IET Communications

Research Article

IET Journals
The Institution of Engineering and Technology

Single branch search based pilot allocation for multi-cell massive multiple-input multiple-output systems

ISSN 1751-8628
Received on 24th September 2016
Revised 7th November 2016
Accepted on 11th November 2016
E-First on 2nd March 2017
doi: 10.1049/iet-com.2016.1148
www.ietdl.org

Anzhong Hu[1] ✉, *Haiquan Wang*[1]

[1]*School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, People's Republic of China*

✉ *E-mail: huaz@hdu.edu.cn*

Abstract: The performance of multi-cell massive multiple-input multiple-output systems is constrained by inter-cell interference. In this study, a pilot allocation approach is proposed to mitigate interference in such systems. The asymptotic spectral efficiency in the limit of the number of antennas is derived first. By analysing the derivatives of the spectral efficiency with the pilot allocation, a single branch search algorithm for pilot allocation is formed. In contrast, the existing approaches allocate pilots for each cell separately or with brute-force search. The analysis and numerical results show that the proposed approach is of better compromise in mitigating the interference and reducing the computational complexity.

1 Introduction

Massive multiple-input multiple-output (MIMO) can achieve extremely high spectral efficiency and greatly reduce the transmission power [1, 2]. Currently, massive MIMO is extensively investigated and is a candidate technology for future wireless communication systems [3–6]. In time-division duplexing (TDD) multi-cell massive MIMO systems, the length of the pilots is constrained by the channel coherence interval. Thus, the pilots employed for uplink channel estimation are non-orthogonal rather than orthogonal and cause interference in the channel estimates. The interference in the channel estimates continues to degrade the system performance in the data transmission phases. This degradation of the system performance is also known as pilot contamination.

For the purpose of mitigating the interference in massive MIMO systems, various aspects of system design have been investigated in [7–14]. Despite these aspects of system design, pilot allocation [15–19] is also important in enhancing the signal and mitigating the interference. Since the pilot contamination is caused by reused pilots, it is obvious that a proper allocation of the pilots is also beneficial for massive MIMO systems. In [15], the optimal pilot allocation is found with a brute-force search algorithm. Due to the high computational complexity of this greedy algorithm, suboptimal approaches have been proposed. For example, the graph colouring and the game theory have been employed in [16–18]. In [19], a propagation loss dependent allocation approach has been proposed. However, these approaches either improve the performance of each cell separately [17–19] or still necessitates brute-force search [16].

To mitigate the interference in massive MIMO systems, the pilot allocation is investigated in this paper. The motivation is that the pilot allocation affects the signal power or the interference power but the existing optimisations of pilots are not good enough for massive MIMO systems. Moreover, the existing approaches for mitigating the pilot allocation in massive MIMO systems rely on other conditions, such as channel covariances, shifted frames and so on [20], while the pilot allocation can be done without these conditions. In this paper, the base station (BS) uses the minimum mean square error (MMSE) principle to estimate the channels, detect the signals, and precode the signals. Utilising the asymptotic orthogonality of the channel vectors in the limit of the BS antennas, the asymptotic spectral efficiency of the considered massive MIMO system is derived. Then, the derivative of the spectral efficiency with respect to the pilot correlation value is used to derive a metric for adjusting the pilot allocation. Based on this metric, a single branch search algorithm is proposed to search for the optimal pilot allocation. For the purpose of clarity, the main contributions of this paper are summarised as follows:

i. The asymptotic spectral efficiency of a massive MIMO system with MMSE processing is derived, which reveals how the pilot allocation affects the spectral efficiency. As can be seen, this result provides the guideline for allocating pilots in such systems.
ii. A metric for deciding whether to change the pilot correlation is derived, with which a single branch search algorithm is proposed to choose the best branch in each stage of the decision process. Since the proposed pilot allocation is designed for the whole system and the search process is simplified, it is expected to achieve a better compromise between the performance and the computational complexity than the approaches in [16, 19].

This paper is organised as follows. In Section 2, the system model is described. The pilot allocation is detailed in Section 3. Simulation results are provided in Section 4. Finally, conclusions are drawn in Section 5.

Notations: Lower-case (upper-case) boldface symbols denote vectors (matrices); $\boldsymbol{I}_K$ is the dimensional-K identity matrix, and $\mathbf{0}_{M\times K}$ represents an $M \times K$ zero matrix; $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$, and $(\cdot)^{\mathrm{H}}$ denote the conjugate, the transpose, and the conjugate transpose, respectively; $[\cdot]_{i,:}$, $[\cdot]_{:,j}$, and $[\cdot]_{i,j}$ are the ith row, the jth column, and the (i, j)th entry of a matrix, respectively; $\|\cdot\|$ is the Euclidean norm of a vector; $\mathcal{CN}(\mu, \sigma^2)$ is a circularly-symmetric complex Gaussian random variable with mean μ and variance σ^2; $a \propto b$ means that a is proportional to b.

2 System model

As shown in Fig. 1, the considered massive MIMO system is composed of L cells. In each cell, there is one BS and K randomly distributed mobile stations (MSs). All the MSs in the system occupy the same time and frequency resources for the uplink or the downlink transmission. The BS is equipped with an array of M antennas, each MS has only one antenna. This system is a typical massive MIMO system shown in [1, 2] and is expected to have severe interference because of the correlated pilots employed.

In one coherence interval, the channels are taken as invariant. The MSs transmit pilots to the BSs simultaneously, the BSs then

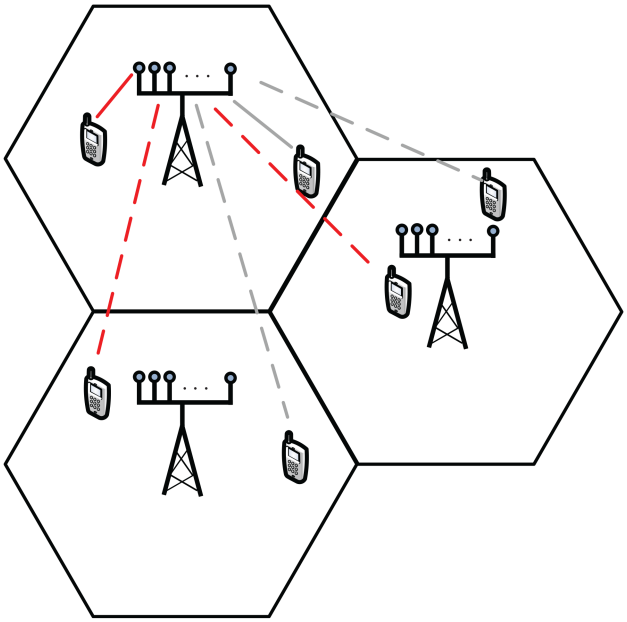

Fig. 1 *Multi-cell massive MIMO system. MSs that use the same pilot sequence interfere with each other*

estimate the channels for two-way transmissions. The received pilot matrix at the BS in the jth cell can be denoted as

$$\boldsymbol{Y}_j = \sqrt{\rho_{\mathrm{ul}}}\sum_{l=1}^{L}\boldsymbol{H}_{jl}\tilde{\boldsymbol{D}}_{jl}\boldsymbol{\Phi}_l^{\mathrm{T}} + \boldsymbol{N}_j \in \mathbb{C}^{M\times N_{\mathrm{P}}},$$

where ρ_{ul} is the transmission power of the pilots from the MSs; $\boldsymbol{H}_{jl}\in\mathbb{C}^{M\times K}$ is the channel matrix that is composed of the small-scale fading coefficients from the MSs in the lth cell to the BS in the jth cell, which are independent and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$ variables; $\tilde{\boldsymbol{D}}_{jl}\in\mathbb{R}^{K\times K}$ is composed of the large-scale fading coefficients and the power control coefficients and is a diagonal matrix, i.e. $\tilde{\boldsymbol{D}}_{jl}=\boldsymbol{D}_{jl}\boldsymbol{P}_l^{\mathrm{up}}$, where $\boldsymbol{D}_{jl}\in\mathbb{R}^{K\times K}$ is composed of the large-scale fading coefficients from the MSs in the lth cell to the BS in the jth cell, and $\boldsymbol{P}_l^{\mathrm{up}}\in\mathbb{R}^{K\times K}$ is composed of the power control coefficients of the MSs in the lth cell. $\boldsymbol{\Phi}_l\in\mathbb{C}^{N_{\mathrm{P}}\times K}$ is the matrix of pilots from the MSs in the lth cell, where N_{p} is the length of the pilot sequence of each MS; $\boldsymbol{N}_j\in\mathbb{C}^{M\times N_{\mathrm{P}}}$ is the noise matrix and is composed of i.i.d. $\mathcal{CN}(0,1)$ variables.

Similar to the idea of zero-forcing, the received pilot matrix can be firstly processed into

$$\begin{aligned}\tilde{\boldsymbol{Y}}_j &= \frac{1}{\sqrt{\rho_{\mathrm{ul}}}}\boldsymbol{Y}_j\boldsymbol{\Phi}_j^{*}(\boldsymbol{\Phi}_j^{\mathrm{T}}\boldsymbol{\Phi}_j^{*})^{-1}\tilde{\boldsymbol{D}}_{jj}^{-1}\\ &= \boldsymbol{H}_{jj} + \sum_{\substack{l=1\\ l\neq j}}^{L}\boldsymbol{H}_{jl}\tilde{\boldsymbol{D}}_{jl}\boldsymbol{\Psi}_{jl}^{\mathrm{H}}\boldsymbol{\Psi}_{jj}^{-1}\tilde{\boldsymbol{D}}_{jj}^{-1} + \frac{1}{\sqrt{\rho_{\mathrm{ul}}}}\boldsymbol{N}_j\boldsymbol{\Phi}_j^{*}\boldsymbol{\Psi}_{jj}^{-1}\tilde{\boldsymbol{D}}_{jj}^{-1}\in\mathbb{C}^{M\times K},\end{aligned} \tag{1}$$

where $\boldsymbol{\Psi}_{jl}=\boldsymbol{\Phi}_j^{\mathrm{T}}\boldsymbol{\Phi}_l^{*}\in\mathbb{R}^{K\times K}$ is the correlation matrix of the pilots. In addition, the following assumptions are made on the correlation of the pilots:

i. For the case of $j=l$, the auto-correlation matrix of the pilots is $\boldsymbol{\Psi}_{ll}=\boldsymbol{I}_K, \forall l=1,2,\ldots,L$.
ii. For the case of $j\neq l$, the cross-correlation of the pilots, $\boldsymbol{\Psi}_{jl}$, is a permutation matrix, i.e. there is only one entry of 1 in each row and each column and the other entries are 0 s.
iii. The pilots are re-assigned by the BSs in each channel coherence interval.

Note that the above assumption is similar to that in [2]. The difference is that [2] assumes fixed allocation of orthogonal pilots, which makes $\boldsymbol{\Psi}_{jl}$ equal $\boldsymbol{I}_K$. However, the orthogonal pilots are not fixed here, which makes $\boldsymbol{\Psi}_{jl}$ a permutation matrix.

With the above assumptions of the pilots, (1) can be simplified into

$$\tilde{\boldsymbol{Y}}_j = \boldsymbol{H}_{jj} + \sum_{\substack{l=1\\ l\neq j}}^{L}\boldsymbol{H}_{jl}\tilde{\boldsymbol{D}}_{jl}\boldsymbol{\Psi}_{jl}^{\mathrm{H}}\tilde{\boldsymbol{D}}_{jj}^{-1} + \frac{1}{\sqrt{\rho_{\mathrm{ul}}}}\boldsymbol{N}_j\boldsymbol{\Phi}_j^{*}\tilde{\boldsymbol{D}}_{jj}^{-1}. \tag{2}$$

For the kth MS in the jth cell, the corresponding processed pilot vector is

$$\begin{aligned}\tilde{\boldsymbol{y}}_{jk} &\triangleq [\tilde{\boldsymbol{Y}}_j]_{:,k}\\ &= \boldsymbol{h}_{jjk} + \tilde{d}_{jjk}^{-1}\sum_{\substack{l=1\\ l\neq j}}^{L}\boldsymbol{H}_{jl}\tilde{\boldsymbol{D}}_{jl}[\boldsymbol{\Psi}_{jl}^{\mathrm{H}}]_{:,k} + \tilde{d}_{jjk}^{-1}\rho_{\mathrm{ul}}^{-1/2}\boldsymbol{N}_j[\boldsymbol{\Phi}_j^{*}]_{:,k}\in\mathbb{C}^{M\times 1},\end{aligned} \tag{3}$$

where $\boldsymbol{h}_{jjk}\triangleq[\boldsymbol{H}_{jj}]_{:,k}\in\mathbb{C}^{M\times 1}$ is the channel vector from the kth MS in the jth cell to the BS in the jth cell, and is composed of small-scale fading coefficients; $\tilde{d}_{jlk}\triangleq[\tilde{\boldsymbol{D}}_{jl}]_{k,k}$ is the multiplication of the large-scale fading coefficient and the power control coefficient. Since these coefficients vary slowly in comparison to the small-scale fading coefficients, it is assumed that they are known a priori; similar assumptions can be found in [13, 14].

By employing the MMSE principle, the estimate of $\boldsymbol{H}_{jj}$ can be expressed as [9, 21]

$$\hat{\boldsymbol{H}}_{jj} = \tilde{\boldsymbol{Y}}_j\boldsymbol{D}_j^{a}\in\mathbb{C}^{M\times K}, \tag{4}$$

where $\boldsymbol{D}_j^{a}\in\mathbb{R}^{K\times K}$ and

$$[\boldsymbol{D}_j^{a}]_{k,k} = 1/\left(1+\tilde{d}_{jjk}^{-2}\sum_{\substack{l=1\\ l\neq j}}^{L}\|\,\tilde{\boldsymbol{D}}_{jl}[\boldsymbol{\Psi}_{jl}^{\mathrm{H}}]_{:,k}\,\|^2 + \tilde{d}_{jjk}^{-2}\rho_{\mathrm{ul}}^{-1}\right).$$

With the estimated channels, the uplink detection and downlink precoding can be conducted at the BSs.

2.1 Uplink detection

For the uplink transmission, the received symbol vector at the jth BS is written as

$$\boldsymbol{y}_j^{\mathrm{up}} = \sqrt{\rho_{\mathrm{ul}}}\sum_{l=1}^{L}\boldsymbol{H}_{jl}\tilde{\boldsymbol{D}}_{jl}\boldsymbol{s}_l + \boldsymbol{n}_j\in\mathbb{C}^{M\times 1}, \tag{5}$$

where $\boldsymbol{s}_l\in\mathbb{C}^{K\times 1}$ is vector composed of the transmitted symbols from the MSs in the lth cell. The elements of $\boldsymbol{s}_l$ are i.i.d. random variables of mean zero and variance one; $\boldsymbol{n}_j\in\mathbb{C}^{M\times 1}$ is the received noise vector with elements being i.i.d. $\mathcal{CN}(0,1)$ variables. With the knowledge of the channel estimates $\hat{\boldsymbol{H}}_{jj}$ and the MMSE detection principle, the received symbol for the kth MS in the jth cell is $\tilde{y}_{jk}^{\mathrm{up}}=\boldsymbol{f}_{jk}(\sqrt{\rho_{\mathrm{ul}}}\sum_{l=1}^{L}\boldsymbol{H}_{jl}\tilde{\boldsymbol{D}}_{jl}\boldsymbol{s}_l+\boldsymbol{n}_j)$, where $\boldsymbol{f}_{jk}\in\mathbb{C}^{1\times M}$ is the MMSE detection vector for the kth MS in the jth cell and is expressed as [7]

$$\boldsymbol{f}_{jk} = \rho_{\mathrm{ul}}^{-1/2}\tilde{d}_{jjk}\hat{\boldsymbol{h}}_{jjk}^{\mathrm{H}}\left(\hat{\boldsymbol{H}}_{jj}\tilde{\boldsymbol{D}}_{jj}^{2}\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}} + \alpha_j\boldsymbol{I}_M\right)^{-1}, \tag{6}$$

in which the expression of α_j is omitted for brevity.

Then, a proposition about the spectral efficiency in the limit of the number of the BS antennas is presented as follows.

Proposition 1: When the number of the BS antennas, M, tends to infinity and the number of the MSs, K, is fixed, the limiting signal-to-interference-plus-noise ratio (SINR) is

$$\bar{\gamma}_{jk}^{\text{up}} = \frac{1}{\sum_{\substack{l=1 \\ l \neq j}}^{L} \sum_{k'=1}^{K} \left[\boldsymbol{\Psi}_{jl}\right]_{k,k'} \tilde{\beta}_{jlk'k}}, \tag{7}$$

where $\tilde{\beta}_{jlk'k} = \tilde{d}_{jlk'}^{4} \tilde{d}_{jjk}^{-4}$, $\boldsymbol{\Psi}_{jl}$ is the correlation matrix of the pilots and is defined below (2).

Proof: Refer to Appendix 1 for details. □

Based on this proposition, the limit of the spectral efficiency of the uplink transmission is

$$\bar{R}_{\text{up}} = \sum_{j=1}^{L} \sum_{k=1}^{K} \log_2\left(1 + \bar{\gamma}_{jk}^{\text{up}}\right). \tag{8}$$

2.2 Downlink precoding

For the downlink transmission, the received symbol vector at the K MS of the jth BS is written as

$$\boldsymbol{y}_j^{\text{down}} = \frac{1}{\xi_j}\left[\sum_{l=1}^{L} \boldsymbol{D}_{lj} \boldsymbol{H}_{lj}^{\text{H}} \boldsymbol{G}_l \boldsymbol{P}_l^{\text{down}} \boldsymbol{d}_l + \boldsymbol{w}_j\right] \in \mathbb{C}^{K \times 1}, \tag{9}$$

where $\boldsymbol{d}_l \in \mathbb{C}^{K \times 1}$ is a vector composed of the transmitted symbols from the lth cell to the MSs in that cell; $\boldsymbol{P}_l^{\text{down}} \in \mathbb{R}^{K \times K}$ is composed of the power allocation coefficients of the MSs in the lth cell. The elements of $\boldsymbol{d}_l$ are i.i.d. random variables of mean zero and variance one; $\boldsymbol{w}_j \in \mathbb{C}^{K \times 1}$ is the received noise of the K MSs in the jth cell and are composed of i.i.d. $\mathcal{CN}(0, 1)$ variables; $\boldsymbol{G}_l = \xi_l \boldsymbol{F}_l \in \mathbb{C}^{M \times K}$ is the precoding matrix of the BS in the lth cell. In addition, the transmission power is ρ_{down}. With the knowledge of the channel estimates $\hat{\boldsymbol{H}}_{jj}$, the MMSE precoding matrix is $\boldsymbol{F}_j = \left(\hat{\boldsymbol{H}}_{jj} \boldsymbol{D}_{jj}^2 \hat{\boldsymbol{H}}_{jj}^{\text{H}} + \delta_j \boldsymbol{I}_M\right)^{-1} \hat{\boldsymbol{H}}_{jj} \boldsymbol{D}_{jj}$, where the expression of δ_j is omitted for brevity.

Similar to the uplink analysis, it can found that the limit of the spectral efficiency of the downlink transmission is

$$\bar{R}_{\text{down}} = \sum_{j=1}^{L} \sum_{k=1}^{K} \log_2\left(1 + \bar{\gamma}_{jk}^{\text{down}}\right), \tag{10}$$

where

$$\bar{\gamma}_{jk}^{\text{down}} \simeq \frac{1}{\sum_{\substack{l=1 \\ l \neq j}}^{L} \sum_{k'=1}^{K} \left[\boldsymbol{\Psi}_{jl}\right]_{k,k'} \beta_{ljkk'}} \tag{11}$$

is an approximation of the actual SINR, $\beta_{ljkk'} = d_{ljk}^4 d_{llk'}^{-4} [\boldsymbol{P}_l^{\text{down}}]_{k',k'}^2$, and $d_{ljk} \triangleq [\boldsymbol{D}_{lj}]_{k,k}$ is the large-scale fading coefficient. Note that the approximation is based on the similar statistical properties of the large-scale fading and the powers in each cell, which approximately holds as the number of MSs is large in massive MIMO systems.

According to the expressions of the limiting spectral efficiencies in (8) and (10), the limiting spectral efficiency of the two-way transmission can be denoted as

$$\bar{R} \triangleq \eta_{\text{up}} \bar{R}_{\text{up}} + \eta_{\text{down}} \bar{R}_{\text{down}} = \sum_{j=1}^{L} \sum_{k=1}^{K} \left(\eta_{\text{up}} \log_2\left(1 + \bar{\gamma}_{jk}^{\text{up}}\right) + \eta_{\text{down}} \log_2\left(1 + \bar{\gamma}_{jk}^{\text{down}}\right)\right), \tag{12}$$

where η_{up} and η_{down} are the ratios of the uplink transmission interval and the downlink transmission interval to the channel coherence interval, respectively.

From the above analysis, it can be seen that the allocation of the pilots directly affect the correlation of the pilot, which in turn affect the spectral efficiency. Hence, the allocation of the pilots is of importance and will be explicitly presented in the following section.

3 Optimisation of the pilot allocation

Owing to the pilot allocation influences the spectral efficiency at the same time, it is reasonable to optimise the pilot allocation. In this section, the allocation of the pilots is investigated and an algorithm is proposed.

3.1 Pilot allocation

The existing pilot allocation schemes either greedily search for the solution or improve the performance of each cell separately [17–19]. In this section, the pilot allocation is sequentially optimised for each cell with the gradient descent algorithm. Moreover, the overall system performance rather than the performance of each cell is optimised in the allocation process.

To maximise the spectral efficiency, the pilot correlation matrices, $\boldsymbol{\Psi}_{jl}, \forall j = 1, 2, \ldots, L, l = 1, 2, \ldots, L, j \neq l$, should be carefully designed. The design of the pilot correlation matrices is also the design of the pilot allocation. In fact, only the partial pilot correlation matrices need to be designed according to the property of the pilot correlation matrices, which is stated as follows.

Property 1: The correlation matrix, $\boldsymbol{\Psi}_{l_1 l_2}$, is determined by the correlation matrices, $\boldsymbol{\Psi}_{jl_1}$ and $\boldsymbol{\Psi}_{jl_2}$.

Proof: According to the definition of the pilot correlation matrix below (2), we have

$$\boldsymbol{\Psi}_{l_1 l_2} = \boldsymbol{\Phi}_{l_1}^{\text{T}} \boldsymbol{\Phi}_{l_2}^{*} = \boldsymbol{\Phi}_{l_1}^{\text{T}} (\boldsymbol{\Phi}_j^{*} \boldsymbol{\Phi}_j^{\text{T}}) \boldsymbol{\Phi}_{l_2}^{*} = \boldsymbol{\Psi}_{jl_1}^{\text{T}} \boldsymbol{\Psi}_{jl_2}, \tag{13}$$

where the second equation is based on $\boldsymbol{\Psi}_{jj} = \boldsymbol{I}_K$ and the third equation is derived by using the property of a permutation matrix, which is $\boldsymbol{\Psi}_{lj} = \boldsymbol{\Psi}_{jl}^{\text{T}}$. □

With the aid of this property, only the pilot correlation matrices, $\boldsymbol{\Psi}_{j_0 l}, \forall l = 1, 2, \ldots, L, l \neq j_0$, need to be designed, where the j_0th cell is the reference cell and can be any cell in the L cells. Assume the pilot correlation matrices are initialised to identity matrices, i.e. $\boldsymbol{\Psi}_{j_0 l} = \boldsymbol{I}_K, \forall l = 1, 2, \ldots, L, l \neq j_0$. When the pilot allocation is changed, the correlation matrix $\boldsymbol{\Psi}_{j_0 l_0}$ is in accordance changed into

$$\tilde{\boldsymbol{\Psi}}_{j_0 l_0} = \boldsymbol{\Psi}_{j_0 l_0} + t_{l_0} \boldsymbol{Z}_{l_0}, \tag{14}$$

where t_{l_0} is the step size in the domain $[0, 1]$, $\boldsymbol{Z}_{l_0} \in \mathbb{R}^{K \times K}$ is the unknown gradient matrix. According to the fact that the pilot correlation matrices are constrained to be permutation matrices, we have the following properties of the step size and the gradient matrix.

Property 2: (i) The value of the step size is either 0 or 1 in the optimisation process; (ii) the values of the elements of $\boldsymbol{Z}_{l_0}$ can only be 0, 1, or −1; (iii) the sums of every column and every row are 0 s; and (iv) only the diagonal elements of $\boldsymbol{Z}_{l_0}$ can be −1.

It is known that each row or each column of the permutation matrix is constituted by one element with value 1 and other elements with value 0. If we change one permutation matrix $\boldsymbol{\Psi}_{j_0 l_0}$ into another permutation matrix $\tilde{\boldsymbol{\Psi}}_{j_0 l_0}$ as (14), the step size t_{l_0} is set as either 0 or 1 to represent the change of the pilot allocation or not. This will benefit the following design as we only need to

Algorithm 1

1: **for all** ($\bar{\mathbf{Z}}^{(\tilde{k})} \in \mathcal{Z}_{\tilde{k}}$, $b_{l_0}^{(\tilde{k})} = 1$) and $b_{l_0}^{(\tilde{k})} = 0$ **do**
2: Calculate $\mathbf{Z}_{l_0}$ according to (19)
3: **if** $\mathbf{Z}_{l_0}$ may fit Property 2 in the following update or fits Property 2 for $\tilde{k} = K$
4: Calculate $\tilde{f}(\mathbf{Z}_{l_0})|_{t_{l_0}=1}$ for $\tilde{k} < K$ and $f(\mathbf{Z}_{l_0})|_{t_{l_0}=1}$ according to (24) for $\tilde{k} = K$
5: **end if**
6: **end for**

Fig. 2 *Algorithm 1: Construction of the Gradient Matrix and Calculation of the Metric*

decide whether or not to change the pilot correlation matrix $\mathbf{\Psi}_{j_0 l_0}$. As long as the step size is set as the first condition of Property 2, the gradient matrix $\mathbf{Z}_{l_0}$ should satisfy the other conditions of Property 2 to make $\tilde{\mathbf{\Psi}}_{j_0 l_0}$ a permutation matrix.

Correspondingly, $\mathbf{\Psi}_{jl}$ is changed into $\tilde{\mathbf{\Psi}}_{jl}$, and the spectral efficiency in (12) can be changed into

$$\tilde{R} = \sum_{j=1}^{L}\sum_{k=1}^{K}\eta_{\text{up}}\log_2\left(1+\tilde{\gamma}_{jk}^{\text{up}}\right) + \sum_{j=1}^{L}\sum_{k=1}^{K}\eta_{\text{down}}\log_2\left(1+\tilde{\gamma}_{jk}^{\text{down}}\right), \tag{15}$$

where

$$\tilde{\gamma}_{jk}^{\text{up}} = \frac{1}{\sum_{\substack{l=1\\ l\neq j}}^{L}\sum_{k'=1}^{K}\left[\tilde{\mathbf{\Psi}}_{jl}\right]_{k,k'}\tilde{\beta}_{jlk'k}}, \tag{16}$$

$$\tilde{\gamma}_{jk}^{\text{down}} = \frac{1}{\sum_{\substack{l=1\\ l\neq j}}^{L}\sum_{k'=1}^{K}\left[\tilde{\mathbf{\Psi}}_{jl}\right]_{k,k'}\beta_{ljkk'}} \tag{17}$$

are the uplink and the downlink SINRs.

Before presenting the pilot allocation algorithm, a proposition should be given first.

Proposition 2: The metric function in (24) satisfies

$$f(\mathbf{Z}_{l_0})|_{t_{l_0}=1} > 0 \tag{18}$$

is an essential condition for $\tilde{R}|_{t_{l_0}=0} < \tilde{R}|_{t_{l_0}=1}$.

Proof: Refer to Appendix 2 for details. □

Note that the value of the metric function $f(\mathbf{Z}_{l_0})$ is proportional to the derivative of the spectral efficiency $\tilde{R}$. Thus, the positiveness of the metric indicates the increase of the spectral efficiency with the increase of the step size t_{l_0}. With this proposition, the performance may be improved by searching toward the direction that makes $f(\mathbf{Z}_{l_0})|_{t_{l_0}=1} > 0$ rather than searching all the realisations of the pilot correlation matrices. Obviously, this strategy will reduce the computational complexity in pilot allocation. The detailed searching process with be presented in the following subsection.

3.2 Single branch search algorithm for pilot allocation

There are many searching algorithms, such as brute-force search, sphere decoding [22], tabu search [23] and so on. However, the decision tree is composed of too many branches, which make those existing search algorithms inefficient. In this paper, a single branch search algorithm is proposed to reduce the computational complexity.

3.2.1 Gradient matrix and metric: According to Property 2, the gradient matrix, i.e. $\mathbf{Z}_{l_0}$ in (18), is constructed as

$$\mathbf{Z}_{l_0} = \sum_{k=1}^{K} b_{l_0}^{(\tilde{k})}\bar{\mathbf{Z}}^{(\tilde{k})}, \tag{19}$$

where $b_{l_0}^{(\tilde{k})} \in \{0,1\}$ is the selection parameter, $\bar{\mathbf{Z}}^{(\tilde{k})} \in \mathbb{C}^{K\times K}$ is the basic component matrix. According to Property 2, $\bar{\mathbf{Z}}^{(\tilde{k})}$ is designed in the manner that only the $\tilde{k}$th row of this matrix is composed of non-zero elements. Moreover, $[\bar{\mathbf{Z}}^{(\tilde{k})}]_{\tilde{k},\tilde{k}} = -1$, one of the other $K-1$ entries in the $\tilde{k}$th row of this matrix is 1, and the left $K-2$ entries are 0 s. Obviously, there are $K-1$ such realisations of this component matrix. The set composing these different realisations of this component matrix is denoted as $\mathcal{Z}_{\tilde{k}} \triangleq \{\bar{\mathbf{Z}}^{(\tilde{k})}\}$ and the size of $\mathcal{Z}_{\tilde{k}}$ is $K-1$. Depending on the realisations of the component matrix, the selection parameter is set to make the gradient matrix fit the constraints in Property 2.

After the gradient matrix $\mathbf{Z}_{l_0}$ is constructed, the corresponding metric function $f(\mathbf{Z}_{l_0})|_{t_{l_0}=1}$ can be calculated. This process is summarised in the algorithm (see Fig. 2).

Remark 1: According to the previous analysis, the gradient matrix $\mathbf{Z}_{l_0}$ is the sum of parts of the component matrices $\bar{\mathbf{Z}}^{(\tilde{k})}, k = 1,2,\ldots,K$. Therefore, in step 1, we construct all the possible realisations of the component matrix ($\bar{\mathbf{Z}}^{(\tilde{k})} \in \mathcal{Z}_{\tilde{k}}$, $b_{l_0}^{(\tilde{k})} = 1$) as well as no component matrix $b_{l_0}^{(\tilde{k})} = 0$. In step 2, we calculate the gradient matrix with the construction given in step 1.

In step 3, we check whether the construction given in step 1 fits Property 2 or not. If $\tilde{k} < K$, we only need to make sure that the temporary construction may fit Property 2 in the following update. If $\tilde{k} = K$, then the construction cannot be changed anymore, we should make sure that the construction fits Property 2.

In step 4, we calculate the metric function for deciding which construction of the gradient matrix is the optimal. The function $\tilde{f}(\mathbf{Z}_{l_0})|_{t_{l_0}=1}$ is a modified metric function, in which $\tilde{\gamma}_{jk}^{\text{up}}$ and $\tilde{\gamma}_{jk}^{\text{down}}$ in (16) and (17) are replaced with $\bar{\gamma}_{jk}^{\text{up}}$ and $\bar{\gamma}_{jk}^{\text{down}}$ in (7) and (11). Since the gradient matrix $\mathbf{Z}_{l_0}$ may not fit Property 2 for $\tilde{k} < K$, the original SINRs in (7) and (11) are used for $\tilde{k} < K$.

3.2.2 Update and selection: With the constructed gradient matrices and the corresponding values of the metric function, Proposition 2 can be employed to select one of the gradient matrices and update the gradient matrix.

By sequentially choosing $\bar{\mathbf{Z}}^{(\tilde{k})}, k = 1,2,\ldots,K$ in the sets $\mathcal{Z}_{\tilde{k}}, k = 1,2,\ldots,K$, respectively, the gradient matrix in (19) can be updated in each stage, so is the metric function in (18). As can be seen, the construction process is exactly a decision tree, the node is the value of the metric function, and the branches are the selections of the component matrix, and the nodes in the last stage are the leaves. In each selection stage except the final stage, the proposed searching approach discards the nodes with smaller values of the metric function. In the last stage, only one leave is left; if the leave is of negative value, no branch is selected. Obviously, only one branch is kept in each stage, the proposed approach is thus named as the single branch search algorithm.

Algorithm 2

Initialize: $\mathbf{\Psi}_{j_0 l} = \mathbf{I}_K, l = 1, 2, \cdots, L, l \neq j_0, b_{l_0}^{(\tilde{k})} = 0, l_0 = 1, 2, \cdots, L, l_0 \neq j_0, \tilde{k} = 1, 2, \cdots, K$

1: **for all** $l_0 = 1, 2, \cdots, L, l_0 \neq j_0$ $\tilde{k} = 1, 2, \cdots, K$ **do**
2: Execute Algorithm 1
3: **if** $\tilde{k} < K$ or ($\tilde{k} = K$ and he maximum of $f(\mathbf{Z}_{l_0})|_{t_{l_0}=1}$ is positive)
4: Keep the branch with the maximal $\tilde{f}(\mathbf{Z}_{l_0})|_{t_{l_0}=1}$ and discard other branches
5: **else** Discard all the branches
6: **end if**
7: **end for**
8: **Calculate (15) for all** $t_{l_0}, \forall l_0 \neq j_0$, **set** t_{l_0} **for the maximal** $\tilde{R}$

Fig. 3 *Algorithm 2: Proposed single branch search algorithm*

Table 1 Parameters in the simulations

Number of cells L	7	Number of MSs K	10
number of antennas M	196	length of the pilot sequence N_{p}	10
decay exponent in path loss	3.8	shadow fading standard deviation	8 dB
height of the BS array	32 m	centre to edge distance in each cell	1600 m
cell-hole radius	100 m	SNR	20 dB

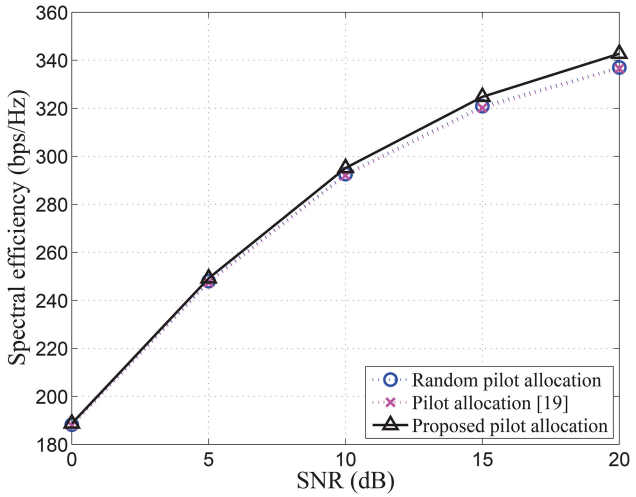

Fig. 4 *Comparison of spectral efficiency versus SNR for different kinds of pilot allocation schemes*

In the above statements, the gradient matrix of one correlation matrix is designed, i.e. $\mathbf{Z}_{l_0}$ in (19). In addition, it can be seen that the gradient matrix, $\mathbf{Z}_{l_0}$, is designed in the condition that other pilot correlation matrices are not updated, i.e. t_l in (14) is $t_l = 0, \forall l \neq l_0, l \neq j_0$. Therefore, the spectral efficiencies corresponding to all the realisations of $t_{l_0}, \forall l_0 \neq j_0$ should be compared to set t_{l_0}. For the sake of clarity, the proposed pilot allocation algorithm is summarised in Fig. 3.

In this algorithm, steps from steps 1 to 7 are used to select the single branch for each cell l_0 and each MS $\tilde{k}$ sequentially. Moreover, step 2 is used to calculate metric functions. Steps from steps 3 to 6 are used to select the single branch. Step 8 is used to calculate the spectral efficiency for all the possible realisations of the pilot correlation matrix and then select the realisation with the maximum spectral efficiency.

It can be seen that the proposed approach allocates the pilots of all the cells to maximise system spectral efficiency, while the approaches in [17–19] only improve the performance of each cell separately, and the system performance may not be improved. In the following, the computational complexity of the proposed approach will be analysed.

3.3 Complexity analysis

In this subsection, the computational complexities of the proposed approach and other existing pilot allocation approaches are analysed. The big O notation, $O(n)$, is used to denote the computational complexity that is linear in $n \in \mathbb{R}^+$ [24].

According to the presentation of the proposed pilot allocation approach, there are $(L-1)(K+1)K/2$ iterations overall. In each iteration, the multiplications in the order of $O(LK^3)$ are needed. Hence, the computational complexity of the iterations is $O(L^2K^5)$. In the final stage, multiplications in the order of $O(L^2K^2)$ are needed in each calculation of the spectral efficiency. The overall computational complexity of the proposed approach is $O(2^{L-1}L^2K^2 + L^2K^5)$.

In contrast, the optimal pilot allocation approach implements brute-force search. The number of iterations is $(K!)^{L-1}$. In each iteration, multiplications in the order of $O(L^2K^2)$ are needed. Thus, the computational complexity is $O((K!)^{L-1}L^2K^2)$. It can be easily found that the proposed pilot allocation approach is of much lower computational complexity than the optimal pilot allocation approach. Besides, the computational complexities of the random pilot assignment and the approach in [19] are negligible.

Till now, the proposed pilot allocation approach has been presented. In addition, the computational complexity of the proposed pilot allocation approach has been analysed and compared with the existing approaches. In the following, the performance of the pilot allocation approach will be demonstrated by simulation results.

4 Simulation results

In this section, simulation results are provided to demonstrate the performance of the proposed pilot allocation. In addition, the proposed approach is compared with the existing approaches to verify the analysis. The simulation parameters are listed in Table 1. Moreover, the MSs distribute uniformly in the cell. The ratio of the uplink/downlink transmission interval to the channel coherence interval is 3/7. With these simulation parameters, the overall computational complexity of the proposed approach is $O(10^6)$, the computational complexity of the optimal pilot allocation is $O(10^{43})$. As can be seen, the proposed approach is much more realisable for massive MIMO systems. In the following simulations, the actual spectral efficiency rather than the asymptotic spectral efficiency is used for comparison.

In Fig. 4, the spectral efficiency versus the signal-to-noise ratio (SNR) for different kinds of pilot allocation schemes is illustrated. In this simulation, the proposed approach, the random pilot allocation, and the pilot allocation in [19] are compared. It can be seen that the spectral efficiency of the proposed pilot allocation is higher than that of the other two approaches. The reason is that the proposed approach allocates the pilots according to the large-scale fading of the MSs for all the cells, while the random pilot allocation does not use any information and the approach in [19] only allocates the pilots for each cell separately. In addition, the performance gap between the proposed approach and the other approaches increases as the SNR increases, which is because the interference becomes a dominating factor as the SNR increases.

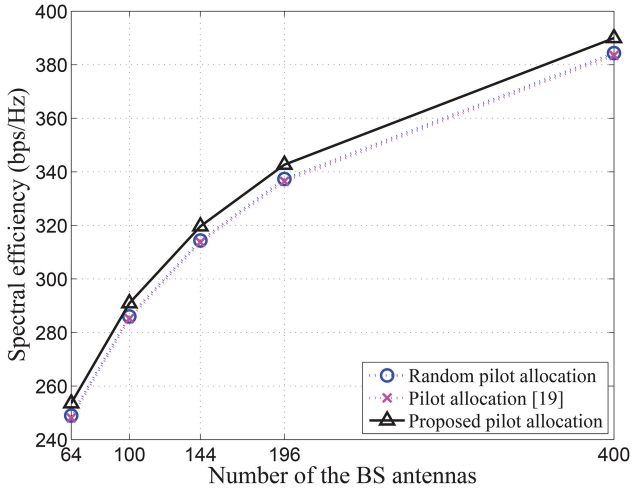

Fig. 5 *Comparison of spectral efficiency versus the number of the BS antennas for different kinds of pilot allocation schemes*

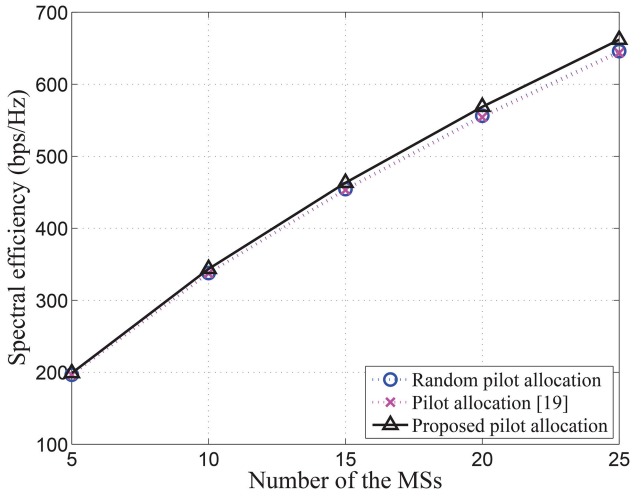

Fig. 6 *Comparison of spectral efficiency versus the number of the MSs for different kinds of pilot allocation schemes*

Fig. 5 illustrates the spectral efficiencies with different number of the BS antennas. The proposed approach has higher spectral efficiency than the other approaches and the performance gap increases as the number of the BS antennas increases. This is because the proposed approach is designed to maximise the spectral efficiency in the limit of the number of the BS antennas and is able to perform better when the number of the BS antennas increases. As can be seen, the increased number of the BS antennas corresponds to the configuration of the massive MIMO systems.

In Fig. 6, the spectral efficiency versus the number of the MSs is depicted. As can be seen, the proposed approach is of higher spectral efficiency than the other approaches and the superiority of the proposed approach is more evident when the number of the MSs increases. Since the interference increases as the number of the MSs increases, this result also demonstrates that the proposed approach can mitigate more interference than the other approaches.

5 Conclusion

In this paper, a pilot allocation approach has been proposed for TDD multi-cell massive MIMO systems. According to the MMSE principle employed in the channel estimation and the signal detection, the spectral efficiency as well as its asymptotic value in the infinite number of the BS antennas are analysed. By analysing the spectral efficiency, a simple metric is derived for discarding unnecessary branches in the decision tree of the pilot allocation. Correspondingly, a search algorithm is proposed to keep the best branch in each decision stage. The simulation results verify the analysis of the spectral efficiency and demonstrate the superiority of the proposed pilot allocation scheme.

6 Acknowledgment

The authors thank the editor Prof. Xuemin Shen and the anonymous reviewers for the helpful comments provided in the revision of this paper. This research was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LQ16F010007, Project 61601152 and Project 61372093 supported by National Natural Science Foundation of China.

7 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**, (1), pp. 40–60

[2] Marzetta, T.L.: 'Noncooperative cellular wireless with unlimited numbers of base station antennas', *IEEE Trans. Wirel. Commun.*, 2010, **9**, (11), pp. 3590–3600

[3] Larsson, E.G., Edfors, O., Tufvesson, F., *et al.*: 'Massive MIMO for next generation wireless systems', *IEEE Commun. Mag.*, 2014, **52**, (2), pp. 186–195

[4] Boccardi, F., Heath, R.W., Lozano, A., *et al.*: 'Five disruptive technology directions for 5G', *IEEE Commun. Mag.*, 2014, **52**, (2), pp. 74–80

[5] Andrews, J.G., Buzzi, S., Choi, W., *et al.*: 'What will 5G Be?', *IEEE J. Sel. Areas. Commun.*, 2014, **32**, (6), pp. 1065–1082

[6] Marzetta, T.L.: 'Massive MIMO: an introduction', *Bell Labs Tech. J.*, 2015, **20**, pp. 11–22

[7] Krishnan, N., Yates, R., Mandayam, N.B.: 'Uplink linear receivers for multi-cell multiuser MIMO with pilot contamination: large system analysis', *IEEE Trans. Wirel. Commun.*, 2014, **13**, (8), pp. 4360–4373

[8] Ngo, H.Q., Matthaiou, M., Larsson, E.G.: 'Performance analysis of large scale MU-MIMO with optimal linear receivers'. Proc. 2012 Swedish Communication Technologies Workshop (Swe-CTW), Lund, Sweden, October 2012, pp. 59–64

[9] Hoydis, H., Brink, S.t., Debbah, M.: 'Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?', *IEEE J. Sel. Areas. Commun.*, 2013, **31**, (2), pp. 160–171

[10] Jose, J., Ashikhmin, A., Marzetta, T.L., *et al.*: 'Pilot contamination and precoding in multi-cell TDD systems', *IEEE Trans. Wirel. Commun.*, 2011, **10**, (8), pp. 2640–2651

[11] Zhang, Q., Jin, S., McKay, M., *et al.*: 'Power allocation schemes for multicell massive MIMO systems', *IEEE Trans. Wirel. Commun.*, 2015, **14**, (11), pp. 5941–5955

[12] Huh, H., Tulino, A.M., Caire, G.: 'Network MIMO with linear zero-forcing beamforming: large system analysis, impact of channel estimation, and reduced-complexity scheduling', *IEEE Trans. Inf. Theory*, 2012, **58**, (5), pp. 2911–2934

[13] Yin, H., 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**, (2), pp. 264–273

[14] Ngo, H.Q., Larsson, E.G.: 'EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays'. Proc. 2012 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, March 2012, pp. 3249–3252

[15] Wang, H., Huang, Y., Jin, S., *et al.*: 'Performance analysis on precoding and pilot scheduling in very large MIMO multi-cell systems'. Proc. 2013 IEEE Wireless Communications and Networking Conf. (WCNC), Shanghai, China, April 2013, pp. 2722–2726

[16] Zhu, X., Dai, L., Wang, Z.: 'Graph coloring based pilot allocation to mitigate pilot contamination for multi-cell massive MIMO systems', *IEEE Commun. Lett.*, 2015, **19**, (10), pp. 1842–1845

[17] Mochaourab, R., Bjornson, E., Bengtsson, M.: 'Pilot clustering in asymmetric massive MIMO networks'. Proc. 2015 IEEE 16th Int. Workshop Signal Processing Advances Wireless Communications (SPAWC), Stockholm, Sweden, June 2015, pp. 231–235

[18] Ahmadi, H., Farhang, A., Marchetti, N., *et al.*: 'A game theoretic approach for pilot contamination avoidance in massive MIMO', *IEEE Wirel. Commun. Lett.*, 2016, **5**, (1), pp. 12–15

[19] Zhu, X., Wang, Z., Dai, L., *et al.*: 'Smart pilot assignment for massive MIMO', *IEEE Commun. Lett.*, 2015, **19**, (9), pp. 1644–1647

[20] Elijah, O., Leow, C.Y., Rahman, T.A., *et al.*: 'A comprehensive survey of pilot contamination in massive MIMO-5G system', *IEEE Commun. Surv. Tutor.*, 2016, **18**, (2), pp. 905–923

[21] Kay, S.M.: '*Fundamentals of statistical signal processing: estimation theory*' (Prentice-Hall, 1993)

[22] Agrell, E., Eriksson, T., Vardy, A., *et al.*: 'Closest point search in lattices', *IEEE Trans. Inf. Theory*, 2002, **48**, (8), pp. 2201–2214

[23] Srinidhi, N., Mohammed, S.K., Chockalingam, A., *et al.*: 'Low-complexity near-ML decoding of large non-orthogonal STBCs using reactive tabu search'. Proc. 2009 IEEE Int. Symp. Information Theory (ISIT), Seoul, Korea, June 2009, pp. 1993–1997

[24] Golub, G.H., Loan, C.F.V.: '*Matrix computations*' (The Johns Hopkins University Press, 1996)

8 Appendix

8.1 Appendix 1: Proof of Proposition 1

First, the signal power is calculated. According to (6), the signal power is

$$S_{jk}^{\mathrm{up}} = \tilde{d}_{jjk}^{4}\left|\left[\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}}\left(\hat{\boldsymbol{H}}_{jj}\tilde{\boldsymbol{D}}_{jj}^{2}\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}}+\alpha_j\boldsymbol{I}_M\right)^{-1}\boldsymbol{H}_{jj}\right]_{k,k}\right|^2, \tag{20}$$

By utilising the matrix inversion identity and the limiting property $\lim_{M\to\infty}(1/M)\boldsymbol{H}_{jj}^{\mathrm{H}}\boldsymbol{H}_{jj}=\boldsymbol{I}_K$, we have

$$\lim_{M\to\infty}\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}}\left(\hat{\boldsymbol{H}}_{jj}\tilde{\boldsymbol{D}}_{jj}^{2}\hat{\boldsymbol{H}}_{jj}^{\mathrm{H}}+\alpha_j\boldsymbol{I}_M\right)^{-1}\boldsymbol{H}_{jj}=\tilde{\boldsymbol{D}}_{jj}^{-2}.$$

Substituting the above equation into (20), we have

$$\lim_{M\to\infty} S_{jk}^{\mathrm{up}} = 1. \tag{21}$$

The interference plus noise power is

$$IN_{jk}^{\mathrm{up}} = \rho_{\mathrm{ul}}\sum_{\substack{k'=1\\k'\neq k}}^{K}\left|\boldsymbol{f}_{jk}\boldsymbol{h}_{jjk'}\right|^2\tilde{d}_{jjk'}^{2}+\rho_{\mathrm{ul}}\sum_{\substack{l=1\\l\neq j}}^{L}\|\,\boldsymbol{f}_{jk}\boldsymbol{H}_{jl}\tilde{\boldsymbol{D}}_{jl}\,\|^2+\|\,\boldsymbol{f}_{jk}\,\|^2 \tag{22}$$

.

By utilising the limiting property above, we have

$$\lim_{M\to\infty} IN_{jk}^{\mathrm{up}} = \tilde{d}_{jjk}^{-4}\sum_{\substack{l=1\\l\neq j}}^{L}\|\left[\boldsymbol{\Psi}_{jl}\tilde{\boldsymbol{D}}_{jl}^{2}\right]_{k,:}\|^2. \tag{23}$$

Thus, we have the limiting SINR shown in (7).

8.2 Appendix 2: Proof of Proposition 2

According to (13), we have

$$\tilde{\boldsymbol{\Psi}}_{jl}=\begin{cases}\boldsymbol{\Psi}_{j_0j}^{\mathrm{T}}\boldsymbol{\Psi}_{j_0l}, & j\neq l_0, l\neq l_0\\ \boldsymbol{\Psi}_{j_0j}^{\mathrm{T}}\tilde{\boldsymbol{\Psi}}_{j_0l_0}, & j\neq l_0, l=l_0\\ \boldsymbol{I}_K, & j=l_0, l=l_0\\ \tilde{\boldsymbol{\Psi}}_{j_0l_0}^{\mathrm{T}}\boldsymbol{\Psi}_{j_0l}, & j=l_0, l\neq l_0\end{cases}.$$

According to the expression of $\tilde{R}$ in (15) and the above property, we have

$$\begin{aligned}\frac{\partial\tilde{R}}{\partial t_{l_0}} \propto & \sum_{\substack{j=1\\j\neq l_0}}^{L}\sum_{k=1}^{K}\eta_{\mathrm{up}}\frac{-(\tilde{\gamma}_{jk}^{\mathrm{up}})^2}{(1+\tilde{\gamma}_{jk}^{\mathrm{up}})}\sum_{k'=1}^{K}\left[\boldsymbol{\Psi}_{j_0j}^{\mathrm{T}}\boldsymbol{Z}_{l_0}\right]_{k,k'}\tilde{\beta}_{jl_0k'k}\\ &+\sum_{k=1}^{K}\eta_{\mathrm{up}}\frac{-(\tilde{\gamma}_{l_0k}^{\mathrm{up}})^2}{(1+\tilde{\gamma}_{l_0k}^{\mathrm{up}})}\sum_{\substack{l=1\\l\neq l_0}}^{L}\sum_{k'=1}^{K}\left[\boldsymbol{Z}_{l_0}^{\mathrm{T}}\boldsymbol{\Psi}_{j_0l}\right]_{k,k'}\tilde{\beta}_{l_0lk'k}\\ &+\sum_{\substack{j=1\\j\neq l_0}}^{L}\sum_{k=1}^{K}\eta_{\mathrm{down}}\frac{-(\tilde{\gamma}_{jk}^{\mathrm{down}})^2}{(1+\tilde{\gamma}_{jk}^{\mathrm{down}})}\sum_{k'=1}^{K}\left[\boldsymbol{\Psi}_{j_0j}^{\mathrm{T}}\boldsymbol{Z}_{l_0}\right]_{k,k'}\beta_{l_0jkk'}\\ &+\sum_{k=1}^{K}\eta_{\mathrm{down}}\frac{-(\tilde{\gamma}_{l_0k}^{\mathrm{down}})^2}{(1+\tilde{\gamma}_{l_0k}^{\mathrm{down}})}\sum_{\substack{l=1\\l\neq l_0}}^{L}\sum_{k'=1}^{K}\left[\boldsymbol{Z}_{l_0}^{\mathrm{T}}\boldsymbol{\Psi}_{j_0l}\right]_{k,k'}\beta_{ll_0kk'}\\ \triangleq & f(\boldsymbol{Z}_{l_0}).\end{aligned} \tag{24}$$

Furthermore, it can be easily found that the second-order derivative $\partial^2\tilde{R}/\partial t_{l_0}^2$ is positive, which means that the spectral efficiency, $\tilde{R}$, is a convex function of the variable t_{l_0}.

Based on the first-order and second-order derivatives, it can be found that $\partial\tilde{R}/\partial t_{l_0}|_{t_{l_0}=1}>0$ is an essential condition of $\tilde{R}\big|_{t_{l_0}=0}<\tilde{R}\big|_{t_{l_0}=1}$. In addition, the bigger $\partial\tilde{R}/\partial t_{l_0}|_{t_{l_0}=1}$, the more likely that $\tilde{R}\big|_{t_{l_0}=0}<\tilde{R}\big|_{t_{l_0}=1}$ holds. According to the definition of $f(\boldsymbol{Z}_{l_0})$ in (24), $f(\boldsymbol{Z}_{l_0})\big|_{t_{l_0}=1}$ is a metric that is proportional to $\partial\tilde{R}/\partial t_{l_0}\big|_{t_{l_0}=1}$. Therefore, $f(\boldsymbol{Z}_{l_0})\big|_{t_{l_0}=1}>0$ is an essential condition of $\tilde{R}\big|_{t_{l_0}=0}<\tilde{R}\big|_{t_{l_0}=1}$.