Joint Tilt Angle Adaptation and Beamforming in Multicell Multiuser Cellular Networks

Soheil Khavari Moghaddam and S. Mohammad Razavizadeh

School of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

khavari@elec.iust.ac.ir, smrazavi@iust.ac.ir

Abstract

3D beamforming is a promising approach for interference coordination in cellular networks which brings significant improvements in comparison with conventional 2D beamforming techniques. This paper investigates the problem of joint beamforming design and tilt angle adaptation of the BS antenna array for maximizing energy efficiency (EE) in downlink of multi-cell multi-user coordinated cellular networks. An iterative algorithm based on fractional programming approach is introduced to solve the resulting non-convex optimization problem. In each iteration, users are clustered based on their elevation angle. Then, optimization of the tilt angle is carried out through a reduced complexity greedy search to find the best tilt angle for a given placement of the users. Numerical results show that the proposed algorithm achieves higher EE compared to the 2D beamforming techniques.

Keywords: 3D beamforming (3DBF), energy efficiency, tilt angle optimization, vertical beamforming, convex optimization, coordinated beamforming, fractional programming, 2D beamforming.

1. Introduction

Tilt angle adaptation which is also known as three dimensional beamforming (3DBF), full dimension multiple-input multiple-output (FD-MIMO) or vertical beamforming is a promising technology for interference management and performance improvement in fifth generation (5G) cellular networks [1, 2]. In this technique by deploying active antenna systems (AAS) at the base station (BS) of cellular networks, it is possible to dynamically adapt the antenna tilt angle in each transmission interval [3]. This can be done by modifying parameters of symmetrical 3D pattern introduced in [4]. The antenna arrays radiates a fan-shaped beam with a large half power beam width (HPBW) in horizontal plane, while radiating a sharp beam with a small HPBW in the vertical plane which prevents signal from leaking to adjacent cells. In contrast to the horizontal plane in which the direction of antenna main lobe is fixed (since the BS orientation is fixed for each sector), the antenna main lobe in the vertical plane can be steered to a desired direction by changing the tilt angle of the BS's antenna array [2]. To this end, we need information about users' locations and

angle of arrival (AoA) at the BS. This tilt angle adjustment can improve some performance metrics such as spectral efficiency (SE), coverage probability or energy efficiency (EE) [1, 2].

Most of the previous works on 3DBF study the problem of maximizing of the spectral efficiency in various scenarios. In [5] a game theory-based approach is proposed that maximizes the sum SE of a multi-cell network. In addition the authors approximate the signal-to-interference-plus-noise ratio (SINR) by its asymptotic value when the number of the BS antennas goes to infinity. In [6, 7, 8] the authors extract conditional ergodic sum SE of the network in terms of the BS antenna array tilt angle and then approximate it with complicated mathematical expressions. Hence the optimum tilt angles are found using exhaustive search over different possible tilt angles.

In general there are two 3DBF strategies: passive and active. In the passive 3DBF, the antenna pattern cannot be dynamically changed, and thus the antenna array tilt angle is fixed during several transmission intervals. This technique is suitable for applications like self organizing networks [9]. However in the active 3DBF, the antenna pattern can be dynamically changed i.e. the tilt angle is optimized by utilizing instantaneous user's location information and AoA in each transmission interval [2, 3]. Thus, information about the users' locations is playing a vital role in designing of the 3DBF technique. The users' locations also can be modeled with the help of stochastic geometry (SG) [3, 9, 10]. To this end the users' locations is modeled by a uniform distribution and the BSs' are assumed to be located on the vertices of a hexagonal grid [10]. In addition, the BSs' locations can also be modeled with Poisson Point Process (PPP) [9]. Since the location of the users is a slow varying parameter, it can be estimated almost exactly and the assumption of knowing these informations is a reasonable assumption [5, 11]. The subject of AoA estimation is well studied in literature [12]. Also there are some valuable works which concern with AoA estimation using novel low energy consuming machine learning approaches such as Bayesian compressive sensing algorithms [13, 14]. Here similar to [3, 5, 8, 11, 15, 22] we assume the information of the AoA of the users are accurately available at the BSs.

It is shown that the active 3DBF can achieve higher performance compared with the passive 3DBF since it adapts the antenna pattern to the instantaneous location of the users instead of using an average location of the users [2, 3]. The problem of maximizing the sum SE in the active and passive 3DBF for a single cell network and two scenarios of single-user and multi-user is presented in [3]. Although in [3] potential of the 3DBF in performance improvement and advantage of active 3DBF over passive 3DBF are well studied, however because of the single cell arrangement the interference management property of the 3DBF is not fully exploited.

Although SE is a very important performance metric in the wireless networks, recently EE is also becoming more important from the network service provider's point of view. The 3DBF is one of the techniques that have been proposed to facilitate energy efficient communication in next generation of cellular networks. In this technique, by proper managing of interference and reducing the transmit power, the EE can be increased [2]. In spite of this fact, the problem of EE maximization by utilizing 3DBF has not been thoroughly investigated in the literature. In [9] by exploiting 3DBF, optimization of the sum EE in a two tier heterogeneous network (HetNet) is addressed.

In this paper, we investigate the problem of joint beamforming and tilt angle optimization at the BS antenna array for maximizing the total EE in a multi-cell multi-user network. In fact, on contrary to the previous works on 3DBF which have considered ergodic sum SE averaged over the users' locations, we investigate maximizing the total EE in a more realistic system with information on the users' locations in the BS. In addition, it is worthwhile to note that in all previous works simple linear beamforming such as eigen beamforming or zero-forcing beamforming were considered while we design the optimum beamforming vectors through a non-linear optimization problem. To the best of our knowledge, there is no similar work which jointly design beamfroming vectors and optimize tilt angle.

In addition, in our scenario the BSs cooperate with each other and share their channel vectors to each user. The objective function is total instantaneous EE subject to sum transmit power constraint at the BS. The resulting optimization problem is non-convex and therefore, a fractional programming approach is employed to convert it to a convex problem. To convexify this problem, lemma 1 is introduced and then to solve the convex problem using block coordinate descent method, theorem 1 is introduced. Also to reduce the complexity of searching for finding the optimal tilt angle, the clustering algorithm is proposed in theorem 2.

Although the new problem is convex in terms of the beamforming vectors, it is still non-convex in terms of the tilt angle. To solve this problem, we divide the users into clusters and then in each cluster the optimum beamforming vectors are calculated. Finally, the optimum tilt angle is found according to the cluster in which the beamforming vectors maximize the total instantaneous EE. It will be shown that the optimal beamforming vectors are dependent on both users' elevation angle and channel gains. Simulation results demonstrate that our method outperforms the conventional 2D beamforming in which the antenna array pattern is omnidirectional and only precoding vectors are designed.

Rest of the paper is organized as follows. In Section 2, the system model and problem formulation are introduced. In Section 3, the proposed algorithm for solving the optimization problem is presented. In Section 4, the clustering algorithm is introduced. Simulation results are presented in Section 5. Finally, Section 6 concludes the paper.

In the paper scalars are denoted by lower-case letters. Vectors and matrices are denoted by bold-face lower-case and upper-case letters, respectively. $\prec$ stands for the generalized inequality. $(.)^H$ is the complex conjugate transpose. $\mathbb{E}\{.\}$ denotes statistical expectation and $||.||$ is the Euclidean distance (also known as L^2-norm). also $(.)^*$ represents the complex conjugate norm. I_M is the Identity matrix of size M and $\mathbb{C}^N$ is the $N-$dimensional complex vector space. In addition, a list of all variables that are used in our paper is given in table 1.

2. System Model

We consider a multi-cell network consisting of L cells. Each cell has K users that are uniformly distributed over the cell coverage region and a BS equipped with a linear antenna array consisting of M active elements. Each user is only served by its own cell's BS.

Parameter	Description
θ_j, ϕ_j	Tilt and boresight angles of the j-th BS respectively.
SLL_{el}, SLL_{az}	Side lobe levels in the elevation and azimuth domain respectively.
$\boldsymbol{x}_i$	Transmitted signal vector of the i-th BS.
$\boldsymbol{\omega}_{in}, d_{in}, y_{in}$	Beamforming vector, data symbol and received signal of the n-th user in the i-th cell.
$\boldsymbol{g}_{ijm}, \beta_{ijm}$	Channel vector and large scale factor between the m-th user in the j-th cell and the i-th BS.
θ_{3dB}, ϕ_{3dB}	Half power beamwidth of the elevation and azimuth patterns, respectively.
θ_{ijk}, ϕ_{ijk}	AoA of the k-th user in the j-th cell and the i-th BS
R_{jm}	Instantaneous rate of the m-th user in the j-th cell.
R_{max}	Maximum rate achievable if each user receives maximum power of its serving BS and no interference exists.
$\boldsymbol{W}, \boldsymbol{\theta}$	Collection of all beamforming vectors and tilt angle of all BSs.
P, P_c, P_0	Maximum transmit power available at the BS, RF-chain and constant power consumption of the BS site, respectively.
ξ	Power amplifier inefficiency
M, L, K	The number of antennas at the BS, number of cells and the number of the users in each cell.
η	Energy efficiency variable.
ϵ, δ	Thresholds for stop criterion of the inner and outer layer algorithms.
$\tilde{d}_{jm}, \mu_{jm}, s_{jm}, e_{jm}$	Estimated symbol, filter coefficient, slack variable and symbol estimation error at the UE of the m-th user in the j-th cell

Table 1: List of all parameters throughout the paper.

The user equipment's (UE's) antenna pattern is assumed to be isotropic and the 3D pattern of the BS antennas is modeled as [4]

$$\alpha(\theta_{BS,i}) = G_{\max} - \min\left[12\left(\frac{\phi}{\phi_{3dB}}\right)^2, \text{SLL}_{az}\right] - \min\left[12\left(\frac{\theta_{BS,i}-\theta_{ijk}}{\theta_{3dB}}\right)^2, \text{SLL}_{el}\right]. \tag{1}$$

where as depicted in Fig. 1, $\theta_{BS,i}$ is the tilt angle of the i-th BS which is defined as the angle between the horizon and the main lobe of vertical pattern of the BS antenna array. θ_{ijk} is the vertical AoA of the k-th user in the j-th cell to the i-th BS. In addition $\phi = \phi_{BS,i} - \phi_{ijk}$ is the horizontal angle difference between boresight of the i-th BS's antenna and the horizontal AoA of the k-th user in the j-th cell to the i-th BS. It should be noted that $\phi_{BS,i}$ is added to the generic model in [4] to suitably encompass the base stations with nonzero boresight angle. Furthermore, $\text{SLL}_{az} = 25$ dB and $\text{SLL}_{el} = 20$ dB are side lobe levels in the azimuth and elevation planes, respectively. Also $\phi_{3dB} = 65°$ and $\theta_{3dB} = 6°$ are HPBWs in the

azimuth and elevation planes, respectively [7]. For a conventional 2D beamforming system, the array's gain in the vertical plane is ignored and only horizontal gain and the first constant term is considered.

During the signal transmission in the downlink, each BS broadcasts its signal to all users. This signal is obtained by multiplying the users' symbols in the corresponding beamforming vectors. Hence, the transmit signal from the i-th BS is obtained as follows

$$\boldsymbol{x}_i = \sum_{n=1}^{K} \boldsymbol{\omega}_{in} d_{in} \tag{2}$$

where $\boldsymbol{\omega}_{in} \in \mathbb{C}^{M \times 1}$ and d_{in} are the beamforming vector and data symbol of the n-th user in the i-th cell, respectively. Moreover, the symbols d_{in} are assumed to be uncorrelated with $\mathbb{E}\{|d_{in}|^2\} = 1$. The received signal of the m-th user in the j-th cell can be written as

$$y_{jm} = \sum_{i=1}^{L} \sqrt{\alpha\left(\theta_{BS,i}\right)} \boldsymbol{g}_{ijm}^{H} \boldsymbol{x}_i = \sum_{i=1}^{L} \sum_{n=1}^{K} \sqrt{\alpha\left(\theta_{BS,i}\right)} \boldsymbol{g}_{ijm}^{H} \boldsymbol{\omega}_{in} d_{in} + n_{jm}$$

$$= \underbrace{\sqrt{\alpha\left(\theta_{BS,j}\right)} \boldsymbol{g}_{jjm}^{H} \boldsymbol{\omega}_{jm} d_{jm}}_{\text{desired}} + \underbrace{\sum_{\substack{i=1 \\ (i,n) \neq (j,m)}}^{L} \sum_{n=1}^{K} \sqrt{\alpha\left(\theta_{BS,i}\right)} \boldsymbol{g}_{ijm}^{H} \boldsymbol{\omega}_{in} d_{in}}_{\text{interference}} + \underbrace{n_{jm}}_{\text{noise}}, \tag{3}$$

where $\boldsymbol{g}_{ijm} \sim \mathcal{CN}\left(0, \beta_{ijm} \boldsymbol{I}_M\right)$ is the channel vector between the m-th user in the j-th cell and the BS in the i-th cell. β_{ijm} is the large scale fading factor that includes both path-loss and shadow fading effects. The noise term n_{jm} is a circularly symmetric complex gaussian random variable with zero mean and normalized variance i.e. $\mathbb{E}\{|n_{jm}|^2\} = 1$. Instantaneous rate of the m-th user in the j-th cell can be written as

$$\hat{R}_{jm} = \log_2 \left(1 + \frac{\alpha\left(\theta_{BS,j}\right) \left|\boldsymbol{g}_{jjm}^{H} \boldsymbol{\omega}_{jm}\right|^2}{\sum\limits_{\substack{i=1 \\ (i,n) \neq (j,m)}}^{L} \sum\limits_{n=1}^{K} \alpha\left(\theta_{BS,i}\right) \left|\boldsymbol{g}_{ijm}^{H} \boldsymbol{\omega}_{in}\right|^2 + 1} \right). \tag{4}$$

Using logarithm in base e, we have $R_{jm} = \ln(1 + \text{SINR}) = \frac{\hat{R}_{jm}}{\ln 2}$. The overall EE of the network is defined as the total weighted sum SE divided by the total energy consumption in the network as follows

$$\text{EE} = \frac{\sum\limits_{j,m} b_{jm} R_{jm}}{\xi \sum\limits_{j,m} \|\boldsymbol{\omega}_{jm}\|^2 + MLP_c + LP_0} = f\left(\boldsymbol{W}, \boldsymbol{\theta}\right). \tag{5}$$

In (5), $\boldsymbol{W}$ is the collection of all beamforming vectors and $\boldsymbol{\theta} = [\theta_{BS,1}, \theta_{BS,2}, ..., \theta_{BS,L}]$ is the vector containing the tilt angles of all BSs. b_{jm} is weight of the m-th user in the j-th cell and in fact it indicates each user's priority with respect to other users. P_c is power consumption of an active RF-chain and P_0 is constant power consumption for maintenance of BS equipments [17]. $\xi \geq 0$ is a constant that accounts for the power amplifiers' inefficiency. Now we can formulate optimization problem of the network as follow

$$
\begin{aligned}
\arg \max_{\boldsymbol{W}, \boldsymbol{\theta}} f\left(\boldsymbol{W}, \boldsymbol{\theta}\right) &= \frac{f_1\left(\boldsymbol{W}, \boldsymbol{\theta}\right)}{f_2\left(\boldsymbol{W}, \boldsymbol{\theta}\right)}, \\
\text{subject to} \quad &\sum_{m=1}^{K} \|\boldsymbol{\omega}_{jm}\|^2 \leqslant P \quad \forall j, \quad j = 1, 2, ..., L, \\
&0_{1 \times L} \prec \boldsymbol{\theta} \prec 90^o \times [1, 1, ..., 1]_{1 \times L}.
\end{aligned}
\tag{6}
$$

where $f_1(\boldsymbol{W}) = \sum_{j,m} b_{jm} R_{jm}$ and $f_2(\boldsymbol{W}) = \xi \sum_{j,m} \|\boldsymbol{\omega}_{jm}\|^2 + MLP_c + LP_0$ which are numerator and denominator of equation (6). The first constraint in (6) is the maximum transmit power available at the BS and the second constrain is the dynamic range of the tilt angles.

It is possible to show that the following bounds are true for the objective function in (6)

$$
\begin{cases}
0 < f_1\left(\boldsymbol{W}, \boldsymbol{\theta}\right) \leqslant R_{\max} \\
f_2\left(\boldsymbol{W}, \boldsymbol{\theta}\right) \geqslant MLP_c + LP_0 \\
f_2\left(\boldsymbol{W}, \boldsymbol{\theta}\right) \leqslant LP + MLP_c + LP_0 \\
R_{\max} = \sum_{j=1}^{L} \sum_{m=1}^{K} \log_2\left(1 + \frac{P\|\boldsymbol{g}_{jjm}^H\|^2}{1}\right)
\end{cases}
\tag{7}
$$

where R_{max} is the maximum rate for the case that the maximum power is allocated for each user and no interference exists.

3. The EE Maximization Problem Solution

Since the objective function in (6) is fractional and non-convex, fractional programming technique is applied to solve this problem [18]. First consider the following equation

$$
F(\eta) = \max_{(\boldsymbol{W}, \boldsymbol{\theta}) \in \mathbb{D}} \{f_1\left(\boldsymbol{W}, \boldsymbol{\theta}\right) - \eta f_2\left(\boldsymbol{W}, \boldsymbol{\theta}\right)\}
\tag{8}
$$

where $\mathbb{D} = \{(W, \theta) | \sum_{m=1}^{K} \|\boldsymbol{\omega}_{jm}\|^2 \leqslant P \quad \forall j, \ 0_{1 \times L} \prec \boldsymbol{\theta} \prec 90^o \times [1, 1, ..., 1]_{1 \times L}\}$ is feasible set of the problem. It is shown that $F(\eta)$ in equation (8) has the following properties [18]

(a) $F(\eta)$ is convex over $\mathbb{R}$,

(b) $F(\eta)$ is continuous over $\mathbb{R}$,

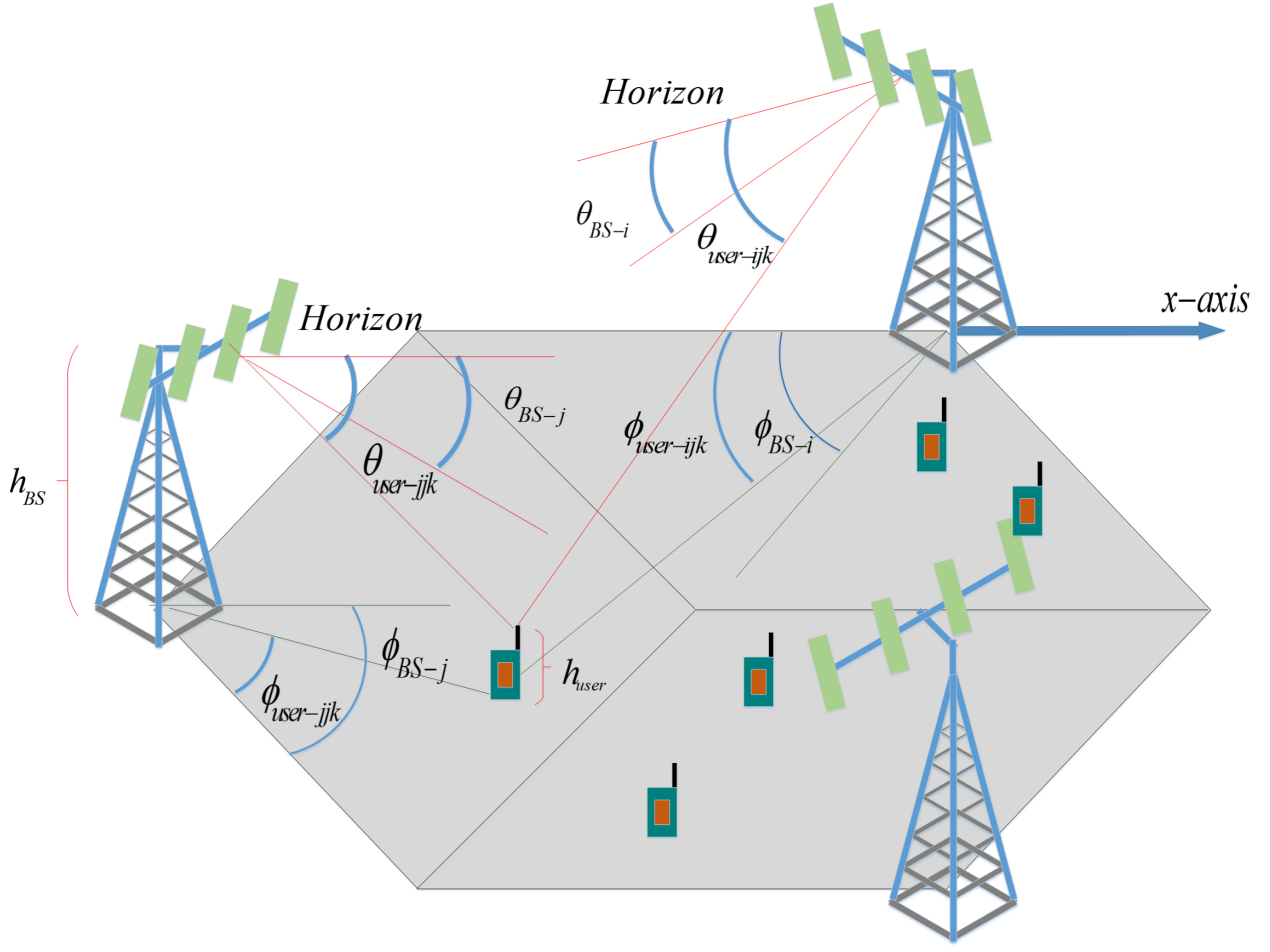(c) $F(\eta)$ is strictly decreasing,

Figure 1: Multi-cell setup with three cells and illustration of 3D angles.

(d) The equation $F(\eta) = 0$ always has a unique solution.

Thus, we can conclude that the following statements are equivalent [18]:

(a) $\displaystyle\max_{(\boldsymbol{W},\boldsymbol{\theta})\in\mathbb{D}} f(\boldsymbol{W},\boldsymbol{\theta}) = \max_{(\boldsymbol{W},\boldsymbol{\theta})\in\mathbb{D}} \frac{f_1(\boldsymbol{W},\boldsymbol{\theta})}{f_2(\boldsymbol{W},\boldsymbol{\theta})} = \eta$

(b) $\displaystyle F(\eta) = \max_{(\boldsymbol{W},\boldsymbol{\theta})\in\mathbb{D}} \{f_1(\boldsymbol{W},\boldsymbol{\theta}) - \eta f_2(\boldsymbol{W},\boldsymbol{\theta})\} = 0.$

In other words, solving univariate equation $F(\eta) = 0$ is equivalent to the maximization problem in (6). It means that if one can find an η such that the optimum value of problem (8) is zero, then this optimum value will be also a solution to the problem (6). To solve it, the problem can be divided into two parts or layers. The so-called outer layer searches for the optimum value of η while the inner layer solves the equation (9) for each value of η to find the optimum beamforming vectors and tilt angle.

For solving the univariate equation (8), we can employ the bisection algorithm as follows

Algorithm 1 Outer Layer Algorithm
___
1: **Initialize:** $\eta_{\min} = 0$, $\eta_{\max} = \frac{R_{\max}}{MLP_c + LP_0}$, $\varepsilon = 10^{-3}$

2: **while** $(|\eta_{\max} - \eta_{\min}| \geqslant \varepsilon)$ **do**

3: assign $\eta = \frac{\eta_{\min} + \eta_{\max}}{2}$, now solve (9) to find the optimum values of $(\boldsymbol{W}^{opt}, \boldsymbol{\theta}^{opt})$ and calculate $F(\eta)$.

4: **if** $(F(\eta) > 0)$ **then** $\eta_{max} = \eta$

5: **else** $\eta_{\min} = \eta$
  end while
___

As we see in algorithm 1, for each value of η, the problem (9) needs to be solved. This optimization problem can also be written as

$$\arg \max_{\boldsymbol{W}, \boldsymbol{\theta}} G(\boldsymbol{W}, \boldsymbol{\theta}) = \left(\sum_{j,m} b_{jm} R_{jm} - \eta \xi \sum_{j,m} \|\boldsymbol{\omega}_{jm}\|^2 \right)$$

$$\text{subject to} \quad \sum_{m=1}^{K} \|\boldsymbol{\omega}_{jm}\|^2 \leqslant P \quad \forall j, \quad j = 1, 2, ..., L \tag{9}$$

$$0_{1 \times L} \prec \boldsymbol{\theta} \prec 90^o \times [1, 1, ..., 1]_{1 \times L}$$

In order to solve (9), we use the method introduced in [19]. According to this method, the data rate of users R_{jm} can be written in terms of achievable minimum mean square error (MMSE) of estimated data symbols in the UE's. This can be done by introducing optimum filter variables μ_{jm} which are defined by the following equation

$$\tilde{d}_{jm} = \mu_{jm}^* y_{jm} \tag{10}$$

where the $\tilde{d}_{jm}$ is the estimated data symbol in the m-th user in the j-th cell. For deriving the MSE, let us introduce $\tilde{e}_{jm}$ as

$$\tilde{e}_{jm} = \left(\mu_{jm}^* \sum_{i=1}^{L} \sum_{n=1}^{K} \boldsymbol{g}_{ijm}^H \boldsymbol{\omega}_{in} d_{in} \sqrt{\alpha(\theta_{BS,i})} + n_{jm} - d_{jm} \right). \tag{11}$$

Then we can write MSE in terms of the receiving filter coefficients and channel parameters as

$$e_{jm} = \mathbb{E} \left\{ \left(\tilde{d}_{jm} - d_{jm} \right) \left(\tilde{d}_{jm} - d_{jm} \right)^* \right\} = \mathbb{E} \left\{ \tilde{e}_{jm} \tilde{e}_{jm}^* \right\}$$

$$= |\mu_{jm}|^2 \left(\sum_{i=1}^{L} \sum_{n=1}^{K} \left| \boldsymbol{g}_{ijm}^H \boldsymbol{\omega}_{in} \right|^2 \alpha(\theta_{BS,i}) + 1 \right) \tag{12}$$

$$- \mu_{jm}^* \boldsymbol{g}_{jjm}^H \boldsymbol{\omega}_{jm} \sqrt{\alpha(\theta_{BS,j})} - \mu_{jm} \boldsymbol{\omega}_{jm}^H \boldsymbol{g}_{jjm} \sqrt{\alpha(\theta_{BS,j})} + 1.$$

In the sequel, we will use the following lemma [20]:

Lemma 1. Maximum value of function $f(\boldsymbol{S}) = -\text{Tr}(\boldsymbol{S}\boldsymbol{E}) + \log(\det(\boldsymbol{S})) + d$ which is the solution of following optimization problem. The solution of

$$\max_{\boldsymbol{S} \in \mathbb{C}^{d \times d}, \boldsymbol{S} \succ 0} f(\boldsymbol{S})$$

exists at

$$\boldsymbol{S}^{opt} = \boldsymbol{E}^{-1}$$

and the optimum value is

$$f(\boldsymbol{S}^{opt}) = \log(\det(\boldsymbol{E}^{-1})).$$

By using lemma 1, the optimization problem in (9) can be rewritten as

$$\max_{\boldsymbol{W}, \boldsymbol{U}, \boldsymbol{S}, \boldsymbol{\theta}} \text{H}(\boldsymbol{W}, \boldsymbol{U}, \boldsymbol{S}, \boldsymbol{\theta})$$

$$\text{subject to} \sum_{m=1}^{K} \|\boldsymbol{\omega}_{jm}\|^2 \leqslant P \quad \forall j, j = 1, 2, ..., L$$

$$\boldsymbol{0}_{1 \times L} \prec \boldsymbol{\theta} \prec 90^o \times [1, 1, ..., 1]_{1 \times L} \tag{13}$$

where new objective function $\text{H}(\boldsymbol{W}, \boldsymbol{U}, \boldsymbol{S}, \boldsymbol{\theta})$ and set of filter variables $\boldsymbol{U}$ slack variables $\boldsymbol{S}$ are defined as

$$\text{H}(\boldsymbol{W}, \boldsymbol{U}, \boldsymbol{S}, \boldsymbol{\theta}) = \sum_{j,m} \left(-b_{jm} e_{jm} s_{jm} + b_{jm} \log(s_{jm}) \right)$$

$$+ \sum_{j,m} \left(b_{jm} - \eta \xi \|\boldsymbol{\omega}_{jm}\|^2 \right)$$

$$\boldsymbol{U} = \{\mu_1, \mu_2, ..., \mu_L\} \quad \text{and} \quad \mu_j = \{\mu_{j1}, \mu_{j2}, ..., \mu_{jK}\}$$

$$\boldsymbol{S} = \{s_1, s_2, ..., s_L\} \quad \text{and} \quad s_j = \{s_{j1}, s_{j2}, ..., s_{jK}\} \tag{14}$$

9

Now by substituting (12) in (13), we get

$$
\begin{aligned}
\max_{\boldsymbol{W}_j, \theta_j} (&- \sum_{m=1}^{K} \sum_{i=1}^{L} \sum_{n=1}^{K} b_{in} s_{in} |\mu_{in}|^2 |\boldsymbol{g}_{jin}^H \boldsymbol{\omega}_{jm}|^2 \alpha\left(\theta_{BS,j}\right) \\
&+ \sum_{m=1}^{K} b_{jm} s_{jm} \mu_{jm}^* \boldsymbol{g}_{jjm}^H \boldsymbol{\omega}_{jm} \sqrt{\alpha\left(\theta_{BS,j}\right)}) \\
&+ (\sum_{m=1}^{K} b_{jm} s_{jm} \mu_{jm} \boldsymbol{\omega}_{jm}^H \boldsymbol{g}_{jjm} \sqrt{\alpha\left(\theta_{BS,j}\right)} - \eta \xi \|\boldsymbol{\omega}_{jm}\|^2)
\end{aligned}
$$

$$
\text{subject to } \sum_{m=1}^{K} \|\boldsymbol{\omega}_{jm}\|^2 \leqslant P \quad \forall j, j = 1, 2, ..., L
$$

$$
\mathbf{0}_{1 \times L} \prec \boldsymbol{\theta} \prec 90^o \times [1, 1, ..., 1]_{1 \times L}
$$

(15)

By substituting (12) in (13) and (14), the optimization problem of (13) decouples among the BSs and each BS must solve its own problem. However, sharing of filter coefficients and slack variables through backhaul links is still necessary.

Theorem 1. For each value of $\boldsymbol{W}$ and $\boldsymbol{\theta}$, the optimum values of $\boldsymbol{U}$ and $\boldsymbol{S}$ which maximizes (15) are

$$
\mu_{jm}^{opt} = \frac{\boldsymbol{g}_{jjm}^H \boldsymbol{\omega}_{jm} \sqrt{\alpha\left(\theta_{BS,j}\right)}}{\sum_{i=1}^{L} \sum_{n=1}^{K} |\boldsymbol{g}_{ijm}^H \boldsymbol{\omega}_{in}|^2 \sqrt{\alpha\left(\theta_{BS,i}\right)} + 1}
$$

$$
\frac{1}{\hat{e}_{jm}} = 1 - \frac{|\boldsymbol{g}_{jjm}^H \boldsymbol{\omega}_{jm}|^2 \alpha\left(\theta_{BS,j}\right)}{\sum_{i=1}^{L} \sum_{n=1}^{K} |\boldsymbol{g}_{ijm}^H \boldsymbol{\omega}_{in}|^2 \sqrt{\alpha\left(\theta_{BS,i}\right)} + 1}
$$

$$
s_{jm}^{opt} = \frac{1}{\hat{e}_{jm}}
$$

(16)

Proof. Similar to lemma 1, to find the optimum value, two equations $\frac{\partial \mathrm{H}(\boldsymbol{W}, \boldsymbol{U}, \boldsymbol{S}, \boldsymbol{\theta})}{\partial s_{j,k}} = 0$ and $\frac{\partial \mathrm{H}(\boldsymbol{W}, \boldsymbol{U}, \boldsymbol{S}, \boldsymbol{\theta})}{\partial \mu_{j,k}^*} = 0$ must be solved which results in (16). $\qquad \square$

Clearly as Theorem 1 states, if we have optimum values of $\boldsymbol{W}$ and $\boldsymbol{\theta}$ then the optimum values of $\boldsymbol{U}$ and $\boldsymbol{S}$ are derived. The optimization problem of (13) is equivalant to finding the optimum values of $\boldsymbol{W}$ and $\boldsymbol{\theta}$ in (15). Now with help of Theorem 1, we aim to solve (15) using its dual problem (Lagrangian method) and block coordinate descent method.

Now we construct Lagrangian dual problem as

$$
L\left(\boldsymbol{W}_j, \lambda_j, \theta_{BS,j}\right) =
$$
$$
-\sum_{m=1}^{K}\left(\sum_{i=1}^{L}\sum_{n=1}^{K} b_{in}s_{in}|\mu_{in}|^2\left|\boldsymbol{g}_{jin}^H\boldsymbol{\omega}_{jm}\right|^2\alpha\left(\theta_{BS,j}\right) - b_{jm}s_{jm}\mu_{jm}^*\boldsymbol{g}_{jjm}^H\boldsymbol{\omega}_{jm}\sqrt{\alpha\left(\theta_{BS,j}\right)}\right)
$$
$$
+\left(\sum_{m=1}^{K}\left(b_{jm}s_{jm}\mu_{jm}\boldsymbol{\omega}_{jm}^H\boldsymbol{g}_{jjm}\sqrt{\alpha\left(\theta_{BS,j}\right)} - \eta\zeta\|\boldsymbol{\omega}_{jm}\|^2\right)\right) - \lambda_j\left(\sum_{m=1}^{K}\|\boldsymbol{\omega}_{jm}\|^2 - P\right). \tag{17}
$$

Using these two equalities [21]

$$
\frac{\partial\|\boldsymbol{\omega}_{jm}\|^2}{\partial\boldsymbol{\omega}_{jm}} = 2\boldsymbol{\omega}_{jm},
$$
$$
\frac{\partial\left|\boldsymbol{g}_{jin}^H\boldsymbol{\omega}_{jm}\right|^2}{\partial\boldsymbol{\omega}_{jm}} = \frac{\partial\left(\boldsymbol{\omega}_{jm}^H\boldsymbol{g}_{jin}\boldsymbol{g}_{jin}^H\boldsymbol{\omega}_{jm}\right)}{\partial\boldsymbol{\omega}_{jm}} = 2\boldsymbol{g}_{jin}\boldsymbol{g}_{jin}^H\boldsymbol{\omega}_{jm}
$$

and also using KKT conditions, the optimum solution of (17) for the beamforming vectors at a given θ can be calculated as follows

$$
\frac{\partial L\left(\boldsymbol{W}_j, \lambda_j, \theta_j\right)}{\partial\boldsymbol{\omega}_{jm}} = -2\sum_{i=1}^{L}\sum_{n=1}^{K} b_{in}s_{in}|\mu_{in}|^2\boldsymbol{g}_{jin}\boldsymbol{g}_{jin}^H\boldsymbol{\omega}_{jm}\alpha\left(\theta_{BS-j}\right)
$$
$$
+ 2b_{jm}s_{jm}\mu_{jm}\boldsymbol{g}_{jjm}\sqrt{\alpha\left(\theta_{BS-j}\right)} - 2\eta\zeta\boldsymbol{\omega}_{jm} = 0
$$
$$
\Rightarrow \left(\sum_{i=1}^{L}\sum_{n=1}^{K} b_{in}s_{in}|\mu_{in}|^2\boldsymbol{g}_{jin}\boldsymbol{g}_{jin}^H\alpha\left(\theta_{BS-j}\right) + \eta\zeta\boldsymbol{I}_M\right)\boldsymbol{\omega}_{jm} = b_{jm}s_{jm}\mu_{jm}\boldsymbol{g}_{jjm}\sqrt{\alpha\left(\theta_{BS-j}\right)}.
$$
$$
\tag{18}
$$

This yields a solution as follows

$$
\frac{\partial L\left(\boldsymbol{W}_j, \lambda_j, \theta_j\right)}{\partial\boldsymbol{\omega}_{jm}} = 0
$$
$$
\Rightarrow \boldsymbol{\omega}_{jm} = b_{jm}s_{jm}\mu_{jm}(\boldsymbol{A}_j + \lambda_j\boldsymbol{I}_M)^\dagger\boldsymbol{g}_{jjm}\sqrt{\alpha\left(\theta_{BS,j}\right)}
$$
$$
\Rightarrow \boldsymbol{\omega}_{jm} = b_{jm}s_{jm}\mu_{jm}(\boldsymbol{A}_j + \lambda_j^*\boldsymbol{I}_M)^\dagger\boldsymbol{g}_{jjm}\sqrt{\alpha\left(\theta_{BS,j}\right)} \tag{19}
$$

where $\boldsymbol{A}_j = \sum_{i=1}^{L}\sum_{n=1}^{K} b_{in}s_{in}|\mu_{in}|^2\boldsymbol{g}_{jin}\boldsymbol{g}_{jin}^H\alpha\left(\theta_{BS,j}\right) + \eta\zeta\boldsymbol{I}_M$. In (19) the optimum value of Lagrangian's multiplier λ_j must be found through one dimensional search which is a common method in optimization via dual problem [16, 19]. We denote this optimum value with λ_j^*. As stated earlier, we see in (19) that optimum values of beamforming vectors are dependent on both channel gains and users' elevations. Therefore finding optimum value of the tilt

angle is inevitable. Unfortunately (17) is a non-convex function in terms of the tilt angle. One way to find the optimum value of the tilt angle is through one dimensional search but this induces a heavy computational task on network. In the next section a new algorithm is proposed for reducing the complexity significantly.

4. Clustering Algorithm

There are some local maximums for θ in the maximization problem (17). Finding the global maximum is possible by searching feasible set of tilt angle $[\theta_{\max}, \theta_{\min}]$ where $\theta_{\max}$ and $\theta_{\min}$ are the maximum and minimum AoA of the users in the cell, respectively [3, 6, 7, 8]. This approach is reasonable when we optimize an ergodic utility which varies very slowly with time. However this is not the case in our setup since channel of the users is varying with time and we need to maximize (17) in each transmission interval. In the following theorem, we consider reducing this search interval.

Theorem 2. Optimum value of $\theta_{BS,j}$ in (17) exists in a symmetric interval of length $2 \times \frac{\theta_{3dB}}{\sqrt{2.4 \ln 10}}$ centered around one of users' AoA.

Proof. As we know, at the location of maximum of a function, its second derivative is negative. Furthermore (17) is weighted sum of shifted versions of (1). By taking the second derivative of (1) with respect to $\theta_{BS,j}$, we can find the interval in which the second derivative of the antenna pattern is negative and hence this interval contains the maximum of (17). This is due to the fact that only in a specific interval around AoA of each user the second derivative of (17) is negative. In othe words we have

$$
\begin{aligned}
\frac{d^2 a(\theta_{BS})}{d\theta_{BS}^2} &= A_{max} \left(\frac{(2.4 \ln 10)^2}{\theta_{3dB}^4} (\theta_{BS} - \theta_{user})^2 - \frac{2.4 \ln 10}{\theta_{3dB}^2} \right) \\
&\times \exp\left(-1.2 \times \frac{(\theta_{BS} - \theta_{user})^2}{\theta_{3dB}^2} \log_2 10 \right) \leq 0 \\
&\Rightarrow \frac{(2.4 \ln 10)^2}{\theta_{3dB}^4} (\theta_{BS} - \theta_{user})^2 - \frac{2.4 \ln 10}{\theta_{3dB}^2} \leq 0 \\
&\Rightarrow \theta_{user} - \frac{\theta_{3dB}}{\sqrt{2.4 \ln 10}} \leq \theta_{BS} \leq \theta_{user} + \frac{\theta_{3dB}}{\sqrt{2.4 \ln 10}}
\end{aligned}
\tag{20}
$$

$\square$

Fig. 3 represents this fact for $\theta_{3dB} = 6°$ and two user case with equal gains in which the array's maximum gain is normalized to unity. Instead of searching whole feasible set for θ, only intervals spanned by clusters will be investigated. Clustering algorithm is presented as follow:

Outputs of algorithm (2) are tilt angle spanning interval of each cluster. It is enough to investigate spanning interval of clusters instead of whole feasible region of the tilt angle

---

Algorithm 2 Clustering Algorithm

1: Sort angles of users in each cell.

2: Put all users with angle difference less than $2 \times \frac{\theta_{3dB}}{\sqrt{2.4 \ln 10}}$ in the same cluster.

3: In each cluster users with maximum and minimum AoA represent cluster interval

4: **if** (a cluster contains just a single user) **then** only AoA of that user represents that cluster

---

to find global maximum of (17) for tilt angle. To further reduce computational task, first we consider the tilt angle be equal to AoA of users and we choose the AoA of the user that maximizes (17). We call the user that its AoA maximizes (17) compared to other users as chosen user and the cluster containing him as chosen cluster. Then global maximum for tilt angle exists in the chosen. In order to find this global maximum, chosen cluster must be investigated. with this simplification it is enough to search spanning interval of the chosen cluster instead of whole feasible set of tilt angle. The overall algorithm for solving (9) is shown in Algorithm 3 and its flowchart is presented in Fig. 2.

Algorithm 3

---

chosen_ user_ cluster(i) is the i-th angle in the cluster containing chosen user (we examine each cluster spanning with known steps, for example 0.1° steps)

1: **Initialize:** initialize step counter $n = 0$, beamforming vectors such that $\sum\limits_{m=1}^{K} \left\| \omega_{jm}^n \right\|^2 \leq P$, filter and slack variables $\mu_{jm}^n = 0, \quad s_{jm}^n = 0$, quit threshold $\delta = 10^{-3}$

2: run clustering algorithm and acquire clusters

3: **while** $\left(\left| G\left(\boldsymbol{W}^n, \boldsymbol{\theta}^n\right) - G\left(\boldsymbol{W}^{n-1}, \boldsymbol{\theta}^{n-1}\right) \right| \geq \delta \right)$ **do**

4:      $(n+1) \to n$ and update filters μ_{jm}^n using (16)

5:      update slack variables s_{jm}^n using (16)

6:      find chosen user (the user that considering its AoA as tilt, maximizes $G\left(\boldsymbol{W}, \boldsymbol{\theta}\right)$ compared to other user) and

7:      **for** $(i = 0$ to $i = \text{length(chosen\_user\_cluster)})$ **do**

8:          $\theta_j = \text{chosen\_user\_cluster}(i)$

9:          update bemforming vectors $\boldsymbol{\omega}_{jm}^n$ using (19)

10:         update G_{new} using (9)

11:         **if** $(G_{new} > G\left(\boldsymbol{W}^n, \boldsymbol{\theta}^n\right))$ **then** $\text{tilt}_j^{opt} = \quad \text{possible\_tilts}_j(i)$ and $\boldsymbol{W}_j^{opt} = \boldsymbol{W}^n$
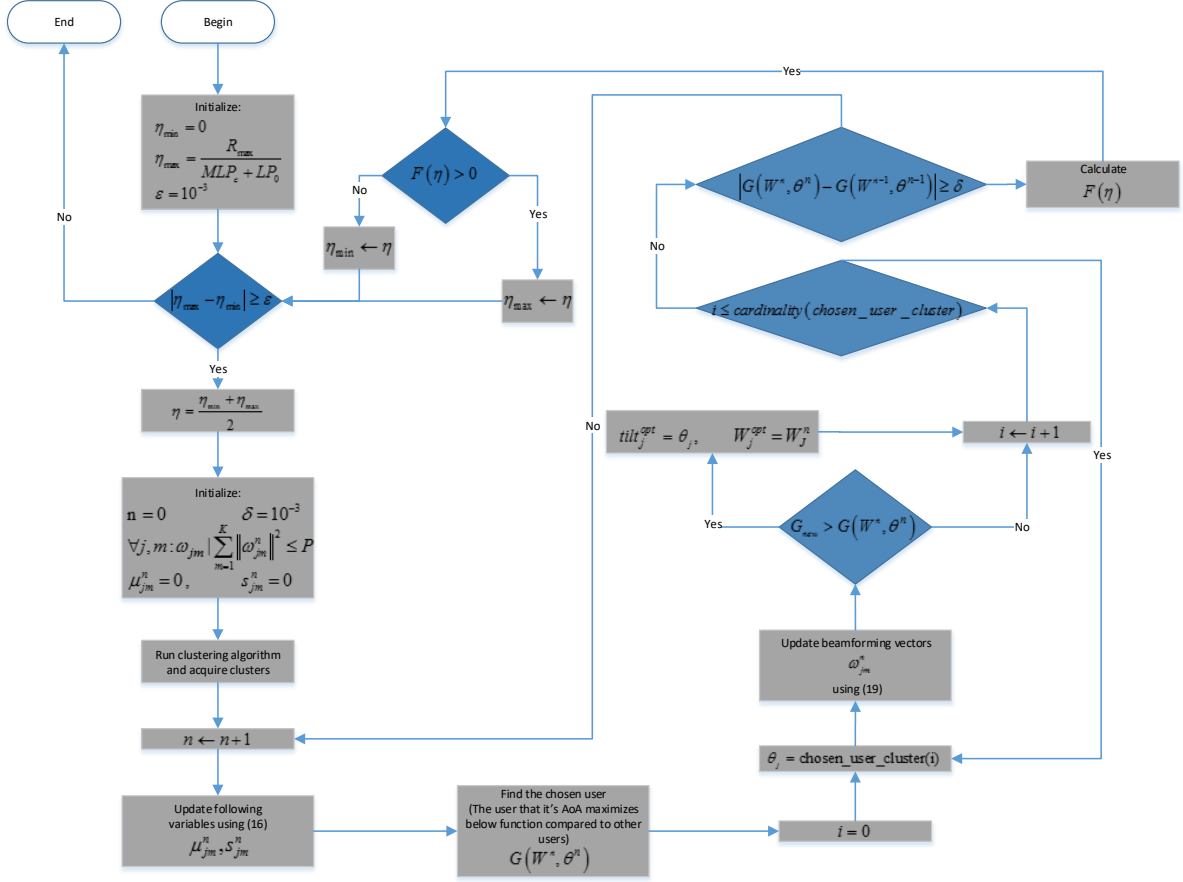   **end for**
 end while

Figure 2: Flowchart of the proposed algorithms.

5. Simulation Results

We consider three cooperative adjacent cells as main interferers, i.e. $L = 3$. Parameters of simulations are extracted from [4, 17]. Height of BSs and UEs are $h_{BS} = 32m$ and $h_{UE} = 1.5m$, respectively. Cell radius is set to be $R = 500m$. Channel vector model is $\boldsymbol{g}_{ijm} = \sqrt{\beta_{ijm}}\boldsymbol{h}_{ijm}$ where $\boldsymbol{h}_{ijm}$ is i.i.d. white gaussian random vector that models small scale fading and large scale fading coefficient $\beta_{ijm} = \frac{z_{ijm}}{d_{ijm}^v}$ compromise pathloss part d_{ijm}^v with $v = 3.8$ and shadow fading effect with log-normal random variable z_{ijm} with zero mean and 8 dB standard deviation. The RF-chain power consumption is $P_c = 30$ dB and BS site maintenance power consumption is $P_0 = 40$. Power amplifier inefficiency is considered to be $\xi = 1$ and all users have identical priority i.e. $b_{jm} = 1 \ \forall j, m$. Threshold values for stopping criteria of algorithms are set to $\delta = 10^{-3}$ and $\varepsilon = 10^{-3}$. Also for taking average over user configuration and random channel realization, 2500 Monte-Carlo simulations are done and the results are average of these simulations. Transmit power constraint of the BSs' is considered to vary in the interval $22 \sim 50$ dBm.

In Fig. 4 the total instantaneous EE for 3D case and its corresponding 2D case is presented. As it is expected by increasing the maximum transmit power in the BSs, the EE increases until it reaches a saturated level and after that by increasing available power the EE does not show variations. This is due to this fact that in low SNR regime (noise limited regime) logarithm function grows faster than linear function, thus by increasing transmit power in BS the EE will increase. On the other hand in high SNR regime (interference limited regime), logarithm function growth is lower than that of linear function, hence algorithm will not use the excess transmit power available at the BS and EE converges to a saturated level. It is observed, the 3D beamforming outperforms conventional 2D beamforming with isotropic BS's antenna array gain in the vertical plane.

Moreover by increasing the transmit power of the BSs, 3D beamforming performance gain over 2D beamforming increases as shown in Fig. 5. This shows the potential of the 3D beamforming method for mitigating intercell interference. After increasing the available transmit power to the level needed to overcome noise (crossing noise limited regime), limiting
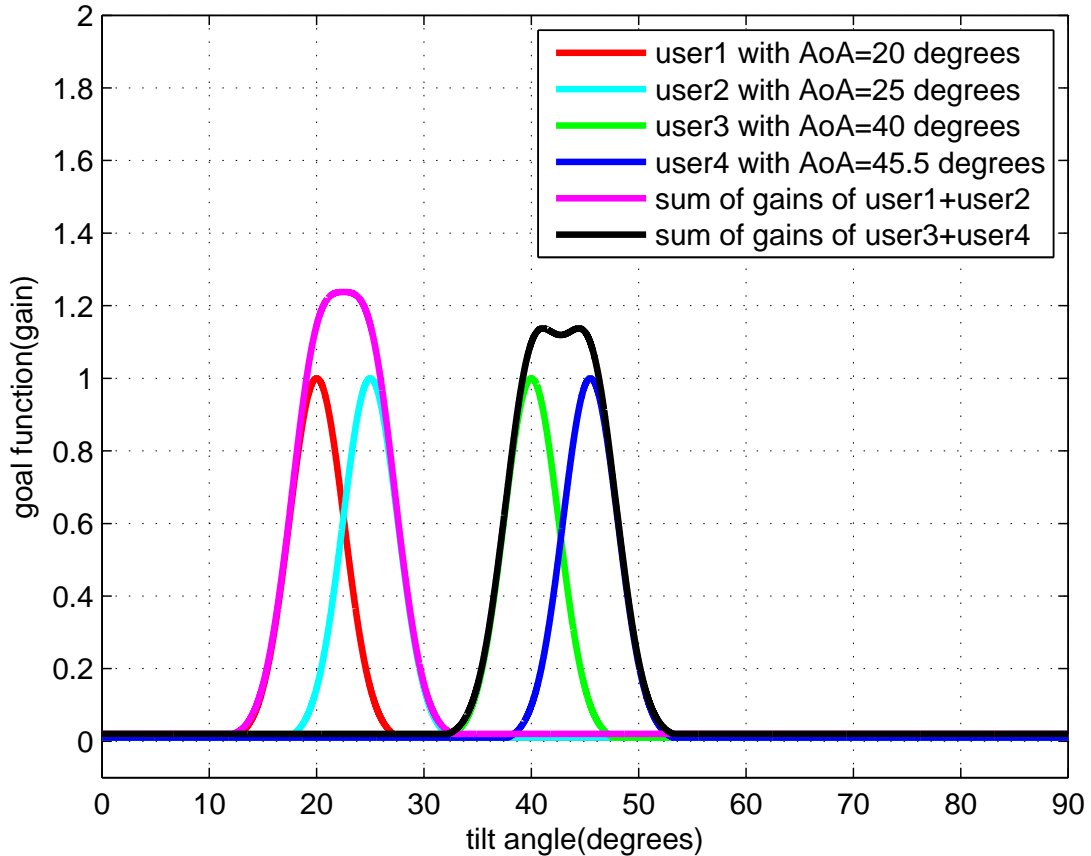

Figure 3: Clustering users based on their angle of arrival difference. (For example with $\theta_{3dB} = 6°$ clustering interval length will be $\frac{\theta_{3dB}}{\sqrt{2.4 \ln 10}} \approx 5.3$)
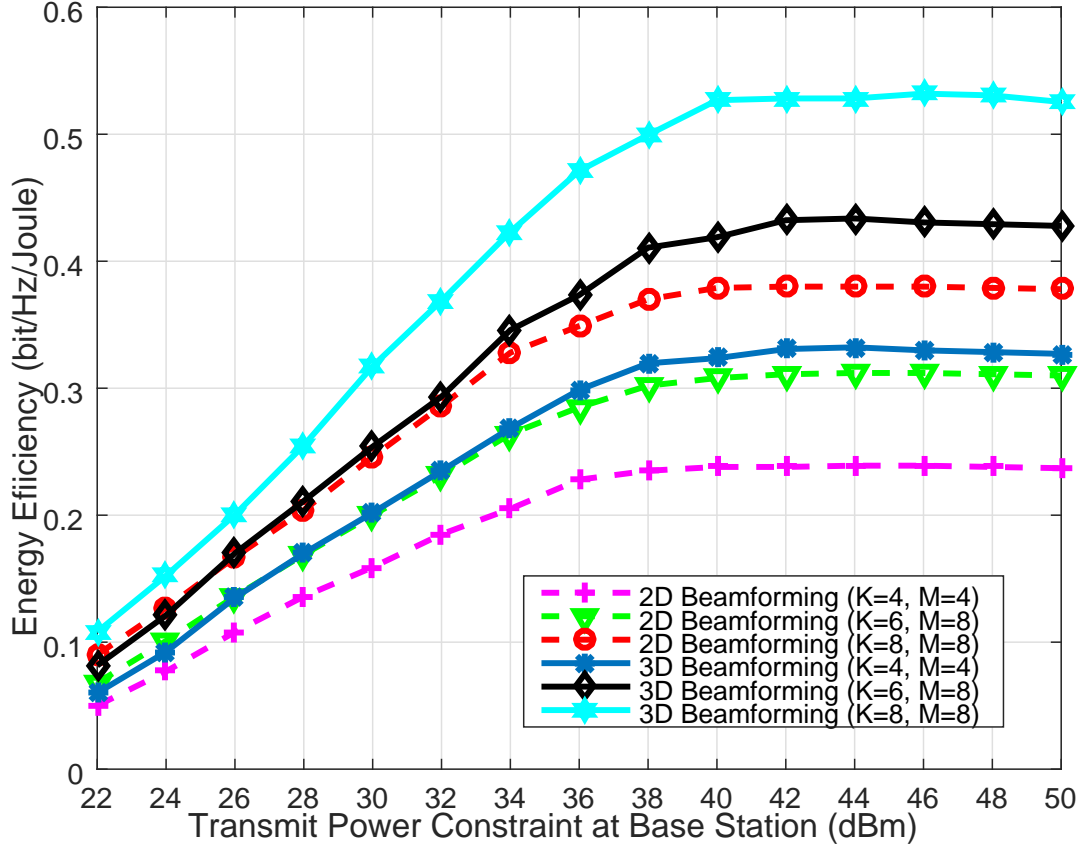
15

Figure 4: Comparison of 3D beamforming and conventional 2D beamforming for $M = 4$ and $M = 8$ number of antennas.

factor will become interference. The 3D beamforming gain over 2D beamforming increases because of the ability of interference mitigation. In other words 3D beamforming is more robust to intercell interference compared with 2D beamforming.

Finally Fig. 6 shows performance comparison of the proposed clustering algorithm and exhaustive search over tilt angle feasible region. The exhaustive search algorithm solves the beamforming problem by exploring over all possible values of tilt angles in all the BSs which results an exponential complexity. By contrast, the clustering algorithm reduces the complexity of the searches by limiting the interval of possible tilt angles in the BSs. In fact, in the clustering algorithm only some candidate angles are investigated.

We see that their performance is almost equal and a little gap appears between performances in high transmit power region. This is due to approximation considered in clustering algorithm when the chosen cluster contains only one user. The optimum tilt angle in clustering algorithm in this case is considered to be tilt angle of the chosen user, while in fact the optimum tilt angle can be around this user and an interval around this user should be searched. But as we see this approximation causes a little gap compared to the optimum

value but the computational complexity gain is considerable. It is notable that the performance gap appears in the interference limited regime where objective function is more sensitive to deviations from the optimum value while in noise limited regime sensitivity to deviations is low and hence there is no gap and no performance loss between two strategies.

Fig. 7 shows the effect of increasing the number of the users in each cell. As it can be seen, by increasing the number of the users total EE will increase in system. This implies that by adding more users, interference management capability of proposed method over the 2D technique will increase. However it is worth mentioning that by increasing the number of cells, performance gain will shrink. This is due to the fact that the number of degrees of freedom is controlled by the number of antennas and increasing number of cells only increases interference. Hence the system performance degrades. But this performance loss is insignificant due to sectoring cells and side lobe level loss of antenna patterns which heavily suppresses interference power.
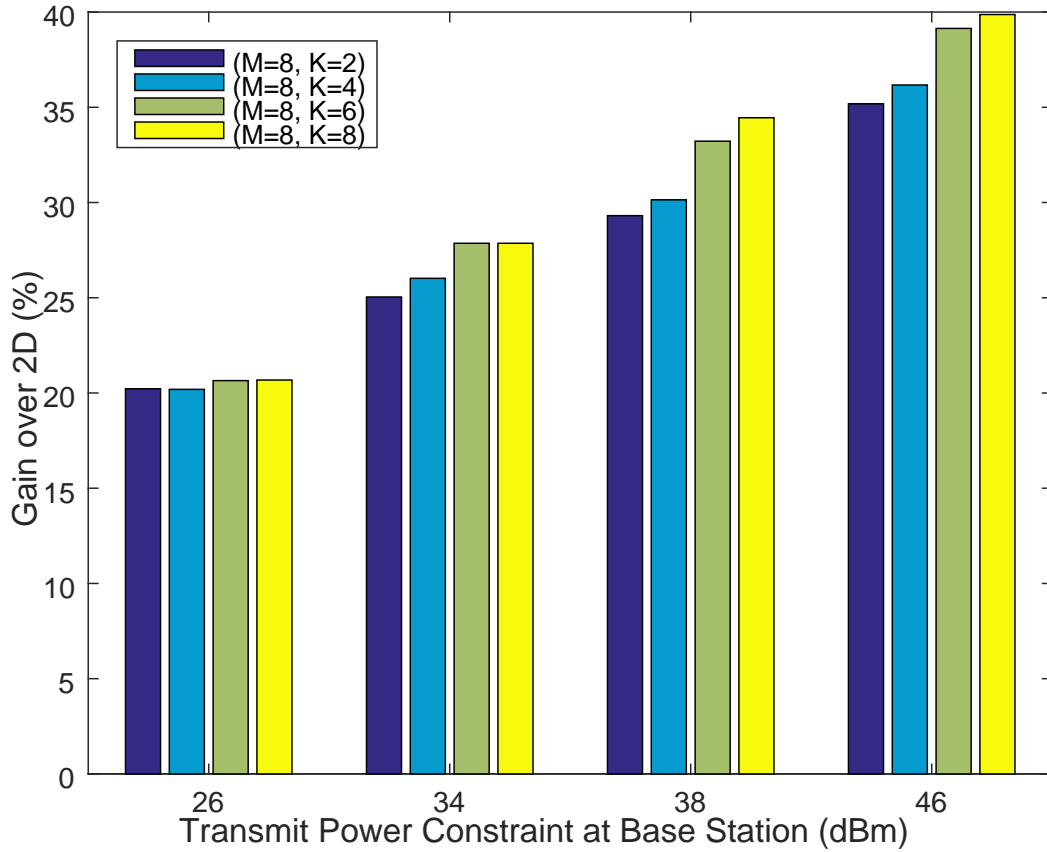

Figure 5: Gain of 3D beamforming over 2D beamforming expressed in percents for different values of transmit power available in BS.
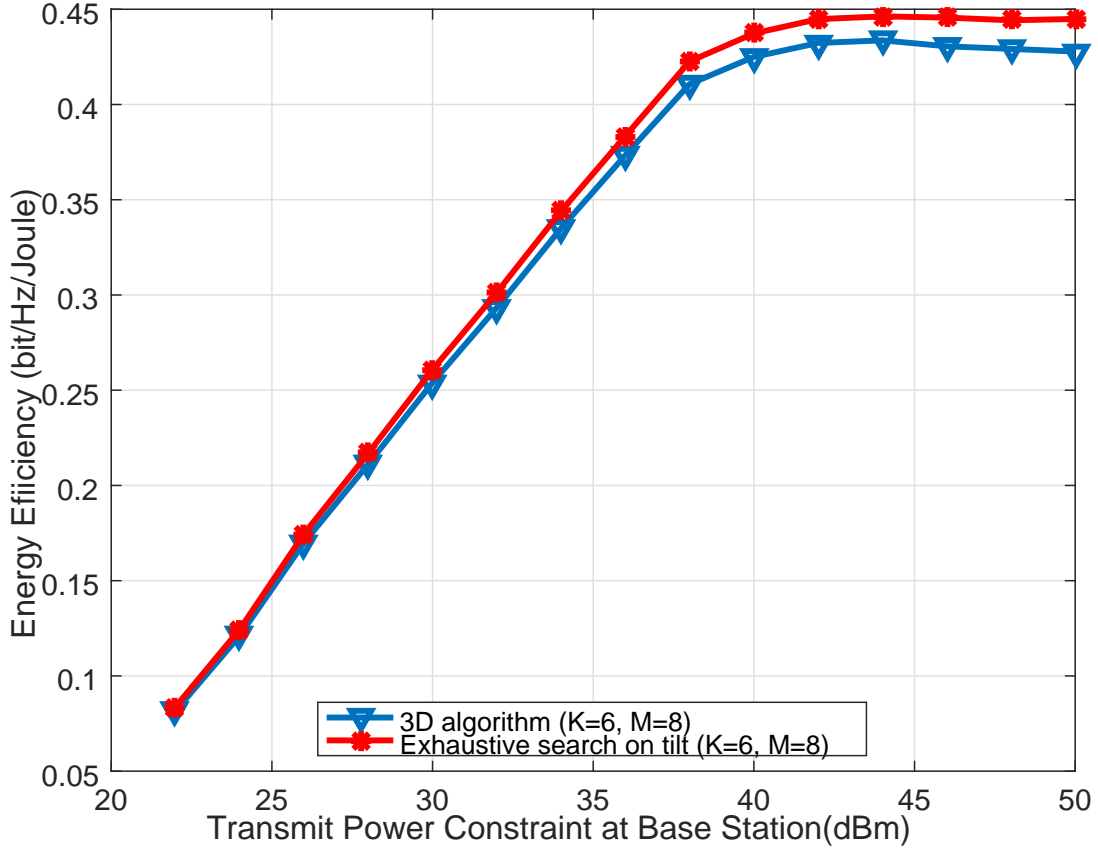
Figure 6: Performance comparison between clustering algorithm and exhaustive search over tilt angle of BS.

6. Conclusion

Downlink transmission in a cellular network with hexagonal shaped cells that simultaneously employs beamforming and tilt angle adaptation of the BSs' antenna array has been investigated. First by employing fractional programming and introducing new variables, problem of maximizing total instantaneous EE with respect to beamforming vectors and tilt angle of BSs' of the network has been converted into multiple tractable problems which are distributed between BSs. Closed form solution of the optimum beamforming vectors have been found which are functions of channel gain of users, users' angle of arrival and tilt angle of the BS. Then finding optimum tilt angle has been investigated which is a non-convex problem and has multiple maximums. To overcome this difficulty, the optimum tilt angle has been found efficiently using a clustering algorithm. This algorithm classifies users based on their angle of arrival to the BS's antenna array and then aims the array's beam to a specific cluster that maximizes total instantaneous EE of the network. The clustering algorithm has significantly less complexity compared to exhaustive search over tilt angle but its performance is near the same as exhaustive search.
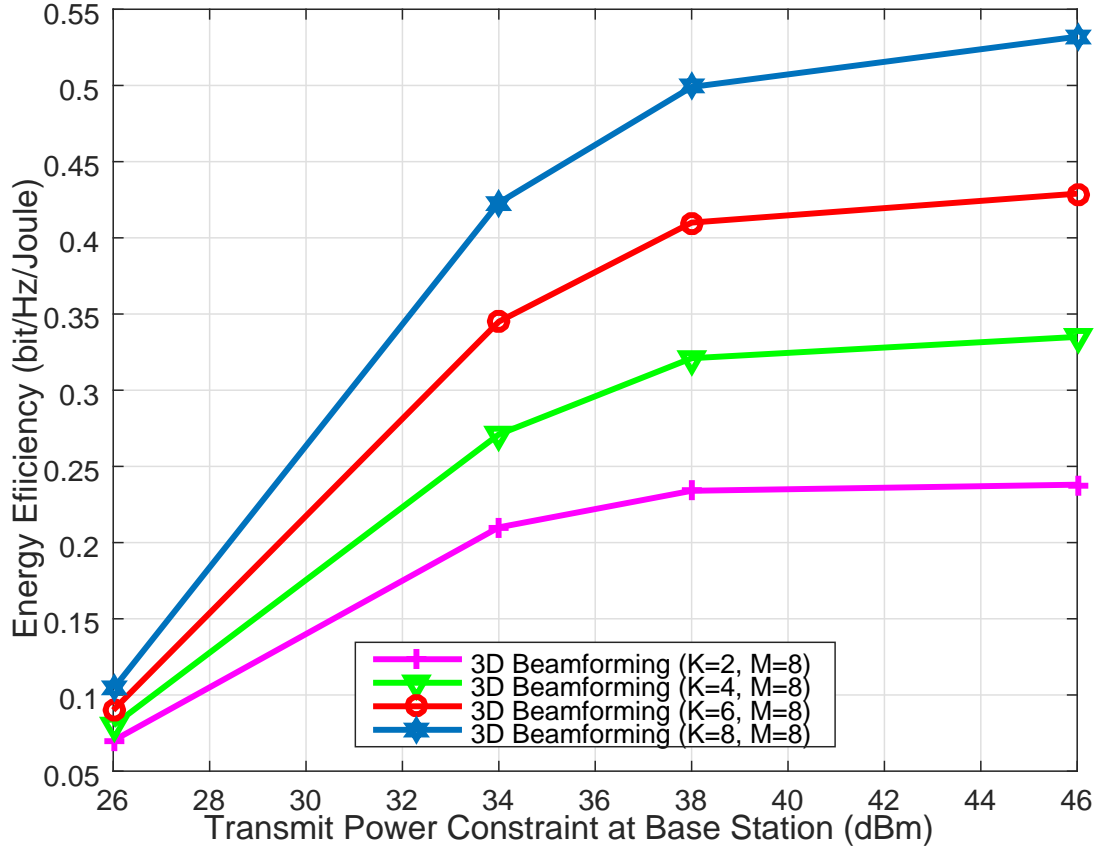
Figure 7: Impact of increasing the number of users in each cell.

References

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong and J. C. Zhang, "What Will 5G Be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, June 2014.

[2] S. M. Razavizadeh, M. Ahn and I. Lee, " Three Dimensional Beamforming: New Enabling Technology for Fifth Generation Wireless Networks," *IEEE Signal Proces. Mag.*, Vol.31, No.6, pp.94-101, Nov. 2014.

[3] W. Lee, S.-R. Lee, H.-B. Kong, S. Lee and I. Lee, "Downlink vertical beamforming designs for active antenna systems," *IEEE Trans. Commun.* , vol. 62, pp. 1897-1907, Jun. 2014.

[4] 3GPP TR 36.814 V9.0, "Further advancements for E-UTRA physical layer aspects," Tech. Rep., Mar. 2010.

[5] T. X. Tran and K. C. Teh, "Network Sum-Rate Optimization for Multicell Massive MIMO Downlink Based on Coordinated Tilt Adaptation and Game Theory Approach," *IEEE Wireless Commun. Letter*, vol. 5, no. 1, pp. 64-67, Feb. 2016.

[6] N. Seifi, J. Zhang, R. W. Heath, T. Svensson and M. Coldrey, "Coordinated 3D Beamforming for Interference Management in Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5396-5410, Oct. 2014.

[7] N. Seifi, R. Heath, M. Coldrey and T. Svensson, "Adaptive Multicell 3D Beamforming in Multi-Antenna Cellular Networks," *IEEE Trans. Veh. Technol.* , vol.PP, no.99, pp.1-1

[8] N. Seifi, M. Coldrey, and M. Viberg, "Throughput optimization in MISO interference channels via coordinated user-specific tilting," *IEEE Comm. Letter* , vol. 16, no. 8, pp. 1248-1251, August 2012.

[9] R. Hernandez-Aquino, S. A. R. Zaidi, D. McLernon, M. Ghogho and A. Imran, "Tilt Angle Optimization in Two-Tier Cellular NetworksA Stochastic Geometry Approach," *IEEE Trans. on Commun.* , vol. 63, no. 12, pp. 5162-5177, Dec. 2015.

[10] S. Jin, J. Wang, Q. Sun, M. Matthaiou and X. Gao, "Cell Coverage Optimization for the Multicell Massive MIMO Uplink," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5713-5727, Dec. 2015.

[11] Xiao Li, Shi Jin, Himal A. Suraweera, Jia Hou and Xiqi Gao, "Statistical 3-D Beamforming for Large-Scale MIMO Downlink Systems Over Rician Fading Channels," *IEEE Trans. on Commun.* , vol. 7, no. 12, pp. 4792 4799, Dec. 2008.

[12] F. Haddadi, M.M. Nayebi, and M.R. Aref, "Direction-of-Arrival Estimation for Temporally Correlated Narrowband Signals," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 600-609, Apr. 2009.

[13] M. Carlin, P. Rocca, G. Oliveri, F. Viani, and A. Massa, "Directions-of-Arrival Estimation Through Bayesian Compressive Sensing Strategies," *IEEE Trans. Antennas Propag.*, vol. 61, pp. 3828-3838, 2013.

[14] Y. Wang, X. Duan, D. Tian, J. Zhou, Y. Lu and G. Lu, "A Bayesian Compressive Sensing Vehicular Location Method Based on Three-Dimensional Radio Frequency," International Journal of Distributed Sensor Networks, vol. 2014, pp. 1-13, 2014.

[15] X. Li, S. Jin, X. Gao and R. W. Heath, "3D Beamforming for Large-scale FD-MIMO Systems Exploiting Statistical Channel State Information," *IEEE Trans. Veh. Technol.*, vol.PP, no.99, pp.1-1

[16] S. He, Y. Huang, S. Jin, and L. Yang, "Coordinated beamforming for energy efficient transmission in multicell multiuser systems," *IEEE Trans. Commun.* , vol. 61, no. 12, pp. 49614971, Dec. 2013.

[17] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types inheterogeneous cellular networks," *2010 Future Network & Mobile Summit* , Florence, 2010, pp. 1-8.

[18] J. Crouzeix, and J. Ferland, "Algorithms for generalized fractional programming," *Mathematical Programming* , vol. 52, pp. 191-207, Jun. 1991.

[19] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted sum rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.* , vol. 7, no. 12, pp. 4792 4799, Dec. 2008.

[20] J. Borwein and A. Lewis, *Convex analysis and nonlinear optimization: theory and examples.* Springer Verlag, 2006.

[21] S. Boyd and L. Vandenberghe, *Convex Optimization* Cambridge Univ. Press, 2004.

[22] A. Adhikary, J. Nam, J. Y. Ahn and G. Caire, "Joint Spatial Division and Multiplexing-The Large-Scale Array Regime," *IEEE Trans. Inform. Theory.*, vol. 59, no. 10, pp. 6441-6463, Oct. 2013.

Soheil Khavari received his B.Sc. from Amirkabir University of Technology (AUT), Tehran, Iran, in 2013 and his M.Sc. from Iran University of Science and Technology (IUST), Tehran, Iran in 2016, all in electrical engineering. His current research interests include wireless communication and signal processing.

S. Mohammad Razavizadeh is an assistant professor in the school of Electrical Engineering at Iran University of Science and Technology (IUST), Tehran, Iran. His research interests are in the area of signal processing for wireless communications and cellular networks. He is a Senior Member of the IEEE.