SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2018, Vol. 61 022311:1–022311:18
doi: 10.1007/s11432-017-9187-5

Multi-pair massive MIMO amplify-and-forward relaying system with low-resolution ADCs: performance analysis and power control

Qimei CUI[1*], Yinjun LIU[1], Yinsheng LIU[2], Weiliang XIE[3] & Yong ZHAO[3]

[1]*National Engineering Laboratory for Mobile Network Security, Beijing University of Posts and Telecommunications, Beijing* 100876*, China;*
[2]*State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing* 100044*, China;*
[3]*China Telecom Corporation Limited, Beijing* 100033*, China*

Received 22 May 2017/Revised 11 July 2017/Accepted 26 July 2017/Published online 26 October 2017

Abstract In this paper, we focus on a general multi-pair massive MIMO amplify-and-forward (AF) relaying system where the relay antennas employ low-resolution analog-to-digital converters (ADCs) to reduce the hardware cost. First, considering the effect of low quantization on channel estimation, a tight closed form approximation of the system ergodic achievable rate is derived. Second, some asymptotic analysis is presented to reveal the impacts of the system parameters on the achievable rate. Particularly, the generalized power scaling schemes are characterized. The results indicate that in some cases, when the number of relay antennas grows without bound, the impact of the finite resolution ADCs on data transmission can be eliminated. To enhance the achievable rate of the quantized systems, the optimal user and relay power control schemes are proposed. Furthermore, to reap all the benefits of low-resolution ADCs, another power control scheme is also designed to minimize the total power consumption while guaranteeing the quality-of-service (QoS) requirement of each user, which can help draw some useful insights into the optimal ADC resolution from power saving perspectives. The simulation results confirm the accuracy of our theoretical analysis and the effectiveness of the proposed power control schemes.

Keywords massive MIMO, relaying system, imperfect CSI, low quantization, power control

Citation Cui Q M, Liu Y J, Liu Y S, et al. Multi-pair massive MIMO amplify-and-forward relaying system with low-resolution ADCs: performance analysis and power control. Sci China Inf Sci, 2018, 61(2): 022311, doi: 10.1007/s11432-017-9187-5

1 Introduction

Massive MIMO has been regarded as one of the most promising techniques of the fifth generation (5G) communication systems due to its potential to significantly enhance the system spectral efficiency (SE) and energy efficiency (EE) [1–6]. When the base stations (BSs) employ massive antennas, a large number of users can be served simultaneously at the same time-frequency resource, and the linear precoding schemes, such as maximal ratio combining (MRC) and zero forcing (ZF), can already achieve the near-optimal performance, leading to significant reduction of the computational complexity [4]. On the other hand, relay is also recognized as one of the key technologies in the future communication systems since it can enhance the network capacity and coverage [7, 8]. To take advantages of both massive MIMO and

* Corresponding author (email: cuiqimei@bupt.edu.cn)

relay, massive MIMO enabled relaying systems with different transmission protocols and operating modes have been widely investigated in [9–15].

All the aforementioned works assume that the full-precision ADCs are employed at the relay. However, the increasing number of antennas contributes to high hardware cost and power consumption, which poses new challenges on the practical implementation of massive MIMO. To address the challenge, the low-resolution ADCs, which are considered as an alternative to its power hungry high-resolution counterpart, have attracted much attention [16–20]. By applying the additive quantization noise model (AQNM), the uplink achievable rate of the quantized massive MIMO systems with low-resolution ADCs for Rayleigh and Rician channels have been investigated in [16] and [17], respectively. It is revealed in [16] that increasing the number of antennas can compensate for the rate loss introduced by the finite resolution ADCs. Therefore, finite resolution ADCs can also be applied for the relaying systems to reduce the hardware power consumption [21–23].

In [21], the uplink achievable rate of a multi-user AF relaying system under perfect CSI conditions has been discussed, where the receiver employs the mixed-resolution ADCs. However, the relay still employs high-resolution ADCs. The secrecy EE of a multi-pair AF relaying system with double-resolution ADCs has been discussed in [22]. Although the imperfect CSI from the source users to the relay is considered, the effect of low quantization on channel estimation is ignored. For a pilot-assisted quantized MIMO system, the system performance will depend on the selected pilot sequences [24, 25]. Therefore, the impact of few-bit ADCs on channel estimation should be considered. In addition, for the asymptotic analysis, the power scaling law of the quantized massive MIMO enabled relaying systems has not been revealed in [22].

Recently, by using the identity matrix as the pilot matrix, the performance and the optimal relay power allocation for the full-duplex massive MIMO relaying system with low-resolution ADCs have been investigated in [23]. When the number of relay antennas grows without bound, the asymptotic analysis is also presented. However, only one special case for power scaling law has been considered in [23]. When the scaling parameters of source users and relay are various, more useful insights into the impact of the quantization at the relay on the system performance can be obtained. In addition, the authors assume that all the source users have the same transmit power in [21–23]. Since the quantization noise at the relay is related to the input signals, the diverse user transmit power will induce a more general case.

On the other hand, power control is closely related to the relaying system performance [7, 8], which should be investigated thoroughly to enhance the performance of the quantized systems. However, none of [21] and [22] have considered the power allocation scheme of the quantized massive MIMO relaying systems. Although the relay power control has been discussed in [23], the user transmit power control has been ignored. When the transmit power of each user is different, the user and relay power can be jointly optimized to further improve the system performance. In addition, from energy efficiency perspectives, the previous works have investigated the unconstraint EE for one-hop quantized systems and only the receiver power consumption has been considered [16–20]. However, the energy-efficient design can only make sense without sacrificing the achievable rate of users [7, 8, 26, 27]. To reap all the benefits of low-resolution ADCs, for the purpose of power savings, the total power consumption of the quantized massive MIMO relaying systems can be optimized while satisfying the QoS requirements of users.

Motivated by the above concerns, in this paper, we present the performance analysis and power control schemes of a multi-pair AF massive MIMO relaying system with maximal ratio combining/maximal ratio transmission (MRC/MRT) beamforming scheme under the imperfect CSI. The relay antennas employ low-precision ADCs to reduce the hardware cost. Different from [23], in this paper, the pilot sequences, which are applied for channel estimation, are selected from Hadamard matrix. In addition, the asymptotic analysis as well as the power control schemes of the quantized massive MIMO enabled relaying system are also investigated thoroughly. The main contributions of this paper can be summarized as follows.

• Considering the impact of low-resolution ADCs on channel estimation and diverse user transmit power, we derived a closed-form approximation of the ergodic achievable rate, whose tightness is demonstrated by the simulation results. Based on the derived expression, the joint user transmit power and relay power optimization is proposed to enhance the sum achievable rate of the quantized MIMO relaying

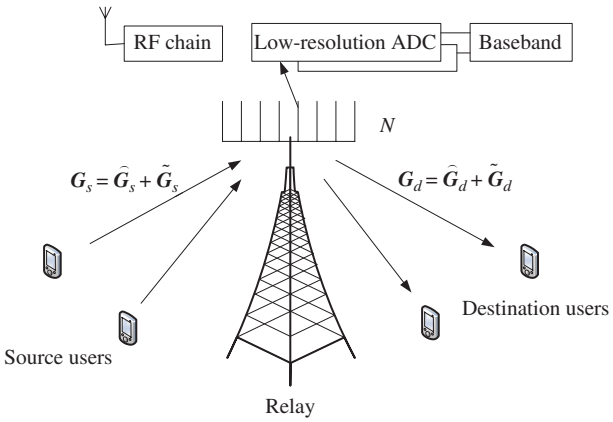

Figure 1 Multi-pair half-duplex relay system with low-resolution ADCs.

systems. To reduce the computational complexity, we also propose an equal user power allocation scheme, which can induce the closed forms of the joint user and relay power allocation.

• To draw more useful insights, some asymptotic results are provided to reveal the impact of system parameters, such as the number of relay antennas and relay transmit power. Based on the asymptotic analysis, the various power scaling schemes are characterized to help analyze the impacts of few bit ADCs on system performance in theory. Particularly, the results show that by selecting certain power scaling scheme, when the number of relay antennas grows without bound, although the effect of low quantization on channel estimation still remains, the effect of low-resolution ADCs on data transmission will disappear.

• Different from the previous literatures, we also focus on the energy-efficient design of low-resolution ADC systems with QoS requirement. The optimal power control strategy is also proposed to minimize the total power consumption while satisfying the user QoS requirement. The simulation results can help to get new insights into the optimal ADC resolution from power saving perspectives, which show that the optimal ADC resolution depends on the QoS requirement and the number of relay antennas.

The rest of this paper is organized as follows. In Section 2, the multi-pair massive MIMO AF relaying system is described. The approximation of ergodic achievable rate is derived and the power control schemes to maximize the sum achievable rate are discussed in Section 3. In Section 4, the power control scheme aiming to minimize the total power consumption is proposed. The simulation results are illustrated in Section 5. Finally, the paper is concluded in Section 6.

Notations: The matrices are denoted by the uppercase boldface letters while the vectors are represented by the lowercase boldface ones. $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$ and $(\cdot)^{\mathrm{H}}$ denote the transpose, conjugate and conjugate transpose of a matrix, respectively. $\|\cdot\|_F^2$ represents the Frobenius norm of a matrix, and $\|\cdot\|^2$ is the Euclidean norm of a vector. $\boldsymbol{I}_N$ is the identity matrix with an $N \times N$ demission. At last, $\mathrm{E}\{\cdot\}$ is the expectation operator.

2 System model

As shown in Figure 1, we consider a multi-pair AF relaying system consisting of K pairs of users and a multiple antenna relay. The kth source user exchanges information with the kth destination user with the help of the relay since no available direct links exist between them. The number of relay antennas is sufficiently larger than that of user pairs, i.e., $N \gg K$. Besides, the relay antennas employ the low-resolution ADCs.

Denote the channel matrix from source users to the relay as $\boldsymbol{G}_s$ with a demission $N \times K$, which can be expressed as

$$\boldsymbol{G}_s = \boldsymbol{H}_s \boldsymbol{D}_s^{1/2}, \tag{1}$$

where $\boldsymbol{G}_s = [\boldsymbol{g}_{s,1}, \boldsymbol{g}_{s,2}, \ldots, \boldsymbol{g}_{s,K}]$ and $\boldsymbol{g}_{s,k}$ denotes the channel vector between the kth source user and relay. $\boldsymbol{H}_s$ is an $N \times K$ matrix representing the small scale fading, and each entry of $\boldsymbol{H}_s$ is the independent and identically distributed (i.i.d.) complex Gaussian random variable with zero mean and unit variance,

i.e., $\mathcal{CN}(0,1)$. The $K \times K$ diagonal matrix $\boldsymbol{D}_s = \mathrm{diag}\{\beta_{s,1},\beta_{s,2},\ldots,\beta_{s,K}\}$ denotes the large scale fading matrix and the entry $\beta_{s,k}$ is the large scale fading between the kth source user and the relay.

Similarly, the channel matrix between the destination users and the relay can be modeled as

$$\boldsymbol{G}_d = \boldsymbol{H}_d \boldsymbol{D}_d^{1/2}, \tag{2}$$

where $\boldsymbol{H}_d$, which is also an $N \times K$ matrix, contains elements following i.i.d. $\mathcal{CN}(0,1)$, and is assumed to be independent of $\boldsymbol{H}_s$. $\boldsymbol{D}_d = \mathrm{diag}\{\beta_{d,1},\beta_{d,2},\ldots,\beta_{d,K}\}$ is the the large scale fading matrix and $\beta_{d,k}$ represents the large scale fading between the kth destination user and the relay. We assume that $\boldsymbol{g}_{d,k}$ is the kth column of $\boldsymbol{G}_d$ denoting the channel vector between the kth destination user and the relay.

2.1 Uplink channel estimation with the impact of ADC

In order to perform beamforming at the relay, the relay needs to know the CSI from the source users to the relay and from the relay to the destination users. Thanks to the channel reciprocity of the time-division-duplexing (TDD) systems, the downlink CSI from the relay to the destination users can be acquired based on the uplink training. Therefore, the CSI matrices $\boldsymbol{G}_s$ and $\boldsymbol{G}_d$ need to be estimated.

During each coherence interval T (symbols), K pairs of users including source and destination users simultaneously send their pilot symbols to the relay [10]. Then, the relay can acquire the CSI based on the MMSE method. The length of the pilot symbols is τ, to avoid the pilot contamination, τ should not be smaller than the number of users, i.e., $\tau \geqslant 2K$. Let $\boldsymbol{G} = [\boldsymbol{G}_s\ \boldsymbol{G}_d]$ denote the channel matrix from all the users to the relay, the received signals at the relay can be described as

$$\boldsymbol{Y}_P = \sqrt{\rho_p}\boldsymbol{G}\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{N}_R, \tag{3}$$

where ρ_p is the transmit power of uplink pilot symbols, and $\boldsymbol{\Phi}$ is the $\tau \times 2K$ pilot matrix. $\boldsymbol{N}_R$ is the additive white Gaussian noise matrix at the relay, which contains i.i.d $\mathcal{CN}(0,1)$ entries. For simplicity, $\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}^{*} = \tau \boldsymbol{I}_{2K}$.

By using the AQNM model, the output signals of the ADCs can be given by [16]

$$\boldsymbol{Y}_P^Q = Q\{\boldsymbol{Y}_P\} = \alpha\sqrt{\rho_p}\boldsymbol{G}\boldsymbol{\Phi}^{\mathrm{T}} + \alpha\boldsymbol{N}_R + \boldsymbol{N}_q, \tag{4}$$

where $Q\{\cdot\}$ is the quantization function, $\alpha = 1-\rho$ where ρ denotes the distortion factor with respect to the ADC resolution. The values of ρ for different quantization bits can be found in [16]. $\boldsymbol{N}_q = \left[\boldsymbol{n}_q^1, \boldsymbol{n}_q^2, \ldots, \boldsymbol{n}_q^\tau\right]$ is the additive Gaussian quantization noise matrix. According to [16], the covariance matrix of $\boldsymbol{n}_q^j$ can be expressed as

$$\boldsymbol{R}_{\boldsymbol{n}_q^j} = \alpha(1-\alpha)\,\mathrm{diag}\left\{\rho_p \boldsymbol{G}\boldsymbol{R}_{\boldsymbol{x}_\phi^j}\boldsymbol{G}^{\mathrm{H}} + \boldsymbol{I}_N\right\}, \tag{5}$$

where $\boldsymbol{x}_\phi^j = [\phi_{j,1},\phi_{j,2},\ldots,\phi_{j,2K}]$ represents the jth pilot symbols vector transmitted by all the users.

As can be seen from (5), the covariance of the quantization noise depends on the uplink pilot sequences, which is different from the unquantized MIMO systems. In this paper, we assume that the pilot sequences of all the users are the different columns of the $\tau \times \tau$ Hadamard matrix, whose element is either 1 or -1. Due to the properties of Hadamard matrix, when $\tau \geqslant 2K$, the pilot sequences of each user can be mutually orthogonal. Therefore, Eq. (5) is derived as

$$\boldsymbol{R}_{\boldsymbol{n}_q^j} = \alpha(1-\alpha)\left[\rho_p \sum_{k=1}^{K}(\beta_{s,k}+\beta_{d,k}) + 1\right]\boldsymbol{I}_N. \tag{6}$$

Denote $\boldsymbol{D} = \mathrm{diag}\{\beta_{s,1},\ldots,\beta_{s,K},\beta_{d,1},\ldots,\beta_{d,K}\}$, then, the estimated channel from all the users to the relay can be given as

$$\hat{\boldsymbol{G}} = \frac{1}{\alpha\tau\sqrt{\rho_p}}\boldsymbol{Y}_P^Q\boldsymbol{\Phi}^{*}\tilde{\boldsymbol{D}} = \boldsymbol{G}\tilde{\boldsymbol{D}} + \frac{1}{\tau\sqrt{\rho_p}}\boldsymbol{N}_R\boldsymbol{\Phi}^{*}\tilde{\boldsymbol{D}} + \frac{1}{\alpha\tau\sqrt{\rho_p}}\boldsymbol{N}_q\boldsymbol{\Phi}^{*}\tilde{\boldsymbol{D}}, \tag{7}$$

where $\tilde{\boldsymbol{D}} = (\boldsymbol{D}^{-1}\frac{\alpha^2+\sigma_q^2}{\alpha^2\tau\rho_p} + \boldsymbol{I}_{2K})^{-1}$, and $\sigma_q^2 = \alpha(1-\alpha)[\rho_p \sum_{k=1}^{K}(\beta_{s,k}+\beta_{d,k})+1]$. Due to the orthogonal property of MMSE based channel estimation, $\boldsymbol{g}_{s,k}$ and $\boldsymbol{g}_{d,k}$ are given as [28]

$$\boldsymbol{g}_{s,k} = \hat{\boldsymbol{g}}_{s,k} + \tilde{\boldsymbol{g}}_{s,k}, \tag{8}$$

$$\boldsymbol{g}_{d,k} = \hat{\boldsymbol{g}}_{d,k} + \tilde{\boldsymbol{g}}_{d,k}, \tag{9}$$

where $\hat{\boldsymbol{g}}_{s,k} \sim \mathcal{CN}(0, \sigma_{s,k}^2\boldsymbol{I}_N)$, $\tilde{\boldsymbol{g}}_{s,k} \sim \mathcal{CN}(0, e_{s,k}^2\boldsymbol{I}_N)$, $\hat{\boldsymbol{g}}_{d,k} \sim \mathcal{CN}(0, \sigma_{d,k}^2\boldsymbol{I}_N)$, and $\tilde{\boldsymbol{g}}_{d,k} \sim \mathcal{CN}(0, e_{d,k}^2\boldsymbol{I}_N)$. Since $\hat{\boldsymbol{g}}_{s,k}$, $\tilde{\boldsymbol{g}}_{s,k}$, $\hat{\boldsymbol{g}}_{d,k}$ and $\tilde{\boldsymbol{g}}_{d,k}$ are mutually independent, according to (7)–(9), we have

$$\sigma_{s,k}^2 = \frac{\alpha^2\tau\rho_p\beta_{s,k}^2}{\alpha^2\tau\rho_p\beta_{s,k}+\alpha^2+\sigma_q^2}, \quad e_{s,k}^2 = \frac{\left(\alpha^2+\sigma_q^2\right)\beta_{s,k}}{\alpha^2\tau\rho_p\beta_{s,k}+\alpha^2+\sigma_q^2}, \tag{10}$$

$$\sigma_{d,k}^2 = \frac{\alpha^2\tau\rho_p\beta_{d,k}^2}{\alpha^2\tau\rho_p\beta_{d,k}+\alpha^2+\sigma_q^2}, \quad e_{d,k}^2 = \frac{\left(\alpha^2+\sigma_q^2\right)\beta_{d,k}}{\alpha^2\tau\rho_p\beta_{d,k}+\alpha^2+\sigma_q^2}. \tag{11}$$

Then, the relay will send the estimated CSI from relay to the destination users as well as the CSI from sources users to the relay to the destination users through the feedback channel. In practice, to utilize the bandwidth more efficiently and save more relay power, there are some other mechanisms for the destination users in AF relaying systems to obtain the CSI [29,30], such as using the downlink pilots. In this paper, we make an assumption that the destination users can acquire the estimated CSI, which is similar as some priori works [9,11,14,15].

2.2 Data transmission

During the data transmission phase, in the first time slot, all the source users send their signals to the relay and the signals received by the relay are expressed as

$$\boldsymbol{y}_R = \boldsymbol{G}_s\boldsymbol{x} + \boldsymbol{n}_R, \tag{12}$$

where $\boldsymbol{x} = [\sqrt{p_{s,1}}x_1, \sqrt{p_{s,2}}x_2, \ldots, \sqrt{p_{s,K}}x_K]$ denotes the signal vector transmitted by the source users. We assume that $\mathrm{E}\left\{\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\right\} = \boldsymbol{P} = \mathrm{diag}\left\{p_{s,1}, p_{s,2}, \ldots, p_{s,K}\right\}$, where $p_{s,k}$ denotes the transmit power of the kth source user. $\boldsymbol{n}_R$ represents the relay noise and $\boldsymbol{n}_R \sim \mathcal{CN}(0, \boldsymbol{I}_N)$.

Since the relay employs the low-resolution ADCs, the output signals of quantizer are expressed as [16]

$$\boldsymbol{y}_R^q = \alpha\boldsymbol{G}_s\boldsymbol{x} + \alpha\boldsymbol{n}_R + \boldsymbol{n}_q, \tag{13}$$

where $\boldsymbol{n}_q$ is the additive Gaussian quantization noise vector, whose covariance matrix can be written as

$$\boldsymbol{R}_{\boldsymbol{n}_q} = \alpha\left(1-\alpha\right)\mathrm{diag}\left\{\boldsymbol{G}_s\boldsymbol{P}\boldsymbol{G}_s^{\mathrm{H}} + \boldsymbol{I}_N\right\}. \tag{14}$$

In the second time slot, the relay amplifies the quantized signals by the beamforming matrix $\boldsymbol{W}_R$ and then broadcast them to the destination users. Hence, the received signals by the kth destination user can be written as

$$y_k = \alpha\boldsymbol{g}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\boldsymbol{G}_s\boldsymbol{x} + \alpha\boldsymbol{g}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\boldsymbol{n}_R + \boldsymbol{g}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\mathbf{n}_q + n_k, \tag{15}$$

where n_k denotes the additive Gaussian white noise at the kth destination user, and $n_k \sim \mathcal{CN}(0,1)$. With the uplink channel estimation errors, by substituting (8) and (9) into (15), y_k can be rewritten as

$$\begin{aligned} y_k = & \underbrace{\alpha\sqrt{p_{s,k}}\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\hat{\boldsymbol{g}}_{s,k}x_k}_{\text{desierd signals}} + \underbrace{\sum_{i=1,i\neq k}^{K}\alpha\sqrt{p_{s,i}}\boldsymbol{g}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\boldsymbol{g}_{s,i}x_i}_{\text{multi-user interference}} \\ & + \underbrace{\alpha\sqrt{p_{s,k}}\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\tilde{\boldsymbol{g}}_{s,k}x_k + \alpha\alpha\sqrt{p_{s,k}}\tilde{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\boldsymbol{g}_{s,k}x_k}_{\text{channel estimation errors}} + \underbrace{\alpha\boldsymbol{g}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\boldsymbol{n}_R}_{\text{noise from relay}} + \underbrace{\boldsymbol{g}_{d,k}^{\mathrm{H}}\boldsymbol{W}_R\boldsymbol{n}_q}_{\text{quantization noise}} + \underbrace{n_k}_{\text{noise at the user}}. \end{aligned} \tag{16}$$

In this paper, the relay employs the MRC/MRT beamforming matrix, which is easy for the practical implementation. Based on the MRC/MRT criterion, $\boldsymbol{W}_R$ is written as

$$\boldsymbol{W}_R = \lambda \hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}}, \tag{17}$$

where λ denotes the relay power normalization factor. To satisfy the relay power constraint P_R, we have

$$\lambda = \sqrt{\frac{P_R}{\alpha^2 \mathrm{E}\{\|\hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}} \boldsymbol{G}_s \boldsymbol{x}\|_F^2\} + \alpha^2 \mathrm{E}\{\|\hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}}\|_F^2\} + \mathrm{E}\{\|\hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}} \boldsymbol{n}_q\|_F^2\}}}. \tag{18}$$

Then, substituting (17) and (18) into (16), the signal-to-interference and noise ratio (SINR) of the kth destination user can be given as

$$\mathrm{SINR}_k = \frac{S_k}{I_{1,k} + I_{2,k} + I_{3,k} + I_{4,k} + I_{5,k}}, \tag{19}$$

where

$$\begin{aligned}
&S_k = \alpha^2 p_{s,k} \|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}} \hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}} \hat{\boldsymbol{g}}_{s,k}\|^2, \quad I_{1,k} = \alpha^2 \sum_{i=1, i\neq k}^{K} p_{s,i} \|\boldsymbol{g}_{d,k}^{\mathrm{H}} \hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}} \boldsymbol{g}_{s,i}\|^2, \\
&I_{2,k} = \alpha^2 p_{s,k} \|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}} \hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}} \hat{\boldsymbol{g}}_{s,k}\|^2 + \alpha^2 p_{s,k} \|\tilde{\boldsymbol{g}}_{d,k}^{\mathrm{H}} \hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}} \boldsymbol{g}_{s,k}\|^2, \\
&I_{3,k} = \alpha^2 \|\boldsymbol{g}_{d,k}^{\mathrm{H}} \hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}}\|^2, \quad I_{4,k} = \|\boldsymbol{g}_{d,k}^{\mathrm{H}} \hat{\boldsymbol{G}}_d \hat{\boldsymbol{G}}_s^{\mathrm{H}} \boldsymbol{n}_q\|^2, \quad I_{5,k} = \frac{1}{\lambda^2}.
\end{aligned}$$

Considering the worst-case that the additive noise is uncorrelated Gaussian, the achievable rate of the kth destination user can be expressed as

$$R_k = \frac{T-\tau}{2T} \log_2 \left(1 + \mathrm{SINR}_k\right), \tag{20}$$

where the factor $1/2$ is because that two time slots are utilized for the data transmission.

3 Performance analysis

In this section, a closed form approximation of the ergodic achievable rate is derived. To draw more conclusions on the effects of system parameters on the system performance, some asymptotic results are also obtained.

3.1 Ergodic achievable rate

In this part, the closed form approximation of the ergodic achievable rate is derived. According to [31], the ergodic achievable rate of the kth destination user can be approximated as

$$\begin{aligned}
\bar{R}_k &= \mathrm{E}\left\{\frac{T-\tau}{2T} \log\left(1 + \mathrm{SINR}_k\right)\right\} \\
&\approx \frac{T-\tau}{2T} \log\left\{1 + \frac{\mathrm{E}\left\{S_k\right\}}{\mathrm{E}\left\{I_{1,k} + I_{2,k} + I_{3,k} + I_{4,k} + I_{5,k}\right\}}\right\} \\
&= \frac{T-\tau}{2T} \log\left\{1 + \frac{\bar{S}_k}{\bar{I}_{1,k} + \bar{I}_{2,k} + \bar{I}_{3,k} + \bar{I}_{4,k} + \bar{I}_{5,k}}\right\}.
\end{aligned} \tag{21}$$

By applying random matrix theory, the closed form expression of (21) will be derived, which is shown in the following theorem.

Theorem 1. For a multi-pair massive MIMO AF relaying system with imperfect CSI and low-resolution ADCs, when the number of relay antennas is sufficiently large, the ergodic achievable rate of the kth destination user can be approximated as

$$\bar{R}_k \approx \frac{T-\tau}{2T}\log\left(1+\frac{a_k p_{s,k} P_R}{\sum_{i=1,i\neq k}^{K} b_{k,i} p_{s,i} P_R + c_k p_{s,k} P_R + \sum_{i=1}^{K} d_i p_{s,i} + e_k P_R + f_k}\right), \tag{22}$$

where

$$a_k = \alpha\left[\left(N^2+2N\right)\sigma_{s,k}^4\sigma_{d,k}^4 + \sigma_{s,k}^2\sigma_{d,k}^2\sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{d,j}^2\right], \tag{23}$$

$$\begin{aligned} b_{k,i} &= \alpha\left(N\beta_{s,i}\sigma_{s,k}^2\sigma_{d,k}^4 + N\sigma_{s,i}^4\sigma_{d,i}^2\beta_{d,k} + \beta_{d,k}\beta_{s,i}\sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{s,j}^2\right) \\ &\quad + (1-\alpha)\,\beta_{s,i}\left[N\sigma_{s,k}^2\sigma_{d,k}^4 + \sigma_{s,i}^4\sigma_{d,i}^2\beta_{d,k} + \beta_{d,k}\sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{d,j}^2\right], \end{aligned} \tag{24}$$

$$\begin{aligned} c_k &= \alpha\left[Ne_{s,k}^2\sigma_{s,k}^2\sigma_{d,k}^4 + N\sigma_{s,k}^4\sigma_{d,k}^2e_{d,k}^2 + \left(e_{s,k}^2\sigma_{d,k}^2 + \beta_{s,k}e_{d,k}^2\right)\sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{d,j}^2\right] \\ &\quad + (1-\alpha)\left[N\sigma_{s,k}^4\sigma_{d,k}^4 + N\sigma_{s,k}^2\sigma_{d,k}^4\beta_{s,k} + \beta_{d,k}\sigma_{s,k}^4\sigma_{d,k}^2 + \beta_{s,k}\beta_{d,k}\sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{d,j}^2\right], \end{aligned} \tag{25}$$

$$d_i = \alpha N\sigma_{s,i}^4\sigma_{d,i}^2 + (1-\alpha)\,\sigma_{s,i}^4\sigma_{d,i}^2 + \beta_{s,i}\sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{d,j}^2, \tag{26}$$

$$e_k = N\sigma_{s,k}^2\sigma_{d,k}^4 + \beta_{d,k}\sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{d,j}^2, \tag{27}$$

$$f_k = \sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{d,j}^2. \tag{28}$$

Proof. See Appendix A.

As shown in Theorem 1, the ergodic achievable rate of the kth destination user is a function with respect to the transmit power of each source user and relay, the number of relay antennas, the ADC resolution as well as the CSI.

3.2 Asymptotic analysis

Since Theorem 1 reveals the performance of a general case where imperfect CSI and low precision ADCs are considered, some special cases can be obtained.

Remark 1. When the number of ADC bits tends to infinite, i.e., $b \to \infty$, the relay is equipped with full-precision ADCs. By substituting $\alpha = 1$ into (22)–(28), the performance of the relaying system with full-precision ADCs is derived. In addition, the imperfect CSI is still considered.

Remark 2. Let $\sigma_{s,k}^2 = \beta_{s,k}, e_{s,k}^2 = 0, \sigma_{d,k}^2 = \beta_{d,k}, e_{d,k}^2 = 0$, the perfect CSI can be acquired at the relay. In this case, only the impact of low-resolution ADCs on data transmission is taken into account for the performance analysis.

Remark 3. Let $\sigma_q^2 = 0$, the variance of estimated channels can be derived as $\sigma_{s,k}^2 = \frac{\tau\rho_p\beta_{s,k}^2}{\tau\rho_p\beta_{s,k}+1}$, $e_{s,k}^2 = \frac{\beta_{s,k}}{\tau\rho_p\beta_{s,k}+1}$, $\sigma_{d,k}^2 = \frac{\tau\rho_p\beta_{d,k}^2}{\tau\rho_p\beta_{d,k}+1}$, $e_{d,k}^2 = \frac{\beta_{d,k}}{\tau\rho_p\beta_{d,k}+1}$, which are the same as the results of the MMSE based channel estimation of the unquantized MIMO systems. In this case, the impact of low-resolution ADCs on channel estimation is ignored, which is not practical.

To get more useful insights, the following asymptotic results, which can reveal the effects of the system parameters on the system performance, can also be derived based on Theorem 1.

Proposition 1. When $N \to \infty$, the asymptotic achievable rate of the kth destination user is given as

$$\bar{R}_k \to \frac{T-\tau}{2T} \log\left(1 + \frac{\alpha p_{s,k} P_R N \sigma_{s,k}^4 \sigma_{d,k}^4}{\sum_{i=1,i\neq k}^{K} p_{s,i} P_R \eta_{k,i} + p_{s,k} P_R \psi_k + \sum_{i=1}^{K} p_{s,i} \xi_{k,i} + P_R \zeta_k}\right), \tag{29}$$

where

$$\eta_{k,i} = \alpha \beta_{s,i}^2 \sigma_{s,k}^2 \sigma_{d,k}^4 + \alpha \sigma_{s,i}^4 \sigma_{d,i}^2 \beta_{d,k}^2, \tag{30}$$

$$\psi_k = \alpha \left(e_{s,k}^2 \sigma_{s,k}^2 \sigma_{d,k}^4 + \sigma_{s,k}^4 \sigma_{d,k}^2 e_{d,k}^2\right) + (1-\alpha)\left(\sigma_{s,k}^4 \sigma_{d,k}^4 + \sigma_{s,k}^2 \sigma_{d,k}^4 \beta_{s,k}\right), \tag{31}$$

$$\xi_{k,i} = \alpha \sigma_{s,i}^4 \sigma_{d,i}^2, \tag{32}$$

$$\zeta_k = \sigma_{s,k}^2 \sigma_{d,k}^4. \tag{33}$$

Proposition 1 indicates that when the number of relay antennas grows without bound, the achievable rate of the kth destination user will tend to infinity. Therefore, the achievable rate loss introduced by the low quantization can be compensated when the relay employs more antennas.

Proposition 2. When N is fixed, if $p_{s,1} \to \infty, \ldots, p_{s,K} \to \infty$, and $P_R \to \infty$, the asymptotic achievable rate of the kth destination user can be derived as

$$\bar{R}_k \to \frac{T-\tau}{2T} \log\left(1 + \frac{\alpha N \sigma_{s,k}^4 \sigma_{d,k}^4}{\sum_{i=1,i\neq k}^{K} \eta_{k,i} + \psi_k}\right), \tag{34}$$

where $\eta_{k,i}$ and ψ_k are the same as (30) and (31), respectively.

It can be observed from Proposition 2 that the ergodic rate of kth destination user will trend to be a constant by increasing the power of each source user and relay. However, the fixed value still be limited by the number of ADC bits. Therefore, the rate loss introduced by the low quantization cannot be compensated by increasing the transmit power of each user and relay.

Proposition 3. Let $p_{s,1} = \frac{E_{s,1}}{N^p}$, $p_{s,2} = \frac{E_{s,2}}{N^p}, \ldots, p_{s,K} = \frac{E_{s,K}}{N^p}$ and $P_R = \frac{E_R}{N^q}$, when $0 < p, q \leqslant 1$, for fixed $E_{s,1}$, $E_{s,2}, \ldots, E_{s,K}$, E_R, when $N \to \infty$, the asymptotic achievable rate of the kth destination user is derived as follows

- Case 1: if $p = 1$ and $q = 1$,

$$\bar{R}_k \to \frac{T-\tau}{2T} \log\left(1 + \frac{\alpha E_{s,k} E_R \sigma_{s,k}^4 \sigma_{d,k}^4}{E_R \sigma_{s,k}^2 \sigma_{d,k}^4 + \alpha \sum_{i=1}^{K} E_{s,i} \sigma_{s,i}^4 \sigma_{d,i}^2 + \sum_{i=1}^{K} \sigma_{s,i}^2 \sigma_{d,i}^2}\right). \tag{35}$$

- Case 2: if $p = 1$ and $q = 0$,

$$\bar{R}_k \to \frac{T-\tau}{2T} \log\left(1 + \alpha E_{s,k} \sigma_{s,k}^2\right). \tag{36}$$

- Case 3: if $p = 0$ and $q = 1$,

$$\bar{R}_k \to \frac{T-\tau}{2T} \log\left(1 + \frac{E_{s,k} E_R \sigma_{s,k}^4 \sigma_{d,k}^4}{\sum_{i=1}^{K} E_{s,i} \sigma_{s,i}^4 \sigma_{d,i}^2}\right). \tag{37}$$

- Case 4: if $0 < p < 1$ and $0 < q < 1$,

$$\bar{R}_k \to \infty. \tag{38}$$

Proposition 3 presents the generalized power scaling law of the quantized MIMO relaying systems with imperfect CSI. When $N \to \infty$, the achievable rate will tend to a constant, which depends on the various scaling parameters, p and q. Some conclusions can be drawn from Proposition 3. Case 1 shows that the transmit power of each user and relay can be both scaled down proportionally to $\frac{1}{N}$ to maintain the fixed achievable rate. Case 2 indicates that when P_R is fixed and $p_{s,k}$ is scaled down proportionally to $\frac{1}{N}$, the

multi-user interference, the noise at the relay or destination users as well as the channel estimation errors will disappear. Case 3 shows that the effect of low-resoluation ADCs on data transmission can disappear when P_R is scaled down proportionally to $\frac{1}{N}$ and $p_{s,k}$ is fixed. However, during the channel estimation phase, the pilot symbols are also quantized by the ADCs. Therefore, when the CSI is imperfect, the impact of few-bit ADCs on channel estimation still remains.

3.3 Power control to maximize the sum achievable rate

Since the transmit power of each source user and relay is different, the joint user and relay power control scheme can be proposed to maximize the system sum achievable rate with the individual transmit power constraint, which can enhance the sum rate of the quantized MIMO systems.

Based on (22), the optimization problem, whose objective is to maximize the system sum achievable rate, is formulated as

$$\max_{p_{s,k},p_R} \sum_{k=1}^{K} \bar{R}_k, \tag{P1}$$

$$\text{s.t.} \quad 0 < p_{s,k} \leqslant P_{s,k}^{\max}, \forall k, \tag{39}$$

$$0 < P_R \leqslant P_R^{\max}, \tag{40}$$

where constraints (39) and (40) represent the individual transmit power constraint of each source user and relay, respectively.

The optimization problem (P1) is the non-convex problem, and the global optimum is difficult to obtain. According to the property of logarithmic function, Eq. (P1) can be transformed as

$$\min_{p_{s,k},P_R,\gamma_k} \prod_{k=1}^{K} (1+\gamma_k)^{-1}, \tag{P2}$$

$$\text{s.t. } \gamma_k \leqslant \frac{a_k p_{s,k}}{\sum_{i=1,i\neq k}^{K} b_{k,i} p_{s,i} + c_k p_{s,k} + \sum_{i=1}^{K} d_i p_{s,i} P_R^{-1} + e_k + f_k P_R^{-1}}, \tag{41}$$

$$0 < p_{s,k} \leqslant P_{s,k}^{\max}, \forall k, \tag{42}$$

$$0 < P_R \leqslant P_R^{\max}. \tag{43}$$

Since the constrains (41)–(43) are polynomial or monomial functions, if (P2) can be further transformed into polynomial function, the optimization problem can be regarded as the geometric programming (GP). Generally, GP can be transformed into a convex form and solved by the optimization softwares, such as CVX, efficiently [32]. Fortunately, according to [31], $1+\gamma_k$ can be approximated near a a point $\hat{\gamma}_k$ by $\xi_k \gamma_k^{\zeta_k}$, where $\zeta_k = \hat{\gamma}_k (1+\hat{\gamma}_k)^{-1}$ and $\xi_k = \hat{\gamma}_k^{-\zeta_k}(1+\hat{\gamma}_k)$. Hence, the objective function will be transformed as the monomial function, $\min_{p_{s,k},P_R,\gamma_k} \prod_{k=1}^{K} (\xi_k \gamma_k^{\zeta_k})^{-1}$. Since the primary optimization problem can be approximated as GP, the Algorithm 1 can be proposed to search the closely local optimum of (P1).

The value of μ in Algorithm 1 can be chosen to strike a comprise between convergence speed and the accuracy of approximation [33].

To reduce the computational complexity, we also propose an equal user power allocation algorithm, where we assume that $p_{s,1} = p_{s,2} = \cdots = p_{s,K} = p_s$ and $P_R = K p_s$, then the original optimization problem can be transformed as

$$\max_{p_s} \sum_{k=1}^{K} \bar{R}_k, \tag{P3}$$

$$\text{s.t.} \quad 0 < p_s \leqslant \min\left(P_{s,1}^{\max}, P_{s,2}^{\max}, \ldots, P_{s,K}^{\max}, \frac{P_R^{\max}}{K}\right), \tag{44}$$

According to the property of logarithmic function, the original objective function of (P3) can be rewritten as

$$\max_{p_s} \prod_{k=1}^{K} \delta_k, \tag{P4}$$

Algorithm 1 Power control scheme to maximize the sum achievable rate

1: **Initialize** $\hat{\gamma}_k^{(0)}$ for $k = 1, 2, \ldots, K$, the convergence conditions ε and the maximal iterative steps θ.

2: **Repeat:**

3: Let $t = 1$, and calculating

$$\zeta_k^{(t)} = \hat{\gamma}_k^{(t-1)} \left(1 + \hat{\gamma}_k^{(t-1)}\right)^{-1}$$

and

$$\xi_k^{(t)} = \left[\hat{\gamma}_k^{(t-1)}\right]^{-\zeta_k^{(t)}} \left(1 + \hat{\gamma}_k^{(t-1)}\right).$$

4: Solving the following optimization problem, and obtain the optimal power $p_{s,k}^{(t)}$ and $P_R^{(t)}$.

$$\min_{p_{s,k}, P_R, \hat{\gamma}_k^{(t)}} \prod_k \left[\xi_k^{(t)} \left(\hat{\gamma}_k^{(t)}\right)^{\zeta_k^{(t)}}\right]^{-1},$$
$$\text{s.t. } \mu^{-1} \hat{\gamma}_k^{(t-1)} \leqslant \hat{\gamma}_k^{(t)} \leqslant \mu \hat{\gamma}_k^{(t-1)}, \ (41), \ (42), \ (43).$$

5: Update $\hat{\gamma}_k^{(t)}$ based on the solutions and calculate $\max\limits_k |\hat{\gamma}_k^{(t)} - \hat{\gamma}_k^{(t-1)}|$.

6: **Until** $\max_k |\hat{\gamma}_k^{(t)} - \hat{\gamma}_k^{(t-1)}| < \varepsilon$ **or** $t = \theta$.

where

$$\delta_k = \frac{(a_k K + \sum_{i=1, i\neq k}^{K} b_{k,i} K + c_k K) p_s^2 + (\sum_{i=1}^{K} d_i + e_k K) p_s + f_k}{(\sum_{i=1, i\neq k}^{K} b_{k,i} K + c_k K) p_s^2 + (\sum_{i=1}^{K} d_i + e_k K) p_s + f_k}. \tag{45}$$

Since the first order derivative of δ_k with respect to p_s, $\frac{\partial \delta_k}{\partial p_s} > 0$, δ_k is a increasing function of p_s. As $\delta_k > 0$ and $\frac{\partial \delta_k}{\partial p_s} >$ for arbitrary k, $\partial \prod_{k=1}^{K} \delta_k / \partial p_s > 0$. Therefore, $\prod_{k=1}^{K} \delta_k$ is also a increasing function with respect to p_s. Therefore, when $p_{s,1} = p_{s,2} = \cdots = p_{s,K} = p_s$ and $P_R = K p_s$, the optimal power allocation can be derived as

$$p_s^* = \min\left(P_{s,1}^{\max}, P_{s,2}^{\max}, \ldots, P_{s,K}^{\max}, \frac{P_R^{\max}}{K}\right), \tag{46}$$

and

$$P_R^* = K \min\left(P_{s,1}^{\max}, P_{s,2}^{\max}, \ldots, P_{s,K}^{\max}, \frac{P_R^{\max}}{K}\right). \tag{47}$$

Since we can know the closed forms of p_s^* and P_R^* for the equal user power allocation algorithm, the computational complexity can be remarkably reduced.

4 Total power consumption minimization

The low-resolution ADCs can reduce the consumed power of the receiver at the cost of degrading the system achievable rate. Meanwhile, the transmit power of each user and relay will increase with the decreasing number of quantization bits to maintain the same achievable rate. To achieve high energy-efficient design without sacrificing the system achievable rate, considering the transmit power consumption and hardware cost, the optimization problem to minimize the total power consumption while satisfying the QoS requirement of each user is formulated and solved. Similar as [19], the circuit power consumption of the relay receiver is modeled as

$$P_{\text{CIRCUIT}} = N\left(2P_{\text{ADC}} + P_{\text{LNA}} + P_{\text{MIXER}} + P_{\text{FILTERS}}\right) + P_{\text{SYN}}, \tag{48}$$

where P_{LNA}, P_{MIXER}, P_{FILTERS}, P_{SYN} denote the consumed power of low noise amplifier, mixer, filters as well as frequency synthesizer, respectively. The power consumption of ADC, P_{ADC}, grows linearly

with the sampling rate and increases exponentially with the number of ADC bits, which can be modeled as [34]

$$P_{\mathrm{ADC}} = \mathrm{FOM} \cdot f_{\mathrm{sampling}} \cdot 2^b, \tag{49}$$

where FOM denotes figure of merit, which describes the power efficiency of ADCs and is measured in fj/conversion-step. When the Nyquist sampling is applied, $f_{\mathrm{sampling}} = 2f_{\mathrm{signal}}$, where f_{signal} is the signal bandwidth. Hence, the total power consumption can be modeled as

$$P_{\mathrm{total}} = \sum_{k=1}^{K} \frac{p_{s,k}}{\eta_{s,k}} + \frac{P_R}{\eta_R} + N\left(2 \cdot \mathrm{FOM} \cdot 2f_{\mathrm{signal}} \cdot 2^b + P_{\mathrm{LNA}} + P_{\mathrm{MIXER}} + P_{\mathrm{FILTERS}}\right) + P_{\mathrm{SYN}}, \tag{50}$$

where $\eta_{s,k}$ and η_R are the power amplifier efficiencies of the kth source user and relay, respectively. Then, the optimization problem is formulated as

$$\min_{p_{s,k}, P_R} P_{\mathrm{total}}, \tag{P5}$$

$$\text{s.t. } \frac{T-\tau}{2T} \log\left(1 + \frac{a_k p_{s,k} P_R}{\sum_{i=1, i\neq k}^{K} b_{k,i} p_{s,i} P_R + c_k p_{s,k} P_R + \sum_{i=1}^{K} d_i p_{s,i} + e_k P_R + f_k}\right) \geqslant r_{s,k}, \forall k, \tag{51}$$

$$\text{s.t.} \quad 0 < p_{s,k} \leqslant P_{s,k}^{\max}, \forall k, \tag{52}$$

$$0 < P_R \leqslant P_R^{\max}, \tag{53}$$

where $r_{s,k}$ denotes the QoS requirement of the kth destination user. By some basic manipulations, Eq. (51) is transformed as

$$\sum_{i=1, i\neq k}^{K} b_{k,i} p_{s,i} (p_{s,k})^{-1} + \sum_{i=1}^{K} d_i p_{s,i} (p_{s,k})^{-1} (P_R)^{-1} + e_k (p_{s,k})^{-1} + f_k (p_{s,k})^{-1} (P_R)^{-1} \leqslant g_k, \tag{54}$$

where $g_k = (2^{\frac{2Tr_{s,k}}{T-\tau}} - 1)^{-1} \cdot a_k - c_k$. Therefore, the original optimization problem can be regarded as the GP and can be solved by the optimization softwares efficiently.

5 Simulation results

In this section, the Monte Carlo simulations are presented. We consider a circle with a radius of 500 m, where a relay is located at the center. The K user pairs are randomly located in the circle. The large scale fading coefficients are modeled as $\beta_{s,k} = z_k/(r_{s,k}/r_0)^v$ and $\beta_{d,k} = z_k/(r_{d,k}/r_0)^v$ [31], respectively, where $r_{s,k}(r_{d,k})$ is the distance between the kth source (destination) user and the relay, v denotes the path loss exponent and z_k represents the log-normal random variable with the standard deviation 8 dB. We assume that $r_0 = 100$ m and $v = 3.8$ [26]. The coherence interval is assumed to be $T = 200$ (symbols) and the pilot length τ depends on the number of user pairs and will vary in different scenarios. The uplink pilot power ρ_p is assumed to be 10 dB.

5.1 Performance analysis

In this section, the derived approximation of ergodic achievable rate of destination users will be evaluated and compared to the simulation results. We assume that $K = 4$, and $\tau = 8$. The transmit power of uplink pilot and all the source users is equal to be 10 dB, i.e., $\rho_p = p_{s,1} = p_{s,2} = \cdots = p_{s,K}$. We choose $\beta_{s,k} = [0.0287, 0.003, 0.0199, 0.0456]$ and $\beta_{d,k} = [0.1187, 0.0854, 0.343, 0.0225]$.

Figure 2 describes the sum ergodic achievable rate with different numbers of relay antennas and quantization bits. As shown in Figure 2, the derived closed form approximation (22) is tight compared to the simulation results. Therefore, the accuracy of our theoretical analysis is demonstrated. The sum rate increases with the growing number of relay antennas and the higher ADC resolution. In addition, the

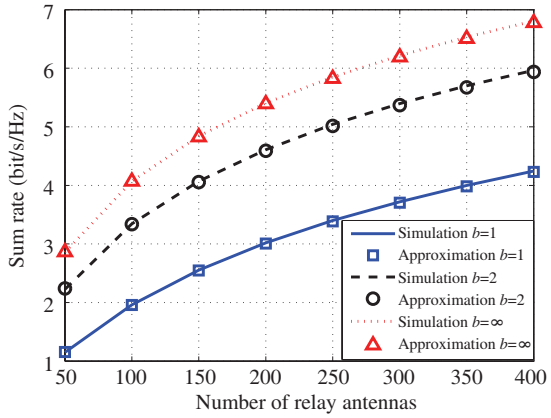

Figure 2 (Color online) Sum achievable rate versus the number of relay antennas with $\rho_p = p_{s,k} = 10$ dB and $P_R = Kp_{s,k}$.

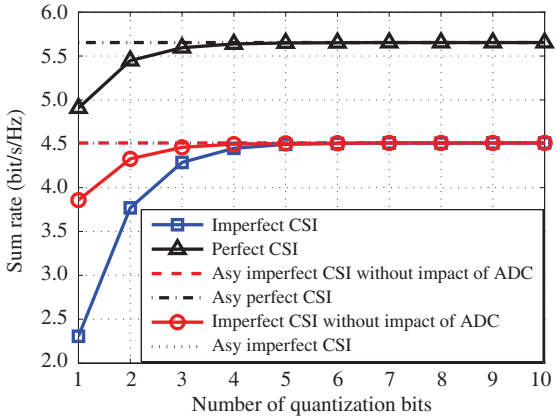

Figure 3 (Color online) Sum achievable rate versus the number of quantization bits with $\rho_p = p_{s,k} = 10$ dB, $P_R = Kp_{s,k}$ and $N = 128$.

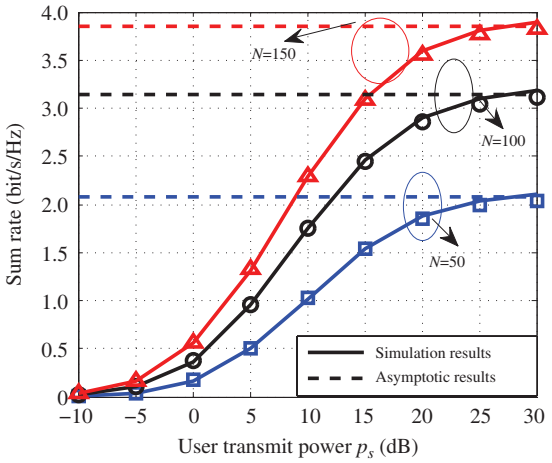

Figure 4 (Color online) Sum achievable rate versus the transmit power of source users and relay with $p_{s,k} = P_R$ and $b = 1$. The markers are the approximated results.

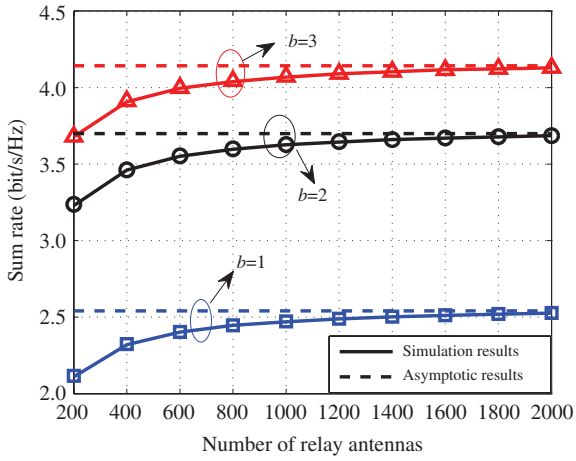

Figure 5 (Color online) Sum achievable rate versus the number of relay antennas with $E_{s,k} = E_R = 30$ dB, $p_{s,k} = E_{s,k}/N$, $P_R = E_R/N$. The markers are the approximated results.

sum rate increases significantly from $b = 1$ to $b = 2$. However, the gap between $b = 2$ and $b \to \infty$ becomes narrow.

Figure 3 compares the sum achievable rate for perfect and imperfect CSI cases with various numbers of quantization bits. The "asy" means the asymptotic results when $b \to \infty$. The sum achievable rate of the case where the CSI is estimated without considering the impact of quantization is also presented. As can be seen from Figure 3, all the curves convergence to fixed values with the increasing ADC resolution, and the coarse quantization (4–5 bits) can achieve the comparable performance to the full-precision ADCs. Furthermore, the few-bit ADCs degrade the accuracy of channel estimation, and consequently affects the sum rate, especially when the resolution is very low (1–2 bits). Particularly, the achievable rate decreases almost 40% and 13% for $b = 1$ and $b = 2$, respectively.

Figure 4 describes the relationship between sum rate and the transmit power of each source user and relay. The asymptotic results for different numbers of relay antennas are derived according to (34). As shown in Figure 4, by increasing the transmit power of users and relay, the sum rate grows rapidly first and then slowly tends to a constant, which agrees with Proposition 2. Besides, Figure 4 also validates the tightness of our derived closed form approximation. Figure 5 shows the power scaling scheme for $p = q = 1$ shown in Proposition 3. For fixed $E_{s,k}, E_R$, when the number of relay antennas tend to infinity, the sum rate tends to a fixe value, which depends on the ADC resolution. Figure 6 illustrates case 3

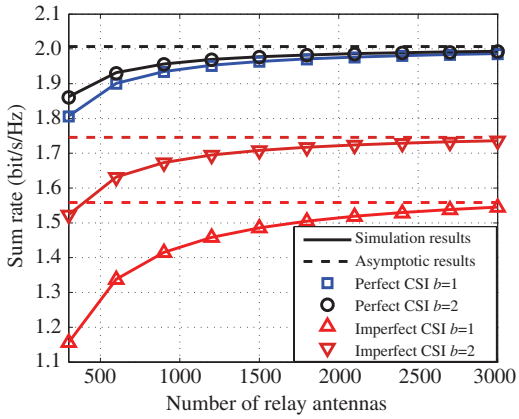

Figure 6 (Color online) Sum achievable rate versus the different numbers of relay antennas with $E_{s,k} = 10$ dB, $E_R = 15$ dB, $p_{s,k} = E_{s,k}$, $P_R = E_R/N$.

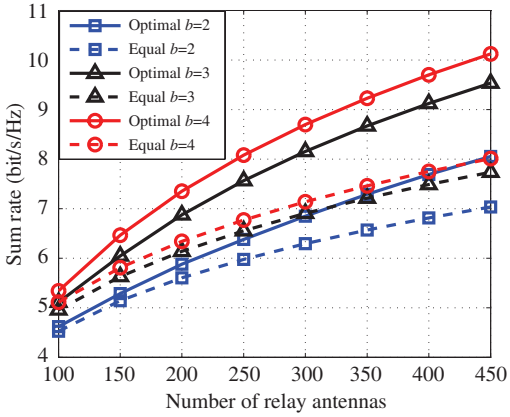

Figure 7 (Color online) Sum achievable rate versus different numbers of relay antennas and power control schemes.

shown in Proposition 3, where $p = 0$ and $q = 1$. As shown in Figure 6, when the relay can acquire the perfect CSI, the curves for $b = 1$ and $b = 2$ cases can convergence to the same constant, which indicates that the effect of few-bit ADC on data transmission will disappear when $N \to \infty$. However, for the imperfect CSI case, the curves for $b = 1$ and $b = 2$ still tends to the different values. This is because that the ADC resolution will affect the accuracy of channel estimation and the impact cannot be eliminated.

5.2 Optimal power control

In this section, the proposed power control schemes are illustrated. We assume $K = 6$, $\tau = 12$. The large scale fading coefficients are chosen as $\beta_{s,k} = [0.2761, 0.0216, 0.5390, 0.0573, 0.0103, 0.0365]$ and $\beta_{d,k} = [0.3277, 0.5051, 0.1048, 0.1360, 0.2512, 0.0472]$, respectively. The maximum transmit power of each source user and relay is 10 dB and 20 dB, respectively.

5.2.1 *Power control scheme to maximize the sum achievable rate*

Figure 7 shows the sum achievable rate with different power control schemes. As shown in Figure 7, the proposed power control algorithm outperforms the equal user power allocation one and the gap becomes more remarkable when the number of relay antennas grows. Based on the optimal power control scheme, when $b = 3$, the sum rate is larger than that of $b = 4$ based on equal power allocation. Therefore, the optimal power control schemes plays a critical role in the quantized systems.

Figure 8 illustrates the convergence of the proposed iterative algorithm with different initial points. We set the different initial transmit power $p_{s,k}^0$ and P_R^0 to get the different values of $\hat{\gamma}_k^{(0)}$. For simplicity, we assume that the initial values of transmit power $p_{s,1}^0 = p_{s,2}^0 = \cdots = p_{s,K}^0$. As can be found in Figure 8, although the initial points are different, the proposed algorithm will achieve the same sum rate, which demonstrates the effectiveness of the proposed iterative algorithm. In addition, the algorithm can always converge within 20 iterations for different initial points.

5.2.2 *Power control to minimize the total power consumption*

Figure 9 shows the minimum total transmit power consumption with various ADC resoltion. It is worthy noting that only the transmit power is considered. The similar equal user power allocation scheme when $p_{s,1} = p_{s,2} = \cdots = p_{s,K}$ and $P_R = Kp_{s,k}$ is proposed for comparison. For simplicity, all the destination users have the same QoS requirement 0.5 Mbps. As expected, the proposed power control scheme can achieve more power savings than the equal user power allocation one. Besides, the total transmit power decreases with the growing number of relay antennas and higher ADC resolution while guaranteeing the same QoS requirement. However, the hardware will consume more power when deploying more antennas and higher precision ADCs at the relay.

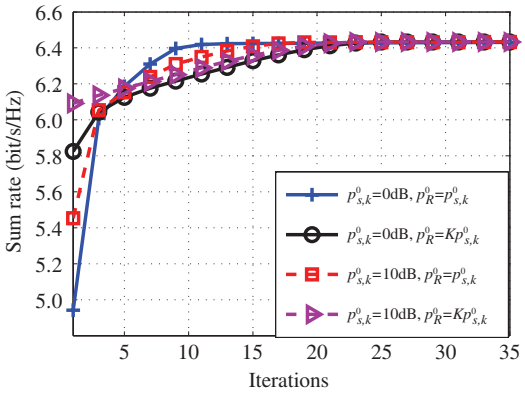

Figure 8 (Color online) Convergence of Algorithm 1 with different initial points when $b = 2$, $N = 256$ and $\rho_p = 10$ dB.

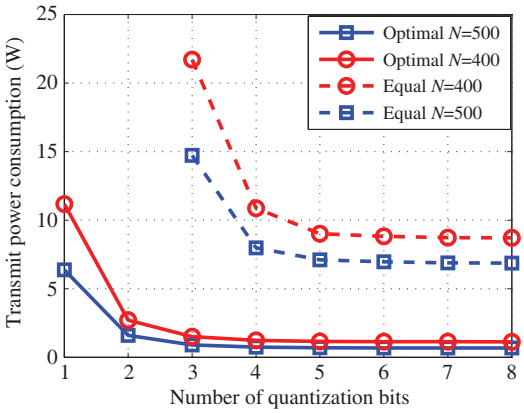

Figure 9 (Color online) Total transmit power consumption versus different numbers of quantization bits when the QoS requirement is 0.5 Mbps.

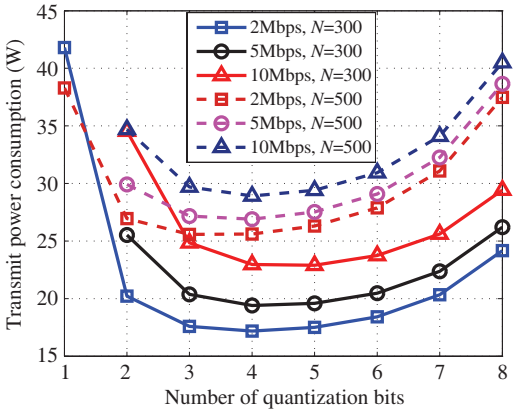

Figure 10 (Color online) Total power consumption versus different numbers of quantization bits.

Figure 10 illustrates the minimum total power consumption to maintain the same QoS requirement with different numbers of ADC bits based the proposed power consumption model (50). We assume $\eta_{s,k} = 0.3$, $\eta_R = 0.39$, $f_{\text{signal}} = 50$ MHz, FOM = 500 fj/conversion-step [35], $P_{\text{MIXER}} = 21$ mW [19], $P_{\text{LNA}} = 20$ mW [19], $P_{\text{FILTERS}} = 5$ mW [19], $P_{\text{SYN}} = 67.5$ mW [19]. As shown in Figure 10, the minimal power consumption grows with the increasing user QoS requirement since more transmit power is needed to satisfy the high date transmission rate. Besides, the minimal total power consumption will first decrease and then increase with the higher ADC resolution. This is because that the transmit power of each user and relay will decrease since increasing the number of ADC bits can enhance the system achievable rate. However, the higher resolution ADCs consumes a large proportion of the total power consumption, which will cause the curves tending to increase. Therefore, from power saving perspectives, the optimal number of quantization bits depends on the user QoS requirement and the number of relay antennas.

6 Conclusion

In this paper, we consider a generalized multi-pair massive MIMO AF relaying system where the relay antennas employ low-precision ADCs. The closed form approximation of the ergodic achievable rate is derived when the channel estimation of the quantized MIMO systems is considered. The impacts of the number of relay antennas, the transmit power of users and relay, the ADC resolution, on the achievable rate are presented by the asymptotic analysis. The generalized power scaling schemes are also illustrated. Furthermore, the optimal joint user and relay power control schemes are designed to enhance the performance of the quantized systems. The results show that to achieve the fixed achievable rate, the

optimal ADC resolution to minimize the total power consumption depends on the user QoS requirements and the number of relay antennas.

Acknowledgements The work was supported by National Nature Science Foundation of China Project (Grant Nos. 61471058, 61501376, 61601018), Key National Science Foundation of China (Grant No. 61461136002), Hong Kong, Macao and Taiwan Science and Technology Cooperation Projects (Grant No. 2016YFE0122900), 111 Project of China (Grant No. B16006), Fundamental Research Funds of Central University (Grant No. 2015RC035), and Shenzhen Science and Technology Project (Grant No. JSGG20150512153045135).

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

References

1 Yang A, Xing W C, Fei Z S, et al. Performance analysis for uplink massive MIMO systems with a large and random number of UEs. Sci China Inf Sci, 2016, 59: 022312
2 Guo Z S, Xing W C, Fei Z S, et al. Robust capacity maximization transceiver design for MIMO OFDM systems. Sci China Inf Sci, 2016, 59: 062301
3 Lu L, Li Y G, Swindlehurst A L, et al. An overview of massive MIMO: benefits and challenges. IEEE J Sel Top Signal Process, 2014, 8: 742–758
4 Bjornson E, Larsson E G, Marzetta T L. Massive MIMO: ten myths and one critical question. IEEE Commun Mag, 2016, 54: 114–123
5 Liu Y J, Lu L, Li Y G, et al. Joint user association and spectrum allocation for small cell networks with wireless backhauls. IEEE Wirel Commun Lett, 2016, 5: 496–499
6 Cai M M, Gao K, Nie D, et al. Effect of wideband beam squint on codebook design in phased-array wireless systems. In: Proceedings of IEEE Global Communications Conference, Washington DC, 2016. 1–6
7 Cui Q M, Yuan T P, Ni W. Energy-efficient two-way relaying under non-ideal power amplifiers. IEEE Trans Veh Technol, 2017, 66: 1257–1270
8 Cui Q M, Zhang Y H, Ni W, et al. Energy efficiency maximization of full-duplex two-way relay with non-ideal power amplifiers and non-negligible circuit power. IEEE Trans Wirel Commun, 2017, 16: 6264–6278
9 Jin S, Liang X S, Wong K K, et al. Ergodic rate analysis for multipair massive MIMO two-way relay networks. IEEE Trans Wirel Commun, 2015, 14: 1480–1491
10 Dai Y Y, Dong X D. Power allocation for multi-pair massive MIMO two-way AF relaying with linear processing. IEEE Trans Wirel Commun, 2016, 15: 5932–5946
11 Wang Y, Li S D, Li C G, et al. Ergodic rate analysis for massive MIMO relay systems with multi-pair users under imperfect CSI. In: Proceedings of IEEE Global Conference on Signal and Information Processing, Orlando, 2015. 33–37
12 Shen H, Xu W, Zhao C M. Transceiver optimization for full-duplex massive MIMO AF relaying with direct link. IEEE Access, 2016, 4: 8857–8864
13 Xu W, Liu J, Jin S. Spectral and energy efficiency of multi-pair massive MIMO relay network with hybird processing. IEEE Trans Commun, 2017, 65: 3794–3809
14 Feng J J, Ma S D, Yang G H, et al. Power scaling of full-duplex two-way massive MIMO relay systems with correlated antennas and MRC/MRT processing. IEEE Trans Wirel Commun, 2017, 16: 4738–4753
15 Wang Q, Jing Y D. Performance analysis and scaling law of MRC/MRT relaying with CSI error in multi-pair massive MIMO systems. IEEE Trans Wirel Commun, 2017, 16: 5882–5896
16 Fan L, Jin S, Wen C K, et al. Uplink achievable rate for massive MIMO systems with low-resolution ADC. IEEE Commun Lett, 2015, 19: 2186–2189
17 Zhang J Y, Dai L L, He Z Y, et al. Performance analysis of mixed-ADC massive MIMO systems over rician fading channels. IEEE J Sel Areas Commun, 2017, 35: 1327–1338
18 Tan W Q, Jin S, Wen C K, et al. Spectral efficiency of mixed-ADC receivers for massive MIMO Systems. IEEE Access, 2016, 4: 7841–7846
19 Liang N, Zhang W Y. Mixed-ADC massive MIMO. IEEE J Sel Areas Commun, 2016, 34: 983–997
20 Mo J, Alkhateeb A, Surra S A, et al. Hybrid architectures with few-bit ADC receivers: achievable rates and energy-rate tradeoffs. IEEE Trans Wirel Commun, 2017, 16: 2274–2287
21 Liu J, Xu J D, Xu W, et al. Multiuser massive MIMO relaying with mixed-ADC receiver. IEEE Signal Process Lett, 2017, 24: 76–80
22 Jia X D, Zhou M, Xie M G, et al. Optimal design of secrecy massive MIMO amplify-and-forward relaying systems with double resolution ADCs antenna array. IEEE Access, 2016, 4: 8757–8774
23 Kong C L, Zhong C J, Jin S, et al. Full-duplex massive MIMO relaying systems with low-resolution ADCs. IEEE Trans Wirel Commun, 2017, 16: 5033–5047
24 Bai Q, Mittmann U, Mezghani A, et al. Minimizing the energy per bit for pilot-assisted data transmission over quantized channels. In: Proceedings of IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communication, Washington DC, 2014. 81–85

25 Fan L, Qiao D, Jin S, et al. Optimal pilot length for uplink massive MIMO systems with low-resolution ADC. In: Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop, Rio de Janerio, 2016. 1–5
26 Amin O, Gedik B, Uysal M. Channel estimation for amplify-and-forward relaying: cascaded against disintegrated estimators. IET Commun, 2010, 4: 1207–1216
27 Patel C, Stuber G. Channel estimation for amplify and forward relay based cooperation diversity systems. IEEE Trans Wirel Commun, 2007, 6: 2348–2356
28 Liu Y S, Tan Z H, Hu H L, et al. Channel estimation for OFDM. IEEE Commun Surv Tutor, 2014, 16: 1891–1908
29 Cui Q M, Gu Y, Ni W, et al. Effective capacity of licensed-assisted access in unlicensed spectrum for 5G: from theory to application. IEEE J Sel Areas Commun, 2017, 35: 1754–1767
30 Yan S, Peng M G, Ababa M, et al. An evolutionary game for user access mode selection in fog radio access networks. IEEE Access, 2017, 5: 2200–2210
31 Zhang Q, Jin S, Wong K K. Power scaling of uplink massive MIMO systems with arbitrary-rank channel means. IEEE J Sel Top Signal Process, 2014, 8: 966–981
32 Boyd S, Vandenberghe L. Convex Optimization. Cambridge: Cambridge University Press, 2004. 160–167
33 Weeraddana P C, Codreanu M, Latva-aho M, et al. Resource allocation for cross-layer utility maximization in wireless networks. IEEE Trans Veh Technol, 2011, 60: 2790–2809
34 Lee H S, Sodini C G. Analog-to-digital converters: digitizing the analog world. Proc IEEE, 2008, 96: 323–334
35 Murmann B. Adc performance survey 1997–2016. http://web.stanford.edu/ murmann/adcsurvey.html

Appendix A Proof of the Theorem 1

According to (21), we need to calculate $\bar{S}_k$, $\bar{I}_{1,k}$, $\bar{I}_{2,k}$,$\bar{I}_{3,k}$,$\bar{I}_{4,k}$ and $\bar{I}_{5,k}$, respectively, which will be calculated in the following

1. First, we calculate the desired signal power $\bar{S}_k$. We have

$$\bar{S}_k = \mathrm{E}\left\{S_k\right\} = \mathrm{E}\left\{\alpha^2 p_{s,k}\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right\|^2\right\} = \alpha^2 p_{s,k}\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right\|^2\right\}.$$

Then, we have

$$\begin{aligned}\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right\|^2\right\} &= \mathrm{E}\left\{\left\|\sum_{i=1}^{K}\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right\|^2\right\}\\ &= \mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right\|^2\right\} + \mathrm{E}\left\{\left\|\sum_{i=1,i\neq k}^{K}\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right\|^2\right\}\\ &= \mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{s,k}\right\|^4\left\|\hat{\boldsymbol{g}}_{d,k}\right\|^4\right\} + \sum_{i=1,i\neq k}^{K}\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right\|^2\right\}. \end{aligned} \tag{A1}$$

Since $\hat{\boldsymbol{g}}_{s,k}$ contains i.i.d. $\mathcal{CN}(0,\sigma_{s,k}^2)$ entries, $\|\hat{\boldsymbol{g}}_{s,k}\|^2$ follows the gamma distribution with shape parameter N and scale parameter $\sigma_{s,k}^2$. Therefore, $\mathrm{E}\|\hat{\boldsymbol{g}}_{s,k}\|^4 = (N^2+N)\sigma_{s,k}^4$. Similarly, $\|\hat{\boldsymbol{g}}_{d,k}\|^2 \sim \Gamma(N,\sigma_{d,k}^2)$ and we have $\mathrm{E}\|\hat{\boldsymbol{g}}_{d,k}\|^4 = (N^2+N)\sigma_{d,k}^4$. Since $\hat{\boldsymbol{g}}_{s,k},\hat{\boldsymbol{g}}_{d,k}$ are mutually independent, we have

$$\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{s,k}\right\|^4\left\|\hat{\boldsymbol{g}}_{d,k}\right\|^4\right\} = \left(N^2+N\right)^2\sigma_{s,k}^4\sigma_{d,k}^4. \tag{A2}$$

Due to the fact that $\hat{\boldsymbol{g}}_{s,k},\hat{\boldsymbol{g}}_{s,i},\hat{\boldsymbol{g}}_{d,i},\hat{\boldsymbol{g}}_{d,k}$ are mutually independent, the second term of (A1) can be written as

$$\sum_{i=1,i\neq k}^{K}\mathrm{E}\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right\|^2 = \sum_{i=1,i\neq k}^{K}\mathrm{E}\|\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{d,i}\|^2\cdot\mathrm{E}\|\hat{\boldsymbol{g}}_{s,k}\hat{\boldsymbol{g}}_{s,i}\|^2 = N^2\sigma_{s,k}^2\sigma_{d,k}^2\sum_{i=1,i\neq k}^{K}\sigma_{s,i}^2\sigma_{d,i}^2. \tag{A3}$$

By substituting (A2) and (A3) into (A1), we can obtain

$$\bar{S}_k = \alpha^2 p_{s,k}\left(N^4+2N^3\right)\sigma_{s,k}^4\sigma_{d,k}^4 + \alpha^2 p_{s,k}N^2\sigma_{s,k}^2\sigma_{d,k}^2\sum_{j=1}^{K}\sigma_{s,j}^2\sigma_{d,j}^2. \tag{A4}$$

2. $\bar{I}_{1,k}$ denotes the multi-user interference, which will be derived as follows. By using the similar derivation methods as $\bar{S}_k$, we have

$$\begin{aligned}&\alpha^2\sum_{i=1,i\neq k}^{K}p_{s,i}\left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\boldsymbol{g}_{s,i}\right\|^2 = \alpha^2\sum_{i=1,i\neq k}^{K}p_{s,i}\sum_{j=1}^{K}\left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,j}\hat{\boldsymbol{g}}_{s,j}^{\mathrm{H}}\boldsymbol{g}_{s,i}\right\|^2\\ &= \alpha^2\sum_{i=1.i\neq k}^{K}p_{s,i}\left(\left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\boldsymbol{g}_{s,i}\right\|^2 + \left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\boldsymbol{g}_{s,i}\right\|^2 + \sum_{j=1,j\neq i,j\neq k}^{K}\left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,j}\hat{\boldsymbol{g}}_{s,j}^{\mathrm{H}}\boldsymbol{g}_{s,i}\right\|^2\right). \end{aligned} \tag{A5}$$

Then, we can have

$$\mathrm{E}\left\{\left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\boldsymbol{g}_{s,i}\right\|^2\right\} = N^3\sigma_{d,k}^4\sigma_{s,k}^2\beta_{s,i} + N^2\sigma_{s,k}^2\beta_{s,i}\sigma_{d,k}^2\beta_{d,k}, \tag{A6}$$

$$\mathrm{E}\left\{\left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\boldsymbol{g}_{s,i}\right\|^2\right\}=N^3\sigma_{s,i}^4\sigma_{d,i}^2\beta_{d,k}+N^2\beta_{d,k}\sigma_{d,i}^2\sigma_{s,i}^2\beta_{s,i}, \tag{A7}$$

$$\mathrm{E}\left\{\sum_{j=1,j\neq i,j\neq k}^{K}\left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,j}\hat{\boldsymbol{g}}_{s,j}^{\mathrm{H}}\boldsymbol{g}_{s,i}\right\|^2\right\}=\sum_{j=1,j\neq i,j\neq k}^{K}N^2\beta_{d,k}\beta_{s,i}\sigma_{s,j}^2\sigma_{d,j}^2. \tag{A8}$$

By substituting (A6), (A7) and (A8) into (A5), $\bar{I}_{1,k}$ can be obtained.

3. The channel estimation errors $\bar{I}_{2,k}$ can be derived similarly as $\bar{I}_{1,k}$, which is omitted due to the limited space.

4. The noise power $\bar{I}_{3,k}$ can be obtained by the similar approach as above ones, which is also omitted due to the limited space.

5. The quantization noise $\bar{I}_{4,k}$ is calculated as follows

$$\begin{aligned}\mathrm{E}\left\{\left\|\boldsymbol{g}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}&=\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}+\mathrm{E}\left\{\left\|\tilde{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}\\&=\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}+\sum_{i=1,i\neq k}^{K}\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}+\sum_{i=1}^{K}\mathrm{E}\left\{\left\|\tilde{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}.\end{aligned} \tag{A9}$$

Due to the fact that $\mathrm{tr}\,(\boldsymbol{AB})=\mathrm{tr}\,(\boldsymbol{BA})$, the first term of (A9) can be written as

$$\begin{aligned}\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}&=\mathrm{E}\left\{\mathrm{tr}\left(\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\boldsymbol{n}_q\boldsymbol{n}_q^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\right)\right\}\\&=\mathrm{E}\left\{\mathrm{tr}\left(\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\boldsymbol{n}_q\boldsymbol{n}_q^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,k}\right)\right\}\\&=\left(N^2+N\right)\sigma_{d,k}^4\cdot\alpha\left(1-\alpha\right)\mathrm{E}\left\{\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\mathrm{diag}\left(\boldsymbol{G}_s\boldsymbol{P}\boldsymbol{G}_s^{\mathrm{H}}+\boldsymbol{I}_N\right)\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\right\}.\end{aligned} \tag{A10}$$

According to [16] and by using the fact $\sigma_{s,k}^2+e_{s,k}^2=\beta_{s,k}$,

$$\mathrm{E}\left\{\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\mathrm{diag}\left(\boldsymbol{G}_s\boldsymbol{P}\boldsymbol{G}_s^{\mathrm{H}}+\boldsymbol{I}_N\right)\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\right\}=N\left(\sigma_{s,k}^2+\sigma_{s,k}^2\sum_{i=1}^{K}p_{s,i}\beta_{s,i}+p_{s,k}\sigma_{s,k}^4\right). \tag{A11}$$

Therefore, we can have

$$\mathrm{E}\left\{\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,k}\hat{\boldsymbol{g}}_{s,k}^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}=\alpha\left(1-\alpha\right)\left(N^3+N^2\right)\sigma_{d,k}^4\left(\sigma_{s,k}^2+\sigma_{s,k}^2\sum_{i=1}^{K}p_{s,i}\beta_{s,i}+p_{s,k}\sigma_{s,k}^4\right). \tag{A12}$$

Based on the above approach, we can get

$$\sum_{i=1,i\neq k}^{K}\mathrm{E}\left\|\hat{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\boldsymbol{n}_q\right\|^2=\sigma_{d,k}^2\sum_{i=1,i\neq k}^{K}N^2\sigma_{d,i}^2\left(\sigma_{s,i}^2+\sigma_{s,i}^2\sum_{j=1}^{K}p_{s,j}\beta_{s,j}+p_{s,i}\sigma_{s,i}^4\right), \tag{A13}$$

and

$$\sum_{i=1}^{K}\mathrm{E}\left\{\left\|\tilde{\boldsymbol{g}}_{d,k}^{\mathrm{H}}\hat{\boldsymbol{g}}_{d,i}\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\boldsymbol{n}_q\right\|^2\right\}=N^2e_{d,k}^2\sum_{i=1}^{K}\sigma_{d,i}^2\left(\sigma_{s,i}^2+\sigma_{s,i}^2\sum_{j=1}^{K}p_{s,j}\beta_{s,j}+p_{s,i}\sigma_{s,i}^4\right), \tag{A14}$$

respectively. By substituting (A12), (A13) and (A14) into (A9), $\bar{I}_{4,k}$ can be derived.

6. $\bar{I}_{5,k}$ has a relationship with the relay power normalization factor λ. According to (18), λ consists of three terms we need to calculate, which are presented in the following.

(a) First, we have

$$\begin{aligned}\mathrm{E}\left\{\left\|\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\boldsymbol{G}_s\boldsymbol{x}\right\|_F^2\right\}&\overset{(a)}{=}\mathrm{E}\left\{\left\|\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\boldsymbol{x}\right\|_F^2\right\}+\mathrm{E}\left\{\left\|\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\tilde{\boldsymbol{G}}_s\boldsymbol{x}\right\|_F^2\right\}\\&\overset{(b)}{=}\mathrm{E}\left\{\mathrm{tr}\left(\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\boldsymbol{P}\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\hat{\boldsymbol{G}}_d^{\mathrm{H}}\right)\right\}+\mathrm{E}\left\{\mathrm{tr}\left(\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\tilde{\boldsymbol{G}}_s\boldsymbol{P}\tilde{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\hat{\boldsymbol{G}}_d^{\mathrm{H}}\right)\right\}\\&\overset{(c)}{=}\mathrm{E}\left\{\mathrm{tr}\left(\hat{\boldsymbol{G}}_d^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\boldsymbol{P}\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\right)\right\}+\mathrm{E}\left\{\mathrm{tr}\left(\hat{\boldsymbol{G}}_d^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\tilde{\boldsymbol{G}}_s\boldsymbol{P}\tilde{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\right)\right\},\end{aligned} \tag{A15}$$

where (a) follows the results of the channel estimation, (b) is due to the fact $\|\boldsymbol{A}\|_F^2=\mathrm{tr}(\boldsymbol{AA}^{\mathrm{H}})$ and (c) follows $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}(\boldsymbol{BA})$. By using the fact that $\mathrm{E}\|\hat{\boldsymbol{g}}_{d,k}\|^2=N\sigma_{d,k}^2$, and due to the fact that $\mathrm{E}\|\hat{\boldsymbol{g}}_{s,k}\|^4=(N^2+N)\sigma_{s,k}^4$ and when $i\neq j$, $\mathrm{E}\|\hat{\boldsymbol{g}}_{s,i}^{\mathrm{H}}\hat{\boldsymbol{g}}_{s,j}\|^2=N\sigma_{s,i}^2\sigma_{s,j}^2$, we have

$$\mathrm{E}\left\{\mathrm{tr}\left(\hat{\boldsymbol{G}}_d^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\boldsymbol{P}\hat{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\right)\right\}=\sum_{k=1}^{K}\sigma_{d,k}^2\left(N^3p_{s,k}\sigma_{s,k}^4+N^2\sigma_{s,k}^2\sum_{i=1}^{K}p_{s,i}\sigma_{s,i}^2\right). \tag{A16}$$

Following the similar steps, we can get

$$\mathrm{E}\left\{\mathrm{tr}\left(\hat{\boldsymbol{G}}_d^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\tilde{\boldsymbol{G}}_s\boldsymbol{P}\tilde{\boldsymbol{G}}_s^{\mathrm{H}}\hat{\boldsymbol{G}}_s\right)\right\}=N^2\sum_{k=1}^{K}\sigma_{d,k}^2\sigma_{s,k}^2\sum_{i=1}^{K}p_{s,i}e_{s,i}^2. \tag{A17}$$

Substituting (A16) and (A17) into (A15), it yields

$$\alpha^2\mathrm{E}\Big\{\left\|\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\boldsymbol{G}_s\boldsymbol{x}\right\|_F^2\Big\}=\alpha^2\sum_{k=1}^{K}\sigma_{d,k}^2\left(N^3p_{s,k}\sigma_{s,k}^4+N^2\sigma_{s,k}^2\sum_{i=1}^{K}p_{s,i}\beta_{s,i}\right). \tag{A18}$$

(b) Following the similar approach as the above derivation, we can have

$$\alpha^2\mathrm{E}\Big\{\left\|\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\right\|_F^2\Big\}=\alpha^2N^2\sum_{k=1}^{K}\sigma_{s,k}^2\sigma_{d,k}^2. \tag{A19}$$

(c) To calculate the quantization noise term, due to the fact that $\|\boldsymbol{A}\|_F^2=\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\right)$ and $\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{B}\right)=\mathrm{tr}\left(\boldsymbol{B}\boldsymbol{A}\right)$, we can rewrite

$$\mathrm{E}\Big\{\left\|\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\boldsymbol{n}_q\right\|_F^2\Big\}=\alpha\left(1-\alpha\right)\mathrm{E}\left[\mathrm{tr}\left(\hat{\boldsymbol{G}}_d^{\mathrm{H}}\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\mathrm{diag}\left\{\boldsymbol{G}_s\boldsymbol{P}\boldsymbol{G}_s^{\mathrm{H}}+\boldsymbol{I}_N\right\}\hat{\boldsymbol{G}}_s\right)\right]. \tag{A20}$$

Due to the fact that $\mathrm{E}\Big\{\left\|\hat{\boldsymbol{g}}_{d,k}\right\|^2\Big\}=N\sigma_{d,k}^2$ and (A11), we can obtain

$$\mathrm{E}\Big\{\left\|\hat{\boldsymbol{G}}_d\hat{\boldsymbol{G}}_s^{\mathrm{H}}\boldsymbol{n}_q\right\|_F^2\Big\}=\alpha\left(1-\alpha\right)N^2\sum_{k=1}^{K}\sigma_{d,k}^2\left(p_{s,k}\sigma_{s,k}^4+\sigma_{s,k}^2\sum_{i=1}^{K}p_{s,i}\beta_{s,i}+\sigma_{s,k}^2\right). \tag{A21}$$

By substituting (A18), (A19) and (A21) into (18), $\bar{I}_{5,k}$ will be achieved.

Finally, based on the above derivations of $\bar{S}_k$, $\bar{I}_{1,k}$, $\bar{I}_{2,k}$, $\bar{I}_{3,k}\bar{I}_{4,k}$, $\bar{I}_{5,k}$, and by some basic manipulations, the proof of Theorem 1 is completed.