Multi-pair Massive MIMO Relay Networks: Power Scaling Laws and User Scheduling Strategy

Xuesong Liang[1,*], Shi Jin[2], Kai-Kit Wong[3], Tao Hong[1], Hongbo Zhu[1]

[1]School of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China
[2]National Mobile Communications Research Laboratory, Southeast University, Nanjing, China
[3]Department of Electronic and Electrical Engineering, University College London, London, United Kingdom
*liangxs@njupt.edu.cn

Abstract: This paper studies a multi-pair massive multiple-input multiple-output (MIMO) relaying network, where multiple pairs of users are served by a single relay station with a large number of antennas, and the amplify-and-forward (AF) protocol and zero-forcing (ZF) beamforming are used at the relay. We investigate the ergodic achievable rates for the users and obtain tight approximations in closed form for finite number of antennas. The rate performance and power efficiency are studied based on the analytical results for asymptotic scenarios, and the effect of scaling factors of transmit powers for users and relay are discussed. The closed-form expressions enable us to determine the optimal user scheduling which maximizes the ergodic sum-rate for the selected pairs. A simplified user scheduling algorithm is proposed which greatly reduces the average complexity of the optimal use pair search without any rate loss. Moreover, the complexity reduction for the proposed algorithm increases nonlinearly with the increase of the number of user pairs, which indicates that the simplified scheduling algorithm has notable advantages when the number of users is increased. The tightness for the analytical approximations and the superiority of the proposed algorithm are verified by Monte-Carlo simulation results.

1. Introduction

In recent years, massive multiple-input multiple-output (MIMO) antennas have emerged as a key wireless communications technology that can deliver extraordinary gains in throughput, energy efficiency and more, by scaling up conventional MIMO by orders of magnitude, e.g., [1–3], and have already been identified as one major component for the 5th generation (5G) wireless systems [4, 5]. Recent developments also have shown strong interest to combine massive MIMO with cooperative relaying to realise energy-efficient mobile networks [6, 7].

By employing a large number of antennas at the relay, the performance for multi-user massive MIMO relaying systems has been investigated in recent literatures. The asymptotic performance was studied for multi-pair one-way relaying in [7] and two-way relaying networks in [8], by adopting linear signal processing at the relay with a large antenna array. Both [7] and [8] pointed out that the transmit powers of users and/or relay can decrease inverse-proportionally to the number of antennas which grows to infinity, while maintaining a given transmission rate. The performance analysis was also extended to the distributed amplify-and-forward (AF) relay networks with an

unlimited number of single-antenna relays, and the spectral and energy efficiencies for distributed large-scale relay networks were analyzed in [9]. Recent studies for utilizing massive MIMO in multi-user relaying systems also addressed pilot allocation [10, 11], power control [12, 13], multi-way relaying [14] and relay-base station (BS) architecture [15]. In addition, full-duplex radios were also actively looked at in tandem with massive MIMO [16, 17], and massive MIMO relaying systems [13, 18, 19], due to their combined advantages.

In the studies for multi-user massive MIMO relaying systems so far, the rate analysis was mostly considered for asymptotic scenarios where the number of antennas at the relay is unlimited. In practice, since the number of antennas has to be limited (not exactly asymptotic), the achievable rate analysis thus is often only approximate [20], which motivates this work. In this paper, our aim is to study the ergodic rates of a multi-pair one-way massive MIMO AF relay system where zero-forcing (ZF) is used to process the signals. Using random matrices analysis, closed-form approximations for the ergodic achievable rates are obtained which are remarkably accurate compared to the actual rates for finite number of relay antennas. Based on the results, a simplified user scheduling algorithm is proposed to maximize the system sum-rate. With the proposed algorithm, the optimal user pairs are chosen while greatly reducing the complexity for user pair searching. Simulation results are provided to validate the analytical results, gain useful insights, and illustrate the superiority of the proposed user scheduling algorithm.

Notations—Throughout this paper, we use capital boldface letters to denote matrices and small boldface letters to denote column vectors while $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $\|\cdot\|^2$ and $\mathrm{tr}\{\cdot\}$ represent the operations of transposition, conjugation, conjugate transposition, Euclidean norm and trace, respectively. Additionally, $\mathrm{E}\{\cdot\}$ stands for the expectation of an input random variable.

2. System Model

We consider the relaying system where M pairs of source nodes, i.e., T_{Ai}, want to communicate with their destination nodes, i.e., T_{Bi}, through an N-antenna one-way relay, i.e., T_R for $i = 1, \ldots, M$. We assume that only K user pairs are allowed to communicate with each other at any one time and the user scheduling algorithm to select K pairs out of M pairs will be discussed in Section 5. In this system, all the source and destination nodes are equipped with single antenna and we assume $N \gg K$. It is also assumed that the direct links between T_{Ai} and T_{Bi} are broken. The channel between T_{Ai} and T_R is denoted as $\boldsymbol{g}_i$ and that between T_{Bi} and T_R is denoted as $\boldsymbol{h}_i$, and they are statistically independent Rayleigh random vectors with independent and identically distributed (i.i.d.) entries, i.e., $\boldsymbol{g}_i \sim \mathcal{CN}\left(\boldsymbol{0}, \sigma_{g_i}^2 \boldsymbol{I}_N\right)$, $\boldsymbol{h}_i \sim \mathcal{CN}\left(\boldsymbol{0}, \sigma_{h_i}^2 \boldsymbol{I}_N\right)$, for $i = 1, \ldots, M$. Moreover, the transmit power for the relay and that at each source terminal is represented, respectively, by P_R and P_i, for $i = 1, \ldots, M$.

Communication for this multi-pair one-way relay system takes place in two phases. In the first phase, T_{Ai} transmits the information-bearing signals (assumed Gaussian variables), i.e., x_{Ai}, for $i = 1, \ldots, K$, to T_R. Thus, the received signals at T_R is expressed as

$$\boldsymbol{y}_r = \boldsymbol{G}\boldsymbol{x}_A + \boldsymbol{n}_R, \tag{1}$$

where $\boldsymbol{G} \triangleq [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K]$, $\boldsymbol{x}_A \triangleq [x_{A1}, \ldots, x_{AK}]^T$ and $\boldsymbol{n}_R \sim \mathcal{CN}\left(\boldsymbol{0}, \sigma_R^2 \boldsymbol{I}_N\right)$ is the noise at T_R.

In the next phase, after receiving $\boldsymbol{y}_r$, T_R multiplies it with the coefficient ρ, and processes the received signal with the matrix $\boldsymbol{F}$, before the following signals are broadcast to all the users:

$$\boldsymbol{y}_t = \rho \boldsymbol{F} \boldsymbol{y}_r. \tag{2}$$

2

It is assumed that T_R has the global channel state information (CSI) and is also in line with previous researches such as [8, 22], In such a situation, T_R uses this knowledgei.e., $\{\boldsymbol{g}_i, \boldsymbol{h}_i\}, \forall i$, to construct the processing matrix $\boldsymbol{F}$ while the fixed ρ is used to constrain the transmit power of T_R and is given by

$$\rho = \sqrt{\frac{P_R}{P_U \|\boldsymbol{F}\boldsymbol{G}\|^2 + \sigma_R^2 \|\boldsymbol{F}\|^2}}, \tag{3}$$

where we assume that the transmit power for all users are identical and represented by P_U.

In this paper, ZF for large-scale MIMO systems is used [4] at T_R. Therefore, $\boldsymbol{F}$ is constructed by the ZF receiver and the ZF precoding matrices, and is expressed as

$$\boldsymbol{F} = \boldsymbol{F}_B \boldsymbol{F}_A, \tag{4}$$

in which the ZF receiver matrix $\boldsymbol{F}_A$ is given by

$$\boldsymbol{F}_A = \left(\boldsymbol{G}^H \boldsymbol{G}\right)^{-1} \boldsymbol{G}^H \tag{5}$$

and the ZF precoding matrix $\boldsymbol{F}_B$ is found as

$$\boldsymbol{F}_B = \boldsymbol{H}^* \left(\boldsymbol{H}^T \boldsymbol{H}^*\right)^{-1} \tag{6}$$

where $\boldsymbol{H} \triangleq [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]$. Finally, T_{Bi} receives

$$z_{Bi} = \boldsymbol{h}_i^T \boldsymbol{y}_t + n_{Bi} \tag{7}$$

with $n_{Bi} \sim \mathcal{CN}\left(0, \sigma_{Bi}^2\right)$, for $i = 1, \ldots, K$.

To analyze the rates, we write $\boldsymbol{F}_A = [\boldsymbol{F}_{A,1}, \ldots, \boldsymbol{F}_{A,2K}]$ and consider the expansion of (7) by

$$\begin{aligned} z_{Bi} &= \rho x_{Ai} + \rho \boldsymbol{F}_{A,i} \boldsymbol{n}_R + n_{Bi} \\ &= \underbrace{\rho x_{Ai}}_{\text{Signal}} + \underbrace{\rho \left[\left(\boldsymbol{G}^H \boldsymbol{G}\right)^{-1} \boldsymbol{G}^H\right]_i \boldsymbol{n}_R + n_{Bi}}_{\text{Noise}}, \text{ for } i = 1, \ldots, K. \end{aligned} \tag{8}$$

Then the representation of the signal-to-noise ratio (SNR) for T_{Bi} is shown as

$$\mathrm{SNR}_{Bi} = \frac{P_U}{\sigma_R^2 \left[\left(\boldsymbol{G}^H \boldsymbol{G}\right)^{-1}\right]_{i,i} + \sigma_{Bi}^2/\rho^2}, \text{ for } i = 1, \ldots, K, \tag{9}$$

and the ergodic achievable rate for T_{Bi} is therefore given by

$$R_{Bi} = \frac{1}{2} \mathrm{E} \left\{\log_2 \left(1 + \mathrm{SNR}_{Bi}\right)\right\}, \text{ for } i = 1, \ldots, K. \tag{10}$$

3. Ergodic Rate and Asymptotic Analysis

However, the derivation of (10) is extremely challenging and an exact expression appears to be infeasible. Therefore, we adopt the following approximation for R_{Bi}:

$$R_{Bi} \approx \tilde{R}_{Bi} \equiv \frac{1}{2} \log_2 \left(1 + \left[\mathrm{E}\left\{\frac{1}{\mathrm{SNR}_{Bi}}\right\}\right]^{-1}\right). \tag{11}$$

Then we obtain an approximation for R_{Bi} in the following theorem.

3

Theorem 1. An approximation for the ergodic achievable rate for T_{Bi} is given by

$$\tilde{R}_{Bi} = \frac{1}{2}\log_2\left(1 + \left[\frac{\sigma_R^2}{(N-K)\,P_U\sigma_{g_i}^2} + \frac{\sigma_{Bi}^2}{(N-K)\,P_R}\sum_{k=1}^K\frac{1}{\sigma_{h_k}^2} + \frac{\sigma_R^2\sigma_{Bi}^2}{(N-K)^2\,P_UP_R}\sum_{k=1}^K\frac{1}{\sigma_{g_k}^2\sigma_{h_k}^2}\right]^{-1}\right),$$

$$\text{for } i = 1,\dots,K. \quad (12)$$

Proof. With the definition of (11), we then have

$$\tilde{R}_{Bi} = \frac{1}{2}\log_2\left(1 + \left[\frac{\sigma_R^2}{P_U}\cdot\mathrm{E}\left\{\left[\left(\boldsymbol{G}^H\boldsymbol{G}\right)^{-1}\right]_{i,i}\right\} + \frac{\sigma_{Bi}^2}{P_U}\cdot\mathrm{E}\left\{\frac{1}{\rho^2}\right\}\right]^{-1}\right), \quad (13)$$

where the components of $\mathrm{E}\left\{\left[\left(\boldsymbol{G}^H\boldsymbol{G}\right)^{-1}\right]_{i,i}\right\}$ and $\mathrm{E}\left\{\frac{1}{\rho^2}\right\}$ are calculated as follows.

First, as $\boldsymbol{G}^H\boldsymbol{G} \sim CW_K\left(\boldsymbol{\Sigma_G}, N\right)$, which is a complex Wishart matrix with $\boldsymbol{\Sigma_G} = \text{diag}\left\{\sigma_{g_1}^2,\dots,\sigma_{g_K}^2\right\}$, then with the properties for Wishart matrix [21], we have

$$\mathrm{E}\left\{\left(\boldsymbol{G}^H\boldsymbol{G}\right)^{-1}\right\} = \frac{1}{N-K}\boldsymbol{\Sigma_G}^{-1} = \frac{1}{N-K}\text{diag}\left\{\frac{1}{\sigma_{g_1}^2},\dots,\frac{1}{\sigma_{g_K}^2}\right\} \quad (14)$$

and

$$\mathrm{E}\left\{\left[\left(\boldsymbol{G}^H\boldsymbol{G}\right)^{-1}\right]_{i,i}\right\} = \frac{1}{N-K}\cdot\frac{1}{\sigma_{g_i}^2}. \quad (15)$$

Secondly, as

$$\mathrm{E}\left\{\frac{1}{\rho^2}\right\} = \frac{P_U}{P_R}\mathrm{E}\left\{\|\boldsymbol{FG}\|^2\right\} + \frac{\sigma_R^2}{P_R}\mathrm{E}\left\{\|\boldsymbol{F}\|^2\right\}, \quad (16)$$

in which

$$\mathrm{E}\left\{\|\boldsymbol{FG}\|^2\right\} = tr\left\{\mathrm{E}\left[\left(\boldsymbol{H}^T\boldsymbol{H}^*\right)^{-1}\right]\right\}, \quad (17)$$

where $\boldsymbol{H}^T\boldsymbol{H}^* \sim CW_K\left(\boldsymbol{\Sigma_H}, N\right)$ with $\boldsymbol{\Sigma_H} = \text{diag}\left\{\sigma_{h_1}^2,\dots,\sigma_{h_K}^2\right\}$.

With the properties for Wishart matrix [21], we have

$$\mathrm{E}\left\{\|\boldsymbol{FG}\|^2\right\} = \frac{1}{N-K}tr\left\{\boldsymbol{\Sigma_H}^{-1}\right\}$$

$$= \frac{1}{N-K}\sum_{i=1}^K\frac{1}{\sigma_{h_i}^2}, \quad (18)$$

while

$$\mathrm{E}\left\{\|\boldsymbol{F}\|^2\right\} = tr\left\{\mathrm{E}\left[\left(\boldsymbol{G}^H\boldsymbol{G}\right)^{-1}\right]\cdot\mathrm{E}\left[\left(\boldsymbol{H}^T\boldsymbol{H}^*\right)^{-1}\right]\right\}$$

$$= \frac{1}{(N-K)^2}tr\left\{\boldsymbol{\Sigma_G}^{-1}\cdot\boldsymbol{\Sigma_H}^{-1}\right\}$$

$$= \frac{1}{(N-K)^2}\sum_{i=1}^K\frac{1}{\sigma_{g_i}^2\sigma_{h_i}^2}. \quad (19)$$

4

Therefore, we have

$$\mathrm{E}\left\{\frac{1}{\rho^2}\right\} = \frac{P_U}{(N-K)\,P_R}\sum_{i=1}^{K}\frac{1}{\sigma_{h_i}^2} + \frac{\sigma_R^2}{(N-K)^2\,P_R}\sum_{i=1}^{K}\frac{1}{\sigma_{g_i}^2\sigma_{h_i}^2}. \tag{20}$$

To the end, we substitute (15) and (20) into (13), and obtain Theorem 1. $\qquad\square$

Theorem 1 gives closed-form expressions for the approximate ergodic rates, and it is obvious that the ergodic rate for the i-th user pair approximately increases with the logarithm of $(N - K)$, which is the rate gain for large antennas at the relay using ZF processing criterion. However, the correlation between user rate and the powers for users and relay, i.e., P_U and P_R is not clear. To understand this, we study the asymptotic scenario when N is sufficiently large, and for simplicity we assume that P_U, P_R and K are all fixed. To begin with, the asymptotic user rate is given below.

Corollary 1. When $N \to \infty$, an upper bound for $\tilde{R}_{Bi}$ is given by

$$\bar{R}_{Bi} = \frac{1}{2}\log_2\left(1 + N\left[\frac{\sigma_R^2}{P_i\sigma_{g_i}^2} + \frac{\sigma_{Bi}^2}{P_R}\sum_{k=1}^{K}\frac{P_k}{P_i\sigma_{h_k}^2}\right]^{-1}\right). \tag{21}$$

Proof. As $K \ll N$, we substitute N for $(N - K)$ in (12), while we omit the third term, which is in the order of N^{-2}, and keep the remaining terms which are in the order of N^{-1} in the square brackets in (12). Therefore we obtain (21). $\qquad\square$

From Corollary 1, it is also deduced that $\tilde{R}_{Bi}$ approaches $\bar{R}_{Bi}$ when N is sufficiently large. Then, two conclusions are drawn based on Corollary 1 as follows.

(a) When P_U and P_R are fixed, $\bar{R}_{Bi}$ increases with $(\log_2 N)/2$, which means that the user rate rises logarithmically by the increase of the number of antennas without enlarging the transmit powers of users and the relay. As a result, the rate gain for massive antennas also applies for the massive MIMO relay system.

(b) As $N \gg 1$, it is deduced that

$$\bar{R}_{Bi} \approx \frac{1}{2}\log_2 N + \min\left\{\frac{1}{2}\log_2\left(\frac{P_U\sigma_{g_i}^2}{\sigma_R^2}\right), \frac{1}{2}\log_2\left(\frac{1}{K}\cdot\frac{P_R\bar{\sigma}_{h_K}^2}{\sigma_{Bi}^2}\right)\right\}, \tag{22}$$

where $\bar{\sigma}_{h_K}^2 = K\left[\sum_{k=1}^{K}\frac{1}{\sigma_{h_k}^2}\right]^{-1}$, which is the harmonic mean of $\sigma_{h_k}^2$, for $i = 1, \ldots, K$.

The result (22) shows that $\bar{R}_{Bi}$ is mainly effected by by two terms, i.e., $P_U\sigma_{g_i}^2/\sigma_R^2$ and $P_R\bar{\sigma}_{h_K}^2/(K\sigma_{Bi}^2)$, despite the spectral efficiency of $(\log_2 N)/2$. The physical meaning of this is quite straightforward: $P_U\sigma_{g_i}^2/\sigma_R^2$ and $P_R\bar{\sigma}_{h_K}^2/(K\sigma_{Bi}^2)$ represent the upper bounds for signal-to-interference plus noise ratio (SINR) of the first hop and the second hop respectively, irrespective of the rate gain for antennas. To further discuss the effect of P_U and P_R on $\bar{R}_{Bi}$, we set $\sigma_{g_i}^2 = \sigma_{h_i}^2 = 1$ and $\sigma_R^2 = \sigma_{Bi}^2 = 1$, $\forall i$. Hence we can recast (22) into

$$\bar{R}_{Bi} \approx \frac{1}{2}\log_2 N + \frac{1}{2}\log_2\left(\min\left\{P_U, \frac{P_R}{K}\right\}\right). \tag{23}$$

5

The result (23) shows that if we enlarge P_R unboundedly, then $\bar{R}_{Bi}$ is limited by $\log_2 P_U$. On the contrary, if we enlarge P_U unboundedly, then $\bar{R}_{Bi}$ is limited by $\log_2 (P_R/K)$. In other words, the ergodic rate for user is limited logarithmically by both P_U and P_R/K, which means that the best ratio for setting P_U and P_R is $P_R/P_U = K$ in this typical scenario.

4. Power Scaling Laws

As above, Corollary 1 reveals the asymptotic performance for user rate and the effect of the given parameters, e.g., N, K, P_R, P_U on $\bar{R}_{Bi}$. In the sequel, we will discuss the power scaling laws of the system when P_R and P_U are scaled inversely with respect to N when N grows infinitely. To begin with, we define $P_U = E_U/N^\alpha$ and $P_R = E_R/N^\beta$ with nonnegative scaling factors $\alpha, \beta \geq 0$, while E_U and E_R are fixed regardless of N. Then we can approximate the ergodic rate in (12) by

$$\hat{R}_{Bi} = \frac{1}{2} \log_2 \left(1 + \frac{abcN^{2-\alpha-\beta}E_U E_R}{acN^{1-\alpha}E_U + bcN^{1-\beta}E_R + ab} \right) \tag{24}$$

as $N \gg K$, where $a = \frac{\sigma_{g_i}^2}{\sigma_R^2}$, $b = \left(\sigma_{Bi}^2 \sum_{k=1}^{K} \frac{1}{\sigma_{h_k}^2} \right)^{-1}$ and $c = \left(\sigma_R^2 \sigma_{Bi}^2 \sum_{k=1}^{K} \frac{1}{\sigma_{g_k}^2 \sigma_{h_k}^2} \right)^{-1}$.

Noting that the value of (24) will not decrease with the increase of N only when $\max\{1-\alpha, 1-\beta, 0\} \leq 2-\alpha-\beta$. That is to say, $\alpha, \beta \leq 1$ must be fulfilled to discuss the power scaling laws of the system. Therefore we consider three cases as follows:

- Case I: $\alpha = 1$, $0 \leq \beta \leq 1$,

- Case II: $0 \leq \alpha \leq 1$, $\beta = 1$,

- Case III: $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$ and $\alpha + \beta = 1$

when $N \to \infty$ and the analytical results are given as below.

Corollary 2. When $N \to \infty$, $\hat{R}_{Bi}$ of Case I, Case II and Case III are given respectively as

$$\hat{R}_{Bi}^{(I)} \to \begin{cases} \frac{1}{2} \log_2 (1 + aE_U), & 0 \leq \beta < 1 \\ \frac{1}{2} \log_2 \left(1 + \frac{abcE_U E_R}{bcE_R + acE_U + ab} \right), & \beta = 1 \end{cases} \tag{25}$$

$$\hat{R}_{Bi}^{(II)} \to \begin{cases} \frac{1}{2} \log_2 (1 + bE_R), & 0 \leq \alpha < 1 \\ \frac{1}{2} \log_2 \left(1 + \frac{abcE_U E_R}{bcE_R + acE_U + ab} \right), & \alpha = 1 \end{cases} , \tag{26}$$

$$\hat{R}_{Bi}^{(III)} \to \begin{cases} \frac{1}{2} \log_2 \left(1 + N^\beta \cdot aE_U \right), & \alpha > \beta \\ \frac{1}{2} \log_2 \left(1 + N^\alpha \cdot bE_R \right), & \alpha < \beta \\ \frac{1}{2} \log_2 \left(1 + N^{\frac{1}{2}} \cdot \frac{abE_U E_R}{aE_U + bE_R} \right), & \alpha = \beta = \frac{1}{2} \end{cases} , \tag{27}$$

where $a = \frac{\sigma_{g_i}^2}{\sigma_R^2}$, $b = \left(\sigma_{Bi}^2 \sum_{k=1}^{K} \frac{1}{\sigma_{h_k}^2} \right)^{-1}$ and $c = \left(\sigma_R^2 \sigma_{Bi}^2 \sum_{k=1}^{K} \frac{1}{\sigma_{g_k}^2 \sigma_{h_k}^2} \right)^{-1}$.

Proof. We substitute the conditions for Case I, Case II and Case III respectively into (24), and reform the expressions as $N \to \infty$. Skipping the tedious details of derivation, the results for (25)-(27) are obtained. $\square$

From the above results, the details of the three cases are illustrated as follows.

(a) For Case I, Corollary 2 indicates that the user rate converges to a fixed value related to E_U as $N \to \infty$, when P_U is scaled down by $1/N$ and P_U is scaled down by $1/N^\beta$ with $0 \le \beta < 1$. In other words, the upper limit of user rate is determined only by E_U, while the value of β only influences the rate of convergence to the limit for this case. However, when $\alpha = 1$, the upper limit is determined by both E_U and E_R, which is less than the former one.

(b) For Case II, it is shown that the user rate approaches to the upper limit determined by E_R when $P_R = E_R/N$ and $N \to \infty$, and the value of limit is effected by E_U when α is 1. Different from Case I, the upper limit of user rate determined by E_R will decrease when K increases. It means that the upper limit for every user rate is weakened when the number of users is enhanced when P_R is scaled down by $1/N$.

(c) For Case III, by fixing the sum of α and β to 1, Corollary 2 shows that both the antenna gain and the rate gain are obtained. We also note that if the part of antenna gain in Case III is omitted, the formulas will degenerate to that in Case I or Case II. Furthermore, when α and β are not equivalent, the antenna gain is determined by the minimum of them and therefore is maximized when they are equal. On the contrary, the rate gain is minimized when $\alpha = \beta$. Considering that the antenna gain is sufficiently larger than the rate gain when $N \to \infty$, the rate performance for users is better in the case of $\alpha = \beta$ than in the case of $\alpha \neq \beta$.

To sum up, Corollary 2 illustrates the power scaling laws we can obtain from P_U and P_R when N is sufficiently large. Both P_U and P_R can be scaled down by factor among $[1, 1/N]$ with non-decreasing user rate when $N \to \infty$. It is shown that at least one of P_U and P_R is scaled down by $1/N$, the upper limit for user rate is fixed and affected by the respective E_U or E_R. Furthermore, the upper limit decreases as the user number increases when P_R is scaled down by $1/N$. On other hand, when both P_U and P_R are scaled by factors less than $1/N$, the antenna gain is obtained while the value of that is determined by the smaller one.

5. User Scheduling Strategy

With the conclusion of Theorem 1, the rate performance and power scaling laws have been discussed for the asymptotic case. Now, we discuss the optimal user scheduling algorithm in order to maximize the sum-rate for the selected pairs. In general, we consider the centralized scheduling scheme performed by the relay, in which the best K user pairs are selected among total M user pairs based on the statistical CSI, and this problem is formulated as

$$\boldsymbol{S}^* = \arg\max_{\boldsymbol{S} \subset \boldsymbol{\Omega}} \tilde{R}_{sum}, \tag{28}$$

where $\boldsymbol{S}$ represents the set of selected user pairs and $\boldsymbol{\Omega}$ denotes the set of all user pairs. An straightforward approach for obtaining the best-K user pairs, which is called the algorithm of greedy user selection (GUS), is to use an exhaustive search over all possible selections among M user pairs. Although the GUS guarantees the best performance for (28), it is an inefficient method as the complexity grows exponentially with M. Therefore, utilizing the analytical results for the ergodic rate, we attempt to find an alternative method which is more efficient without compromising the rate performance.

Prior to this, a criterion for user scheduling is firstly established, that is, any of the available user pairs to be selected must fulfill the condition of $P_U \sigma_{g_i}^2 \ge \sigma_R^2$, for $i = 1, \ldots, K$. To achieve

better rate performance, this criterion removes those user pair that the statistical received SNR at the relay is less than 1, thus does little good to sum-rate but probably brings more interference to other use pairs. With the user scheduling criterion, $\tilde{R}_{Bi}$ is simplified by Lemma 1 below.

Lemma 1. When $P_U \sigma_{g_i}^2 > \sigma_R^2$, for $i = 1, \ldots, K$, then $\tilde{R}_{Bi}$ is approximated by

$$\hat{R}_{Bi} = \frac{1}{2} \log_2 \left(1 + N \left[\frac{\sigma_R^2}{P_U \sigma_{g_i}^2} + \frac{\sigma_{Bi}^2}{P_R} \sum_{k=1}^{K} \frac{1}{\sigma_{h_k}^2} \right]^{-1} \right). \tag{29}$$

Proof. As $N \gg K$, we substitute N for $N - K$ into (12), and when $P_U \sigma_{g_i}^2 > \sigma_R^2$, for $i = 1, \ldots, K$, the third term in the square brackets of (12) is much smaller than the second terms, which is

$$\frac{\sigma_{Bi}^2}{N P_R} \sum_{k=1}^{K} \frac{1}{\sigma_{h_k}^2} > N \cdot \frac{\sigma_R^2 \sigma_{Bi}^2}{N^2 P_U P_R} \sum_{k=1}^{K} \frac{1}{\sigma_{h_k}^2 \sigma_{g_k}^2}. \tag{30}$$

Therefore with $N \gg 1$, we omit the third term and obtain (29). $\qquad\square$

With the analytical result for rate performance in (29), we then consider the problem of choosing a subset of users that maximizes the sum rate performance, and it is formulated as

$$\boldsymbol{S^*} = \arg\max_{\boldsymbol{S} \subset \boldsymbol{\Omega}} \sum_{i=s(1)}^{s(K)} \hat{R}_{Bi}, \tag{31}$$

where $\boldsymbol{\Omega}$ represents the index set for all user pairs, i.e., $\boldsymbol{\Omega} = \{1, 2, \ldots, M\}$ and $\boldsymbol{S}$ represents the index subset for the selected user pairs, i.e., $\boldsymbol{S} = \{s(1), s(2), \ldots, s(K)\}$.

From the expression of (29) it is shown that the optimal user scheduling is irrelevant to N and the problem of (31) is reformulated as

$$\boldsymbol{S^*} = \arg\min_{\boldsymbol{S} \subset \boldsymbol{\Omega}} \prod_{i=s(1)}^{s(K)} r_{s(i)}, \tag{32}$$

where $r_{s(i)} = \frac{\sigma_R^2}{P_U \sigma_{g_{s(i)}}^2} + \sum_{k=s(1)}^{s(K)} \frac{\sigma_{Bs(i)}^2}{P_R \sigma_{h_{s(k)}}^2}$ when $N \gg 1$. Obviously the optimal subset of user pairs for (32) can be solved by brute-force searching. However, the number of evaluations for $r_{s(i)}$ in the exhaustive search is $\binom{M}{K}$. To reduce the burden of complexity, we therefore propose a simplified algorithm which greatly reduces the complexity shown as follows.

Algorithm Simplified User Selection (SUS) Algorithm

1. Initially, rearrange the elements in $\boldsymbol{\Omega}$ as $\boldsymbol{\Omega} = \{\omega(1), \omega(2), \ldots, \omega(M)\}$ to satisfy $\sigma_{h_{\omega(1)}}^2 \geq \sigma_{h_{\omega(2)}}^2 \geq \cdots \geq \sigma_{h_{\omega(M)}}^2$, and let $\boldsymbol{S} = \{\omega(1), \omega(2), \ldots, \omega(K)\}$, $\boldsymbol{U} = \boldsymbol{\Omega}/\boldsymbol{S}$.

2. For any $\omega(i) \in \boldsymbol{S}$, construct a $\boldsymbol{U_i}$ with all elements in $\boldsymbol{U}$ with $\sigma_{g_{u(l)}}^2 \geq \sigma_{g_{\omega(i)}}^2$, i.e., $\boldsymbol{U_i} = \left\{ u(l) \in \boldsymbol{U} \, \middle| \, \sigma_{g_{u(l)}}^2 \geq \sigma_{g_{\omega(i)}}^2 \right\}$. Solve the problem

$$\boldsymbol{S^*} = \{s^*(1), s^*(2), \ldots, s^*(K)\} = \arg\min_{s(i) \in \boldsymbol{U_i} \cup \{\omega(i)\}} \left\{ \prod_{i=s(1)}^{s(K)} r_{s(i)} \right\} \tag{33}$$

and the optimal user pairs subset $\boldsymbol{S^*}$ is yielded.

Computational Complexity	GUS	SUS (average)	Reduced Rate
$M = 4, K = 2$	107	67.93	0.3651
$M = 6, K = 2$	269	153.49	0.4294
$M = 8, K = 2$	503	272.66	0.4579
$M = 10, K = 5$	22679	7104.82	0.6867
$M = 15, K = 5$	270269	65524.99	0.7576
$M = 20, K = 5$	1395359	306535.46	0.7803

Table 1. *Comparison of the computational complexity (real-valued operations) for the GUS and SUS algorithms*

The key of the SUS algorithm is to use the characteristics of the results in (32) for removing the redundancy in the exhaustive search for the GUS. It can be shown[1] that the proposed SUS algorithm guarantees the same sum-rate performance as the GUS algorithm, while the complexity of SUS is notably reduced when compared to the GUS. The computational complexity analysis for the algorithms of GUS and SUS is given as follows.

Based on the expression of (32), it is derived that $\binom{M}{K} (3K^2 + 3K) - 1$ real-valued operations are required for the GUS to complete an exhaustive search, which is approximately $\mathcal{O}\left(M^K\right)$ when $M \gg K$ and may be unaffordable in practice. Besides, the computational complexity for the SUS does not maintain a constant level because it depends on the variance distribution of the channels, i.e., the ranking of both $\sigma_{g_i}^2$ and $\sigma_{h_i}^2$, $\forall i = 1, \ldots, M$. In the best case, the computational complexity for the SUS is $M(M - 1)$ real-valued operations where $\boldsymbol{U_i} = \emptyset$, for $i = 1 \cdots K$ in step 2 of the proposed algorithm. In the worst case, the complexity of the proposed algorithm is $M(M-1) + \binom{M}{K}(3K^2 + 3K) - 1$ where $\boldsymbol{U_i} = \boldsymbol{U}$, for $i = 1 \cdots K$ in step 2 of the proposed algorithm. It should be mentioned that both the best and worst cases are events of small probability, and in general, the computation complexity for the SUS can be expressed as $M(M-1) + x(3K^2 + 3K) - 1$, where x denotes the searching times in (33) and depends on the distribution with $\sigma_{g_i}^2$ and $\sigma_{h_i}^2$ and $2 \le x \le \binom{M}{K}$. It is shown that a substantial reduction of complexity can be achieved by the SUS when $\sigma_{g_i}^2$ and $\sigma_{h_i}^2$ both obey uniform random distribution, as illustrated in Table 1.

To validate the superiority for the proposed algorithm, we show the computational complexity results for the SUS in Table 1 by averaging 10^4 samples of $\sigma_{g_i}^2$ and $\sigma_{h_i}^2$, which is uniformly distributed between 0.1 and 1. We compare the average computational complexity for SUS to that for GUS with different values of K and M in Table 1, which reveals that the reduction ratio of the complexity is increased with the increase of M and K, especially as M and K both increase with an equal proportion. For example, for $K = 2$, the complexity reduction is at least 36% and is increased with the increase of M, while for $K = 5$, the complexity reduction is at least 68% and is increased with the increase of M as well. It should be mentioned that compared to the improvement for complexity reduction by enhancing M, a remarkable improvement for that of nearly 32% is achieved when K and M are both enhanced by 2.5-fold. Therefore, it is shown that a remarkable reduction of complexity for the SUS can be achieved in practical assumption that all channels fading are with random uniform distribution, and the proposed algorithm becomes more efficient as the number of user pairs (both M and K) increases. Furthermore, the rate performance for the SUS compared with the GUS and other algorithm will be shown in Section 6.

---

[1]The details for the proof can be easily established through comparison with the exhaustive search, and are thus omitted.

Remark 1. From the expression of $r_{s(i)}$ we can see that the relative magnitude of P_R and P_U does act on the scheme. In particular, the scheme will degrade to picking out the subset of the user pairs with max-K $\sigma_{g_i}^2$ when $P_R \gg P_U$, and the subset of the max-K $\sigma_{h_i}^2$'s user pairs would be chosen if $P_U \gg P_R$ (However, this assumption is impractical and is thus not further discussed).

6. Numerical Results

In the Monte-Carlo simulations, the results were generated by the average of 10^5 samples and it was assumed that $\sigma_R^2 = \sigma_{Bi}^2 = 1$, $\forall i$ for simplicity. The cases of $K = 2$ and $K = 4$ were considered and σ_{g_i} and σ_{h_i} are set to be random between 0.1 and 1, unless otherwise specified.
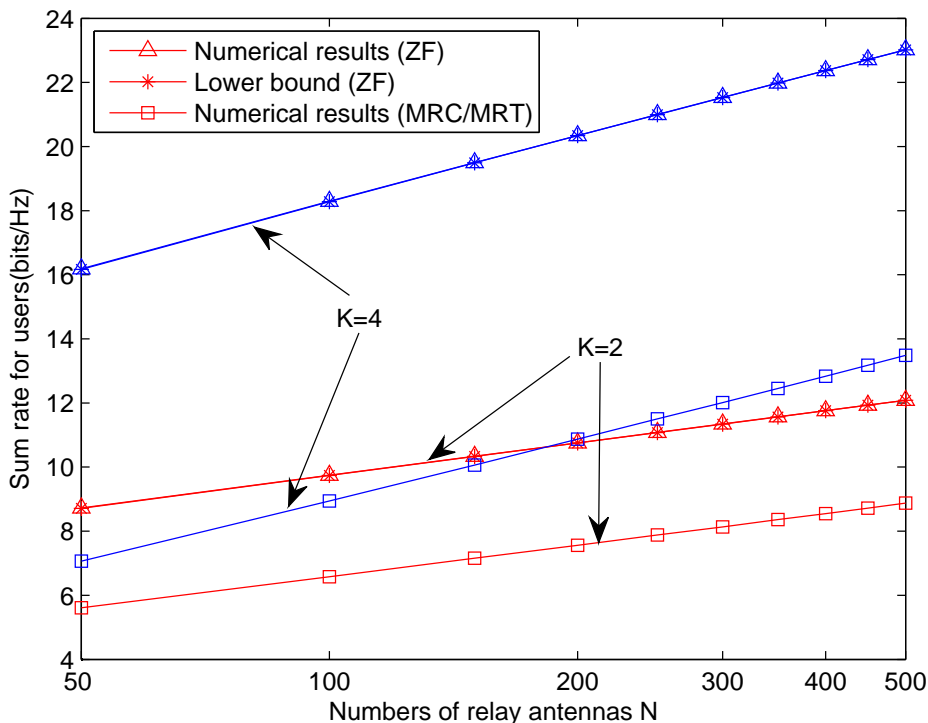
Fig. 1. Sum rate versus the number of relay antennas N (for K = 2 and K = 4).

6.1. Ergodic Achievable Rates

In this subsection, we evaluate the validity of approximation given by (12) as well as the upper bounds as shown in (22). We first compare the simulated sum-rate for all user pairs with the analytical results given by (12) where the powers are set by $P_U = 20$ and $P_R = 100$. Fig. 1 shows the simulated ergodic sum-rate versus the number of relay antennas N for $K = 2$ and $K = 4$. We can see that the approximations for (12) are very tight, even in the scenario of finite number of relay antennas, e.g., $N = 50$. This means that the proposed approximations derived from finite N are a good predictor of the ergodic achievable rates for users wherever N is large enough or not. The simulation results also show that the spectral efficiency increases logarithmically with N.
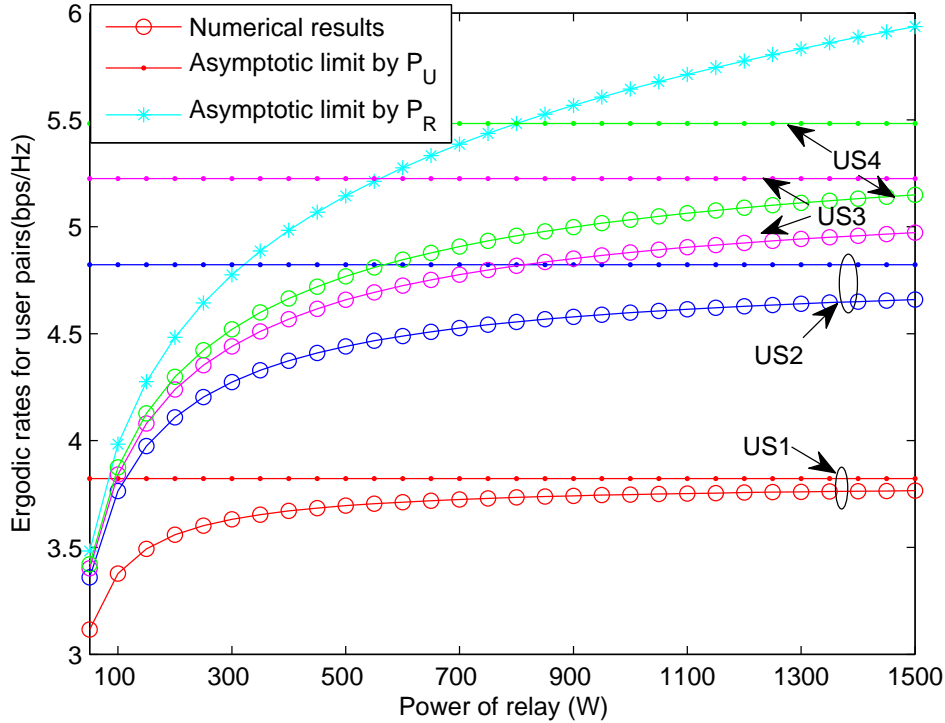
Fig. 2. *User-pair rates and the asymptotic limits related to P_U or P_R (for $K = 4$).*

This has verified the conclusion for Corollary 1. Meanwhile, Fig. 1 also compares the numerical results for relaying massive MIMO system by using ZF processing method with that by using MRC/MRT processing method. The figure shows that ZF relaying provides better performance than MRC/MRT relaying in massive MIMO systems. In particular, it is shown in Fig. 1 that ZF outperforms MRC/MRT significantly for higher numbers of user pairs.

Next, we evaluate the validity of the asymptotic limits given by (22). For easier comparison, we set $\left\{\sigma^2_{g_i}\right\} = [0.1, 0.4, 0.7, 1.0]$ and $\bar{\sigma}^2_{h_K} = 0.1$, whilst P_R varies from 50 to 1500. Fig. 2 shows that each user pair rate approaches to the respective upper bound for the fixed P_U when N grows to infinity and also they are below the upper bound determined by P_R. Furthermore, the larger the user pair rate, the more slowly the rate increases, which confirms (22).

6.2. Power Scaling Laws

In this subsection, we examine the power scaling laws for the three cases: Case I, Case II and Case III and we set $E_U = 20$ and $E_R = 100$. Firstly, the average ergodic rates for Cases I and II are exhibited with $K = 2$ and $K = 4$ as shown in Fig. 3, in which we set $(\alpha, \beta) = \{(1, 1/4), (1, 1/2), (1/4, 1), (1/2, 1), (1, 1)\}$. It is shown that the user rate approaches to a fixed upper limit for the case of $\alpha = 1$ or/and $\beta = 1$,when N is sufficiently large. The limits are the same for $(1, 1/4), (1, 1/2)$, and also for $(1/4, 1), (1/2, 1)$, which confirms the results for (25) and (26) respectively. Meanwhile, the upper limits decrease when K is enhanced when $\beta = 1$, while it does not when $\alpha = 1$ only, which also verifies the conclusions for Corollary 2.

Next, the average ergodic rates for Case III are exhibited with $K = 2$ as shown in Fig. 4 where
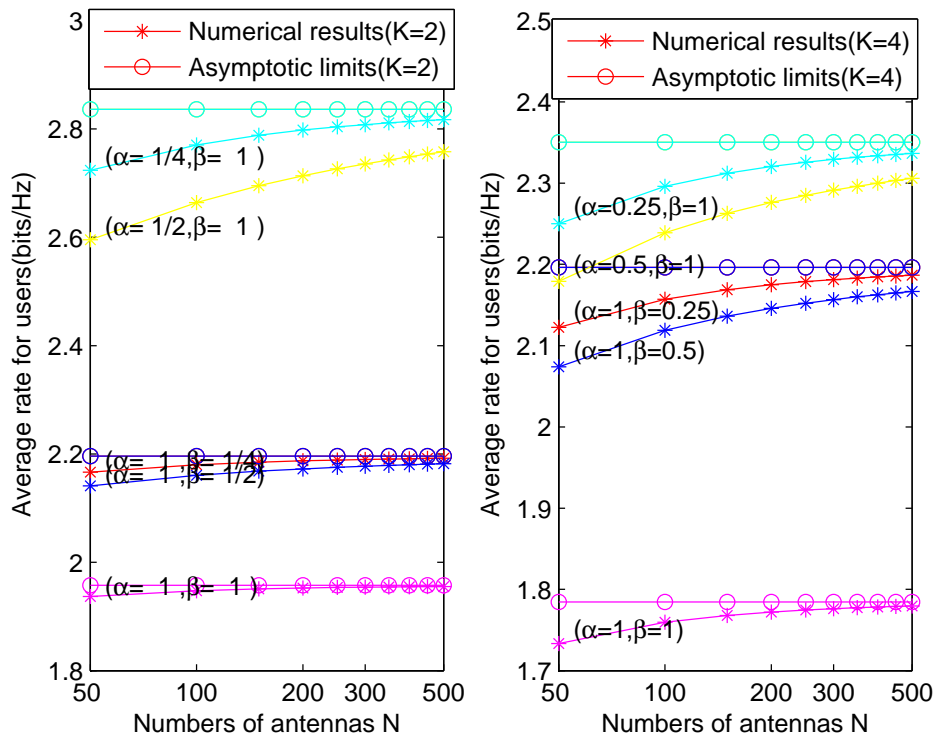
Fig. 3. *Average rate for Case I and II versus the number of relay antennas N (for $K = 2$ and $K = 4$).*
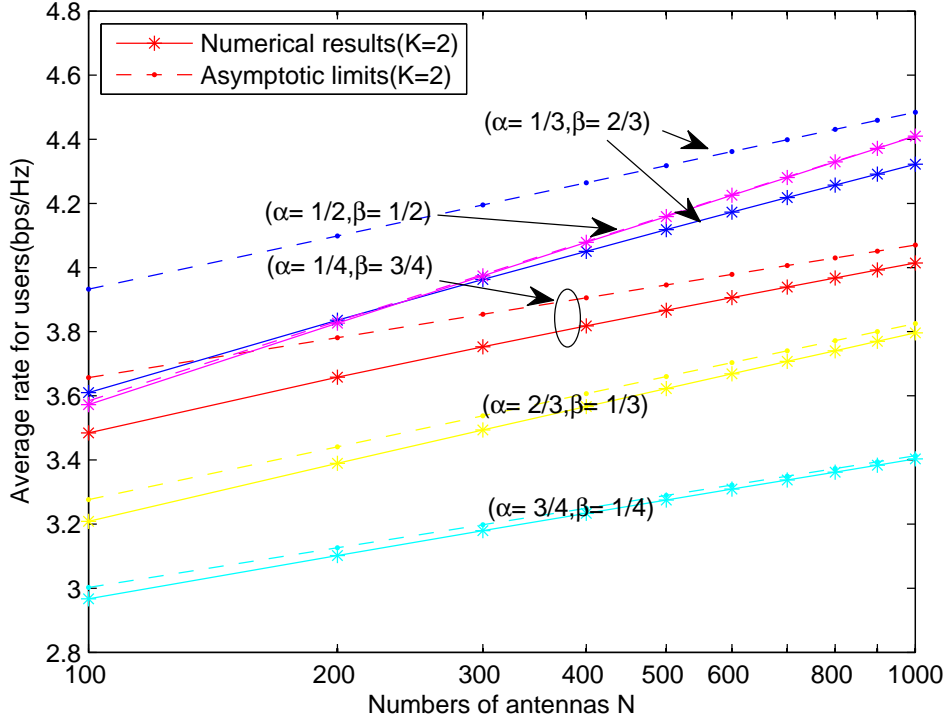
Fig. 4. *Average rate for Case III versus the number of relay antennas N (for K = 2).*

$(\alpha, \beta) = \{(1/4, 3/4), (1/3, 2/3), (1/2, 1/2), (2/3, 1/3), (3/4, 1/4)\}$. It is shown that all the average user rates increase when N grows, which means that antenna gain is obtained for all these cases. Meanwhile, the antenna gain for the case of $(1/2, 1/2)$ is the largest among all the cases, so that the average rate of which becomes the largest one when N is more than 400. When we compare the results between the case of $(1/4, 3/4)$ and $(1/3, 2/3)$, it is shown that the average rate increases when the smaller one between α and β enhances, which is indicated by (27). This conclusion is also verified for the comparison of $(2/3, 1/3)$ and $(3/4, 1/4)$, while the average rates of them are much smaller than that of former pairs because E_U is much smaller than E_R.

6.3. User Scheduling

Here, we will examine the validity of the SUS algorithm, by comparing to the GUS and random user selection (RUS) algorithms, and also to the user selection (US-G) algorithm, for which the user pairs with the max-K of $\sigma_{g_i}^2$ among all user pairs are selected. We set $\sigma_{g_i}^2$ and $\sigma_{h_i}^2$ to be uniformly distributed between 0.1 and 1, $\forall i$, and thus the RUS algorithm is performed by choosing the first K pairs for all user pairs. We also set $P_R = 100$, $K = 4$ and $M = 10$ for Fig. 5, and the rate comparison of the above algorithms is shown in Fig. 5-a and Fig. 5-b for $P_U = 20$ and $P_U = 50$, respectively. First of all Fig. 5 shows a perfect agreement between the sum-rate for the SUS algorithm and the GUS algorithm, which means the proposed algorithm can achieve the optimal sum-rate and a substantial complexity reduction. In addition, we see that there is an apparent gap between the former two algorithms and the other algorithms, in which the rate performance for the US-G is better than that for the RUS. Moreover, Fig. 5 also illustrates that the US-G performs
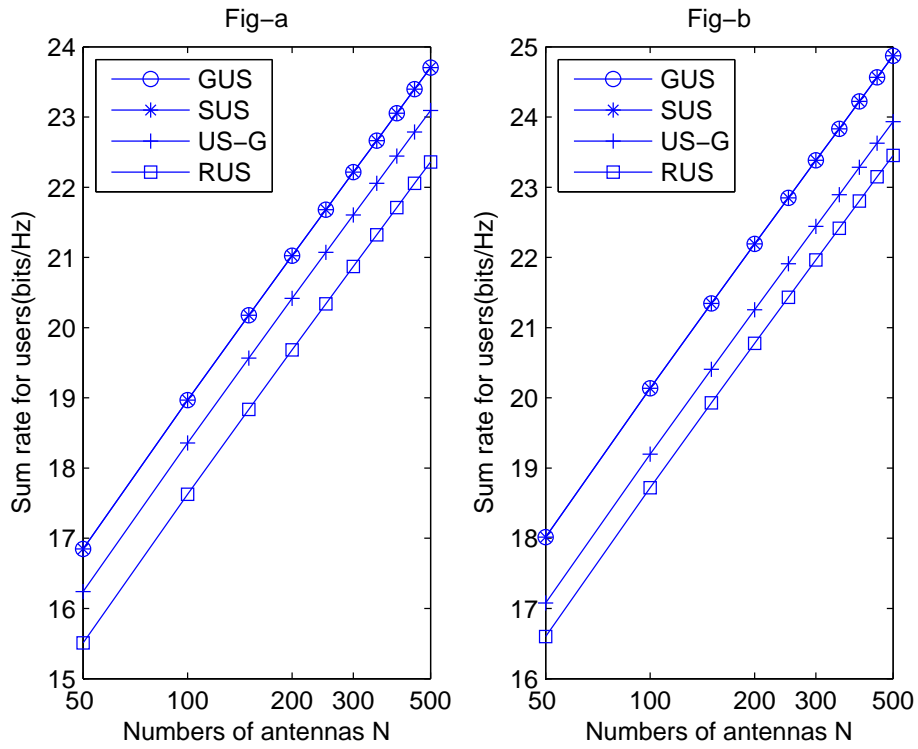
Fig. 5. _Comparison of sum rate for different user scheduling algorithms (for $P_U = 20$ and $P_U = 50$)._
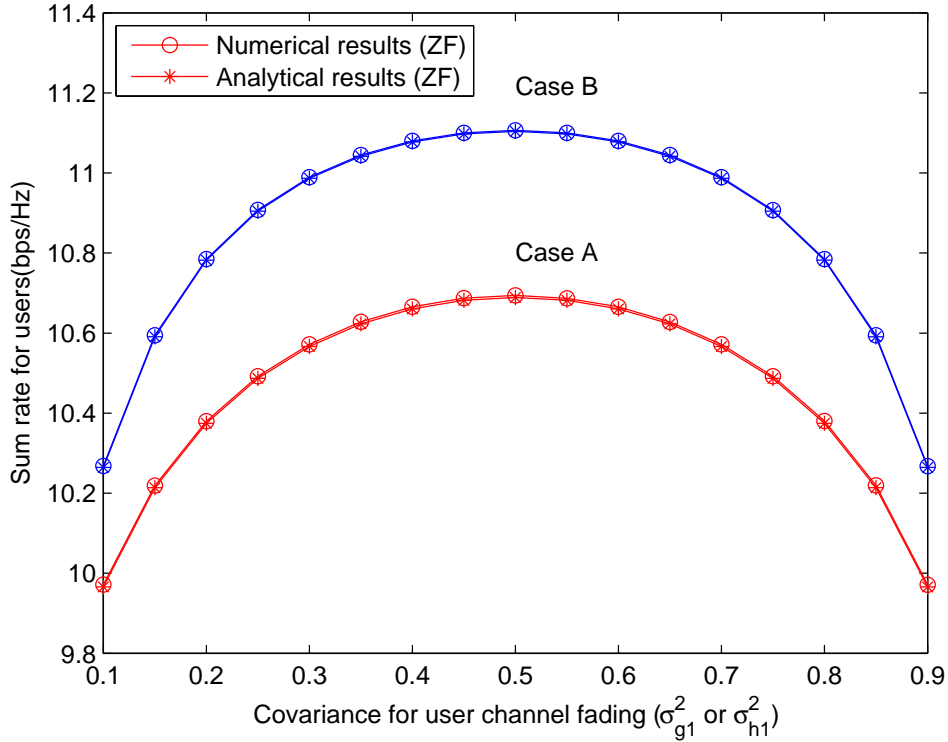
Fig. 6. *Region for achievable sum rate for user pairs (for $K = 2$ and $N = 200$).*

better when P_R is much larger than P_U, which validates the conclusion in Remark 1.

To the end, we also discuss the region of sum rate for each user pairs with different channel fading and in particular, the region is investigated for $K = 2$ and $N = 200$. Given that the sum-rate of system is effected by all the channel fading, we fix half of them for simplicity and consider the following two cases: Case A for $\sigma_{h_1}^2 = \sigma_{h_2}^2 = 1$ and $\sigma_{g_1}^2 + \sigma_{g_2}^2 = 1$, and Case B for $\sigma_{g_1}^2 = \sigma_{g_2}^2 = 1$ and $\sigma_{h_1}^2 + \sigma_{h_2}^2 = 1$ whilst $\sigma_{g_i}^2$ for Case A and $\sigma_{h_i}^2$ for Case B both vary between 0.1 and 0.9, for $i = 1, 2$. The numerical results for rate region for Cases A and B are shown in Fig. 6 where the powers are set by $P_U = 20$ and $P_R = 100$. Fig. 6 reveals that the sum rate is maximized in both two cases when the coefficients for the channel fading is equalized. This has given some insights for user scheduling strategy of $K = 2$ while it will be more complicated when K is enhanced.

7. Conclusions

This paper studied the multi-pair one-way relay system where a large number of antennas are equipped at the relay. Adopting the ZF scheme at the relay, the closed-form approximations of the ergodic rates have been derived for arbitrary number of relay antennas and the accuracy of that is verified by simulation results. Based on the analytical results, the asymptotic properties for the rate gain are discussed and power scaling laws are investigated. With the analysis in the asymptotic scenario, we revealed that the ergodic rate for user is upper limited by the equivalent SNRs for the two-hop relaying, where the former is enlarged by the power of users and the later is

enlarged by the power of relay but divided by user number. We also derived the asymptotic limits for user rate when the power of users and relay are scaled down by different factors, and discussed the power scaling law and antenna gain in each case. Through the analytical results for users, the user scheduling methods with the aim to maximize the sum-rate of system have been studied. We proposed a simplified algorithm for optimal user pair search, without the loss of rate performance. Furthermore, the complexity of scheduling decreases significantly by the proposed algorithm, and the advantage is more obvious when the number of users is large.

8. References

[1] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[2] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[3] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, May 2013.

[4] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[5] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, 2014.

[6] X. Liang, S. Jin, X. Gao, and K.-K. Wong, "Ergodic rate analysis for multi-pair two-way relay large-scale antenna system," in *Proc. IEEE International Conference on Communications (ICC)*, Sydney, Australia, Jun. 2014, pp. 5819–5824.

[7] H. Suraweera, H. Q. Ngo, T. Q. Duong, C. Yuen, E. G. Larsson *et al.*, "Multi-pair amplify-and-forward relaying with very large antenna arrays," in *Proc. IEEE International Conference on Communications (ICC)*, Budapest, Hungary, Jun. 2013, pp. 4635–4640.

[8] H. Cui, L. Song, and B. Jiao, "Multi-pair two-way amplify-and-forward relaying with very large number of relay antennas," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2636–2645, May 2014.

[9] H. Q. Ngo and E. Larsson, "Large-scale multipair two-way relay networks with distributed AF beamforming," *IEEE Commun. Lett.*, vol. 17, no. 12, pp. 1–4, Dec. 2013.

[10] H. Yang and T. Marzetta, "Performance of pilot reuse in multi-cell massive MIMO," in *Proc. IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, Constanta, Romania, May 2015, pp. 157–161.

[11] X. Zheng, E. Liu, Z. Zhang, X. Qu, R. Wang, X. Yin, and F. Liu, "An efficient pilot scheme in large-scale two-way relay systems," *IEEE Commun. Lett.*, vol. 19, no. 6, pp. 1061–1064, Jun. 2015.

[12] R. Rolny, C. Dunner, and A. Wittneben, "Power control for cellular networks with large antenna arrays and ubiquitous relaying," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Toronto, Canada, Jun. 2014, pp. 65–69.

[13] J. S. Lemos, F. Rosário, F. A. Monteiro, J. Xavier, and A. Rodrigues, "Massive MIMO full-duplex relaying with optimal power allocation for independent multipairs," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Stockholm, Sweden, Jun. 2015, pp. 306–310.

[14] G. Amarasuriya and H. Poor, "Impact of channel aging in multi-way relay networks with massive MIMO," in *Proc. IEEE International Conference on Communications (ICC)*, London, UK, Jun. 2015, pp. 1951–1957.

[15] X. Yang, X. Liang, X. Zhang, S. Jin, and T. S. Luo, "Power Scaling Laws for Massive MIMO Relay Systems with Linear Transceivers," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Orlando, Florida, USA, Dec. 2015, (to appear).

[16] B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and J. Lilleberg, "Full-duplex in large-scale wireless systems," in *Proc. Asilomar Conf. Signals, Syst., Comput. (ASILOMAR)*. Pacific Grove, CA, USA: IEEE, Nov. 2013, pp. 1623–1627.

[17] Y. Jang, K. Min, S. Park, and S. Choi, "Spatial resource utilization to maximize uplink spectral efficiency in full-duplex massive MIMO," in *Proc. IEEE International Conference on Communications (ICC)*, London, UK, Jun. 2015, pp. 1583–1588.

[18] H. Q. Ngo, H. Suraweera, M. Matthaiou, and E. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Sep. 2014.

[19] Z. Zhang, Z. Chen, M. Shen, B. Xia, and L. Luo, "Achievable rate analysis for multi-pair two-way massive MIMO full-duplex relay systems," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Hong Kong, China, Jun. 2015, pp. 2598–2602.

[20] S. Jin, X. Liang, K.-K. Wong, X. Gao, and Q. Zhu, "Ergodic rate analysis for multipair massive MIMO two-way relay networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1480–1491, Mar. 2015.

[21] J. A. Tague and C. I. Caldwell, "Expectations of useful complex Wishart forms," *Multidimensional Systems and Signal Processing*, vol. 5, no. 3, pp. 263–279, Jul. 1994.

[22] Zhang, Zhanzhan, et al., "Spectral and energy efficiency of multipair two-way full-duplex relay systems with massive MIMO," *IEEE journal on selected areas in communications*, vol 34, no. 4, pp. 848–863, Mar. 2016.