Coordinated Scheduling in Downlink Multi-cell OFDMA Networks

Ling Zhang†, Xiangming Wen†, Ting Xu, Zhaoming Lu†, Jun Zhao†, Shenghua He†
†Beijing Key Laboratory of Network System Architecture and Convergence
†Beijing University of Posts and Telecommunications, Beijing, China
Email: sing19901101@hotmail.com

***Abstract*—In this paper, we propose a coordinated scheduling scheme in downlink multi-cell Orthogonal Frequency Division Multiplexing Access (OFDMA) networks. Firstly, a novel decision tree algorithm is introduced to categorize users according to certain attribute parameters to improve system performance. To suppress inter-cluster interference, a filter is used at the receive. To mitigate intra-cluster interference and enhance system performance, we jointly form precoding matrix (using Block Diagonalization precoding) and distribute power for coordinative users. Considering the fact that there is a constraint on number of coordinative users simultaneously served on one resource slot and there may be a large number of users needing coordinative transmission, it is necessary to allocate proper resource slots to them, seeking a tradeoff between coordinative user sum rate and system total capacity. Thus, a performance-based slot distribution is utilized. What needs to be pointed out is that our power control is a per BS power constraint. Simulation results show that our proposed scheme can significantly improve the performance of multi-cell networks.**

I. Introduction

Technological revolutions have hit the wireless communication world over the past decades. Nevertheless, the contradiction between the demand of heterogeneous business and the scarce spectrum resource is still there. To overcome this problem, spectrum reuse and dense BSs deployment in addition to OFDMA techniques are utilized to improve the spectrum efficiency. Yet, the performance of network is still seriously impaired by inter-cell interference (ICI) [1].

In order to suppress the ICI, the 3rd Generation Partnership Project (3GPP) and other organizations have proposed a series of inter-cell interference suppression technology including interference randomization, interference coordination and interference cancellation [2]. However, all of them have drawbacks. Thus, it is not practical to support a large number of users and provide big data service [3] through above schemes. Cooperative schemes have recently been proposed as an effective solutions to suppress ICI, especially for User Equipments (UEs) at the edge of cell. Considering OFDMA has desirable ascendancy and the fact that it has been adopted in the downlinks of LTE-A [4], this paper will concentrate on Multi-BS cooperation problems in downlink multi-cell OFDMA networks.

Multi-BS cooperation is theoretically attractive, nevertheless Multi-BS cooperation faces many challenges. In this paper, to improve the spectrum efficiency, we group them into coordinative users and non-coordinative users by decision tree. There are several literatures about coordinative user selection. In [3] the author utilizes the geometry law or RSVP (Reference Signal Receiving Power) law to distinguish coordinative users and non-coordinative users. It usually classifies them unilaterally according to a certain user's location or RSVP threshold. The author in [5] also proposes a metric based on user location which groups the users into cluster interior and cluster edge users. Such approaches that only rely on signal and interference plus noise ratio (SINR) or UE location, usually resulting a user far away from BS to the need-type, are not accurate enough and may cause waste, impelling us to explore a more precise method to make the classification. Decision Tree, which has a wide range of applications from data mining to economic statistical data processing, has received much attention in intelligent learning recently.

The remainder of this paper is organized as follows. Section II describes the system model. Section III then introduces decision tree into our target problem. Joint precoding and power allocation are given as well as a performance-based slots distribution. Some simulation results are followed in section IV. Finally, some conclusions are drawn in section V.

Notations: For a matrix X, $(X)^*$ denotes the hermitian of matrix. For a set Y, we use $|Y|$ to denote the number of elements and E means the action of expectation. $C^{a\times b}$ refers to a complex matrix with a row and b column.

II. System Model

We consider a system with C clusters (**C**=$\{1,2,..,C\}$), each cluster comprising K cells (**K**=$\{1,2,..K\}$. The system model is as Fig. 1 shows. Every cell consists of one Base Station (BS_k) and a plenty of users (denoted as set U_k^c). Assuming that each BS is equipped with N_t transmit antennas and UEs are equipped with N_r receive antenna. Supposing there are N resource slots (**N**=$1,2,..,N$) for each BS and the maximum transmit per-BS power for coordinative users is P_{max}.

To efficiently accommodate all the users, we classify them into two types: coordinative users and non-coordinative users. Coordinative users do need multiple points transmission while non-coordinative users do not. The user classification algorithm will be related in Section III. Now We denote with $\mathbf{S_k^c}$ and $\mathbf{I_k^c}$ ($\mathbf{I_k^c} = \mathbf{U_k^c} - \mathbf{S_k^c}$) the coordinative users and non-coordinative users respectively in $cell_k$ cluster c. The

978-1-4799-4449-1/14/$31.00

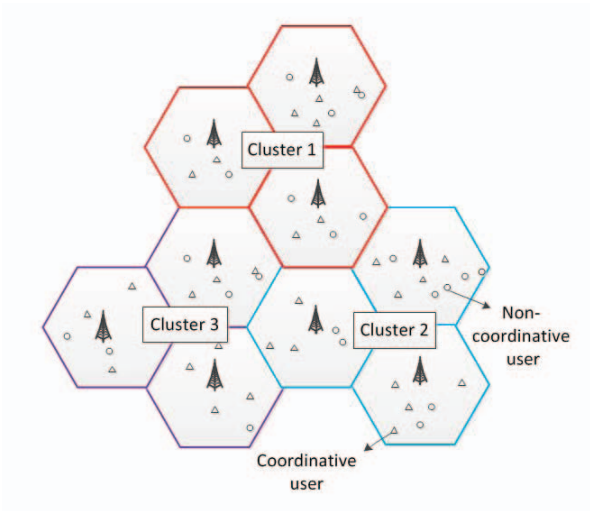

Fig. 1: The system model with $C = 3$ and $K = 3$

congregation of S_k^c is denoted with $\mathbf{\hat{S}_c}$, i.e.$\mathbf{\hat{S}_c} = \bigcup_{k=1}^{k=K} \mathbf{S_k^c}$.

Since system's affordability is limited, only several coordinative users can be served simultaneously. Assuming maximum number of coordinative users to be served on one slot is $S_{\max}$ and we have got $\bar{N}$ slots for coordinative users. We just let $\mathbf{\bar{S}_c^n}$, $\mathbf{\bar{S}_c^n} \subseteq \mathbf{\hat{S}_c}$, represents the selected coordinative users on slot n and $s \in \mathbf{\bar{S}_c^n}$. The received signal $\mathbf{y_s^{c,n}}$ can be written

$$\begin{aligned}\mathbf{y_s^{c,n}} = &\underbrace{\sum_{k=1}^{\mathbf{K}} \mathbf{H_{k,s}^{c,n} W_{k,s}^{c,n} x_s^{c,n}}}_{desired\ signal} + \underbrace{\sum_{k=1}^{\mathbf{K}} \mathbf{H_{k,s}^{c,n}} \sum_{s' \neq s, s' \in \mathbf{\bar{S}_c^n}} \mathbf{W_{k,s'}^{c,n} x_{s'}^{c,n}}}_{intra-cluster\ interference} \\ &+ \underbrace{\sum_{c'=1, c' \neq c}^{\mathbf{C}} \sum_{k'=1}^{\mathbf{K}} \mathbf{H_{k',s}^{c',n}} \sum_{s'' \in \mathbf{\bar{S}_{c'}^n}} \mathbf{W_{k',s''}^{c',n} x_{s''}^{c',n}}}_{inter-cluster\ interference} + \mathbf{z_s^{c,n}}\end{aligned} \tag{1}$$

where $\mathbf{H_{k,s}^{c,n}} \in C^{N_r \times N_t}$ denotes the channel gain matrix from BS_k on slot n in cluster c to coordinative user s. $\mathbf{W_{k,s}^{c,n}} \in C^{N_t \times l}$ denotes the precoding matrix and $\mathbf{x_s^{c,n}} = (a_1, a_2 ... a_l)^* \in C^{l \times 1}$ represents the signal stream. We just assume the number of spatial streams is $l = N_r$. $E[(\mathbf{x_s^{c,n} x_s^{c,n}})^*] = \mathbf{I_l}$. $\mathbf{z_s^{c,n}}$ is additive white Gaussian noise with $E[\mathbf{z_s^{c,n}}(\mathbf{z_s^{c,n}})^*] = \sigma^2 \mathbf{I_{N_r}}$.

The received signal can be rewritten as follow:

$$\begin{aligned}\mathbf{y_s^{c,n}} = \mathbf{H_s^{c,n}} &\sum_{s=1, s \in \mathbf{\bar{S}}_c^n} \mathbf{W_s^{c,n} x_s^{c,n}} + \; + \mathbf{z_s^{c,n}} \\ &\sum_{c'=1, c' \neq c}^{\mathbf{C}} \mathbf{H_s^{c',n}} \sum_{s' \in \mathbf{\bar{S}}_{c'}^n} \mathbf{W_{s'}^{c',n} x_{s'}^{c',n}}\end{aligned} \tag{2}$$

where $\mathbf{H_s^{c,n}} = [\mathbf{H_{1,s}^{c,n}}, \mathbf{H_{2,s}^{c,n}}, .., \mathbf{H_{K,s}^{c,n}}]$ is a $C^{N_r \times KN_t}$ channel matrix, and $\mathbf{W_s^{c,n}} = [(\mathbf{W_{1,s}^{c,n}})^*, (\mathbf{W_{2,s}^{c,n}})^*, .., (\mathbf{W_{K,s}^{c,n}})^*]^*$ is precoder for user s. We plus inter-cluster interference and noise, using $\mathbf{R_s^{c,n}}$ to represent:

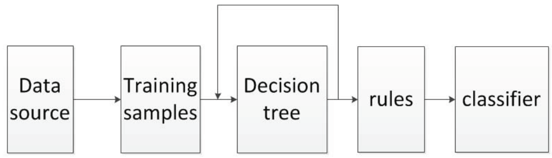

Fig. 2: The process of using a decision tree

$$\mathbf{y_s^{c,n}} = \mathbf{H_s^{c,n}} \sum_{s=1, s \in \mathbf{\bar{S}_c^n}} \mathbf{W_{s'}^{c,n} x_{s'}^{c,n}} + \mathbf{R_s^{c,n}} \tag{3}$$

Denote the covariance matrix of $\mathbf{R_s^{c,n}}$ with $\mathbf{T_s^{c,n}}$. To suppress inter-cell interference, we apply an $N_r \times N_r$ whitening filter $\mathbf{Q_s^{c,n}}$ at the receiver ends, which is shown to be related with $\mathbf{T_s^{c,n}}$ as

$$[\mathbf{T_s^{c,n}}]^{-1} = \mathbf{Q_s^{c,n}}[\mathbf{Q_s^{c,n}}]^*$$

With this whitening filter, the received signal for user s after post-processing is

$$\begin{aligned}\mathbf{y_s^{c,n}} &= \mathbf{Q_s^{c,n} H_s^{c,n}} \sum_{s=1, s \in \mathbf{\bar{S}_c^n}} \mathbf{W_{s'}^{c,n} x_{s'}^{c,n}} + \mathbf{Q_s^{c,n} R_s^{c,n}} \\ &= \mathbf{\hat{H}_s^{c,n}} \sum_{s=1, s \in \mathbf{\bar{S}_c^n}} \mathbf{W_{s'}^{c,n} x_{s'}^{c,n}} + \mathbf{\hat{R}_s^{c,n}}\end{aligned} \tag{4}$$

where $\mathbf{\hat{H}_s^{c,n}} = \mathbf{Q_s^{c,n} H_s^{c,n}}$ and $\mathbf{\hat{R}_s^{c,n}} = \mathbf{Q_s^{c,n} R_s^{c,n}}$ are equivalent channel matrix and noise vector.

III. Coordinated Scheduling Based on Decision Tree

A. Decision Tree Classification

Decision Tree is a flow-chart like structure in which internal node represents test on an attribute and each branch represents outcome of test while each leaf node represents class label (decision taken after computing all attributes). Figure 2 demonstrate the model.

Assuming that groups of samples are obtained and we denote them with $\mathbf{T}$, denote with C_0 and C_1 the coordinative users and the non-coordinative users respectively. Thus we create a proper decision tree which we use to judge user types in real communication. C4.5 algorithm is a classical decision tree building method proposed by Quinlan. It uses information gain ratio rather than information gain to select testing attributes and is suitable to discrete attributes [6].

We pluck four attributes of a service, A={SINR, LOCATION, PACKET LOSS, DELAY}, which may indicate the Quality of Service (QoS) most. Denote the information entropy of samples as $I(\mathbf{T})$ and it can be calculated as follow:

$$I(\mathbf{T}) = -\frac{|\mathbf{P}|}{|\mathbf{T}|}\log_2\frac{|\mathbf{P}|}{|\mathbf{T}|} - \frac{|\mathbf{Q}|}{|\mathbf{T}|}\log_2\frac{|\mathbf{Q}|}{|\mathbf{T}|} \tag{5}$$

where $\mathbf{P}$ and $\mathbf{Q}$ stand for the set of C_0 and C_1 in $\mathbf{T}$ respectively. Assuming $|\mathbf{T}| = M$. A_k is the *k-th* attribute of $\mathbf{A}$. Sort all the samples in ascending order by $[A_{k0}, .., A_{ki}, .., A_{kM}]$. $\mathbf{T_i}$ means the set of samples in interval $\left[A_{k(i-1)}, A_{ki}\right]$. $\mathbf{C_{0i}}$

and $\mathbf{C_{1i}}$ mean the number of C_0 type and C_1 type in $\mathbf{T_i}$ respectively. Split every $\left[A_{k(i-1)}, A_{ki}\right]$ into 2 subinterval uniformly by $\bar{A}_{k_i}$, (i=1,2,...,M-1). If A_k is discrete, we calculate information gain ratio by following:

$$Entropy(A_k) = -\sum_{i=1}^{M-1} \frac{|\mathbf{T_i}|}{|\mathbf{T}|} \times I(\mathbf{C_{0i}}, \mathbf{C_{1i}}) \tag{6}$$

$$I(\mathbf{C_{0i}}, \mathbf{C_{1i}}) = \frac{|\mathbf{C_{0i}}|}{|\mathbf{T_i}|} log_2 \frac{|\mathbf{C_{0i}}|}{|\mathbf{T_i}|} + \frac{|\mathbf{C_{1i}}|}{|\mathbf{T_i}|} log_2 \frac{|\mathbf{C_{1i}}|}{|\mathbf{T_i}|} \tag{7}$$

$$Gain(A_k) = I(\mathbf{T}) - Entropy(A_k) \tag{8}$$

$$Gainratio(A_k) = \frac{Gain(A_k)}{\sum_{i=1}^{M-1} \frac{|\mathbf{T_i}|}{|\mathbf{T}|} \log_2 \frac{|\mathbf{T_i}|}{|\mathbf{T}|}} \tag{9}$$

Applying C4.5 algorithm to our target problem, we build a decision tree as Algorithm 1 goes.

Algorithm 1: Building Decision Tree by C4.5

1 Obtain samples
2 **A**={SINR, LOCATION, PACKET LOSS, DELAY}
3 **While** $\mathbf{A} \neq \Phi$
4 **For** $k \leq |\mathbf{A}|$
5 Sort A_k in ascending order by $[A_{k0}, .., A_{ki}, .., A_{kM}]$.
6 Split every interval $\left[A_{k(i-1)}, A_{ki}\right]$ into 2 subinterval uniformly by $\bar{A}_{k_i}$, (i=1,2,...,M-1).
7 **For** $0 \prec i \prec M$
8 Group samples by $[A_{k0}, \bar{A}_{k_i}]$ and $[\bar{A}_{k_i}, A_{kM}]$, dividing continuous attribute into two discrete intervals ; calculate $Gainratio(A_k)$ as (4)-(7) show as if attributes are discrete except that $M = 2$ now.
9 **End for**
10 Select $\bar{A}_{k_i}$ that has the maximum information gain ratio as threshold and record its $Gainratio(A_k)$
11 **End for**
12 Select A_j that has the maximum information gain ratio as testing attribute
13 Update A by $\mathbf{A} = \mathbf{A} - A_j$
14 **End while**
15 Return the decision tree, construct a training pattern.

In algorithm, Step (4-10) is to discrete continuous attribute. Here we discrete attribute into two subintervals.

In conclusion, given training samples, we first make a decision tree as Algorithm 1 goes. And then UEs feed back four attributes to their own server BS. BS classify users by the tree. After that we can get $\mathbf{S_k^c}$, $\mathbf{I_k^c}$ and $\hat{\mathbf{S}}_{\mathbf{c}}$ in section II.

B. Joint Precoding Method And Power Allocation

To mitigate intra-cluster interference, a classical precoding method Block Digonalization precoding (BD) scheme is utilized. First, construct the aggregate interference matrix for user s in cluster c on slot n as $\bar{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}} = [(\hat{\mathbf{H}}_{\mathbf{1}}^{\mathbf{c,n}})^*, .., (\hat{\mathbf{H}}_{\mathbf{s-1}}^{\mathbf{c,n}})^*, (\hat{\mathbf{H}}_{\mathbf{s+1}}^{\mathbf{c,n}})^*, .., (\hat{\mathbf{H}}_{|\bar{\mathbf{S}}_{\mathbf{c}}^{\mathbf{n}}|}^{\mathbf{c,n}})^*]^*$ The principle idea of BD is to find the precoding matrix $\mathbf{W_s^{c,n}}$ such that $\bar{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}}\mathbf{W_s^{c,n}} = 0$, which means there is no inter-user interference. Thus $\mathbf{W_s^{c,n}}$ lies in the null space of $\bar{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}}$. A sufficient condition for the existence of a nonzero effective channel matrix for user s, $\bar{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}}\mathbf{W_s^{c,n}}$, is that at least one row of $\bar{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}}$ is linearly independent of the rows of $\mathbf{W_s^{c,n}}$. This introduces the constraint that the number of total transmit antennas (KN_t) is no smaller than the number of total receive antennas $(|\bar{\mathbf{S}}_{\mathbf{c}}^{\mathbf{n}}|N_r)$. Therefore, there is a constraint on the total number of users that can be served simultaneously in each cluster, specified as follows:

Lemma 1 (User constraint for multi-cell BD): For a clustered MIMO network with K BSs per cluster, the maximum number of users that can be supported simultaneously on one slot in each cluster by multi-cell BD is bounded by

$$S_{\max} \leq \left\lfloor \frac{KN_t}{N_r} \right\rfloor$$

where $\lfloor x \rfloor$ is the maximum integer less than or equal to x.

We describe the precoding matrix design as follows. Let $\tilde{N}_s^{(c,n)} = rank(\bar{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}})$, and denote the singular value decomposition (SVD) of $\bar{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}}$ as

$$\bar{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}} = \tilde{\mathbf{U}}_{\mathbf{s}}^{\mathbf{c,n}} \tilde{\mathbf{\Lambda}}_{\mathbf{s}}^{\mathbf{c,n}} [\tilde{\mathbf{V}}_{\mathbf{s}}^{(\mathbf{c,n,1})} \tilde{\mathbf{V}}_{\mathbf{s}}^{(\mathbf{c,n,0})}]^* \tag{10}$$

where $\tilde{\mathbf{U}}_{\mathbf{s}}^{\mathbf{c,n}}$ and $\tilde{\mathbf{\Lambda}}_{\mathbf{s}}^{\mathbf{c,n}}$ represents left singular matrix and diagonal matrix of the singular value respectively. $\tilde{\mathbf{V}}_{\mathbf{s}}^{(\mathbf{c,n,1})}$ comprises the first $\tilde{N}_s^{(c,n)}$ right singular vectors while $\tilde{\mathbf{V}}_{\mathbf{s}}^{(\mathbf{c,n,0})}$ comprises the last $(KN_t - \tilde{N}_s^{(c,n)})$ right singular vectors which consists of a set of orthogonal basis from which we get $\mathbf{W_s^{c,n}}$.

With the derived $\mathbf{W_s^{c,n}}$, the received signal becomes

$$\mathbf{y_s^{c,n}} = \hat{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}} \mathbf{W_s^{c,n}} \mathbf{x_s^{c,n}} + \hat{\mathbf{R}}_{\mathbf{s}}^{\mathbf{c,n}} \tag{11}$$

Thus, the achievable sum rate per cluster is given by [6]

$$C_c = \max_{\mathbf{W_{k,s}^{c,n}}} \sum_{n=1}^{n \in \bar{\mathbf{N}}} \sum_{s \in \bar{\mathbf{S}}_c^{\mathbf{n}}} \log_2 \det(\mathbf{I}_{N_r} + \mathbf{E_s^{c,n}}) \tag{12}$$

Since $E[\mathbf{x_s^{c,n}}(\mathbf{x_s^{c,n}})^*] = \mathbf{I_l}$, Then $\mathbf{E_s^{c,n}}$ is as follow:

$$\begin{aligned} \mathbf{E_s^{c,n}} &= \hat{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}} \mathbf{W_s^{c,n}} \mathbf{x_s^{c,n}} (\mathbf{x_s^{c,n}})^* (\mathbf{W_s^{c,n}})^* (\hat{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}})^* \\ &= \hat{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}} \mathbf{W_s^{c,n}} (\mathbf{W_s^{c,n}})^* (\hat{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}})^* \end{aligned} \tag{13}$$

Denote the cluster power allocation matrix is

$$\mathbf{W_k^c} = \begin{bmatrix} \mathbf{W_{k,1}^{c,1}} & \cdots & \mathbf{W_{k,|\bar{s}_c^1|}^{c,1}} \\ \cdots & & \cdots \\ \mathbf{W_{k,1}^{c,|\bar{N}|}} & \cdots & \mathbf{W_{k,|\bar{s}_c^{|\bar{N}|}|}^{c,|\bar{N}|}} \end{bmatrix}, \quad k \in \mathbf{K}$$

Then the transmit power constraint for each BS is

$$sum(\mathbf{W_k^c}) = \sum_{n \in \bar{\mathbf{N}}} \sum_{s \in \bar{\mathbf{S}}_c} \mathbf{W_{k,s}^{c,n}} \leq P_{\max}, \quad k \in \mathbf{K} \tag{14}$$

Denote the SVD of the efficient channel $\hat{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}}\mathbf{W_s^{c,n}}$ as

$$\hat{\mathbf{H}}_{\mathbf{s}}^{\mathbf{c,n}} \mathbf{W_s^{c,n}} = \mathbf{U_s^{c,n}} \begin{bmatrix} \mathbf{\Lambda_s^{c,n}} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{V_s^{c,n}} \tag{15}$$

where $\mathbf{\Lambda_s^{c,n}} = diag(\lambda_{s,1}, \lambda_{s,2}, .., \lambda_{s,r_k})$ and $r_k = rank(\mathbf{\hat{H}_s^{c,n} W_s^{c,n}})$. Let $\mathbf{\Lambda^{c,n}} = blockdiag\{\mathbf{\Lambda_1^{c,n}}, .., \mathbf{\Lambda_{|\bar{S}_c^n|}^{c,n}}\}$. Denote $\mathbf{V^{c,n}} = blockdiag\{\mathbf{V_1^{c,n}}, \mathbf{V_2^{c,n}}, .., \mathbf{V_{|\bar{S}_c^n|}^{c,n}}\}$.Then the cluster sum rate optimal problem can be written as

$$\begin{cases} C_c = \max\limits_{\mathbf{W_k^c}} \sum\limits_{n=1}^{n \in \mathbf{\bar{N}}} log_2 \det(\mathbf{I} + \mathbf{\Lambda^{c,n} \hat{Q}^{c,n}} (\mathbf{\Lambda^{c,n}})^*) \\ s.t.\ \ sum(\mathbf{W_k^c}) \leq P_{\max},\ \ k \in \mathbf{K} \end{cases} \quad (16)$$

. where $\mathbf{\hat{Q}^{c,n}} = (\mathbf{V^{c,n}})^* \mathbf{V^{c,n}}$.

To simplify the problem, we divide it to $\bar{N}$ low dimensional optimal problem and find a suboptimal solution. We try to optimize the following:

$$\begin{cases} C_{(c,n)} = \max\limits_{\mathbf{W_{k,s}^{c,n}}} log_2 \det(\mathbf{I} + \mathbf{\Lambda^{c,n} Q^{c,n}} (\mathbf{\Lambda^{c,n}})^*) \\ s.t.\ \ \sum\limits_{s \in \mathbf{\bar{S}}_c^n} \mathbf{W_{k,s}^{c,n}} \leq \frac{P_{\max}}{\bar{N}},\ \ n \in \mathbf{\bar{N}}, k \in \mathbf{K} \end{cases}, \quad n \in \mathbf{\bar{N}} \quad (17)$$

We know that the optimal power allocation to (17) is a diagonal matrix [7], making the optimization problem an algebra problem. It was shown in [8] that for the special case of $K = 1$ or 2, this problem can be efficiently solved by the waterfilling algorithm or by characterizing the intersections of the hyperplane constraints, which, however, cannot be easily extended to the more general case. To the best of our knowledge, at this point no closed-form solutions such as waterfilling are available for the optimization problem. The objective function, however, is concave and the constraint functions are linear, so this is a convex optimization problem and can be solved numerically, e.g. with the interiorpoint method. Recently, a more efficient algorithm was proposed in [9], where the dual problem is solved by a B-dimensional subgradient iteration which solves our problem.

C. User Scheduling

As we know there are N resource slots for total users at every BS. Now, we need allocate slots properly. To strike a balance between coordinative user sum rate and system total performance, we can distribute resource slots as algorithm 2 goes. Denote $\hat{N}$ as the slots for non-coordinative users. We first divide the users in $\mathbf{\hat{S}_c}$ into M groups, each have S_{max} users. Denote $C_{coor}^{i,n}, i = \{1, 2, .., M\}$ as the *i-th* group coordinative sum rate when slot n is allocated. C_k^n means non-coordinative user rate when allocated with slot n. Once the slots are distributed, users are also selected. At every scheduling interval, such a slots distribution is conducted.

Algorithm 2: Performance-based Slots Distribution

$\bar{N} = \emptyset, \hat{N} = \emptyset$
For $n = \{1, 2, .., N\}$
  For $i = \{1, 2, .., M\}$
   Calculate $C_{coor}^{i,n}$
  End for
   Select the maximum $C_{coor}^{i,n}$, $C_{coor}^{i,n} \rightarrow C_{coor}^{max}$
  For $k = \{1, 2, .., K\}$
   For every BS_k, search exhaustively for best non-coordinative users for slot n
   Calculate C_k^n
  End for
 Calculate $C_{non} = \sum_{k=1}^{K} C_k^n$
  If $C_{non} \leq C_{coor}^{\max}$
   $n \rightarrow \bar{N}$
  Else $n \rightarrow \hat{N}$
  End if
End for

IV. Simulation Results

A. Decision Tree Simulations

We utilize a group of data tested by other objects in our laboratory as in Table I which evaluates the Quality of Service according to several attribute parameters. If a service is good enough to a user, it was defined as '1.0'. Otherwise, it was '0.0'. Here we declare that '0.0' means C_0 type while '1.0' means C_1 type.

Then we use Table I to build a decision tree classifier by Weka 7.0. And the result is as Figure 3 shows (label 'yes' means C_1 type and label 'no' means C_0 type). Here we can see that only two attributes are selected and this is reasonable since SINR and DELAY are decisive for users' experience of certain service (such as real time transmission video). The number after labels is the quantity of samples classified to that type and '17.0/1.0' means that there is one sample being judged wrongly.

TABLE I: Training Data

No.	SINR (dB)	DELAY (ms)	PACKET LOSS	LOCATION (km)	CLASS
1	22.9416	9.6274	0.0556	699.3316	0.0
2	20.8948	6.3867	0.0571	800.4586	0.0
3	22.6839	5.8296	0.0945	598.1755	1.0
...	....				
28	22.8927	5.5716	0.0612	600.2993	1.0
29	19.7783	11.1725	0.0705	901.0142	0.0
30	25.5531	5.1537	0.0441	499.1056	1.0

B. System Performance Simulations

In this section, the performance of the proposed coordination strategy is shown via monte carlo simulation. We choose the number of antennas to be $N_t = 4$ and $N_r = 2$ and the

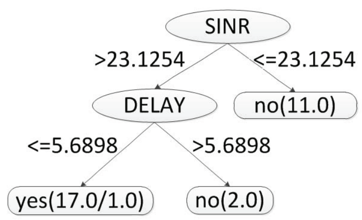

Fig. 3: Built decision tree

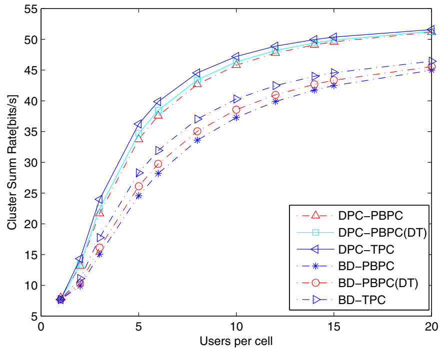

Fig. 4: Cluster sum rate for different systems, with cluster size $K = 3$.

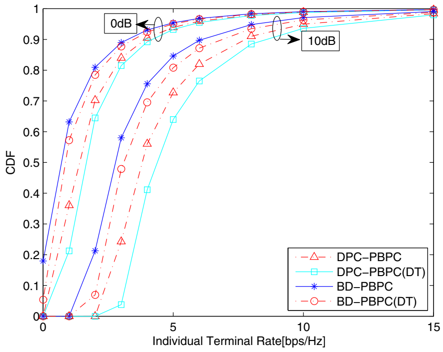

Fig. 5: CDF of the rates for coordinative users in the cluster. $K = 3$

resource slots $N = 10$. The standard deviation of shadowing is 8 dB, the path loss exponent is 3.7, and the cell radius is 1 km.

In order to illustrate the excellent performance of the proposed scheme with decision tree, we compare it with different systems as Fig. 4 shows. Dirty paper coding(DPC) is a excellent precoding method except that it is nonlinear and extremely complicated. TPC means total power control while PBPC means per BS power control. From Fig. 4, we see that DPC systems have better performance than BD systems. PBPC schemes are generally a little bit no-good than TPC schemes. Last but not least, when added with decision tree user classification, both DPC-PBPC and BD-PBPC show a progress than that without decision tree.

Figure 5 shows the cumulative distribution function (CDF) of mean rates of users. We totally use 3000 samples of user rates, with which we can plot the CDF. As Fig. 5 shows, two group of lines are plot respectively. One is under the condition of transmit power being 0dB and the other is simulated with transmit power being 10dB. Obviously, more transmit power lead to better capacity. Furthermore, we can see that the rate with 70% outrage when transmit power being 10dB for DPC-PBPC is 4 bps/Hz, for DPC-PBPC (DT) is 4.2 bps/Hz and for BD-PBPC with or without DT are 3.8 bps/Hz and 3.7 bps/Hz, respectively. For BD schemes, 80% users have mean rate larger than 4bps/Hz.

V. Conclusion

We described a coordinated scheduling scheme which use a decision tree classification to classify users. To maximum the system capacity, we divided resource slots based on system performance. A whitening filter to cancel inter-cluster interference is used. Then, we jointly precode and distribute power for all selected coordinative users to mitigate intra-cluster interference. We break up the object problem which is extremely complicated. Cooperation can be conducted between femtocells [10-11]. Mobility robust optimization in [12] may also be our future work.

VI. Acknowledgement

This research is supported by the Co-building Project of Beijing Municipal Commission of Education (The development of the always online application detection and monitoring equipment for mobile internet) and the fundamental research funds for the central universities (2013RC0110).

References

[1] 3GPP TS 36.300 V8.9.0. "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN)", Overall description; Stage 2[S]. 2011

[2] Rahman M, et al. "Enhancing cell-edge performance: a downlink dynamic interference avoidance scheme with inter-cell coordination[J]", Wireless Communications, IEEE Transactions on, 2010, 9(4): 1414-1425.

[3] Fu W, et al. "The inter-cell interference suppression algorithm based on the JP-CoMP and performance simulation[C]", Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on. IEEE, 2013: 528-533.

[4] Mogensen P E. Pedersen K I, et al. "LTE-advanced: the path towards gigabits in wireless mobile communications[C]", Wireless Communication, Vehicular Technology, Information Theory and Aerospace and Electronic Systems Technology, 2009. Wireless VITAE 2009. 1st International Conference on. IEEE, 2009: 147-151.

[5] Zhang J, et al. "Networked MIMO with clustered linear precoding[J]", Wireless Communications, IEEE Transactions on, 2009, 8(4): 1910-1921.

[6] Chang Z, et al. "The application of C4. 5 algorithm based on SMOTE in financial distress prediction model[C]", Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2011 2nd International Conference on. IEEE, 2011: 5852-5855.

[7] I. E. Telatar, et al. "Capacity of mulit-antenna Gaussian channels", Europ. Trans. Telecommun., vol. 10, pp. 585C595, Nov. 1999.

[8] F. Boccardi, et al. "Zero-forcing precoding for the MIMO broadcast channel under per-antenna power constraints", in Proc. IEEE SPAWC, Cannes, July 2006, pp. 1C5.

[9] G. Caire, et al. "Multiuser MIMO downlink with limited inter-cell cooperation: approximate interference alignment in time, frequency, and space", in Proc. Allerton Conf. on Comm. Control and Comp., Monticello, IL, Sept. 2008.

[10] Haijun Zhang, et al. "Resource Allocation in Spectrum-Sharing Femtocells with Heterogeneous Services", to appear in IEEE Trans. on Communications.

[11] H. Zhang, et al. "Resource Allocation with Interference Mitigation in OFDMA Femtocells for Co-channel Deployment", EURASIP Journal on Wireless Communications and Networking: Special Issue on Femtocells in 4G Systems, 2012:289, doi:10.1186/1687-1499-2012-289.

[12] H. Zhang, et al. "Mobility Robustness Optimization in Femtocell Networks based on Ant Colony Algorithm", IEICE Trans. Commun., Vol.E95-B No.4 pp.1455-1458.