SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •
Special Focus on 5G Wireless Communication Networks

February 2016, Vol. 59 022307:1–022307:15
doi: 10.1007/s11432-015-5503-7

Sectorization based pilot reuse for improving net spectral efficiency in the multicell massive MIMO system

Jiahui LI[1,2,3], Limin XIAO[1,2,3], Xibin XU[1,2,3] & Shidong ZHOU[1,2,3]*

[1] *State Key Laboratory on Microwave and Digital Communications, Beijing* 100084*, China;*
[2] *Tsinghua National Laboratory for Information Science and Technology, Beijing* 100084*, China;*
[3] *Department of Electronic Engineering, Tsinghua University, Beijing* 100084*, China*

Received August 11, 2015; accepted November 24, 2015; published online January 5, 2016

Abstract A sectorization method using the uniform circular array (UCA) is proposed to improve the net spectral efficiency (SE) of the multicell massive MIMO system, which is an important index for evaluating the performance of a communication system. We derive the ergodic achievable uplink net SE per cell of a general sectorized system and obtain its deterministic approximation based on the large random matrix theory. Different weight matrices are considered for the sectorized system and the one with the best performance is utilized for further analysis. The consistency of the deterministic approximation with the result of Monte-Carlo simulation is proved at the same time. At last, numerical results indicate that the net SE per cell can be greatly improved compared to the conventional multicell massive MIMO system, which validates the effectiveness of the sectorization method. Moreover, comparisons with other pilot reuse methods are also made in this paper.

Keywords massive MIMO, multicell, net SE, pilot reuse, sectorization, UCA

Citation Li J H, Xiao L M, Xu X B, et al. Sectorization based pilot reuse for improving net spectral efficiency in the multicell massive MIMO system. Sci China Inf Sci, 2016, 59(2): 022307, doi: 10.1007/s11432-015-5503-7

1 Introduction

In a large scale multiple-input multiple-output (MIMO) or massive MIMO system, the base stations (BSs) are equipped with hundreds of antennas and tens of single-antenna users can be simultaneously served by each BS. Compared with conventional MIMO, massive MIMO can greatly improve the system throughput or link reliability [1], so it has attracted much greater attention recently.

Channel state information (CSI) plays a key role in the multiuser MIMO system, since it is utilized to get the transmission data for both uplink and downlink. Under time division duplex (TDD) mode, due to channel reciprocity, CSI can be obtained by only uplink training. To guarantee more accurate CSI for each user, orthogonal pilots are generally used in each cell, thus the length of pilots should be at least equal to the number of users in each cell [2]. As the number of users increases, a great proportion of channel coherence interval (defined as the product of coherence time and coherence bandwidth) will be spent on uplink training, which finally limits the system net spectral efficiency (SE).

* Corresponding author (email: zhousd@mail.tsinghua.edu.cn)

On the other hand, considering the multicell scenario [3], when different cells share the same set of pilots, it will cause pilot contamination among different cells, which introduces a bad effect on the accuracy of CSI. A simple solution is to use different sets of pilots which are mutually orthogonal among adjacent cells [4]. When the pilot reuse factor among different cells is 3 or bigger, the home cell can be surrounded by non-contaminating cells, but a much greater proportion of channel coherence interval will be spent on uplink training, so it is urgent to reduce the pilot overhead. In this paper, we focus on the pilot reuse among users in the same cell.

In [5], a pilot reuse scheme was proposed for users in the same cell, which utilizes the channel covariance matrix of all the users, and the users with the most different spatial features will share the same pilot. The spatial feature is corresponding to the signal subspace of the covariance matrix. The method can reduce the pilot overhead a lot and guarantee the channel estimation accuracy, thus improving the net SE. However, it requires to estimate the channel covariance matrices of all users in advance and execute the user grouping algorithm, which is of high computational complexity.

Under the condition when the signal angle spread of all users is small enough, there exists an exact relationship between the location of each user and its channel covariance matrix, and the covariance matrix is sensitive to the location of the user. A sectorization method proposed by [6] utilizes this feature. It partitions the cell equally into several sectors and each sector serves multiple users simultaneously. The net SE of each cell can be greatly improved due to pilot reuse among different sectors in each cell. An analysis under multicell scenario was done in [7], which considered full pilot reuse among different cells. However, it is not feasible for a practical massive MIMO system to utilize the uniform linear array (ULA) as described in [6, 7].

In this paper, the uniform circular array (UCA) will be used in the sectorized system, which has a much better symmetric property and can save much space for the deployment of the BS compared with the ULA. The basic idea of the sectorization method is introduced. Then we derive the ergodic achievable net SE per cell of the uplink under the general condition that the pilots are reused among different sectors in one cell and different cells, and its deterministic approximation is obtained by using the large random matrix analysis methods introduced in [8], which is validated using the Monte-Carlo simulation results under the condition of full pilot reuse. Several specific sectorization weight matrices are considered and the one with the best performance will be utilized for further analysis. At last, the performance of the sectorized system under multicell scenario is evaluated by comparing with that of other systems, including the conventional multicell massive MIMO system and the systems introduced in [5, 7].

Notations: Vectors are column vectors and denoted by lower case boldface and italic: $\boldsymbol{x}$. Matrices are upper case boldface: $\boldsymbol{A}$. $\boldsymbol{I}_M$ is the size-M identity matrix. The trace, transpose and Hermitian transpose are denoted by $\mathrm{tr}(\cdot)$, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$, respectively. The 2-norm of a vector $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\|$. $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ stands for the circular symmetric complex Gaussian distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$. $\mathrm{E}[\cdot]$ is the expectation operator and $\mathrm{E}[\cdot|\cdot]$ denotes the conditional expectation operator. $\xrightarrow[M\to\infty]{\text{a.s.}}$ denotes almost sure convergence of a stochastic sequence. $\Re(\cdot)$ is used to obtain the real part of a complex value.

2 System model

A multicell massive MIMO TDD system consisting of $L(L > 1)$ cells is considered. The BSs equipped with large M-antenna UCAs are deployed at the center of each cell, and each cell serves K single-antenna users. Assume that all the BSs and users are perfectly synchronized and channels for different users are independent of each other. Transmissions over flat fading channels are under consideration and the spatial correlation channel model [9, 10] is utilized.

2.1 Channel model

The uplink channel gain vector between the k-th $(k = 1, \ldots, K)$ user in the l-th $(l = 1, \ldots, L)$ cell and the j-th $(j = 1, \ldots, L)$ BS is modeled as

$$\boldsymbol{h}_{jlk} = \boldsymbol{R}_{jlk}^{\frac{1}{2}} \boldsymbol{v}_{jlk}, \tag{1}$$

where $\boldsymbol{v}_{jlk} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I}_M)$ is the spatially white channel vector, and $\boldsymbol{R}_{jlk} = \mathrm{E}[\boldsymbol{h}_{jlk}\boldsymbol{h}_{jlk}^{\mathrm{H}}]$ is the channel covariance matrix. For a UCA, we have $\boldsymbol{R}_{jlk} = \beta_{jlk}\mathrm{E}\{\boldsymbol{a}(\theta_{jlk})[\boldsymbol{a}(\theta_{jlk})]^{\mathrm{H}}\}$ (see [10] for details), where $\beta_{jlk} = zd_{jlk}^{-\gamma}$ is the slow fading coefficient, including pathloss $d_{jlk}^{-\gamma}$ (γ is the pathloss exponent) and shadow fading z, θ_{jlk} is a random variable with a certain distribution, describing the signal angle of arrival (AOA) from the k-th user in the l-th cell to the j-th BS, and

$$[\boldsymbol{a}(\theta)]_m = a_m(\theta) = a_1\left(\theta - \frac{2\pi}{M}(m-1)\right)\mathrm{e}^{-\mathrm{j}kr\cos\left(\theta - \frac{2\pi}{M}(m-1)\right)} \tag{2}$$

is the active element response of the m-th element ($m = 1, \ldots, M$) in the array [11], where $k = 2\pi/\lambda$ is the wavenumber and r is the radius of the UCA [1)].

2.2 Benchmark system

In the conventional system, users in each cell transmit mutually orthogonal pilots to obtain the CSI at the BS, and full pilot reuse is assumed among different cells. Such a system is referred to as the benchmark system in the following analysis.

2.2.1 *Uplink transmission*

For each use of the channel in uplink transmission, the signal vector received by the j-th BS is

$$\boldsymbol{y}_j = \sqrt{p_{\mathrm{ul}}}\boldsymbol{H}_{jj}\boldsymbol{x}_j + \sqrt{p_{\mathrm{ul}}}\sum_{l\neq j}\boldsymbol{H}_{jl}\boldsymbol{x}_l + \boldsymbol{n}_j, \tag{3}$$

where $\boldsymbol{H}_{jl} = [\boldsymbol{h}_{jl1}, \ldots, \boldsymbol{h}_{jlK}] \in \mathbb{C}^{M\times K}$ is the channel matrix between users in the l-th cell and the j-th BS, $\boldsymbol{x}_j \in \mathbb{C}^{K\times 1}$ is the transmitted signal vector of users in the j-th cell with i.i.d. zero-mean and unit-variance elements, and $\boldsymbol{n}_j \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I}_M)$ is the normalized noise vector at the receiver. The scalar p_{ul} denotes the uplink receive signal-to-noise ratio (SNR).

2.2.2 *Channel estimation*

During the uplink training phase, the users in each cell transmit mutually orthogonal pilots so as to estimate the channel matrix at the corresponding BS. The same set of pilots are reused in all the cells, so the channel estimate is interfered by training signals from other cells.

By correlating the received training signal with the pilot of the k-th user in the j-th cell, the j-th BS estimates the channel vector $\boldsymbol{h}_{jjk}$ based on the observation $\boldsymbol{y}_{jk}^{\mathrm{tr}} \in \mathbb{C}^{M\times 1}$, given as

$$\boldsymbol{y}_{jk}^{\mathrm{tr}} = \boldsymbol{h}_{jjk} + \sum_{l\neq j}\boldsymbol{h}_{jlk} + \frac{1}{\sqrt{p_{\mathrm{tr}}}}\boldsymbol{n}_{jk}^{\mathrm{tr}}, \tag{4}$$

where $\boldsymbol{n}_{jk}^{\mathrm{tr}} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I}_M)$ is the normalized training noise vector and p_{tr} is the effective training SNR, which depends on the pilot transmit power and the length of pilots. Assume that p_{tr} is a given parameter. The MMSE estimate of $\boldsymbol{h}_{jjk}$ is given by [8]

$$\hat{\boldsymbol{h}}_{jjk} = \boldsymbol{R}_{jjk}\boldsymbol{Q}_{jk}\boldsymbol{y}_{jk}^{\mathrm{tr}}, \tag{5}$$

which is distributed as $\hat{\boldsymbol{h}}_{jjk} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Phi}_{jjk})$. Here we have

$$\boldsymbol{\Phi}_{jlk} = \boldsymbol{R}_{jjk}\boldsymbol{Q}_{jk}\boldsymbol{R}_{jlk}, \ \forall j, l, k, \tag{6}$$

$$\boldsymbol{Q}_{jk} = \left(\sum_{l}\boldsymbol{R}_{jlk} + \frac{1}{p_{\mathrm{tr}}}\mathbf{I}_N\right)^{-1}, \ \forall j, k. \tag{7}$$

Define estimation error as $\tilde{\boldsymbol{h}}_{jjk} = \boldsymbol{h}_{jjk} - \hat{\boldsymbol{h}}_{jjk}$. Due to the joint Gaussianity of both the estimation and the estimation error vectors, $\tilde{\boldsymbol{h}}_{jjk} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{R}_{jjk} - \boldsymbol{\Phi}_{jjk})$ is independent of $\hat{\boldsymbol{h}}_{jjk}$.

1) In this paper, we only consider the azimuth domain, and the extension to both the azimuth and elevation domains can be easily achieved by replacing the UCA with a uniform cylindrical array.

2.2.3 *Achievable uplink spectral efficiency*

Consider the MMSE detection of $\boldsymbol{x}_j$. The detector for the k-th user in the j-th cell is

$$\boldsymbol{r}_{jk} = \left(\hat{\boldsymbol{H}}_{jj}\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}} + \boldsymbol{Z}_j + \frac{1}{p_{\mathrm{ul}}}\boldsymbol{I}_M \right)^{-1} \hat{\boldsymbol{h}}_{jjk}, \tag{8}$$

where $\boldsymbol{Z}_j$ is the covariance matrix of channel estimation error and inter-cell interference, given as

$$\begin{aligned} \boldsymbol{Z}_j &= \mathrm{E}\left[\tilde{\boldsymbol{H}}_{jj}\tilde{\boldsymbol{H}}_{jj}^{\mathrm{H}} + \sum_{l \neq j} \boldsymbol{H}_{jl}\boldsymbol{H}_{jl}^{\mathrm{H}} \right] \\ &= \sum_{m} (\boldsymbol{R}_{jjm} - \boldsymbol{\Phi}_{jjm}) + \sum_{l \neq j}\sum_{m} \boldsymbol{R}_{jlm}. \end{aligned} \tag{9}$$

Using a standard bound given by [12] based on the worst-case uncorrelated Gaussian noise, the ergodic achievable uplink SE of the k-th user in the j-th cell is given by

$$C_{jk}^{\mathrm{bk}} = \mathrm{E}\left[\log_2\left(1 + \gamma_{jk}^{\mathrm{bk}}\right) \right], \tag{10}$$

where $\gamma_{jk}^{\mathrm{bk}}$ denotes the signal-to-interference-plus-noise ratio (SINR), given by

$$\gamma_{jk}^{\mathrm{bk}} = \frac{|\boldsymbol{r}_{jk}^{\mathrm{H}}\hat{\boldsymbol{h}}_{jjk}|^2}{\mathrm{E}\left[\boldsymbol{r}_{jk}^{\mathrm{H}} \left(\tilde{\boldsymbol{h}}_{jjk}\tilde{\boldsymbol{h}}_{jjk}^{\mathrm{H}} + \sum_l \boldsymbol{H}_{jl}\boldsymbol{H}_{jl}^{\mathrm{H}} - \boldsymbol{h}_{jjk}\boldsymbol{h}_{jjk}^{\mathrm{H}} + \frac{1}{p_{\mathrm{ul}}}\boldsymbol{I}_N \right) \boldsymbol{r}_{jk} \middle| \hat{\boldsymbol{H}}_{jj} \right]}. \tag{11}$$

The superscript "bk" refers to the benchmark system. The deterministic approximation of $\gamma_{jk}^{\mathrm{bk}}$, denoted by $\bar{\gamma}_{jk}^{\mathrm{bk}}$, can be found in [8, Eq. (25)]. Considering the overhead of pilot, we can get the achievable net SE of the j-th cell,

$$R_j^{\mathrm{bk}} = \left(1 - \frac{\tau^{\mathrm{bk}}}{T_{\mathrm{c}}}\right)\sum_{k=1}^{K} C_{jk}^{\mathrm{bk}}, \tag{12}$$

where τ^{bk} is the pilot length for the benchmark system and T_{c} is the length of coherence interval.

3 Sectorization method

Instead of using the ULA like [6, 7], the UCA is considered in the sectorized system. Because the UCA is more symmetrical than the ULA, which means a more uniform angular resolution can be achieved by the UCA.

Consider a spatial filter for the UCA with the weight vector $\boldsymbol{\omega}_1 = [\omega_0, \omega_1, \ldots, \omega_{M-1}]^{\mathrm{H}} \in \mathbb{C}^{M\times 1}$, which is called the basic weight vector. The spatial response of the basic weight vector can be denoted as

$$u_1(\theta) = \boldsymbol{\omega}_1^{\mathrm{H}}\boldsymbol{a}(\theta) = \sum_{m=0}^{M-1} \omega_m a_1\left(\theta - \frac{2\pi}{M}(m-1)\right)\mathrm{e}^{-\mathrm{j}kr\cos\left(\theta - \frac{2\pi}{M}(m-1)\right)}. \tag{13}$$

Shift the elements of $\boldsymbol{\omega}_1$ by b times with shift step length s, another weight vector can be obtained:

$$\boldsymbol{\omega}_{b+1} = [\omega_{M-bs}, \ldots, \omega_{M-1}, \omega_0, \omega_1, \ldots, \omega_{M-bs-1}]^{\mathrm{H}}. \tag{14}$$

Thanks to the symmetric property of the UCA, the corresponding spatial response of $\boldsymbol{\omega}_{b+1}$ is given by

$$u_{b+1}(\theta) = u_1\left(\theta - \frac{2\pi bs}{M}\right), 1 \leqslant b \leqslant M/s, \tag{15}$$

which rotates in the azimuth domain and the rotation angle is proportional to the shift length bs. Therefore, by designing $\boldsymbol{\omega}_1$, we can get several beams uniformly distributed in the azimuth domain. Then we divide the azimuth domain into several sectors, and the beams are assigned to the corresponding sectors. In order to control inter-sector interference, the beams at the edge of each sector will be discarded. Assume that the weight matrix containing the weight vectors of the effective beams in the q-th ($q = 1, \ldots, Q$) sector is $\boldsymbol{W}_q = [\boldsymbol{\omega}_{(q-1)M/Q+i_{\min}}, \ldots, \boldsymbol{\omega}_{(q-1)M/Q+i_{\max}}] \in \mathbb{C}^{M\times B}(B = i_{\max}-i_{\min}+1, 1 \leqslant i_{\min} \leqslant i_{\max} \leqslant M/Q)$, where Q is the number of sectors, $i_{\min}$ is the minimum index for the effective beams in the first sector, and $i_{\max}$ is the maximum index for the effective beams in the first sector. Each sector serves K' users simultaneously, and the total user number is $K = QK'$. Other parameters such as the cell number and the BS antenna number are the same as described in Section 2. The signal received by the q-th sector of the j-th BS can be written as

$$\boldsymbol{y}_{jq} = \underbrace{\sqrt{p_{\mathrm{ul}}}\boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{H}_{jjq}\boldsymbol{x}_{jq}}_{\text{desired signal}} + \underbrace{\sum_{p\neq q}\sqrt{p_{\mathrm{ul}}}\boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{H}_{jjp}\boldsymbol{x}_{jp}}_{\text{inter-sector interference}} + \underbrace{\sum_{l\neq j}\sum_{p}\sqrt{p_{\mathrm{ul}}}\boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{H}_{jlp}\boldsymbol{x}_{lp}}_{\text{inter-cell interference}} + \boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{n}_j, \tag{16}$$

where $\boldsymbol{H}_{jlp} = [\boldsymbol{h}_{jlp1}, \ldots, \boldsymbol{h}_{jlpK'}] \in \mathbb{C}^{M\times K'}$ is the channel matrix between users in the p-th sector of the l-th cell and the j-th BS, $\boldsymbol{x}_{jp} \in \mathbb{C}^{K'\times 1}$ is the transmitted signal vector of users in the p-th sector of the j-th cell with i.i.d. zero-mean and unit-variance elements, and $\boldsymbol{n}_j \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M)$ is the normalized noise vector at the receiver.

4 Uplink spectral efficiency derivation with sectorization

In this section, the ergodic achievable uplink net SE of each cell in the sectorized system and its deterministic approximation will be derived. Considering the general condition, we assume that the pilot reuse factor among different cells is α and the pilot reuse factor among different sectors in the same cell is β. Users in each sector use orthogonal pilots, so the minimum length of the pilots is $\alpha\beta K'$. Let $\mathcal{C}_a$ ($a = 1, 2, \ldots, \alpha$) and $\mathcal{C}$ represent the index set of the cells in the a-th group and the index set of all the cells, respectively. Let $\mathcal{S}_b$ ($b = 1, 2, \ldots, \beta$) and $\mathcal{S}$ represent the index set of the sectors in the b-th group and the index set of all the sectors, respectively.

4.1 Channel estimation

During the training phase, all users transmit their pilot signals synchronously. The j-th BS estimates the channel vector $\boldsymbol{h}_{jjqk}$ of the k-th ($k = 1, 2, \ldots, K'$) user in the q-th ($q \in \mathcal{S}_b$) sector of the j-th ($j \in \mathcal{C}_a$) cell based on the observation $\boldsymbol{y}_{jqk}^{\mathrm{tr}} \in \mathbb{C}^N$, given as

$$\boldsymbol{y}_{jqk}^{\mathrm{tr}} = \boldsymbol{h}_{jjqk} + \sum_{p\in\mathcal{S}_b, p\neq q}\boldsymbol{h}_{jjpk} + \sum_{l\in\mathcal{C}_a, l\neq j}\sum_{p\in\mathcal{S}_b}\boldsymbol{h}_{jlpk} + \frac{1}{\sqrt{p_{\mathrm{tr}}}}\boldsymbol{n}_{jqk}^{\mathrm{tr}}, \tag{17}$$

where $\boldsymbol{n}_{jqk}^{\mathrm{tr}} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_N)$ denotes the normalized training noise vector. Under the MMSE estimation, the estimate of $\boldsymbol{h}_{jjqk}$ is written as

$$\hat{\boldsymbol{h}}_{jjqk} = \boldsymbol{R}_{jjqk}\boldsymbol{Q}_{jbk}\boldsymbol{y}_{jqk}^{\mathrm{tr}}, \tag{18}$$

here $\hat{\boldsymbol{h}}_{jjqk} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{\Phi}_{jqjqk})$ and we have

$$\boldsymbol{\Phi}_{jqlpk} = \boldsymbol{R}_{jjqk}\boldsymbol{Q}_{jbk}\boldsymbol{R}_{jlpk}, \ \forall a, b, k, \forall j, l \in \mathcal{C}_a, \forall q, p \in \mathcal{S}_b, \tag{19}$$

$$\begin{cases} \boldsymbol{R}_{jlpk} = \mathrm{E}[\boldsymbol{h}_{jlpk}\boldsymbol{h}_{jlpk}^{\mathrm{H}}], & \forall j, l, p, k, \\ \boldsymbol{Q}_{jbk} = \left(\displaystyle\sum_{l\in\mathcal{C}_a}\sum_{p\in\mathcal{S}_b}\boldsymbol{R}_{jlpk} + \frac{1}{p_{\mathrm{tr}}}\boldsymbol{I}_M\right)^{-1}, & \forall a, b, k, \forall j \in \mathcal{C}_a. \end{cases} \tag{20}$$

The channel estimation error is denoted by $\tilde{\boldsymbol{h}}_{jjqk} = \boldsymbol{h}_{jjqk} - \hat{\boldsymbol{h}}_{jjqk}$. Due to the joint Gaussianity of both vectors, $\tilde{\boldsymbol{h}}_{jjqk} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_{jjqk} - \boldsymbol{\Phi}_{jqjqk})$ is independent of $\hat{\boldsymbol{h}}_{jjqk}$.

4.2 Achievable uplink spectral efficiency

For simplicity, we define the effective channel matrix of the sectorized system as $\boldsymbol{G}_{jqlp} = \boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{H}_{jlp} = [\boldsymbol{g}_{jqlp1}, \dots, \boldsymbol{g}_{jqlpK'}] \in \mathbb{C}^{B\times K'}$. Therefore, the estimate of the effective channel matrix is denoted by $\hat{\boldsymbol{G}}_{jqlp} = \boldsymbol{W}_q^{\mathrm{H}}\hat{\boldsymbol{H}}_{jlp}$ and the estimation error of the effective channel matrix is denoted by $\tilde{\boldsymbol{G}}_{jqlp} = \boldsymbol{W}_q^{\mathrm{H}}\tilde{\boldsymbol{H}}_{jlp}$. Then the receive signal in (16) can be rewritten as

$$\boldsymbol{y}_{jq} = \sqrt{p_{\mathrm{ul}}}\hat{\boldsymbol{G}}_{jqjq}\boldsymbol{x}_{jq} + \boldsymbol{z}_{jq}, \tag{21}$$

where $\boldsymbol{z}_{jq}$ contains the estimation error, the interference and the noise, given by

$$\boldsymbol{z}_{jq} = \sqrt{p_{\mathrm{ul}}}\tilde{\boldsymbol{G}}_{jqjq}\boldsymbol{x}_{jq} + \sum_{p\neq q}\sqrt{p_{\mathrm{ul}}}\boldsymbol{G}_{jqjp}\boldsymbol{x}_{jp} + \sum_{l\neq j}\sum_{p}\sqrt{p_{\mathrm{ul}}}\boldsymbol{G}_{jqlp}\boldsymbol{x}_{lp} + \boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{n}_j. \tag{22}$$

During the uplink transmission phase, each BS detects the signals of the users in its cell. The detector for the k-th user in the q-th sector of the j-th cell is similar to the MMSE detector, which is written as

$$\boldsymbol{r}_{jqk} = \left(\hat{\boldsymbol{G}}_{jqjq}\hat{\boldsymbol{G}}_{jqjq}^{\mathrm{H}} + \boldsymbol{Z}_{jq} + \frac{c_q}{p_{\mathrm{ul}}}\boldsymbol{I}_N\right)^{-1}\hat{\boldsymbol{g}}_{jqjqk}, \tag{23}$$

where $c_q = \mathrm{tr}(\boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{W}_q)/B$ is a normalization factor and $\boldsymbol{Z}_{jq}$ is the covariance matrix of the channel estimation error, the inter-sector interference and the inter-cell interference, given by

$$\begin{aligned}\boldsymbol{Z}_{jq} =& \mathrm{E}\left[\tilde{\boldsymbol{G}}_{jqjq}\tilde{\boldsymbol{G}}_{jqjq}^{\mathrm{H}} + \sum_{p\neq q}\boldsymbol{G}_{jqjp}\boldsymbol{G}_{jqjp}^{\mathrm{H}} + \sum_{l\neq j}\sum_{p}\boldsymbol{G}_{jqlp}\boldsymbol{G}_{jqlp}^{\mathrm{H}}\right]\\ =& \sum_{m}(\boldsymbol{\Theta}_{jqjqm} - \Xi_{jqjqm}) + \sum_{p\neq q}\sum_{m}\boldsymbol{\Theta}_{jqjpm} + \sum_{l\neq j}\sum_{p}\sum_{m}\boldsymbol{\Theta}_{jqlpm}.\end{aligned} \tag{24}$$

The matrices used above are defined as follows:

$$\begin{cases}\boldsymbol{\Theta}_{jqlpk} = \boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{R}_{jlpk}\boldsymbol{W}_q, \quad \forall j,l,q,p,k,\\ \Xi_{jqlpk} = \boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{\Phi}_{jqlpk}\boldsymbol{W}_q, \ \forall a,b,k,\forall j,l\in\mathcal{C}_a,\forall q,p\in\mathcal{S}_b,\\ \qquad\quad = \boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{R}_{jjqk}\boldsymbol{Q}_{jbk}\boldsymbol{R}_{jlpk}\boldsymbol{W}_q.\end{cases} \tag{25}$$

Similar to (10), the ergodic achievable uplink SE of the k-th user in the q-th sector of the j-th cell is calculated by

$$C_{jqk}^{\mathrm{sr}} = \mathrm{E}\left[\log_2\left(1+\gamma_{jqk}^{\mathrm{sr}}\right)\right], \tag{26}$$

where $\gamma_{jqk}^{\mathrm{sr}}$ denotes the SINR, given by

$$\gamma_{jqk}^{\mathrm{sr}} = \frac{|\boldsymbol{r}_{jqk}^{\mathrm{H}}\hat{\boldsymbol{g}}_{jqjqk}|^2}{\mathrm{E}\left[\boldsymbol{r}_{jqk}^{\mathrm{H}}\left(\tilde{\boldsymbol{g}}_{jqjqk}\tilde{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}} + \sum_l\sum_p\boldsymbol{G}_{jqlp}\boldsymbol{G}_{jqlp}^{\mathrm{H}} - \boldsymbol{g}_{jqjqk}\boldsymbol{g}_{jqjqk}^{\mathrm{H}} + \frac{1}{p_{\mathrm{ul}}}\boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{W}_q\right)\boldsymbol{r}_{jqk}\middle|\hat{\boldsymbol{G}}_{jqjq}\right]}. \tag{27}$$

The superscript "sr" refers to the sectorized system. Considering the overhead of pilot, we can get the achievable net SE of the j-th cell:

$$R_j^{\mathrm{sr}} = \left(1-\frac{\tau^{\mathrm{sr}}}{T_{\mathrm{c}}}\right)\sum_{q=1}^{Q}\sum_{k=1}^{K'}C_{jqk}^{\mathrm{sr}}, \tag{28}$$

where τ^{sr} is the pilot length for the sectorized system.

4.3 Asymptotic analysis

In order to compute the net SE in closed form, we would like to obtain the deterministic approximation of the SINR, which is asymptotically tight when large system limit is considered [13]. Assume that $M, K \to \infty$ and the ratio of them is finite, then a tight approximation of the SINR can be obtained, which only depends on large scale fading, uplink SNR and pilot allocation scheme, while the SINR in (27) also depends on the stochastic small scale fading.

Before the asymptotic analysis, two related results of the large random matrix theory are recalled.

Theorem 1 (Theorem 1 in [13]). Let $\boldsymbol{D} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{S} \in \mathbb{C}^{M\times M}$ be Hermitian nonnegative definite and let $\boldsymbol{H} \in \mathbb{C}^{M\times K}$ be random with independent column vectors $\boldsymbol{h}_k \sim \mathcal{CN}(\boldsymbol{0}, \frac{1}{M}\boldsymbol{R}_k)$. Assume that $\boldsymbol{D}$ and the matrices $\boldsymbol{R}_k (k = 1, 2, \ldots, K)$, have uniformly bounded spectral norms (with respect to M). Then, for any $\rho > 0$,

$$\frac{1}{M}\mathrm{tr}\left(\boldsymbol{D}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{S} + \rho\boldsymbol{I}_M)^{-1}\right) - \frac{1}{M}\mathrm{tr}\left(\boldsymbol{D}\boldsymbol{T}(\rho)\right) \xrightarrow[M\to\infty]{\text{a.s.}} 0, \tag{29}$$

where $\boldsymbol{T}(\rho) \in \mathbb{C}^{M\times M}$ is defined as

$$\boldsymbol{T}(\rho) = \left(\frac{1}{M}\sum_{k=1}^{K}\frac{\boldsymbol{R}_k}{1+\delta_k(\rho)} + \boldsymbol{S} + \rho\boldsymbol{I}_M\right)^{-1}, \tag{30}$$

and the elements of $\boldsymbol{\delta}(\rho) \triangleq [\delta_1(\rho), \ldots, \delta_K(\rho)]^{\mathrm{T}}$ are defined as $\delta_k(\rho) = \lim_{t\to\infty}\delta_k^{(t)}(\rho)$, where for $t = 1, 2, \cdots$.

$$\delta_k^{(t)}(\rho) = \frac{1}{M}\mathrm{tr}\left(\boldsymbol{R}_k\left(\frac{1}{M}\sum_{j=1}^{K}\frac{\boldsymbol{R}_j}{1+\delta_j^{(t-1)}(\rho)} + \boldsymbol{S} + \rho\boldsymbol{I}_M\right)^{-1}\right), \tag{31}$$

with initial values $\delta_k^{(0)}(\rho) = 1/\rho$ for all k.

Theorem 2 ([13]). Let $\boldsymbol{\Theta} \in \mathbb{C}^{M\times M}$ be Hermitian nonnegative definite uniformly bounded spectral norm (with respect to M). Under the same conditions for $\boldsymbol{D}$ and $\boldsymbol{H}$ as described in Theorem 1,

$$\frac{1}{M}\mathrm{tr}\left(\boldsymbol{D}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{S} + \rho\boldsymbol{I}_M)^{-1}\boldsymbol{\Theta}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{S} + \rho\boldsymbol{I}_M)^{-1}\right) - \frac{1}{M}\mathrm{tr}\left(\boldsymbol{D}\boldsymbol{T}'(\rho)\right) \xrightarrow[M\to\infty]{\text{a.s.}} 0, \tag{32}$$

where $\boldsymbol{T}'(\rho) \in \mathbb{C}^{M\times M}$ is defined as

$$\boldsymbol{T}'(\rho) = \boldsymbol{T}(\rho)\boldsymbol{\Theta}\boldsymbol{T}(\rho) + \boldsymbol{T}(\rho)\frac{1}{M}\sum_{k=1}^{K}\frac{\boldsymbol{R}_k\delta_k'(\rho)}{(1+\delta_k(\rho))^2}\boldsymbol{T}(\rho). \tag{33}$$

$\boldsymbol{T}(\rho)$ and $\boldsymbol{\delta}(\rho)$ are given by Theorem 1, and $\boldsymbol{\delta}'(\rho) \triangleq [\delta_1'(\rho), \ldots, \delta_K'(\rho)]^{\mathrm{T}}$ is calculated as

$$\boldsymbol{\delta}'(\rho) = (\boldsymbol{I}_K - \boldsymbol{J}(\rho))^{-1}\boldsymbol{v}(\rho), \tag{34}$$

where $\boldsymbol{J}(\rho) \in \mathbb{C}^{K\times K}$ and $\boldsymbol{v} \in \mathbb{C}^{K\times 1}$ are defined as

$$[\boldsymbol{J}(\rho)]_{kl} = \frac{\frac{1}{M}\mathrm{tr}\left(\boldsymbol{R}_k\boldsymbol{T}(\rho)\boldsymbol{R}_l\boldsymbol{T}(\rho)\right)}{M(1+\delta_l(\rho))^2}, 1 \leqslant k, l \leqslant K, \tag{35}$$

$$[\boldsymbol{v}(\rho)]_k = \frac{1}{M}\mathrm{tr}\left(\boldsymbol{R}_k\boldsymbol{T}(\rho)\boldsymbol{\Theta}\boldsymbol{T}(\rho)\right), 1 \leqslant k \leqslant K. \tag{36}$$

Then we derive the deterministic approximation of $\gamma_{jqk}^{\mathrm{sr}}$, denoted by $\bar{\gamma}_{jqk}^{\mathrm{sr}}$, such that

$$\bar{\gamma}_{jqk}^{\mathrm{sr}} - \gamma_{jqk}^{\mathrm{sr}} \xrightarrow[M\to\infty]{\text{a.s.}} 0. \tag{37}$$

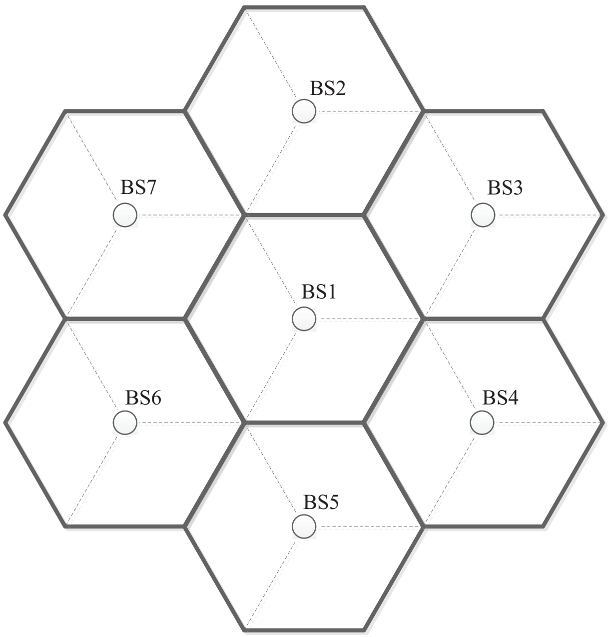

Figure 1 A 7-cell system, each cell with 3 sectors.

Theorem 3. For the uplink detector in (23), we have $\bar{\gamma}_{jqk}^{\mathrm{sr}} - \gamma_{jqk}^{\mathrm{sr}} \xrightarrow[M\to\infty]{\text{a.s.}} 0$, where $\bar{\gamma}_{jqk}^{\mathrm{sr}}$ is given by

$$\bar{\gamma}_{jqk}^{\mathrm{sr}} = \frac{\delta_{jqk}^2}{\frac{1}{Bp_{\mathrm{ul}}}\omega_{jqk} + \frac{1}{B}\sum\limits_{l\in\mathcal{C}_a, p\in\mathcal{S}_b, m\neq k}\mu_{jqlpmk} + \frac{1}{B}\sum\limits_{l\in\mathcal{C}\backslash\mathcal{C}_a \text{ or } p\in\mathcal{S}\backslash\mathcal{S}_b, m\neq k}\nu_{jqlpmk} + \sum\limits_{l\in\mathcal{C}_a, p\in\mathcal{S}_b, (l,p)\neq(j,q)}|\vartheta_{jqlpk}|^2}, \tag{38}$$

with

$$\mu_{jqlpmk} = \nu_{jqlpmk} - \frac{2\Re(\vartheta_{jqlpm}^* \vartheta_{jqlpmk}')}{1+\delta_{jqm}} + \frac{|\vartheta_{jqlpm}|^2 \delta_{jqmk}'}{(1+\delta_{jqm})^2},$$

$$\nu_{jqlpmk} = \frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Theta}_{jqlpm}\boldsymbol{T}_{jqk}'\right), \omega_{jqk} = \frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\bar{\boldsymbol{T}}_{jq}'\right),$$

$$\vartheta_{jqlpm} = \frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqlpm}\boldsymbol{T}_{jq}\right), \vartheta_{jqlpmk}' = \frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqlpm}\boldsymbol{T}_{jqk}'\right),$$

where

(1) $\boldsymbol{T}_{jq} = \boldsymbol{T}\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ and $\boldsymbol{\delta}_{jq} = [\delta_{jq1}, \ldots, \delta_{jqK'}]^{\mathrm{T}} = \boldsymbol{\delta}\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ are given by Theorem 1, for $\boldsymbol{S} = \boldsymbol{Z}_{jq}/B$, $\boldsymbol{D} = \boldsymbol{I}_B$ and $\boldsymbol{R}_m = \boldsymbol{\Xi}_{jqjqm}$;

(2) $\bar{\boldsymbol{T}}_{jq}' = \boldsymbol{T}'\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ is given by Theorem 2, for $\boldsymbol{S} = \boldsymbol{Z}_{jq}/B$, $\boldsymbol{\Theta} = \boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{W}_q$, $\boldsymbol{D} = \boldsymbol{I}_B$ and $\boldsymbol{R}_m = \boldsymbol{\Xi}_{jqjqm}$;

(3) $\boldsymbol{T}_{jqk}' = \boldsymbol{T}'\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ and $\boldsymbol{\delta}_{jqk}' = [\delta_{jq1k}', \ldots, \delta_{jqK'k}']^{\mathrm{T}} = \boldsymbol{\delta}'\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ are given by Theorem 2, for $\boldsymbol{S} = \boldsymbol{Z}_{jq}/B$, $\boldsymbol{\Theta} = \boldsymbol{\Xi}_{jqjqk}$, $\boldsymbol{D} = \boldsymbol{I}_B$ and $\boldsymbol{R}_m = \boldsymbol{\Xi}_{jqjqm}$.

Proof. The main idea is to obtain the deterministic approximation of each term in (27). Then the deterministic approximation of (27) is derived as (38), with a full proof shown in Appendix B.

5 Numerical results

In this section, we study the performance of the sectorized system using different sectorization weight matrices. Then the one with the best performance is utilized for further performance evaluation.

Consider a 7-cell massive MIMO system. Each cell has a BS equipped with the UCA located in the center and is equally partitioned into $Q = 3$ sectors (see Figure 1). Here we assume that all the cells and

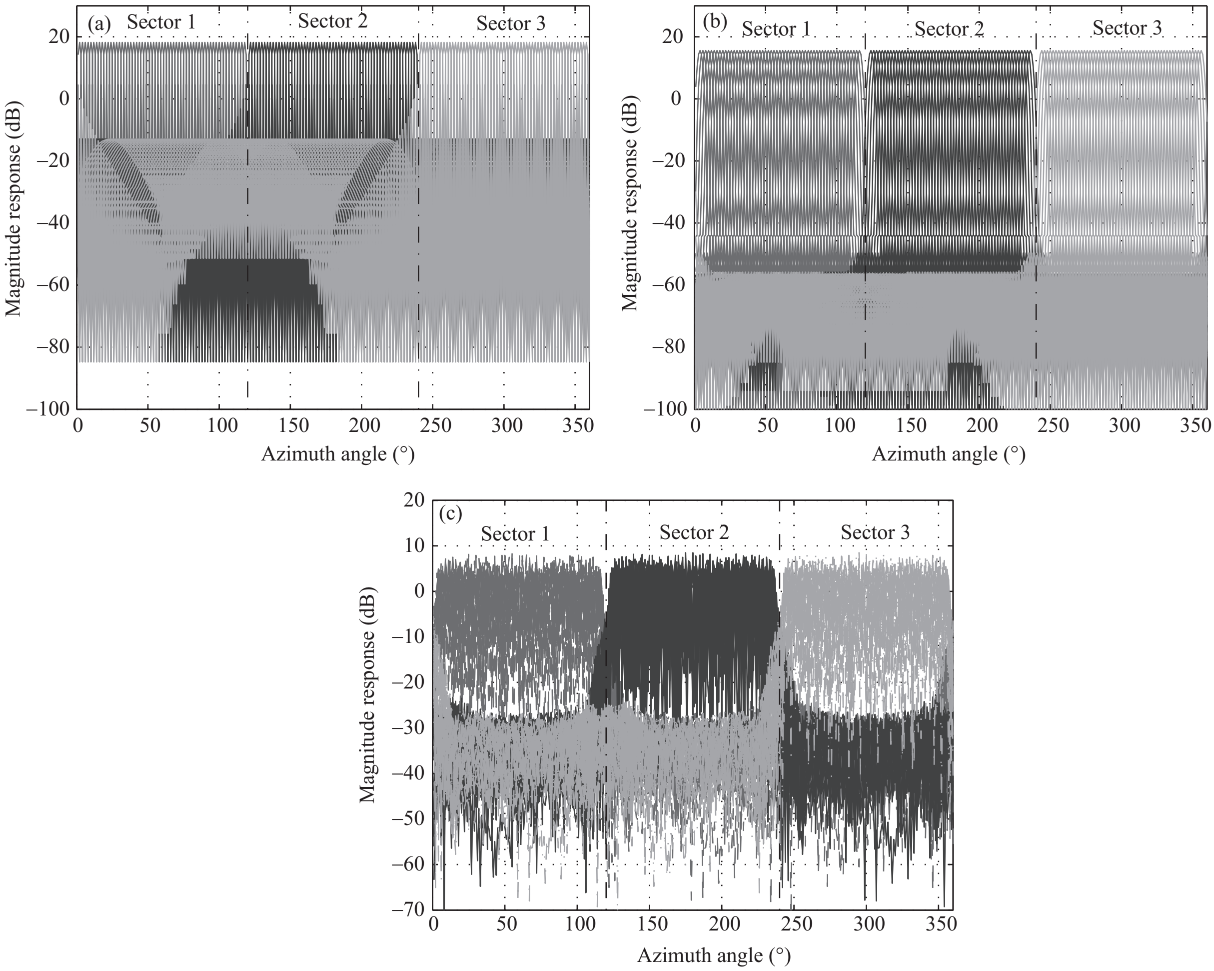

Figure 2 Magnitude responses of spatial filters for all the sectors. (a) Narrow beam; (b) medium beam; (c) wide beam.

sectors fully reuse the same set of pilots [2)]. Let $N = 180$, $r = N\lambda/4\pi$ (the distance between adjacent antennas is about half the wavelength), and $p_{\rm tr} = \tau^{\rm sr} p_{\rm ul}$. Here $\tau^{\rm sr}$ is K'. The antenna pattern used in this paper is the same as the element response described in [11], which is not omnidirectional and is beneficial to sectorization with the UCA.

5.1 Performance comparison using different weight matrices

Three different weight matrices are utilized in the sectorized system, which are corresponding to narrow beam, medium beam and wide beam, with magnitude responses shown in Figure 2(a)(b)(c), respectively. Here beams in different sectors are denoted by curves in different gray scales. The narrow beam means only adjacent beams in the same sector are overlapped with each other, the medium beam means several neighboring beams in the same sector are overlapped with each other, and the wide beam means all beams in the same sector are overlapped with each other. The basic weight vectors of the former two matrices are obtained using the method introduced in [11], and the last weight matrix is obtained using the main eigenvectors.

Considering the channel vector $\boldsymbol{h}_{jjqk}$, its corresponding AOA spread is denoted as $\Delta\theta_{jjqk}$, which is determined by one ring model [14], and the mean value of the AOAs (denoted by $\bar{\theta}_{jjqk}$) is assumed to fall into the q-th sector of the j-th cell. Assume that signals from all AOAs are independent of each other and uniformly distributed over $[\bar{\theta}_{jjqk} - \Delta\theta_{jjqk}/2, \bar{\theta}_{jjqk} + \Delta\theta_{jjqk}/2]$. The covariance matrix of $\boldsymbol{h}_{jjqk}$ can

2) Due to similar conclusions and limited space, other pilot reuse factors are not considered in this part.

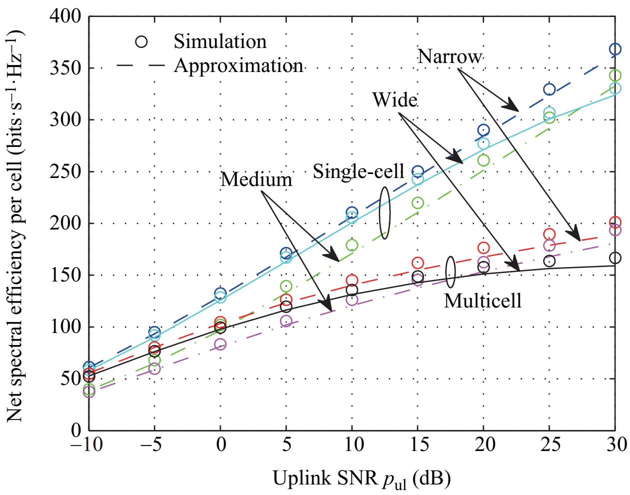

Figure 3 (Color online) Net SE comparison of different weight matrices under different uplink SNR.

be computed by

$$[\boldsymbol{R}_{jjqk}]_{mn} = \frac{zd_{jjqk}^{-\gamma}}{\Delta\theta_{jjqk}} \int_{\bar{\theta}_{jjqk}-\frac{\Delta\theta_{jjqk}}{2}}^{\bar{\theta}_{jjqk}+\frac{\Delta\theta_{jjqk}}{2}} a_1\left(\theta - \frac{2\pi}{M}(m-1)\right) a_1^*\left(\theta - \frac{2\pi}{M}(n-1)\right) \mathrm{e}^{-\mathrm{j}kr\left[\cos\left(\theta-\frac{2\pi}{M}(m-1)\right)-\cos\left(\theta-\frac{2\pi}{M}(n-1)\right)\right]}\mathrm{d}\theta. \tag{39}$$

Without loss of generality, we assume that users in each sector are uniformly distributed on a circle of 1/3 inter-site distance (set as 500 m) around the BS, so AOA spreads of users seen at their serving BSs are the same, denoted by $\Delta\theta$. Based on $\Delta\theta_{jjqk} = \Delta\theta(\forall j, q, k)$ and distance between the user and its serving BS, the radius of scattering ring can be gotten, then the AOA spreads of the users seen at their non-serving BSs can be obtained. With such AOA spreads, $\boldsymbol{R}_{jlpk}(\forall l \neq j, \forall p, k)$ can be computed similarly as (39).

Let $K' = 10$, $\gamma = 3.7$, $10\log_{10}(z) \sim \mathcal{N}(0, \delta_{sf}^2)$ with $\delta_{sf}^2 = 5$, $T_{\rm c} = 60$ and $\Delta\theta = 8°$. We compare the net SE and its deterministic approximation using such three weight matrices under different uplink SNR for both single-cell scenario and multicell scenario. As shown in Figure 3, it can be seen that the performance of narrow beam is better than the other two beams for both scenarios and all uplink SNR. The main reason for this is that the narrow beam can provide a better beamforming gain and have the same number of degrees of freedom for each sector as the other two beams. Thus we will utilize the narrow beam in the sectorized system for further analysis. What's more, it is verified that the results of deterministic approximation fit well with those of Monte-Carlo simulation. So only the deterministic approximation is shown in the following analysis.

5.2 Performance evaluation of the sectorization method

In order to evaluate the performance of the sectorized system, the benchmark system introduced in Section 2, the system using the pilot reuse method introduced in [5] and the system introduced in [7] are utilized as comparisons. To be fair, each system has the same user distribution and the systems that reuse pilots have the same sector numbers (or group numbers [5]) per cell. The effects of several different parameters are considered, including the number of users per cell, uplink SNR, coherence interval length and AOA spread. For all conditions below, we have $\gamma = 3.7$ and $10\log_{10}(z) \sim \mathcal{N}(0, \delta_{sf}^2)$ with $\delta_{sf}^2 = 5$.

As shown in Figure 4, we consider the net SE per cell versus uplink SNR with $T_{\rm c} = 60$, $K = 30$, $\Delta\theta = 8°$. The performance of the three systems with pilot reuse is always better than the benchmark system. As $p_{\rm ul}$ increases, the performance improvement of the sectorized system compared with the system in [7] becomes greater, which proves that the UCA can achieve a better performance than the ULA. On the other hand, compared with the system in [5], the sectorized system in this paper achieves

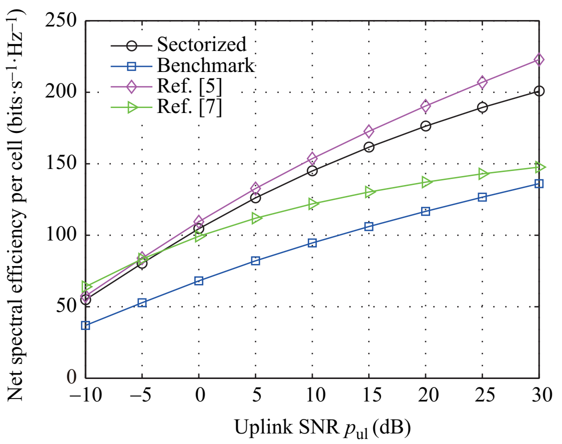

Figure 4 (Color online) Net SE per cell vs. p_{ul} with $T_{\mathrm{c}} = 60$, $K = 30$, $\Delta\theta = 8°$.

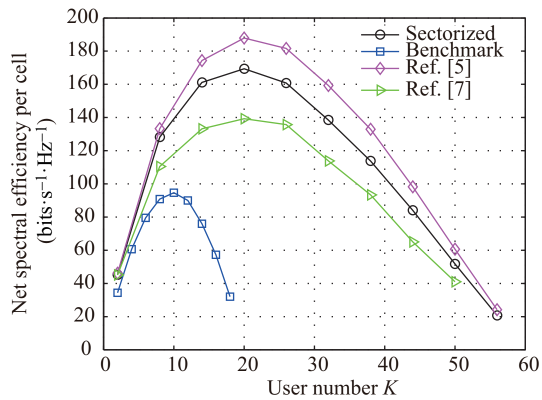

Figure 5 (Color online) Net SE per cell vs. K with $p_{\mathrm{ul}} = 10$ dB, $T_{\mathrm{c}} = 60$, $\Delta\theta = 8°$.

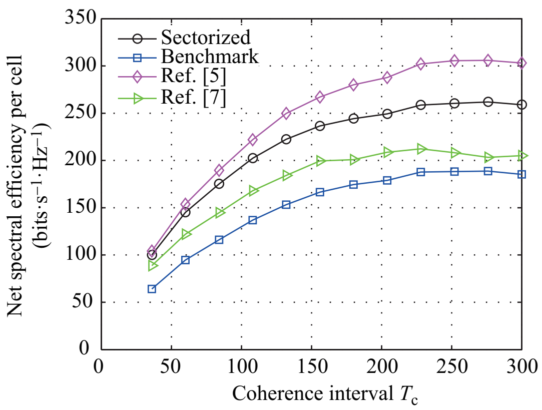

Figure 6 (Color online) Net SE per cell vs. T_{c} with $p_{\mathrm{ul}} = 10$ dB, $K = T_{\mathrm{c}}/2$, $\Delta\theta = 8°$.

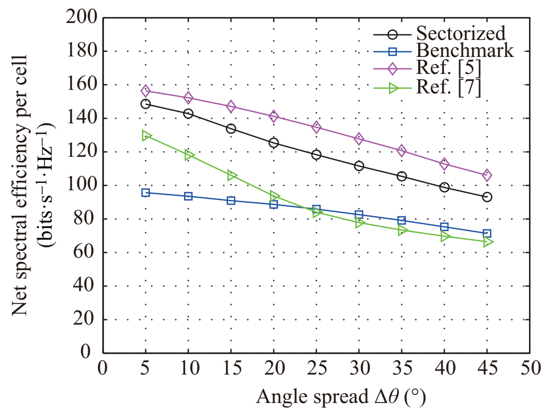

Figure 7 (Color online) Net SE per cell vs. $\Delta\theta$ with $p_{\mathrm{ul}} = 10$ dB, $T_{\mathrm{c}} = 60$, $K = 30$.

a similar performance when p_{ul} is low, and when p_{ul} increases, their performance gap becomes wider, which is caused by the inter-beam interference and is acceptable if considering the overhead of covariance matrix estimation and user grouping algorithm.

In Figure 5, the net SE per cell versus user number per cell is analyzed with $p_{\mathrm{ul}} = 10$ dB, $T_{\mathrm{c}} = 60$, $\Delta\theta = 8°$. It is obvious that the sectorized system outperforms the benchmark system and the system in [7]. Compared with the benchmark system, due to the reduction of pilot overhead, the sectorized system can support more users simultaneously and can greatly improve the maximum net SE. When the user number is small, the performance gap between the sectorized system and the the system in [5] is narrow. But as user number increases, there is a higher probability that different users in the same sector encounter high inter-beam interference, which restricts the performance of the sectorized system.

In Figure 6, we show the net SE per cell versus coherence interval length with $p_{\mathrm{ul}} = 10$ dB, $\Delta\theta = 8°$. Since the net SE of the benchmark system is maximized when it spends half of the coherence interval on pilot training [3], so here we set $K = T_{\mathrm{c}}/2$. The sectorized system has a better performance than the benchmark system and the system in [7] for all T_{c}. What's more, the sectorized system can achieve more than 85% of the performance of the system in [5].

At last, Figure 7 shows the net SE per cell versus angle spread with $p_{\mathrm{ul}} = 10$ dB, $T_{\mathrm{c}} = 60$, $K = 30$. Since a larger angle spread can cause more serious inter-user interference, the performance gets worse as the angle spread increases. The sectorized system outperforms the benchmark system and the system in [7] and its performance deterioration rate is much lower than the system in [7], whose main reason is that narrow beam in this paper can control the inter-user interference much better than the wide beam utilized in [7]. Moreover, as the angle spread increases, the performance gap between the sectorized system and the system in [5] remains almost constant.

6 Conclusion

Based on the sectorization method using the UCA, we have obtained a sectorized multicell massive MIMO system, whose ergodic achievable uplink net SE was derived with its deterministic approximation. The weight matrix of narrow beam for the sectorized system was proved to have the best performance, and it has been utilized in the sectorized system for further comparisons with other systems. Moreover, the consistency of the deterministic approximation with the Monte-Carlo simulation was confirmed at the same time. The deterministic approximation can be utilized for system design and optimization, which is left for future research. At last, the performance of the sectorized system was compared with those of the conventional system and two systems with pilot reuse. It outperforms the conventional system and the system in [7], and has a performance slightly lower than that of the system in [5], which validates the effectiveness of the sectorization method using the UCA.

Acknowledgements This work was supported by National Basic Research Program of China (Grant No. 2012CB316002), National Natural Science Foundation of China (Grant No. 61201192), National High Technology Research and Development Program of China (863 Program) (Grant No. 2014AA01A704), Science Fund for Creative Research Groups of NSFC (Grant No. 61321061), International Science and Technology Cooperation Program (Grant No. 2012DFG12010), National S&T Major Project (Grant No. 2014ZX03003003-002), Key Grant Project of Chinese Ministry of Education (Grant No. 313005), Tsinghua University Initiative Scientific Research (Grant No. 2015Z02-3), Open Research Fund of National Mobile Communications Research Laboratory, Southeast University (Grant No. 2012D02), and Tsinghua-Qualcomm Joint Research Program. We would like to thank the reviewers for their valuable comments.

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Rusek F, Persson D, Lau B K, et al. Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process Mag, 2013, 30: 40–60
2 Larsson E, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless systems. IEEE Commun Mag, 2014, 52: 186–195
3 Marzetta T L. Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans Wirel Commun, 2010, 9: 3590–3600
4 Marzetta T L. Massive MIMO: an introduction. Bell Labs Tech J, 2015, 20: 11–22
5 You L, Gao X Q, Xia X G, et al. Pilot reuse for massive MIMO transmission over spatially correlated rayleigh fading channels. IEEE Trans Wirel Commun, 2015, 14: 3352–3366
6 He Q, Xiao L M, Zhong X F, et al. Increasing the sum-throughput of cells with a sectorization method for massive MIMO. IEEE Commun Lett, 2014, 18: 1827–1830
7 Li J H, He Q, Xiao L M, et al. Uplink Sum-throughput evaluation of sectorized multi-cell massive MIMO system. In: Proceedings of IEEE International Conference on Communication Workshop (ICCW), London, 2015. 1143–1148
8 Hoydis J, Ten Brink S, Debbah M. Massive MIMO in the UL/DL of cellular networks: how many antennas do we need?. IEEE J Sel Areas Commun, 2013, 31: 160–171
9 Nam J, Adhikary A, Ahn J Y, et al. Joint spatial division and multiplexing: opportunistic beamforming, user grouping and simplified downlink scheduling. IEEE J Sel Topics Signal Process, 2014, 8: 876–890
10 Yin H F, Gesbert D, Filippou M, et al. A coordinated approach to channel estimation in large-scale multiple-antenna systems. IEEE J Sel Areas Commun, 2013, 31: 264–273
11 Alrabadi O N, Tsakalaki E, Huang H, et al. Beamforming via large and dense antenna arrays above a clutter. IEEE J Sel Areas Commun, 2013, 31: 314–325
12 Hassibi B, Hochwald B M. How much training is needed in multiple-antenna wireless links? IEEE Trans Inf Theory, 2003, 49: 951–963
13 Wagner S, Couillet R, Debbah M, et al. Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback. IEEE Trans Inf Theory, 2012, 58: 4509–4537

14 Xu Y, Yue G, Prasad N, et al. User grouping and scheduling for large scale MIMO systems with two-stage precoding. In: Proceedings of the IEEE International Conference on Communication (ICC), Sydney, 2014. 5197–5202

Appendix A Useful lemmas

Lemma A1 (Matrix inversion lemma(I), (2.2) in [A1]). Let $\boldsymbol{A} \in \mathbb{C}^{M\times M}$ be Hermitian invertible. Then, for any vector $\boldsymbol{x} \in \mathbb{C}^{M\times 1}$ and any scalar $\tau \in \mathbb{C}$ such that $\boldsymbol{A} + \tau\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}$ is invertible,

$$\boldsymbol{x}^{\mathrm{H}}(\boldsymbol{A} + \tau\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}})^{-1} = \frac{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{-1}}{1+\tau\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{-1}\boldsymbol{x}}.$$

Lemma A2 (Matrix inversion lemma(II), Lemma 2 in [8]). Let $\boldsymbol{A} \in \mathbb{C}^{M\times M}$ be Hermitian invertible. Then, for any vector $\boldsymbol{x} \in \mathbb{C}^{M\times 1}$ and any scalar $\tau \in \mathbb{C}$ such that $\boldsymbol{A} + \tau\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}$ is invertible,

$$(\boldsymbol{A} + \tau\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}})^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\tau\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{-1}}{1+\tau\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{-1}\boldsymbol{x}}.$$

Lemma A3 (Generalized rank-1 perturbation lemma, Lemma 14.3 in [A2]). Let $\boldsymbol{A} \in \mathbb{C}^{M\times M}$ be deterministic with uniformly bounded spectral norm (with respect to M) and $\boldsymbol{B} \in \mathbb{C}^{M\times M}$ be a random Hermitian matrix, with eigenvalues $\lambda_1^{\boldsymbol{B}} \leqslant \cdots \leqslant \lambda_M^{\boldsymbol{B}}$ such that, with probability on, there exist $\varepsilon > 0$ and M_0 such that $\lambda_1^{\boldsymbol{B}} > \varepsilon$ for all $M > M_0$. Then for any vector $\boldsymbol{x} \in \mathbb{C}^{M\times 1}$,

$$\frac{1}{M}\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{B}^{-1}\right) - \frac{1}{M}\mathrm{tr}\left(\boldsymbol{A}\left(\boldsymbol{B}+\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\right)^{-1}\right) \xrightarrow[M\to\infty]{\text{a.s.}} 0,$$

where $\boldsymbol{B}^{-1}$ and $\left(\boldsymbol{B}+\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\right)^{-1}$ exist with probability one.

Lemma A4 (Lemma B.26 in [A3], Theorem 3.7 in [A2], Lemma 12 in [A4]). Let $\boldsymbol{A} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{x}, \boldsymbol{y} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M)$. Assume that $\boldsymbol{A}$ has uniformly bounded spectral norm (with respect to M) and that $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{A}$ are mutually independent. Then, for all $p \geqslant 1$,

(1) $\mathrm{E}\left\{\left|\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x} - \frac{1}{M}\mathrm{tr}(\boldsymbol{A})\right|^p\right\} = \mathcal{O}\left(\frac{1}{M^{\frac{p}{2}}}\right)$,

(2) $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x} - \frac{1}{M}\mathrm{tr}(\boldsymbol{A}) \xrightarrow[M\to\infty]{\text{a.s.}} 0$,

(3) $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{y} \xrightarrow[M\to\infty]{\text{a.s.}} 0$,

(4) $\mathrm{E}\left\{\left|\left(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}\right)^2 - \left(\frac{1}{M}\mathrm{tr}(\boldsymbol{A})\right)^2\right|\right\} \xrightarrow[M\to\infty]{} 0$.

Appendix B Proof of Theorem 3

Define the following matrices for $j = 1, \ldots, L$, $q = 1, \ldots, Q$ and $k = 1, \ldots, K'$:

$$\begin{cases} \boldsymbol{\Sigma}_{jq} = \left(\hat{\boldsymbol{G}}_{jqjq}\hat{\boldsymbol{G}}_{jqjq}^{\mathrm{H}} + \boldsymbol{Z}_{jq} + \frac{c_q}{p_{\mathrm{ul}}}\boldsymbol{I}_B\right)^{-1}, \\ \boldsymbol{\Sigma}_{jqk} = \left(\hat{\boldsymbol{G}}_{jqjq}\hat{\boldsymbol{G}}_{jqjq}^{\mathrm{H}} - \hat{\boldsymbol{g}}_{jqjqk}\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}} + \boldsymbol{Z}_{jq} + \frac{c_q}{p_{\mathrm{ul}}}\boldsymbol{I}_B\right)^{-1}, \\ \boldsymbol{\Sigma}'_{jq} = B\boldsymbol{\Sigma}_{jq}, \boldsymbol{\Sigma}'_{jqk} = B\boldsymbol{\Sigma}_{jqk}. \end{cases} \tag{B1}$$

In the following proof, we use $a \asymp b$ to represent $a - b \xrightarrow[M\to\infty]{\text{a.s.}} 0$.

1) Signal power

$$\begin{aligned} \boldsymbol{r}_{jqk}^{\mathrm{H}}\hat{\boldsymbol{g}}_{jqjqk} &= \hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\hat{\boldsymbol{g}}_{jqjqk} \\ &= \hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\left(\boldsymbol{\Sigma}_{jqk}^{-1} + \hat{\boldsymbol{g}}_{jqjqk}\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\right)\hat{\boldsymbol{g}}_{jqjqk} \overset{\text{(a)}}{=} \frac{\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\hat{\boldsymbol{g}}_{jqjqk}}{1+\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\hat{\boldsymbol{g}}_{jqjqk}} \\ &\overset{\text{(b)}}{\asymp} \frac{\frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}'_{jqk}\right)}{1+\frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}'_{jqk}\right)} \overset{\text{(c)}}{\asymp} \frac{\frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}'_{jq}\right)}{1+\frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}'_{jq}\right)} \\ &\overset{\text{(d)}}{\asymp} \frac{\frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\boldsymbol{T}_{jq}\right)}{1+\frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\boldsymbol{T}_{jq}\right)} \overset{\text{(e)}}{=} \frac{\delta_{jqk}}{1+\delta_{jqk}}, \end{aligned} \tag{B2}$$

where (a) follows from Lemma A1, (b) follows from Lemma A4 (2), (c) follows from Lemma A3, (d) results from Theorem 1 for $\boldsymbol{D} = \boldsymbol{\Xi}_{jqjqk}$, $\boldsymbol{S} = \boldsymbol{Z}_{jq}/B$, $\boldsymbol{T}_{jq} = \boldsymbol{T}_{jq}\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ with $\boldsymbol{R}_m = \boldsymbol{\Xi}_{jqjqm}$ and (e) uses the definition $\delta_{jqk} = \frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\boldsymbol{T}_{jq}\right)$. By using the continuous mapping theorem [B1], we can obtain

$$\left|\boldsymbol{r}_{jqk}^{\mathrm{H}}\hat{\boldsymbol{g}}_{jqjqk}\right|^2 - \left(\frac{\delta_{jqk}}{1+\delta_{jqk}}\right)^2 \xrightarrow[M\to\infty]{\text{a.s.}} 0. \tag{B3}$$

2) Channel uncertainty

$$\begin{aligned}\boldsymbol{r}_{jqk}^{\mathrm{H}}\tilde{\boldsymbol{g}}_{jqjqk} &= \hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\tilde{\boldsymbol{g}}_{jqjqk} \\ &\overset{(a)}{=} \frac{\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\tilde{\boldsymbol{g}}_{jqjqk}}{1+\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\hat{\boldsymbol{g}}_{jqjqk}} = \frac{\frac{1}{B}\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}'_{jqk}\tilde{\boldsymbol{g}}_{jqjqk}}{1+\frac{1}{B}\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}'_{jqk}\hat{\boldsymbol{g}}_{jqjqk}} \overset{(b)}{\asymp} 0,\end{aligned} \tag{B4}$$

where (a) follows from Lemma A1 and (b) follows from Lemma A4 (3). Thus by using the dominated convergence theorem [B2] and the continuous mapping theorem [B1], we have

$$\mathrm{E}\left\{\left|\boldsymbol{r}_{jqk}^{\mathrm{H}}\tilde{\boldsymbol{g}}_{jqjqk}\right|^2 \middle| \hat{\boldsymbol{G}}_{jqjq}\right\} \xrightarrow[M\to\infty]{\text{a.s.}} 0. \tag{B5}$$

3) Noise power

$$\begin{aligned}\left\|\boldsymbol{W}_q\boldsymbol{r}_{jqk}^{\mathrm{H}}\right\|^2 &= \hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{W}_q\boldsymbol{\Sigma}_{jq}\hat{\boldsymbol{g}}_{jqjqk} \\ &= \frac{\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{W}_q\boldsymbol{\Sigma}_{jqk}\hat{\boldsymbol{g}}_{jqjqk}}{\left(1+\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\hat{\boldsymbol{g}}_{jqjqk}\right)^2} \overset{(a)}{\asymp} \frac{1}{\left(1+\delta_{jqk}\right)^2}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jq}\boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{W}_q\boldsymbol{\Sigma}_{jq}\right) \\ &\overset{(b)}{\asymp} \frac{1}{\left(1+\delta_{jqk}\right)^2}\frac{1}{B^2}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\bar{\boldsymbol{T}}'_{jq}\right) \overset{(c)}{=} \frac{1}{\left(1+\delta_{jqk}\right)^2}\frac{1}{B}\omega_{jqk},\end{aligned} \tag{B6}$$

where (a) results from (B2), Lemma A3 and A4 (2), (b) results from Theorem 2 for $\bar{\boldsymbol{T}}'_{jq} = \boldsymbol{T}'\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$, $\boldsymbol{S} = \boldsymbol{Z}_{jq}/B$, $\boldsymbol{\Theta} = \boldsymbol{W}_q^{\mathrm{H}}\boldsymbol{W}_q$, $\boldsymbol{D} = \boldsymbol{\Xi}_{jqjqk}$ and $\boldsymbol{R}_m = \boldsymbol{\Xi}_{jqjqm}$, and (c) uses the definition $\omega_{jqk} = \frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqk}\bar{\boldsymbol{T}}'_{jq}\right)$.

By using the dominated convergence theorem [B2], we can get

$$\mathrm{E}\left\{\left\|\boldsymbol{W}_q\boldsymbol{r}_{jqk}^{\mathrm{H}}\right\|^2 \middle| \hat{\boldsymbol{G}}_{jqjq}\right\} \xrightarrow[M\to\infty]{\text{a.s.}} \frac{1}{\left(1+\delta_{jqk}\right)^2}\frac{1}{B}\omega_{jqk}. \tag{B7}$$

4) Interference power

The interference power from the m-th user in the p-th sector of the l-th cell is $\mathrm{E}\{|\boldsymbol{r}_{jqk}^{\mathrm{H}}\tilde{\boldsymbol{g}}_{jqlpm}|^2|\hat{\boldsymbol{G}}_{jqjq}\}$, which depends on the pilot reuse pattern.

a) $m = k, j, l \in \mathcal{C}_a, q, p \in \mathcal{S}_b$. In this case, we have

$$\begin{aligned}\left|\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\boldsymbol{g}_{jqlpm}\right|^2 &= \boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\hat{\boldsymbol{g}}_{jqjqk}\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\boldsymbol{g}_{jqlpk} \\ &\overset{(a)}{=} \frac{\boldsymbol{g}_{jqlpk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\hat{\boldsymbol{g}}_{jqjqk}\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\boldsymbol{g}_{jqlpk}}{\left(1+\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\tilde{\boldsymbol{g}}_{jqjqk}\right)^2} \overset{(b)}{\asymp} \frac{1}{\left(1+\delta_{jqk}\right)^2}\left|\boldsymbol{g}_{jqlpk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\hat{\boldsymbol{g}}_{jqjqk}\right|^2 \\ &\overset{(c)}{\asymp} \frac{1}{\left(1+\delta_{jqk}\right)^2}\left|\frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqlpk}\boldsymbol{T}_{jq}\right)\right|^2 \overset{(d)}{=} \frac{1}{\left(1+\delta_{jqk}\right)^2}\left|\vartheta_{jqlpk}\right|^2,\end{aligned} \tag{B8}$$

where (a) follows from Lemma A1, (b) uses the results in (B2), (c) results from Lemma A4 (2), Lemma A3, Theorem 1 for $\boldsymbol{D} = \boldsymbol{\Xi}_{jqlpk}$, $\boldsymbol{S} = \boldsymbol{Z}_{jq}/B$, $\boldsymbol{T}_{jq} = \boldsymbol{T}\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ with $\boldsymbol{R}_m = \boldsymbol{\Xi}_{jqjqm}$ and (d) uses the definition $\vartheta_{jqlpk} = \frac{1}{B}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqlpk}\boldsymbol{T}_{jq}\right)$.

b) $m \neq k, j, l \in \mathcal{C}_a, q, p \in \mathcal{S}_b$. In this case, we have

$$\begin{aligned}\left|\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\boldsymbol{g}_{jqlpm}\right|^2 &= \boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\hat{\boldsymbol{g}}_{jqjqk}\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\boldsymbol{g}_{jqlpm} \\ &\overset{(a)}{=} \frac{\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\boldsymbol{g}_{jqlpm}\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\hat{\boldsymbol{g}}_{jqjqk}}{\left(1+\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\tilde{\boldsymbol{g}}_{jqjqk}\right)^2} \overset{(b)}{\asymp} \frac{\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqk}\boldsymbol{g}_{jqlpm}}{\left(1+\delta_{jqk}\right)^2},\end{aligned} \tag{B9}$$

where (a) follows from Lemma A1, (b) follows from Lemma A4 (2) and uses the results in (B2). Then Lemma A2 is used:

$$\boldsymbol{\Sigma}_{jqk} = \boldsymbol{\Sigma}_{jqkm} - \frac{\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm}\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}}{1+\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm}}, \tag{B10}$$

where

$$\boldsymbol{\Sigma}_{jqkm} = \left(\hat{\boldsymbol{G}}_{jqjq}\hat{\boldsymbol{G}}_{jqjq}^{\mathrm{H}} - \hat{\boldsymbol{g}}_{jqjqk}\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}} - \hat{\boldsymbol{g}}_{jqjqm}\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}} + \boldsymbol{Z}_{jq} + \frac{c_q}{p_{\mathrm{ul}}}\boldsymbol{I}_B\right)^{-1}. \tag{B11}$$

Here $\boldsymbol{\Sigma}_{jqkm}$ is independent of $\boldsymbol{g}_{jqlpm}$, which means Lemma A4 can be used. By plugging (B10) into the numerator of (B9), we can get

$$\begin{aligned}\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqk}\boldsymbol{g}_{jqlpm} = \; &\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{g}_{jqlpm} \\ &- 2\Re\left\{\frac{\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{g}_{jqlpm}\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm}}{1+\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm}}\right\} \\ &+ \frac{\left|\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{g}_{jqlpm}\right|^2\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm}}{\left(1+\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm}\right)^2}.\end{aligned} \tag{B12}$$

As proved in (B2)(B8), we have $\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm} \asymp \delta_{jqm}$ and $\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{g}_{jqlpm} \asymp \vartheta_{jqlpm}^{*}$. Similarly, we can use Lemma A3, A4 and Theorem 2 to obtain the following results:

$$\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{g}_{jqlpm} \asymp \frac{1}{B^2}\mathrm{tr}\left(\boldsymbol{\Theta}_{jqlpm}\boldsymbol{T}_{jqk}'\right) = \frac{\nu_{jqlpmk}}{B}, \tag{B13}$$

$$\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm} \asymp \frac{1}{B^2}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqlpm}\boldsymbol{T}_{jqk}'\right) = \frac{\vartheta_{jqlpmk}'}{B}, \tag{B14}$$

$$\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm} \asymp \frac{1}{B^2}\mathrm{tr}\left(\boldsymbol{\Xi}_{jqjqm}\boldsymbol{T}_{jqk}'\right) = \frac{\delta_{jqmk}'}{B}, \tag{B15}$$

where $\boldsymbol{T}_{jqk}' = \boldsymbol{T}'\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ and $\boldsymbol{\delta}_{jqk}' = [\delta_{jq1k}', \ldots, \delta_{jqK'k}']^{\mathrm{T}} = \boldsymbol{\delta}'\left(\frac{c_q}{Bp_{\mathrm{ul}}}\right)$ are given by Theorem 2 for $\boldsymbol{S} = \boldsymbol{Z}_{jq}/B$, $\boldsymbol{\Theta} = \boldsymbol{\Xi}_{jqjqk}$, and $\boldsymbol{R}_m = \boldsymbol{\Xi}_{jqjqm}$. Then we can get

$$\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqk}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqk}\boldsymbol{g}_{jqlpm} \asymp \frac{\nu_{jqlpmk}}{B} - \frac{2\Re\{\vartheta_{jqlpm}^{*}\vartheta_{jqlpmk}'\}}{B(1+\delta_{jqm})} + \frac{\left|\vartheta_{jqlpm}\right|^2\delta_{jqmk}'}{B\left(1+\delta_{jqm}\right)^2} = \frac{\mu_{jqlpmk}}{B}. \tag{B16}$$

c) $m \neq k, l \in \mathcal{C}\backslash\mathcal{C}_a$ or $p \in \mathcal{S}\backslash\mathcal{S}_b$. Similar to b), we have

$$\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{g}_{jqlpm} \asymp \frac{1}{B^2}\mathrm{tr}\left(\boldsymbol{\Theta}_{jqlpm}\boldsymbol{T}_{jqk}'\right) = \frac{\nu_{jqlpmk}}{B}, \tag{B17}$$

$$\boldsymbol{g}_{jqlpm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm} \asymp 0, \tag{B18}$$

$$\hat{\boldsymbol{g}}_{jqjqm}^{\mathrm{H}}\boldsymbol{\Sigma}_{jqkm}\boldsymbol{\Xi}_{jqjqk}\boldsymbol{\Sigma}_{jqkm}\hat{\boldsymbol{g}}_{jqjqm} \asymp 0. \tag{B19}$$

Therefore, by using the dominated convergence theorem [B2], we have

$$\mathrm{E}\left\{\left|\hat{\boldsymbol{g}}_{jqjqk}^{\mathrm{H}}\boldsymbol{\Sigma}_{jq}\boldsymbol{g}_{jqlpm}\right|^2\middle|\hat{\boldsymbol{G}}_{jqjq}\right\} \xrightarrow[M\to\infty]{\text{a.s.}} \begin{cases} \dfrac{1}{\left(1+\delta_{jqk}\right)^2}\left|\vartheta_{jqlpk}\right|^2, & m=k, j,l \in \mathcal{C}_a, q,p \in \mathcal{S}_b, \\ \dfrac{1}{\left(1+\delta_{jqk}\right)^2}\dfrac{\mu_{jqlpmk}}{B}, & m \neq k, j,l \in \mathcal{C}_a, q,p \in \mathcal{S}_b, \\ \dfrac{1}{\left(1+\delta_{jqk}\right)^2}\dfrac{\nu_{jqlpmk}}{B}, & m \neq k, l \in \mathcal{C}\backslash\mathcal{C}_a \text{ or } p \in \mathcal{S}\backslash\mathcal{S}_b, \\ 0, & \text{others.} \end{cases} \tag{B20}$$

Finally, by using the continuous mapping theorem [B1] and combining (B3)(B5)(B7)(B20), we complete the proof with

$$\bar{\gamma}_{jqk}^{\mathrm{sr}} - \frac{\delta_{jqk}^2}{\frac{1}{Bp_{\mathrm{ul}}}\omega_{jqk} + \frac{1}{B}\sum\limits_{l\in\mathcal{C}_a, p\in\mathcal{S}_b, m\neq k}\mu_{jqlpmk} + \frac{1}{B}\sum\limits_{l\in\mathcal{C}\backslash\mathcal{C}_a \text{ or } p\in\mathcal{S}\backslash\mathcal{S}_b, m\neq k}\nu_{jqlpmk} + \sum\limits_{l\in\mathcal{C}_a, p\in\mathcal{S}_b, (l,p)\neq(j,q)}|\vartheta_{jqlpk}|^2} \xrightarrow[M\to\infty]{\text{a.s.}} 0. \tag{B21}$$

Appendix References

A1 Silverstein J W, Bai Z D. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. J Multivariate Anal, 1995, 54: 175–192

A2 Couillet R, Debbah M. Random Matrix Methods for Wireless Communications. Cambridge: Cambridge University Press, 2011. 48–49, 352–353

A3 Bai Z, Silverstein J W. Spectral Analysis of Large Dimensional Random Matrices. New York: Springer, 2010. 530

A4 Hoydis J. Random matrix methods for advanced communication systems. Dissertation for Ph.D. degree. France: Supélec, 2012. 18

B1 van der Vaart A W. Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge: Cambridge University Press, 2000. 7–8

B2 Billingsley P. Probability and Measure. Hoboken: John Wiley & Sons. 1995. 209–210