methods for joint designs of transmit beamforming and RRH-user association in the downlink C-RAN. In particular, we overcome the difficulty of combination by transforming the considered MINP into continuous problems based on two approaches: penalty method and ℓ_0-approximation. In the first approach, we represent the discrete set by a set of constraints involving continuous variables, and then apply the penalty method to solve the resulting problem. In the second, the binary variables are approximated by an ℓ_0 approximation function [8]. In both approaches, the obtained continuous problems are nonconvex, which are then solved by the SCA technique. The proposed methods are guaranteed to converge. The numerical results demonstrate that the proposed methods are capable of achieving near-optimal performance. In addition, they are superior to the aforementioned methods in both terms of achieved sum rate and computational complexity.

II. System Model and Problem Formulation

A. System Model

We consider the downlink of a multiuser C-RAN system in which a set of B RRHs, each equipped with M antennas, cooperatively serves a set of K single-antenna users following the *coordinated multipoint joint transmission* (CoMP-JT) scheme, i.e., a user simultaneously receive data from multiple RRHs. Let us denote by $\mathcal{B} \triangleq \{1,\dots,B\}$ the set of RRHs and by $\mathcal{K} \triangleq \{1,\dots,K\}$ the set of users. All RRHs are connected to a common BBU pool which is responsible BB signal processing. It is assumed that the BBU pool has perfect channel state information (CSI) associated with all the users in the network. Data for each user is shared from BBU pool to its serving RRHs through fronthaul links. Let s_k denote the data symbol intended for user k which has unit-energy, i.e., $\mathbb{E}[|s_k|^2] = 1$, and assume that linear beamforming scheme is adopted to form transmit signals. Considering flat fading channels, the received signal at user k can be written as

$$\begin{aligned} y_k &= \underbrace{\Big(\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k}\mathbf{w}_{b,k}\Big)s_k}_{\text{desired signal}} + \underbrace{\sum_{j\in\mathcal{K}\setminus k}\Big(\sum_{b\in\mathcal{B}} \mathbf{h}_{b,k}\mathbf{w}_{b,j}\Big)s_j}_{\text{interference}} + n_k \\ &= \mathbf{h}_k\mathbf{w}_k s_k + \sum_{j\in\mathcal{K}\setminus k}\Big(\sum_{b\in\mathcal{B}} \mathbf{h}_k\mathbf{w}_j\Big)s_j + n_k \end{aligned} \qquad (1)$$

where $\mathbf{w}_{b,k} \in \mathbb{C}^{M\times 1}$ denote the beamforming vector from RRH b to user k, $\mathbf{h}_{b,k} \in \mathbb{C}^{1\times M}$ is the channel between RRH b and user k, and $n_k \sim \mathcal{CN}(0,\sigma_k^2)$ is the additive white Gaussian noise at user k. In (1), we have denoted $\mathbf{h}_k \triangleq [\mathbf{h}_{1,k},\mathbf{h}_{2,k},\ldots,\mathbf{h}_{B,k}] \in \mathbb{C}^{1\times MB}$ and $\mathbf{w}_k \triangleq [\mathbf{w}_{1,k}^{\mathrm{T}},\mathbf{w}_{2,k}^{\mathrm{T}},\ldots,\mathbf{w}_{B,k}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{MB\times 1}$, for notational convenience. Assuming single-user decoding, i.e. interference among users is treated as Gaussian noise, the SINR at user k can be written as

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

where $\mathbf{w}$ is the beamforming vector stacking all $\mathbf{w}_k$.

B. Problem Formulation

We jointly design beamforming vectors and RRH-user association with the aim of maximizing the sum data rate. The problem of interest reads

$$\begin{aligned}
\underset{\mathbf{w},\mathbf{x},\mathbf{r}}{\text{maximize}}\quad & \sum_{k\in\mathcal{K}} r_k && \text{(2a)}\\
\text{subject to}\quad & r_k \geq Q_k,\ \forall k\in\mathcal{K} && \text{(2b)}\\
& \sum_{k\in\mathcal{K}}\|\mathbf{w}_{b,k}\|_2^2 \leq \bar{P}_b,\ \forall b\in\mathcal{B} && \text{(2c)}\\
& \|\mathbf{w}_{b,k}\|_2^2 \leq x_{b,k}\bar{P}_b,\ \forall k\in\mathcal{K}, b\in\mathcal{B} && \text{(2d)}\\
& r_k \leq \log(1+\gamma_k(\mathbf{w})),\ \forall k\in\mathcal{K} && \text{(2e)}\\
& \sum_{k\in\mathcal{K}} x_{b,k} r_k \leq \bar{C}_b,\ \forall b\in\mathcal{B} && \text{(2f)}
\end{aligned}$$

where $\mathbf{x} \triangleq [x_{1,1}\ldots,x_{b,k},\ldots x_{B,K}]^{\mathrm{T}} \in \{0,1\}^{BK}$, $\mathbf{r} \triangleq [r_1,\ldots,r_K]^{\mathrm{T}}$. Here, r_k is the achievable data rate transmitted to user k; $x_{b,k} \in \{0,1\}$ represents the association between RRH b and user k, i.e., $x_{b,k}=1$ indicates that user k receives data from RRH b and $x_{b,k}=0$ otherwise. The constraints in (2b) are to guarantee that the data rate of user k always meets its required quality-of-service Q_k. The constraints in (2c) represents the total transmit power at each individual RRH. The constraints in (2d) ensure that $\|\mathbf{w}_{b,k}\|_2^2=0$ when RRH b does not serve user k, i.e., $x_{b,k}=0$. The constraints in (2e) are for feasible transmission on wireless channels. Finally, the constraints in (2f) mean that the total data rate transmitted over a fronthaul link does not exceed its capacity $\bar{C}_b$.

Problem (2) is an MINP generally known to be NP-hard. Globally optimal method for this problem was investigated in [7], but it is for benchmarking purposes only. In the following, we present two approaches efficiently solving (2).

III. Proposed Algorithms

A. Penalty Method

In the first method, we use a set of continuous functions to equivalently represent binary variables, then apply a penalty method. In particular, we recall the following well-known relaxation of binary variables to represent $\mathbf{x}$, i.e.,

$$\sum_{b\in\mathcal{B}}\sum_{k\in\mathcal{K}} x_{b,k}^2 - x_{b,k} \geq 0,\ x_{b,k}\in[0,1],\forall b,k. \tag{3}$$

It can easily be seen that (3) holds true if and only if $x_{b,k},\forall b,k$ admits binary value $\{0,1\}$. With (3), we can treat $\mathbf{x}$ as continuous variable. Next, to facilitate the application of the SCA framework in solving (2), we rewrite the problem as

$$\begin{aligned}
\underset{\substack{\mathbf{w},\mathbf{x},\mathbf{r},\\ \mathbf{g},\mathbf{q}}}{\text{maximize}}\quad & \sum_{k\in\mathcal{K}} r_k && \text{(4a)}\\
\text{subject to}\quad & \log(1+g_k)\geq r_k\ \forall k\in\mathcal{K} && \text{(4b)}\\
& q_k \geq \textstyle\sum_{j\in\mathcal{K}\setminus k}|\mathbf{h}_k\mathbf{w}_j|^2+\sigma_k^2,\forall k\in\mathcal{K} && \text{(4c)}\\
& q_k g_k \leq |\mathbf{h}_k\mathbf{w}_k|^2\ \forall k\in\mathcal{K} && \text{(4d)}\\
& \|\mathbf{w}_{b,k}\|_2^2 \leq x_{b,k}^p\bar{P}_b, && \text{(4e)}\\
& \text{(2b)},\text{(2c)},\text{(2f)},\text{(3)} && \text{(4f)}
\end{aligned}$$

where $\mathbf{g} \triangleq [g_1,\ldots,g_K]^{\mathrm{T}}$ and $\mathbf{q} \triangleq [q_1,\ldots,q_K]^{\mathrm{T}}$ are newly introduced slack variables. We note that (4) is the epigraph of (2) and thus it maintains the feasible set of the original problem. Here, we have replaced (2d) by (4e) in the above transformation, where $p>1$. This maneuver is motivated by the expectation that solving (4) will eventually return binary solutions, as discussed in [9].

At this point, we can apply the SCA technique to solve the continuous problem (4). However, finding an initial point to start the iterative process is usually challenging due to the association vector and the fronthaul constraint. To tackle this issue, we apply a penalty method which results in the following regularized problem

$$\begin{aligned}
\underset{\mathbf{\Lambda}}{\text{maximize}}\quad & \phi(\mathbf{\Lambda},\alpha,\xi) \triangleq \sum_{k\in\mathcal{K}} r_k + \alpha\sum_{b\in\mathcal{B},k\in\mathcal{K}}(x_{b,k}^2-x_{b,k})\\
& \qquad + \xi\sum_{b\in\mathcal{B}}\min\{0,\bar{C}_b-v_b\} && \text{(5a)}\\
\text{subject to}\quad & \sum_{k\in\mathcal{K}} x_{b,k}r_k \leq v_b,\ v_b\in[0,\bar{C}_b],\ \forall b\in\mathcal{B} && \text{(5b)}\\
& \text{(2b)},\text{(2c)},\text{(4b)}-\text{(4e)}, && \text{(5c)}
\end{aligned}$$

where $\mathbf{v} \triangleq [v_1,\ldots,v_B]^{\mathrm{T}}$ is the vector of newly introduced variables, $\mathbf{\Lambda} \triangleq \{\mathbf{w},\mathbf{x},\mathbf{r},\mathbf{g},\mathbf{q},\mathbf{v}\}$, and $\alpha,\xi>0$ are the penalty parameters. In $\phi(\mathbf{\Lambda},\alpha,\xi)$, we have added two penalty terms to minimize the costs when $\mathbf{x}$ is not a binary vector, and when the fronthaul constraints are violated, respectively.

Now the difficulty in solving (5) lies in the nonconvex constraints (4d), (4e), (5b), and the objective function. In light of the SCA principle [10], (4d), (4e) and (5b) can be approximated as

$$\frac{\lambda_k^{(n)}g_k^2}{2}+\frac{q_k^2}{2\lambda_k^{(n)}} \leq 2\Re(\mathbf{w}_k^{(n)}\mathbf{h}_k^H\mathbf{h}_k\mathbf{w}_k)-\|\mathbf{h}_k\mathbf{w}_k^{(n)}\|_2^2, \tag{6}$$

$$\|\mathbf{w}_{b,k}\|_2^2 \leq (p(x_{b,k}^{(n)})^{p-1}x_{b,k}+(1-p)(x_{b,k}^{(n)})^p)\bar{P}_b,\ \forall b,k \tag{7}$$

$$\begin{aligned}\sum_{k\in\mathcal{K}}(x_{b,k}+r_k)^2 \leq \sum_{k\in\mathcal{K}}\big(&2x_{b,k}^{(n)}x_{b,k}+2r_k^{(n)}r_k\\ &-(x_{b,k}^{(n)})^2-(r_k^{(n)})^2+2v_b\big)\end{aligned} \tag{8}$$

where $\lambda_k^{(n)} = \frac{g_k^{(n)}}{q_k^{(n)}}$, and the superscript n is the iteration counter. Remark that (6) follows the result in [11], and

Algorithm 1 Proposed method for solving (4)

1: **Initialization:** Set $n := 0$, initialize $\mathbf{\Lambda}^{(0)}$ and set $\alpha^{(0)}$ small
2: **repeat** $\{n := n+1\}$
3: Solve (9) and achieve $\mathbf{\Lambda}^{*}$, then update $\mathbf{\Lambda}^{(n)} := \mathbf{\Lambda}^{*}$
4: Update $\alpha^{(n)} := \min\{\alpha_{\max}; \alpha^{(n-1)} + \varepsilon\}$ for small ε
5: **until** Convergence

(8) is due to the fact that $2x_{b,k}r_k = (x_{b,k} + r_k)^2 - x_{b,k}^2 - r_k^2$. In addition, the objective function $\phi(\mathbf{\Lambda}, \alpha, \xi)$ can also be convexified using the first order approximation, i.e., $\phi(\mathbf{\Lambda}, \alpha, \xi; \mathbf{\Lambda}^{(n)}) \triangleq \sum_{k\in\mathcal{K}} r_k + \alpha \sum_{b\in\mathcal{B},k\in\mathcal{K}}(2x_{b,k}x_{b,k}^{(n)} - (x_{b,k}^{(n)})^2 - x_{b,k}) + \xi \sum_{b\in\mathcal{B}} \min\{0, \bar{C}_b - \vartheta_b\}$. To summarize, the following approximate convex program will be solved at every iteration of the proposed SCA procedure

$$\max_{\mathbf{\Lambda}} \ \phi(\mathbf{\Lambda}, \alpha, \xi; \mathbf{\Lambda}^{(n)}) \text{ s.t. } \{(2b), (2c), (4b), (4c), (6)-(7)\} \quad (9)$$

Algorithm 1 describes the SCA procedure solving (4), and its convergence can be proved following the arguments in [12, Section 2].

In Algorithm 1, we note that the value of penalty parameter α is increased at each iteration, i.e., Step 4. This is inspired by the fact that α provides the tightness of the relaxation, and a high value of α will encourage $x_{b,k}$ to take on binary values. Thus we let Algorithm 1 start with a small value of α to focus on maximizing the sum rate objective. Then α is increased in subsequent iterations to gradually force $x_{b,k}$ to be binary. We will illustrate by the simulation that $x_{b,k}$ approaches to binary value at the convergence (see Fig. 3).

B. ℓ_0-Approximation Method

In the second method, we formulate the joint design of beamforming and RRH-user association as finding a sparse solution of beamformer vector $\mathbf{w}$. In other words, the RRH-user association policy here is derived from the values of beamformers $\mathbf{w}_{b,k}$, instead of represented by binary variables $x_{b,k}$. To be specific, we consider the following beamforming design problem

$$\begin{aligned}
&\underset{\mathbf{w},\mathbf{r},\mathbf{u}}{\text{maximize}} \ \sum_{k\in\mathcal{K}} r_k && (10a)\\
&\text{subject to } \|\mathbf{w}_{b,k}\|_2 \le u_{b,k}, \ \sum_{k\in\mathcal{K}} u_{b,k}^2 \le \bar{P}_b, \ \forall b\in\mathcal{B} && (10b)\\
&\sum_{k\in\mathcal{K}} \psi(u_{b,k}) r_k \le \bar{C}_b, \ \forall b\in\mathcal{B} && (10c)\\
&(2b), (2e), && (10d)
\end{aligned}$$

where $u_{b,k}$ is the slack variable associated to the power of beamformer $\mathbf{w}_{b,k}$, and $\mathbf{u} \triangleq [u_{1,1}, \ldots, u_{b,k}, \ldots, u_{B,K}]^{\mathrm{T}}$; $\psi(u_{b,k})$ is the step function, i.e.

$$\psi(u_{b,k}) \triangleq \begin{cases} 1 & u_{b,k} > 0, \\ 0 & u_{b,k} = 0. \end{cases}$$

It is clear that RRH b serves user k if $u_{b,k} > 0$, and otherwise if $u_{b,k} = 0$. Thus problem (10) is in fact equivalent to (2).

The central idea of the second proposed method is to approximate $\psi(u_{b,k})$ by a continuous function as such continuous optimization technique can be applied to solve (10). To this purpose, we consider the (concave) approximation functions listed in Table I (see the top of the next page), which have often been used in sparse optimization problems in wireless communications [5], [6], [9]. It can easily be seen that $\psi(u_{b,k}) \approxeq \psi_\beta(u_{b,k})$ when β is sufficiently *large*. Therefore, we can approximate (10) by replacing $\psi(u_{b,k})$ by $\psi_\beta(u_{b,k})$ as

$$\begin{aligned}
&\underset{\mathbf{w},\mathbf{r},\mathbf{u}}{\text{maximize}} \ \sum_{k\in\mathcal{K}} r_k && (11a)\\
&\text{subject to } \sum_{k\in\mathcal{K}} \psi_\beta(u_{b,k}) r_k \le \bar{C}_b, \ \forall b\in\mathcal{B} && (11b)\\
&(2b), (2e), (10b) && (11c)
\end{aligned}$$

Toward solving (11), we use the following transformation

$$\begin{aligned}
&\underset{\substack{\mathbf{w},\mathbf{r},\mathbf{u},\tilde{\mathbf{x}}\\ \mathbf{g},\mathbf{q},}}{\text{maximize}} \ \sum_{k\in\mathcal{K}} r_k && (12a)\\
&\text{subject to } \sum_{k\in\mathcal{K}} \tilde{x}_{b,k} r_k \le \bar{C}_b, \forall b\in\mathcal{B} && (12b)\\
&\tilde{x}_{b,k} \ge \psi_\beta(u_{b,k}), \ \forall b\in\mathcal{B}, k\in\mathcal{K} && (12c)\\
&(2b), (4b)-(4d), (10b) && (12d)
\end{aligned}$$

where $\tilde{\mathbf{x}} \triangleq [\tilde{x}_{1,1} \ldots, \tilde{x}_{b,k}, \ldots \tilde{x}_{B,K}]^{\mathrm{T}}$, and $\mathbf{g}, \mathbf{q}$ are introduced exactly as those in the previous subsection. Note that the nonconvexity of (12) is due to constraints (4d), (12b) and (12c). Now we can readily apply the SCA to solve (12). Specifically, constraints (4d) and (12b) can be approximated as in (6) and (8), respectively. For (12c), its convex approximation is given by

$$\tilde{x}_{b,k} \ge \bar{\psi}_\beta(u_{b,k}; u_{b,k}^{(n)}) \quad (13)$$

where $\bar{\psi}_\beta(.)$ is provided in Table I.[1] Finally, we have the following approximate convex program of (10), i.e.,

$$\max_{\tilde{\mathbf{\Lambda}}} \ \sum_{k\in\mathcal{K}} r_k \text{ s.t. } \{(2b), (4b), (4c), (6), (8), (10b), (13)\}, \quad (14)$$

where $\tilde{\mathbf{\Lambda}} \triangleq \{\mathbf{w}, \mathbf{r}, \mathbf{u}, \mathbf{g}, \mathbf{q}, \tilde{\mathbf{x}}\}$.

The second proposed suboptimal method is outlined in Algorithm 2. Since the approximation parameter β provides the tightness of the approximate step functions in Table I, we also update β after every iteration which follows the similar idea as with Algorithm 1. In addition, to circumvent the initial guess issue, a penalty of violating the fronthaul constraints can be added to the objective of (14). Convergence of Algorithm 2 is guaranteed, which can be proved using the same arguments as in [9].

We remark that the value of $\tilde{x}_{b,k}$ returned by the SCA in Algorithm 2 is not ensured to be close to a binary value. Therefore, a simple post-processing scheme is applied at the output of the SCA to derive a feasible solution for problem

[1] For Capped-ℓ_1 function, $\psi_\beta(y)$ is concave and continuous but not smooth at $y = \frac{1}{\beta}$. However its convex upper bound can be derived based on the sub-differential $\psi_\beta(y)$ [8].

Table I
ℓ_0-APPROXIMATION FUNCTION $\psi_\beta(y)$, AND THE CORRESPONDING SUBGRADIENT $\partial\psi_\beta(y)$ AND FIRST ORDER APPROXIMATIONS $\bar{\psi}_\beta(y;y^{(n)})$ [8]

Approximation	Function $\psi_\beta(y)$	Subgradient $\partial\psi_\beta(y)$	First-order approximation $\bar{\psi}_\beta(y;y^{(n)})$
Exponential function (Exp)	$1-\exp(-\beta y)$	$\beta\exp(-\beta y)$	$1-\exp(-\beta y^{(n)})+\beta\exp(-\beta y^{(n)})(y-y^{(n)})$
Logarithmic function (Log)	$\frac{\log(1+\beta y)}{\log(1+\beta)}$	$\frac{\beta}{\log(1+\beta)(1+\beta y)}$	$\frac{1}{\log(1+\beta)}\left(\log(1+\beta y^{(n)})+\frac{\beta(y-y^{(n)})}{(1+\beta y^{(n)})}\right)$
Capped-ℓ_1 function	$\min\{1,\beta y\}$	$\begin{cases}0 & \text{if } y\geq\frac{1}{\beta}\\ \beta & \text{if otherwise}\end{cases}$	$\begin{cases}1 & \text{if } y^{(n)}\geq\frac{1}{\beta}\\ \beta y & \text{otherwise}\end{cases}$

Algorithm 2 Proposed method for solving (11)

1: **Initialization:** Set $n:=0$, choose initial values for $\tilde{\mathbf{\Lambda}}^{(0)}$ and set $\beta^{(0)}$ small
2: **repeat** $\{n:=n+1\}$
3: Solve (14) and achieve $\tilde{\mathbf{\Lambda}}^{*}$, then update $\tilde{\mathbf{\Lambda}}^{(n)}:=\tilde{\mathbf{\Lambda}}^{*}$
4: Update $\beta^{(n)}:=\min\{\beta_{\max};\beta^{(n-1)}+\varepsilon\}$ for small ε
5: **until** Convergence and output $\tilde{\mathbf{\Lambda}}^{*}$
6: Apply the post-processing procedure

(2). In particular, we first map $\tilde{x}_{b,k}$ to nearest binary value, denoted as $\bar{x}_{b,k}\in\{0,1\}$, and then recalculate beamforming vector and achieved data rate accordingly by (2d) and (2e), respectively. As shown at the end of the next section, the proposed approach outputs feasible solutions to (2) after the post-processing procedure with high probability. However, in the worst case where the fronthaul constraint is violated, we fix $\tilde{x}_{b,k}=\bar{x}_{b,k}$ and carry out the second loop of the SCA to determine feasible $\mathbf{w}$ and $\mathbf{r}$.

IV. Numerical Results

This section provides numerical illustrations to evaluate the effectiveness of the proposed methods. Here we consider a network with $B=3$ RRHs and $K=4$ users. The inter-RRH distance is 200 m. Each RRH is equipped with $M=2$ antennas. The channel $\mathbf{h}_{b,k}$ is assumed to be flat fading which is generated following Gaussian distribution, i.e., $\mathbf{h}_{b,k}\sim\mathcal{CN}(0,\rho_{b,k}\mathbf{I}_M)$, where $\rho_{b,k}$ represents the large-scale fading and is calculated as $\rho_{b,k}[\text{dB}]=30\log_{10}(D_{b,k}[\text{m}])+38+\mathcal{N}(0,8)$. The bandwidth is 10 MHz and the noise power is -143 dBW. Parameters $\bar{P}_b$ and $\bar{C}_b$ are set to be same for all RRHs, i.e., $\bar{P}_b=\bar{P}$ and $\bar{C}_b=\bar{C},\forall b$. The minimum required data rate for each user is $Q_k=1$ nat/s/Hz. To generate an initial point for starting the proposed methods, we fix all elements of the selection vector as one (i.e., full-cooperation) and create beamforming vector $\mathbf{w}^{(0)}$ satisfying power constraint (2c); then the values for the remaining variables are determined based on (4b)–(4d). For the penalty parameters, we take $\xi=1$, $p=15$, $\alpha_{\max}=10^3$, $\beta_{\max}=10^7$, and initialize $\alpha^{(0)}=10^{-2}$ and $\beta^{(0)}=0.1$. For comparison purposes, we provide performances of the existing methods in [5] (dubbed as '*Sparse-WMMSE*') and [6] (dubbed as '*Sparse-SCA*'). All the convex programs in this paper are solved by MOSEK solver in MATLAB environment.

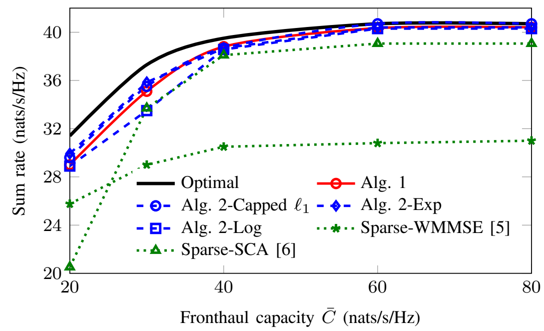

Fig. 1. Average sum rate performances of the considered schemes with $\bar{P}=30$ dBm.

In Fig. 1, we compare the achieved average performance of the proposed approaches to the optimal one in [7], and that of the existing schemes. The figure plots the average sum rate of the considered schemes versus the fronthaul capacity $\bar{C}$. We can see that the proposed methods achieve the near-optimal performance and remarkably outperform the existing schemes, especially in small regime of $\bar{C}$. This validates the effectiveness of the proposed schemes in terms of achieved sum rate.

Fig. 2 shows the complexity of the considered methods in terms of the average total run time for convergence as a function of $\bar{C}$. We first observe that the proposed algorithms take smaller run time to output the solution compared to the existing ones. This result again confirms the advantage of the proposed approaches. We can also see that, except Sparse-WMMSE, the run time for all considered schemes reduces as $\bar{C}$ increases. This can be explained as follows. A higher value of $\bar{C}$ allows more users to be served by an RRH which facilitates the association process. As a result, the optimization procedure takes fewer operations to arrive at the optimal selection vector, and the run time is reduced. For the Sparse-WMMSE scheme, we note that the RRH-user association is not simultaneously updated with the beamforming at each iteration due to the alternative optimization procedure. Hence it may require more iterations for convergence, especially when the feasible set is expanded as $\bar{C}$ grows. Another observation is that run time for Algorithm 1 may be larger than that of Algorithm 2 in some cases. On the other hand, among the considered approximation functions in Algorithm 2, Capped-ℓ_1 provides the best performances, followed by Exp and Log functions. Furthermore, our proposed optimization approach

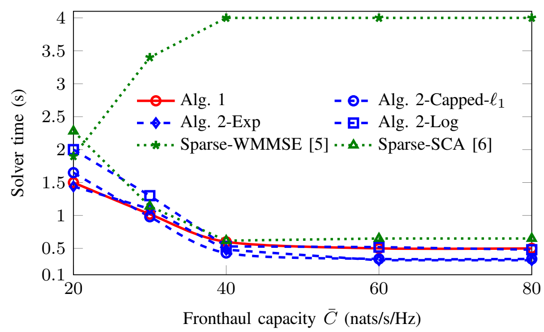

Fig. 2. Average solver run time of the considered schemes with $\bar{P} = 30$ dBm.

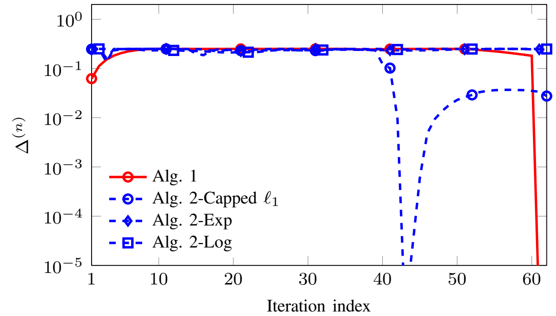

Fig. 3. Gap to the binary of relaxed variables obtained by the proposed algorithms for one channel realization with $C_b = 20$ nats/s/Hz, $\bar{P} = 30$ dBm.

based on ℓ_0-approximation has lower complexity compared to the one in [6]. These observations can be explained by the next experiment.

In Fig. 3, we show how close the obtained values of the relaxed variables are to 0 or 1. In particular, we plot the gap to the binary of relaxed variables obtained by the proposed algorithms, denoted by $\Delta^{(n)}$, for one channel realization where $\Delta^{(n)}$ is defined as

$$\Delta^{(n)} \triangleq \begin{cases} \max_{b,k}\{x_{b,k}^{(n)} - (x_{b,k}^{(n)})^2\} & \text{for Algorithm 1} \\ \max_{b,k}\{\tilde{x}_{b,k}^{(n)} - (\tilde{x}_{b,k}^{(n)})^2\} & \text{for Algorithm 2} \end{cases}$$

Here, a smaller $\Delta^{(n)}$ indicates a closer gap between $\{x_{b,k}^{(n)}\}_{b,k}$ (or $\{\tilde{x}_{b,k}^{(n)}\}_{b,k}$) and binary values. For Algorithm 1, it is seen that $\Delta^{(n)} \approx 0$ at convergence. This implies that the penalty method can achieve binary solutions. However, it takes a few more iterations to reach such values, which may result in higher solver time in some case, as seen in Fig. 2. On the other hand, the ℓ_0-approximation based method cannot derive exact binary solutions for relaxed variables in general. Among the considered approximation functions, Capped-ℓ_1 function can yield $\{x_{b,k}^{(n)}\}_{b,k}$ very close to 0 or 1 at convergence (the maximum gap is about 10^{-2}), and is superior to the other two functions. This indicates that Algorithm 2 using Capped-ℓ_1 function can return a feasible solutions after the simple mapping process (i.e., without carrying out the second SCA loop) with higher probability than the two other ones. In fact, we have observed when $\bar{C} = 20$ nats/s/Hz, Algorithm 2 with Capped-ℓ_1 successfully outputs feasible solutions for up to 77% of the number of channel realizations. The corresponding percentages for Algorithm 2 adopting Exp and Log functions are 61% and 10%, respectively. When increasing $\bar{C}$, i.e., $\bar{C} = 60$ nats/s/Hz, the percentage of the successful channel realizations when using Capped-ℓ_1, Exp and Log functions is 100%, 100% and 91%, respectively. This result also implies that Algorithm 2 often carries out one optimization stage, and thus it has lower computational complexity compared to the two-stage approach in [6].

V. Conclusion

This paper has studied the joint design of beamforming and RRH-user association to maximize sum rate in fronthaul-constrained C-RANs. Two new methods has been developed for the design problem, which can achieve very close to optimal performance with reasonable complexity. Numerical results have shown that the proposed schemes outperform the other known methods in both terms of sum rate performance, and computational complexity.

Acknowledgment

This work was supported by the Academy of Finland under projects WiConIE (Grant 297803), FURMESFuN (Grant 31089), and 6Genesis Flagship (Grant 318927).

References

[1] E. Hossain and M. Hasan, "5G cellular: key enabling technologies and research challenges," *IEEE Instrum. Meas. Mag.*, vol. 18, no. 3, pp. 11–21, June 2015.

[2] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud ran for mobile networks – A technology overview," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, Firstquarter 2015.

[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] Q. Vu, K. Nguyen, and M. Juntti, "Weighted max-min fairness for C-RAN multicasting under limited fronthaul constraints," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1534–1548, April 2018.

[5] 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.

[6] L. Liu and W. Yu, "Cross-layer design for downlink multihop cloud radio access networks with network coding," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1728–1740, April 2017.

[7] K. G. Nguyen, Q. D. Vu, M. Juntti, and L. N. Tran, "Globally optimal beamforming design for downlink CoMP transmission with limited backhaul capacity," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017.

[8] H. A. Le Thi, T. P. Dinh, H. M. Le, and X. T. Vo, "DC approximation approaches for sparse optimization," *European Journal of Operational Research*, vol. 244, no. 1, pp. 26–46, 2015.

[9] K.-G. Nguyen, Q.-D. Vu, M. Juntti, and L.-N. Tran, "Energy efficiency maximization for C-RANs: Discrete monotonic optimization, penalty, and ℓ_0-approximation methods," *IEEE Trans. Signal Process.*, vol. 66, no. 17, pp. 4435–4449, Sept 2018.

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

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

[12] H. A. Le Thi, T. P. Dinh *et al.*, "DC programming and DCA for general DC programs," in *Advanced Computational Methods for Knowledge Engineering*. Springer, 2014, pp. 15–35.