9-1-2012

Mobile Association for Wireless Heterogeneous Networks with Cooperative Relays: Optimal Framework and Implementation Schemes

Q. Li

Rose Qingyang Hu
Utah State University

G. Wu

Optimal Intra-Cell Cooperation With Precoding in Wireless Heterogeneous Networks

Yiran Xu, Rose Qingyang Hu, Qian (Clara) Li
Department of Electrical and Computer Engineering
Utah State University, Logan, UT, USA
Emails: yiran.xu@ieee.org, rosehu@ieee.org, liqian@ieee.org

Yi Qian
Department of Computer and Electronics Engineering
University of Nebraska-Lincoln, NE USA
Email: yqian@ieee.org

Abstract—**Wireless heterogeneous networks have emerged as a new paradigm to meet the fast increasing wireless capacity and coverage demands. Coordinated Multipoint Processing (CoMP) and Precoding are two promising techniques to further improve the network capacity and spectral efficiency. This paper presents an optimal intra-cell CoMP resource allocation scheme in a wireless heterogeneous network and explores the Tomlinson-Harashima Precoding (THP) in the physical layer to reduce the inter-user interference. The objective is to maximize the aggregate proportional rates in the system. We derive an asymptotically optimal solution for resource allocation by using a gradient descent based scheduling and KKT conditions for optimality. Simulation results demonstrate the system proportional fairness capacity gain of proposed resource allocation scheme and this resource management framework provides a guideline for future radio resource management in wireless heterogeneous networks.**

I. Introduction

Driven by the proliferation of wireless devices and applications, future wireless systems are required to support various applications at a much higher capacity and a higher spectral efficiency. Heterogeneous network is a key technology adopted in 3GPP LTE-Advanced to realize a higher spectrum efficiency. Deploying base stations (BSs) of diverse sizes and capacities, heterogeneous network can provide end users various accessing capabilities and coverage needs [1]. In contrast to a macro-BS (MBS), a pico-BS (PBS) has a smaller transmit power and thus a smaller footprint for capacity expansion or coverage enhancement in a hotspot or in an indoor environment. PBSs overlay with MBSs and usually have full radio resource management capabilities. The coexistence of PBSs and MBSs forms a wireless heterogeneous network. Without proper mobile association, interference management and resource allocation schemes, the PBSs may suffer severe interference from neighboring MBSs and thus their resources may not be effectively utilized [2].

In this paper, we consider an orthogonal-frequency-division-multiplexing access (OFDMA) based wireless heterogeneous network. A user equipment (UE) can be either associated with a MBS or a PBS based on different association criteria. In a range-expansion based mobile association scheme [3], a UE may attach to the BS that is closest. However, by doing so, UEs located in the cell edge of the PBSs could be exposed to the strong interference from nearby high power MBSs, which will need advanced interference management schemes or otherwise may result in the degradation of user experience and spectral efficiency. Recently, intra-cell Coordinated Multipoint Processing (CoMP) has been considered as an important technique to improve the performance for UEs at cell edge. Its application in a wireless heterogeneous network resource allocation has been addressed in [4]- [9]. The MBSs and PBSs can coordinate on scheduling and data transmission among adjacent cells to improve the coverage and cell edge signal quality [2], [10]. In addition, precoding applies an appropriate weight to the signal emitted from each of the transmit antennas so that the signal power is maximized at the receivers. Precoding can be combined with CoMP technique to further improve the cell edge user performance. In our proposed resource allocation framework, we employ Tomlinson-Harashima Precoding [11] so that a MBS and a PBS can serve two UEs on the same resource simultaneously and achieve more efficient utilization of radio resources.

The remainder of the paper is organized as follows. In Section II, we introduce the wireless heterogeneous network model and CoMP + THP precoder design. We then formulate an optimal CoMP + precoding based scheduling problem to maximize the aggregate proportional system throughput in Section III. In Section IV, we develop an asymptotically optimal solution to the problem we formulate in Section III. Numerical analysis and discussions are presented in Section V. The paper is concluded in Section VI.

II. Network Model and Precoder Design

In this paper we consider downlink communication in a wireless heterogeneous network shown in Fig.1. Each cell is divided into several sectors and each sector is equipped with one MBS and N_r PBSs. We use $\mathcal{N}_{0,i}$ to represent the MBS in sector i and $\mathcal{N}_{j,i}$ to represent the jth PBS in sector i. In total N_u UEs are uniformly distributed in the network which has N_c sectors. We assume all the MBSs have the same transmit power P_m, which is higher than the PBS's transmit power P_p. UEs falling into a MBS's coverage range are associated with the MBS and are denoted as macro-UE (MUE) (Fig.2-a). The second type, denoted as pico-UE (PUE), locates closely to a PBS and is associated with a PBS. The third type, denoted as cooperative UE (CUE), locates at the extended coverage area of a PBS and is associated with the PBS. It receives cooperative transmissions from the MBS and the PBS (Fig.2-b,c) in order to mitigate the interference from the MBS that causes the strongest interference. We divide the total frequency band into F resource blocks (RBs) and each UE can be assigned with an integer number of RBs at time t. Each node (MBS or PBS) is equipped with only one antenna. However, for the CUEs that receive cooperative transmissions or joint processing from both MBS and PBS, we can view the two transmitting nodes as two antennas, thus forming a network MIMO for the CUEs. In order to maximize the sum throughput in a network MIMO system, we allow two CUEs to simultaneously receive from the same MBS and PBS (Fig.2-c) by applying THP precoding algorithm [11], which can eliminate co-user interference between the two CUEs that receive from the same nodes and that use the same radio resource. We first formulate a 2×2 channel matrix $\mathbf{H}$

$$\mathbf{H} = \begin{bmatrix} \sqrt{P_m} h_{\mathrm{MUE1}} & \sqrt{P_p} h_{\mathrm{PUE1}} \\ \sqrt{P_m} h_{\mathrm{MUE2}} & \sqrt{P_p} h_{\mathrm{PUE2}} \end{bmatrix} \quad (1)$$

Based on [11], the precoding matrix is expressed as

$$\mathbf{W} = \mathbf{F}\mathbf{B}^{-\mathbf{1}}\mathbf{J}. \quad (2)$$

Matrices $\mathbf{F}$, $\mathbf{B}$ and $\mathbf{J}$ are computed in the following,

$$\begin{aligned} &\mathbf{H}^* = \mathbf{Q}\mathbf{R}^*; \quad \mathbf{B} = \mathbf{G}\mathbf{R}; \quad \mathbf{F} = \mathbf{Q}; \\ &\mathbf{G} = \mathrm{diag}[\frac{1}{|r_{11}|} \quad \frac{1}{|r_{22}|}]; \quad \mathbf{J} = \mathrm{diag}[\frac{r_{11}}{|r_{11}|} \quad \frac{r_{22}}{|r_{22}|}], \end{aligned} \quad (3)$$

Where $\mathbf{Q}$ is a unitary matrix or semi-unitary matrix and $\mathbf{R}$ is a lower triangle matrix. r_{kk} is the diagonal element of $\mathbf{R}$ in the kth row and $\mathbf{J}$ is a local phase adjustment matrix which is used to combine the channel gains coherently at the UE. The received signal $\mathbf{y} = [y_1 \quad y_2]^T$ at the CUE is

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{n} = \mathbf{H}\mathbf{F}\mathbf{B}^{-\mathbf{1}}\mathbf{J}\mathbf{x} + \mathbf{n} \\ &= \begin{bmatrix} r_{11} & 0 \\ 0 & r_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}. \end{aligned} \quad (4)$$

The corresponding downlink received signal-to-interference-noise-ratio (SINR) for MUE, PUE and CUE can be evaluated as follows.

$$\mathrm{SINR}^f_{k,0,i}(t) = \frac{P_m |h^f_{k,0,i}(t)|^2}{N_0 + \sum\limits_{i'=1, i' \neq i}^{N_c} |h^f_{k,0,i'}(t)|^2 P_m + \sum\limits_{i'=1}^{N_c} \sum\limits_{j'=1}^{N_r} |h^f_{k,j',i'}(t)|^2 P_p} \quad (5)$$

$$\mathrm{SINR}^f_{k,j,i}(t) = \frac{P_p |h^f_{k,j,i}(t)|^2}{N_0 + \sum\limits_{\substack{i'=1 \\ (i',j') \neq (i,j)}}^{N_c} \sum\limits_{j'=1}^{N_r} |h^f_{k,j',i'}(t)|^2 P_p + \sum\limits_{i'=1}^{N_c} |h^f_{k,0,i'}(t)|^2 P_m} \quad (6)$$

$$\mathrm{SINR}^{c,f}_{k,j,i}(t) = \frac{|r^f_{11(22)}(t)|^2}{N_0 + \sum\limits_{\substack{i'=1 \\ (i',j') \neq (i,j)}}^{N_c} \sum\limits_{j'=1}^{N_r} |h^f_{k,j',i'}(t)|^2 P_p + \sum\limits_{i'=1, i' \neq i}^{N_c} |h^f_{k,0,i'}(t)|^2 P_m} \quad (7)$$

$$i = 1, \cdots, N_c; \; j = 1, \cdots, N_r; \; k = 1, \cdots, N_u.$$

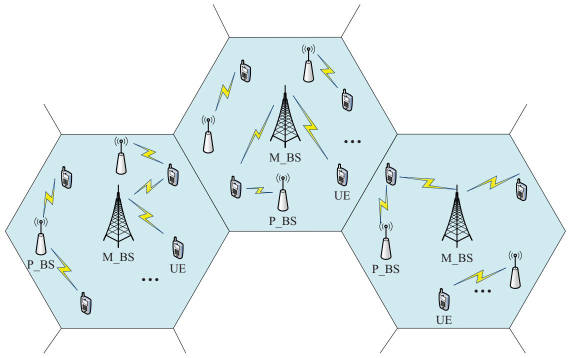

Fig. 1. Heterogeneous Networks Model.

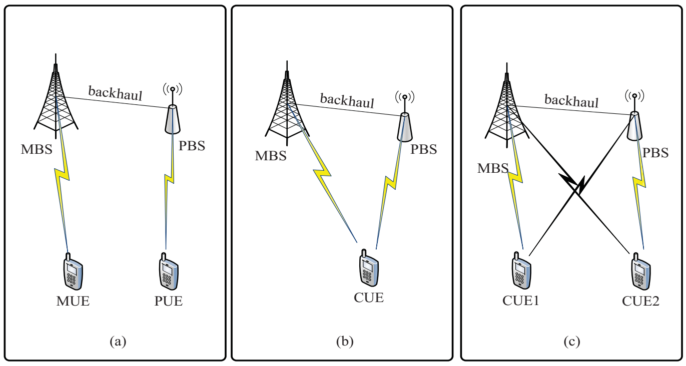

Fig. 2. Service Modes: (a) No CoMP; (b) CoMP without precoding; (c) CoMP with precoding

Here, $h^f_{k,0,i}(t)$ is the channel gain of the fth RB at time t between node $\mathcal{N}_{0,i}$ and UE k, and $h^f_{k,j,i}(t)$ is the channel gain of the fth RB at time t between node $\mathcal{N}_{j,i}$ and UE k. They both include long-term path loss, shadowing and short term fading due to multipath and mobility. r^f_{11} is the equivalent channel gain of the fth RB between node $\mathcal{N}_{0,i}$ and CUE1. r^f_{22} is the equivalent channel gain of the fth RB between node $\mathcal{N}_{j,i}$ and CUE2. N_0 is the variance of the additive white Gaussian noise. Given SINR, the unit achievable data rate in terms of bit/s/Hz for each UE can be calculated using Shannon formula.

$$\begin{aligned} &R^b_a = \log(1 + \mathrm{SINR}^b_a), \\ &\text{for } a = ((k,0,i),(k,j,i)) \text{ and } b = (f,(c,f)). \end{aligned} \quad (8)$$

III. Problem Formulation

Our objective is to optimize the network long-term spectrum efficiency and service fairness. Towards that end, we need to 1) properly decide the association for each UE, and 2) properly allocate RBs to the UEs. In order to expand the PBS's coverage range so that each PBS's resource can serve more UEs, we choose a range-expansion based mobile association scheme [3]. The kth UE will attach to the node $\mathcal{N}_{(i^*,j^*)}$ based on the following criterion

$$(i^*, j^*)_k = \arg \max_{i\in\{1,\cdots,N_c\}, j\in\{0,1,\cdots,N_r\}} (|h_{k,j,i}|^2/\theta_{i,j}), \quad (9)$$

where $\theta_{i,0} = 1$ and $1 < \theta_{i,j} < (P_m/P_p)$, for $j > 0$. We denote $x_{k,0,i}$ as the decision variable to indicate the association status between the kth UE and $\mathcal{N}_{0,i}$. Specifically,

$$x_{k,0,i} = \begin{cases} 1 & \text{if } k\text{th UE is associated with } \mathcal{N}_{0,i} \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

$x_{k,j,i}$ is similarly defined for UEs associated with PBSs. Each UE can only be associated with one node, i.e., $\sum_{i=1}^{N_c}\sum_{j=0}^{N_r} x_{k,j,i} \le 1, \ \forall k$. Furthermore, $x^{c,f}_{k,j,i}(t)$ is used to denote if CoMP is used or not at each scheduling cycle t. $x^{c,f}_{k,j,i}(t) = 1$ indicates that UE k is jointly served on RB f by $\mathcal{N}_{j,i}$ and $\mathcal{N}_{0,i}$ while $x^{c,f}_{k,j,i}(t) = 0$ indicates UE k receives transmission only from $\mathcal{N}_{j,i}$ on RB f. Unlike $x_{k,j,i}$ or $x_{k,0,i}$, which are decided during the mobile association stage, $x^{c,f}_{k,j,i}(t)$ is decided at t based on the instantaneous channel state. We denote $\mathcal{K}_{0,i}$ as the set of MUEs associated with $\mathcal{N}_{0,i}$, $\mathcal{K}_{j,i}$ as the set of PUEs associated with $\mathcal{N}_{j,i}$. At t, we also decide the set of CUEs $\mathcal{K}^{c,f}_{j,i}(t)$ that are associated with $\mathcal{N}_{j,i}$ but are the candidates for joint processing by $\mathcal{N}_{j,i}$ and $\mathcal{N}_{0,i}$ on RB f.

$$\mathcal{K}^{c,f}_{j,i}(t) = \{k \in \mathcal{K}_{j,i} | \mathrm{SINR}^f_{k,j,i}(t) < \alpha\}, \quad (11)$$

where α is the SINR threshold that decides the CoMP set.

In order to formulate the scheduling problem, we introduce the following variables. $n^f_{k,0,i}(t) = 1(0)$ denotes that the fth RB is (is not) assigned to the kth MUE by $\mathcal{N}_{0,i}$ at time t. $n^f_{k,j,i}(t) = 1(0)$ denotes that the fth RB is (is not) assigned to the kth PUE by $\mathcal{N}_{j,i}$ at time t, $n^{c,f}_{k,j,i}(t) = 1(0)$ denotes that the fth RB is (is not) assigned to the kth CUE served by $\mathcal{N}_{j,i}$ and $\mathcal{N}_{0,i}$ at time t. Let $\mathbf{n}(t) = \{n^f_{k,0,i}(t), n^f_{k,j,i}(t), n^{c,f}_{k,j,i}(t)\}$.

We use proportional fairness as the performance metric to ensure a good tradeoff between spectrum efficiency and fairness. The optimization problem with a long-term proportional fair resource allocation is thus formulated as

$$[\mathbf{P}_1] \quad U(\mathbf{R}(t)) := \max_{\mathbf{n}(t)} \sum_k \log(R_k(t)) \quad (12)$$

subject to

$$\sum_{k=1}^{N_u} x_{k,0,i} n^f_{k,0,i}(t) + \sum_{j=1}^{N_r}\sum_{k=1}^{N_u} x_{k,j,i} x^{c,f}_{k,j,i}(t) n^{c,f}_{k,j,i}(t) \le 1$$
$$\text{for } i = 1,\cdots,N_c, f = 1,\cdots,F \quad (13)$$

$$\sum_{k=1}^{N_u} x_{k,j,i}(1 - x^{c,f}_{k,j,i}(t)) n^f_{k,j,i}(t) + \sum_{k=1}^{N_u} x_{k,j,i} x^{c,f}_{k,j,i}(t) n^{c,f}_{k,j,i}(t) \le 1$$
$$\text{for } i = 1,\cdots,N_c, \ j = 1,\cdots,N_r, f = 1,\cdots,F \quad (14)$$

$$n^f_{k,j,i}(t) = 0 \text{ or } 1 \quad \text{and} \quad n^{c,f}_{k,j,i}(t) = 0 \text{ or } 1, \ \forall i,j,k,f, \quad (15)$$

where

$$R_k(t) = \frac{1}{T_c}\sum_{\tau=t-T_c+1}^{t} S_k(\tau), \quad (16)$$

T_c is the size of time window for moving average, and S_k is the moving average system throughput, which is expressed as

$$\begin{aligned} S_k(\tau) &= \sum_{f=1}^{F}\sum_{i=1}^{N_c}\sum_{j=0}^{N_r} \big(x_{k,j,i}(1 - x^{c,f}_{k,j,i}(\tau)) R^f_{k,j,i}(\tau) n^f_{k,j,i}(\tau) \\ &+ x_{k,j,i} x^{c,f}_{k,j,i}(\tau) R^{c,f}_{k,j,i}(\tau) n^{c,f}_{k,j,i}(\tau)\big), \end{aligned} \quad (17)$$

where we set $x^{c,f}_{k,0,i}(\tau) = 0$ for notational consistency.

By solving $n^f_{k,j,i}(t)$ and $n^{c,f}_{k,j,i}(t)$ in $\mathbf{P}_1$, we can find the allocated UE for each RB. The problem is a 0-1 Knapsack problem and is NP-hard. We first relax the integers $n^f_{k,j,i}(t)$ and $n^{c,f}_{k,j,i}(t)$ into real numbers, i.e., $n^f_{k,j,i}(t) \in [0,1]$ and $n^{c,f}_{k,j,i}(t) \in [0,1]$. By doing so, (13) is the resource constraint for each RB at the MBSs. The first term in (13) represents the portion of RB f used by MUEs, and the second term in (13) computes the portion of the RB f used at the MBS to serve the CUEs. Similarly, (14) gives the resource constraint for each RB at the PBS. The first term represents the portion of RB f used by PUEs with no CoMP while the second term represents the portion of RB f used by the PUEs with CoMP, i.e., CUEs. As a multicarrier proportional fair scheduling problem, it is hard to find the optimal solution of $\mathbf{P}_1$ directly [12]. Considering practical implementation, we apply the gradient descent based scheduling algorithm in [13], which proved that the gradient descent based scheduling algorithm asymptotically converges to the optimal solution. In the next section, based on the gradient descent based scheduling algorithm, we show how to optimally allocate resources in such a heterogeneous network.

IV. An Asymptotically Optimal Radio Resource Scheduling Scheme

Using the gradient-based scheduling framework, the system parameters are chosen to maximize the drift of the objective function at each subframe, given as

$$\begin{aligned} &U(\mathbf{R}(t+1)) - U(\mathbf{R}(t)) \\ &= \sum_{k=1}^{N_u}\Big(\log\Big(R_k(t) + \epsilon\big(S_k(t+1) - S_k(t-T_c+1)\big)\Big) \\ &- \log\big(R_k(t)\big)\Big) = \sum_{k=1}^{N_u}\frac{1}{R_k(t)} S_k(t+1)\epsilon \\ &- \sum_{k=1}^{N_u}\frac{1}{R_k(t)} S_k(t-T_c+1)\epsilon + O(\epsilon^2), \end{aligned} \quad (18)$$

where $\epsilon = 1/T_c$ and the second equality is obtained using first order Taylor expansion. Since only the first term in (18) depends on future decisions and constraints (13)-(15) are set on a per RB basis, we can formulate the gradient-based scheduling problem for each RB f as $[\mathbf{P}_2]$:

$$\begin{aligned} &\max_{\mathbf{n}(t)} \sum_{k=1}^{N_u}\sum_{i=1}^{N_c}\sum_{j=0}^{N_r} [x_{k,j,i}(1 - x^{c,f}_{k,j,i}(t)) R^f_{k,j,i}(t) n^f_{k,j,i}(t) \\ &+ (x_{k,j,i} x^{c,f}_{k,j,i}(t) R^{c,f}_{k,j,i}(t) n^{c,f}_{k,j,i}(t)]/R_k(t-1) \end{aligned} \quad (19)$$

subject to (13)-(15). By gradient-based scheduling, multi-carrier scheduling problem $\mathbf{P}_1$ can be decomposed into multiple single-carrier scheduling problem $\mathbf{P}_2$. $\mathbf{P}_2$ only consists of linear objective function and linear constraints with variables $n^f_{k,j,i}(t)$ and $n^{c,f}_{k,j,i}(t)$. Thus it is a convex problem.

A. Optimal Resource Scheduling Scheme by Solving the KKT Conditions

For convex optimization problems, the KKT conditions are necessary and sufficient for optimality. Optimal solution for the convex optimization problem $\mathbf{P}_2$ can thus be solved from the KKT conditions. By introducing Lagrangian multipliers $\lambda^f_i(t)$, $\mu^f_{j,i}(t)$, $\nu^f_{k,j,i}(t)$ and $\nu^{c,f}_{k,j,i}(t)$, the Lagrangian function of $\mathbf{P}_2$ is shown in (20).

Our goal is to find at time t, for each RB f in each sector i, the optimal MUE index k^*_0, the optimal PUE index $k^*_{1,j}$, and the optimal CUE index k^*_2 for $\mathcal{N}_{0,i}$, $\mathcal{N}_{j,i}$, and their joint processing, respectively. It also needs to decide the corresponding optimal $n^{f*}_{k,j,i}(t)$ and $n^{c,f*}_{k,j,i}(t)$ values. Towards this end, we first formulate the KKT conditions stated in [9] and [14], then we solve the optimal Lagrangian multipliers as follows.

$$\lambda^{f*}_i(t) = \max\{\lambda^f_{i,A}(t), \lambda^f_{i,j*,B}(t)\}, \tag{21}$$

$$\mu^{f*}_{j,i}(t) = \max_{k_{1,j}\in\mathcal{K}_{j,i}} \frac{R^f_{k_{1,j},j,i}(t)}{R_{k_{1,j}}(t-1)} \tag{22}$$

$$\nu^{f*}_{k_0,j,i}(t) = \lambda^{f*}_i(t) - \frac{R^f_{k_0,0,i}(t)}{R_{k_0}(t-1)} \quad \text{for } k_0 \in \mathcal{K}_{0,i}, \tag{23}$$

$$\nu^{f*}_{k_{1,j},j,i}(t) = \mu^{f*}_{j,i}(t) - \frac{R^f_{k_{1,j},j,i}(t)}{R_{k_{1,j}}(t-1)} \quad \text{for } k_{1,j} \in \mathcal{K}_{j,i}, \tag{24}$$

$$\nu^{c,f*}_{k_2,j,i}(t) = \lambda^{f*}_i(t) - \Big[\Big(\sum_{k_2\in\mathcal{Q}}\frac{R^{c,f}_{k_2,j,i}(t)}{R_{k_2}(t-1)}\Big) - \max_{k_{1,j}\in\mathcal{K}_{j,i}}\frac{R^f_{k_{1,j},j,i}(t)}{R_{k_{1,j}}(t-1)}\Big] \quad \text{for } k_2 \in \mathcal{K}^{c,f}_{j,i}(t), \tag{25}$$

where

$$\lambda^f_{i,A}(t) = \max_{k_0\in\mathcal{K}_{0,i}} \frac{R^f_{k_0,0,i}(t)}{R_{k_0}(t-1)}, \tag{26}$$

$$\lambda^f_{i,j*,B}(t) = \max_{j\in\mathcal{J}}\Big(\max_{\mathcal{Q}\subseteq\mathcal{K}^{c,f}_{j,i}}\sum_{k_2\in\mathcal{Q}}\frac{R^{c,f}_{k_2,j,i}(t)}{R_{k_2}(t-1)} - \max_{k_{1,j}\in\mathcal{K}_{j,i}}\frac{R^f_{k_{1,j},j,i}(t)}{R_{k_{1,j}}(t-1)}\Big). \tag{27}$$

Here, $\mathcal{J}$ is a set of $\mathcal{N}_{j,i}$'s in sector i, $\mathcal{Q}^{c,f}_{j,i}(t)$ ($\dim \mathcal{Q}^{c,f}_{j,i}(t) = 2$ and $\mathcal{Q}^{c,f}_{j,i}(t) \subseteq \mathcal{K}^{c,f}_{j,i}(t)$) consist of the two CUEs that are jointly processed by $\mathcal{N}_{j,i}$ and $\mathcal{N}_{0,i}$ on RB f by using the precoding technique.

It can be seen that $\lambda^f_{i,A}(t)$ and $\lambda^f_{i,j*,B}(t)$ represent the gains in proportional fairness value at node $\mathcal{N}_{0,i}$ by different strategies in assigning the fth RB at time t. Specifically, $\lambda^f_{i,A}(t)$ calculates the gain in assigning RB f to the best MUE. $\lambda^f_{i,j*,B}(t)$ calculates the gain in assigning RB f to the best CUEs in the coverage range of the $\mathcal{N}_{j^*,i}$. The value of $\lambda^f_i(t)$ is chosen to be the highest among all the gains under different strategies, and the corresponding UE is assigned with the RB.

Based on the obtained $\lambda^f_i(t)$ value, the optimal value of $\mu^f_{j,i}(t)$ can be calculated from (22). It can be considered as the proportional fairness gain at $\mathcal{N}_{j,i}$ by serving the selected PUE. Specifically, the term $R^f_{k_{1,j},j,i}(t)/R_{k_{1,j}}(t-1)$ is the gain of serving the k_1th PUE by $\mathcal{N}_{j,i}$.

Based on the derived optimal Lagrangian multiplier values, we consider the following two cases in finding the optimal k^*_0, $k^*_{1,j}$, k^*_2, $n^{f*}_{k,j,i}(t)$, and $n^{c,f*}_{k,j,i}(t)$ values for each RB f in each sector i at time t .

Case − 1: $\lambda^f_{i,A}(t) \geq \lambda^f_{i,j*,B}(t)$

In this case, we have

$$\lambda^f_i(t) = \max_{k_0\in\mathcal{K}_{0,i}} \frac{R^f_{k_0,0,i}(t)}{R_{k_0}(t-1)}, \tag{28}$$

and

$$\max_{\mathcal{Q}\subseteq\mathcal{K}^{c,f}_{j^*_1,i}}\sum_{k_2\in\mathcal{Q}}\frac{R^{c,f}_{k_2,j^*_1,i}(t)}{R_{k_2}(t-1)} < \max_{k_0\in\mathcal{K}_{0,i}}\frac{R^f_{k_0,0,i}(t)}{R_{k_0}(t-1)} + \frac{R^f_{k^*_{1,j},j^*_1,i}(t)}{R_{k^*_{1,j}}(t-1)}, \tag{29}$$

where

$$j^*_1 = \arg\max_{j\in\mathcal{J}}\Big(\max_{\mathcal{Q}\subseteq\mathcal{K}^{c,f}_{j,i}}\sum_{k_2\in\mathcal{Q}}\frac{R^{c,f}_{k_2,j,i}(t)}{R_{k_2}(t-1)} - \frac{R^f_{k^*_{1,j},j,i}(t)}{R_{k^*_{1,j}}(t-1)}\Big). \tag{30}$$

The left hand side of the inequalities (29) is the proportional fairness gain by serving the best two CUEs in the fth RB of $\mathcal{N}_{0,i}$. The right hand side of (29) is the proportional fairness gain by serving the best MUE on RB f of $\mathcal{N}_{0,i}$ and the best PUE associated with the $\mathcal{N}_{j^*_1,i}$ on RB f separately. From (29), we know that the case with $\lambda^f_{i,A}(t) \geq \lambda^f_{i,j*,B}(t)$ corresponds to the scenario where serving the CUE cooperatively on RB f by the MBS and the PBS receives a less gain than using the RB for the MUE and the PUE separately. In another word, CoMP and precoding shall not be used on RB f for node $\mathcal{N}_{0,i}$.

Substituting (28) into (22)-(25), we have $\nu^{c,f*}_{k_2,j,i}(t) > 0$ for all $k_2 \in \mathcal{K}^{c,f}_{j,i}(t)$, and we can get the optimal indices for different UEs.

The optimal MUE index is

$$k^*_0 = \arg\max_{k_0\in\mathcal{K}_{0,i}} \frac{R^f_{k_0,0,i}(t)}{R_{k_0}(t-1)}, \tag{31}$$

and the optimal PUE index is

$$k^*_{1,j} = \arg\max_{k_{1,j}\in\mathcal{K}_{j,i}} \frac{R^f_{k_{1,j},j,i}(t)}{R_{k_{1,j}}(t-1)}. \tag{32}$$

In the case with $\lambda^f_{i,A}(t) \geq \lambda^f_{i,j*,B}(t)$, the optimal strategy in allocating the fth RB at the tth subframe at node $\mathcal{N}_{0,i}$ is to let $\mathcal{N}_{0,i}$ transmit to the k^*_0th MUE on the entire RB f, $\mathcal{N}_{j_1,i}$ transmit to the $k^*_{1,j}$th PUE on the entire RB f.

The optimal $n^f_{k,j,i}(t)$ and $n^{c,f}_{k,j,i}(t)$ values for the virtual resource allocation problem can be solved from (13) and (14) as

$$n^{f*}_{k,0,i} = 0 \text{ for } k \neq k^*_0,\ n^{f*}_{k,j,i} = 0 \text{ for } k \neq k^*_{1,j} \tag{33}$$

$$
\begin{aligned}
&L(n^f_{k,0,i}(t), n^f_{k,j,i}(t), n^{c,f}_{k,j,i}(t), \lambda^f_i(t), \mu^f_{j,i}(t), \nu^f_{k,j,i}(t), \nu^{c,f}_{k,j,i}(t)) \\
&= -\sum_{k=1}^{N_u}\sum_{i=1}^{N_c}\sum_{j=0}^{N_r} \frac{1}{R_k(t-1)}\Big(x_{k,j,i}(1-x^{c,f}_{k,j,i}(t))R^f_{k,j,i}(t)n^f_{k,j,i}(t) + x_{k,j,i}x^{c,f}_{k,j,i}(t)R^{c,f}_{k,j,i}(t)n^{c,f}_{k,j,i}(t)\Big) \\
&+\sum_{i=1}^{N_c}\lambda^f_i(t)\Big(\sum_{k=1}^{N_u}x_{k,0,i}n^f_{k,0,i}(t) + \sum_{j=1}^{N_r}\sum_{k=1}^{N_u}x_{k,j,i}x^{c,f}_{k,j,i}(t)n^{c,f}_{k,j,i}(t) - 1\Big) \\
&+\sum_{i=1}^{N_c}\sum_{j=1}^{N_r}\mu^f_{j,i}(t)\Big(\sum_{k=1}^{N_u}x_{k,j,i}(1-x^{c,f}_{k,j,i}(t))n^f_{k,j,i}(t) + \sum_{k=1}^{N_u}x_{k,j,i}x^{c,f}_{k,j,i}(t)n^{c,f}_{k,j,i}(t) - 1\Big) \\
&-\sum_{i=1}^{N_c}\sum_{j=0}^{N_r}\sum_{k=1}^{N_u}\nu^f_{k,j,i}(t)n^f_{k,j,i}(t) - \sum_{i=1}^{N_c}\sum_{j=1}^{N_r}\sum_{k=1}^{N_u}\nu^{c,f}_{k,j,i}(t)n^{c,f}_{k,j,i}(t).
\end{aligned} \tag{20}
$$

$$n^{c,f*}_{k,j,i} = 0, \quad n^{f*}_{k^*_0,0,i}(t) = 1, \quad n^{f*}_{k^*_{1,j},j,i}(t) = 1. \tag{34}$$

Case − 2: $\lambda^f_{i,A}(t) < \lambda^f_{i,j*,B}(t)$

In this case, we have

$$\lambda^{f*}_i(t) = \max_{j\in\mathcal{J}}\Big(\max_{\mathcal{Q}\subseteq\mathcal{K}^{c,f}_{j,i}}\sum_{k_2\in\mathcal{Q}}\frac{R^{c,f}_{k_2,j,i}(t)}{R_{k_2}(t-1)} - \frac{R^f_{k^*_{1,j},j,i}(t)}{R_{k^*_{1,j}}(t-1)}\Big). \tag{35}$$

Following the same analysis in Case-1, it is known that $\lambda^f_{i,A}(t) < \lambda^f_{i,j*,B}(t)$ corresponds to a scenario where the gain in proportional fairness value by serving the CUEs cooperatively on RB f by the MBS and the PBS is higher than the gain in using RB f to serve the MUE and the PUE separately. We can obtain the indices for optimal CUE

$$k^*_2 = \arg\max_{\mathcal{Q}\subseteq\mathcal{K}^{c,f}_{j^*,i}}\sum_{k_2\in\mathcal{Q}}\frac{R^{c,f}_{k_2,j^*,i}(t)}{R_{k_2}(t-1)} \tag{36}$$

and

$$j^* = \arg\max_{j\in\mathcal{J}}\Big(\max_{\mathcal{Q}\subseteq\mathcal{K}^{c,f}_{j,i}}\sum_{k_2\in\mathcal{Q}}\frac{R^{c,f}_{k_2,j,i}(t)}{R_{k_2}(t-1)} - \frac{R^f_{k^*_{1,j},j,i}(t)}{R_{k^*_{1,j}}(t-1)}\Big), \tag{37}$$

where $k^*_{j,1}$ value is given in (34).

In the case with $\lambda^f_{i,A}(t) < \lambda^f_{i,j*,B}(t)$, the optimal resource allocation strategy at the tth subframe is to allocate the fth RB of $\mathcal{N}_{0,i}$ and the $\mathcal{N}_{j^*,i}$ to jointly serve the k^*_2th CUE, allocate the fth RB of the PBS with index $j \in \mathcal{J}, j \neq j^*$ to serve the $k^*_{1,j}$th RUE.

The optimal $n^f_{k,j,i}(t)$ and $n^{c,f}_{k,j,i}(t)$ values for the virtual resource allocation problem can be solved from (13)-(14) as

$$n^{c,f*}_{k,j,i} = 0 \text{ for } k \neq k^*_2,\ n^{f*}_{k,j,i} = 0 \text{ for } k \neq k^*_{1,j} \tag{38}$$

$$n^{f*}_{k,0,i} = 0, \quad n^{f*}_{k^*_{1,j},j,i}(t) = 1, \quad n^{c,f*}_{k^*_2,j,i}(t) = 1. \tag{39}$$

B. Summary of Optimization Procedure

So far, we derive the index for optimal MUE, PUE and CUE. For the CUE with the indices $k_2 \in \mathcal{K}^{c,f}_{j,i}(t)$, we need to try different combinations to find the best combination of two CUEs $\mathcal{Q}$ to be served simultaneously, which is very complicated and unfeasible in practice. In order to tackle the computational complexity, we propose a three-step optimization procedure.

Step − 1: Determine UE's associations status
1 − 1): Based on the bias value θ, all the UEs in the system can be decided as either MUEs or PUEs.
Step − 2: At time t, form the CoMP candidate set $\mathcal{K}^{c,f}_{j,i}(t)$. The CUEs in the same CoMP set $\mathcal{K}^{c,f}_{j,i}(t)$ should be associated with the same PBS.
2 − 1): Given an SINR threshold α, all the PUEs whose SINR values are less than α are marked as CUEs and form the CoMP candidate set.
Step − 3: At time t, apply the proposed resource scheduling in two rounds
3 − 1): In the 1st round, following the discussion in Case-1 and Case-2, we assign the best UE (MUE , PUE or CUE) for each RB at each MBS and PBS.
3 − 2): In the 2nd round, for each RB allocated to CUE at the corresponding MBS and PBS, identify the second CUE in the same CoMP set to share the same RB. Apply (4) to obtain the equivalent channel gains. Then evaluate the SINR and R to find the second CUE via (36).

V. Numerical Results

In this section, we conduct the simulation study in a 19-cell 3-sector three-ring hexagonal cellular network by using the extended typical urban (ETU) channel model. Simulation setup follows the guidelines described in 3GPP technical report [1]. The total bandwidth is 10MHz, which is partitioned into 50 resource blocks (RB). The transmit power of the MBS is 46dBm (40W) and the transmit power of the PBS is 30dBm (1W). 50 UEs are uniformly distributed in each sector and travel at a speed of 3 km/h.

In Fig.3, we evaluate the performance of the systems with and without CoMP, as well as with and without precoding. We express the system throughput as the relative percentage of the throughput of the homogeneous network, which consists of only one MBS per sector. It is worth of mentioning that the system throughput does not change with SINR threshold α if there is no CoMP. With CoMP, the system throughput reaches the maximum when $\alpha^* = 0$dB. The throughput goes lower when α is either higher or lower than 0dB. When α is lower, some poor SINR UEs do not have CoMP and their low throughput lead to overall low system throughput. When α is higher, more UEs in the PBS cells will become CUEs. For CUEs with poor SINRs, they have a very good chance to be served cooperatively by MBS and PBS. However, if α is too high, UEs with good SINRs could be unnecessarily served by MBS and PBS cooperatively, leading to a waste of radio resources. Hence, we can see the throughput falloff when α exceeds 0dB. We also compare the CoMP schemes with and without precoding. With precoding, MBS and PBS can support two CUEs simultaneously on the same RB. The system with CoMP + precoding achieves a much larger capacity gain than with CoMP only, approximately 830% vs 680% when $\alpha^* = 0$dB.

In Fig.4, we compare the system throughput performances with CoMP + precoding under different mobile association biases (θ values). $\theta = 0$dB corresponds to a pathloss-based mobile association while $\theta = 16$dB represents a best-power based mobile association. With different θ values, the system can always achieve the highest throughput gain at $\alpha^* = 0$dB. If we evaluate the system in terms of log scale throughput, which is the objective function of $\mathbf{P}_2$, the system log scale throughput is maximized at $\alpha^* = -5$dB. The system can benefit more from a high α threshold but also suffers more from a low α when θ value is low. Comparing $\theta = 2$dB and $\theta = 8$dB, more UEs will fall into the coverage area of PBSs when $\theta = 2$dB and it will lead to a larger pool of CUEs participating CoMP. A too low α will leave a large number of UEs at the PBS cell edge in the low SINR region, thus leading to a low system throughput. This particularly hurts when a low mobile association bias θ is used. Therefore, SINR threshold α needs to be selected appropriately in order to realize a high capacity gain.

VI. Conclusions

In this paper, we proposed a resource allocation scheme for OFDMA-based heterogeneous networks with intra-cell precoding-based cooperation. The proposed scheme can improve the network capacity and coverage considerably. System simulation demonstrates the performance gains by using the proposed CoMP and precoding techniques.

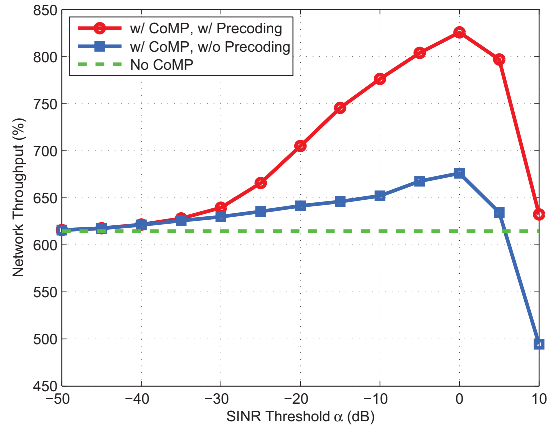

Fig. 3. Network throughput comparison with bias value $\theta = 0$ dB.

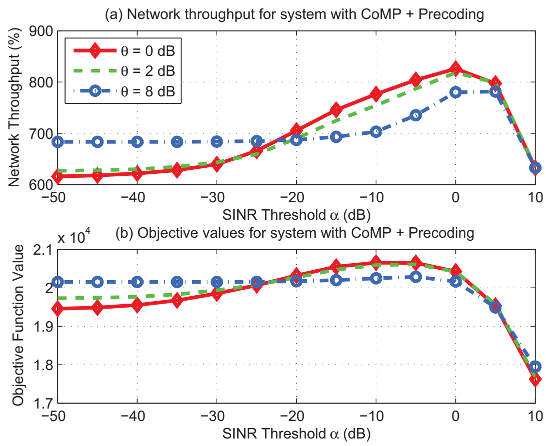

Fig. 4. Performance comparison of system with CoMP and Precoding under different θ.

References

[1] 3GPP TR36.814, "Further advancements for E-UTRA physical layer aspects," v9.0.0, Mar. 2010.

[2] Y. Yu, R. Q. Hu, C. Bontu, Z. Cai, "Mobile Association and Load balancing in a Cooperative Relay Enabled Cellular Network," *IEEE Communications Magazine*, vol. 49, no. 5, pp. 83-89, May 2011.

[3] Qualcomm Europe, "Range expansion for efficient support of heterogeneous networks," TSG-RAN WG1 #54bis R1-083813, Sept. 2008.

[4] Y. Xu, R. Q. Hu, "Optimal Intra-cell Cooperation in the Heterogeneous Relay Networks," in *Proc. IEEE GLOBECOM 2012*, Dec. 2012.

[5] S. Hua, P. Liu, S. Panwar, "The urge to merge: When cellular service providers pool capacity,"in *Proc. IEEE ICC 2012*, Jun. 2012.

[6] Y. Q. Zhou, J. Wang and M. Sawahashi, "Downlink transmission of broadband OFCDM systems Part I: hybrid detection," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 718-729, April 2005.

[7] J. Wang, H. Zhu, and N. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *IEEE J. Select. Areas Commun.*, vol. 30, pp. 675-683, May 2012.

[8] S. Hua, H. Liu, M. Wu, S. Panwar, "Exploiting MIMO antennas in cooperative cognitive radio networks,"in *Proc. IEEE INFOCOM 2011*, pp. 2714-2722, Apr. 2011.

[9] Q. Li, R. Q. Hu, Y. Qian, and G. Wu, "Intra-cell Cooperation and Resource Allocation in a Heterogeneous Network with Relays and Intra-Cell CoMP," accepted by *IEEE Trans. Vel. Tech.*, Jul. 2012

[10] Q. Li, R. Q. Hu, Y. Qian, and G. Wu, "Cooperative Communication for Wireless Networks: Techniques and Applications in LTE-Advanced Systems," *IEEE Wireless Communications Magazine*, vol. 19, no. 2, pp. 22-29, April 2012.

[11] B. Wang, B. Li, M. Liu "A novel precoding method for joint processing in CoMP," in *Proc. ICNCIS 2011*.

[12] H. Kim, Y. Han, "A proportional fair scheduling for multicarrier transmission systems,"*IEEE Commun. Letters*, vol.9, pp. 210-212, Mar. 2005.

[13] A. L. Stolyar, "On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation," *Operations Research*, vol. 53, no. 1, pp. 12-25, 2005.

[14] S. Boyd and V. Vandenberghe, "Convex Optimization," Cambridge University Press, 2004.