Backhaul-Aware Multicell Beamforming for Downlink Cloud Radio Access Network

Binbin Dai and Wei Yu

Department of Electrical and Computer Engineering

University of Toronto, Toronto, Ontario M5S 3G4, Canada

Emails: {bdai, weiyu}@comm.utoronto.ca

Abstract—This paper considers a heterogeneous downlink cloud radio access network (C-RAN) where all the base stations (BSs) in the network are connected to a central processor (CP) via capacity-limited backhaul links. Under this model, we investigate the message-sharing transmission strategy where the CP shares each user's message with a fixed subset of BSs, which then serve the user through joint beamforming. In this setting, although the overall long-term average backhaul consumption is limited by the fixed cluster size, the instantaneous backhaul consumption at each BS may vary significantly depending on the data rates of the scheduled users at each time slot. To avoid such large fluctuations in backhaul consumption, this paper proposes a backhaul-aware multicell scheduling and beamforming strategy that explicitly accounts for backhaul consumption. Specifically, a beamforming design algorithm is proposed to maximize the network utility for a downlink C-RAN under both per-BS power constraints and per-BS backhaul constraints in each time slot. Although this problem has already been considered in our previous work, this present paper proposes a new beamforming design algorithm that not only has guaranteed convergence but also achieves better system performance. This paper also shows that the performance of the proposed algorithm can be further improved by iterating with an additional power minimization step.

I. Introduction

Cloud radio access network (C-RAN) is an emerging network architecture that shows significant promise in supporting the ever increasing data traffic. In C-RAN, all the base stations (BSs) are connected to a centralized baseband unit (BBU) through digital backhaul links. The centralized BBU jointly processes the user messages and forwards them to the BSs, which in turn form the radio frequency signals to be transmitted to the users. The benefits of the C-RAN architecture are that it can better manage inter-user interference, better balance traffic load, enhance energy efficiency, and improve the overall network throughput.

The backhaul links in practical C-RAN implementations are of finite capacity. This limits the amount of information transfer from the central processor (CP) to the BSs. In this domain, many works in the literature have considered the implementation of C-RAN with limited backhaul [1]–[4].

This paper focuses on one specific transmission strategy in the downlink C-RAN known as *message-sharing*. In message-sharing strategy, the CP simply shares each user's original message with a selected subset of BSs, known as *BS cluster*, which then jointly beamform the signals to the intended users. The overall long-term average backhaul requirement for message-sharing strategy is determined by the cluster size: larger cluster size leads to higher backhaul consumption.

Previous works on downlink C-RAN have considered limiting the backhaul consumption either by limiting the cluster size [5]–[7] or adding a backhaul-related penalized term to the objective function [8], [9]. However, these previous approaches control only the long-term average backhaul; the instantaneous backhaul consumption may still have large variations since different users may be scheduled and served with different data rates at each time slot.

This paper differs from previous works in that we consider a backhaul-aware beamforming design problem which explicitly incorporates the instantaneous per-BS backhaul constraint in the network optimization framework. With explicit instantaneous backhaul constraints, the backhaul consumption at each time slot is accounted for at the scheduling and beamforming stage. As shown in [10], explicit consideration of per-BS backhaul constraints can also provide more freedom in choosing the cluster sizes and lead to full utilization of backhaul resources, resulting in significant improvement in network utility.

Explicit backhaul constraints have been considered in our previous work [10] under both dynamic clustering and fixed clustering. This paper focuses on fixed BS clustering only but provides an alternative and potentially better approach for instantaneous backhaul control through beamforming. Specifically, we propose a beamforming design algorithm which is guaranteed to converge to a stationary point solution to the utility maximization problem for downlink C-RAN under both per-BS power constraints and per-BS backhaul constraints. This is in contrast to the previous algorithm developed in [10], which has no theoretical convergence guarantee.

This paper also differs from [10] in user scheduling strategy. In [10], user scheduling is implicitly done within its beamforming design algorithm. However, implicit user scheduling cannot be directly applied to the new algorithm in this paper because the computational complexity would be too high. Instead, we propose to first select a subset of candidate users using the weighted minimum-mean-squared-error (WMMSE) approach [7], then select which subset of users to be scheduled among those candidate users using the proposed beamforming design algorithm. Our simulation results reveal that under such a two-stage user scheduling policy, the proposed algorithm provides considerable performance improvement as compared to the previous algorithm in [10] at comparable complexity.
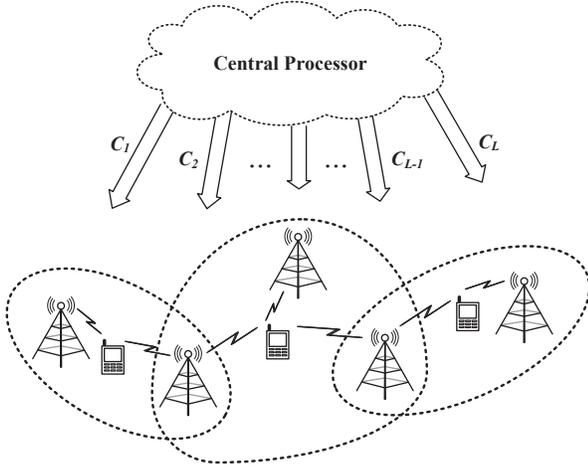
Fig. 1. Downlink C-RAN with limited per-BS backhaul capacity, in which each user is cooperatively served by a fixed subset of BSs.

We also show in this paper that the performances of both the previously proposed algorithm in [10] and the newly proposed algorithm in this paper can be further improved by iterating with an additional power minimization step. Further, the performance gain achieved by the power minimization step in the newly proposed algorithm is larger than that in the previously proposed algorithm .

II. SYSTEM MODEL AND PROBLEM FORMULATION

A downlink C-RAN architecture with a total number of L BSs and K users is considered. Each BS is equipped with M transmit antennas while each user has N receive antennas. As depicted in Fig. 1, each BS is connected to the CP via a backhaul link of capacity $C_l, l \in \mathcal{L} = \{1, 2, \cdots, L\}$.

This paper considers the message-sharing strategy between the CP and the BSs. Due to limited capacities of the backhaul links, the CP routes user k's message only to an individually selected and fixed subset of BSs, denoted as $\mathcal{L}_k = \{l_1, l_2, \cdots, l_{|\mathcal{L}_k|}\} \subseteq \mathcal{L}, k \in \mathcal{K} = \{1, 2, \cdots, K\}$, which then cooperatively serve user k through joint beamforming.

Let the beamforming vector formed by the serving cluster $\mathcal{L}_k$ for user k be denoted as

$$\mathbf{w}_k = \left[\mathbf{w}_{l_1 k}^H, \mathbf{w}_{l_2 k}^H, \cdots, \mathbf{w}_{l_{|\mathcal{L}_k|} k}^H\right]^H \quad (1)$$

where $\mathbf{w}_{l_m k} \in \mathbb{C}^{M \times 1}$ is the beamformer from BS l_m, $m \in \{1, 2, \cdots, |\mathcal{L}_k|\}$, to user k. The received signal at user k, $\mathbf{y}_k \in \mathbb{C}^{N \times 1}$, can be written as

$$\mathbf{y}_k = \mathbf{H}_{kk}\mathbf{w}_k s_k + \sum_{j \neq k, j \in \mathcal{K}} \mathbf{H}_{jk}\mathbf{w}_j s_j + \mathbf{n}_k, \quad (2)$$

where $\mathbf{H}_{jk} \in \mathbb{C}^{N \times |\mathcal{L}_j|M}$ is the matrix of channel gains from user j's serving cluster $\mathcal{L}_j$ to user k, $\mathbf{n}_k \in \mathbb{C}^{N \times 1}$ is the received noise at user k and is assumed to be distributed as $\mathcal{CN}(0, \sigma^2\mathbf{I})$, s_k is user k's message and is drawn independently from the signal constellation with zero mean and unit variance. With linear receive beamformer applied at the user

side, the signal-to-interference-and-noise ratio (SINR) for user k can be written as

$$\text{SINR}_k = \frac{\left\|\mathbf{u}_k^H \mathbf{H}_{kk}\mathbf{w}_k\right\|^2}{\sum_{j \neq k, j \in \mathcal{K}} \left\|\mathbf{u}_k^H \mathbf{H}_{jk}\mathbf{w}_j\right\|^2 + \sigma^2 \left\|\mathbf{u}_k\right\|_2^2} \quad (3)$$

where $\mathbf{u}_k \in \mathbb{C}^{N \times 1}$ is user k's receive beamformer.

It is assumed that the CP has access to global channel state information (CSI) for designing the beamforming vector $\mathbf{w}_k$. Once $\mathbf{w}_k$ is determined, the CP transmits user k's message, along with the beamforming coefficients, to user k's serving cluster $\mathcal{L}_k$, which then forms the beamformed signals to user k. We also assume that the channels are slow varying and only consider the backhaul consumption due to user message sharing and ignore the backhaul required for sharing CSI and delivering beamforming coefficients. Under this assumption, the backhaul consumption of the lth BS at current time slot is simply $\sum_{k \in \mathcal{K}_l} R_k$, where $\mathcal{K}_l \subseteq \mathcal{K}$ is the fixed subset of users associated with BS l and R_k is user k's data rate given by

$$R_k \leq$$

$$\log_2 \left(1 + \mathbf{w}_k^H \mathbf{H}_{kk}^H \left(\sum_{\substack{j \neq k \\ j \in \mathcal{K}}} \mathbf{H}_{jk}\mathbf{w}_j\mathbf{w}_j^H\mathbf{H}_{jk}^H + \sigma^2\mathbf{I}\right)^{-1} \mathbf{H}_{kk}\mathbf{w}_k\right)$$

$$\triangleq \log_2 \left(1 + \text{SINR}_{k,mmse}\right). \quad (4)$$

In (4), the $\text{SINR}_{k,mmse}$ is obtained by substituting the minimum-mean-squared-error (MMSE) receiver

$$\mathbf{u}_{k,mmse} = \left(\sum_{j \in \mathcal{K}} \mathbf{H}_{jk}\mathbf{w}_j\mathbf{w}_j^H\mathbf{H}_{jk}^H + \sigma^2\mathbf{I}\right)^{-1} \mathbf{H}_{kk}\mathbf{w}_k \quad (5)$$

into the SINR experession (3).

Considering weighted sum rate (WSR) as the utility function, we formulate the utility maximization problem for downlink C-RAN under per-BS power constraints P_l and instantaneous per-BS backhaul constraints C_l as

$$\underset{\{R_k, \mathbf{w}_k\}}{\text{maximize}} \quad \sum_{k \in \mathcal{K}} \alpha_k R_k \quad (6a)$$

$$\text{subject to} \quad \sum_{k \in \mathcal{K}_l} \|\mathbf{w}_{lk}\|_2^2 \leq P_l, \ \forall l \quad (6b)$$

$$\sum_{k \in \mathcal{K}_l} R_k \leq C_l, \ \forall l \quad (6c)$$

$$R_k \leq \log_2\left(1 + \text{SINR}_{k,mmse}\right), \forall k \quad (6d)$$

where α_k is the priority weight associated with user k. In (6b), $\|\cdot\|_2$ denotes the ℓ_2-norm of a vector. It is worth noting that the proposed algorithm in this paper can also be generalized to deal with the per-BS backhaul constraint averaged over multiple time slots. The objective function (6a) can also be replaced by other types of utility functions as discussed later in Remark 4.2.

In [10], a similar optimization problem as (6) is considered but with the inequality constraint (6d) replaced by equality. It can be shown that both problems are equivalent in achieving

the same optimal objective value. This is because for any optimal solution $\left\{\widetilde{\mathbf{w}}_k^*, \tilde{R}_k^*\right\}$ found in problem (6), there always exist another set of beamformers that can achieve $\left\{\tilde{R}_k^*\right\}$ with equality. This point is further elaborated in Section VI.

Problem (6) with either inequality or equality in constraint (6d) is a non-convex optimization problem, for which finding global optimum is quite challenging. In our previous work [10], (6d) is regarded as equality constraint and is substituted into the objective function. Then, the per-BS backhaul constraint (6c) is reformulated as a weighted sum power constraint so that the WMMSE approach [11] can be generalized to solve this problem. This algorithm is reviewed in Section III. In Section IV, we develop a new algorithm which solves problem (6) with rate inequality constraint in (6d). The newly proposed algorithm adopts an approach first used in [12] that applies the WMMSE technique in the constraint for solving a max-min rate optimization problem for multiple-input multiple-output (MIMO) interfering broadcast channel. The major difference between the existing algorithm in [10] and the newly proposed algorithm is that the latter has guaranteed convergence and also has better performance.

III. REVIEW OF EXISTING ALGORITHM IN [10]

This section reviews an existing algorithm for solving (6) as proposed in [10]. The essential idea in [10] is to approximate the per-BS backhaul constraint (6c) as a *weighted* sum power constraint so that the WMMSE approach in [11] can be generalized to solve this problem.

Specifically, it is noted in [10] that only the subset of users with nonzero beamforming vector need to be considered in the summations of the per-BS backhaul constraint (6c). Thus, constraint (6c) can be reformulated as

$$\sum_{k \in \mathcal{K}_l} R_k = \sum_{k \in \mathcal{K}_l} \left\| \|\mathbf{w}_k\|_2^2 \right\|_0 R_k \leq C_l, \ \forall l. \tag{7}$$

Further, [10] proposes to approximate the ℓ_0-norm in (7) by its weighted ℓ_1-norm:

$$\sum_{k \in \mathcal{K}_l} \beta_k \|\mathbf{w}_k\|_2^2 R_k \leq C_l, \ \forall l, \tag{8}$$

where β_k is the weight iteratively updated according to

$$\beta_k = \frac{1}{\|\mathbf{w}_k\|_2^2 + \tau}, \forall k, \tag{9}$$

with a small positive regularization factor $\tau > 0$.

Furthermore, [10] proposes to approximate the rate variable R_k in (8) using the achievable rate from previous iteration and reformulates problem (6) as

$$\begin{aligned} \underset{\{\mathbf{w}_k\}}{\text{maximize}} \quad & \sum_{k \in \mathcal{K}} \alpha_k \log_2 \left(1 + \text{SINR}_{k,mmse}\right) && \text{(10a)} \\ \text{subject to} \quad & \sum_{k \in \mathcal{K}_l} \|\mathbf{w}_{lk}\|_2^2 \leq P_l, \ \forall l && \text{(10b)} \\ & \sum_{k \in \mathcal{K}_l} \beta_k \hat{R}_k \|\mathbf{w}_k\|_2^2 \leq C_l, \ \forall l && \text{(10c)} \end{aligned}$$

Algorithm 1 WMMSE with Reweighted ℓ_1-norm [10]

Initialization: $\left\{\beta_k, \hat{R}_k\right\}$;
Repeat:
  1) Fix $\left\{\beta_k, \hat{R}_k\right\}$, use the generalized WMMSE approach to solve problem (10);
  2) At the end of each WMMSE routine, compute the achievable rate R_k according to (4), and update $\hat{R}_k = R_k$ and β_k according to (9), $\forall k$.
Until convergence

---

where $\hat{R}_k$ in constraint (10c) is the achievable rate from previous iteration. The approximated per-BS backhaul constraint (10c) now can be regarded as a weighted sum power constraint with weights $\beta_k \hat{R}_k$.

Although the reformulated problem (10) is still non-convex, the conventional WMMSE approach [11] can be generalized to solve problem. The complete algorithm is summarized in Algorithm 1. One issue with Algorithm 1 is that it has no theoretical convergence guarantee due to the heuristic update of $\hat{R}_k$ and β_k. In the subsequent section, we propose an alternative algorithm which has provable convergence guarantee.

IV. PROPOSED BEAMFORMING ALGORITHM

The existing algorithm in [10] focuses on transforming the per-BS backhaul constraint (6c) into a format so that the WMMSE approach can be applied. In contrast, this paper focuses on dealing with the non-convex constraint (6d) and tries to transform (6d) into a convex format so that the standard convex optimization method can be used to solve the problem.

The key transformation is to rewrite $\log_2 \left(1 + \text{SINR}_{k,mmse}\right)$ in (6d) in the following equivalent form:

$$\log_2 \left(1 + \text{SINR}_{k,mmse}\right) = \max_{\{\mathbf{u}_k, \rho_k\}} \left(\frac{\ln \rho_k - \rho_k e_k + 1}{\ln 2} \right) \tag{11}$$

where $e_k \in \mathbb{R}$ is the mean-squared-error (MSE) for user k defined as

$$\begin{aligned} e_k = \ & \mathsf{E}\left[\left\|\mathbf{u}_k^H \mathbf{y}_k - s_k\right\|_2^2\right] \\ = \ & \mathbf{u}_k^H \left(\sum_{j \in \mathcal{K}} \mathbf{H}_{jk} \mathbf{w}_j \mathbf{w}_j^H \mathbf{H}_{jk}^H + \sigma^2 \mathbf{I}\right) \mathbf{u}_k \\ & - 2\text{Re}\left\{\mathbf{u}_k^H \mathbf{H}_{kk} \mathbf{w}_k\right\} + 1 \end{aligned} \tag{12}$$

and $\rho_k \in \mathbb{R}$ is a scalar variable associated with user k.

The proof of the relationship in (11) is based on the first order optimality condition. By taking partial derivatives of the function in the right-hand side of (11) with respect to $\mathbf{u}_k$ and ρ_k respectively and setting them to zeros, it is easy to see that the unconstrained optimization problem in (11) achieves its optimum when $\mathbf{u}_k = \mathbf{u}_{k,mmse}$, as defined in (5), and $\rho_k = 1/e_k^*$, where e_k^* is the MMSE under $\mathbf{u}_{k,mmse}$. Substituting $\mathbf{u}_k = \mathbf{u}_{k,mmse}$ and $\rho_k = 1/e_k^*$ into the right-hand side of (11) produces the left-hand side.

With the above equivalent rate expression, we can now reformulate the rate inequality constraint (6d) as

$$R_k \le \max_{\{\mathbf{u}_k, \rho_k\}} \left(\ln \rho_k - \rho_k e_k + 1 \right) / \ln 2, \qquad (13)$$

where e_k is a function of $\mathbf{u}_k$ as defined in (12).

Note that problem (6) is non-convex due to the rate inequality constraint (6d). However, for fixed $\{\mathbf{u}_k, \rho_k\}$, the above equivalent rate inequality constraint (13) is a convex quadratic constraint in $\{\mathbf{w}_k, R_k\}$ jointly. This observation motivates us to utilize the block coordinate descent (BCD) method to solve problem (6) by iterating between the optimization of $\{\mathbf{u}_k, \rho_k\}$ and $\{\mathbf{w}_k, R_k\}$. We summarize the proposed algorithm in Algorithm 2. The most computationally intensive step in Algorithm 2 is Step 3, which finds the optimal $\{\mathbf{w}_k, R_k\}$ jointly by solving the following convex optimization problem under fixed $\{\mathbf{u}_k, \rho_k\}$:

$$\underset{\{R_k, \mathbf{w}_k\}}{\text{maximize}} \quad \sum_{k \in \mathcal{K}} \alpha_k R_k \qquad (14a)$$

$$\text{subject to} \quad \sum_{k \in \mathcal{K}_l} \|\mathbf{w}_{lk}\|_2^2 \le P_l, \ \forall l \qquad (14b)$$

$$\sum_{k \in \mathcal{K}_l} R_k \le C_l, \ \forall l \qquad (14c)$$

$$\sum_{j \in \mathcal{K}} \mathbf{w}_j^H \mathbf{H}_{jk}^H \mathbf{u}_k \mathbf{u}_k^H \mathbf{H}_{jk} \mathbf{w}_j + \frac{\ln 2}{\rho_k} R_k$$
$$- 2\text{Re}\left\{ \mathbf{u}_k^H \mathbf{H}_{kk} \mathbf{w}_k \right\} \le t_k, \ \forall k \quad (14d)$$

where t_k is a constant for given $\{\mathbf{u}_k, \rho_k\}$ and is defined as $t_k = \frac{\ln \rho_k + 1}{\rho_k} - \sigma^2 \|\mathbf{u}_k\|_2^2 - 1$.

The usual convergence proof for BCD requires that the constraints are separable in the block variables, which is not the case for problem (6) as the rate inequality constraint (13) is coupled among the optimization variables. However, we can show that the proposed Algorithm 2 converges to a Karush-Kuhn-Tucker (KKT) point of (6).

Proposition 4.1: Starting with any initial point, the sequence $\{\mathbf{x}^n\}_{n=1}^{\infty} \triangleq \{\mathbf{w}_k^n, R_k^n, \mathbf{u}_k, \rho_k\}_{n=1}^{\infty}$ generated by Algorithm 2 converges to the KKT solution of (6) in the sense that $\lim_{n \to \infty} d(\mathbf{x}^n, \mathcal{S}) = 0$, where $\mathcal{S}$ is the set of KKT points of (6) and $d(\mathbf{x}^n, \mathcal{S}) \triangleq \inf_{\mathbf{s} \in \mathcal{S}} \|\mathbf{x}^n - \mathbf{s}\|_2$.

Proof Sketch: Since the objective function in problem (6) monotonically decreases under Algorithm 2 and that the constrained set of (6) is compact, we can use the technique of [12] to show that any cluster point of $\{\mathbf{x}\}_{n=1}^{\infty}$ satisfies the KKT conditions of (6). Although the KKT points may not be unique, the distance between $\{\mathbf{x}\}_{n=1}^{\infty}$ and the KKT solution set of (6) nevertheless goes to zero in the limit. ∎

Remark 4.2: Algorithm 2 only relies on the fact that for given $\{\mathbf{u}_k, \rho_k\}$, the optimization problem (14) for jointly finding $\{\mathbf{w}_k\}$ and $\{R_k\}$ is a convex optimization problem. Since all the constraints in problem (14) are convex, this implies that Algorithm 2 can be readily generalized to any utility function that is a concave function of $\{R_k\}$. This includes the WSR utility $\sum_k \alpha_k R_k$, proportional fairness utility $\sum_k \log R_k$, weighted harmonic mean rate utility $-\sum_k \alpha_k R_k^{-1}$, etc. In

---

Algorithm 2 Block Coordinate Descent (BCD) Method

Initialization: $\{\mathbf{w}_k\}$;

Repeat:
1) Fix $\{\mathbf{w}_k\}$, compute the MMSE receiver $\{\mathbf{u}_k\}$ and the corresponding MSE $\{e_k\}$ according to (5) and (12);
2) Update the MSE weight $\{\rho_k\}$ as $\rho_k = 1/e_k$;
3) Fix $\{\mathbf{u}_k, \rho_k\}$, find the optimal transmit beamformer $\{\mathbf{w}_k\}$ and rate $\{R_k\}$ jointly through solving the convex optimization problem (14) using the standard convex optimization tool [13].

Until convergence

---

addition, Algorithm 2 can also be used to solve the minimum rate fairness utility $(\min_k R_k)$ maximization problem, which is not differentiable and cannot be handled by Algorithm 1 as the WMMSE equivalence does not hold in this case.

The computational complexities for both Algorithm 1 and 2 are dominated by the transmit beamformer design step. In our simulations, the optimization problems for designing the transmit beamformer in both Algorithm 1 and 2 are reformulated as second order cone programming (SOCP) and are solved by the interior-point method. The complexities for Algorithm 1 and 2 are thus given as $O(K^3 L^4 M^3)$ and $O(K^4 L^3 M^3)$, respectively. The difference in these two complexity orders is due to the different total number of constraints in (10) and (14). Typically, the number of users K is larger than the number of BSs L. Hence, Algorithm 2 is at most $O(K)$ more complex than Algorithm 1. In the next section, we propose an efficient user scheduling strategy that can significantly reduce the total number of users needed to be considered in Algorithm 2.

V. USER SCHEDULING

In Algorithm 1, the scheduled users at each time slot can be implicitly determined by considering all the K users in the optimization problem but only choosing those users with nonzero rates. It has been reported in [10] that such implicit user scheduling brings considerable utility improvement compared to round-robin user scheduling. Due to the per-user rate constraint (6d), implicit user scheduling adds significant complexity to Algorithm 2.

To reduce the computational complexity, we propose a two-stage user scheduling procedure for Algorithm 2. In the first stage, the WMMSE approach [7] is used to solve the WSR maximization problem (6) with only power constraints (6b) for a fixed number of iterations, which is set to 10 iterations in our simulations. By the end of the first stage, a large number of users will have zero rate and can be removed from the user scheduling pool. In the second stage, only those surviving users from the first stage with nonzero rates are considered in Algorithm 2, which makes the final scheduling decision. Since the computational complexity of Algorithm 2 is dominated by the number of considered users, such two-stage user scheduling policy significantly improves the efficiency of Algorithm 2. Our simulation results show that under such two-stage user scheduling, Algorithm 2 outperforms Algorithm 1
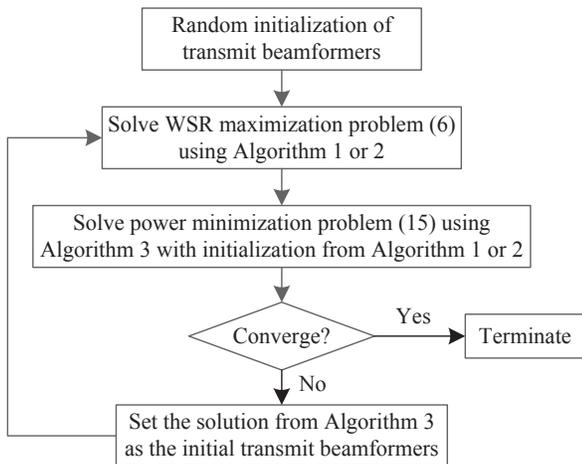
Fig. 2. Block diagram of the optimization framework which further improves the performances of Algorithm 1 and 2.

under implicit user scheduling at comparable complexity for a realistic network.

VI. POWER MINIMIZATION

In this section, we show that a power minimization step can further improve the performance of the proposed algorithm. Let $\left\{\widetilde{\mathbf{w}}_k^*, \tilde{R}_k^*\right\}$ denote the solution to problem (6) obtained from Algorithm 2. Due to the rate inequality constraint (6d), the user rates $\log_2\left(1 + \mathrm{SINR}_{k,mmse}\right)$ achieved by $\{\widetilde{\mathbf{w}}_k^*\}$ may be strictly larger than $\tilde{R}_k^*$. We can further improve the beamforming strategy $\{\widetilde{\mathbf{w}}_k^*\}$ by solving the following power minimization problem to reduce total power consumption

$$\underset{\{\mathbf{w}_k,\mathbf{u}_k\}}{\text{minimize}} \quad \sum_{k \in \mathcal{K}} \|\mathbf{w}_k\|_2^2 \qquad (15a)$$

$$\text{subject to} \quad \sum_{k \in \mathcal{K}_l} \|\mathbf{w}_{lk}\|_2^2 \leq P_l, \ \forall l \qquad (15b)$$

$$\mathrm{SINR}_k \geq 2^{\tilde{R}_k^*} - 1, \ \forall k \qquad (15c)$$

where SINR_k is defined in (3).

For the above total power minimization problem, it is well known that at the optimal point the SINR constraints (15c) are achieved with equality. The key advantage here is that by reducing transmission power, less interference is generated; the performance of the overall system improves. In fact, this motivates us to propose a new optimization framework that iteratively solves problem (6) and (15) to further improve the performance of Algorithm 1 and 2. Such framework is shown in Fig. 2 and is guaranteed to converge as each iteration generates a non-decreasing network utility. As shown in the simulation results, the optimization framework in Fig. 2 can produce a considerably better solution to problem (6) than the solution produced by Algorithm 1 or Algorithm 2 alone.

However, note that the power minimization problem (15) is non-convex when there are multiple receive antennas at the user side. Global optimum can only be achieved if the receive beamformer $\{\mathbf{u}_k\}$ is fixed. With fixed receive beamformer,

Cellular Layout	Hexagonal 7-cell wrapped-around
Channel bandwidth	10 MHz
Distance between cells	0.8 km
Num. of (macro-BSs, pico-BSs, users)/cell	$(1, 3, 30)$
Num. of antennas/(macro-BS, pico-BS, user)	$(4, 2, 2)$
Max. Tx power for (macro-BS, pico-BS)	$(43, 30)$ dBm
Antenna gain	15 dBi
Background noise	-169 dBm/Hz
Path loss from macro-BS to user	$128.1 + 37.6 \log_{10}(d)$
Path loss from pico-BS to user	$140.7 + 36.7 \log_{10}(d)$
Log-normal shadowing	8 dB
Rayleigh small scale fading	0 dB
Reweighting function parameter	$\tau = 10^{-12}$

problem (15) can be transformed into a SOCP problem [14]. Also, with fixed transmit beamformer, the optimal receive beamformer is the MMSE receiver. Based on these facts, we propose to seek a locally optimal solution to (15) by iterating between the MMSE receiver design and optimal transmitter design under fixed receiver, as illustrated in Algorithm 3.

Algorithm 3 Power Minimization

Initialization: Set the solution to problem (6) as the initial value for $\{\mathbf{w}_k\}$.
Repeat:
 1) Fix $\{\mathbf{w}_k\}$, update $\{\mathbf{u}_k\}$ as the MMSE receiver defined in (5);
 2) Fix receive beamformer $\{\mathbf{u}_k\}$, find the optimal transmit beamformer $\{\mathbf{w}_k\}$ for problem (15) by solving a SOCP problem [14].
Until convergence

VII. SIMULATION RESULTS

This section evaluates the effectiveness of the proposed algorithms through numerical simulations for a heterogeneous C-RAN environment with the parameters listed in Table I. The BS clustering strategy adopts the biased-signal-strength-based clustering scheme proposed in [10]. For simplicity, we set all the macro-BSs to have equal backhaul constraints and likewise for the pico-BSs. The backhaul constraints are denoted as $(C_{\mathrm{macro}}, C_{\mathrm{pico}}) = (690, 107)$ Mbps respectively, which are extracted from the baseline scheme considered in [10].

In Fig. 3, the convergence behavior of Algorithm 2 is plotted under round-robin user scheduling. As we can see, the sum rate generated by Algorithm 2 monotonically increases as the iteration goes on and eventually converges. The convergence behavior of the optimization framework proposed in Fig. 2 is also included in Fig. 3, in which the power minimization step is added to Algorithm 2 at the 15th iteration. This is when the objective function only improves by 1% per iteration with Algorithm 2 by itself. It can be seen that with power minimization step, the sum rate obtained is strictly higher than that obtained from Algorithm 2 alone and eventually converges
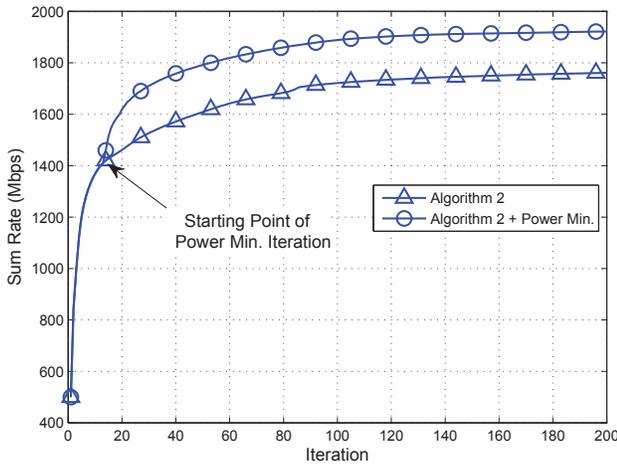
Fig. 3. Convergence behavior of the proposed algorithms under round-robin user scheduling and $\alpha_k = 1, \forall k$.

Fig. 4. CDF of user data rates comparison with α_k updated according to proportional fairness criterion.

to a higher optimal value. This illustrates the effectiveness of Algorithm 3 in better refining the solution obtained from Algorithm 2.

In Fig. 4, we compare the cumulative distribution functions (CDF) of user data rates between Algorithm 1 and 2 under different scenarios. As we can see, with round-robin scheduling, Algorithm 1 and 2 exhibit similar performance. However, we observe that the power minimization step improves the performance of Algorithm 2 more than Algorithm 1. This is mainly because, when Algorithm 1 converges, R_k is typically equal to $\log_2(1 + \text{SINR}_{k,mmse})$, but when Algorithm 2 converges, $\log_2(1 + \text{SINR}_{k,mmse})$ may be strictly larger than R_k. Hence, the transmit beamformers obtained from Algorithm 2 can benefit more from the power minimization step in Algorithm 3. In Fig. 4, we also plot the performance of the optimization framework in Fig. 2 under the user scheduling strategy discussed in Section V. With a total number of 210 users in the network, we note that under the optimization framework in Fig. 2, Algorithm 2 with the proposed two-stage user scheduling has comparable complexity as Algorithm 1 with implicit scheduling. However, the new algorithm significantly improves the overall system performance particularly for the low-to-medium rate users.

VIII. CONCLUSION

This paper proposes a new algorithm to solve the network utility maximization problem for downlink C-RAN with fixed BS clustering under both per-BS power constraints and per-BS backhaul constraints in each time slot. Unlike the previous algorithm, the new algorithm is guaranteed to converge to a stationary point of the network utility maximization problem. We also show that by iterating with an additional power minimization step, the performance of the proposed algorithm can be further improved. Simulation results reveal that the proposed algorithm with power minimization iteration improves over the existing algorithm with comparable complexity.
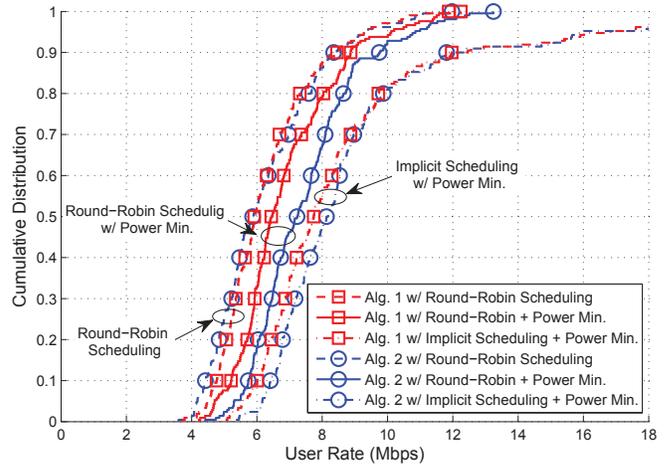
REFERENCES

[1] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai, "Joint precoding and multivariate backhaul compression for the downlink of cloud radio access networks," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5646–5658, Nov. 2013.

[2] ——, "Performance evaluation of multiterminal backhaul compression for cloud radio access networks," in *Proc. IEEE Conf. on Inform. Sci. and Syst. (CISS)*, March 2014, pp. 1–6.

[3] P. Marsch and G. Fettweis, "On base station cooperation schemes for downlink network MIMO under a constrained backhaul," in *Proc. IEEE Global Commun. Conf. (Globecom)*, 2008, pp. 1–6.

[4] R. Zakhour and D. Gesbert, "Optimized data sharing in multicell MIMO with finite backhaul capacity," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6102–6111, Dec. 2011.

[5] J. Zhang, R. Chen, J. Andrews, A. Ghosh, and R. Heath, "Networked MIMO with clustered linear precoding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1910–1921, 2009.

[6] C. Ng and H. Huang, "Linear precoding in cooperative MIMO cellular networks with limited coordination clusters," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1446–1454, 2010.

[7] S. Kaviani, O. Simeone, W. Krzymien, and S. Shamai, "Linear precoding and equalization for network MIMO with partial cooperation," *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, pp. 2083–2096, 2012.

[8] M. Hong, R.-Y. Sun, H. Baligh, and Z.-Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogenous networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 226–240, Feb. 2013.

[9] A. Chowdhery, W. Yu, and J. Cioffi, "Cooperative wireless multicell OFDMA network with backhaul capacity constraints," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2011, pp. 1–6.

[10] B. Dai and W. Yu, "Sparse beamforming and user-centric clustering for downlink cloud radio access network," *IEEE Access, Special Issue on Recent Advances in Cloud Radio Access Networks*, vol. 2, pp. 1326–1339, 2014.

[11] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331 –4340, Sep. 2011.

[12] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "Linear transceiver design for a MIMO interfering broadcast channel achieving max-min fairness," *Signal Process.*, vol. 93, no. 12, pp. 3327–3340, Dec. 2013.

[13] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

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