SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

August 2018, Vol. 61 082302:1–082302:10
https://doi.org/10.1007/s11432-017-9353-8

# User scheduling for downlink FD-MIMO systems under Rician fading exploiting statistical CSI

Xiao LI*, Tingting SUN, Nana QIN, Shi JIN & Xiqi GAO

*National Mobile Communications Research Laboratory, Southeast University, Nanjing* 210096*, China*

Received 14 November 2017/Revised 28 December 2017/Accepted 24 January 2018/Published online 27 April 2018

**Abstract** In this paper, we consider the user scheduling algorithm for downlink full-dimension multiple-input multiple-output (FD-MIMO) system under Rician fading channels. We assume that two-dimensional (2D) large-scale antenna array is deployed at base station (BS). An approximation of user's signal-to-interference-plus-noise ratio (SINR) and a lower bound of user's average signal-to-leakage-plus-noise ratio (SLNR) are derived. Based on these, two user scheduling algorithms exploiting only statistical channel state information (CSI) are proposed. The proposed algorithms take both the achievable sum rate and fairness into account. Simulation results reveal that the proposed user scheduling algorithms can make good trade-off between the achievable rate and fairness.

**Keywords** scheduling, FD-MIMO, fairness, Rician fading

**Citation** Li X, Sun T T, Qin N N, et al. User scheduling for downlink FD-MIMO systems under Rician fading exploiting statistical CSI. Sci China Inf Sci, 2018, 61(8): 082302, https://doi.org/10.1007/s11432-017-9353-8

## 1 Introduction

In recent years, massive multiple-input multiple-output (MIMO), which adopts large numbers of antennas at base station (BS), has been recognized as one of the key techniques in future wireless communication systems [1–4]. It was shown that the effect of fast fading vanishes as the number of antennas grows large [5], and the transmit power of each user on the uplink can be reduced [6]. However, in practice, there remains some challenges for massive MIMO systems. As pointed out in [7], the number of deployed antennas is often limited by the physical space of BS and the carrier frequencies. This limitation has motivated full-dimension MIMO (FD-MIMO) — a kind of systems that places a large number of antennas in a two-dimensional (2D) grid [7, 8]. With this antenna configuration, three-dimensional (3D) beamforming method [9,10], which can exploit the vertical dimension as well as the horizontal dimension, are proposed.

The acquisition of channel state information (CSI) at BS is another challenge [11]. Since most cellular systems today are in frequency-division duplexing (FDD) mode [12, 13], it has become a subject of considerable interest in making massive MIMO work for systems in FDD mode. In FDD systems, the BSs get access to the CSI through feedback link. The transmission and scheduling algorithms based on instantaneous CSI cause great overhead in the feedback link as the number of transmit antennas and users grow large. Moreover, the complexity of user scheduling also becomes large. To overcome these problems, some transmission algorithms exploiting the statistical information of the channel were proposed. In [9],

* Corresponding author (email: li_xiao@seu.edu.cn)

 

a two-stage precoding scheme using both statistical and part of instantaneous CSI was proposed and corresponding user scheduling algorithm was proposed in [14] for massive MIMO system under Rayleigh fading channels. To further reduce the information required at the BS for FD-MIMIO system, a 3D beamforming transmission and scheduling algorithm exploiting only statistical CSI was proposed in [15], also under Rayleigh fading channels.

However, the Rayleigh fading channel model cannot capture the characteristics ofthe line-of-sight (LOS) component between the transmitter and the receiver. Taking into account the LOS component, Ref. [16] investigates the downlink transmission and scheduling algorithm for FD-MIMO systems under Rician fading channels. In this algorithm, the BS schedules the users which are dominated by the LOS component and orthogonal either in horizontal or vertical LOS component. Although this algorithm can achieve high throughput, the fairness between users is totally ignored, which leads to poor user experience for users with relatively small Rician $K$-factors.

Motivated by the above observations, we investigate the user scheduling algorithm for downlink FD-MIMO systems in this paper, taking into account both the achievable rate and fairness. With the assumption that only statistical CSI is available at BS, an approximation of user's signal-to-interference-plus-noise ratio (SINR) and a lower bound of user's average signal-to-leakage-plus-noise ratio (SLNR) are derived. Based on these, we propose two scheduling algorithms exploiting only statistical CSI. Simulation results reveal that the proposed algorithms can make good trade-off between throughput and fairness.

The remainder of this paper is organized as follows. The system model is introduced in Section 2. In Section 3, an approximation of the SINR and a lower bound of the average SLNR are derived, and then an SINR based and an SLNR based user scheduling algorithms are proposed. The results of simulation are provided in Section 4, and Section 5 concludes the paper.

Notation. Vectors are represented as columns and are denoted in lower-case bold-faced characters, and matrices are represented in upper-case bold-faced. The superscript $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$, and $(\cdot)^\dagger$ indicate the matrix transpose, conjugate, and conjugate-transpose, respectively. The complex number field is denoted as $\mathbb{C}$, $\mathrm{E}\{\cdot\}$ represents the expectation of the input random variable (RV), $\mathrm{Var}\{\cdot\}$ evaluates the variance of the input RV, $\mathrm{Cov}\{\cdot,\cdot\}$ measures the joint variability of two RVs, $\otimes$ represents the Kronecker product, and $\mathbf{I}_d$ denotes a $d \times d$ identity matrix.

## 2 System model

In this paper, We consider the downlink of a single-cell FD-MIMO system. It has $L$ single-antenna users. The BS is equipped with a rectangular $M \times N$ antenna array which has $M$ rows in the vertical dimension and each row has $N$ antennas. The distances between two neighboring antennas in each row and column are both half of the carrier wavelength. The BS can serve at most $U$ users simultaneously.

Assume that after user scheduling, a total number of $U_t$ ($U_t \leqslant U$) users are scheduled. The signal received at user $u$ can be expressed as

$$y_u = \sqrt{p_u}\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u x_u + \sum_{i=1, i\neq u}^{U_t} \sqrt{p_i}\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_i x_i + n_u, \tag{1}$$

where $\boldsymbol{h}_u^{\mathrm{T}} \in \mathbb{C}^{1\times MN}$ represents the channel vector from the BS to user $u$, $\boldsymbol{w}_i \in \mathbb{C}^{MN\times 1}$ is the normalized beamforming vector of user $i$, $x_i$ is the data symbol for user $i$ satisfying $\mathrm{E}\{|x_i|^2\} = 1$, $n_u \sim CN(0, \sigma_u^2)$ is the complex noise, $p_i$ is transmit power to user $i$ satisfying total power constraint $\sum_{i=1}^{U_t} p_i \leqslant P$. Here, we assume that the transmit power is allocated equally to the scheduled users, i.e., $p_i = P/U_t$.

For the channel vector, Rician fading is considered. Under this model, the channel vector $\boldsymbol{h}_u$ consists of a deterministic LOS component and a Rayleigh-distributed random component. It can be written as

$$\boldsymbol{h}_u = \sqrt{\frac{K_u}{K_u+1}}\overline{\boldsymbol{h}}_u + \sqrt{\frac{1}{K_u+1}}\boldsymbol{h}_{w,u}, \tag{2}$$

where Rician-$K$ factor $K_u$ is the ratio between the channel power of LOS and non-LOS part, $\overline{\boldsymbol{h}}_u$ is the deterministic component, $\boldsymbol{h}_{w,u}$ denotes the random component with i.i.d. zero mean and unit variance complex Gaussian entries, and $\boldsymbol{h}_{w,u}, u = 1, \ldots, L$ are independent from each other. For the considered antenna array, the deterministic component $\overline{\boldsymbol{h}}_u$ of user $u$ can be given by [17]

$$\overline{\boldsymbol{h}}_u = \boldsymbol{a}_t^{(v)}(\theta_u) \otimes \boldsymbol{a}_t^{(h)}(\theta_u, \varphi_u), \tag{3}$$

with

$$\boldsymbol{a}_t^{(v)}(\theta_u) = \left[1, \mathrm{e}^{-\mathrm{j}\pi \sin\theta_u}, \ldots, \mathrm{e}^{-\mathrm{j}\pi(M-1)\sin\theta_u}\right]^{\mathrm{T}}, \tag{4}$$

and

$$\boldsymbol{a}_t^{(h)}(\theta_u, \varphi_u) = \left[1, \mathrm{e}^{-\mathrm{j}\pi\phi_u}, \cdots, \mathrm{e}^{-\mathrm{j}\pi(N-1)\phi_u}\right]^{\mathrm{T}}, \tag{5}$$

where $\phi_u = \cos\theta_u \sin\varphi_u$, $\theta_u$ and $\varphi_u$ are the vertical and horizontal angles of departure (AoD) of user $u$. Assume that each user has its own perfect effective CSI, while the BS has only statistical CSI, i.e., $K_u$, $\sigma_u^2$, $\overline{\boldsymbol{h}}_u$ or $\boldsymbol{a}_t^{(v)}(\theta_u)$ and $\boldsymbol{a}_t^{(h)}(\theta_u, \varphi_u)$. These statistical information varies relatively slow, and therefore can be obtained at the BS by long-term feedback [18].

Here, we adopt the 3D beamforming algorithm proposed in [16]. For the scheduled user $u$, let us define

$$\boldsymbol{A}_u^{(v)} = \boldsymbol{F}_M \left(\boldsymbol{a}_t^{(v)}(\theta_u)\right)^* \left(\boldsymbol{a}_t^{(v)}(\theta_u)\right)^{\mathrm{T}} \boldsymbol{F}_M^{\dagger}, \tag{6}$$

and

$$\boldsymbol{A}_u^{(h)} = \boldsymbol{F}_N \left(\boldsymbol{a}_t^{(h)}(\theta_u, \varphi_u)\right)^* \left(\boldsymbol{a}_t^{(h)}(\theta_u, \varphi_u)\right)^{\mathrm{T}} \boldsymbol{F}_N^{\dagger}, \tag{7}$$

where $\boldsymbol{F}_M \in \mathbb{C}^{M\times M}$ and $\boldsymbol{F}_N \in \mathbb{C}^{N\times N}$ are both unitary DFT matrices. Assume that the largest diagonal element of $\boldsymbol{A}_u^{(v)}$ is the $\overline{m}_u$-th diagonal element, and the largest diagonal element of $\boldsymbol{A}_u^{(h)}$ is the $\overline{n}_u$-th diagonal element. Then, the beamforming vector of user $u$ is

$$\boldsymbol{w}_u = \left(\boldsymbol{F}_M^{\dagger}\right)_{\overline{m}_u} \otimes \left(\boldsymbol{F}_N^{\dagger}\right)_{\overline{n}_u}. \tag{8}$$

## 3 User scheduling algorithms

The user scheduling algorithm in [16] divides the users into $MN$ clusters, so that users in different clusters are orthogonal in either horizontal or vertical LOS component, i.e., $\overline{m}_i \neq \overline{m}_u$ or $\overline{n}_i \neq \overline{n}_u$, if user $i$ and user $u$ are in different clusters. Then, select the user whose Rician-$K$ factor is the largest in its cluster. It is obvious that in this algorithm, only the users who have the largest Rician $K$-factor can get the opportunity to be served while the others may never be served. This is unfair for the users whose Rician $K$-factors are relatively small. A straight-forward method to achieve good fairness is to select the user in a Round-Robin fashion in each cluster so that every user can be served in turn. However, it does not take into account the interference from other clusters. Since the number of antennas is usually limited in practical systems, there might be certain interference between the adjacent clusters. Here we will take Round-Robin user scheduling as a performance baseline.

In this section, we investigate the scheduling algorithm that takes into account both the achievable rate and fairness, and propose two algorithms, based on the SINR and SLNR respectively.

### 3.1 SINR based user scheduling

Inspired by [19], we propose a user scheduling algorithm taking into account both the fairness and the interference between different users. This algorithm works by maintaining a priority vector, i.e., $C = \{c_1, \ldots, c_L\}$, where $c_i \in [0, Q]$ is an integer representing the priority of user $i$. We first initialize every user's priority to 0, i.e., $c_i = 0$, $i = 1, \ldots, L$. At the end of every scheduling slot, increase the priorities of the unselected users by 1 until they reach the highest priority $Q$, while let the priorities of the selected users be 0. At every scheduling slot, we select users in such an order that the level of priority

goes from high to low. When we reach a certain level of priority, we find all the users with this level of priority and satisfying some certain interference constraints with respect to the already selected users in this scheduling slot, and select the user which can achieve the highest rate among them. Repeat this until no user with this level of priority can be selected. Then, we move on to the next lower level of priority and repeat the searching process until we reach the lowest level of priority or $U$ users has been selected. Note that for the first selected user at every scheduling slot, we just select the one that can achieve the highest rate among the users with the highest priority, since no user has been selected to cause interference.

What follows next is to determine the interference constraints and which user among the candidate users can achieve the highest rate. To achieve high sum rate, we would like that the SINR of the selected user will not decrease much due to the interference from the candidate user, and the candidate user could have high SINR. From (1), the SINR of user $u$ can be expressed as

$$\mathrm{SINR}_u = \frac{\frac{P}{U_t}\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2}{\sigma_u^2 + \frac{P}{U_t}\sum_{i=1,i\neq u}^{U_t}\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_i\right|^2}. \tag{9}$$

However, user scheduling based on the instantaneous SINR requires instantaneous CSI at the BS, which leads to large amount of feedback overhead in FDD systems. The following theorem gives an approximation of the SINR of user $u$, which depends only on the statistical CSI. We will derive user scheduling algorithm based on the approximation of SINR.

**Theorem 1.** When $MN \to \infty$, we have that $\frac{1}{MN}\left(\mathrm{SINR}_u - \overline{\mathrm{SINR}}_u\right) \xrightarrow{\mathrm{P}} 0$, where

$$\overline{\mathrm{SINR}}_u = \frac{\frac{P}{U_t}D_u}{\sigma_u^2 + \frac{P}{U_t}\sum_{i=1,i\neq u}^{U_t} I_{u,i}}, \tag{10}$$

$$D_u = \frac{K_u}{K_u+1}\left(\alpha_u^{(v)}\right)_{\overline{m}_u}\left(\alpha_u^{(h)}\right)_{\overline{n}_u} + \frac{1}{K_u+1}, \tag{11}$$

$$I_{u,i} = \frac{K_u}{K_u+1}\left(\alpha_u^{(v)}\right)_{\overline{m}_i}\left(\alpha_u^{(h)}\right)_{\overline{n}_i} + \frac{1}{K_u+1}, \tag{12}$$

$(\alpha_u^{(v)})_i$ and $(\alpha_u^{(h)})_j$ are the $i$-th diagonal element of $\boldsymbol{A}_u^{(v)}$ and $j$-th diagonal element of $\boldsymbol{A}_u^{(h)}$, and $\xrightarrow{\mathrm{P}}$ represents convergence in probability.

*Proof.* See Appendix A for the proof of Theorem 1.

Based on the above theorem, we use $\overline{\mathrm{SINR}}_u$ which is more tractable and depends only on statistical CSI, to approximate $\mathrm{SINR}_u$ when $M$ and $N$ become large. Then we derive user scheduling algorithm based on $\overline{\mathrm{SINR}}_u$. Let us rewrite (10) as

$$\overline{\mathrm{SINR}}_u = \frac{1}{\frac{\sigma_u^2 U_t}{PD_u} + \sum_{i=1,i\neq u}^{U_t}\frac{I_{u,i}}{D_u}}, \tag{13}$$

and define

$$J_{u,i} = I_{u,i}/D_u. \tag{14}$$

It can be seen that the SINR approximation (13) increases as $J_{u,i}$ decreases. To make sure that $\overline{\mathrm{SINR}}_i$ of the selected user $i$ will not decreases much due to the candidate user $u$ and the SINR of the candidate user $u$ could be high, we would like that $J_{u,i}$ and $J_{i,u}$ are both below some certain threshold $\zeta$, so that these are negligible. Therefore, we choose the interference constraints to be $J_{u,i} \leqslant \zeta$ and $J_{i,u} \leqslant \zeta$. Moreover, from (13), we can see that when the interference constraints are satisfied so that $J_{u,i}$, $i \neq u$ is negligible, $\overline{\mathrm{SINR}}_u$ increases as $D_u$ increases. Therefore, the user that achieve the highest rate among the candidate users can be considered as the one with the largest $D_u$ satisfying the constraints of $J_{u,i}$ and $J_{i,u}$ should be scheduled.

Based on these analysis, we propose the following SINR based user scheduling algorithm. Define $\Omega_c$ to be the set of unselected users with priority $c$ and satisfying the interference constraints with respect to

the already selected users, and $S$ to be the set of already selected users. At every scheduling slot, firstly initialize $S$ to be empty, find the corresponding priority denoted as $c_{\max}$, and let $c = c_{\max}$. Then, find the corresponding $\Omega_c$, and select the user with the largest $D_u$ from $\Omega_c$, i.e., $s = \arg\max_{u\in\Omega_c} D_u$. Add user $s$ into the set of selected users $S$, and removed it from $\Omega_c$. Next, remove all the users in $\Omega_c$ that does not satisfy the interference constraints with any selected user in $S$. After that, add the user with the largest $D_u$ in $\Omega_c$ into $S$ and remove it from $\Omega_c$. Then, repeat the previous process for users in $\Omega_c$ until there is no user left in $\Omega_c$. Then, move on to the next lower level of priority, i.e., $c = c - 1$, and repeat the previous process until $U$ users have been selected or we reach the lowest level of priority and no user can be selected. The proposed algorithm is summarized as follows.

**Step 1.** Initialize $S = \emptyset$, the number of already selected users $l = 0$, $c = \max_{k=1,\ldots,L} c_k$.

**Step 2.** Find $\Omega_c = \{k|c_k = c, k = 1, \ldots, L\}$.

**Step 3.** If $S \neq \emptyset$, compute $J_{u,i}$ and $J_{i,u}$ for each user $u \in \Omega_c$ and user $i \in S$, and remove user $u$ that does not satisfy $J_{u,i} \leqslant \zeta$ and $J_{i,u} \leqslant \zeta$ from $\Omega_c$.

**Step 4.** Find $s = \arg\max_{u\in\Omega_c} D_u$, and update $S = S \cup s$, $\Omega_c = \Omega_c \backslash s$, $l = l + 1$. Repeat Steps 3 and 4 until $\Omega_c = \emptyset$ or $l = U$. Go to Step 5 when $\Omega_c = \emptyset$; and go to Step 6 when $l = U$.

**Step 5.** Let $c = c - 1$. Go to Step 2 when $c \geqslant 0$, and go to Step 6 when $c < 0$.

**Step 6.** Set $c_k = 0$ for $k \in S$, and $c_k = \min\{c_k + 1, Q\}$ for $k \notin S$.

### 3.2 SLNR based user scheduling

In the above algorithm, SINR is used to determine the interference constraints between the selected users. Besides SINR, SLNR can also be used [20,21]. From (1), the SLNR of user $u$, which measures the amount of power leaked from its beamforming direction to other users' channel direction, can be given by

$$\mathrm{SLNR}_u = \frac{\frac{P}{U_t}\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2}{\sigma_u^2 + \frac{P}{U_t}\sum_{i=1,i\neq u}^{U_t}\left|\boldsymbol{h}_i^{\mathrm{T}}\boldsymbol{w}_u\right|^2}. \tag{15}$$

Similar to SINR, the instantaneous SLNR also depends on the instantaneous CSI which is difficult to acquire at BS for massive MIMO systems working at FDD mode. Note that the numerator and denominator of (15) are independent. Based on Mullen's inequality [22], we can obtain that

$$\mathrm{E}\{\mathrm{SLNR}_u\} \geqslant \frac{\frac{P}{U_t}\boldsymbol{w}_u^{\dagger}\mathrm{E}\left\{\boldsymbol{h}_u^{*}\boldsymbol{h}_u^{\mathrm{T}}\right\}\boldsymbol{w}_u}{\sigma_u^2 + \frac{P}{U_t}\sum_{i=1,i\neq u}^{U_t}\boldsymbol{w}_u^{\dagger}\mathrm{E}\left\{\boldsymbol{h}_i^{*}\boldsymbol{h}_i^{\mathrm{T}}\right\}\boldsymbol{w}_u}. \tag{16}$$

From (2), we can get that

$$\mathrm{E}\left\{\boldsymbol{h}_u^{*}\boldsymbol{h}_u^{\mathrm{T}}\right\} = \frac{K_u}{K_u+1}\overline{\boldsymbol{h}}_u^{*}\overline{\boldsymbol{h}}_u^{\mathrm{T}} + \frac{1}{K_u+1}\mathbf{I}_{MN}. \tag{17}$$

Substitute (3), (6)–(8), and (17) into (16). After some manipulation, we get

$$\mathrm{E}\left\{\mathrm{SLNR}_u\right\} \geqslant \mathrm{SLNR}_u^L = \frac{\frac{P}{U_t}D_u}{\sigma_u^2 + \frac{P}{U_t}\sum_{i=1,i\neq u}^{U_t} L_{u,i}}, \tag{18}$$

where

$$L_{u,i} = \frac{K_i}{K_i+1}\left(\alpha_i^{(v)}\right)_{\overline{m}_u}\left(\alpha_i^{(h)}\right)_{\overline{n}_u} + \frac{1}{K_i+1}. \tag{19}$$

From (18), we see that the lower bound $\mathrm{SLNR}_u^L$ only depends on statistical CSI. Therefore, instead of the instantaneous SLNR, we will derive user scheduling algorithm based on $\mathrm{SLNR}_u^L$. Similar to SINR based user scheduling, this algorithm also maintains a priority vector and updates it at the end of every scheduling slot. The difference is, in this part, the lower bound of SLNR is used to determine the interference constraints instead of the approximation of SINR. Let us rewrite $\mathrm{SLNR}_u^L$ as

$$\mathrm{SLNR}_u^L = \frac{1}{\frac{\sigma_u^2 U_t}{P D_u} + \sum_{i=1,i\neq u}^{U_t}\frac{L_{u,i}}{D_u}}, \tag{20}$$

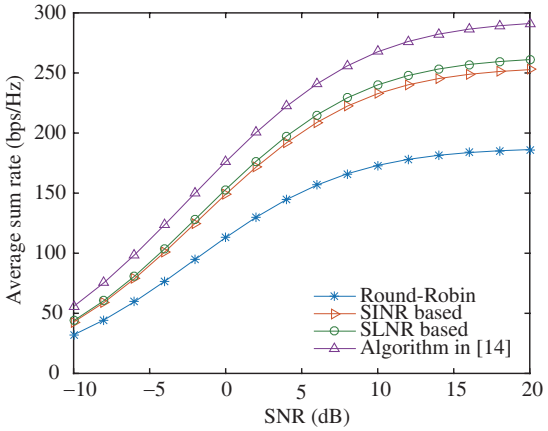


**Figure 1** (Color online) Average sum rate per slot under different user scheduling algorithms.

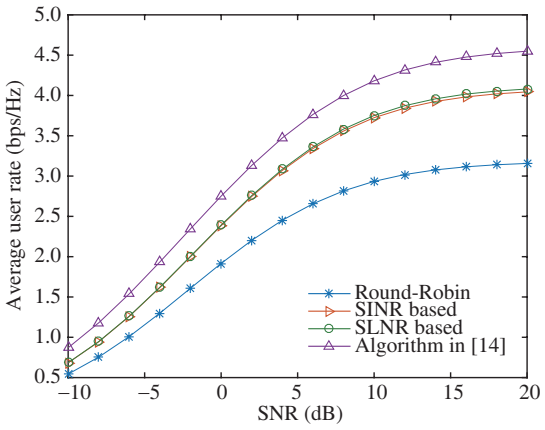


**Figure 2** (Color online) Average user rate under different user scheduling algorithms.

and define

$$G_{u,i} = L_{u,i}/D_u. \tag{21}$$

It can be seen that $\mathrm{SLNR}_u^L$ increases as $G_{u,i}$ decreases. We would like that $G_{u,i}$ and $G_{i,u}$ are below some certain threshold $\gamma$, to ensure that the SLNR lower bound of candidate user $u$ would be high and the SLNR lower bound of selected user $i$ will not decrease much due to candidate user $u$. Therefore, changing the interference constraints in the SINR based user scheduling algorithm to $G_{u,i} \leqslant \gamma$ and $G_{i,u} \leqslant \gamma$, we have the following SLNR based user scheduling algorithm.

**Step 1.** Initialize $S = \emptyset$, the number of already selected users $l = 0$, $c = \max_{k=1,\ldots,L} c_k$.

**Step 2.** Find $\Omega_c = \{k | c_k = c, k = 1, \ldots, L\}$.

**Step 3.** If $S \neq \emptyset$, compute $G_{u,i}$ and $G_{i,u}$ for each user $u \in \Omega_c$ and user $i \in S$, and remove user $u$ that does not satisfy $G_{u,i} \leqslant \gamma$ and $G_{i,u} \leqslant \gamma$ from $\Omega_c$.

**Step 4.** Find $s = \arg\max_{u \in \Omega_c} D_u$, and update $S = S \cup s$, $\Omega_c = \Omega_c \backslash s$, $l = l + 1$. Repeat Steps 3 and 4 until $\Omega_c = \emptyset$ or $l = U$. Go to Step 5 when $\Omega_c = \emptyset$; and go to Step 6 when $l = U$.

**Step 5.** Let $c = c - 1$. Go to Step 2 when $c \geqslant 0$, and go to Step 6 when $c < 0$.

**Step 6.** Set $c_k = 0$ for $k \in S$, and $c_k = \min\{c_k + 1, Q\}$ for $k \notin S$.

## 4 Simulation results

In this section, numerical results are presented to validate the performance of the proposed user scheduling algorithms. In all simulations, the total number of scheduling slot is 1000, $L = 300$, $M = 16$, $N = 64$, $\sigma_u^2 = \sigma^2, u = 1, \ldots, L$. The BS can serve 64 users at most. We assume that $K_u, u = 1, \ldots, L$ is uniformly distributed in $[K_{\min}, K_{\max}]$, where $K_{\min} = -10$ dB, $K_{\max} = 10$ dB, the vertical and horizontal AoD of each user are all uniformly distributed in $(-90°, 90°)$, and $Q = 3$. The performance of the scheduling algorithm in [16] and the "Round-Robin" algorithm are also shown. In the "Round-Robin" scheduling algorithm, users are divided into $4 \times 16$ clusters, so that user $u$ in cluster $(i, j)$ satisfying $\bar{m}_u \in [4(i-1)+1, 4i]$ and $\bar{n}_u \in [4(j-1)+1, 4j]$. Therefore, users in the same cluster have similar horizontal and vertical LOS component, and users in different clusters are orthogonal in horizontal or vertical LOS component. Then, select one user in each cluster in a Round-Robin fashion. The threshold for SINR based user scheduling algorithm is $\zeta = 0.01$ and the threshold for SLNR based user scheduling algorithm is $\gamma = 0.01$. Note that although we set the thresholds of the SINR and SLNR based algorithms to be the same, they could be different according to the needs of the system.

Figure 1 shows the average sum rate of the FD-MIMO downlink transmission systems under different user scheduling algorithms. Figure 2 shows the average user rate performance of different scheduling algorithms. From these figures, we see that the scheduling algorithm in [16] performs best in terms of

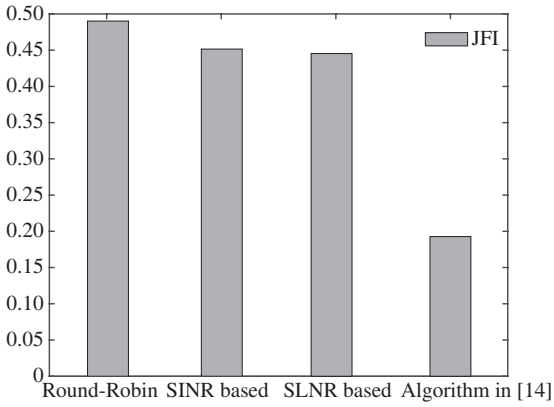


**Figure 3** JFI under different user scheduling algorithms.

both the average sum rate and average user rate, since this scheduling algorithm was designed to maximize the throughput. The "Round-Robin" algorithm performs worst in both average sum rate and average user rate performance, since it focuses on the fairness between different users and lessens the constraint of the interference between them. The proposed SINR based and SLNR based scheduling algorithm improve the average sum rate and average user rate performance of the system a lot, compared to the "Round-Robin" algorithm. Under the same interference threshold, the performance of the SINR based and SLNR based scheduling algorithms are almost the same, while the SLNR based algorithm performs slight better.

To quantify the fairness of the scheduling algorithms, Jain's fairness index (JFI) [23] is used. It takes values ranging from $1/L$ to 1. The larger the JFI, the fairer the scheduling algorithm. JFI $= 1$ represents the fairest case when the throughput of each user are the same. In this paper, JFI is defined as

$$\text{JFI} = \frac{\left(\sum_{i=1}^{L} r_i\right)^2}{L\left(\sum_{i=1}^{L} r_i^2\right)}, \tag{22}$$

where $r_i$ is the throughput for user $i$.

Figure 3 shows the JFI of the transmission system under different user scheduling algorithms when SNR $= 20$ dB. It can be seen that the "Round-Robin" scheduling algorithm is the fairest among the four scheduling algorithms, and the scheduling algorithm in [16] is the most unfair one. The JFIs of the two proposed scheduling algorithms and the JFI of the "Round-Robin" scheduling algorithm are very close. Compared with Figures 1 and 2, we can see that the scheduling algorithm in [16] is the best in achieving high rate but the worst in achieving good fairness. On the contrary, the "Round-Robin" scheduling algorithm is the worst in achieving high rate but the best in providing good fairness. In contrast, the proposed scheduling algorithms can make good trade-off between the achievable rate and fairness. Note that the performance of the SINR based and SLNR based scheduling algorithms are almost the same. It is flexible to use either SINR based or SLNR based scheduling algorithm.

## 5 Conclusion

In this paper, we investigated the user scheduling algorithms for downlink FD-MIMO transmission systems with only statistical CSI at BS. The channel was modeled as Rician fading. We derived an approximation of user's SINR and a lower-bound of each user's average SLNR. Based on these and under the help of a priority vector, we proposed two user scheduling algorithms that take into account both the achievable rate and fairness. Simulation results showed that the proposed algorithms work well and can achieve good trade-off between throughput and fairness.

**Acknowledgements** This work was supported by National Natural Science Foundation of China (Grant Nos. 61571112, 61531011, 61625106, 61320106003, 61521061), Foundation for the Author of National Excellent Doctoral

Dissertation of PR China (FANEDD) (Grant No. 201446), and Fundamental Research Funds for the Central Universities (Grant No. 2242015R30006).

## 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 You X H, Pan Z W, Gao X Q, et al. The 5G mobile communication: the development trends and its emerging key techniques (in Chinese). Sci Sin Inform, 2014, 44: 551–563
3 Wang D M, Zhang Y, Wei H, et al. An overview of transmission theory and techniques of large-scale antenna systems for 5G wireless communications. Sci China Inf Sci, 2016, 59: 081301
4 Wang C-X, Wu S B, Bai L, et al. Recent advances and future challenges for massive MIMO channel measurements and models. Sci China Inf Sci, 2016, 59: 021301
5 Marzetta T L. Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans Wireless Commun, 2010, 9: 3590–3600
6 Ngo H Q, Larsson E G, Marzetta T L. Energy and spectral efficiency of very large multiuser MIMO systems. IEEE Trans Commun, 2013, 61: 1436–1449
7 Nam Y H, Ng B L, Sayana K, et al. Full-dimension MIMO (FD-MIMO) for next generation cellular technology. IEEE Commun Mag, 2013, 51: 172–186
8 Kim Y, Ji H, Lee J, et al. Full dimension MIMO (FD-MIMO): the next evolution of MIMO in LTE systems. IEEE Wireless Commun, 2014, 21: 26–33
9 Adhikary A, Nam J, Ahn J-Y, et al. Joint spatial division and multiplexing-the large-scale array regime. IEEE Trans Inform Theor, 2013, 59: 6441–6463
10 Xu W. Capacity improvement analysis of 3D-beamforming in small cell systems. Sci China Inf Sci, 2018, 61: 022305
11 Mao J L, Gao J C, Liu Y A, et al. Robust multiuser MIMO scheduling algorithms with imperfect CSI. Sci China Inf Sci, 2012, 55: 815–826
12 Marzetta T L, Caire G, Debbah M, et al. Special issue on massive MIMO. J Commun Netw, 2013, 59: 333–337
13 Chan P W C, Lo E S, Wang R R, et al. The evolution path of 4G networks: FDD or TDD? IEEE Commun Mag, 2006, 44: 42–50
14 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, 2013, 8: 876–890
15 Li X, Jin S, Gao X, et al. Three-dimensional beamforming for large-scale FD-MIMO systems exploiting statistical channel state information. IEEE Trans Veh Technol, 2016, 65: 8992–9005
16 Li X, Jin S, Suraweera H A, et al. Statistical 3-D beamforming for large-scale MIMO downlink systems over Rician fading channels. IEEE Trans Commun, 2016, 64: 1529–1543
17 Han Y, Zhang H, Jin S, et al. Investigation of transmission schemes for millimeter-wave massive MU-MIMO systems. IEEE Syst J, 2017, 11: 72–83
18 Maddah-Ali M A, Tse D. Completely stale transmitter channel state information is still very useful. IEEE Trans Inform Theor, 2012, 58: 4418–4431
19 Adhikary A, Dhillon H S, Caire G. Massive-MIMO meets HetNet: interference coordination through spatial blanking. IEEE J Sel Areas Commun, 2015, 33: 1171–1186
20 Sadek M, Tarighat A, Sayed A. A leakage-based precoding scheme for downlink multi-user MIMO channels. IEEE Trans Wireless Commun, 2007, 6: 1711–1721
21 Xia X, Wu G, Liu J, et al. Leakage-based user scheduling in MU-MIMO broadcast channel. Sci China Ser F-Inf Sci, 2009, 52: 2259–2268
22 Mullen K. A note on the ratio of two independent random variables. Amer Stat, 1967, 21: 30–31
23 Sediq A B, Gohary R H, Schoenen R, et al. Optimal tradeoff between sum-rate efficiency and Jain's fairness index in resource allocation. IEEE Trans Wireless Commun, 2013, 12: 3496–3509


## Appendix A Proof of Theorem 1

According to (9), we can obtain that

$$\frac{1}{MN}\left(\mathrm{SINR}_u - \overline{\mathrm{SINR}}_u\right) = \frac{a_1 + a_2}{\left(\sigma_u^2 + \frac{P}{U_t}\sum_{i=1,i\neq u}^{U_t}|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_i|^2\right)\left(\sigma_u^2 + \frac{P}{U_t}\sum_{i=1,i\neq u}^{U_t} I_{u,i}\right)}, \tag{A1}$$

where

$$a_1 = \frac{1}{MN}\frac{P\sigma_u^2}{U_t}\left(\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2 - D_u\right), \tag{A2}$$

$$a_2 = \frac{1}{MN}\frac{P^2}{U_t^2}\left(\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2\sum_{i=1,i\neq u}^{U_t} I_{u,i} - D_u\sum_{i=1,i\neq u}^{U_t}\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_i\right|^2\right). \tag{A3}$$

First, consider the term $a_1$. From (2) and (8), $|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u|^2 - D_u$ can be rewritten as

$$\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2 - D_u = \frac{K_u}{K_u+1}\left|\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2 + \frac{\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2}{K_u+1} + \frac{\sqrt{K_u}}{K_u+1}\left(\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\boldsymbol{h}_{w,u}^{*} + \left(\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\boldsymbol{h}_{w,u}^{*}\right)^{\dagger}\right) - D_u. \tag{A4}$$

(1) $\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\boldsymbol{h}_{w,u}^{*}$: For the sake of simplicity, we define $\widetilde{\boldsymbol{h}}^{\dagger}=\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}=[\widetilde{h}_1^{*},\ldots,\widetilde{h}_{MN}^{*}]$ and $\boldsymbol{h}_{w,u}=[h_{w,u}^{(1)},\ldots,h_{w,u}^{(MN)}]^{\mathrm{T}}$. Then, we can obtain that $\widetilde{\boldsymbol{h}}^{\dagger}\boldsymbol{h}_{w,u}^{*}=\sum_{i=1}^{MN}\widetilde{h}_i^{*}(h_{w,u}^{(i)})^{*}$. Since $h_{w,u}^{(i)},i=1,\ldots,MN$ are i.i.d. complex Gaussian variables with zero mean and unit variance, it can be obtained that $\widetilde{h}_i^{*}(h_{w,u}^{(i)})^{*},\ i=1,\ldots,MN$ are uncorrelated and $\mathrm{Var}\{\widetilde{h}_i^{*}(h_{w,u}^{(i)})^{*}\}=|\widetilde{h}_i^{*}|^2$. Then, $\sum_{i=1}^{MN}\mathrm{Var}\{\widetilde{h}_i^{*}(h_{w,u}^{(i)})^{*}\}=\sum_{i=1}^{MN}|\widetilde{h}_i^{*}|^2=\widetilde{\boldsymbol{h}}^{\dagger}\widetilde{\boldsymbol{h}}=\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\overline{\boldsymbol{h}}_u^{*}$. Note that $\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\overline{\boldsymbol{h}}_u^{*}\leqslant MN$. Therefore, we can get $\frac{1}{(MN)^2}\sum_{i=1}^{MN}\mathrm{Var}\{\widetilde{h}_i^{*}(h_{w,u}^{(i)})^{*}\}\leqslant\frac{1}{MN}$, and then $\sum_{i=1}^{MN}\mathrm{Var}\{\widetilde{h}_i^{*}(h_{w,u}^{(i)})^{*}\}=\mathrm{o}((MN)^2)$. According to Chebyshev Theorem[1)], we can get that when $MN\to\infty$,

$$\frac{1}{MN}\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\boldsymbol{h}_{w,u}^{*}-\frac{1}{MN}\mathrm{E}\left\{\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\boldsymbol{h}_{w,u}^{*}\right\}\xrightarrow{\mathrm{p}}0. \tag{A5}$$

Note that $\mathrm{E}\{\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\boldsymbol{h}_{w,u}^{*}\}=0$. We have that when $MN\to\infty$,

$$\frac{1}{MN}\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}\boldsymbol{h}_{w,u}^{*}\xrightarrow{\mathrm{p}}0. \tag{A6}$$

(2) $\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2$: Denote $\boldsymbol{W}_u=\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger}$, then

$$\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2=\sum\nolimits_{i=1}^{MN}\left|h_{w,u}^{(i)}\right|^2W_{i,i}+\sum\nolimits_{i=1}^{MN-1}\sum\nolimits_{j=i+1}^{MN}2\mathrm{Re}\left(h_{w,u}^{(i)}\left(h_{w,u}^{(j)}\right)^{*}W_{i,j}\right), \tag{A7}$$

where $W_{i,j}$ is the $(i,j)$-th element of $\boldsymbol{W}_u$ and $\mathrm{Re}(\cdot)$ represents the real part of the complex number. Assume that $h_{w,u}^{(i)}=h_{w,u,R}^{(i)}+\tilde{\mathrm{j}}h_{w,u,I}^{(i)}$, $W_{i,j}=W_{i,j,R}+\tilde{\mathrm{j}}W_{i,j,I}$, where $h_{w,u,R}^{(i)}$, $h_{w,u,I}^{(i)}$, $W_{i,j,R}$, and $W_{i,j,I}$ are real numbers, and $\tilde{\mathrm{j}}^2=-1$. Therefore, it can be obtained that

$$\left|h_{w,u}^{(i)}\right|^2W_{i,i}=\left(\left(h_{w,u,R}^{(i)}\right)^2+\left(h_{w,u,I}^{(i)}\right)^2\right)W_{i,i}, \tag{A8}$$

and

$$\mathrm{Re}\left(h_{w,u}^{(i)}\left(h_{w,u}^{(j)}\right)^{*}W_{i,j}\right)=\left(h_{w,u,R}^{(i)}h_{w,u,R}^{(j)}+h_{w,u,I}^{(i)}h_{w,u,I}^{(j)}\right)W_{i,j,R}-\left(h_{w,u,I}^{(i)}h_{w,u,R}^{(j)}-h_{w,u,I}^{(j)}h_{w,u,R}^{(i)}\right)W_{i,j,I}. \tag{A9}$$

Substituting (A8) and (A9) into (A7) yields

$$\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2=\sum\nolimits_{i=1}^{MN}b_i, \tag{A10}$$

where

$$b_i=\begin{cases}\left(\left(h_{w,u,R}^{(i)}\right)^2+\left(h_{w,u,I}^{(i)}\right)^2\right)W_{i,i}+\sum_{j=i+1}^{MN}\left(c_{i,j}^{(1)}+c_{i,j}^{(2)}\right),\ i=1,\ldots,MN-1,\\ \left(\left(h_{w,u,R}^{(MN)}\right)^2+\left(h_{w,u,I}^{(MN)}\right)^2\right)W_{MN,MN},\ i=MN,\end{cases} \tag{A11}$$

$$c_{i,j}^{(1)}=2\left(h_{w,u,R}^{(i)}h_{w,u,R}^{(j)}+h_{w,u,I}^{(i)}h_{w,u,I}^{(j)}\right)W_{i,j,R}, \tag{A12}$$

$$c_{i,j}^{(2)}=2\left(h_{w,u,I}^{(j)}h_{w,u,R}^{(i)}-h_{w,u,I}^{(i)}h_{w,u,R}^{(j)}\right)W_{i,j,I}. \tag{A13}$$

Note that $h_{w,u,x}^{k}$, $k=1,\ldots,MN$, $x\in\{R,I\}$ are i.i.d. complex Gaussian variables with zero mean and 0.5 variance. It can be obtained that $\mathrm{Cov}\{b_{i_1},b_{i_2}\}=0,\ i_1\neq i_2$, and then $b_i$, $i=1,\ldots,MN$ are uncorrelated. Moreover, it is easy to get that

$$\mathrm{Var}\{c_{i,j}^{(1)}\}=2W_{i,j,R}^2, \tag{A14}$$

$$\mathrm{Var}\{c_{i,j}^{(2)}\}=2W_{i,j,I}^2, \tag{A15}$$

$$\mathrm{Var}\left\{\left(\left(h_{w,u,R}^{(i)}\right)^2+\left(h_{w,u,I}^{(i)}\right)^2\right)W_{i,i}\right\}=W_{i,i}^2. \tag{A16}$$

Then, it can be obtained that

$$\mathrm{Var}\left\{b_i\right\}=\begin{cases}W_{i,i}^2+\sum_{j=i+1}^{MN}2\left(W_{i,j,R}^2+W_{i,j,I}^2\right),\ i=1,\ldots,MN-1,\\ W_{MN,MN}^2,\ i=MN.\end{cases} \tag{A17}$$

Therefore, we can get

$$\sum\nolimits_{i=1}^{MN}\mathrm{Var}\left\{b_i\right\}=\sum\nolimits_{i=1}^{MN}\sum\nolimits_{j=1}^{MN}|W_{i,j}|^2=\mathrm{tr}\left(\boldsymbol{W}_u\boldsymbol{W}_u^{\dagger}\right)=1. \tag{A18}$$

It can be seen that $\sum_{i=1}^{MN}\mathrm{Var}\{b_i\}=\mathrm{o}((MN)^2)$. According to Chebyshev Theorem, we can get that when $MN\to\infty$,

$$\frac{1}{MN}\sum\nolimits_{i=1}^{MN}b_i-\frac{1}{MN}\mathrm{E}\Big\{\sum\nolimits_{i=1}^{MN}b_i\Big\}\xrightarrow{\mathrm{p}}0. \tag{A19}$$

Since $\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2=\sum_{i=1}^{MN}b_i$, then we have

$$\frac{1}{MN}\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2-\frac{1}{MN}\mathrm{E}\Big\{\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2\Big\}\xrightarrow{\mathrm{p}}0. \tag{A20}$$

Note that $\mathrm{E}\{\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2\}=\mathrm{tr}(\boldsymbol{w}_u\boldsymbol{w}_u^{\dagger})=1$. Therefore, we can get that, when $MN\to\infty$,

$$\frac{1}{MN}\left|\boldsymbol{h}_{w,u}^{\mathrm{T}}\boldsymbol{w}_u\right|^2-\frac{1}{MN}\xrightarrow{\mathrm{p}}0. \tag{A21}$$

1) Robert J S. Approximation Theorems of Mathematical Statistics. New York: John Wiley and Sons Inc, 2008. 27.

Moreover, from (3) and (6)–(8), it can be obtained that $\left|\overline{\boldsymbol{h}}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2 = \left(\alpha_u^{(v)}\right)_{\overline{m}_u}\left(\alpha_u^{(h)}\right)_{\overline{n}_u}$. Combining this with (A4), (A6) and (A21), we can get that

$$\frac{1}{MN}\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2 - \frac{1}{MN}D_u \xrightarrow{\mathrm{p}} 0 \tag{A22}$$

and therefore $a_1 \xrightarrow{\mathrm{p}} 0$, when $MN \to \infty$.

Next, consider the term $a_2$. It can be rewritten as

$$a_2 = \frac{1}{MN}\frac{P^2}{U_t^2}\sum\nolimits_{i=1,i\neq u}^{U_t}\left(\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_u\right|^2 - D_u\right)I_{u,i} - \frac{1}{MN}\frac{P^2}{U_t^2}\sum\nolimits_{i=1,i\neq u}^{U_t}\left(\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_i\right|^2 - I_{u,i}\right)D_u. \tag{A23}$$

Using the similar method, and after some manipulations, we can get that

$$\frac{1}{MN}\left|\boldsymbol{h}_u^{\mathrm{T}}\boldsymbol{w}_i\right|^2 - \frac{1}{MN}I_{u,i} \xrightarrow{\mathrm{p}} 0, \tag{A24}$$

when $MN \to \infty$. Combining (A22)–(A24), we have that $a_2 \xrightarrow{\mathrm{p}} 0$, when $MN \to \infty$. Therefore, combining (A1) with $a_1 \xrightarrow{\mathrm{p}} 0$ and $a_2 \xrightarrow{\mathrm{p}} 0$ when $MN \to \infty$, we get Theorem 1.