Sparse Beamforming and User-Centric Clustering for Downlink Cloud Radio Access Network

BINBIN DAI, (Student Member, IEEE), AND WEI YU, (Fellow, IEEE)
The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada

Corresponding author: W. Yu (weiyu@comm.utoronto.ca)

ABSTRACT This paper considers a downlink cloud radio access network (C-RAN) in which all the base-stations (BSs) are connected to a central computing cloud via digital backhaul links with finite capacities. Each user is associated with a user-centric cluster of BSs; the central processor shares the user's data with the BSs in the cluster, which then cooperatively serve the user through joint beamforming. Under this setup, this paper investigates the user scheduling, BS clustering, and beamforming design problem from a network utility maximization perspective. Differing from previous works, this paper explicitly considers the per-BS backhaul capacity constraints. We formulate the network utility maximization problem for the downlink C-RAN under two different models depending on whether the BS clustering for each user is dynamic or static over different user scheduling time slots. In the former case, the user-centric BS cluster is dynamically optimized for each scheduled user along with the beamforming vector in each time-frequency slot, whereas in the latter case, the user-centric BS cluster is fixed for each user and we jointly optimize the user scheduling and the beamforming vector to account for the backhaul constraints. In both cases, the nonconvex per-BS backhaul constraints are approximated using the reweighted ℓ_1-norm technique. This approximation allows us to reformulate the per-BS backhaul constraints into weighted per-BS power constraints and solve the weighted sum rate maximization problem through a generalized weighted minimum mean square error approach. This paper shows that the proposed dynamic clustering algorithm can achieve significant performance gain over existing naive clustering schemes. This paper also proposes two heuristic static clustering schemes that can already achieve a substantial portion of the gain.

INDEX TERMS Cloud radio access network (C-RAN), network multiple-input multiple-output (MIMO), coordinated multi-point (CoMP), limited backhaul, user scheduling, base-station clustering, beamforming, weighted sum rate maximization, weighted minimum mean square error (WMMSE).

I. INTRODUCTION

The fifth-generation (5G) wireless system is expected to support an ever increasing number of mobile devices with ubiquitous service access. To realize this 5G vision, ultra-dense small cell deployments and cloud computing are recognized as the two key enabling technologies [2]. With small cells, the received signal strength is enhanced at the user's side due to the reduced distance to the serving base-stations (BSs). However, as the neighboring BSs are also located closer in distance, the users are exposed to more inter-cell interference, which limits the performance of the cellular network. Cloud radio access network (C-RAN) is an emerging network architecture that is capable of dealing with this inter-cell interference issue. In C-RAN, the BSs are connected to a central processor (CP) via digital backhaul links. This allows the CP to jointly encode the user messages using linear

precoding or beamforming techniques for interference mitigation purpose in the downlink. The C-RAN architecture can be thought of as a platform for the practical implementation of network multiple-input multiple-output (MIMO) and coordinated multi-point (CoMP) transmission concepts [3].

This paper studies the optimization of the C-RAN architecture focusing on the effect of finite-capacity backhaul links on the overall network capacity. In this realm, several practical transmission strategies have been proposed for the downlink C-RAN to account for the finite backhaul. In one such strategy, the CP performs the beamforming operation, then compresses and forwards the beamformed signals to the BSs. Compression is needed because of the capacity limits of the backhaul links. This strategy is investigated in [4] and [5] and is referred to as the *compression* strategy in this paper. In an alternative strategy, each user is associated with a cluster

of multiple BSs and the CP simply shares each user's message directly with its serving BS cluster. The BSs form the beamformed signals locally, then cooperatively transmit the signals to the users. This strategy is studied in [4], [6], and [7] and is referred to as the *data sharing* strategy in this paper. In the compression strategy, the amount of available backhaul capacity determines the resolution of the compressed signals: higher-resolution compression requires larger backhaul capacity. In the data sharing strategy, the amount of required backhaul capacity is related to the BS cluster size: larger cluster size leads to higher backhaul consumption.

This paper focuses on the data sharing strategy. The performance of the data sharing strategy depends crucially on the choice of BS cooperation cluster for each user. Broadly speaking, there are two types of BS clustering schemes for data sharing: *disjoint clustering* and *user-centric clustering*. In disjoint clustering scheme, the entire network is divided into non-overlapping clusters and the BSs in each cluster jointly serve all the users within the coverage area [8]. Although disjoint clustering scheme has already been shown to be effective in mitigating the inter-cell interference [9], [10], users at the cluster edge still suffer from considerable inter-cluster interference. Differently, in user-centric clustering, each user is served by an individually selected subset of neighboring BSs and different clusters for different users may overlap. The benefit of user-centric clustering is that there exists no explicit cluster edge. This paper adopts the user-centric clustering scheme and further considers two different implementations of user-centric clustering depending on whether BS clustering is *dynamic* or *static* over the different user scheduling time slots. In dynamic clustering, the BS cluster for each user can change over time, allowing for more freedom to fully utilize the backhaul resources. However, dynamic clustering scheme also requires more signaling overhead as new BS-user associations need to be established continuously. In static clustering, the BS-user association is fixed over time and may only need to be updated as the user location changes.

This paper considers both dynamic and fixed clustering schemes, and proposes joint clustering, user scheduling and beamforming designs for the downlink C-RAN with user-centric data sharing strategy. We explicitly take per-BS backhaul capacity constraints into account in the network utility maximization framework, and use the ℓ_1-norm reweighting technique in compressive sensing and a generalized weighted minimum mean square error (WMMSE) [11], [12] approach to solve the problem. We show that dynamic clustering can significantly outperform the naive channel strength based clustering strategy, while the proposed heuristic static clustering schemes can already achieve a substantial portion of the performance gain.

A. RELATED WORK

The information-theoretical capacity of the C-RAN model has been considered extensively in the literature. However, most of the theoretical analysis on C-RAN is restricted to simplified channel models [6], [7], [13], [14]. Specifically, the achievable rate regions derived in [6] and [7] are based on a two-BS-two-user channel model, while [13] and [14] consider Wyner-like channel models and report the achievable rates and capacity bounds, respectively. In [15] and [16], large-system analysis of network MIMO system is carried out. Although based on simplified models, these previous information-theoretical results already reveal the benefits of C-RAN in significantly improving the system performance.

This paper focuses on practical system design for the downlink C-RAN. This design problem has been considered in the literature under various performance metrics. For instance, under the signal-to-interference-and-noise ratio (SINR) constraints at the receivers, [17] considers backhaul minimization while [18] and [19] consider network power minimization as the objective. Furthermore, the optimal tradeoff between the backhaul capacity and transmit power is investigated in [20]–[22], while several other performance measures like mean square error (MSE) and energy efficiency (bits/Joule delivered to the users) are considered in [23]–[26], respectively.

In this paper, we consider the network utility as the performance measure for the downlink C-RAN. Differing from the utility maximization problems in conventional wireless networks with only transmit power constraints, the additional backhaul constraints in C-RAN make the problem more challenging as the backhaul consumption at a particular BS is a function of not only the (continuous) user rates but also the (discrete) number of associated users. To tackle this mixed continuous and discrete optimization problem, existing literature mostly take the limited backhaul capacities into account *implicitly* either by fixing the BS clusters [27]–[29] or by adding the backhaul as a penalized term into the objective function [30]–[32]. Specifically, [27] considers sum rate maximization under fixed and disjoint clustering scheme while [28] and [29] maximize a more general utility function under user-centric but predetermined BS clusters. Dynamic user-centric clustering design is considered in [30] by penalizing the objective function with an ℓ_2-norm approximation of the cluster size. Alternatively, [31] and [32] choose the backhaul rate as the penalized term but solve the problem heuristically. For fixed clustering scheme assumed in [27]–[29], the backhaul consumption is only known afterwards by evaluating the rates of user messages delivered in each backhaul link. For dynamic clustering designs considered in [30]–[32], one has to optimally choose the price associated with each penalized term to ensure that the overall backhaul stays within the budget, which is not easy.

In contrast to all the above existing works in network utility maximization for the downlink C-RAN, this paper *explicitly* formulates the per-BS backhaul constraints in the optimization framework. With explicit per-BS backhaul constraints, we show that the backhaul resources can be more efficiently utilized and that the network utility can be significantly improved.

The BS clustering problem for C-RAN with limited backhaul capacity is combinatorial in nature, for which finding the global optimum is expected to be quite challenging. Several suboptimal cluster formation algorithms have already been proposed in the literature. In [33], the BS cluster is assumed to be selected from a set of predetermined candidate clusters, and a greedy cluster selection algorithm is proposed to maximize the network utility. Alternatively, [34] models the cluster formation problem using graph theory, while [35] treats the problem from queuing theory perspective.

This paper differs from previous work in that we propose a dynamic clustering scheme by optimizing a *sparse* beamforming vector for each user, where the nonzero entries in the beamforming vector correspond to the user's serving cluster. This allows the formulations of the BS clustering problems as an ℓ_0-norm optimization problem and its subsequent solution via an application of the ℓ_1-norm reweighting technique in compressive sensing.

This paper proposes a novel application of WMMSE approach to jointly optimize the user scheduling and beamforming vectors under either dynamic or fixed BS clustering. This is in contrast to [28], which uses the first-order Taylor expansion to approximate the the nonconvex rate expression. Although the WMMSE approach has been applied to the C-RAN setup in the past [29], [30], these previous works do not explicitly take the backhaul constraints into consideration. As related work, the WMMSE approach has also been adapted to solve the max-min fairness problem for MIMO interfering broadcast channel [36], a link flow rate control problem for the radio access network [37] and a power minimization problem under time-averaged user rate constraints for the CoMP architecture [38]. Recently, the WMMSE technique is generalized in [39] to a wider class of network setups using the successive convex approximation idea. Finally, we mention that the WMMSE is a numerical approach for solving for a stationary point to the weighted sum rate (WSR) maximization problem over the beamformers, which is known to be a challenging problem. Recent progress for finding globally optimal solution to the WSR maximization problem under various conditions has been reported in [40]–[42].

B. MAIN CONTRIBUTIONS

This paper considers the user scheduling, user-centric BS clustering and beamforming design problem for the downlink C-RAN. The main contributions in this paper are summarized as follows:

1) Per-BS backhaul constraints are explicitly considered in the network utility maximization problem for the downlink C-RAN under data sharing strategy. A key novel technique proposed in this paper is that the per-BS backhaul constraint can be formulated in a weighted ℓ_0-norm format and approximated using the reweighted ℓ_1-norm.

2) A novel application of the WMMSE approach is proposed to solve the utility maximization problem with backhaul constraints. The proposed algorithm can be applied to the cases where the BS clustering for each user can be either dynamic or static.

3) We show numerically that with explicit per-BS backhaul constraints, the proposed algorithm is able to utilize the backhaul resources more efficiently, as well as to offer more flexibilities in choosing the cluster size. Simulation results also show that as compared with the naive clustering schemes, both the dynamic and the static clustering schemes proposed in this paper achieve significant performance improvement.

C. PAPER ORGANIZATION AND NOTATIONS

The rest of the paper is organized as follows. Section II introduces the system model. Section III considers dynamic BS clustering and proposes a joint scheduling, beamforming and clustering design algorithm together with two additional techniques to further reduce the computational complexity of the proposed algorithm. In Section IV, the user scheduling and beamforming vectors are jointly optimized under fixed BS clustering and two heuristic static clustering algorithms are proposed. Numerical results are provided in Section V. Conclusions are drawn in Section VI.

Throughout this paper, lower-case bold letters (e.g. $\mathbf{w}$) denote vectors and upper-case bold letters (e.g. $\mathbf{H}$) denote matrices. We use $\mathbb{R}$ and $\mathbb{C}$ to denote real and complex domain, respectively. The matrix inverse, conjugate transpose and ℓ_p-norm of a vector are denoted as $(\cdot)^{-1}$, $(\cdot)^H$ and $|\cdot|_p$ respectively. The complex Gaussian distribution is represented by $\mathcal{CN}(\cdot, \cdot)$ while $\text{Re}\{\cdot\}$ stands for the real part of a scalar. The expectation of a random variable is denoted as $\mathsf{E}[\cdot]$. Calligraphy letters are used to denote sets while $|\cdot|$ stands for either the size of a set or the absolute value of a real scalar, depending on the context.

II. SYSTEM MODEL

Consider a downlink C-RAN with L BSs and K users, where each BS has M transmit antennas while each user has N receive antennas. Each BS l is connected to a CP with a backhaul link with capacity limit $C_l, l \in \mathcal{L} = \{1, 2, \ldots, L\}$, as depicted in Fig. 1. We assume that the CP has access to all users' data and distributes each user's data to an individually selected cluster of BSs via the backhaul links. Each user is then cooperatively served by its serving cluster through joint beamforming.

In order to represent the BS cluster and transmit beamformer in a compact form, we introduce a network-wide beamforming vector $\mathbf{w}_k = [\mathbf{w}_k^1, \mathbf{w}_k^2, \ldots, \mathbf{w}_k^L] \in \mathbb{C}^{M_t \times 1}$ for each user $k \in \mathcal{K} = \{1, 2, \ldots, K\}$, where $M_t = LM$ and $\mathbf{w}_k^l \in \mathbb{C}^{M \times 1}$ is the transmit beamformer from BS l to user k. Suppose that BS l is not part of user k's serving cluster, then the corresponding beamformer block $\mathbf{w}_k^l$ is set to $\mathbf{0}$. Since each user is expected to be served by only a small number of BSs, the network-wide beamforming vector $\mathbf{w}_k$ is group sparse.
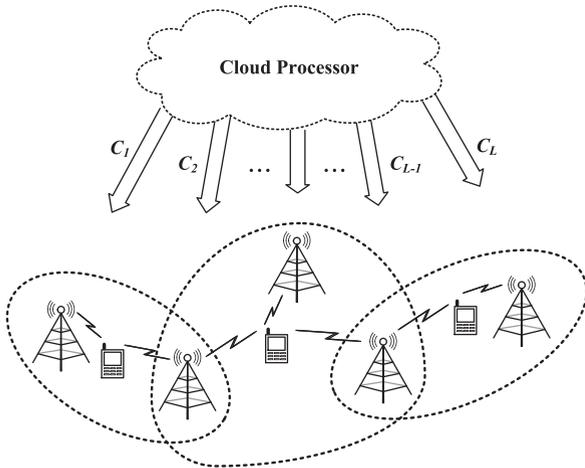
FIGURE 1. Downlink C-RAN with per-BS backhaul capacity limits, where each user is cooperatively served by a user-centric and potentially overlapping subset of BSs.

Here, we assume that all the L BSs can potentially serve each scheduled user in order to simplify the notations. However, the proposed algorithms in this paper can be readily applied to the situation where only a subset of BSs are considered as each user's candidate serving BSs.[1]

With linear transmit beamforming scheme at the BSs, the received signal at user k, denoted as $\mathbf{y}_k \in \mathbb{C}^{N \times 1}$, can be written as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{w}_k s_k + \sum_{j \neq k, j \in \mathcal{K}} \mathbf{H}_k \mathbf{w}_j s_j + \mathbf{n}_k, \qquad (1)$$

where $\mathbf{H}_k \in \mathbb{C}^{N \times M_t}$ denotes the channel state information (CSI) matrix from all the M_t transmit antennas 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 $\mathcal{CN}(0, \sigma^2 \mathbf{I})$. In this paper, we consider the case where each user has only a single data stream for simplicity and assume that user k's message s_k is independent and identically distributed according to $\mathcal{CN}(0, 1)$. Under this consideration, the achievable rate for user k can be written as:

$$R_k = \log \left(1 + \mathbf{w}_k^H \mathbf{H}_k^H \left(\sum_{\substack{j \neq k \\ j \in \mathcal{K}}} \mathbf{H}_k \mathbf{w}_j \mathbf{w}_j^H \mathbf{H}_k^H + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{H}_k \mathbf{w}_k \right). \qquad (2)$$

Note that the rate expression (2) can also account for the user scheduling operation. A user k is scheduled, i.e. R_k is nonzero, if and only if its beamformer vector $\mathbf{w}_k$ is nonzero. In other words, the scheduling choice is determined by the indicator function:

$$\mathbb{1}\left\{ \|\mathbf{w}_k\|_2^2 \right\} = \begin{cases} 0, & \text{if } \|\mathbf{w}_k\|_2^2 = 0 \\ 1, & \text{otherwise.} \end{cases} \qquad (3)$$

In this manner, the user scheduling, BS clustering and beamforming design for the downlink C-RAN is unified within this

---

[1]In the simulation part of this paper, only the strongest few BSs around each user are considered as the candidate serving BSs in order to reduce the computational complexity of the proposed algorithms.

single task of determining the sparse beamforming vector $\mathbf{w}_k$ for each user.

In this paper, we assume that the CP has access to global CSI for designing the sparse beamforming vector $\mathbf{w}_k$. Once $\mathbf{w}_k$ is determined, the CP transmits user k's message, along with the beamforming coefficients, to those BSs corresponding to the nonzero entries in $\mathbf{w}_k$ through the backhaul links. We also assume that the channels are slow varying and only consider the backhaul consumption due to the user data sharing and ignore the backhaul required for sharing CSI and delivering beamforming coefficients.

Intuitively, the backhaul consumption at the lth BS is the accumulated data rates of the users served by BS l. Notationally, we can characterize whether or not user k is served by BS l using the indicator function $\mathbb{1}\left\{ \|\mathbf{w}_k^l\|_2^2 \right\}$ and cast the per-BS backhaul constraint as:

$$\sum_{k \in \mathcal{K}} \mathbb{1}\left\{ \|\mathbf{w}_k^l\|_2^2 \right\} R_k \leq C_l, \quad \forall l. \qquad (4)$$

Specifically, in the case where the serving cluster for each user is fixed, or equivalently the set of users associated with each BS is predetermined, the backhaul consumption at BS l is also equal to the accumulated data rates of users associated with BS l, which can be formulated as

$$\sum_{k \in \mathcal{K}_l} R_k \leq C_l, \quad \forall l \qquad (5)$$

where $\mathcal{K}_l \subseteq \mathcal{K}$ denotes the fixed subset of users associated with BS l. Note that in each time-frequency slot, only the subset of users scheduled to be served have nonzero rates. So in (5), summing over the set of users associated with BS l is equivalent to summing over the set of scheduled users.

From (4) and (5), we see that the backhaul consumption is a function of both the cluster size and the user rate, where in addition the user rate is a function of user scheduling and beamforming operation. This observation provides us with different degrees of freedom in controlling the backhaul consumption depending on whether the BS clustering is dynamic or fixed in different user scheduling time slots. When the BS clustering is dynamic, we can jointly design the clustering, scheduling and beamforming to satisfy the per-BS backhaul constraint expressed in (4). But even when the BS clustering is fixed (or static), we can still control the user rates through scheduling and beamforming vectors to make sure that the backhaul constraint expressed in (5) is satisfied. In the following two sections, we discuss in detail how to incorporate the per-BS backhaul constraints in network utility maximization framework for downlink C-RAN under the above two different situations.

III. UTILITY MAXIMIZATION WITH DYNAMIC BS CLUSTERING

In this section, we propose a joint dynamic clustering, user scheduling and beamforming design strategy for the downlink C-RAN. The proposed algorithm designs a group sparse beamforming vector $\mathbf{w}_k$ for each user in each scheduling

slot using an ℓ_1-norm reweighting technique followed by a WMMSE approach under explicit per-BS backhaul capacity constraints.

A. PROBLEM FORMULATION

This paper considers network utility maximization as the objective. Among the family of utility functions, WSR has been widely applied to network control and optimization problems. In this paper, we also adopt the WSR utility but point out that the proposed scheme can be readily extend to any utility function that holds an equivalence relationship with the WMMSE problem (see [12] for a sufficient condition on the utility functions holding such an equivalence).

In dynamic BS clustering, the serving cluster for each user is a variable to be optimized in each scheduling slot. We adopt the per-BS backhaul constraint formulation (4) and formulate the WSR maximization problem under per-BS power constraints and per-BS backhaul constraints as:

$$\underset{\{\mathbf{w}_k^l | l \in \mathcal{L}, k \in \mathcal{K}\}}{\text{maximize}} \sum_{k \in \mathcal{K}} \alpha_k R_k \tag{6a}$$

$$\text{subject to} \sum_{k \in \mathcal{K}} \left\| \mathbf{w}_k^l \right\|_2^2 \leq P_l, \quad \forall l \tag{6b}$$

$$\sum_{k \in \mathcal{K}} \mathbb{1}\left\{ \left\| \mathbf{w}_k^l \right\|_2^2 \right\} R_k \leq C_l, \quad \forall l \tag{6c}$$

where α_k denotes the priority weight associated with user k at the current scheduling slot which can be updated according to, for example, the proportional fairness criterion. Here, P_l and C_l represent the transmit power budget and the backhaul capacity limit for BS l, respectively. The rate R_k shown in (6a) and (6c) is defined in (2), which is a function of the set of sparse beamforming vectors $\mathbf{w}_k$'s only. Note that $\mathbf{w}_k$ is comprised of the beamforming vectors $\mathbf{w}_k^l$'s. Thus, the optimization variables for problem (6) are the set of $\mathbf{w}_k^l$'s.

B. PROPOSED ALGORITHM

The conventional WSR maximization problem is a well-known nonconvex optimization problem, for which finding the global optimal solution is already quite challenging even without the additional mixed discrete and continuous backhaul constraint (6c). This paper focuses on heuristic algorithms for approaching a local optimum solution to the problem (6) only. Our main contribution is a new way of dealing with the discrete indicator function in constraint (6c).

A key observation made in this paper is that the indicator function in (6c) can also be equivalently expressed as an ℓ_0-norm of a scalar. The ℓ_0-norm is the number of nonzero entries in a vector. So it reduces to an indicator function in the scalar case. This equivalent expression allows us to use ideas from the compressive sensing literature [43], where a nonconvex ℓ_0-norm optimization objective can often be approximated by a convex reweighted ℓ_1-norm, i.e.

$$\|\mathbf{x}\|_0 \approx \sum_i \beta_i |x_i| \tag{7}$$

where x_i denotes the ith component in the vector $\mathbf{x}$ and β_i is the weight associated with x_i. By properly choosing weights β_i's, the minimization of $\|\mathbf{x}\|_0$ can be effectively solved through the minimization of $\sum_i \beta_i |x_i|$ instead.

This paper goes one step further in that we extend the reweighted ℓ_1-norm approximation technique (7) originally proposed for minimizing the ℓ_0-norm in the objective to dealing with the ℓ_0-norm in the constraint. In particular, we rewrite the indicator function $\mathbb{1}\left\{ \left\| \mathbf{w}_k^l \right\|_2^2 \right\}$ as

$$\mathbb{1}\left\{ \left\| \mathbf{w}_k^l \right\|_2^2 \right\} = \left\| \left\| \mathbf{w}_k^l \right\|_2^2 \right\|_0, \tag{8}$$

and reformulate the backhaul constraint (6c) as:

$$\sum_{k \in \mathcal{K}} \beta_k^l \left\| \mathbf{w}_k^l \right\|_2^2 R_k \leq C_l \tag{9}$$

where β_k^l is a constant weight associated with BS l and user k and is updated iteratively according to

$$\beta_k^l = \frac{1}{\left\| \mathbf{w}_k^l \right\|_2^2 + \tau}, \quad \forall k, l \tag{10}$$

with some small constant regularization factor $\tau > 0$ and $\left\| \mathbf{w}_k^l \right\|_2^2$ from the previous iteration.

The heuristic weight updating rule (10) is motivated by the fact that by choosing β_k^l to be inversely proportional to the transmit power level $\left\| \mathbf{w}_k^l \right\|_2^2$, those BSs with lower transmit power to user k would have higher weights and would be forced to further reduce its transmit power and encouraged to drop out of the BS cluster eventually. Note that not only the BS cluster formation, but also the user scheduling can be controlled through $\left\| \mathbf{w}_k^l \right\|_2^2$ since the user k is scheduled if and only if there exists at least one BS $l \in \mathcal{L}$ such that $\left\| \mathbf{w}_k^l \right\|_2^2 \neq 0$.

However, even with the above approximation, the optimization problem (6) with the backhaul constraint (6c) replaced by (9) is still difficult to deal with, due to the fact that the rate R_k appears in both the objective function and the constraints. To address this difficulty, we propose to solve the problem (6) iteratively with fixed rate $\hat{R}_k$ in (9) obtained from the previous iteration. Under fixed β_k^l and $\hat{R}_k$, problem (6) now reduces to

$$\underset{\{\mathbf{w}_k^l | l \in \mathcal{L}, k \in \mathcal{K}\}}{\text{maximize}} \sum_{k \in \mathcal{K}} \alpha_k R_k \tag{11a}$$

$$\text{subject to} \sum_{k \in \mathcal{K}} \left\| \mathbf{w}_k^l \right\|_2^2 \leq P_l, \quad \forall l \tag{11b}$$

$$\sum_{k \in \mathcal{K}} \beta_k^l \hat{R}_k \left\| \mathbf{w}_k^l \right\|_2^2 \leq C_l, \quad \forall l \tag{11c}$$

where the approximated backhaul constraint (11c) can be interpreted as a *weighted* per-BS power constraint bearing a resemblance to the traditional per-BS power constraint (11b).

Although the approximated problem (11) is still nonconvex, we can reformulate it as an equivalent WMMSE problem and use the block coordinate descent method to reach a stationary point of (11). The equivalence between WSR maximization and WMMSE is first established in [11] for MIMO broadcast channel and later generalized to MIMO interfering channel in [12] and MIMO interfering channel with partial cooperation in [29].

It is not difficult to see that the generalized WMMSE equivalence established in [12] also extends to the problem (11) with the newly introduced weighted per-BS power constraint (11c). We explicitly state the equivalence as follows:

Proposition 3.1: The WSR maximization problem (11) has the same optimal solution as the following WMMSE problem:

$$\underset{\{\rho_k, \mathbf{u}_k, \mathbf{w}_k^l | l \in \mathcal{L}, k \in \mathcal{K}\}}{\text{minimize}} \sum_{k \in \mathcal{K}} \alpha_k \left(\rho_k e_k - \log \rho_k\right)$$

$$\text{subject to} \quad \sum_{k \in \mathcal{K}} \left\|\mathbf{w}_k^l\right\|_2^2 \le P_l, \quad \forall l$$

$$\sum_{k \in \mathcal{K}} \beta_k^l \hat{R}_k \left\|\mathbf{w}_k^l\right\|_2^2 \le C_l, \quad \forall l \quad (12)$$

where ρ_k denotes the MSE weight for user k and e_k is the corresponding MSE defined as

$$e_k = \mathsf{E}\left[\left\|\mathbf{u}_k^H \mathbf{y}_k - s_k\right\|_2^2\right]$$

$$= \mathbf{u}_k^H \left(\sum_{j \in \mathcal{K}} \mathbf{H}_k \mathbf{w}_j \mathbf{w}_j^H \mathbf{H}_k^H + \sigma^2 \mathbf{I}\right) \mathbf{u}_k$$

$$-2\text{Re}\left\{\mathbf{u}_k^H \mathbf{H}_k \mathbf{w}_k\right\} + 1 \quad (13)$$

under the receiver $\mathbf{u}_k \in \mathbb{C}^{N \times 1}$.

The advantage of solving WSR maximization problem (11) through its equivalent WMMSE problem (12) is that (12) is convex with respect to each of the individual optimization variables. This crucial fact allows problem (12) to be solved efficiently through the block coordinate descent method by iteratively optimizing over ρ_k, $\mathbf{u}_k$ and $\mathbf{w}_k$:

- The optimal MSE weight ρ_k under fixed $\mathbf{w}_k$ and $\mathbf{u}_k$ is given by

$$\rho_k = e_k^{-1}, \quad \forall k. \quad (14)$$

- The optimal receiver $\mathbf{u}_k$ under fixed $\mathbf{w}_k$ and ρ_k is the MMSE receiver:

$$\mathbf{u}_k = \left(\sum_{j \in \mathcal{K}} \mathbf{H}_k \mathbf{w}_j \mathbf{w}_j^H \mathbf{H}_k^H + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{H}_k \mathbf{w}_k, \quad \forall k. \quad (15)$$

- The optimization problem for finding the optimal transmit beamformer $\mathbf{w}_k$ under fixed $\mathbf{u}_k$ and ρ_k is a quadratically constrained quadratic program-

ming (QCQP) problem:

$$\underset{\{\mathbf{w}_k^l | l \in \mathcal{L}, k \in \mathcal{K}\}}{\text{minimize}} \sum_{k \in \mathcal{K}} \mathbf{w}_k^H \left(\sum_{j \in \mathcal{K}} \alpha_j \rho_j \mathbf{H}_j^H \mathbf{u}_j \mathbf{u}_j^H \mathbf{H}_j\right) \mathbf{w}_k$$

$$- 2 \sum_{k \in \mathcal{K}} \alpha_k \rho_k \text{Re}\left\{\mathbf{u}_k^H \mathbf{H}_k \mathbf{w}_k\right\}$$

$$\text{subject to} \sum_{k \in \mathcal{K}} \left\|\mathbf{w}_k^l\right\|_2^2 \le P_l, \quad \forall l$$

$$\sum_{k \in \mathcal{K}} \beta_k^l \hat{R}_k \left\|\mathbf{w}_k^l\right\|_2^2 \le C_l, \quad \forall l \quad (16)$$

which can be solved using a standard convex optimization solver such as CVX [44].

A straightforward way of applying the above WMMSE algorithm to solve the original problem (6) would involve two loops: an inner loop to solve the approximated WSR maximization problem (11) with fixed weight β_k^l and rate $\hat{R}_k$, and an outer loop to update β_k^l and $\hat{R}_k$. Although such an algorithm can guarantee that the inner loop converges to a stationary point of the problem (11), its computational complexity can be high. Instead, we propose to combine these two loops into a single loop and update the weight β_k^l and rate $\hat{R}_k$ inside the WMMSE algorithm, as summarized in Algorithm 1. Although Algorithm 1 does not have a proof of convergence, numerical simulation shows that it converges reasonably fast, and it significantly outperforms the fixed clustering baseline schemes.

Algorithm 1 WSR Maximization With Per-BS Backhaul Constraints Under Dynamic BS Clustering

---

Initialization: $\beta_k^l, \hat{R}_k, \mathbf{w}_k, \forall l, k$;
Repeat:
1) Fix $\mathbf{w}_k, \forall k$, compute the MMSE receiver $\mathbf{u}_k$ and the corresponding MSE e_k according to (15) and (13);
2) Update the MSE weight ρ_k according to (14);
3) Find the optimal transmit beamformer $\mathbf{w}_k$ under fixed $\mathbf{u}_k$ and ρ_k, $\forall k$, by solving the QCQP problem (16);
4) Compute the achievable rate R_k according to (2), $\forall k$;
5) Update $\hat{R}_k = R_k$ and β_k^l according to (10), $\forall l, k$.

Until convergence

C. COMPLEXITY ANALYSIS

Assuming a typical network with $K > L > M > N$, the computational complexity of Step 1 in Algorithm 1 is $O(K^2 LMN)$, mainly due to the receive covariance matrix computation in (15) and (13). With the MSE e_k obtained from Step 1, the additional computational complexity for Step 2 for updating all the MSE weights ρ_k's is only $O(K)$. Step 3 requires solving a QCQP problem, which can also be equivalently reformulated as a second order cone programming (SOCP) problem as we do in the simulation part of this paper. The total number of variables in the equivalent SOCP problem is

KLM and the computation complexity of using interior-point method to solve such an SOCP problem is approximately $O((KLM)^{3.5})$ [45]. In Step 5, the rate updating procedure requires the computation of the achievable rate in Step 4 according to (2), which has the same computational complexity as computing the MSE, i.e. $O(K^2LMN)$. As we can see, the computational complexity of Algorithm 1 per iteration mainly comes from the optimal transmit beamformer design in Step 3. Suppose Algorithm 1 requires T total number of iterations to converge, the overall computational complexity of Algorithm 1 is therefore $O((KLM)^{3.5}T)$.

D. HEURISTIC COMPLEXITY REDUCTION TECHNIQUES

To improve the efficiency of Algorithm 1 in each iteration, in what follows, we further propose two techniques, *iterative link removal* and *iterative user pool shrinking*. The former aims at reducing the number of potential transmit antennas *LM* serving each user while the latter is intended to decrease the total number of users K to be considered in each iteration.

1) ITERATIVE LINK REMOVAL

Similar to what we observed in [20], the transmit power from some of the candidate serving BSs would drop down rapidly as the iterations go on. By taking advantage of this fact, we propose to iteratively remove the *l*th BS from the *k*th user's candidate cluster once the transmit power from BS l to user k, i.e. $\left\|\mathbf{w}_k^l\right\|_2^2$, is below a certain threshold, say -100 dBm/Hz. This reduces the dimension of the potential transmit beamformer for each user and reduces the complexity of solving SOCP in Step 3 of Algorithm 1.

2) ITERATIVE USER POOL SHRINKING

The proposed algorithm does user scheduling implicitly. We observe from simulations that, it is beneficial for Algorithm 1 to consider a large pool of users. However, to consider all the users in the entire network all the time would incur significant computational burden. Instead, we propose to check the achievable user rate R_k in Step 4 in each iteration and ignore those users with negligible rates (below some threshold, say 0.01 bps/Hz) for subsequent iterations. Our simulations show that, after around 10 iterations, more than half of the total users can be taken out of the consideration with negligible performance loss to the overall algorithm. This significantly reduces the total number of variables to be optimized.

IV. UTILITY MAXIMIZATION WITH STATIC BS CLUSTERING

In the previous section, BS clustering is dynamically determined in each time-frequency slot together with the beamforming vector and user scheduling in a joint fashion. However, dynamic BS clustering may incur significant signaling overhead in practice as new BS-user associations need to be established continuously over time. In this section, we discuss static clustering schemes, where the BS clusters

only need to be updated at much larger time scale, typically only when user locations change.

As discussed previously, backhaul consumption under static clustering can still be controlled by jointly optimizing the user scheduling and beamforming. In this section, we first adapt the sparse beamforming algorithm proposed in previous section to jointly schedule the users and design the beamforming vectors under per-BS backhaul constraints while assuming that the BS clustering is fixed. We then propose two heuristic static clustering schemes: one depends on the maximum number of users each BS can support and the long-term channel condition each user experiences; the other generalizes the SINR-bias technique used for cell range expansion [46] to form a static cluster for each user.

A. JOINT SCHEDULING AND BEAMFORMING DESIGN WITH FIXED BS CLUSTERING

Let $\mathcal{L}_k$ be the fixed cluster of BSs serving user k. The joint scheduling and beamforming design problem is that of determining the scheduled users in each time-frequency slot and the corresponding beamformers from the BSs in $\mathcal{L}_k$ to each scheduled user k while satisfying the per-BS power constraints and per-BS backhaul constraints.

Equivalently, let $\mathcal{K}_l$ be the set of users associated with BS l, the network utility maximization problem can now be formulated as

$$\underset{\{\mathbf{w}_k^l | k \in \mathcal{K}, l \in \mathcal{L}_k\}}{\text{maximize}} \sum_{k \in \mathcal{K}} \alpha_k R_k \tag{17a}$$

$$\text{subject to} \quad \sum_{k \in \mathcal{K}_l} \left\|\mathbf{w}_k^l\right\|_2^2 \le P_l, \quad \forall l \tag{17b}$$

$$\sum_{k \in \mathcal{K}_l} R_k \le C_l, \quad \forall l \tag{17c}$$

Note that the difference between the utility maximization problems (17) and (6) is that the transmit power and backhaul constraint for BS l now only need to take into account the fixed subset of users associated with BS l, $\mathcal{K}_l$. Also, the optimization variable $\mathbf{w}_k^l$ is only over the beamforming vectors from the BSs in each user k's serving cluster $\mathcal{L}_k$ since $\mathbf{w}_k^l = \mathbf{0}$ for $l \notin \mathcal{L}_k$.

Note that user scheduling is implicitly being optimized in (17). Only the subset of users scheduled in the current time-frequency slot would have nonzero rates and need to be considered in the summations (17b) and (17c). With this observation, we can rewrite (17c) in the following equivalent form:

$$\sum_{k \in \mathcal{K}_l} \mathbb{1}\left\{\left\|\mathbf{w}_k^{\mathcal{L}_k}\right\|_2^2\right\} R_k \le C_l, \quad \forall l. \tag{18}$$

where $\mathbf{w}_k^{\mathcal{L}_k} \in \mathbb{C}^{|\mathcal{L}_k|M \times 1}$ is the beamforming vector from user k's serving cluster $\mathcal{L}_k$ to user k. This allows us to utilize a similar idea as in previous section to solve problem (17) approximately by fixing the user rate R_k in the constraint and approximating the indicator function in (18) using the

reweighted ℓ_1-norm technique. The resulting approximated optimization problem to (17) now becomes

$$\underset{\{\mathbf{w}_k^l \mid k \in \mathcal{K}, l \in \mathcal{L}_k\}}{\text{maximize}} \sum_{k \in \mathcal{K}} \alpha_k R_k \tag{19a}$$

$$\text{subject to} \quad \sum_{k \in \mathcal{K}_l} \left\| \mathbf{w}_k^l \right\|_2^2 \leq P_l, \quad \forall l \tag{19b}$$

$$\sum_{k \in \mathcal{K}_l} \beta_k \hat{R}_k \left\| \mathbf{w}_k^{\mathcal{L}_k} \right\|_2^2 \leq C_l, \quad \forall l \tag{19c}$$

where β_k is the constant weight associated with the predetermined BS cluster $\mathcal{L}_k$ for user k and is updated iteratively according to

$$\beta_k = \frac{1}{\left\| \mathbf{w}_k^{\mathcal{L}_k} \right\|_2^2 + \tau}, \quad \forall k, \tag{20}$$

and $\hat{R}_k$ is the achievable rate from previous iteration.

Comparing the weight updating rule (20) with (10), we note that the role of β_k^l in problem formulation (11) is to determine whether or not BS l should serve user k in the current time-frequency slot, while the role of β_k in problem formulation (19) is to decide whether or not user k should be scheduled and served by its predetermined serving cluster $\mathcal{L}_k$ *as a whole* in the current time-frequency slot. Note that β_k^l only appears in the lth BS backhaul constraint in (11c), while in (19c) β_k appears in the approximated per-BS backhaul constraints corresponding to all of the user k's pre-associated cluster of BSs in $\mathcal{L}_k$.

With the approximated problem formulation (19), it becomes straightforward to extend Algorithm 1 for solving problem (17). The only necessary adaptation occurs in Step 5 of Algorithm 1, where the update of weight β_k^l is replaced by updating the weight β_k according to (20). Under given BS cluster $\mathcal{L}_k$, we rewrite the MMSE receiver, MSE and achievable rate for user k as below[2]:

$$\mathbf{u}_k = \left(\sum_{j \in \mathcal{K}} \mathbf{H}_k^{\mathcal{L}_j} \mathbf{w}_j^{\mathcal{L}_j} \left(\mathbf{H}_k^{\mathcal{L}_j} \mathbf{w}_j^{\mathcal{L}_j} \right)^H + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{H}_k^{\mathcal{L}_k} \mathbf{w}_k^{\mathcal{L}_k} \tag{21}$$

$$e_k = \mathbf{u}_k^H \left(\sum_{j \in \mathcal{K}} \mathbf{H}_k^{\mathcal{L}_j} \mathbf{w}_j^{\mathcal{L}_j} \left(\mathbf{H}_k^{\mathcal{L}_j} \mathbf{w}_j^{\mathcal{L}_j} \right)^H + \sigma^2 \mathbf{I} \right) \mathbf{u}_k$$
$$- 2\text{Re} \left\{ \mathbf{u}_k^H \mathbf{H}_k^{\mathcal{L}_k} \mathbf{w}_k^{\mathcal{L}_k} \right\} + 1 \tag{22}$$

$$R_k = \log \left(1 + \left(\mathbf{H}_k^{\mathcal{L}_k} \mathbf{w}_k^{\mathcal{L}_k} \right)^H \left(\sum_{\substack{j \neq k \\ j \in \mathcal{K}}} \mathbf{H}_k^{\mathcal{L}_j} \mathbf{w}_j^{\mathcal{L}_j} \left(\mathbf{H}_k^{\mathcal{L}_j} \mathbf{w}_j^{\mathcal{L}_j} \right)^H \right.$$
$$\left. + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{H}_k^{\mathcal{L}_k} \mathbf{w}_k^{\mathcal{L}_k} \right) \tag{23}$$

---

[2]One can also use the MMSE receiver, MSE and achievable rate defined previously in (15), (13) and (2) respectively by filling those entries corresponding to $\mathcal{L}_k$ in $\mathbf{w}_k$ with $\mathbf{w}_k^{\mathcal{L}_k}$ and the rest entries in $\mathbf{w}_k$ with zero.

where $\mathbf{H}_k^{\mathcal{L}_j}$ is the CSI matrix from user j's serving cluster $\mathcal{L}_j$ to user k. The proposed joint scheduling and beamforming design algorithm under fixed BS clustering is summarized in Algorithm 2.

It is worth noting that both Algorithm 1 and 2 implement user scheduling operation implicitly by optimizing the beamforming vectors for all the users in the entire network but only selecting those users with nonzero beamforming vectors to be served. This is in contrast to the conventional user scheduling approach, where typically a subset of users are pre-selected and only the beamforming vectors corresponding to those pre-selected users are optimized. Simulation results show that the proposed algorithms are able to achieve better performance by scheduling the users implicitly, although the performance gain comes at a complexity cost.

We also note that both Algorithm 1 and 2 require global CSI at the CP in order to schedule the users and to design beamformers accordingly, which may lead to large channel estimation overhead. Performance analysis of C-RAN with partial CSI has been carried on in [47] under a simplified model where the BSs and the users are equipped with a single antenna each and no backhaul constraint is considered. The impact of channel estimation overhead on the system performance of C-RAN under a more realistic model considered in this paper is nontrivial and is left for future work.

Algorithm 2 WSR Maximization With Per-BS Backhaul Constraints Under Static BS Clustering

---

Initialization: $\beta_k, \hat{R}_k, \mathbf{w}_k^{\mathcal{L}_k}, \forall k$;
Repeat:
1) Fix $\mathbf{w}_k^{\mathcal{L}_k}, \forall k$, compute the MMSE receiver $\mathbf{u}_k$ and the corresponding MSE e_k according to (21) and (22);
2) Update the MSE weight ρ_k according to (14);
3) Find the optimal transmit beamformer $\mathbf{w}_k^{\mathcal{L}_k}$ under fixed $\mathbf{u}_k$ and $\rho_k, \forall k$, by solving the QCQP problem (16) with $\mathbf{w}_k$, $\mathbf{H}_j$ and $\mathbf{H}_k$ replaced by $\mathbf{w}_k^{\mathcal{L}_k}$, $\mathbf{H}_j^{\mathcal{L}_k}$ and $\mathbf{H}_k^{\mathcal{L}_k}$ respectively;
4) Compute the achievable rate R_k according to (23), $\forall k$;
5) Update $\hat{R}_k = R_k$ and β_k according to (20), $\forall k$.

Until convergence

B. PROPOSED STATIC CLUSTERING ALGORITHMS

Thus far in this section, we have dealt with the joint user scheduling and beamforming design under per-BS backhaul capacity constraints for any given fixed clustering scheme. We now propose heuristic algorithms to optimize over the clustering strategies.

Differing from the traditional BS-user association problem where each user is only associated with a single BS, in C-RAN each user is served by a cluster of BSs. The optimal fixed BS clustering design for C-RAN is nontrivial as each user wants to be served by as many nearby BSs as possible while each BS can only support a limited number of users due to the limited radio resources, i.e. transmit power, and

limited backhaul capacity. A good clustering strategy for C-RAN should account for not only the channel strength from the BSs to each user but also the available resources at each BS.

This paper adopts the user-centric clustering strategy, in which each user is served by an individually selected and potentially overlapping subset of BSs. A simple way of forming user-centric clusters is to choose an equal number of strongest BSs around each user as its serving cluster. However, such a scheme may produce imbalanced traffic loads across the network, especially in a heterogeneous deployment where the macro-BSs typically have much higher transmit power than the pico-BSs, so more than the optimal number of users would associate with the macro-BSs. This paper proposes two heuristic static clustering schemes to address this load balancing issue. The first scheme is based on imposing a maximum load on each BS. The second scheme is based on introducing a bias term to the received signal strength at each user. These two schemes are described in detail below.

1) MAXIMUM LOADING BASED STATIC CLUSTERING

In order to avoid BS overloading, in this scheme, we propose to set an upper bound on the number of users that can associate with each BS l, denoted as $K_{l,max}$. The value of $K_{l,max}$ depends on the amount of resources available at BS l. For example, in a heterogeneous network, the macro-BSs usually have higher transmit power and more backhaul capacity than the pico-BSs, hence $K_{l,max}$ for macro-BSs should be larger than that of pico-BSs. With $K_{l,max}$, those BSs that have already reached its maximum number of associated users would direct the subsequent users to other underloaded BSs.

At the user's side, each user wants to be served by as many nearby BSs as possible to obtain the highest service rate. However, since a cell-center user typically already experiences good channel condition, it is reasonable to connect it with fewer BSs, whereas a cell-edge user may need more serving BSs to coordinately mitigate inter-cell interference. To capture the difference in the ideal cluster size for different users, which depends on their relative locations, we propose to set a candidate BS cluster for each user based on a threshold on the received signal strength difference. For each user k, only those BSs from which the received signal strength is within η_1 gap to the signal strength from the strongest BS are considered as potential serving BSs. Mathematically, let $s_{l,k}$ be the received signal strength from BS l to user k, defined as the maximum transmit power from BS l compensated by the path loss to user k without accounting for possible antenna beamforming gain, the candidate serving cluster $\mathcal{C}_k$ for user k is a set defined as follows:

$$\mathcal{C}_k = \left\{ l \in \mathcal{L} \mid \max_m s_{m,k} - s_{l,k} \leq \eta_1 \right\}. \quad (24)$$

Since a cell-edge user sees more nearby BSs with similar signal strength than a cell-center user, the candidate cluster size $|\mathcal{C}_k|$ for a cell-edge user would be larger, which potentially results in more serving BSs for the cell-edge user.

Based on the aforementioned parameter $K_{l,max}$ and set $\mathcal{C}_k$, we propose a simple user-centric clustering scheme based on the following two heuristics:

- Each user k sends requests to the BSs in the candidate set $\mathcal{C}_k$ sequentially from the strongest to the weakest;
- Each BS l accepts up to $K_{l,max}$ users.

The details of the first proposed static clustering scheme is shown in Algorithm 3, which requires a multi-round negotiation between the BSs and the users until the BSs are connected with the maximum number of users or the users have exhausted all their BSs in the candidate sets.

Algorithm 3 Maximum Loading Based Static Clustering

Initialization:
1) $\mathcal{K} = \{1, 2, \cdots, K\}$, $\mathcal{L} = \{1, 2, \cdots, L\}$;
2) Let $\mathcal{K}_l$ be the set of users associated with the lth BS. Set $\mathcal{K}_l = \emptyset$, $l = 1, 2, \cdots, L$;
3) Let $K_{l,max}$ be the maximum number of users BS l can support, $l = 1, 2, \cdots, L$;
4) Let $\mathcal{C}_k$ be the candidate serving cluster for user k, $k = 1, 2, \cdots, K$;
5) Set iteration index $i = 1$.

Repeat:
1) Each user $k \in \mathcal{K}$ sends a request to the ith strongest BS in $\mathcal{C}_k$;
2) For each BS $l \in \mathcal{L}$:
   If $K_{l,max} - |\mathcal{K}_l| \geq$ total number of received requests
   2.1) $\mathcal{K}_l = \mathcal{K}_l \cup \{$All the received requests$\}$;
   otherwise
   2.2) $\mathcal{K}_l = \mathcal{K}_l \cup \{$The $\left(K_{l,max} - |\mathcal{K}_l|\right)$ strongest users among all the received requests$\}$;
   2.3) $\mathcal{L} = \mathcal{L} \setminus \{l\}$.
   end if
3) For each user $k \in \mathcal{K}$, if it has exhausted all the candidate BSs in the list $\mathcal{C}_k$, update $\mathcal{K} = \mathcal{K} \setminus \{k\}$;
4) $i = i + 1$.

Until $\mathcal{L} = \emptyset$ or $\mathcal{K} = \emptyset$

We remark that the parameters $K_{l,max}$ and $\mathcal{C}_k$ jointly play an important role in determining the serving clusters in Algorithm 3. Using $K_{l,max}$ or $\mathcal{C}_k$ alone would not have produced a good clustering scheme. For instance, suppose each BS is simply associated with the $K_{l,max}$ strongest users it sees, then the cell-edge users would be at a risk of connecting with no BSs. Or, if each user is served by all the BSs in its candidate list $\mathcal{C}_k$, then the high-power BSs may be overloaded. Only by taking into account both the traffic load of each BS and the channel condition of each user through the parameters $K_{l,max}$ and $\mathcal{C}_k$ jointly would the proposed Algorithm 3 be able to produce a good clustering scheme.

It is worth noting that the use of $(K_{l,max}, \mathcal{C}_k)$ in Algorithm 3 is only one possibility in jointly controlling the traffic load of the BSs and the channel condition of the users.

For example, one can also set the candidate cluster for user k as the strongest L_k BSs instead of the proposed $\mathcal{C}_k$. Such an algorithm also has reasonable performance. However, to find a good L_k for each user in the entire network is not an easy task. Suppose that the L_k is set to be equal for all users to simplify the search, it may result in the cell-center users and the cell-edge users being associated with an equal (or at least a similar) number of BSs, which may lead to inefficient usage of the backhaul resources. The proposed candidate cluster $\mathcal{C}_k$ can be alternatively seen as one way to set a nonuniform L_k for each user (as $L_k = |\mathcal{C}_k|$) through a common parameter η_1 as in (24).

2) BIASED SIGNAL STRENGTH BASED STATIC CLUSTERING

An alternative method for controlling the traffic load of the BSs is to set a received signal strength bias ζ_l for each BS l and to determine the user-centric clusters based on the biased signal strength. By setting a higher bias for the underloaded BSs, users around the overloaded BSs are prompted to connect with the underloaded BSs instead. This biasing idea originates from the SINR-bias technique for cell range expansion in the traditional BS-user association problem for heterogeneous networks [46], where each user is only associated with a single BS. The proposed algorithm generalizes this idea to the case where each user can associate with a cluster of multiple BSs, and further combines the idea of biasing with the idea of using a received signal strength gap threshold, denoted as η_2, to determine the cluster sizes for different users, as done in (24). The proposed biased signal strength based static clustering scheme is described in detail as Algorithm 4 below.

Algorithm 4 Biased Signal Strength Based Static Clustering

Let ζ_l be the received signal strength bias from BS l, $l = 1, 2, \cdots, L$, the serving BS cluster for user k is set as:

$$\mathcal{L}_k = \left\{ l \in \mathcal{L} \mid \max_m \left(s_{m,k} + \zeta_m \right) - \left(s_{l,k} + \zeta_l \right) \leq \eta_2 \right\}$$

$k = 1, 2, \cdots, K$, where $s_{m,k}$ is the received signal strength from BS m to user k.

V. SIMULATION RESULTS

In this section, numerical simulations are conducted to show the effectiveness of the proposed algorithms. We consider a 7-cell wrapped-around two-tier heterogeneous network with the simulation parameters listed in Table 1. Each cell is a regular hexagon with a single macro-BS located at the center and 3 pico-BSs equally separated in space as illustrated in Fig. 2. To simplify the discussion, we set all the macro-BSs to have equal backhaul constraints and likewise for the pico-BSs. The backhaul constraints are denoted as $(C_{\text{macro}}, C_{\text{pico}})$ respectively. The proposed algorithms are simulated under the same power constraints listed in Table 1 but with various sets of $(C_{\text{macro}}, C_{\text{pico}})$ backhaul constraints.

TABLE 1. Simulation parameters.

| 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	$\tau = 10^{-10}$
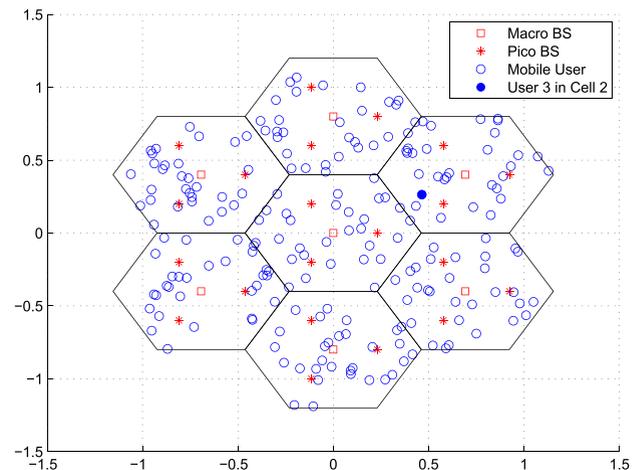
FIGURE 2. 7-cell wrapped around two-tier heterogeneous network.

A. DYNAMIC BS CLUSTERING

We first evaluate the performance of the proposed Algorithm 1 under dynamic BS clustering. As indicated earlier, instead of considering all the L BSs in the entire network as candidates for serving each user, in simulations we only consider the strongest L_c ($L_c \leq L$) BSs around each user as its candidate cluster. To illustrate how the sparse beamforming vector $\mathbf{w}_k$ for each user is formed using Algorithm 1, we plot in Fig. 3 the power evolutions of the strongest 8 BSs for the third user in the second cell as an example. As we can see, after around 20 iterations only the first and third strongest BSs maintain a reasonable transmit power level. They eventually form the cluster to serve user 3 in cell 2. With the proposed iterative link removal technique and by setting the threshold to be -100 dBm/Hz, Algorithm 1 can narrow down the candidate BSs to only the strongest 4 BSs after the 5th iteration, and to the (1st, 3rd, 4th) strongest BSs after the 8th iteration, and finally to the (1st, 3rd) strongest BSs after the 17th iteration.

To demonstrate the effectiveness of the proposed iterative link removal and iterative user pool shrinking techniques in improving the efficiency of Algorithm 1, we plot in Fig. 4 the CPU execution time needed for each iteration in Algorithm 1. As we can see, the per-iteration execution time for Algorithm 1 drops dramatically from around 170 seconds to about 5 seconds within 20 iterations. This is due to the continually shrinking candidate cluster size and user scheduling pool. For instance, at the 20th iteration, the
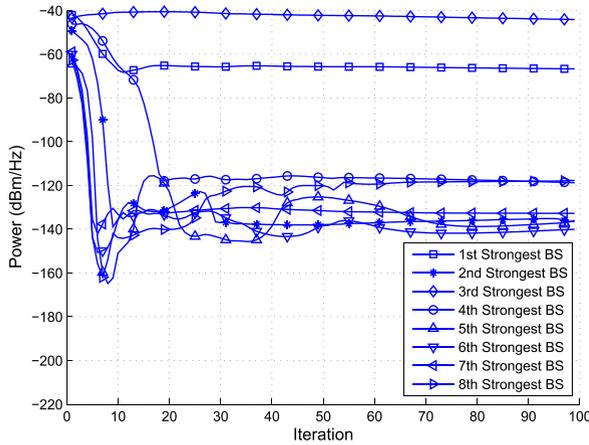
FIGURE 3. Power evolutions of the strongest 8 BSs for user 3 in cell 2, $(C_{\mathrm{macro}}, C_{\mathrm{pico}}) = (245, 70)$ Mbps, $\alpha_k = 1, \forall k, L_c = 8$.

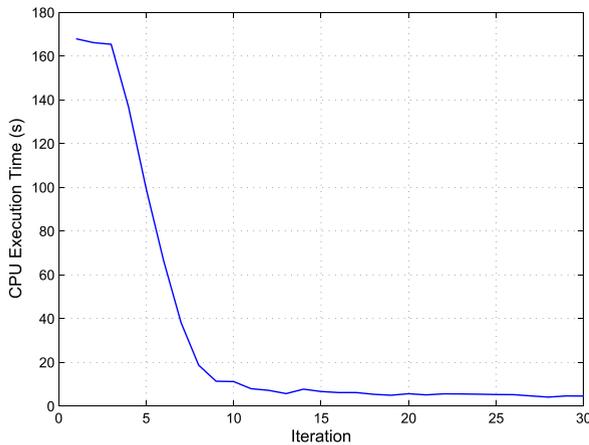**FIGURE 4.** CPU execution time needed for each iteration in a linux x86_64 machine with 2.3 GHz CPU and 2 GB RAM under $(C_{\mathrm{macro}}, C_{\mathrm{pico}}) = (245, 70)$ Mbps, $\alpha_k = 1, \forall k, L_c = 8$.

average candidate cluster size for each user is 1.96 and the number of remaining users in the scheduling pool is 54, which are only about 1/4 of the original cluster size $L_c = 8$ and the total number of users $K = 210$, respectively.

In Fig. 5, we compare the cumulative distributions of the long-term average user rates between Algorithm 1 with dynamic clustering and the baseline scheme where each user is served by the strongest S BSs ($S = 1, 2, 3, 4$). We first run the baseline schemes using the WMMSE algorithm without explicit backhaul constraints as in [29] and evaluate the corresponding backhaul requirement for each BS afterwards. We then set the explicit backhaul constraints $(C_{\mathrm{macro}}, C_{\mathrm{pico}})$ in Algorithm 1 to be the average backhaul requirements over the macro-BSs and the pico-BSs respectively computed from the baseline. The baseline is denoted as "Strongest S BSs ($C_{\mathrm{macro}}, C_{\mathrm{pico}}$) Mbps" while the proposed Algorithm 1 is denoted as "Dynamic Clustering ($C_{\mathrm{macro}}, C_{\mathrm{pico}}$) Mbps." Each curve in Fig. 5 is obtained by iteratively simulating the corresponding scheme with fixed user priority weights α_k's and updating the weights as $\alpha_k = 1/\bar{R}_k$ according to the proportional fairness criterion, where $\bar{R}_k$ is the long term average data rate for user k.
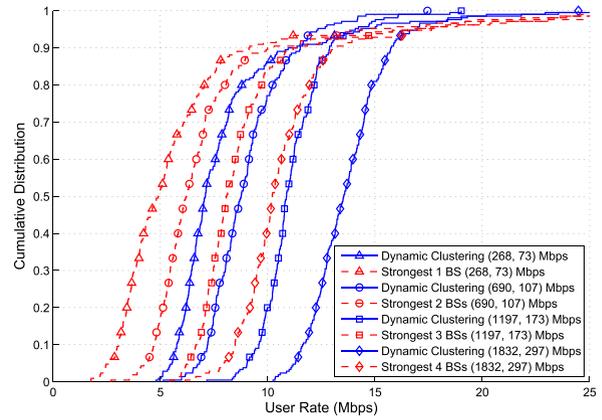

FIGURE 5. Cumulative distribution function of user data rate comparison with $L_c = 8$ and proportionally fair scheduling.

As we can see from Fig. 5, by optimizing with explicit backhaul constraints, Algorithm 1 achieves significant performance gain. We list the 50th percentile user rate comparison between Algorithm 1 and the baseline in Table 2. For instance, around 35% improvement is obtained for the 50th percentile user as compared with the baseline where each user is served by the strongest 3 BSs. Note that since the baseline algorithm connects each user to an equal number of neighboring BSs, it inevitably favors high-rate users. In contrast, the proposed Algorithm 1 can select clusters for each user adaptively, and in particular can choose a larger cluster for the low-rate users, thus achieving a significant overall gain from a network utility perspective.

TABLE 2. 50th percentile user data rate comparison.

	50th Percentile Rate	Gain
Strongest 1 BS (268, 73) Mbps	4.9 Mbps	44.9%
Dynamic Clustering (268, 73) Mbps	7.1 Mbps	
Strongest 2 BSs (690, 107) Mbps	6.2 Mbps	41.9%
Dynamic Clustering (690, 107) Mbps	8.8 Mbps	
Strongest 3 BSs (1197, 173) Mbps	8.1 Mbps	34.6%
Dynamic Clustering (1197, 173) Mbps	10.9 Mbps	
Strongest 4 BS (1832, 297) Mbps	10.2 Mbps	33.3%
Dynamic Clustering (1832, 297) Mbps	13.6 Mbps	

We also see from Fig. 5 that the "Dynamic Clustering (690, 107) Mbps" scheme shows better performance as compared to the baseline where each user is served by the strongest 3 BSs as in the "Strongest 3 BSs (1197, 173) Mbps" scheme, while only requires about 60% of backhaul. To further investigate the reason behind this significant reduction in backhaul consumption, we plot the probability density function (pdf) of the backhaul consumption in each time slot for all the 7 macro-BSs in Fig. 6. As we can see, the macro-BS backhaul consumption for Algorithm 1 has a peak around its constraint[3] and is highly concentrated. However, the backhaul

---

[3]The slight excess over the backhaul constraint in the pdf curve is due to the embedded pdf estimation function in MATLAB, which interpolates more points at the boundary to smooth the estimated curve.
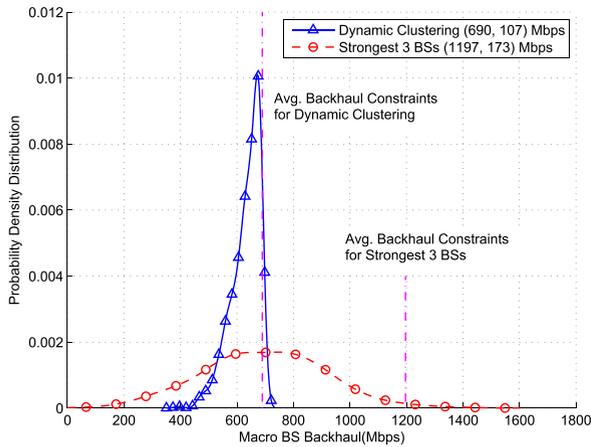
FIGURE 6. Probability density distribution of maco-BS backhaul consumption, where the straight magenta dashed line represents the position of the corresponding C_{macro}.
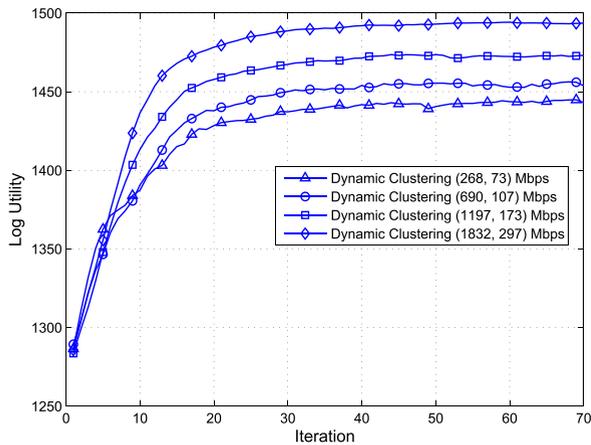

FIGURE 7. Convergence behavior of $\sum_k \log(\bar{R}_k)$ for Algorithm 1 under dynamic clustering, where $\bar{R}_k$ is the long term average rate for user k and $L_c = 8$.

consumption for the baseline is more spread out, whose peak point is only about half of its backhaul requirement. Similar phenomenon also occurs in the backhaul consumption for the pico-BSs. The inefficient usage of the backhaul resource in the baseline is due to fixed clustering and the lack of explicit backhaul constraints. In contrast, the proposed dynamic clustering with explicit backhaul constraints can adaptively form the clusters for each user according to the available backhaul budget, which also has an effect of balancing the data traffic.

Although a rigorous theoretical proof for the convergence of the proposed Algorithm 1 is not yet available, we plot in Fig. 7 the log utility evolutions for Algorithm 1 under the same set of backhaul constraints as in Fig. 5. As we can see, the proposed Algorithm 1 with dynamic BS clustering converges in roughly 40-50 iterations under each of the (C_{macro}, C_{pico}) settings.

B. STATIC BS CLUSTERING
We now evaluate the performance of Algorithm 2 under static BS clustering schemes. In Fig. 8, we set the ''Strongest 2 BSs'' scheme from Fig. 5 as the baseline and set
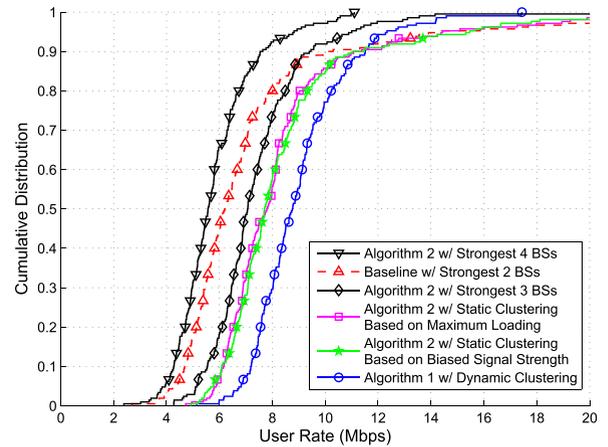

FIGURE 8. Cumulative distribution function of user data rate comparison with backhaul constraint (C_{macro}, C_{pico}) = (690, 107) Mbps under proportionally fair scheduling.

its corresponding backhaul requirement (C_{macro}, C_{pico}) = (690, 107) Mbps as the explicit backhaul constraint for Algorithm 2 under different static BS clustering schemes. We first compare the baseline with Algorithm 2 in which each user is associated with the strongest 3 and 4 BSs, respectively. As we can see, by connecting each user to an additional BS (strongest 3 BSs), the proposed Algorithm 2 with explicit backhaul constraints improves the overall network utility by sacrificing the high-rate users to improve the performance of the low-rate users while still keeping the same backhaul consumption. However, if each user is further connected to one more additional BS (strongest 4 BSs), then the overall performance is even worse than the case where each user is connected to the strongest 2 BSs. This is because the BSs are now overloaded under the given backhaul capacity constraints and are forced to schedule only the low-rate users. This illustrates the importance of choosing BS cluster size.

We also plot in Fig. 8 the performance of Algorithm 2 under the proposed static clustering schemes as listed in Algorithm 3 and Algorithm 4, respectively, again with (C_{macro}, C_{pico}) = (690, 107) Mbps. In the maximum loading based static clustering scheme, the maximum number of users each macro-BS and each pico-BS can support is set to be $K_{macro,max} = 70, K_{pico,max} = 10$ and the received signal strength gap is set as $\eta_1 = 14$ dB. In the biased signal strength based static clustering scheme, the received signal strength bias is set as $\zeta_{macro} = 0$ dB for the macro-BSs and $\zeta_{pico} = 6$ dB for the pico-BSs and the received signal strength gap is set as $\eta_2 = 12$ dB. As we can see, the two proposed clustering schemes have similar performance, and both can achieve a significant portion of the performance gain that the dynamic clustering scheme achieves over the baseline of simply choosing the strongest BSs. For instance, compared with the baseline at the 50th percentile user rate, both of the proposed static clustering schemes show around 26% performance improvement, while the proposed Algorithm 1 under dynamic clustering achieves 41.9% gain as listed in Table 2.
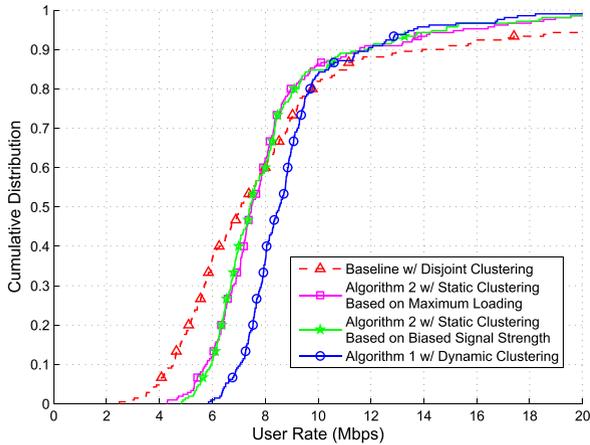
FIGURE 9. Cumulative distribution function of user data rate comparison with backhaul constraint ($C_{\text{macro}}, C_{\text{pico}}) = (453, 453)$ Mbps under proportionally fair scheduling.

In Fig. 9, we compare the proposed static clustering schemes with a different baseline. The baseline scheme is chosen to be a disjoint clustering scheme where each user is jointly served by the 4 BSs, i.e. 1 macro-BS and 3 pico-BSs, within its cell. Similar to Fig. 5, we first run the disjoint clustering scheme using WMMSE algorithm [29], then set the backhaul constraints for the proposed algorithms as the average backhaul consumptions evaluated from the disjoint clustering scheme. Since the macro-BS and the pico-BSs in each cell are associated with the same set of users, the average backhaul consumptions for the macro-BSs and for the pico-BSs are equal in the disjoint clustering scheme, i.e. $(C_{\text{macro}}, C_{\text{pico}}) = (453, 453)$ Mbps. In this case, the setting of $K_{\text{macro},max} = 50$, $K_{\text{pico},max} = 20$ and $\eta_1 = 20$ dB provides the best performance for the proposed static clustering scheme in Algorithm 3, while the setting of $\zeta_{\text{macro}} = 0$ dB, $\zeta_{\text{pico}} = 14$ dB and $\eta_2 = 12$ dB provides the best performance for the proposed static clustering scheme in Algorithm 4. As we can see from Fig. 9, since the cluster-edge users suffer from considerable inter-cluster interference, the disjoint clustering scheme has a substantial number of low-rate users, while the user-centric clustering schemes proposed in this paper can effectively improve the performance of the low-rate users. For example, for the 10th percentile user, both of the proposed static clustering schemes improve the rate over the disjoint clustering scheme by about 32%, while the proposed Algorithm 1 under dynamic clustering achieves around 57% improvement.

VI. CONCLUSION

This paper proposes an ℓ_0-norm formulation for the per-BS backhaul constraint in a downlink C-RAN system. By taking advantage of the ℓ_1-norm reweighting technique, the nonconvex per-BS backhaul constraint is approximated as a convex weighted per-BS power constraint. This approximation allows us to use a generalized WMMSE approach to solve the WSR maximization problem under two different

cases, depending on whether BS clustering is dynamic or static over different time-frequency slots. In the former case, we propose a joint user-centric clustering, user scheduling and beamforming design algorithm; in the latter case we fix the BS cluster for each user and jointly optimize the user scheduling and beamforming vectors. We also provide two static clustering formation algorithms, which can effectively take into account both the traffic load of each BS and the channel condition for each user. Simulation results show that with explicit per-BS backhaul constraints, the proposed dynamic clustering scheme is able to significantly improve the system performance over the naive clustering schemes, while the proposed static clustering schemes can already achieve a substantial portion of the performance gain.
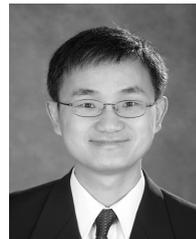
REFERENCES

[1] B. Dai and W. Yu, "Sparse beamforming design for network MIMO system with per-base-station backhaul constraints," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, pp. 294–298, Jun. 2014.
[2] P. Rost, C. J. Bernardos, A. De Domenico, M. Di Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wübben, "Cloud technologies for flexible 5G radio access networks," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 68–76, May 2014.
[3] D. Gesbert, S. Hanly, H. Huang, S. S. Shamai, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.
[4] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Downlink multi-cell processing with limited-backhaul capacity," *EURASIP J. Adv. Signal Process.*, vol. 2009, Feb. 2009, Art. ID 840814.
[5] 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.
[6] 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)*, Nov./Dec. 2008, pp. 1–6.
[7] 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.
[8] "Coordinated multi-point operation for LTE physical layer aspects," Tech. Rep. 3GPP TR 36.819 V11.0.0, 3rd Generation Partnership Project (3GPP), Sep. 2011. [Online]. Available: http://www.qtc.jp/3GPP/Specs/36819-b10.pdf
[9] F. Boccardi and H. Huang, "Limited downlink network coordination in cellular networks," in *Proc. IEEE Int. Symp. Personal, Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2007, pp. 1–5.
[10] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. J. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2983–2989, Jun. 2009.
[11] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
[12] 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.
[13] S. Shamai, O. Simeone, O. Somekh, and H. V. Poor, "Joint multi-cell processing for downlink channels with limited-capacity backhaul," in *Proc. IEEE Inf. Theory Appl. Workshop (ITA)*, Jan. 2008, pp. 345–349.
[14] Z. Jiang, S. Zhou, and Z. Niu, "Capacity bounds of downlink network MIMO systems with inter-cluster interference," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 4612–4617.
[15] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2911–2934, May 2012.

[16] R. Ran, C.-H. An, D. Kim, and V. Lau, "Capacity analysis of the clustered network MIMO with overlap," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2012, pp. 1942–1946.

[17] J. Zhao, T. Q. S. Quek, and Z. Lei, "Coordinated multipoint transmission with limited backhaul data transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2762–2775, Jun. 2013.

[18] Y. Shi, J. Zhang, and K. B. Letaief, "Group sparse beamforming for green Cloud-RAN," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2809–2823, May 2014.

[19] S. Luo, R. Zhang, and T. J. Lim. "Downlink and uplink energy minimization through user association and beamforming in Cloud RAN," 2014. [Online]. Available: http://arxiv.org/abs/1402.4238

[20] B. Dai and W. Yu, "Sparse beamforming for limited-backhaul network MIMO system via reweighted power minimization," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 1962–1967.

[21] 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.

[22] V. N. Ha, L. B. Le, and N.-D. Dao, "Energy-efficient coordinated transmission for Cloud-RANs: Algorithm design and trade-off," in *Proc. IEEE 48th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2014, pp. 1–6.

[23] S. Shi, M. Schubert, N. Vucic, and H. Boche, "MMSE optimization with per-base-station power constraints for network MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2008, pp. 4106–4110.

[24] H. Park, S.-H. Park, H.-B. Kong, and I. Lee, "Weighted sum MSE minimization under per-BS power constraint for network MIMO systems," *IEEE Commun. Lett.*, vol. 16, no. 3, pp. 360–363, Mar. 2012.

[25] M. Sarkiss and M. Kamoun, "Energy efficiency analysis of base station cooperation under limited backhaul capacity," in *Proc. IEEE 3rd Int. Conf. Commun. Netw. (ComNet)*, Mar. 2012, pp. 1–8.

[26] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3412–3427, Jul. 2013.

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

[28] C. T. K. 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, Dec. 2010.

[29] S. Kaviani, O. Simeone, W. A. 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, Jun. 2012.

[30] 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.

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

[32] S. Mehryar, A. Chowdhery, and W. Yu, "Dynamic cooperation link selection for network MIMO systems with limited backhaul capacity," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 4410–4415.

[33] P. Baracca, F. Boccardi, and V. Braun, "A dynamic joint clustering scheduling algorithm for downlink CoMP systems with limited CSI," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2012, pp. 830–834.

[34] J.-M. Moon and D.-H. Cho, "Inter-cluster interference management based on cell-clustering in network MIMO systems," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC)*, May 2011, pp. 1–6.

[35] Y. Cui, Q. Huang, and V. Lau, "Queue-aware dynamic clustering and power allocation for network MIMO systems via distributed stochastic learning," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1229–1238, Mar. 2011.

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

[37] W.-C. Liao, M. Hong, H. Farmanbar, X. Li, Z.-Q. Luo, and H. Zhang, "Min flow rate maximization for software defined radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1282–1294, Jun. 2014.

[38] S. Han, C. Yang, and A. F. Molisch, "Spectrum and energy efficient cooperative base station doze," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 285–296, Feb. 2014.

[39] M. Hong, Q. Li, and Y.-F. Liu. "Decomposition by successive convex approximation: A unifying approach for linear transceiver

[40] C. W. Tan, S. Friedland, and S. H. Low, "Spectrum management in multiuser cognitive wireless networks: Optimality and algorithm," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 421–430, Feb. 2011.

[41] C. W. Tan, M. Chiang, and R. Srikant, "Maximizing sum rate and minimizing MSE on multiuser downlink: Optimality, fast algorithms and equivalence via max-min SINR," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6127–6143, Dec. 2011.

[42] I.-W. Lai, L. Zheng, C.-H. Lee, and C. W. Tan, "Beamforming duality and algorithms for weighted sum rate maximization in cognitive radio networks," accepted in *IEEE J. Sel. Areas Commun.*, 2014.

[43] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, 2008.

[44] M. Grant and S. Boyd. *CVX: Matlab Software for Disciplined Convex Programming, Version 2.0 Beta*, Sep. 2013. [Online]. Available: http://cvxr.com/cvx

[45] Y. Ye, *Interior Point Algorithms: Theory and Analysis*. New York, NY, USA: Wiley, 1997.

[46] I. Guvenc, M.-R. Jeong, I. Demirdogen, B. Kecicioglu, and F. Watanabe, "Range expansion and inter-cell interference coordination (ICIC) for picocell networks," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Sep. 2011, pp. 1–6.

[47] M. M. Mohiuddin, V. Maheshwari, S. Veetil, K. Kuchi, G. V. V. Sharma, and S. Emami. "Performance limits of a cloud radio," 2013. [Online]. Available: http://arxiv.org/abs/1307.4733v1

design in heterogeneous networks," 2013. [Online]. Available: http://arxiv.org/abs/1210.1507

BINBIN DAI (S'12) received the B.E. degree in information science and engineering from Chien-Shiung Wu Honors College, Southeast University, Nanjing, China, in 2011, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2014, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering. His research interests include optimization, wireless communications, and signal processing.

WEI YU (S'97–M'02–SM'08–F'14) received the B.A.Sc. degree in computer engineering and mathematics from the University of Waterloo, Waterloo, ON, Canada, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1998 and 2002, respectively. Since 2002, he has been with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, where he is currently a Professor and the Canada Research Chair (Tier 1) in Information Theory and Wireless Communications. His main research interests include information theory, optimization, wireless communications, and broadband access networks.

Prof. Yu served as an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY (2010–2013), an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS (2009–2011) and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2004–2007), and a Guest Editor for a number of special issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and the EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING. He was a Technical Program Committee Co-Chair of the Communication Theory Symposium at the IEEE International Conference on Communications (ICC) in 2012, and the Technical Program Co-Chair of the IEEE Communication Theory Workshop in 2014. He was a member of the Signal Processing for Communications and Networking Technical Committee of the IEEE Signal Processing Society (2008–2013). He was a recipient of the IEEE ICC Best Paper Award in 2013, the IEEE Signal Processing Society Best Paper Award in 2008, the McCharles Prize for Early Career Research Distinction in 2008, the Early Career Teaching Award from the Faculty of Applied Science and Engineering, University of Toronto, in 2007, and the Early Researcher Award in Ontario in 2006. He is included in the list of Highly Cited Researchers (2014) by Thomson Reuters. Prof. Yu is a registered Professional Engineer in Ontario.

• • •