JOINT POWER ALLOCATION, BASE STATION ASSIGNMENT AND BEAMFORMER DESIGN FOR AN UPLINK SIMO HETEROGENEOUS NETWORK

Jingran Lin, Yubai Li, and Qicong Peng

Department of Communication Engineering
University of Electronic Science and Technology of China, Chengdu, Sichuan 611731

ABSTRACT

Consider the max-min problem for an uplink SIMO heterogeneous network, where the base stations (BS) are coordinated dynamically for joint reception under some backhaul overhead constraints. We formulate this problem in the perspective of joint power allocation, BS assignment and beamformer design, and develop an efficient algorithm based on alternating optimization. In particular, we transfer the joint BS assignment and beamformer design subproblem into a group LASSO problem by applying the alternating direction method of multipliers (ADMM). Consequently, the problem is solved in a partially distributed manner and in each iteration a simple closed-form solution is derived. Numerical simulations demonstrate the effectiveness and efficiency of the proposed algorithm.

***Index Terms*—** SIMO heterogeneous network, Max-min, Base station assignment, Beamformer, Alternating direction method

1. INTRODUCTION

In a heterogeneous network (HetNet), the user may be covered by multiple base stations (BS) simultaneously. BS assignment (i.e., associating users with BSs dynamically based on, e.g., the channel state information (CSI) and interference) is a promising approach to further improve the network performance in interference mitigation or congestion control [1, 2]. Motivated by this observation, we consider in this paper solving the max-min problem for the uplink transmission in a SIMO HetNet by the means of joint power allocation, BS assignment and beamformer design.

Actually, in conventional wireless networks, the max-min problem, since can guarantee fairness among users, has been intensively researched during the past decades. However, so far most works, whether analyzing the complexity status [3, 4] or designing the algorithms [4, 5, 6, 7, 8], are based on the assumption that the BS-user assignment is known and fixed. The situation does not change much as the data communication in a HetNet is concerned, where the BSs can be coordinated for joint transmission or reception. One popular cooperative strategy is the partial coordinated multiple points (CoMP) transmission [9, 10]. In this strategy, multiple BSs form a virtual BS and perform joint processing (JP) by sharing the user data via the backhaul links, while coordinated beamforming (CB) is performed among these virtual BSs to mitigate interference. Unfortunately, most related works are still based on fixed BS assignment. They either simply assume the BS-user association is known, or form the virtual BS (also called as the BS cluster) greedily by an exhaustive search procedure [9, 10, 11, 12].

A few recent works integrate BS assignment into the network optimization process for further performance improvement [13, 14, 15, 16, 17]. For example, [15] considers maximizing the sum rate for an uplink HetNet by joint BS assignment and beamformer design, under the constraint that each user can be served by only one BS. On the other hand, [16] and [17] extend the BS assignment problem into the partial CoMP downlink transmission. They both adopt the sum rate utility function but with different sparse penalty terms to balance the overall spectrum efficiency and, e.g., the size of each virtual BS [16], or the total number of active BSs [17]. WMMSE framework is used to solve these problems. In addition, [17] also consider balancing the transmit power and the active BSs number under some QoS constraints. This work further investigate the distributed implementation of the algorithm. It applies a two-block alternating direction method of multipliers (ADMM) to reformulate the QoS constraints and then design an efficient distributed algorithm.

[***Relation to prior work***] Different from the works maximizing the sum rate in the downlink transmission (e.g., [16, 17]), here we concentrate on the max-min problem for an uplink SIMO HetNet. Note although [15] also considers the uplink scenario, it adopts the sum rate utility function and assumes no BS clustering. In this paper, the BSs adaptively perform coordinated reception based on the instantaneous CSI and the constrained backhaul overhead. To control the backhaul overhead in the uplink scenario, we explore the sparsity of the receive beamformer, while most existing works (e.g., [16, 17]) concentrate on the sparsity of the transmit beamformer in downlink applications. Moreover, inspired by the work in [17], we consider the distributed implementation issues. Based on a different problem formulation, we apply the *two-block* ADMM and transfer the problem to a group LASSO problem. As a consequence, the whole problem is solved efficiently (with closed-form solutions) in a partially distributed manner.

2. SYSTEM MODEL

Consider the uplink transmission in a SIMO HetNet consisting of M users and K BSs. Each user is equipped with a single antenna while each BS with $T > 1$ antennas. For the sake of simplicity, we assume *both the channel vectors and the receive beamformers are real in this paper*. Let $\mathbf{h}_{km} \in \mathbb{R}^{T\times 1}$ denote the channel vector between user m and BS k, and $\mathbf{w}_{km} \in \mathbb{R}^{T\times 1}$ the receive beamformer used by BS k for user m, $k = 1, 2, ..., K$, $m = 1, 2, ..., M$. Define vectors $\bar{\mathbf{p}} = [\bar{p}_1, \bar{p}_2, ..., \bar{p}_M]^T \in \mathbb{R}^{M\times 1}$ and $\mathbf{p} = [p_1, p_2, ..., p_M]^T \in \mathbb{R}^{M\times 1}$ as the system power budget vector and the actual transmit power vector, with $\bar{p}_m$ and p_m the power budget and the actual transmit power of user m, $m = 1, 2, ..., M$, respectively. Let σ^2 denote the power of the additive white Gaussian noise at the BSs.

Assume the BSs perform the coordinated reception and all the BSs can potentially join the virtual BS for any user, then the coordinated receive beamformer for user m and the associated channel, denoted by $\mathbf{w}_m$ and $\mathbf{h}_m \in \mathbb{R}^{KT\times 1}$ respectively, can be defined as

$$\mathbf{w}_m = [\mathbf{w}_{1,m}^T, ..., \mathbf{w}_{K,m}^T]^T,\ \mathbf{h}_m = [\mathbf{h}_{1,m}^T, ..., \mathbf{h}_{K,m}^T]^T,\ \forall m. \quad (1)$$

Then the SINR of user m is calculated by

$$\mathrm{SINR}_m = \frac{p_m \mathbf{w}_m^T \mathbf{h}_m \mathbf{h}_m^T \mathbf{w}_m}{\mathbf{w}_m^T(\sigma^2 \mathbf{I} + \sum_{n \neq m} p_n \mathbf{h}_n \mathbf{h}_n^T)\mathbf{w}_m}, \forall m, \qquad (2)$$

To avoid heavy signaling overhead, we need to control the number of BSs in coordination. Note if BS k is not assigned to user m, the corresponding receive beamformer $\mathbf{w}_{km} = \mathbf{0}$. In other words, $\mathbf{w}_m$ is group sparse if only a small number of BSs are allowed to perform coordinated reception for user m. Obviously, solving $\mathbf{w}_m$ determines the BS assignment and the receive beamformer for user m simultaneously. Hence, the joint power allocation, BS assignment and beamformer design problem can be expressed as

$$\begin{aligned} \text{(P1)} \quad & \max_{\mathbf{w}_m,\, p_m} \min_m \ \mathrm{SINR}_m \\ & \text{s.t.} \quad \sum_{k=1}^K \|\mathbf{w}_{km}\|_2 \leq \beta, \ \mathbf{1}^T \mathbf{w}_m = 1, \\ & \qquad\ 0 \leq p_m \leq \bar{p}_m, \quad \forall m. \end{aligned}$$

where $\beta > 0$ is the threshold for the $l_{1,2}$-norm term $\sum_{k=1}^K \|\mathbf{w}_{km}\|_2$. This constraint is used to control the sparsity of $\mathbf{w}_m$.

Remark 1: A more popular way to address the sparse requirement for $\mathbf{w}_m$ is putting the $l_{1,2}$-norm term in the objective as a penalty term (see, e.g., [16, 17]). However, in this paper we keep it in the constraint region with a threshold of β to guarantee the convergence of the algorithm based on alternating optimization.

Remark 2: The constraint $\mathbf{1}^T \mathbf{w}_m = 1$ is used to avoid the trivial solution $\mathbf{w}_m = \mathbf{0}$, since we plan to apply the SOCP reformulation [7] to deal with the SINR constraints.

3. PROBLEM FORMULATION

In (P1), $\mathbf{p}$ and $\mathbf{w}_m$ are coupled within the SINR terms. A popular method to solve this kind of problems is alternating optimization, i.e., iteratively fixing one variable and solving another. In this framework, (P1) can be separated into two subproblems. Fixing $\mathbf{w}_m$ we get the power allocation (PA) subproblem; fixing $\mathbf{p}$ we get the joint BS assignment and beamformer design (BABF) subproblem.

We claim if the two subproblems can be globally solved in each iteration, the algorithm based on alternating optimization converges to a stationary solution (or a KKT point). The proof follows the same line as that in [4], but with some small modifications since we have a different model. It can be easily checked that if the $l_{1,2}$-norm term is in the objective, we will lose the degree of freedom of the Lagrangian multiplier and then the KKT condition cannot be satisfied. Due to the space limitation, we omit the detail of the proof here.

Since so far numerous centralized or distributed algorithms have been proposed for the PA subproblem [18], we focus on solving the BABF subproblem. Obviously, this subproblem can be further divided into M independent small problems of $\mathbf{w}_m$, represented as (P2) with γ an auxiliary variable. Then (P2) can be solved by a bisection procedure of feasibility checking.

$$\begin{aligned} \text{(P2)} \quad & \max_{\mathbf{w}_m, \gamma} \ \gamma \\ & \text{s.t.} \quad \mathrm{SINR}_m \geq \gamma, \\ & \qquad \sum_{k=1}^K \|\mathbf{w}_{km}\|_2 \leq \beta, \ \mathbf{1}^T \mathbf{w}_m = 1. \end{aligned}$$

Without $\mathbf{1}^T \mathbf{w}_m = 1$, we can apply the SOCP reformulation to transfer (P2) to a convex problem and check the feasibility of (P2) for γ directly. However, $\mathbf{1}^T \mathbf{w}_m = 1$ is sensitive to the sign of $\mathbf{w}_m$, so we need to pay attention to the sign of $\mathbf{h}_m^T \mathbf{w}_m$ in the SOCP reformulation. To handle the sign problem, we first define two subsets for $\mathbf{w}_m$, i.e., $\mathcal{A}_m^+ \triangleq \{\mathbf{w}_m | \ \mathbf{h}_m^T \mathbf{w}_m \geq 0\}$ and $\mathcal{A}_m^- \triangleq \{\mathbf{w}_m | \ \mathbf{h}_m^T \mathbf{w}_m < 0\}$, $\forall m$. Then we solve (P2) with fixed γ in $\mathcal{A}_m^+$ and $\mathcal{A}_m^-$ separately via the SOCP reformulation. If in any subset it is feasible, then (P2) is feasible for γ. Since the procedures solving (P2) in $\mathcal{A}_m^+$ and $\mathcal{A}_m^-$ are exactly the same, we focus on how to solve it in $\mathcal{A}_m^+$.

In the procedure of feasibility checking, we move the $l_{1,2}$-norm term to the objective to force a group sparse solution, i.e.,

$$\begin{aligned} \text{(P3)} \quad & \min_{\mathbf{w}_m} \ \sum_{k=1}^K \|\mathbf{w}_{km}\|_2 \\ & \text{s.t.} \quad \bar{\mathbf{h}}_m^T(p_m, \gamma)\mathbf{w}_m \geq \|\mathbf{Q}^T(\mathbf{p})\mathbf{w}_m\|_2, \ \mathbf{1}^T \mathbf{w}_m = 1. \end{aligned}$$

where $\bar{\mathbf{h}}_m(p_m, \gamma) = \sqrt{p_m(1+\gamma^{-1})}\mathbf{h}_m$ and $\mathbf{Q}(\mathbf{p})\mathbf{Q}^T(\mathbf{p}) = \mathbf{H}(\mathbf{p})$ with $\mathbf{H}(\mathbf{p}) = \sigma^2 \mathbf{I} + \sum_{n=1}^M p_n \mathbf{h}_n \mathbf{h}_n^T$.

Remark 3: Solving (P3) is equivalent to checking the feasibility of (P2) for γ in $\mathcal{A}_m^+$. When (P3) is feasible, we always get an optimal objective value minimizing $\sum_{k=1}^K \|\mathbf{w}_{km}\|_2$. Then we compare this value with β. If it does not exceed β, (P2) is feasible for γ. Otherwise, i.e., when (P3) is infeasible or feasible with an optimal objective value greater than β, we make a *double check* in $\mathcal{A}_m^-$.

Remark 4: Note without $\mathbf{1}^T \mathbf{w}_m = 1$, (P3) is feasible for any γ at $\mathbf{w}_m = \mathbf{0}$.

4. PARTIALLY DISTRIBUTED ALGORITHM

In practice, people prefer implementing the algorithm distributively, i.e., with some necessary data exchange, each BS can decide by itself whether to join the coordinated reception for a specific user or not.

The major obstacle to develop a distributed algorithm for (P3) is the $\{\mathbf{w}_{km}\}$ blocks are coupled within the constraints. To overcome it, we introduce two series of auxiliary variables $\mathbf{u}_m \in \mathbb{R}^{KT\times 1}$ and $\mathbf{v}_m \in \mathbb{R}^{(KT+1)\times 1}$, and define $\mathbf{V}_m(\mathbf{p}, \gamma) = [\bar{\mathbf{h}}_m(p_m, \gamma), \ \mathbf{Q}(\mathbf{p})] \in \mathbb{R}^{KT\times(KT+1)}$, $\forall m$. Then (P3) is equivalent to

$$\begin{aligned} \text{(P4)} \quad & \min_{\mathbf{w}_m, \mathbf{v}_m, \mathbf{u}_m} \ \sum_{k=1}^K \|\mathbf{w}_{km}\|_2 \\ & \text{s.t.} \quad \begin{bmatrix} \mathbf{w}_m \\ \mathbf{v}_m \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{V}_m^T(\mathbf{p}, \gamma) \end{bmatrix} \mathbf{u}_m, \\ & \qquad v_{m,1} \geq \|\mathbf{v}_{m,-1}\|_2, \ \mathbf{1}^T \mathbf{u}_m = 1. \end{aligned}$$

with $v_{m,1} = \bar{\mathbf{h}}_m^T(p_m, \gamma)\mathbf{u}_m$ the first element of $\mathbf{v}_m$, and $\mathbf{v}_{m,-1} = \mathbf{Q}^T(\mathbf{p})\mathbf{u}_m$ the remaining subvector of $\mathbf{v}_m$, or $\mathbf{v}_m = [v_{m,1}, \mathbf{v}_{m,-1}^T]^T$.

Utilizing the method of *augmented Lagrangian Minimization*, (P4) can be reformulated as

$$\begin{aligned} \text{(P5)} \quad & \min_{\mathbf{w}_m, \mathbf{v}_m, \mathbf{u}_m, \boldsymbol{\lambda}_m} \ L_{c_m}(\mathbf{w}_m, \mathbf{v}_m, \mathbf{u}_m, \boldsymbol{\lambda}_m) \\ & \text{s.t.} \qquad v_{m,1} \geq \|\mathbf{v}_{m,-1}\|_2, \ \mathbf{1}^T \mathbf{u}_m = 1. \end{aligned}$$

where c_m is the positive penalty parameter; $\boldsymbol{\lambda}_m \triangleq [\boldsymbol{\lambda}_{m,1}^T, \boldsymbol{\lambda}_{m,2}^T]^T$ with $\boldsymbol{\lambda}_{m,1} \in \mathbb{R}^{KT\times 1}$, $\boldsymbol{\lambda}_{m,2} \in \mathbb{R}^{(KT+1)\times 1}$ the Lagrangian multipliers corresponding to the constraints of $\mathbf{w}_m = \mathbf{u}_m$, $\mathbf{v}_m = \mathbf{V}_m^T \mathbf{u}_m$, respectively. For the sake of conciseness, from now on we simply use $\bar{\mathbf{h}}_m$, $\mathbf{Q}$ and $\mathbf{V}_m$ to denote $\bar{\mathbf{h}}_m(p_m, \gamma)$, $\mathbf{Q}(\mathbf{p})$ and $\mathbf{V}_m(\mathbf{p}, \gamma)$ when there is no ambiguity. The *augmented Lagrangian function* $L_{c_m}(\mathbf{w}_m, \mathbf{v}_m, \mathbf{u}_m, \boldsymbol{\lambda}_m)$ is defined as,

$$\begin{aligned} L_{c_m}(\mathbf{w}_m, \mathbf{v}_m, \mathbf{u}_m, \boldsymbol{\lambda}_m) &= \sum_{k=1}^K \|\mathbf{w}_{km}\|_2 \\ &+ [\boldsymbol{\lambda}_{m,1}^T, \boldsymbol{\lambda}_{m,2}^T] \begin{bmatrix} \mathbf{w}_m - \mathbf{u}_m \\ \mathbf{v}_m - \mathbf{V}_m^T \mathbf{u}_m \end{bmatrix} + \frac{c_m}{2} \left\| \begin{matrix} \mathbf{w}_m - \mathbf{u}_m \\ \mathbf{v}_m - \mathbf{V}_m^T \mathbf{u}_m \end{matrix} \right\|_2^2 \end{aligned}$$

Dividing $\mathbf{w}_m$, $\mathbf{v}_m$ and $\mathbf{u}_m$ into two blocks as $\{\mathbf{w}_m, \mathbf{v}_m\}$ and $\{\mathbf{u}_m\}$, we can apply the following *two-block* ADMM framework to solve (P5) efficiently [19].

$$\begin{cases} \begin{bmatrix} \mathbf{w}_m^{(l+1)} \\ \mathbf{v}_m^{(l+1)} \end{bmatrix} = \arg\min_{\{\mathbf{w}_m,\mathbf{v}_m\}} L_{c_m}(\mathbf{w}_m, \mathbf{v}_m, \mathbf{u}_m^{(l)}, \boldsymbol{\lambda}_m^{(l)}), \\ \mathbf{u}_m^{(l+1)} = \arg\min_{\{\mathbf{u}_m\}} L_{c_m}(\mathbf{w}_m^{(l+1)}, \mathbf{v}_m^{(l+1)}, \mathbf{u}_m, \boldsymbol{\lambda}_m^{(l)}), \\ \boldsymbol{\lambda}_m^{(l+1)} = \begin{bmatrix} \boldsymbol{\lambda}_{m,1}^{(l)} \\ \boldsymbol{\lambda}_{m,2}^{(l)} \end{bmatrix} + c_m \begin{bmatrix} \mathbf{w}_m^{(l+1)} - \mathbf{u}_m^{(l+1)} \\ \mathbf{v}_m^{(l+1)} - \mathbf{V}_m^T \mathbf{u}_m^{(l+1)} \end{bmatrix}, \end{cases}$$

with l the ADMM iteration index.

The main advantage of ADMM is separating (P5) into two simple convex problems of $\{\mathbf{w}_m, \mathbf{u}_m\}$ and $\mathbf{v}_m$. Both can be solved efficiently (sometimes even distributively). Actually, a simple closed-form solution can be derived for each problem.

4.1. Solution to the problem of $\{\mathbf{w}_m, \mathbf{v}_m\}$

Note the problem of $\{\mathbf{w}_m, \mathbf{v}_m\}$ can be further decomposed into two problems of $\mathbf{w}_m$ and $\mathbf{v}_m$, since $\mathbf{w}_m$ and $\mathbf{v}_m$ are separable in both objective and constraints.

The problem of $\mathbf{w}_m$ is a group LASSO problem.

$$\text{(P5}\mathbf{w}\text{)} \quad \min_{\mathbf{w}_m} \; L_{c_m}(\mathbf{w}_m, \mathbf{u}_m^{(l)}, \mathbf{v}_m^{(l)}, \boldsymbol{\lambda}_m^{(l)})$$

Since the $\{\mathbf{w}_{km}\}$ blocks are completely separated in the objective, (P5$\mathbf{w}$) can be divided into K independent problems of $\mathbf{w}_{km}$,

$$\min_{\mathbf{w}_{km}} \|\mathbf{w}_{km}\|_2 + \boldsymbol{\lambda}_{km,1}^{(l)T}(\mathbf{w}_{km} - \mathbf{u}_{km}^{(l)}) + \frac{c_m}{2}\|\mathbf{w}_{km} - \mathbf{u}_{km}^{(l)}\|_2^2$$

where $\boldsymbol{\lambda}_{km,1}$ and $\mathbf{u}_{km} \in \mathbb{R}^{T\times 1}$ are defined similarly as $\mathbf{w}_{km}$, denoting the k^{th} blocks of $\boldsymbol{\lambda}_{m,1}$ and $\mathbf{u}_m$ respectively. Using the first order optimality condition for the optimal solution $\mathbf{w}_{km}^{(l+1)}$, we have

$$-\boldsymbol{\lambda}_{km,1}^{(l)} - c_m(\mathbf{w}_{km}^{(l+1)} - \mathbf{u}_{km}^{(l)}) \in \partial\|\mathbf{w}_{km}^{(l+1)}\|_2, \tag{3}$$

with $\partial\|\mathbf{w}_{km}^{(l+1)}\|_2$ the subgradient of the function $\|\cdot\|_2$, defined as

$$\partial\|\mathbf{w}_{km}\|_2 = \begin{cases} \dfrac{\mathbf{w}_{km}}{\|\mathbf{w}_{km}\|_2}, & \mathbf{w}_{km} \neq \mathbf{0}, \\ \{\mathbf{g} \mid \|\mathbf{g}\|_2 \leq 1, \mathbf{g} \in \mathbb{R}^{T\times 1}\}, & \mathbf{w}_{km} = \mathbf{0}. \end{cases} \tag{4}$$

Inserting eq.(4) into eq.(3), we can easily obtain

$$\mathbf{w}_{km}^{(l+1)} = \begin{cases} \mathbf{0}, & \text{if } \|c_m\mathbf{u}_{km}^{(l)} - \boldsymbol{\lambda}_{km,1}^{(l)}\|_2 \leq 1, \\ \dfrac{c_m\mathbf{u}_{km}^{(l)} - \boldsymbol{\lambda}_{km,1}^{(l)}}{c_m + \delta_{km}}, & \text{otherwise}, \end{cases} \tag{5}$$

with $\delta_{km} = (\|\mathbf{w}_{km}^{(l+1)}\|_2)^{-1} = (\|c_m\mathbf{u}_{km}^{(l)} - \boldsymbol{\lambda}_{km,1}^{(l)}\|_2 - 1)^{-1}c_m$.

Remark 5: Due to the separable structure of (P5$\mathbf{w}$), $\mathbf{w}_{km}$ can be computed distributively in the BSs. That is, with the knowledge of the current $\mathbf{u}_{km}$, $\boldsymbol{\lambda}_{km,1}$ and c_m, each BS can easily make a decision for itself on whether to join the coordinated reception for user m or not. If the answer is positive, the receive beamformer will be determined simultaneously, i.e., we jointly optimize the BS assignment and the receive beamformer.

The problem of $\mathbf{v}_m$ is expressed as

$$\begin{aligned} \text{(P5}\mathbf{v}\text{)} \quad & \min_{\mathbf{v}_m} \; \boldsymbol{\lambda}_{m,2}^{(l)T}(\mathbf{v}_m - \mathbf{V}_m^T\mathbf{u}_m^{(l)}) + \frac{c_m}{2}\|\mathbf{v}_m - \mathbf{V}_m^T\mathbf{u}_m^{(l)}\|_2^2 \\ & \text{s.t.} \quad v_{m,1} \geq \|\mathbf{v}_{m,-1}\|_2. \end{aligned}$$

The first optimality conditions for $\mathbf{v}_m^{(l+1)}$ are listed as follows,

$$\begin{cases} \lambda_{m,2,1}^{(l)} + c_m(v_{m,1}^{(l+1)} - \bar{\mathbf{h}}_m^T\mathbf{u}_m^{(l)}) - \mu_m = 0, \\ -\boldsymbol{\lambda}_{m,2,-1}^{(l)} - c_m(\mathbf{v}_{m,-1}^{(l+1)} - \mathbf{Q}^T\mathbf{u}_m^{(l)}) \in \mu_m\partial\|\mathbf{v}_{m,-1}^{(l+1)}\|_2, \\ 0 \leq \mu_m \perp (v_{m,1}^{(l+1)} - \|\mathbf{v}_{m,-1}^{(l+1)}\|_2) \geq 0. \end{cases} \tag{6}$$

where μ_m is the Lagrangian multiplier for the constraint of $v_{m,1} \geq \|\mathbf{v}_{m,-1}\|_2$; $\lambda_{m,2,1}$ is the first element of $\boldsymbol{\lambda}_{m,2}$ and $\boldsymbol{\lambda}_{m,2,-1}$ is the remaining subvector of $\boldsymbol{\lambda}_{m,2}$, i.e., $\boldsymbol{\lambda}_{m,2} = [\lambda_{m,2,1}, \boldsymbol{\lambda}_{m,2,-1}^T]^T$; the expression $0 \leq a \perp b \geq 0$ indicates the KKT complementarity condition, i.e., $a \geq 0, b \geq 0$ and $a \times b = 0$.

Assuming $\mathbf{v}_{m,-1} \neq \mathbf{0}$, from eq.(6) we can directly get,

$$\begin{cases} v_{m,1}^{(l+1)} = c_m^{-1}(c_m\bar{\mathbf{h}}_m^T\mathbf{u}_m^{(l)} - \lambda_{m,2,1}^{(l)} + \mu_m), \\ \mathbf{v}_{m,-1}^{(l+1)} = (c_m + \mu_m\rho_m)^{-1}(c_m\mathbf{Q}^T\mathbf{u}_m^{(l)} - \boldsymbol{\lambda}_{m,2,-1}^{(l)}), \end{cases} \tag{7}$$

where $\rho_m \triangleq (\|\mathbf{v}_{m,-1}^{(l+1)}\|_2)^{-1} > 0$, and μ_m should be chosen properly such that the KKT complementarity condition is satisfied.

If $(v_{m,1}^{(l+1)} - \|\mathbf{v}_{m,-1}^{(l+1)}\|_2)|_{\mu_m=0} \geq 0$, or equivalently if

$$(c_m\bar{\mathbf{h}}_m^T\mathbf{u}_m^{(l)} - \lambda_{m,2,1}^{(l)}) \geq \|c_m\mathbf{Q}^T\mathbf{u}_m^{(l)} - \boldsymbol{\lambda}_{m,2,-1}^{(l)}\|_2, \tag{8}$$

then $\mu_m = 0$. Otherwise, we have $v_{m,1}^{(l+1)} = \|\mathbf{v}_{m,-1}^{(l+1)}\|_2$ when $\mu_m > 0$. In the case of

$$|c_m\bar{\mathbf{h}}_m^T\mathbf{u}_m^{(l)} - \lambda_{m,2,1}^{(l)}| < \|c_m\mathbf{Q}^T\mathbf{u}_m^{(l)} - \boldsymbol{\lambda}_{m,2,-1}^{(l)}\|_2, \tag{9}$$

combining $v_{m,1}^{(l+1)} = \|\mathbf{v}_{m,-1}^{(l+1)}\|_2$ and $\rho_m\|\mathbf{v}_{m,-1}^{(l+1)}\|_2 = 1$, we obtain

$$\rho_m = \frac{2c_m}{\|c_m\mathbf{Q}^T\mathbf{u}_m^{(l)} - \boldsymbol{\lambda}_{m,2,-1}^{(l)}\|_2 + (c_m\bar{\mathbf{h}}_m^T\mathbf{u}_m^{(l)} - \lambda_{m,2,1}^{(l)})}, \tag{10}$$

$$\mu_m = \frac{\|c_m\mathbf{Q}^T\mathbf{u}_m^{(l)} - \boldsymbol{\lambda}_{m,2,-1}^{(l)}\|_2 - (c_m\bar{\mathbf{h}}_m^T\mathbf{u}_m^{(l)} - \lambda_{m,2,1}^{(l)})}{2}. \tag{11}$$

The conditions in eq.(9) guarantee $\mu_m > 0$ and $\rho_m > 0$. Note there is another possibility besides eq.(8) and eq.(9), i.e.,

$$-(c_m\bar{\mathbf{h}}_m^T\mathbf{u}_m^{(l)} - \lambda_{m,2,1}^{(l)}) \geq \|c_m\mathbf{Q}^T\mathbf{u}_m^{(l)} - \boldsymbol{\lambda}_{m,2,-1}^{(l)}\|_2. \tag{12}$$

In this case, $\mathbf{v}_m^{(l+1)} = \mathbf{0}$ is an optimal solution, which can be easily proved by checking the KKT condition in eq.(6).

4.2. Solution to the problem of $\mathbf{u}_m$

The problem of $\mathbf{u}_m$ is also a very simple problem,

$$\begin{aligned} \text{(P5}\mathbf{u}\text{)} \; & \min_{\mathbf{u}_m} \; L_{c_m}(\mathbf{w}_m^{(l+1)}, \mathbf{v}_m^{(l+1)}, \mathbf{u}_m, \boldsymbol{\lambda}_m^{(l)}) \\ & \text{s.t.} \; \mathbf{1}^T\mathbf{u}_m = 1. \end{aligned}$$

The optimal solution $\mathbf{u}_m^{(l+1)}$ can be easily derived as

$$\begin{cases} \mathbf{u}_m^{(l+1)} = \mathbf{B}^{-1}(\mathbf{b} - \theta_m\mathbf{1}), \; \theta_m = \dfrac{\mathbf{1}^T\mathbf{B}^{-1}\mathbf{b} - 1}{\mathbf{1}^T\mathbf{B}^{-1}\mathbf{1}}, \\ \mathbf{B} \triangleq c_m(\mathbf{V}_m\mathbf{V}_m^T + \mathbf{I}), \\ \mathbf{b} \triangleq \mathbf{V}_m(c_m\mathbf{v}_m^{(l+1)} + \boldsymbol{\lambda}_{m,2}^{(l)}) + (c_m\mathbf{w}_m^{(l+1)} + \boldsymbol{\lambda}_{m,1}^{(l)}), \end{cases} \tag{13}$$

with θ_m the Lagrangian multiplier for $\mathbf{1}^T\mathbf{u}_m^{(l+1)} = 1$.

Then the algorithm is summarized in Tab.1, where we simply use the distributed algorithm in [18] to solve the PA subproblem.

Table 1. The Proposed Partially Distributed Algorithm

Initialization: Set system parameters and initial variable values;
Stage 1: Solve the PA subprob. by the distributed algo. in [18];
Stage 2: Solve the BABF subprob. by bisection;
  In each bisection iteration, for user m, ***Repeat***
    Update $\mathbf{w}_m$ and $\mathbf{v}_m$, as eq.(5) and (7);
    Update $\mathbf{u}_m$ as eq.(13),
    Update $\boldsymbol{\lambda}_m$ as in the ADMM framework;
  Until *converge*, compare $\sum_{k=1}^{K} \|\mathbf{w}_{km}\|_2$ with β;
    If ">", double check, ***If*** *still* ">", (P3) infeasible for γ;
    else, (P3) feasible for γ;
Stop: ***If*** *stop criterion satisfied*, stop alternating optimization;
    else go to **Stage 1**;
Debiasing: discard the sparse constraints and apply alternating optimization again, with $\mathbf{w}_m$ simply the MMSE beamformer;

Remark 6: The algorithm is "partially" distributed, since (P5$\mathbf{w}$) can be solved distributively, while (P5$\mathbf{v}$) and (P5$\mathbf{u}$) are still solved in a centralized manner (with simple closed-form solutions).

Remark 7: Actually, introducing only one auxiliary variable is enough to decouple the $\{\mathbf{w}_{km}\}$ blocks in (P3). However, in this case solving the remaining SOCP problem is still difficult since it is not easy to get a closed-form solution.

Remark 8: When the BS assignment converges, an additional debiasing operation can further improve the min user rate [17, 20]. It discards the sparse constraint and applies alternating optimization again, with the optimal $\mathbf{w}_m$ simply the MMSE beamformer.

5. NUMERICAL RESULTS AND CONCLUSIONS

Consider a single macro cell in a SIMO HetNet. The distance between the adjacent corners of the hexagonal cell is $d = 1000m$. There are $K = 10$ BSs and $M = 10$ users deployed randomly in the cell. Each BS is equipped with $T = 2$ antennas and each user with a single antenna. The elements of the channel $\mathbf{h}_{km}$ are generated according to the distribution $\mathcal{N}(0, \sigma_{km}^2)$, where the variance is given by $\sigma_{km}^2 = L_{km}(\frac{200}{d_{km}})^3$ with d_{km} the distance between user m and BS k, and L_{km} the shadowing effect satisfying $10\log_{10}(L_{km}) \sim \mathcal{N}(0, 64)$. We fix the 0dB environment noise power for all BSs (i.e., $\sigma_k^2 = \sigma^2 = 1,\ \forall k$), and let all the users have the same power budget of 30dB (i.e., $\bar{p}_m = \bar{p} = 1000,\ \forall m$).

Three algorithms are compared here. (1) Algo. PD, the proposed partially distributed algorithm; (2) Algo. CVX, the centralized algorithm which solves the BABF subproblem by the CVX solver; (3) Algo. RS, which randomly selects some BSs (the number is determined by Algo. PD) to form the virtual BS and then alternatively optimizes $\mathbf{p}$ and $\mathbf{w}_m$, with $\mathbf{w}_m$ the optimal MMSE beamformer. The following results are obtained from 100 simulation trials. Each corresponds to a random network configuration including the terminals (BSs and users) positions and the channel gains.

The performance comparison at different β is displayed in Fig.1. Firstly, as expected, β can control the size of the virtual BSs. More BSs tend to join the cooperative reception as β increases. Consequently, the system achieves a higher max-min user rate at the cost of heavy backhaul overhead. Secondly, the debiasing operation apparently improves the max-min user rate, because it uses the optimal MMSE receive beamformer after the BS assignment converges. Thirdly, Algo. PD and Algo. CVX achieve very close performance. It has been observed Algo. PD and Algo. CVX may converge to different stationary solutions. However, it's interesting that Algo. PD usually gets a sparser solution than Algo. CVX. Lastly, Algo. PD outperforms Algo. RS which randomly associates users with BSs.

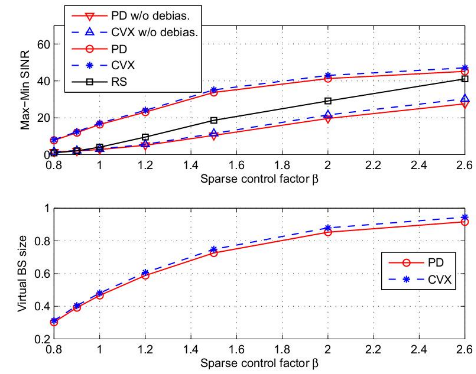

Fig. 1. Performance comparison v.s. β.

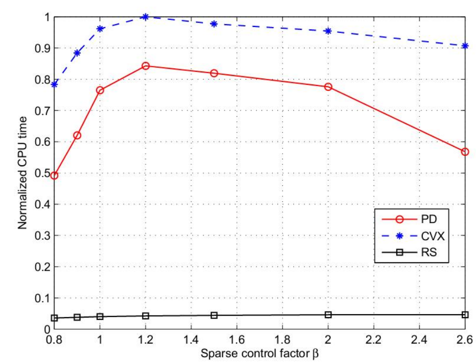

Fig. 2. CPU time comparison v.s. β.

The average CPU time comparison at different β is shown in Fig.2. In order to simulate the distributed implementation, we record the CPU time for each step and divide the CPU time of the steps distributively implemented by the number of terminals. Then summing up these results produces the final CUP time in partially distributed implementation. Note the results in Fig.2 are normalized by the largest CPU time in the simulations. As expected, Algo. PD outperforms Algo. CVX substantially in terms of normalized CPU time. This is reasonable since in each iteration the BABF subproblem is solved efficiently in closed-form by Algo. PD and about one third of the tasks can be implemented distributively. Note Algo. RS has a very low CPU time because there is no BS assignment procedure and the optimal receive beamformer is simply the closed-form MMSE beamformer. It implies the CPU time of the debiasing operation is neglectable in Algo. PD.

In summary, the proposed partially distributed algorithm can effectively control the backhaul overhead in the coordinated reception for an uplink SIMO HetNet. Moreover, the ADMM reformulation apparently improves the algorithm's efficiency, since it can be implemented in a partially distributed manner and in each step the problem is solved in closed-form. Some interesting future directions of this work are under investigation, e.g., how to extend this work into the case of *complex* channels; and how to integrate the BS activation problem into this work, i.e., we need to control not only the size of each virtual BS, but also the total number of active BSs.

6. REFERENCES

[1] Y. Yu, R.Q. Hu, C.S. Bontu, and Z. Cai, "Mobile association and load balancing in a cooperative relay cellular network," *IEEE Communications Magazine*, vol. 49, no. 5, pp. 83–89, May 2011.

[2] R. Madan, J. Borran, A. Sampath, and N. Bhushan, "Cell association and interference coordination in heterogeneous LTE-A cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1479–1489, Dec. 2010.

[3] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[4] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Max-min fairness linear transceiver design for a multi-user MIMO interference channel," *IEEE Trans. Signal Processing*, vol. 61, no. 9, pp. 2413–2423, May 2013.

[5] D. Gerlach and A. Paulraj, "Adaptive transmitting antenna methods for multipath environments," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 28, no. 9, pp. 425–429, San Francisco, Nov. 1994.

[6] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. 37th Allerton Conf. Commun. Con-trol Comput*, pp. 987–996, Allerton, Sept. 1999,

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

[8] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.

[9] A. Papadogiannis, D. Gesbert, and E. Hardouin, "A dynamic clustering approach in wireless networks with multi-cell cooperative processing," in *IEEE International Conference on Communications*, pp. 4033–4037, Beijing, May 2008.

[10] S.-J. Kim, S. Jain, G.B. Giannalkis, "Backhaul-constrainted multi-cell cooperation using compressive sensing and spectral clustering," in *Proc. of IEEE 13th International Workshop on Signal Processing Advances for Wireless Communications*, pp. 165–69, Cesme, Jun. 2012.

[11] 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. 1910C-1921, April 2009.

[12] 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.

[13] L. Gao, X. Wang, G. Sun, and Y. Xu, "A game approach for cell selection and resource allocation in heterogeneous wireless networks," in *Proc. of IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pp. 530–538, Salt Lake City, June 2011.

[14] A.S. Cacciapuoti, I.F. Akyildiz, and L. Paura, "Correlation-aware user selection for cooperative spectrum sensing in cognitive radio ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 2, pp. 297–306, Feb. 2012.

[15] M. Hong and Z.-Q. Luo, "Distributed linear precoder optimization and base station selection for an uplink heterogeneous network," *IEEE Trans. Signal Processing*, vol. 61, no. 12, pp. 3214–3228, June 2013.

[16] M. Hong, R. Sun, H. Baligh, and Z.-Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 226–240, Feb. 2013.

[17] W.-C. Liao, M. Hong, Y.-F. Liu, and Z.-Q. Luo, "Base station activation and linear transceiver design for optimal resource management in heterogeneous networks," Sept. 16, 2013 [online]. Available: http://arxiv.org/abs/1309.4138.

[18] C. Tan, M. Chiang, and R. Srikant, "Fast algorithms and performance bounds for sum rate maximization in wireless networks," in *the Proceedings of the IEEE INFOCOM*, pp. 1350–1358, Rio de Janeiro, April 2009.

[19] M. Fukushima, "Application of the alternating direction method of multipliers to separable convex programming problems," *Computational Optimization and Applications*, vol. 1, no. 1, pp. 93–111, Oct. 1992.

[20] S.J. Wright, R.D. Nowak, and M.A.T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479-C2493, July 2009.