Joint Power Allocation and Relay Grouping for Large MIMO Relay Network with Successive Relaying Protocol

Hong Peng, Changran Su, Yu Zhang(✉), Linjie Xie, and Weidang Lu

College of Information Engineering, Zhejiang University of Technology, Hangzhou, China
{ph, yzhang, luweid}@zjut.edu.cn, schangran@163.com, xielingjie@hotmail.com

Abstract. In this paper, we consider a large relay network with one base station (BS), multiple users and quantities of relays, where the BS and relays are equipped with multiple antennas. Firstly, we propose an amplify-and-forward (AF) transmission scheme with successive relaying protocol for the network. Then both relay grouping and power allocation over multiple data streams at the BS are jointly optimized to further improve the proposed scheme. Numerical results show that the achievable system spectrum efficiency is considerably improved by the proposed optimized relaying scheme.

Keywords: Successive relaying · MIMO · Power allocation · Relay grouping

1 Introduction

Large relay network receives considerable attentions, since the next generation network typically supports a large number of devices which can serve as relays to improve the achievable rate of the network. On the other hand, the technique of multiple-input multiple-output (MIMO) yields great improvements in spectral efficiency through spatial multiplexing, link reliability through space-time coding, and coverage [1, 2] through array gain. Precoding and power allocation over the data streams can further bring performance gain [3, 4].

In conventional relay networks, the source transmits the data frame in the first time slot and the relays forward in the second. It is proved the performance of such one way system with half-duplex protocol will suffer a multiplexing loss (1/2 rate-loss) [5]. Then a new protocol called successive relaying protocol is put forward, in which the relays are divided into two groups to in turns receive and forward the message. Results in [6, 7] have shown that successive relaying protocol performs well in recovering multiplexing loss, so that the achievable rate of network can be markedly improved, especially in large relay networks.

This paper considers a large one-way MIMO relay network in which a BS equipped with M antennas transits messages to M user terminals (each is equipped with single antenna) with the help of quantities of relays. We propose an AF based successive relaying protocol in which each multi-antenna relay is assigned to receive (through

X. Gu et al. (Eds.): MLICOM 2017, Part II, LNICST 227, pp. 273–282, 2018.
https://doi.org/10.1007/978-3-319-73447-7_31

matched filtering) and forward (through amplify-and-forward scheme) one of the M multiplexed data streams. Furthermore, we jointly consider the power allocation over the M multiplexed data streams at the BS as well as the relay grouping scheme. Numerical results show that the achievable rate of the network can be considerably improved by the proposed joint optimization scheme.

The remaining of this paper is organized as follows. In Sect. 2, we introduce the system model. In Sect. 3, we introduce the improved successive relaying protocol. In Sect. 4, we investigate the impact of joint optimization. Section 5 presents the numerical simulation. We conclude in Sect. 6.

2 System Model

As shown in Fig. 1, we consider a large relay network with one BS (base station), M users and quantities of relays (each equipped with N antennas) which are randomly distributed in a given area. We assume no direct links between BS and users due to long distances and obstacles. We pick K relays out to help transmission. We assume perfect local CSI for each relay and statistic global CSI for BS and each user terminal. Since the considered network consists of a large number of relays, it is very beneficial for the BS and users to avoid the acquisition of the global CSI. However, it's easy to get perfect CSI at each relay with training sequences. The BS which is equipped with M antennas first processes the M data streams with power allocation matrix **P** then transmits the data streams with the help of K relays to M user terminals with successive relaying protocol. The channel for each link experiences independent Rayleigh fading across each time slot.

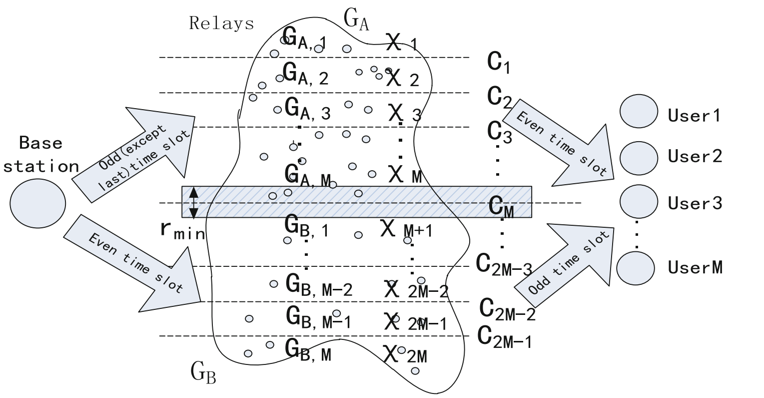

Fig. 1. Successive relaying protocol (when the order of group is in sequential order)

3 Relaying Protocol

As shown in Fig. 1, before transmission, we divide the domain into 2 * M areas denoted as $\chi_m(\mathrm{m} = 1,\ 2\ ,\ldots,\ 2\ *\ \mathrm{M})$. We denote C_m (m = 1, 2, ..., 2 * M−1) as the boundary line between area χ_m and area χ_{m+1} (note that χ_1 and χ_{2M} is bounded by the border of whole area). The relays in area χ_1 to χ_M forms G_A, while the relays in area χ_{M+1} to χ_{2M} forms G_B. We set a minimal distance r_{min} between χ_M and χ_{M+1} to avoid a large inter-relay-interference [7]. Then take G_A as an example, the relays in each area χ_m (m = 1, 2, …, M) help one of the M users, which are denoted as group $G_{A,m}$ if they help the *m*th user. Note that Fig. 1 shows an example of sequential relay group order where the relays in χ_m and χ_{m+M} (m = 1, 2, …, M) serve the *m*th user thus are denoted as $G_{A,m}$ and $G_{B,m}$ respectively. In fact, the boundary of each area and the order of relay groups affect the system performance and need optimization which will be discussed in Sect. 4.

3.1 Successive Relaying Protocol

We assume that a total of $L-1$ sub-frames are transmitted within L (an odd L is assumed) time slots as a transmission round. Relays in G_A and G_B in turns receive and forward sub-frame from the BS.

As shown in Fig. 1, in the first time slot, the BS carries out power allocation to the original data streams, and then transmits the processed sub-frame to the relays of G_A in the network. The relay k in G_A receives the signal:

$$\mathbf{r}_k^{(1)} = \mathbf{H}_k^{(1)}\mathbf{P}^{(1)}\mathbf{s}^{(1)} + \boldsymbol{\omega}_k^{(1)},\ k \in G_A \tag{1}$$

where $\mathbf{s}^{(l)} = [s_1^{(l)} s_2^{(l)} \ldots s_M^{(l)}]^T$ ($\|\mathbf{s}^{(l)}\| = 1$) denotes the original data frame transmitted in l th time slot and $\mathbf{P} = [\mathbf{p}_1\ \mathbf{p}_2 \ldots \mathbf{p}_M]$ ($\mathbf{P}$ is a diagonal matrix) for the $M \times M$ power allocation matrix. Hence $\mathbf{Ps}$ is the transmitted signal from BS. $\mathbf{s}$ and $\mathbf{P}$ satisfies the restriction $\|\mathbf{Ps}\|^2 = P_S$, where P_S is the power constraint at BS. $\mathbf{H}_\mathbf{k}^{(l)} = \left[\mathbf{h}_1^{(l)}\ \mathbf{h}_2^{(l)} \ldots \mathbf{h}_k^{(l)} \ldots\ \mathbf{h}_\mathbf{N}^{(l)}\right]^T$ is the $N \times M$ channel coefficient matrix between BS to the relay k in group G_A and each element in $\mathbf{h}_k^{(l)}$ satisfies $CN(0, \theta_k)$ (complex Gaussian distributed i.i.d.). θ_k is the variance of the channel coefficient between BS and relay k which depends on pathloss. $\boldsymbol{\omega}_k$ is the additive white Gaussian noise (AWGN) sampled at relay k with unit variance.

In the second time slot and the following even time slots l (for $l = 2, 4, \ldots L-1$) the relays in group G_A forward the received sub-frame to the user terminals. The i th user receives the signals from relays in group G_A:

$$y_i^{(l)} = \sum_{m=1}^{M} \sum_{k \in G_{A,m}} \mathbf{g}_{k,i}^{(l)^H} \mathbf{t}_{A,k}^{(l)} + \boldsymbol{\omega}_D^{(l)}, \quad i = 1, 2, \ldots, M \tag{2}$$

where $\mathbf{g}_{k,i}^{(l)}$ is the channel coefficient vector from relay k to the i th user, in which the element satisfies $CN\left(0, \beta_{k,i}\right)$. $\beta_{k,i}$ is the variance of the channel coefficient between the i th user and relay k. $\boldsymbol{\omega}_D^{(l)}$ is AWGN sampled at i th user with unit variance in the l th time slot. $\mathbf{t}_{A,k}^{(l)}$ denotes the transmitted signal from relay k in group G_A in the l th time slot which satisfies $Ef\left\|\mathbf{t}_{A,k}^{(l)}\right\|^2 = \frac{P_R}{M_A}$, where P_R is the total transmission power constraint of all the relays. M_A denotes the number of relays in group G_A. The generation of $\mathbf{t}_{A,k}^{(l)}$ will be discussed in Sect. 3.2.

Meanwhile, the BS transmits new sub-frame to relays in group G_B. The relay k' in group G_B receives:

$$\mathbf{r}_{k'}^{(l)} = \mathbf{H}_{k'}^{(l)}\mathbf{P}^{(l)}\mathbf{s}^{(l)} + \sum_{k \in G_A} \mathbf{F}_{k',k}^{(l)}\mathbf{t}_{A,k}^{(l)} + \boldsymbol{\omega}_{k'}^{(l)}, \quad k' \in G_B \tag{3}$$

where $\mathbf{F}_{k',k}^{(l)}$ denotes the $N \times N$ channel coefficients matrix from relay k' in group G_B to relay k in group G_A in the l th time slot, in which all the entries satisfy $CN\left(0, \eta_{k',k}\right)$. $\eta_{k',k}$ denotes the variance of channel coefficient between relay k' in group G_B and relay k in group G_A.

In the third time slot and the following odd time slots (for $l = 1, 3, \ldots, L$), relays in group G_B transit signal to the i th user. Meanwhile relay k in group G_A receives the signal from the BS:

$$\mathbf{r}_{k}^{(l)} = \mathbf{H}_{k}^{(l)}\mathbf{P}^{(l)}\mathbf{s}^{(l)} + \sum_{k' \in G_B} \mathbf{F}_{k,k'}^{(l)}\mathbf{t}_{B,k'}^{(l)} + \boldsymbol{\omega}_{k}^{(l)}, \ k \in G_A \tag{4}$$

where $\mathbf{t}_{B,k'}^{(l)}$ denotes the signal transmitted by relay k' in group G_B in the l th time slot. Meanwhile, The i th user receive the signals from relays in group G_B:

$$y_i^{(l)} = \sum_{m=1}^{M} \sum_{k \in G_{B,m}} \mathbf{g}_{k,i}^{(l)^H} \mathbf{t}_{B,k'}^{(l)} + \boldsymbol{\omega}_D^{(l)}, k' \in G_B, \ i = 1, 2, \ldots, M \tag{5}$$

In the last time slot L (recall that L is odd), relays in group G_A keep silent, meanwhile the relays in group G_B transmit signal to the user terminals, the situation is similar to (5).

3.2 Amplify-and Forward After Matched Filtering

Recall that we assign each relay groups to serve for one of M users. That's to say, relay group $G_{A,m}$ and relay group $G_{B,m}$ take responsibility to receive and forward the m th user's message. Without loss of generality, we consider relay k in group G_A serves for the n th user. All the relays receive signal with matched filtering. Take the l th (l is even and $l > 2$) time slot as example. According to (4), the relay k in group G_A processes the received signals as follows:

$$\begin{aligned} \mathbf{t}_{A,k}^{(l)} &= \gamma_k^{(l)} \mathbf{g}_{k,n}^{(l)} \widetilde{\mathbf{h}}_{k,n}^{(l-1)^H} \mathbf{r}_k^{(l-1)} \\ &= \gamma_k^{(l)} \mathbf{g}_{k,n}^{(l)} \left(\left\| \widetilde{\mathbf{h}}_{k,n}^{(l-1)} \right\|^2 \mathbf{s}_n^{(l-1)} + \sum_{j=1, j \neq n}^{M} \widetilde{\mathbf{h}}_{k,n}^{(l-1)^H} \widetilde{\mathbf{h}}_{k,j}^{(l-1)} s_j^{(l-1)} + \widetilde{\mathbf{h}}_{k,n}^{(l-1)} \boldsymbol{\omega}_k^{(l-1)} + \sum_{k' \in G_B} \widetilde{\mathbf{h}}_{k,n}^{(l-1)^H} \mathbf{F}_{k,k'}^{(l-1)} \mathbf{t}_{B,k}^{(l-1)} \right), \ k \in G_A \end{aligned} \tag{6}$$

where $\tilde{\mathbf{h}}_{k,n} = [\mathbf{h}_1\mathbf{p}_n \ \mathbf{h}_2\mathbf{p}_n \ \ldots \ \mathbf{h}_N\mathbf{p}_n]^T$ is equivalent channel coefficient matrix between the BS and relay k. The variance of each element in $\tilde{\mathbf{h}}_{k,n}$ denoted as $\alpha_{k,n}$. Thus we can derive that $\alpha_{k,n} = \mathrm{P}_{1,n}^2\theta_k + \mathrm{P}_{2,n}^2\theta_k + \ldots + \mathrm{P}_{\mathrm{Q},n}^2\theta_k$. The amplifying coefficient $\gamma_k^{(l)}$ of relay k in group G_A can be calculated according to the relay power restriction $Ef\left\|\mathbf{t}_{\mathrm{A},k}^{(l)}\right\|^2 = \frac{\mathrm{P}_{\mathrm{R}}}{M_A}$. The amplifying coefficient $\gamma_k^{(l)}$ in even time slot l is derived as:

$$\gamma_k^{(l)} = \sqrt{\frac{\frac{\mathrm{P_R}}{\mathrm{M}_A}}{\frac{1}{M}p_s\sum_{k=1}^N \beta_{k,n}\sum_{k=1}^N 3\alpha_{k,n}^2 + \sum_{j=1,j\neq n}^{\mathrm{M}}\frac{1}{\mathrm{M}}p_s\sum_{k=1}^N \beta_{k,n}\sum_{k=1}^N \alpha_{k,n}\alpha_{k,j} + \sum_{k=1}^N \beta_{k,n}\sum_{k=1}^N \alpha_{k,n} + \frac{\mathrm{P_R}}{\mathrm{NM_B}}\sum_{k=1}^N \beta_{k,n}\sum_{k=1}^N \alpha_{k,n}\sum_{k\in G_B}\sum_{k=1}^N \eta_{k,\mathrm{k}'}}}, k \in \mathrm{G}_A, l = 4, 6\ldots, \mathrm{L}-1 \tag{7}$$

The detailed calculation is similar to the work in [7], which is omitted here. Similarly we can derive the coefficient in the odd and second time slot. Substituting (6) into (2), the received sub-frame at the i th user is given by:

$$y_i^{(l)} = \underbrace{Ef\left(u_{m,k,i}^{(l-1)}\right)s_i^{(l-1)}}_{\mathrm{L}_1} + \underbrace{\left(u_{m,k,i}^{(l-1)} - Ef\left(u_{m,k,i}^{(l-1)}\right)\right)}_{\mathrm{L}_{error}} s_i^{(l-1)} + \underbrace{u_{m,k,j}^{(l-1)}s_j^{(l-1)}}_{\mathrm{L}_2} + \underbrace{v_{m,k}^{(l-1)}\boldsymbol{\omega}_k^{(l-1)}}_{\mathrm{L}_3} + \underbrace{v_{m,k}^{(l-1)}\sum_{k'\in G_B}\mathbf{F}_{k,k'}^{(l-1)}\mathbf{t}_{\mathrm{B},k'}^{(l-1)}}_{L_4} + \boldsymbol{\omega}_D^{(l)} \tag{8}$$

where the following notations are used:

$$\begin{aligned}
u_{m,k,i}^{(l-1)} &= \sum_{k\in \mathrm{G}_{A,i}} \gamma_k^{(l)}\left\|\mathbf{g}_{k,i}^{(l)}\right\|^2\left\|\tilde{\mathbf{h}}_{k,i}^{(l-1)}\right\|^2 + \sum_{m=1,m\neq i}^{M}\sum_{k\in \mathrm{G}_{A,m}}\gamma_k^{(l)}\mathbf{g}_{k,i}^{(l)^H}\mathbf{g}_{k,m}^{(l)}\tilde{\mathbf{h}}_{k,m}^{(l-1)^H}\tilde{\mathbf{h}}_{k,i}^{(l-1)} \\
u_{m,k,j}^{(l-1)} &= \sum_{j=1,j\neq i}^{M}\left(\sum_{k\in \mathrm{G}_{A,i}}\gamma_k^{(l)}\left\|\mathbf{g}_{k,i}^{(l)}\right\|^2\tilde{\mathbf{h}}_{k,i}^{(l-1)^H}\tilde{\mathbf{h}}_{k,j}^{(l-1)} + \sum_{k\in \mathrm{G}_{A,j}}\gamma_k^{(l)}\mathbf{g}_{k,i}^{(l)^H}\mathbf{g}_{k,j}^{(l)}\left\|\tilde{\mathbf{h}}_{k,j}^{(l-1)}\right\|^2 + \sum_{m=1,m\neq i,j}^{M}\sum_{k\in \mathrm{G}_{A,m}}\gamma_k^{(l)}\mathbf{g}_{k,i}^{(l)^H}\mathbf{g}_{k,m}^{(l)}\tilde{\mathbf{h}}_{k,m}^{(l-1)^H}\tilde{\mathbf{h}}_{k,j}^{(l-1)}\right) \\
v_{m,k}^{(l-1)} &= \sum_{k\in \mathrm{G}_{A,i}}\gamma_k^{(l)}\left\|\mathbf{g}_{k,i}^{(l)}\right\|^2\tilde{\mathbf{h}}_{k,i}^{(l-1)^H} + \sum_{m=1,m\neq i}^{M}\sum_{k\in \mathrm{G}_{A,m}}\gamma_k^{(l)}\mathbf{g}_{k,i}^{(l)^H}\mathbf{g}_{k,m}^{(l)}\tilde{\mathbf{h}}_{k,m}^{(l-1)^H}
\end{aligned} \tag{9}$$

where L_1 is the desired data of the i th user, L_{error} is the interference caused by using the statistical channel information for signal detection and decoding. L_2 is the interference from the other sources and L_3 is the noise from the relays. L_4 is the interference from relay group G_B. It should be noted that in the second time slot, L_4 doesn't exist since relays in group G_B are silent in the first time slot.

Each user terminal applies coherent detection with the statistical channel information. The achievable capacity of the i th user in the even time slot is given by:

$$R_i^{(even)} = \log\left(1 + \frac{Ef|\mathrm{L}_1|^2}{\sum_{n=2}^4 Ef|\mathrm{L}_n|^2 + Ef|\mathrm{L}_{error}|^2 + 1}\right), \quad l = 4, 6, \ldots, \mathrm{L}-1 \tag{10}$$

where $Ef|\mathrm{L}_n|^2$(for n = 1, …, 4) and $Ef|\mathrm{L}_{error}|^2$ is calculated as follows:

$$Ef|\mathrm{L}_1|^2 = \frac{\mathrm{P}_S}{\mathrm{M}}\left(\sum_{k\in \mathrm{G}_{A,i}} \gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k,i}}^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,i}}^2 + 2\sum_{\mathrm{k1,k2}\in \mathrm{G}_{A,i}} \gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k1,i}} \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k1,i}} \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k2,i}} \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k2,i}}\right)$$

$$Ef|\mathrm{L}_{error}|^2 = \frac{\mathrm{P}_S}{\mathrm{M}}\left(\sum_{k\in \mathrm{G}_{A,i}} 8\gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k,i}}^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,i}}^2 + \sum_{m=1,\mathrm{m}\neq \mathrm{i}}^{\mathrm{M}} \sum_{k\in G_{A,m}} \gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k,i}}\beta_{\mathrm{k,m}} \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,i}}\alpha_{\mathrm{k,m}}\right)$$

$$Ef|\mathrm{L}_2|^2 = \frac{\mathrm{P}_S}{\mathrm{M}} \sum_{j=1,j\neq i}^{\mathrm{M}} \left(\sum_{m=1,m\neq i,j}^{M} \sum_{k\in G_{A,m}} \gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k,i}}\beta_{\mathrm{k,m}} \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,m}}\alpha_{k,j} + \sum_{k\in G_{A,i}} \gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} 3\beta_{\mathrm{k,i}}^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,i}}\alpha_{k,j} \right.$$

$$\left. + \sum_{k\in G_{A,j}} \gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k,i}}\beta_{\mathrm{k,j}} \sum_{\mathrm{k=1}}^{\mathrm{N}} 3\alpha_{\mathrm{k,j}}^2 \right)$$

$$Ef|\mathrm{L}_3|^2 = \sum_{m=1,m\neq i}^{M} \sum_{k\in G_{A,m}} \gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k,i}}\beta_{\mathrm{k,m}} \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,m}} + \sum_{k_k\in G_{A,i}} \gamma_k^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \beta_{\mathrm{k,i}}^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,i}}$$

$$\begin{aligned} Ef|\mathrm{L}_4|^2 = & \sum_{k\in G_{A,i}} \gamma_k^2 \frac{\mathrm{P_R}}{\mathrm{NM_B}} \sum_{k=1}^{N} 3\beta_{\mathrm{k,i}}^2 \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,i}} \sum_{k'\in G_B} \sum_{k'=1}^{N} \eta_{k,k'} \\ & + \sum_{m=1,m\neq i}^{M} \sum_{k\in G_{A,m}} \gamma_k^2 \frac{\mathrm{P_R}}{\mathrm{NM_B}} \sum_{k=1}^{N} \beta_{k,i}\beta_{k,m} \sum_{\mathrm{k=1}}^{\mathrm{N}} \alpha_{\mathrm{k,m}} \sum_{k'\in G_B} \sum_{k'=1}^{N} \eta_{k,k'} \end{aligned} \quad (11)$$

The detailed calculation of $Ef|\mathrm{L}_n|^2$(for n = 1, …, 5) and $Ef|\mathrm{L}_{error}|^2$ is similar to [7].

Similarly we can derive the i th user's achievable rate in the second time slot and odd time slots denoted by $R_i^{(odd)}$ and $R_i^{(2)}$. The detailed expression is omitted here. We assume the channel varies in different time slot but the variance of channel coefficient is stable from time to time. Thus, we can derive the sum achievable rate of the network as follows:

$$R_{sum} = \frac{1}{\mathrm{L}} \sum_{i=1}^{M} \left(\frac{\mathrm{L}-1}{2} R_i^{(odd)} + \frac{\mathrm{L}-3}{2} R_i^{(even)} + R_i^{(2)} \right) \quad (12)$$

4 Joint Optimization

In this section, we consider a joint optimization on power allocation matrix **P** and relay grouping to maximize the achievable sum rate of the network. Relay grouping includes the order of relay groups and the boundary line between different relay group (as shown in Fig. 1). The flowchart of the joint optimization is shown in Fig. 2. For the inner loop of the algorithm, for each fixed order of relay groups, we in turns optimize the power allocation matrix **P** and the boundary line with differential evolution scheme (DE). The inner loop stops when a pre-set iteration time is reached. Then for the outer loop, we exhaustively search any possible order of relay groups and do the joint optimization (i.e., the inner loop). The algorithm stops when all the orders of relay groups are traversed and the best one to maximize the achievable rate is chosen.

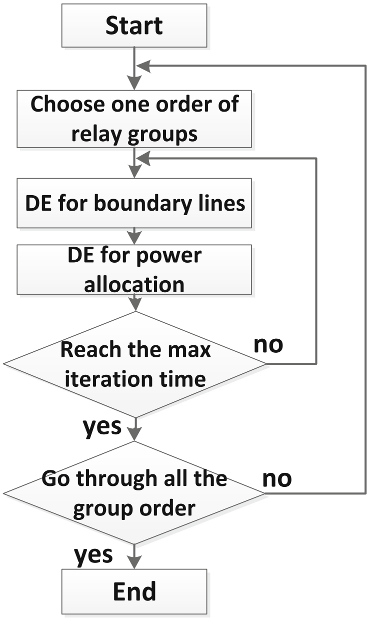

Fig. 2. Flowchart of joint optimization.

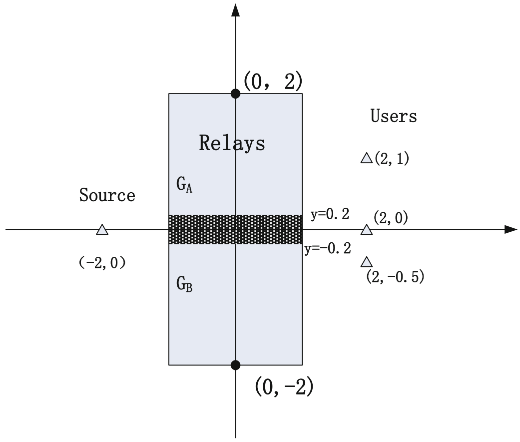

Fig. 3. The simulation scenario.

4.1 Optimization for the Order of Relay Groups

The order of relay groups affects the network rate by two factors: 1. inter-relay interference 2. large scale fading. We use exhaustive method to perform optimization on order of relay groups. We traverse all the relay group sort and then select the best one to maximize the achievable sum rate. Although exhaustive method can traverse all possible situations, the complexity of optimization will be quite high. One can design heuristic methods based on fairness principle to ensure that there is a balance between the interference and distance between the transmitter/user and the relay. We do not discuss heuristic methods here due to limited space.

4.2 Optimization for the Boundary Lines Between Relay Groups and Power Allocation

For the sake of simplicity, we denote the diagonal elements of power allocation matrix as $\tilde{\mathbf{P}} = \left[\mathrm{P}'_1, \quad \mathrm{P}'_2, \ldots, \mathrm{P}'_\mathrm{M}\right]$ and the boundary line matrix as $[\mathrm{C}_1, \ldots \mathrm{C}_\mathrm{M}, \ldots \mathrm{C}_{2\mathrm{M}-1}]$ (recall that C_m (m = 1, 2, …, 2 * M−1) is the boundary line between area χ_m and area χ_{m+1}). The group model is shown in Fig. 1.

First we optimize the boundary line. We determine the max iteration number of optimization and generate D $(D = 20 * \mathrm{M})$ initial solution candidates which meet the system constraints. We denote each solution candidate as $C_{i,G}$, where G is for the generation and i for the candidate number. The element in each candidate is defined as

$C_{j,i,G}$, where j is the element number. The first element in i th candidate in first generation is formed as follows:

$$C_{1,i,0} = rand(0,1) \cdot (C_j^{(U)} - C_j^{(L)}) + C_j^{(L)}, \quad (i = 1, 2..., D, j = 1, 2, ..., 2M - 1) \tag{13}$$

where $C_j^{(U)}$ and $C_j^{(L)}$ represent the upper bound and the lower bound of boundary line.

Similarly we can get all the element in $C_{i,0}$ and then we can have all the candidate in first generation. Next we generate the mutation of the first generation by:

$$V_{i,G+1} = C_{r1,G} + F \cdot (C_{r2,G} - C_{r3,G}) \tag{14}$$

where $r1$, $r2$, $r3$ is different to each other. $F = F_0 \cdot 2^{\lambda}$, $\lambda = e^{1-\frac{G_m}{G_m+1-G}}$, $F_0 \in (0, 0.5)$. Then we get the crossover candidate $U_{i,G+1}$ by:

$$U_{ji,G+1} = \begin{cases} V_{ji,G+1} & \text{if rand(0,1)} \leq 0.9 \ or\ j = \text{randperm(i)} \\ C_{ji,G} & else \end{cases} \tag{15}$$

where randperm (i) denotes random number between 1 and 2M−1. The next step is selection. We generate the new generation by choose better candidate to max the achievable rate between $C_{ji,G}$ and $U_{ji,G+1}$:

$$C_{i,G+1} = \begin{cases} U_{i,G+1}, & if\ f(U_{i,G+1}) > f(C_{i,G}) \\ C_{i,G}, & else \end{cases} \tag{16}$$

If the round of iteration reaches the max times, the new generation will be regarded as the optimal boundary line and then the optimization on boundary comes to an end. Otherwise, the new solution will be taken as initial solution of the next generation and the optimization continues.

Power allocation is similar to the method of optimizing the relay boundary line.

5 Numerical Simulation

In this section, we examine the performance of the proposed AF successive relaying protocol. As for the simulation scenario, we consider the network depicted in Fig. 3, in which with a BS on (−2,0) and three user terminals (M = 3) on (2,1), (2,0) and (2, −0.5). K relays are randomly and independently distributed in a bounded domain. The domain is a rectangle region whose vertices are (1,2), (1,−2), (−1,2) and (−1−2). The minimal distance $r_{\min}$ is set to 0.2 and for DE, F = 0.5, G_m = 5. The variance of channel coefficient is denoted as $1/d^3$ where d is the distance between two nodes. The power constraints on BS and all the relays are set to 100 and 50, respectively. We examine the performances of joint optimization given in Sect. 4 when the relay antenna number is fixed to N = 8 in Fig. 4 (a) and the relay number is fixed to K = 800 in Fig. 4 (b), respectively.

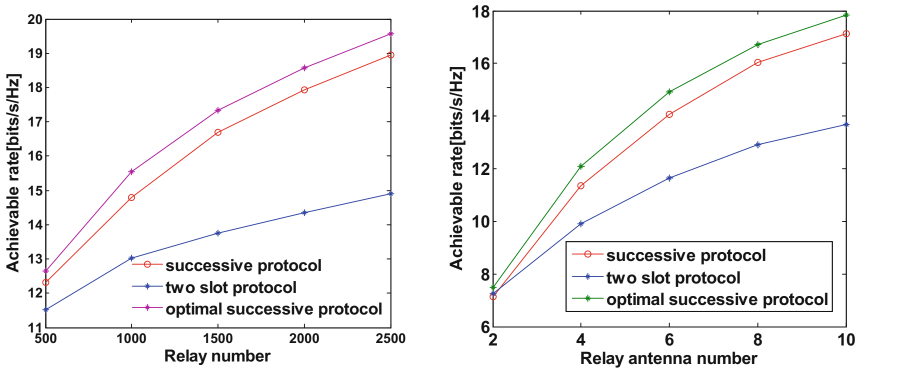

Fig. 4. Achievable sum rate over (a) different relay number; (b) different relay antenna number.

As shown in the two figures, we can see both the achievable rates with successive protocol and with two-slot protocol increase as the relay number and antenna number increase. It is observed that the rate with successive relaying protocol scales quickly than with two-slot protocol. And the performance of the successive relaying protocol is considerably improved due to the joint power allocation and relay grouping optimization.

6 Conclusion

This paper investigates the joint optimization on power allocation and relay grouping in relay network with successive relaying protocol. We consider a scenario where a base station transfers data to the user terminals through multiple relays. Each relay works in amplify-and-forward scheme. We find that the achievable rate of the system can be markedly improved by jointly optimizing the relay grouping and the power allocation matrix **P** with DE.

Acknowledgement. This work was supported by National Natural Science Foundation of China (No. 61401391) and Zhejiang Provincial Natural Science Foundation of China under Grant No. LY17F010014.

References

1. Chiurtu, N., Rimoldi, B., Telatar, E.: On the capacity of multi-antenna Gaussian channels. Eur. Trans. Telecommun. **10**(6), 585–595 (1999)
2. Bolcskei, H., Nabar, R.U., Oyman, O., Paulraj, A.J.: Capacity scaling laws in mimo relay networks. IEEE Trans. Wirel. Commun. **5**(6), 1433–1444 (2006)
3. Lozano, A., Tulino, A.M., Verdu, S.: Optimum power allocation for parallel Gaussian channels with arbitrary input distributions. IEEE Trans. Inf. Theory **52**(7), 3033–3051 (2006)

4. Vojcic, B.R., Jang, W.M.: Transmitter precoding in synchronous multiuser communications. IEEE Trans. Commun. **46**(10), 1346–1355 (1998)
5. Gastpar, M., Vetterli, M.: On the capacity of wireless networks: the relay case. In: Proceedings of Joint Conference of the IEEE Computer and Communications Societies, vol. 3, pp. 1577–1586 (2002)
6. Fan, Y., Wang, C., Thompson, J., et al.: Recovering multiplexing loss through successive relaying using repetition coding. IEEE Trans. Wirel. Commun. **6**(12), 4484–4493 (2007)
7. Zhang, Y., Zhang, Z., Ping, L., Chen, X.: Capacity scaling of relay networks with successive relaying. In: Proceedings of IEEE International Symposium on Information Theory, pp. 176–180 (2015)