# Achievable Rate Analysis and Feedback Design for Multiuser MIMO Relay with Imperfect CSI

Zhangjie Peng, *Student Member, IEEE,* Wei Xu, *Member, IEEE,*
Li-Chun Wang, *Fellow, IEEE,* and Chunming Zhao, *Member, IEEE*

*Abstract*—This paper investigates a multiuser MIMO relay downlink system with imperfect channel estimation and limited feedback. We analyze the achievable rate loss due to channel state information (CSI) mismatch arising from both channel estimation and quantization feedback. We first derive an upper bound to characterize the effect of imperfect CSI, and then present a limited feedback strategy for both the two-hop channels in the relay system. This newly presented feedback strategy reveals the relationship among the CSI quantization levels, the transmit power, and the pilots for channel estimation. An optimized pilot design at the base station and the relay is also presented by applying the derived bounds for the two-hop system.

*Index Terms*—Achievable rate, limited feedback, imperfect CSI, MIMO relay, channel estimation

## I. Introduction

RELAYING has been well acknowledged as a promising technique for improving the cell-edge performance of conventional cellular networks. Meanwhile, applying the multiple-input multiple-output (MIMO) technique can significantly improve the capacity of wireless systems [1] [2]. By combining the relaying and MIMO techniques, MIMO-relay techniques can thus make use of their advantages to increase data rate at the cell edge and extend the network coverage [3]-[5].

The theoretical capacity for a single user MIMO relay system has been widely investigated in [6]-[10]. Specially, in [6]-[8], ergodic capacity and precoding schemes for a MIMO system with a single relay were studied. The average capacity with relay selection for multiple-relay MIMO system was

analyzed in [9], and the capacity of a multi-hop relay channel was characterized in [10]. The performance of a multiuser MIMO relay system with different beamforming methods was studied in [11], while a linear precoding design was presented in [12] for the system by jointly optimizing the precoders at both source and relay sides.

In the above-mentioned literatures, the authors all assumed that channel state information (CSI) at the transmitter (CSIT) and the receiver (CSIR) are perfect. However, in practice, either exact CSIT or exact CSIR is hard to achieve. A common way to deal with this problem is to deploy channel estimation at the receiver and then send a quantized version of CSI back to the transmitter for precoding design [13] [14]. This process degrades the system performance due to imperfect CSI at both sides. Generally, there exist three imperfect CSI scenarios: 1) imperfect CSIR, without quantization errors in the backhaul channel, 2) imperfect CSIT due to quantization errors, but having perfect CSI at receiver, 3) without perfect CSI at both transmitter and receiver, saying that both estimation errors and quantization errors exist.

In the first imperfect CSI scenario, the receiver commonly uses pilots sent by the transmitter to estimate channels with errors. For a point-to-point MIMO relay system, a superimposed training strategy and the bit error rate (BER) performance have been investigated [15]-[18]. The capacity lower bound when using pilot to estimate the source-relay-destination channel was then derived in [19]. The authors of [20] then analyzed the robustness for a multipoint-to-multipoint MIMO relay system with channel estimation errors at both source-relay and relay-destination channels.

In the second imperfect CSI scenario, the transmitter is unable to obtain the CSI perfectly, especially in a frequency division duplexing (FDD) system. The design of an effective limited feedback mechanism is crucial to deal with this issue, where each receiver sends back a finite number of bits of CSI to the corresponding transmitter [14] [21]. In [22], the capacity and an optimized feedback strategy have been presented for a multiuser relay system with imperfect CSIT at both the base station (BS) and the relay.

In the third imperfect CSI scenario which is more practical in applications, the channel estimation and limited feedback techniques are applied to obtain the CSIR and CSIT, respectively. Nevertheless, to our best knowledge, few existing works have considered both channel estimation and limited feedback in MIMO relaying systems. Two noticeable exceptions are [23] and [24]. In [23], the authors analyzed the achievable rate and relay precoder design for a MIMO point-to-point two hop
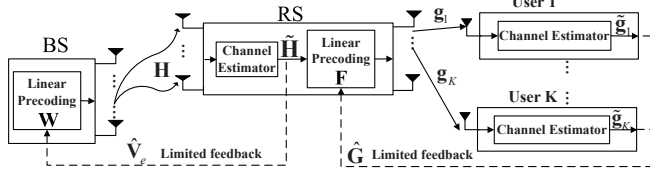
Fig. 1.   MIMO relay system model.

network, and the authors of [24] investigated the throughput for a single user multi-hop MIMO relay system. However, both studies were confined to a single user relay system, and the CSI feedback strategy was not considered.

In this work, we investigate the effects of various CSI mismatches under a multiuser two-hop relay downlink channel. We will first characterize the rate loss due to both channel estimation errors and limited feedback by deriving a rate loss upper bound. This bound is then utilized for developing a scaling CSI feedback strategy to improve the entire system performance, hence the well-known "interference limited" effect [14] for limited feedback systems is avoided. Moreover, we also derive an optimized BS-relay pilot design to assist channel estimation.

The remainder of this paper is organized as follows. Section II presents the multiuser MIMO relaying system model under consideration. Section III analyzes the rate loss as well as the achievable rate bounds for the system. Section IV presents the scaling number of feedback bits in order to maintain a predetermined rate loss and designs the optimal numbers of pilots for channel estimation. Section V provides numerical simulations before conclusions in Section VI.

*Notation*: Boldface upper and lower case letters respectively denote matrices and column vectors; $(.)^H$ stands for Hermitian transpose. The complex number field is represented by $\mathbb{C}$. $\mathbf{I}_M$ is an identity matrix of size $M \times M$. $\mathcal{CN}(\mu, \nu)$ stands for the complex Gaussian distribution with mean $\mu$ and covariance $\nu$. $E\{\cdot\}$ denotes the expectation operator. $\mathbf{A}_{ij}$ denotes the $(i, j)$th element of a matrix $\mathbf{A}$, while $\mathbf{a}_{k,n}$ represents the $n$th element of a vector $\mathbf{a}_k$. $\|\mathbf{a}\|$ and $|a|$ stand for the Euclidean norm of a vector $\mathbf{a}$ and the norm of scalar $a$, respectively. $h(x)$ is the differential entropy of random variable $x$, and $\mathcal{I}(x; y)$ is the mutual information between random variables $x$ and $y$.

## II. SYSTEM MODEL

Consider an amplify-and-forward (AF) MIMO relay wireless system in Fig. 1. There is an $M$-antenna BS serving $K$ single-antenna users through a relay station (RS) equipped with $N$ antennas. In this paper, we neglect the direct link from the BS to users due to severe path loss. Denote that $\mathbf{H} \in \mathbb{C}^{N \times M}$ is the channel matrix of the channel from BS to RS, and $\mathbf{g}_k \in \mathbb{C}^{N \times 1}$ $(k = 1, \cdots K)$ is the channel vector from RS to the $k$th user. Then, the received symbol at the $k$th user for the input symbol vector $\mathbf{x} \in \mathbb{C}^{K \times 1}$ is given by

$$y_k = \sqrt{\rho_1 \rho_2} \mathbf{g}_k^H \mathbf{FHWx} + \sqrt{\rho_2} \mathbf{g}_k^H \mathbf{Fn} + z_k \qquad (1)$$

where $\mathbf{W} \in \mathbb{C}^{M \times K}$ and $\mathbf{F} \in \mathbb{C}^{N \times N}$ are precoding matrices for the BS and the RS, respectively; $\mathbf{n} \in \mathbb{C}^{N \times 1}$ and $z_k$ denote the complex additive white Gaussian noises (AWGNs) at the RS and the $k$th user, respectively; and $\rho_1$ and $\rho_2$ are

scaling factors for the transmit symbols at the BS and the RS, respectively. Note that the channel input symbols $\mathbf{x}$ normalized by $E\{\mathbf{xx}^H\} = \mathbf{I}_K$ are Gaussian independent and identically distributed (i.i.d.) random variables. In this paper, we consider the Rayleigh fading channels with independent fading from block to block, and the entries of $\mathbf{H}$, $\mathbf{g}_k$, $\mathbf{n}$ and $z_k$ are independent complex Gaussian variables with zero-mean and unit variance. Also, we assume that the number of antennas at the BS is not less than that at the RS, which means $M \geq N$. To focus our study on the effect of limited feedback and channel estimation, we consider $K = N$.

### A. Linear Precoding

In this study, we exploit the precoding scheme in [22] with either perfect or limited CSI feedback. First, we briefly review the precoding scheme. The RS performs the singular value decomposition (SVD) of $\mathbf{H}$, which yields

$$\mathbf{H} = \mathbf{U}[\boldsymbol{\Sigma}\ \mathbf{0}][\mathbf{V}\ \mathbf{V}_r]^H \qquad (2)$$

where $\mathbf{U}$ is an $N \times N$ unitary matrix, $\boldsymbol{\Sigma}$ is an $N \times N$ diagonal matrix, and $\mathbf{V}$ is a column unitary matrix with size $M \times N$. The precoding matrix $\mathbf{W}$ at the BS is

$$\mathbf{W} = \mathbf{V}, \qquad (3)$$

and the precoding matrix $\mathbf{F}$ at the RS follows

$$\mathbf{F} = \mathbf{F}_G \mathbf{U}^H \qquad (4)$$

where $\mathbf{F}_G = [\mathbf{f}_{G,1}, \cdots, \mathbf{f}_{G,N}]$ is calculated according to zero-forcing beamforming (ZFBF) [14], $\mathbf{f}_{G,k}$ is the normalized $k$th column of $\mathbf{G}^H(\mathbf{GG}^H)^{-1}$, and $\mathbf{G}$ is defined by $\mathbf{G} = [\mathbf{g}_1, \cdots, \mathbf{g}_N]^H$. By exploiting the precoding matrices in (3) and (4), the system model in (1) becomes

$$y_k = \sqrt{\rho_1 \rho_2} \mathbf{g}_k^H \mathbf{F}_G \boldsymbol{\Sigma} \mathbf{x} + \sqrt{\rho_2} \mathbf{g}_k^H \mathbf{F}_G \bar{\mathbf{n}} + z_k \qquad (5)$$

where $\bar{\mathbf{n}} = \mathbf{U}^H \mathbf{n}$. Since every column of $\mathbf{U}^H$ is an unit-norm vector, the equivalent AWGN $\bar{\mathbf{n}}$ is complex Gaussian distributed with an identity variance $\mathbf{I}_N$. Given the BS transmit power constraint $P_1$ and the RS transmit power constraint $P_2$, the scaling factors $\rho_1$ and $\rho_2$ are determined by [22]

$$\rho_1 = \frac{P_1}{N}, \qquad \rho_2 = \frac{P_2}{MP_1 + N}. \qquad (6)$$

### B. Channel Estimation

In this subsection, we characterize both two-hop channels with estimation error. In order to obtain $\mathbf{H}$ and $\mathbf{g}_k$ at corresponding receivers, we employ channel estimation techniques using pilots at the relay and remote users. Like [25], we assume that the BS uses $\beta_1 M$ orthogonal pilots for $M$ transmit antennas at the BS. The $\beta_1 M$ pilots are then utilized at the relay for estimating $\mathbf{H}$ via minimum mean square error (MMSE) estimation. Meanwhile, a similar procedure is implemented at remote users to estimate $\mathbf{g}_k$ with $\beta_2 N$ orthogonal pilots. Then the relationships between the real channels and their estimations are modeled by [25] [26]

$$\mathbf{H} = \tilde{\mathbf{H}} + \mathbf{E}, \qquad (7)$$

and

$$\mathbf{g}_k = \tilde{\mathbf{g}}_k + \mathbf{e}_k \qquad (8)$$

where $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{g}}_k$ are the estimations of $\mathbf{H}$ and $\mathbf{g}_k$, respectively. $\mathbf{E}$ and $\mathbf{e}_k$ are the corresponding estimation errors that are independent of their estimated channel components. The entries of $\tilde{\mathbf{H}}$ are i.i.d. complex Gaussian distributed with its elements $\tilde{\mathbf{H}}_{ij} \sim \mathcal{CN}\left(0, 1-\delta_1^2\right)$ and $\tilde{\mathbf{g}}_{k,n} \sim \mathcal{CN}(0, 1-\delta_2^2)$. Similarly, the entries of $\mathbf{E}$ and $\mathbf{e}_k$ follow $\mathbf{E}_{ij} \sim \mathcal{CN}(0, \delta_1^2)$ and $\mathbf{e}_{k,n} \sim \mathcal{CN}(0, \delta_2^2)$, respectively. Further, given that the energy of noise is normalized, we have $\delta_1^2 = \frac{1}{1+\beta_1 P_1}$ and $\delta_2^2 = \frac{1}{1+\beta_2 P_2}$.

### C. Channel State Information Feedback

In this subsection, we consider the precoding scheme presented in Subsection II-A to scenarios using estimated and quantized CSI. When a RS obtains $\tilde{\mathbf{H}}$, it performs SVD of $\tilde{\mathbf{H}}$ as follows:

$$\tilde{\mathbf{H}} = \tilde{\mathbf{U}}[\tilde{\boldsymbol{\Sigma}}\ \mathbf{0}][\tilde{\mathbf{V}}\ \tilde{\mathbf{V}}_r]^H \qquad (9)$$

where $\tilde{\mathbf{U}}$ is an $N \times N$ unitary matrix, $\tilde{\boldsymbol{\Sigma}}$ is an $N \times N$ diagonal matrix, and $\tilde{\mathbf{V}}$ is a column unitary matrix with size $M \times N$. We use $B_1$ bits (as an index) to quantize each column of $\tilde{\mathbf{V}}$ via random vector quantization (RVQ) [27], and then the RS sends the index to the BS. Let $\hat{\mathbf{V}}_e = [\hat{\mathbf{v}}_{e,1}, \cdots, \hat{\mathbf{v}}_{e,N}]$ be the quantization of $\tilde{\mathbf{V}}$ where $\hat{\mathbf{v}}_{e,k}$ is column quantization of $\tilde{\mathbf{V}}$ given by

$$\hat{\mathbf{v}}_{e,k} = \arg \max_{\mathbf{v}_{\mathbf{q}_j} \in \mathbf{V}_\mathbf{Q}} \left| \tilde{\mathbf{v}}_k^H \mathbf{v}_{\mathbf{q}_j} \right| \qquad (10)$$

and $\mathbf{V}_\mathbf{Q} = \left[ \mathbf{v}_{\mathbf{q}_1}, \cdots, \mathbf{v}_{\mathbf{q}_{2^{B_1}}} \right]$ is the codebook. The precoding matrix $\mathbf{W}$ is equal to

$$\mathbf{W} = \hat{\mathbf{V}}_e . \qquad (11)$$

Meanwhile at the user side, after obtaining $\tilde{\mathbf{g}}_k$ through channel estimation, the $k$th user quantizes the vector with $B_2$ bits. Let $\hat{\mathbf{g}}_k$ be the quantized version of $\tilde{\mathbf{g}}_k$. We have

$$\hat{\mathbf{g}}_k = \arg \max_{\mathbf{g}_{\mathbf{q}_j} \in \mathbf{G}_\mathbf{Q}} \left| \tilde{\mathbf{g}}_k^H \mathbf{g}_{\mathbf{q}_j} \right| . \qquad (12)$$

where $\mathbf{G}_\mathbf{Q} = \left[ \mathbf{g}_{\mathbf{q}_1}, \cdots, \mathbf{g}_{\mathbf{q}_{2^{B_2}}} \right]$ is the codebook. By defining $\hat{\mathbf{G}} = [\hat{\mathbf{g}}_1, \cdots, \hat{\mathbf{g}}_N]^H$ at the RS, $\mathbf{F}$ is designed by

$$\mathbf{F} = \hat{\mathbf{F}}_G \tilde{\mathbf{U}}^H \qquad (13)$$

where $\hat{\mathbf{F}}_G = [\hat{\mathbf{f}}_{G,1}, \cdots, \hat{\mathbf{f}}_{G,N}]$ and $\hat{\mathbf{f}}_{G,k}$ is the normalized $k$th column of $\hat{\mathbf{G}}^H \left( \hat{\mathbf{G}} \hat{\mathbf{G}}^H \right)^{-1}$. Further by substituting (7), (11), and (13) into (1), it gives

$$\begin{aligned}\tilde{y}_k &= \sqrt{\rho_2}\mathbf{g}_k^H \mathbf{F}\left(\sqrt{\rho_1}\mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{n}\right) + z_k \\ &= \sqrt{\rho_1 \rho_2}\mathbf{g}_k^H \hat{\mathbf{F}}_G \tilde{\mathbf{U}}^H \tilde{\mathbf{H}}\hat{\mathbf{V}}_e\mathbf{x} + \sqrt{\rho_1 \rho_2}\mathbf{g}_k^H \hat{\mathbf{F}}_G \tilde{\mathbf{U}}\mathbf{E}\hat{\mathbf{V}}_e\mathbf{x} \\ &\quad + \sqrt{\rho_2}\mathbf{g}_k^H \hat{\mathbf{F}}_G \tilde{\mathbf{n}} + z_k \end{aligned} \qquad (14)$$

where $\tilde{\mathbf{n}} = \tilde{\mathbf{U}}^H \mathbf{n}$. Note that if there is no estimation and quantization error, implying that $\tilde{\mathbf{H}} = \mathbf{H}$, $\tilde{\mathbf{g}}_k = \mathbf{g}_k$, $\mathbf{E} = \mathbf{0}$, $\hat{\mathbf{V}}_e = \mathbf{V}$, $\hat{\mathbf{F}}_G = \mathbf{F}_G$, and $\tilde{\mathbf{U}} = \mathbf{U}$, the received signal in (14) reduces to (5). It indicates that $y_k$ is a special case of $\tilde{y}_k$.

### III. ACHIEVABLE RATE BOUNDS

In this section, we study the effect of CSI imperfection on the achievable rate in a MIMO relaying system. We first derive a lower bound of the achievable rate for each user, which then generates an upper bound of the rate loss due to both channel estimation error and quantized CSI feedback.

### A. Upper Bound of the Rate Loss

Specially, when perfect CSI are available at both BS and RS, the achievable rate by using linear precoding in Section II-A can be easily formulated. From (5), the SINR for the $k$th user is

$$\gamma_k = \frac{\lambda_k^2 \left| \mathbf{g}_k^H \mathbf{f}_{G,k} \right|^2}{\frac{1}{\rho_1} \left| \mathbf{g}_k^H \mathbf{f}_{G,k} \right|^2 + \frac{1}{\rho_1 \rho_2}} \qquad (15)$$

where $\lambda_k$ is the $k$th diagonal element of $\boldsymbol{\Sigma}$. Thus the ergodic achievable rate for the entire system with perfect CSI equals

$$R = \frac{N}{2} E \left\{ \log_2 \left(1 + \gamma_k\right) \right\} \qquad (16)$$

where the factor $1/2$ results from the fact that data is transmitted over two time-slots.

For the scenario with imperfect CSI considering both channel estimation and quantized CSI feedback, we derive a lower bound on the mutual information for the $k$th user, which is indicated by $\tilde{R}_k \triangleq \frac{1}{2}\mathcal{I}\left(x_k; \tilde{y}_k, \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right)$. The following theorem gives a lower bound on the achievable rate under imperfect CSI.

*Theorem 1:* The ergodic achievable rate for the $k$th user with imperfect CSI can be lower bounded by

$$\tilde{R}_k \geqslant \frac{1}{2} E \left\{ \log_2 \left( 1 + \frac{\left| \tilde{\mathbf{g}}_k^H \hat{\mathbf{F}}_G \tilde{\mathbf{U}}^H \tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k} \right|^2}{T} \right) \right\} \qquad (17)$$

where $T$ is defined by (37) in Appendix A.

*Proof:* See Appendix A. ∎

From *Theorem* 1, it is easy straightforward to obtain the lower bound of the achievable rate for the entire system given by the following lemma.

*Lemma 1:* The achievable rate of the entire system with imperfect CSI can be lower bounded by

$$\begin{aligned}\tilde{R} = \sum_{k=1}^{N} \tilde{R}_k &\geq \frac{N}{2} E \left\{ \log_2 \left( 1 + \frac{\left| \tilde{\mathbf{g}}_k^H \hat{\mathbf{F}}_G \tilde{\mathbf{U}}^H \tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k} \right|^2}{T} \right) \right\} \\ &\triangleq \tilde{R}_{lower}\end{aligned} \qquad (18)$$

Here, from (16) and (18), we can quantify the effects of imperfect CSI by characterizing the rate loss as follows:

$$\Delta R \triangleq \frac{1}{N}\left(R - \tilde{R}\right). \qquad (19)$$

Note that it is difficult to analyze $\Delta R$ with exact closed form expressions. Hence, we resort to deriving an upper bound to $\Delta R$, which will be shown useful to offer some guidelines for the system design. Now, by using *Theorem* 1, we are able to derive a useful upper bound to the rate loss $\Delta R$ in (19). The rate loss is characterized in *Theorem* 2.

*Theorem 2:* The rate loss per user due to the channel estimation error and limited feedback for high SNRs can be upper bounded by

$$\Delta R \leqslant \frac{1}{2}\log_2 \left(1 + \rho_1 \rho_2 \left(\alpha_1 \bar{\varepsilon} + \alpha_2 \bar{\tau} + \alpha_3\right)\right) + O\left(1\right) \triangleq \Delta R_h \qquad (20)$$

where $\bar{\varepsilon}$ and $\bar{\tau}$ are quantization errors, $\bar{\varepsilon} = \frac{M-1}{M}2^{-\frac{B_1}{(M-1)}}$, and $\bar{\tau} = \frac{N-1}{N}2^{-\frac{B_2}{(N-1)}}$. In (20), notations $\alpha_1$, $\alpha_2$, and $\alpha_3$ are

defined by (54), (55), and (56) in Appendix B, respectively, and they are constants depending on fixed system parameters $P_1$, $P_2$, $\beta_1$, $\beta_2$, $M$, and $N$.

*Proof:* See Appendix B. ∎

From *Theorem 2*, we have the following remark:

*Remark 1:* A special case is $\delta_1^2 = \delta_2^2 = 0$ when there is no estimation error. In this case, we obtain that $\alpha_1 = \frac{M(N-1)}{M-1}$, $\alpha_2 = MN + N/\rho_1$, and $\alpha_3 = 0$. Then *Theorem 2* reduces to Theorem 1 in our previous work [22] with perfect channel estimation.

*Remark 2:* The rate loss increases with increasing transmit powers $P_1$ and $P_2$, while it decreases with increasing numbers of feedback bits $B_1$ and $B_2$, and it also decreases with increasing numbers of pilot symbols $\beta_1$ and $\beta_2$. In detail, the quantization errors $\bar{\varepsilon}$ and $\bar{\tau}$ decrease exponentially with $B_1$ and $B_2$; while the variances of estimation errors $\delta_1^2$ and $\delta_2^2$ are inversely proportional to $\beta_1$ and $\beta_2$. Hence, we can observe that the rate loss decreases more quickly with the increment of feedback bits than with that of the pilot symbols.

We can obtain large achievable rate in the high SNR regions for the MIMO relay system, and imperfect CSI mainly affects these regions. However, we can not obtain the exact system performance in the high SNR regions, we will analyze the asymptotic performance in the following lemma.

*Lemma 2:* As $P_2 = \kappa P_1$ grows large, the upper bound of rate loss in the high SNR region can be characterized by

$$\Delta R_\infty = \frac{1}{2}\log_2\left(1 + \varrho\right) + O\left(1\right) \qquad (21)$$

where $\varrho$ is defined by (62) in Appendix C.

*Proof:* See Appendix C. ∎

In order to analyze the effect of channel estimation on the rate loss, we consider the special case with no quantization error. For the special case, the quantization errors $\bar{\varepsilon}$ and $\bar{\tau}$ are equal to zero, thus, we have $\varrho = \frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)$. In the high SNR regions, the term $O\left(1\right)$ can be neglected. Then, the behavior of $\Delta R_\infty$ with no quantization error is characterized by the following lemma.

*Lemma 3:* As $P_2 = \kappa P_1$ grows large, the upper bound of rate loss in the high SNRs with no quantization error can be written by

$$\lim_{\substack{B_1 \to +\infty \\ B_2 \to +\infty}} \Delta R_\infty = \frac{1}{2}\log_2\left(1 + \frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)\right) \triangleq \Delta R_\infty^B . \qquad (22)$$

From (22), it is observed that the upper bound of rate loss in the high SNR regions decreases with $\beta_1$ and $\beta_2$, and increases with $\kappa = \frac{P_2}{P_1}$ when there is no quantization error.

Fig. 2 shows the lower bound of the achievable sum rate for estimated channel with no quantization error and perfect CSI for a system with $\beta_1 = \beta_2 = 1$, $\beta_1 = \beta_2 = 3$, and $P_2 = \kappa P_1$. In the system, both BS and RS deploy four antennas, i.e., $M = N = 4$, and there are four users, i.e., $K = 4$. In this figure, $\tilde{R}_{lower}$ is defined by (18), while in the simulation, the term $T$ in (18) is calculated by its closed-form expression $T_1$ defined by (38) in Appendix B. For this special case, in (38), it is noted that $\hat{\mathbf{v}}_{e,i} = \tilde{\mathbf{v}}_i$, $\tilde{\mathbf{G}} = [\tilde{\mathbf{g}}_1, \cdots, \tilde{\mathbf{g}}_N]^H$, $\hat{\mathbf{F}}_G = \bar{\mathbf{F}}_G = [\tilde{\mathbf{f}}_{G,1}, \cdots, \tilde{\mathbf{f}}_{G,N}]$, and $\tilde{\mathbf{f}}_{G,k}$ is the normalized $k$th column of $\tilde{\mathbf{G}}^H\left(\tilde{\mathbf{G}}\tilde{\mathbf{G}}^H\right)^{-1}$. From this figure, it is observed that the
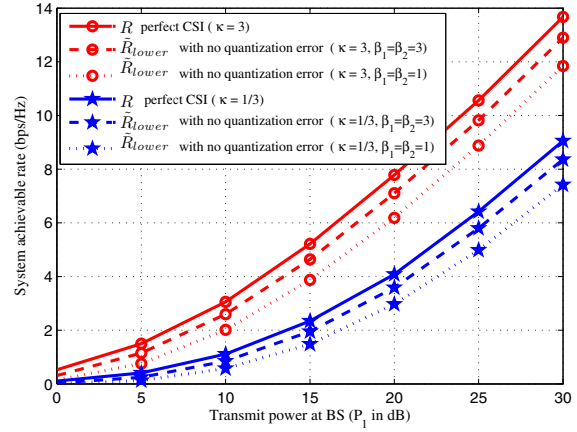


Fig. 2. The achievable rate with perfect CSI and estimated CSI with no quantization error for $M = N = K = 4$ and $P_2 = \kappa P_1$ with increasing transmit power.

rate loss gap between perfect CSI and the estimated channel decreases with $\beta_1$ and $\beta_2$ and increases with $\kappa$ in the whole SNR regions, which coincides with $\Delta R_\infty^B$ in (22). And the gap is upper bounded by $N \cdot \Delta R_\infty^B$ bps/Hz, thus, *Lemma 3* is verified by the numerical results.

### B. Upper Bound of the Achievable Rate

An upper bound of the achievable rate is directly derived by considering perfect CSIR at both RS and users, i.e., $\tilde{\mathbf{H}} = \mathbf{H}$ and $\tilde{\mathbf{g}}_k = \mathbf{g}_k$. The upper bound is given in the following lemma.

*Lemma 4:* The achievable rate for the system with both the channel estimation error and limited feedback can be upper bounded by

$$\tilde{R} \leqslant \frac{N}{2}E\left\{\log_2\left(1 + \bar{\gamma}_k\right)\right\} \triangleq \tilde{R}_{upper} \qquad (23)$$

where $\bar{\gamma}_k$ is the SINR of the $k$th user with limited feedback and no estimation error, and $\bar{\gamma}_k$ is defined by (63) in Appendix D.

*Proof:* See Appendix D. ∎

From [22], it is observed that $\tilde{R}_{upper}$ in above lemma is equal to $R_Q$ in [22, eq.(24)], and it means that $\tilde{R}_{upper}$ is also the achievable sum rate for the scheme proposed in [22]. It means that the achievable sum rate of this work is upper bounded by that of [22].

## IV. FEEDBACK AND PILOT DESIGN

In this section, we will specify the required numbers of feedback bits for the multiuser relay system to maintain a certain rate loss gap according to the rate loss upper bound. Moreover, considering the channel estimation, we also derive an optimized $\beta_1$ and $\beta_2$ for joint BS-RS pilot design via maximizing the lower bound of the achievable rate in the high SNR regions.

## A. Scaling Feedback

It is known in [14] that, due to the "interference-limited effect" phenomenon, the system performance will be upper bounded in a limited feedback system when SNR grows large. In order to let the rate performance always increase with the system SNR, one needs to assign more feedback bits for CSI quantization. *Theorem* 2 has predicted how the rate loss varies with a growing SNR and the numbers of feedback bits. The following theorem allows us to further specify a sufficient scaling of the feedback bit numbers $B_1$ and $B_2$ to maintain a bounded rate loss gap.

*Theorem 3:* In order to maintain the rate offset no larger than $\frac{1}{2}\log_2 b$ per user, it is sufficient to scale $B_1$ and $B_2$ according to (24) and (25), respectively. Specially, as $P_2 = \kappa P_1 \to +\infty$, $B_1$ and $B_2$ can be scaled according to simpler expressions (26) and (27).

$$B_1 = (M-1)\left[\log_2(\rho_1\rho_2\alpha_1) - \log_2\left(\frac{b-1-\rho_1\rho_2\alpha_3}{2}\right)\right]$$
$$+ (M-1)\log_2\left(\frac{M-1}{M}\right) \tag{24}$$

$$B_2 = (N-1)\left[\log_2(\rho_1\rho_2\alpha_2) - \log_2\left(\frac{b-1-\rho_1\rho_2\alpha_3}{2}\right)\right]$$
$$+ (N-1)\log_2\left(\frac{N-1}{N}\right) \tag{25}$$

$$B_1^\infty = (M-1)\left(\log_2 P_2 - \log_2 M\right)$$
$$+ (M-1)\log_2\left(\frac{2(N-1)}{N(b-1) - \frac{N}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)}\right) \tag{26}$$

$$B_2^\infty = (N-1)\log_2 P_2$$
$$+ (N-1)\log_2\left(\frac{2(N-1)}{N(b-1) - \frac{N}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)}\right). \tag{27}$$

*Proof:* See Appendix E. ∎

*Remark 3:* In the high SNR regions, (26) and (27) reveal that the numbers of feedback bits increases with $P_2$ (the SNR at the RS) and $\kappa$ (the ratio of the SNR at the RS and the SNR at the BS), and decrease with $\beta_1$ and $\beta_2$ (the numbers of pilots). And in (26) and (27), the terms $(M-1)\log_2 P_2$ and $(N-1)\log_2 P_2$ dominate $B_1^\infty$ and $B_2^\infty$, respectively. This indicates that the numbers of feedback bits increase approximately linearly with $\log_2 P_2$ in the high SNR regions.

*Remark 4:* Further, we consider a special case when there is no channel estimation error, i.e., $\delta_1^2 = \delta_2^2 = 0$. With this case, we obtain that $\alpha_1 = \frac{M(N-1)}{M-1}$, $\alpha_2 = MN + N/\rho_1$, and $\alpha_3 = 0$, then our derived *Theorem 3* and *Theorem 4* reduce to Theorem 1 and Theorem 3 in [22], respectively.

## B. Optimal Number of Pilot Symbols

By substituting (21) into (19), the lower bound of the achievable rate in the high SNRs is specified as (28) on the top of the next page.

Firstly, we analyze the tightness of $\tilde{R}_{lower}^\infty$. However, it is quite difficult to stringently characterize the tightness of the
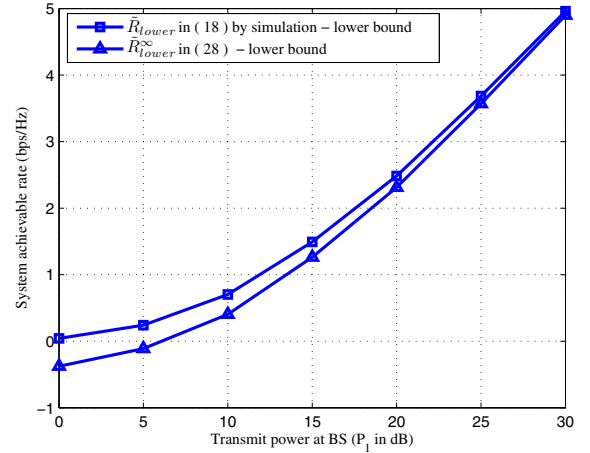


Fig. 3.    The lower bound of achievable rate for $M = N = K = 2$, $P_2 = P_1 - 4$ dB, $B_1 = B_2 = 10$ with increasing transmit power.

bound. Hence, in order to address the tightness in a qualitative way, we focus on the high SNR regime which is the most interest in our work. In (28), only the term $\Delta R_\infty$ is affected by the numbers of feedback bits ($B_1$ and $B_2$) and the numbers of pilot symbols ($\beta_1$ and $\beta_2$). Thus, the main term affecting the tightness of $\tilde{R}_{lower}^\infty$ is $\Delta R_\infty$ in (21).

From the system viewpoint, if $B_1$, $B_2$, $\beta_1$, and $\beta_2$ are large enough in the high SNR regimes, CSIT and CSIR are approximately perfect, and the rate loss is approximately equal to zero. While, for the above similar scenario, the terms $2^{-\frac{B_1}{M-1}}$, $2^{-\frac{B_2}{N-1}}$, and $\frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)$ in (21) converge to zero, and thus, $\Delta R_\infty$ converges to zero. From the above analysis, it is observed that $\Delta R_\infty$ becomes tighter as $B_1$, $B_2$, $\beta_1$, and $\beta_2$ increase. According to (26) and (27) in *Theorem* 3, $B_1$ and $B_2$ scale linearly with $\log_2 P_2$, and they are large in the high SNR regions. Therefore, $\tilde{R}_{lower}^\infty$ is also tight in the high SNR regions.

Then, we provide some simulation results in Fig. 3 to test the tightness. Fig. 3 considers a system with $P_2 = P_1 - 4$ dB, $\beta_1 = \beta_2 = 3$, and $B_1 = B_2 = 10$. In the system, both BS and RS deploy two antennas, i.e., $M = N = 2$, and there are two users, i.e., $K = 2$. In Fig. 3, $\tilde{R}_{lower}$ defined in (18) is the lower bound of the achievable sum rate, and $\tilde{R}_{lower}^\infty$ defined in (28) is the lower bound for the high SNR regions. From this figure, it is observed that $\tilde{R}_{lower}^\infty$ converges to $\tilde{R}_{lower}$ with increasing transmit power. Thus, $\tilde{R}_{lower}^\infty$ becomes tighter for the high SNR regimes.

From *Lemma 2*, we observe that $\Delta R_\infty$ is a function of $\beta_1$ and $\beta_2$. In the following lemma, we will find the optimal $\beta_1$ and $\beta_2$ to obtain better performance.

*Lemma 5:* In order to maximize the lower bound of the achievable rate in the high SNR regions (i.e., $P_2 = \kappa P_1 \to +\infty$) under the condition that $\beta_1 + \beta_2 = C_\beta$, we use the Lagrange multiplier method to obtain the optimal $\beta_1$ and $\beta_2$ as

$$\beta_{1\_opt} = \frac{1}{\sqrt{\frac{M}{\kappa}} + 1}C_\beta , \tag{29}$$

$$\tilde{R}_{lower}^{\infty} \triangleq R - N\Delta R_{\infty} = R - \frac{N}{2}\log_2\left(\frac{P_2}{N}\frac{N-1}{M}2^{-\frac{B_1}{M-1}} + P_2\frac{N-1}{N}2^{-\frac{B_2}{N-1}} + \left(1 + \frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)\right)\right) - O(N) \quad (28)$$
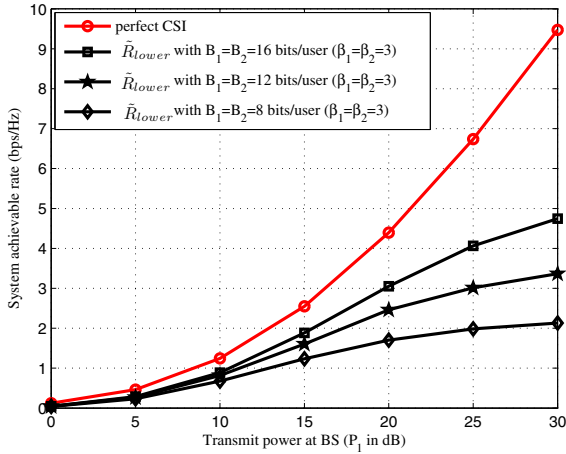


Fig. 4. The achievable rate with different number of feedback bits for $M = N = K = 4$, $\beta_1 = \beta_2 = 3$, $P_2 = P_1 - 4$ dB and $B_1 = B_2$ with increasing transmit power.



Fig. 5. The number of feedback bits calculated by two different equations for $M = N = K = 4$, $\beta_1 = \beta_2 = 3$, $P_2 = P_1 - 4$ dB and $b = 4$ with increasing transmit power.

and

$$\beta_{2\_opt} = \frac{\sqrt{\frac{M}{\kappa}}}{\sqrt{\frac{M}{\kappa}} + 1}C_\beta \quad (30)$$

*Proof:* See Appendix F. ∎

If both $\beta_1$ and $\beta_2$ are integers, pilot symbols can be orthogonal in time, i.e., $\beta_1$ pilots are successively transmitted from each of the $M$ BS antennas for a total of $\beta_1 M$ channel realizations, and $\beta_2$ pilots are successively transmitted from each of the $N$ RS antennas for a total of $\beta_2 N$ channel realizations.

For arbitrary values of $\beta_1$ and $\beta_2$, from [29], it is found sufficient for both $\beta_1 M$ and $\beta_2 N$ to be integers. Therefore, we should set $\beta_{1\_opt}$ or $\beta_{1\_opt} M$ to be an integer. In the same way, we can also set $\beta_{2\_opt}$ or $\beta_{2\_opt} N$ to be an integer for RS. Thus, if $\frac{1}{\sqrt{\frac{M}{\kappa}}+1}C_\beta$ is an integer, we set $\beta_{1\_opt} = \frac{1}{\sqrt{\frac{M}{\kappa}}+1}C_\beta$ and $\beta_{2\_opt} = \frac{\sqrt{\frac{M}{\kappa}}}{\sqrt{\frac{M}{\kappa}}+1}C_\beta$, which are defined by (29) and (30), respectively. If $\frac{1}{\sqrt{\frac{M}{\kappa}}+1}C_\beta$ is not an integer, we set

$$\beta_{1\_opt} = \frac{\left\lceil \frac{\frac{M}{\sqrt{\frac{M}{\kappa}}+1}C_\beta}{1} \right\rceil}{M} \text{ and } \beta_{2\_opt} = \frac{\left\lceil (C_\beta - \beta_{1\_opt})N \right\rceil}{N},$$ where the function $\lceil x \rceil$ stands for the smallest integer that is not less than $x$.

## V. NUMERICAL RESULTS

This section presents some numerical results for the relay system under various scenarios. It is noted that in the following figures, the lower bound of the achievable sum rate $\tilde{R}_{lower}$ is defined by (18), and for the simulation results, the term $T$ in (18) is calculated by its closed-form expression $T_1$ defined by (38) in Appendix B.
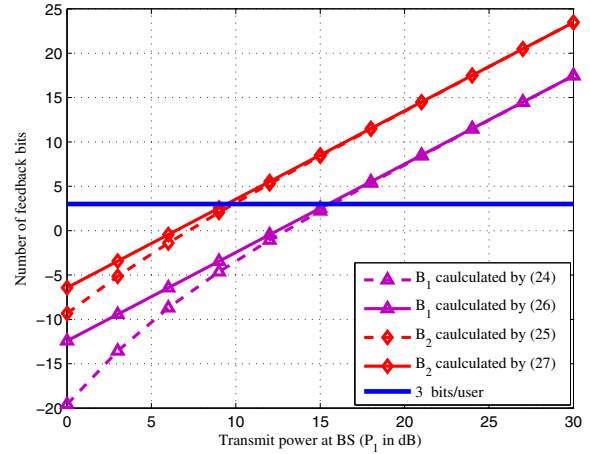
Firstly, we investigate how the quantization level influences the system performance. Regardless of how many feedback bits ($B_1$ and $B_2$) are used, the system eventually becomes interference limited because both interference and signal power scale linearly with $P_1$ and $P_2$. Fig. 4 shows the achievable rate with fixed $B_1$ and $B_2$ for a system with $\beta_1 = \beta_2 = 3$, $P_2 = P_1 - 4$ dB, and $B_1 = B_2$. In the system, both BS and RS deploy four antennas, i.e., $M = N = 4$, and there are four users, i.e., $K = 4$. As shown in this figure, the performance of the limited feedback is very close to the perfect CSI when the SNRs are small, while the rate loss gap between the limited feedback and perfect CSI becomes larger as the SNRs increase. As the numbers of feedback bits increase, the gap becomes smaller. Thus, we can improve the quality of feedback to get larger achievable rate. In Fig. 5, we show the scaled numbers of feedback bits to maintain the rate loss no larger than $\frac{1}{2}\log_2 b$ bps/Hz/user for a system with $\beta_1 = \beta_2 = 3$, $P_2 = P_1 - 4$ dB and $b = 4$. In the system, both BS and RS deploy four antennas, i.e., $M = N = 4$, and there are four users, i.e., $K = 4$. The scaled numbers $B_1$ and $B_2$ are calculated by (24), (25), and simple expressions (26), (27) in *Theorem* 3, respectively. Note that if $B_1$ or $B_2$ is smaller than three, its value is set to three. In the high SNR regions, it is observed that $B_1$ calculated by (24) and (26) are also very close when $P_1$ is larger than 15 dB, while $B_2$ calculated by (25) and (27) are very close when $P_1$ is larger than 9 dB. Therefore, in the high SNR regions, we can use the simple expressions (26) and (27) to respectively calculate $B_1$ and $B_2$.

Furthermore, we show that the achievable rate grows with the system SNRs by using scaled $B_1$ and $B_2$. We assume that $B_1$ and $B_2$ are scaled according to (24) and (25) in *Theorem* 3.
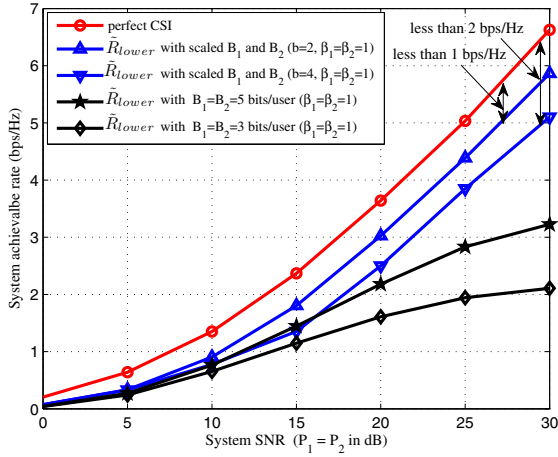
1) $M = N = K$. In Figs. 6–7, the achievable rates

Fig. 6. The achievable rate for $M = N = K = 2$, $P_1 = P_2$, $\beta_1 = \beta_2 = 1$, and $b = 2, 4$ with increasing SNR.
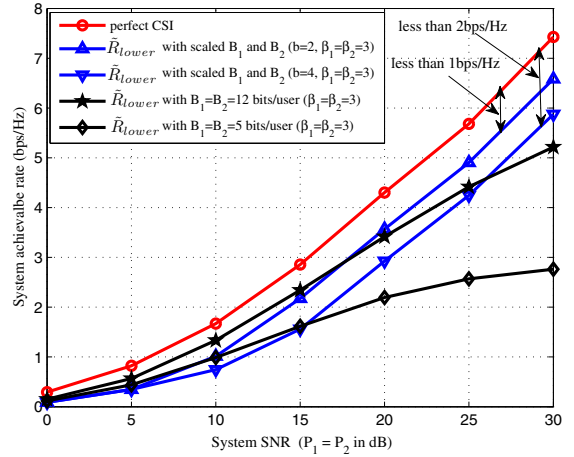


Fig. 8. The achievable rate for $M = 4, N = K = 2$, $P_2 = P_1$, $\beta_1 = \beta_2 = 3$, and $b = 2, 4$ with increasing SNR.
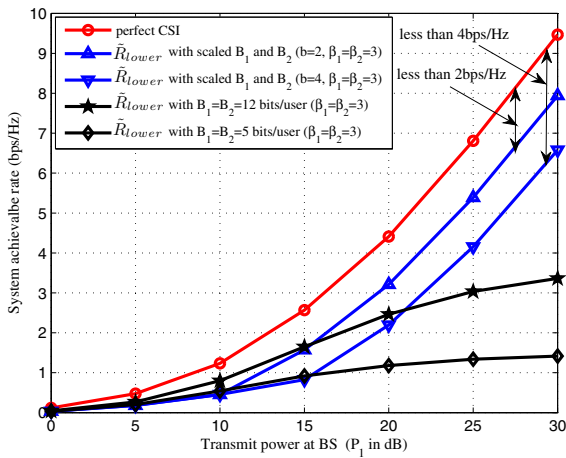


Fig. 7. The achievable rate for $M = N = K = 4$, $P_2 = P_1 - 4$ dB, $\beta_1 = \beta_2 = 3$, and $b = 2, 4$ with increasing transmit power.

are shown to compare perfect CSI, quantized feedback using scaled $B_1$ and $B_2$, and quantized feedback with fixed $B_1$ and $B_2$. Fig. 6 displays the achievable rate for a system with $P_1 = P_2$, $\beta_1 = \beta_2 = 1$, $b = 2$, and $b = 4$. In the system, both BS and RS deploy two antennas, i.e., $M = N = 2$, and there are two users, i.e., $K = 2$. For the system with a fixed $B_1 = B_2 = 3$ and $B_1 = B_2 = 5$, the rate loss due to quantized feedback is unbounded and grows with $P_1$ and $P_2$, thereby resulting in an upper bounded system rate when $P_1$ and $P_2$ increase. However, coinciding with *Theorem* 3, it is observed that the rate loss is smaller than 1 bps/Hz ($b = 2$) and 2 bps/Hz ($b = 4$) in the whole tested SNR regime by using the proposed scaled feedback strategy. Fig. 7 plots the achievable rate for a system with $P_2 = P_1 - 4$ dB, $\beta_1 = \beta_2 = 3$, $b = 2$, and $b = 4$. In the system, both BS and RS deploy four antennas, i.e., $M = N = 4$, and there are four users, i.e., $K = 4$. By using scaled $B_1$ and $B_2$, the system rate grows with the system SNRs and the rate loss is always maintained within 2 bps/Hz ($b = 2$) and 4 bps/Hz ($b = 4$) with increasing transmit power $P_1$ and $P_2$.

2) $M > N = K$. Fig. 8 demonstrates the results for $P_2 = P_1$, $\beta_1 = \beta_2 = 3$, and $b = 2, 4$. In the system, BS deploys four antennas, i.e., $M = 4$, RS deploys two antennas, i.e., $N = 2$, and there are two users, i.e., $K = 2$. By adopting feedback strategy in *Theorem* 3, the system rate loss is always controlled to remain within 1 bps/Hz ($b = 2$) and 2 bps/Hz ($b = 4$) in the whole SNR regions. Note that although the scaled $B_1$ and $B_2$ in *Theorem* 3 are derived by omitting the $O(1)$ term in the whole tested SNR regime, the results in the above three figures verify its accuracy and effectiveness for different tested cases.

Then, we compare the performance of this scheme with the scheme proposed in [22]. Fig. 9 shows the achievable rate with $P_1 = P_2$, $\beta_1 = \beta_2 = 1$, and $B_1 = B_2 = 18$. In the system, both BS and RS deploy four antennas, i.e., $M = N = 4$, and there are four users, i.e., $K = 4$. In Fig. 9, $\tilde{R}_{lower}$ in (18) is the lower bound of the achievable sum rate for our proposed scheme, while $\tilde{R}_{upper}$ in (23) is the upper bound of the achievable sum rate for the scheme. As detailed in *Lemma* 4, $\tilde{R}_{upper}$ is also the achievable sum rate for the scheme proposed in [22]. Due to considering the channel estimation errors, the achievable sum rate of this work is smaller than that of [22] in the low and medium SNR regions. However, in the high SNR regions, the system can obtain the approximately perfect CSIR. Thus, the performance in this work is very close to the performance of [22] in the high SNR regions. As it is shown in Fig. 9, $\tilde{R}_{lower}$ does converge to $\tilde{R}_{upper}$ with increasing SNRs.

In addition, the effect of $\kappa = \frac{P_2}{P_1}$ on the system performance is illustrated in Fig. 10. This figure displays the lower bound of achievable rate for a system with increasing $\kappa$ and constant total power, i.e., $P_1 + P_2 = 20$ dB and $P_1 + P_2 = 25$ dB. In the system, both BS and RS deploy four antennas, i.e., $M = N = 4$, and there are four users, i.e., $K = 4$. From the numerical results, it is observed that there exist different optimal values of $\kappa$ for different scenarios. For the system with $P_1 + P_2 = 20$ dB, $\beta_1 = \beta_2 = 3$, and $B_1 = B_2 = 15$, the optimal $\kappa$ is 1.9, while the optimal $\kappa$ is 2.5 for the system with $P_1 + P_2 = 25$ dB, $\beta_1 = \beta_2 = 3$, and $B_1 = B_2 = 18$.
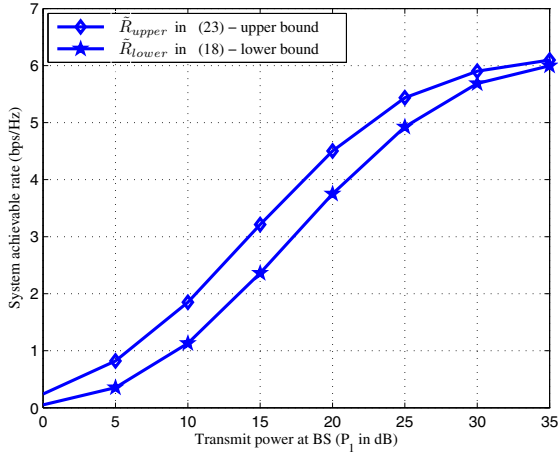
Fig. 9. The lower and upper bound of the achievable rate for $M = N = K = 4$, $P_2 = P_1$, $\beta_1 = \beta_2 = 1$, $B_1 = B_2 = 18$ with increasing SNR.



Fig. 11. The achievable rate of optimal number of pilots for $M = N = K = 4$, $P_2 = P_1/4$, $\kappa = 1/4$, and $C_\beta = \beta_1 + \beta_2 = 5$ with increasing numbers of feedback bits.
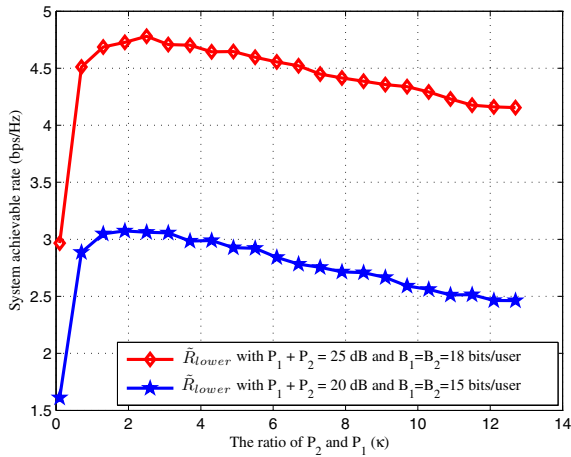


Fig. 10. The achievable rate for $M = N = K = 4$, $P_2 = \kappa P_1$, and $\beta_1 = \beta_2 = 3$ with increasing $\kappa$.

At last, we show that the achievable rate for a system with the optimal number of pilots. In Fig. 11, throughput curves are shown for a system with $P_2 = P_1/4$, $\kappa = 1/4$, and $C_\beta = \beta_1 + \beta_2 = 5$ as feedback bits increase. In the system, both BS and RS deploy four antennas, i.e., $M = N = 4$, and there are four users, i.e., $K = 4$. The optimal numbers of pilots calculated by (29) and (30) are $\beta_1 = 1$ and $\beta_2 = 4$ when $C_\beta = 5$. It is observed that we can obtain more achievable rate by optimizing the number of pilots in the whole tested SNRs regime.

## VI. CONCLUSIONS

In this paper, we have investigated the effects of various imperfect CSI on the performance of the multiuser AF MIMO relay downlink system. Compared with the system using perfect channel state information (CSI), we have analyzed the achievable rate loss due to CSI mismatch arising from both channel estimation and CSI quantization feedback. In order to quantitatively characterize the effect of imperfect CSI, we have derived an upper bound to the rate loss. Given the rate
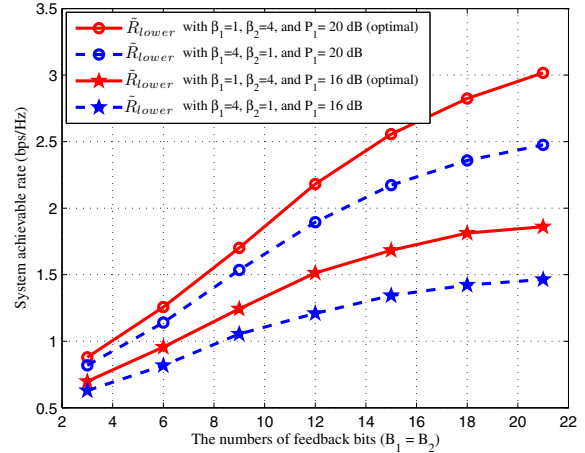
loss bound, we then presented a limited feedback strategy for both two-hop channels in the relay system to guarantee a bounded rate loss for growing SNRs. It's found that this newly presented strategy reveals the relationship among CSI quantization level and other system parameters, including the transmit power and the pilots for channel estimation. As a special case where there is no channel estimation error in the two-hop channels, we found that our derived *Theorem* 2 and *Theorem* 3 respectively reduce to the Theorem 1 and Theorem 3 in [22], and hence the model in this paper is more general. Besides these, we also have derived the optimal numbers of pilots by using the upper bound considering both channel estimation and quantization errors in the high SNR regions. This result can serve as a guideline for optimal channel estimation pilot design in practical applications.

## APPENDIX A
## PROOF OF THEOREM 1

The proof is inspired by Theorem 1 of [25] and Lemma B.0.1 of [30]. From (14), we have

$$\mathcal{I}\left(x_k; \tilde{y}_k, \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right) = h\left(x_k\right) - h\left(x_k \big| \tilde{y}_k, \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right)$$
$$= \log_2\left(\pi e\right) - h\left(x_k \big| \tilde{y}_k, \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right) \quad (31)$$

where the last equality holds because the entropy $h\left(x_k\right)$ equals to $\log_2\left(\pi e\right)$ as $x_k$ is a complex Gaussian input with unity variance [31]. Then, by further concerning the second term of the last equality in (31), the conditional entropy $h\left(x_k \big| \tilde{y}_k, \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right)$ satisfies

$$h\left(x_k \big| \tilde{y}_k, \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right) \overset{(a)}{=} h\left(\left(x_k - \theta \tilde{y}_k\right) \big| \tilde{y}_k, \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right)$$
$$\overset{(b)}{\leqslant} h\left(\left(x_k - \theta \tilde{y}_k\right) \big| \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right) \quad (32)$$

where (a) holds for any deterministic factor $\theta$ depending only on $\tilde{y}_k$, $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{g}}_k$, and (b) follows from the fact that conditioning reduces entropy. By substituting (32) into (31),

we have

$$\mathcal{I}\left(x_k; \tilde{y}_k, \tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right)$$
$$\geq \log_2\left(\pi e\right) - h\left(\left(x_k - \theta\tilde{y}_k\right)\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right)$$
$$\geq \log_2\left(\pi e\right) - E\left\{\log_2\left(\pi e \cdot E\left\{\left|x_k - \theta\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}\right)\right\}.$$
$$(33)$$

In (33), we use fact that Gaussian distributed random variable maximizes the entropy given a fixed variance [31]. In detail, if $\left(\left(x_k - \theta\tilde{y}_k\right)\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right)$ is a complex Gaussian random variable with its variance $E\left\{\left|x_k - \theta\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}$, the entropy $h\left(\left(x_k - \theta\tilde{y}_k\right)\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right)$ is thus maximized to achieve the second term in the last inequality, which yields (33).

In order to obtain a tighter bound in (33), we need to select a proper value of the scaling factor $\theta$ to maximize the right hand side of (33), equivalently to minimize $E\left\{\left|x_k - \theta\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}$. This corresponds to using a linear MMSE estimator to $x_k$ given $\tilde{y}_k$, $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{g}}_k$. According to [25], $\theta$ is determined by

$$\theta = \frac{E\left\{x_k\tilde{y}_k^*\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}}{E\left\{\left|\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}} \qquad (34)$$

where $(\cdot)^*$ denotes the conjugation of the complex input. By substituting (14) in Section II-C into $E\left\{x_k\tilde{y}_k^*\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}$ and $E\left\{\left|\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}$, and using the fact that $\mathbf{E}$, $\mathbf{e}_k$, $\mathbf{x}$, $\tilde{\mathbf{n}}$, and $z_k$ are independent, zero-mean Gaussian random variables, we have

$$E\left\{x_k\tilde{y}_k^*\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\} = \sqrt{\rho_1\rho_2}\left(\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right)^*, \qquad (35)$$

and

$$E\left\{\left|\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\} = \rho_1\rho_2\left(\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right|^2 + T\right) \quad (36)$$

where $T$ is defined by

$$T = \frac{1}{\rho_1}E_{\mathbf{e}_k}\left\{\left\|\mathbf{g}_k^H\hat{\mathbf{F}}_G\right\|^2\right\} + E_{\mathbf{e}_k}\left\{\left|\mathbf{e}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right|^2\right\}$$
$$+ \frac{1}{\rho_1\rho_2} + E_{\mathbf{E},\mathbf{e}_k}\left\{\left\|\mathbf{g}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\mathbf{E}\hat{\mathbf{V}}_e\right\|^2\right\}$$
$$+ \sum_{j\neq k}E_{\mathbf{e}_k}\left\{\left|\mathbf{g}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,j}\right|^2\right\}. \qquad (37)$$

Then, with simple calculations, the closed-form of $T$ is written by

$$T_1 = \delta_2^2\left\|\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right\|^2 + N\delta_1^2\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right\|^2 + N^2\delta_1^2\delta_2^2$$
$$+ \sum_{j\neq k}\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,j}\right|^2 + \sum_{j\neq k}\delta_2^2\left\|\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,j}\right\|^2$$
$$+ \frac{1}{\rho_1}\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right\|^2 + \frac{1}{\rho_1}N\delta_2^2 + \frac{1}{\rho_1\rho_2}. \qquad (38)$$

From (34), (35), and (36), we have the corresponding MMSE as

$$E\left\{\left|x_k - \theta\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}$$
$$= E\left\{\left|x_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\} - \frac{\left|E\left\{x_k\tilde{y}_k^*\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}\right|^2}{E\left\{\left|\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}}$$
$$= 1 - \frac{\rho_1\rho_2\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right|^2}{E\left\{\left|\tilde{y}_k\right|^2\left|\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\right.\right\}} = \frac{T}{\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right|^2 + T}. \qquad (39)$$

Now that, by substituting (39) into (33), we have (17), which completes the proof.

## APPENDIX B
## PROOF OF THEOREM 2

By substituting (16) and (18) in *Theorem* 1 into (19), the rate loss $\Delta R$ is upper bounded by (40) on the top of the next page. In (40), the last equality comes from the fact that $\hat{\mathbf{f}}_{G,k}$ and $\mathbf{f}_{G,k}$ follow the same distribution. In the following, we separately calculate $\Delta_a$ and $\Delta_b$.

From (37) and (40), $\Delta_a$ is bounded by (41) on the top of the next page. In (41), $\hat{\mathbf{V}} = [\hat{\mathbf{v}}_1, \cdots, \hat{\mathbf{v}}_N]$ denotes the quantization of $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_N]$. Inequality (a) holds by dropping some positive terms in the denominator and using the equality $\sum_{j\neq k}E_{\mathbf{e}_k}\left\{\left|\mathbf{g}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,j}\right|^2\right\} = \sum_{j\neq k}\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,j}\right|^2 + \sum_{j\neq k}E_{\mathbf{e}_k}\left\{\left|\mathbf{e}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,j}\right|^2\right\}$, and the facts that $\mathbf{g}_k$ and $\tilde{\mathbf{H}}$ are statistically equivalent to $\frac{1}{\sqrt{1-\delta_2^2}}\tilde{\mathbf{g}}_k$ and $\sqrt{1-\delta_1^2}\mathbf{H}$, respectively. In (b), we use the fact that $\lambda_k^2\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 = \left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\mathbf{\Sigma}\mathbf{V}^H\mathbf{v}_k\right|^2$, and drop some positive terms in the numerator. Inequality (c) comes from the inequality $\mathbf{v}_k\mathbf{v}_k^H - \hat{\mathbf{v}}_k\hat{\mathbf{v}}_k^H \preceq \sqrt{\phi_k}\mathbf{I}_M$ [22], and dropping some positive terms in the denominator. In (41), we obtain (d) by using the generalized Rayleigh quotient theorem [32] and the Jensen's inequality, where $\phi_k = 1 - \left|\hat{\mathbf{v}}_k^H\hat{\mathbf{v}}_k\right|^2$, $\lambda_{\min}(\cdot)$ is the minimum eigenvalue of a matrix, and matrix $\mathbf{Q}$ is defined by

$$\mathbf{Q} = \mathbf{V}^H\hat{\mathbf{V}}\hat{\mathbf{V}}^H\mathbf{V}. \qquad (42)$$

Then we calculate $\Delta_b$ and derive (43) (on the top of the next page) from (37) and (40). In (43), (a) holds for the fact that $\mathbf{g}_k$ is statistically equivalent to $\frac{1}{\sqrt{1-\delta_2^2}}\tilde{\mathbf{g}}_k$, and the inequality follows from $\frac{1}{\rho_1\left(1-\delta_2^2\right)}\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 \geq 0$ and the Jensen's inequality. By respectively deriving the upper bounds to the five variables $\Delta_{b1}$, $\Delta_{b2}$, $\Delta_{b3}$, $\Delta_{b4}$, and $\Delta_{b5}$ in (43), we can obtain an upper bound on $\Delta_b$. Before calculating them, we need to recall some preliminary results as in (44) and (45) from [22], (46), and (47).

$$E_{\mathbf{V}_e}\left\{\mathbf{V}_e^H\hat{\mathbf{v}}_{e,k}\hat{\mathbf{v}}_{e,k}^H\mathbf{V}_e\right\} = (1-\bar{\varepsilon})\mathbf{s}_k\mathbf{s}_k^H + \sum_{j\neq k}\frac{1}{M-1}\bar{\varepsilon}\mathbf{s}_j\mathbf{s}_j^H \qquad (44)$$

$$2\Delta R \leqslant E\left\{\log_2\left(\frac{\frac{1}{\rho_1}\left|\mathbf{g}_k^H\mathbf{f}_{G,k}\right|^2 + \lambda_k^2\left|\mathbf{g}_k^H\mathbf{f}_{G,k}\right|^2 + \frac{1}{\rho_1\rho_2}}{\frac{1}{\rho_1}\left|\mathbf{g}_k^H\mathbf{f}_{G,k}\right|^2 + \frac{1}{\rho_1\rho_2}}\right)\right\} - E\left\{\log_2\left(\frac{T + \left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right|^2}{T}\right)\right\}$$

$$= \underbrace{E\left\{\log_2\left(\frac{\frac{1}{\rho_1}\left|\mathbf{g}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 + \lambda_k^2\left|\mathbf{g}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 + \frac{1}{\rho_1\rho_2}}{T + \left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right|^2}\right)\right\}}_{\Delta_a} + \underbrace{E\left\{\log_2\left(\frac{T}{\frac{1}{\rho_1}\left|\mathbf{g}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 + \frac{1}{\rho_1\rho_2}}\right)\right\}}_{\Delta_b} \tag{40}$$

$$\Delta_a \overset{(a)}{\leqslant} E\left\{\log_2\left(\frac{1}{(1-\delta_2^2)}\left(\frac{1}{\rho_1}\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 + \lambda_k^2\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 + \frac{1-\delta_2^2}{\rho_1\rho_2}\right)\right)\right\}$$
$$- E\left\{\log_2\left((1-\delta_1^2)\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\mathbf{V}^H\hat{\mathbf{V}}\right\|^2 + \frac{1}{\rho_1}\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right\|^2 + \frac{1}{\rho_1\rho_2}\right)\right\}$$

$$\overset{(b)}{\leqslant} \log_2\left(\frac{1}{1-\delta_2^2}\right) + E\left\{\log_2\left(\begin{array}{l} 1 + \dfrac{\delta_1^2\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\mathbf{V}^H\mathbf{v}_k\mathbf{v}_k^H\mathbf{V}\boldsymbol{\Sigma}\left(\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right)^H}{(1-\delta_1^2)\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\mathbf{V}^H\hat{\mathbf{V}}\right\|^2 + \frac{1}{\rho_1}\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right\|^2 + \frac{1}{\rho_1\rho_2}} \\[4ex] + \dfrac{(1-\delta_1^2)\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\mathbf{V}^H\left(\mathbf{v}_k\mathbf{v}_k^H - \hat{\mathbf{v}}_k\hat{\mathbf{v}}_k^H\right)\mathbf{V}\boldsymbol{\Sigma}\left(\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right)^H}{(1-\delta_1^2)\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\mathbf{V}^H\hat{\mathbf{V}}\right\|^2 + \frac{1}{\rho_1}\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right\|^2 + \frac{1}{\rho_1\rho_2}} \end{array}\right)\right\}$$

$$\overset{(c)}{\leqslant} \log_2\left(\frac{1}{1-\delta_2^2}\right) + E\left\{\log_2\left(1 + \frac{\delta_1^2\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\mathbf{V}^H\hat{\mathbf{v}}_k\hat{\mathbf{v}}_k^H\mathbf{V}\boldsymbol{\Sigma}\left(\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right)^H}{(1-\delta_1^2)\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\mathbf{V}^H\hat{\mathbf{v}}_k\right|^2} + \frac{\sqrt{\phi_k}\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\boldsymbol{\Sigma}\left(\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\right)^H}{(1-\delta_1^2)\left\|\tilde{\mathbf{g}}_k^H\hat{\mathbf{F}}_G\boldsymbol{\Sigma}\mathbf{V}^H\hat{\mathbf{V}}\right\|^2}\right)\right\}$$

$$\overset{(d)}{\leqslant} \frac{1}{2}\log_2\left(\frac{1}{1-\delta_1^2}\right) + \log_2\left(\frac{1}{1-\delta_2^2}\right) + \frac{1}{2}\log_2\left(1 + E\left\{\frac{\sqrt{\phi_k}}{\lambda_{\min}(\mathbf{Q})}\right\}\right) \tag{41}$$

$$\Delta_b \overset{(a)}{=} E\{\log_2 T\} - E\left\{\log_2\left(\frac{1}{\rho_1(1-\delta_2^2)}\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 + \frac{1}{\rho_1\rho_2}\right)\right\}$$

$$= E\left\{\log_2\left(1 + \frac{T - \frac{1}{\rho_1(1-\delta_2^2)}\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 - \frac{1}{\rho_1\rho_2}}{\frac{1}{\rho_1(1-\delta_2^2)}\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 + \frac{1}{\rho_1\rho_2}}\right)\right\}$$

$$\leqslant \log_2\left(\begin{array}{l} 1 + \rho_2\underbrace{\sum_{j\neq k}E\left\{\left|\mathbf{g}_k^H\hat{\mathbf{f}}_{G,j}\right|^2\right\}}_{\Delta_{b1}} + \rho_1\rho_2\underbrace{E\left\{\left\|\mathbf{g}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\mathbf{E}\hat{\mathbf{V}}_e\right\|^2\right\}}_{\Delta_{b2}} \\[4ex] + \rho_1\rho_2\underbrace{\sum_{j\neq k}E\left\{\left|\mathbf{g}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,j}\right|^2\right\}}_{\Delta_{b3}} + \rho_1\rho_2\underbrace{E\left\{\left|\mathbf{e}_k^H\hat{\mathbf{f}}_G\tilde{\mathbf{U}}^H\tilde{\mathbf{H}}\hat{\mathbf{v}}_{e,k}\right|^2\right\}}_{\Delta_{b4}} + \rho_2\underbrace{E\left\{\left|\mathbf{e}_k^H\hat{\mathbf{f}}_{G,k}\right|^2\right\}}_{\Delta_{b5}} \end{array}\right) \tag{43}$$

$$\sum_{j\neq k}E_{\mathbf{V}_e}\left\{\mathbf{V}_e^H\hat{\mathbf{v}}_{e,j}\hat{\mathbf{v}}_{e,j}^H\mathbf{V}_e\right\} = \frac{N-1}{M-1}\bar{\varepsilon}\mathbf{s}_k\mathbf{s}_k^H$$
$$+ \sum_{j\neq k}\left(1 - \frac{M-N+1}{M-1}\bar{\varepsilon}\right)\mathbf{s}_j\mathbf{s}_j^H \tag{45}$$

$$E\left\{\left|\mathbf{g}_k^H\hat{\mathbf{f}}_{G,j}\right|^2\right\} = E\left\{\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,j}\right|^2\right\} + E\left\{\left|\mathbf{e}_k^H\hat{\mathbf{f}}_{G,j}\right|^2\right\}$$
$$\overset{(a)}{=} \frac{N(1-\delta_2^2)\bar{\tau}}{N-1} + \delta_2^2 \tag{46}$$

$$E\left\{\left|\mathbf{g}_k^H\hat{\mathbf{f}}_{G,k}\right|^2\right\} = E\left\{\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,k}\right|^2\right\} + E\left\{\left|\mathbf{e}_k^H\hat{\mathbf{f}}_{G,k}\right|^2\right\} = 1 \tag{47}$$

where $\bar{\varepsilon} = E\left\{1 - \left|\mathbf{v}_j^H\hat{\mathbf{v}}_j\right|^2\right\} = \frac{M-1}{M}2^{-\frac{B_1}{(M-1)}}$ and $\bar{\tau} = E\left\{1 - \left|\left(\frac{\tilde{\mathbf{g}}_k}{\|\tilde{\mathbf{g}}_k\|^2}\right)^H\hat{\mathbf{g}}_k\right|\right\} = \frac{N-1}{N}2^{-\frac{B_2}{(N-1)}}$ [22]. In (46), (a) holds because $E\left\{\left|\tilde{\mathbf{g}}_k^H\hat{\mathbf{f}}_{G,j}\right|^2\right\} = \frac{N(1-\delta_2^2)\bar{\tau}}{N-1}$ [22]. Note that, $\mathbf{s}_i$ in (45) represents the vector with 1 at the $i$th element and zeros elsewhere.

By using (46), $\Delta_{b1}$ in (43) is rewritten by

$$\Delta_{b1} = \rho_2 N \left(1 - \delta_2^2\right) \bar{\tau} + \rho_2 \left(N - 1\right) \delta_2^2 . \qquad (48)$$

Concerning $\Delta_{b2}$ in (43), we have (49) at the bottom of the next page. In (49), $\tilde{\mathbf{G}} = [\tilde{\mathbf{g}}_1, \cdots, \tilde{\mathbf{g}}_N]^H$. Due to $E_{\mathbf{E}} \left\{ \left(\mathbf{E}\hat{\mathbf{V}}_e\right) \left(\mathbf{E}\hat{\mathbf{V}}_e\right)^H \right\} = \mathrm{tr}\left(\hat{\mathbf{V}}_e \hat{\mathbf{V}}_e^H\right) \delta_2^2 \mathbf{I}_N = \mathrm{tr}\left(\hat{\mathbf{V}}_e^H \hat{\mathbf{V}}_e\right) \delta_2^2 \mathbf{I}_N = N\delta_2^2 \mathbf{I}_N$, inequality (a) in (49) establishes. By substituting (46) and (47) into (a), we have (b).

Now, we calculate $\Delta_{b3}$ in (43) as (50) at the bottom of the next page. In (50), (a) uses (45), and (b) holds because $E_{\tilde{\boldsymbol{\Sigma}}}\left\{\tilde{\boldsymbol{\Sigma}}^2\right\} = M\left(1 - \delta_1^2\right)\mathbf{I}_N$ and using (46) and (47).

Then, $\Delta_{b4}$ in (43) is calculated by (51) at the bottom of the next page. In (51), (a) uses (44), and (b) holds because $E\left\{\left|\mathbf{e}_k^H \hat{\mathbf{f}}_{G,k}\right|^2\right\} = \delta_2^2$ and $E\left\{\left|\mathbf{e}_k^H \hat{\mathbf{f}}_{G,j}\right|^2\right\} = \delta_2^2$.

At last, by using $E\left\{\left|\mathbf{e}_k^H \hat{\mathbf{f}}_{G,k}\right|^2\right\} = \delta_2^2$, we have

$$\Delta_{b5} = \rho_2 \delta_2^2. \qquad (52)$$

Here, we substitute (48), (49), (50), (51), and (52) into (43). It gives

$$\Delta_b \leq \frac{1}{2}\log_2\left(1 + \rho_1\rho_2\left(\alpha_1\bar{\varepsilon} + \alpha_2\bar{\tau} + \alpha_3\right)\right) \qquad (53)$$

where notations $\alpha_1$, $\alpha_2$, and $\alpha_3$ are defined in (54), (55), and (56), respectively. In (53), the inequality holds for $\rho_1\rho_2 \frac{M(M-N+1)N}{M-1}\left(1 - \delta_1^2\right)\left(1 - \delta_2^2\right)\bar{\varepsilon}\bar{\tau} \geq 0$ and $\rho_1\rho_2 \frac{M\left(MN-N^2\right)}{M-1}\left(1 - \delta_1^2\right)\delta_2^2\bar{\varepsilon} \geq 0$.

$$\alpha_1 = \frac{M(N-1)}{M-1}\left(1 - \delta_1^2\right)\left(1 - \delta_2^2\right) \qquad (54)$$

$$\alpha_2 = N\left(1 - \delta_2^2\right)\left(N\delta_1^2 + M\left(1 - \delta_1^2\right) + \frac{1}{\rho_1}\right) \qquad (55)$$

$$\alpha_3 = N\delta_1^2 + MN\left(1 - \delta_1^2\right)\delta_2^2 + N\left(N - 1\right)\delta_1^2\delta_2^2 + \frac{N\delta_2^2}{\rho_1}. \qquad (56)$$

At this step, by substituting (41) and (53) into (40), the rate loss can be readily bounded by

$$\Delta R \leq \frac{1}{2}\log_2\left(1 + \rho_1\rho_2\left(\alpha_1\bar{\varepsilon} + \alpha_2\bar{\tau} + \alpha_3\right)\right) + \frac{1}{2}\log_2\left(\frac{1}{1 - \delta_1^2}\right)$$
$$+ \frac{1}{2}\log_2\left(\frac{1}{1 - \delta_2^2}\right) + \frac{1}{2}\log_2\left(1 + E\left\{\frac{\sqrt{\phi_k}}{\lambda_{\min}(\mathbf{Q})}\right\}\right) . \qquad (57)$$

In (57), the rate loss upper bound consists of four components. The first component increases with $P_1$ and $P_2$, while the second and third components decrease with $P_1$ and $P_2$, respectively. In the high SNR regions, both the second and the third components converge to 0 since $\delta_1^2 = \frac{1}{1 + \beta_1 P_1}$ and $\delta_2^2 = \frac{1}{1 + \beta_2 P_2}$ go to 0. The fourth component is a function of $\phi_k$ and $\mathbf{Q}$, which only depend on $B_1$ and are independent of SNR. Thus, we can conclude that the rate loss upper bound is dominated by the first component at high SNRs. Therefore, a more concise expression for the rate loss bound in the high SNR regions can be obtained as (20), which completes the proof.

## APPENDIX C
## PROOF OF LEMMA 2

As $P_2 = \kappa P_1 \to +\infty$, three equations can be obtained as (58), (59) (at the bottom of the next page), and (60) at the bottom of the next page.

$$\lim_{P_2 \to +\infty} \rho_1\rho_2\alpha_1 = \lim_{P_2 \to +\infty} \rho_1\rho_2 \frac{M(N-1)}{M-1}\left(1 - \delta_1^2\right)\left(1 - \delta_2^2\right)$$
$$= \frac{P_2}{N}\frac{N-1}{M-1}, \qquad (58)$$

By substituting (58) – (60) into (20), it gives

$$\Delta R_\infty \triangleq \lim_{P_2 \to +\infty} \Delta R_h$$
$$= \frac{1}{2}\log_2\left(1 + \varrho\right) + O\left(1\right)$$
$$+ \frac{1}{2}\log_2\left(1 + \frac{\frac{N(1+\beta_1)\kappa}{M\beta_1}\frac{N-1}{N}2^{-\frac{B_2}{(N-1)}}}{1 + \varrho}\right) \qquad (61)$$

In (61), $\varrho$ is defined by

$$\varrho = \frac{P_2}{N}\frac{N-1}{M}2^{-\frac{B_1}{M-1}} + P_2\frac{N-1}{N}2^{-\frac{B_2}{N-1}} + \frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right) . \qquad (62)$$

From results above, we further consider the following term.

$$\lim_{P_2 \to +\infty} \frac{\frac{N(1+\beta_1)\kappa}{M\beta_1}\frac{N-1}{N}2^{-\frac{B_2}{(N-1)}}}{1 + \varrho}$$
$$< \lim_{P_2 \to +\infty} \frac{(N-1)(1+\beta_1)\kappa 2^{-\frac{B_2}{(N-1)}}}{P_2 M\beta_1} = 0 .$$

Thus, the third term of (61) converges to 0. Then we have (21), which completes the proof.

## APPENDIX D
## PROOF OF LEMMA 4

When there is no estimation error, the RS and the $k$th user can respectively obtain $\mathbf{H}$ and $\mathbf{g}_k$. Let $\hat{\mathbf{V}} = [\hat{\mathbf{v}}_1, \cdots\cdots, \hat{\mathbf{v}}_N]$ and $\bar{\mathbf{g}}_k$ denote the quantizations of $\mathbf{V}$ and $\mathbf{g}_k$, respectively. We define $\bar{\mathbf{G}} = [\bar{\mathbf{g}}_1, \cdots, \bar{\mathbf{g}}_N]^H$ and $\bar{\mathbf{F}}_G = [\bar{\mathbf{f}}_{G,1}, \cdots, \bar{\mathbf{f}}_{G,N}]$, where $\bar{\mathbf{f}}_{G,k}$ is the normalized $k$th column of $\bar{\mathbf{G}}^H(\bar{\mathbf{G}}\bar{\mathbf{G}}^H)^{-1}$. Then, the SINR for the $k$th user is given by

$$\bar{\gamma}_k = \frac{\left|\mathbf{g}_k^H \bar{\mathbf{F}}_G \mathbf{U}^H \mathbf{H}\hat{\mathbf{v}}_k\right|^2}{\sum_{j\neq k}\left|\mathbf{g}_k^H \bar{\mathbf{F}}_G \mathbf{U}^H \mathbf{H}\hat{\mathbf{v}}_j\right|^2 + \frac{1}{\rho_1}\left\|\mathbf{g}_k^H \bar{\mathbf{F}}_G\right\|^2 + \frac{1}{\rho_1\rho_2}} . \qquad (63)$$

By respectively defining sets $\tilde{\mathbf{A}}_k$ and $\bar{\mathbf{A}}_k$ as $\tilde{\mathbf{A}}_k \triangleq \{\tilde{\mathbf{H}}, \tilde{\mathbf{g}}_k\}$ and $\bar{\mathbf{A}}_k \triangleq \{\mathbf{H}, \mathbf{g}_k\}$, it is observed that $\tilde{\mathbf{A}}_k$ is a noisy version of $\bar{\mathbf{A}}_k$. Then, by using the data-processing inequality, we can obtain

$$\tilde{R}_k = \frac{1}{2}\mathcal{I}\left(x_k; \left(\tilde{y}_k, \tilde{\mathbf{A}}_k\right)\right) \leq \frac{1}{2}\mathcal{I}\left(x_k; \left(\bar{y}_k, \bar{\mathbf{A}}_k\right)\right)$$
$$= \frac{1}{2}E\left\{\log_2\left(1 + \bar{\gamma}_k\right)\right\} \qquad (64)$$

Thus, the achievable rate for the system is upper bounded by (23), which completes the proof.

## APPENDIX E
## PROOF OF THEOREM 3

In order to characterize a sufficient scaling law of feedback bits, we set $\Delta R_h$ in (20) equal to $\frac{1}{2}\log_2 b$. Then we have

$$\frac{1}{2}\log_2\left(1 + \rho_1\rho_2\left(\alpha_1\bar{\varepsilon} + \alpha_2\bar{\tau} + \alpha_3\right)\right) + O\left(1\right) = \frac{1}{2}\log_2 b .\quad(65)$$

By omitting the $O\left(1\right)$ term for high SNRs, (65) reduces to

$$\frac{1}{2}\log_2\left(1 + \rho_1\rho_2\left(\alpha_1\bar{\varepsilon} + \alpha_2\bar{\tau} + \alpha_3\right)\right) = \frac{1}{2}\log_2 b \quad(66)$$

which implies that

$$\alpha_1\bar{\varepsilon} + \alpha_2\bar{\tau} = \frac{\left(b - 1 - \rho_1\rho_2\alpha_3\right)}{\rho_1\rho_2} .\quad(67)$$

To get a more meaningful result, we introduce a new variable $\zeta \in (0,1)$ to rewrite (67) by

$$\bar{\varepsilon} = \frac{\zeta(b - 1 - \rho_1\rho_2\alpha_3)}{\rho_1\rho_2\alpha_1} ,\quad(68)$$

and

$$\bar{\tau} = \frac{\left(1 - \zeta\right)\left(b - 1 - \rho_1\rho_2\alpha_3\right)}{\rho_1\rho_2\alpha_2} .\quad(69)$$

where different values of $\zeta$ result in different solutions to $\bar{\varepsilon}$ and $\bar{\tau}$, while keeping the rate loss requirement guaranteed.

By substituting $\bar{\varepsilon} = \frac{M-1}{M}2^{-\frac{B_1}{(M-1)}}$ and $\bar{\tau} = \frac{N-1}{N}2^{-\frac{B_2}{(N-1)}}$ into (68) and (69), respectively, we have

$$B_1 = (M-1)\left(\log_2\left(\rho_1\rho_2\alpha_1\right) - \log_2\left(\zeta(b - 1 - \rho_1\rho_2\alpha_3)\right)\right)$$
$$+ (M-1)\log_2\left(\frac{M-1}{M}\right) ,\quad(70)$$

and

$$B_2 = (N-1)\left(\log_2\left(\rho_1\rho_2\alpha_2\right) - \log_2\left((1-\zeta)(b - 1 - \rho_1\rho_2\alpha_3)\right)\right)$$
$$+ (N-1)\log_2\left(\frac{N-1}{N}\right) .\quad(71)$$

From the above results in (70) and (71), it is found that there are many pairs of $B_1$ and $B_2$ as a function of $\zeta$ to guarantee the required rate loss bound. Therefore, we are able to find

---

$$\Delta_{b2} = \rho_1\rho_2 E_{\tilde{\mathbf{G}},\mathbf{e}_k,\hat{\mathbf{v}}_e,\tilde{\mathbf{U}}}\left\{\left(\mathbf{g}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\right)E_{\mathbf{E}}\left\{\mathbf{E}\hat{\mathbf{V}}_e\left(\mathbf{E}\hat{\mathbf{V}}_e\right)^H\right\}\left(\mathbf{g}_k^H\hat{\mathbf{F}}_G\tilde{\mathbf{U}}^H\right)^H\right\}$$

$$\overset{(a)}{=} \rho_1\rho_2 N\delta_1^2 E_{\tilde{\mathbf{G}},\mathbf{e}_k}\left\{\left|\mathbf{g}_k^H\hat{\mathbf{f}}_{G,k}\right|^2 + \sum_{j\neq k}\left|\mathbf{g}_k^H\hat{\mathbf{f}}_{G,j}\right|^2\right\} \overset{(b)}{=} \rho_1\rho_2 N\left(1 + N\left(1 - \delta_2^2\right)\bar{\tau} + (N-1)\delta_2^2\right)\delta_1^2 \quad(49)$$

---

$$\Delta_{b3} = \rho_1\rho_2 E_{\tilde{\mathbf{G}},\mathbf{e}_k,\tilde{\boldsymbol{\Sigma}}}\left\{\mathbf{g}_k^H\hat{\mathbf{F}}_G\tilde{\boldsymbol{\Sigma}}\left\{\sum_{j\neq k}E_{\tilde{\mathbf{V}}}\left\{\tilde{\mathbf{V}}^H\hat{\mathbf{v}}_{e,j}\hat{\mathbf{v}}_{e,j}^H\tilde{\mathbf{V}}\right\}\right\}\tilde{\boldsymbol{\Sigma}}\left(\mathbf{g}_k^H\hat{\mathbf{F}}_G\right)^H\right\}$$

$$\overset{(a)}{=} \rho_1\rho_2 E_{\tilde{\mathbf{G}},\mathbf{e}_k}\left\{\mathbf{g}_k^H\hat{\mathbf{F}}_G\left(E_{\tilde{\boldsymbol{\Sigma}}}\left\{\tilde{\boldsymbol{\Sigma}}^2\right\}\right)\left(\frac{N-1}{M-1}\bar{\varepsilon}\mathbf{s}_k\mathbf{s}_k^H + \sum_{j\neq k}\left(1 - \frac{M-N+1}{M-1}\bar{\varepsilon}\right)\mathbf{s}_j\mathbf{s}_j^H\right)\left(\mathbf{g}_k^H\hat{\mathbf{F}}_G\right)^H\right\}$$

$$\overset{(b)}{=} \rho_1\rho_2 M\left(1 - \delta_1^2\right)\left(\frac{N-1}{M-1}\bar{\varepsilon} + \left(1 - \frac{M-N+1}{M-1}\bar{\varepsilon}\right)\left(N\left(1 - \delta_2^2\right)\bar{\tau} + (N-1)\delta_2^2\right)\right) \quad(50)$$

---

$$\Delta_{b4} = \rho_1\rho_2 E_{\tilde{\mathbf{G}},\mathbf{e}_k,\tilde{\boldsymbol{\Sigma}}}\left\{\mathbf{e}_k^H\hat{\mathbf{F}}_G\tilde{\boldsymbol{\Sigma}}\left\{E_{\tilde{\mathbf{V}}}\left\{\tilde{\mathbf{V}}^H\hat{\mathbf{v}}_{e,k}\hat{\mathbf{v}}_{e,k}^H\tilde{\mathbf{V}}\right\}\right\}\tilde{\boldsymbol{\Sigma}}\left(\mathbf{e}_k^H\hat{\mathbf{F}}_G\right)^H\right\}$$

$$\overset{(a)}{=} \rho_1\rho_2 E_{\tilde{\mathbf{G}},\mathbf{e}_k}\left\{\mathbf{e}_k^H\hat{\mathbf{F}}_G\left(E_{\tilde{\boldsymbol{\Sigma}}}\left\{\tilde{\boldsymbol{\Sigma}}^2\right\}\right)\left((1 - \bar{\varepsilon})\mathbf{s}_k\mathbf{s}_k^H + \sum_{j\neq k}\frac{1}{M-1}\bar{\varepsilon}\mathbf{s}_j\mathbf{s}_j^H\right)\left(\mathbf{e}_k^H\hat{\mathbf{F}}_G\right)^H\right\}$$

$$\overset{(b)}{=} \rho_1\rho_2 M\left(1 - \delta_1^2\right)\left(1 - \frac{M-N}{M-1}\bar{\varepsilon}\right)\delta_2^2 \quad(51)$$

---

$$\lim_{P_2\to+\infty}\rho_1\rho_2\alpha_2 = \lim_{P_2\to+\infty}\rho_1\rho_2 N\left(1 - \delta_2^2\right)\left(N\delta_1^2 + M\left(1 - \delta_1^2\right) + \frac{1}{\rho_1}\right) = P_2 + \frac{N\left(1 + \beta_1\right)\kappa}{M\beta_1},\quad(59)$$

and

$$\lim_{P_2\to+\infty}\rho_1\rho_2\alpha_3 = \lim_{P_2\to+\infty}\rho_1\rho_2\left(N\delta_1^2 + MN\left(1 - \delta_1^2\right)\delta_2^2 + N\left(N-1\right)\delta_1^2\delta_2^2 + \frac{N\delta_2^2}{\rho_1}\right) = \frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right) \quad(60)$$

an optimized $\zeta$ to minimize the total number of feedback bits $(B_1 + B_2)$ while satisfying the conditions in (68) and (69). It is not hard to obtain the optimal $\zeta = 0.5$ when $M = N$. By substituting $\zeta = 0.5$ into (70) and (71), we have (24) and (25).

A procedure similar as *Lemma* 2 is carried out to analyze the number of feedback bits in the high SNR regions. As $P_2 = \kappa P_1 \to +\infty$, we have (58) – (60) in Appendix C. By substituting (58) – (60) into (70) and (71), $B_1$ and $B_2$ can be scaled according to (26) and (27), respectively.

## APPENDIX F
## PROOF OF LEMMA 5

Under the constraint $\beta_1 + \beta_2 = C_\beta$, the optimization problem is formulated as

$$\underset{\beta_1, \beta_2}{\text{maximize}} \qquad \tilde{R}_{lower}^\infty$$
$$\text{subject to} \qquad \beta_1 + \beta_2 = C_\beta \qquad (72)$$

Then, we use the Lagrangian multiplier method to solve the problem. The Lagrangian function is (73) on the top of the next page. By calculating its partial derivative with respect to $\beta_1$ and $\beta_2$, and equating them to zero, we obtain (74) on the top of the next page. Then, by solving the equations in (74), the proof of *Lemma* 5 completes.

## ACKNOWLEDGMENT

## REFERENCES

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[2] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[3] L.-L. Xie and P. R. Kumar, "An achievable rate for the multiple-level relay channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1348–1358, Apr. 2005.

[4] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, June 2006.

[5] C.-B. Chae, T. Tang, R. W. Health, and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 727–738, Feb. 2008.

[6] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[7] R. Mo and Y. H. Chew, "Precoder design for non-regenerative MIMO relay systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5041–5049, Oct. 2009.

[8] R. Mo and Y. H. Chew, "MMSE-based joint source and relay precoding design for amplify-and-forward MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4668–4676, Sep. 2009.

[9] L. Sun and M. R. McKay, "Opportunistic relaying for MIMO wireless communication: relay selection and capacity scaling laws," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1786–1797, June 2011.

[10] A. I. Sulyman, G. Takahara, H. S. Hassanein, and M. Kousa, "Multihop capacity of MIMO-multiplexing relaying systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3095–3103, June 2009.

[11] J. Joung and A. H. Sayed, "Multiuser two-way amplify-and-forward relay processing and power control methods for beamforming systems," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1833–1846, Mar. 2010.

[12] W. Xu, X. Dong, and W.-S. Lu, "Joint precoding optimization for multiuser multi-antenna relaying downlinks using quadratic programming," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1228–1235, May 2011.

[13] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay network," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[14] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[15] Y. Rong, "Robust design for linear non-regenerative MIMO relays with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2455–2460, May 2011.

[16] B. Gedik and M. Uysal, "Impact of imperfect channel estimation on the performance of amplify-and-forward relaying," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1468–1479, Mar. 2009.

[17] J. Ma, P. Orlik, J. Zhang, and G. Y. Li, "Pilot matrix design for estimating cascaded channels in two-hop MIMO amplify-and-forward relay systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1956–1965, June 2011.

[18] F. Gao, B. Jiang, X. Gao, and X.-D. Zhang, "Superimposed training based channel estimation for OFDM modulated amplify-and-forward relay networks," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 2029–2039, July 2011.

[19] T. A. Lamahewa, P. Sadeghi, and X. Zhou, "On lower bounding the information capacity of amplify and forward wireless relay channels with channel estimation errors," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2075–2079, July 2011.

[20] B. K. Chalise and L. Vandendorpe, "MIMO relay design for multipoint-to-multipoint communications with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2785–2796, July 2009.

[21] D. J. Love, R. W. Heath Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[22] W. Xu, X. Dong, and W.-S. Lu, "MIMO relaying broadcast channels with linear precoding and quantized channel state information feedback," *IEEE Trans. Signal Process.*, vol. 58, no.10, pp. 5233–5245, Oct. 2010.

[23] R. Mo, Y. H. Chew, and C. Yuen, "Information rate and relay precoder design for amplify-and-forward MIMO relay networks with imperfect channel state information," *IEEE Trans. Veh. Technol.*, vol. 61, no. 9, pp. 3958–3968, Nov. 2012.

[24] B. Yi, S. Wang, and S. Y. Kwon, "On MIMO relay with finite-rate feedback and imperfect channel estimation," in *Proc. 2007 IEEE Global Telecommun. Conf.*, pp. 3878–3882.

[25] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no.6, pp. 2845–2866, June 2010.

[26] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[27] W. Santipach and M. L. Honig, "Signature optimization for CDMA with limited feedback," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3475–3492, Oct. 2005.

[28] M. Kobayashi, N. Jindal, and G. Caire, "Training and feedback optimization for multiuser MIMO downlink," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug. 2011.

[29] T. L. Marzetta and B. M. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1268–1278, Apr. 2006.

[30] A. Lapidoth and S. Shamai, "Fading channels: how perfect need "perfect-side information" be?," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[31] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley and Sons, 1991.

[32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

**Zhangjie Peng** (S'10) is currently a Ph.D. student at the National Mobile Communications Research Laboratory, Southeast University, China, where he received his M.S. degree in 2007. He received his B.S. degree from Southwest Jiaotong University, Chengdu, China, in 2004. His current research interests are in the area of wireless communications.

$$L\left(\beta_1, \beta_2, \varphi\right) = R - \frac{N}{2}\log_2\left(\frac{P_2}{N}\frac{N-1}{M}2^{-\frac{B_1}{M-1}} + P_2\frac{N-1}{N}2^{-\frac{B_2}{N-1}} + \left(1 + \frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)\right)\right)$$
$$+ \varphi\left(\beta_1 + \beta_2 - C_\beta\right) - O\left(N\right) \tag{73}$$

$$\begin{cases} \dfrac{N}{2M\ln 2\left(\frac{P_2}{N}\frac{N-1}{M}2^{-\frac{B_1}{M-1}} + P_2\frac{N-1}{N}2^{-\frac{B_2}{N-1}} + \left(1 + \frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)\right)\right)}\dfrac{\kappa}{\beta_1^2} + \varphi = 0 \\[4mm] \dfrac{N}{2M\ln 2\left(\frac{P_2}{N}\frac{N-1}{M}2^{-\frac{B_1}{M-1}} + P_2\frac{N-1}{N}2^{-\frac{B_2}{N-1}} + \left(1 + \frac{1}{M}\left(\frac{\kappa}{\beta_1} + \frac{M}{\beta_2}\right)\right)\right)}\dfrac{M}{\beta_2^2} + \varphi = 0 \\[4mm] \beta_1 + \beta_2 - C_\beta = 0 \end{cases} \tag{74}$$

**Wei Xu** (S'07-M'09) received his B.Sc. degree in electrical engineering in 2003 and his M.S. and Ph.D. degrees in communication and information engineering in 2006 and 2009, respectively, all from Southeast University, Nanjing, China. He is currently an Associate Professor at the National Mobile Communications Research Lab (NCRL), Southeast University. Between 2009 and 2010, he was a postdoctoral research fellow with the Department of Electrical and Computer Engineering, University of Victoria, Canada.

Dr. Xu is an Editor of IEEE COMMUNICATIONS LETTERS. He has been involved in Technical Program Committees for many international conferences including the IEEE Global Communications Conference (Globecom), the IEEE Wireless Communications and Networking Conference (WCNC), the IEEE Vehicular Technology Conference (VTC), etc. He has co-authored more than 60 papers and has over 10 patents pending. He was elected Core Team Member of the Jiangsu Innovation Team in 2012. His research interests include communication theory, cooperative communications, and signal processing and optimization theory for wireless communications.

**Li-Chun Wang** (M'96-SM'06-F'11) received the B.S. degree from National Chiao Tung University, Taiwan, R.O.C., in 1986, the M.S. degree from National Taiwan University in 1988, and the Ms.Sci. and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, in 1995, and 1996, respectively, all in electrical engineering.

From 1990 to 1992, he was with the Telecommunications Laboratories of the Ministry of Transportation and Communications in Taiwan (currently Telecom Labs of Chunghwa Telecom Co.). In 1995, he was affiliated with Bell Northern Research of Northern Telecom, Inc., Richardson, TX. From 1996 to 2000, he was with AT&T Laboratories, where he was a Senior Technical Staff Member in the Wireless Communications Research Department. In August 2000, he joined the Department of Electrical and Computer Engineering of National Chiao Tung University in Taiwan and is currently Chairman of the same department. His current research interests are in the areas of radio resource management, cross-layer optimized techniques for heterogeneous wireless networks, and cloud computing for mobile applications.

Dr. Wang was elected to the IEEE Fellow grade in 2011 for his contributions in cellular architecture and radio resource management in wireless networks. He won the Distinguished Research Award of the National Science Council, Taiwan, in 2012, and was a co-recipient (with Gordon L. Stüber and Chin-Tau Lea) of the 1997 IEEE Jack Neubauer Best Paper Award for his paper "Architecture Design, Frequency Planning, and Performance Analysis for a Microcell/Macrocell Overlaying System," IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, vol. 46, no. 4, pp. 836–848, 1997. He has published over 150 journal and international conference papers. He served as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2001 to 2005, the Guest Editor of the Special Issue on "Mobile Computing and Networking" for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS in 2005, and on "Radio Resource Management and Protocol Engineering in Future IEEE Broadband Networks" for *IEEE Wireless Communications Magazine* in 2006. He holds 10 US patents.
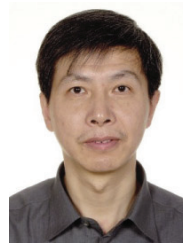
**Chunming Zhao** (M'93) received the B.S. and M.S. degrees from Nanjing Institute of Posts and Telecommunications in 1982 and 1984, respectively. In 1993, he received his Ph.D. degree from the Department of Electrical and Electronic Engineering, University of Kaiserslautern, Germany. He has been a Postdoctoral Researcher at the National Mobile Communications Research Lab, Southeast University, where he is currently a professor and vice director of the lab. He has managed several key projects of the Chinese Communications High Tech. Program and was awarded as an "excellent researcher" from the Ministry of Science and Technology, China. He also won the First Prize of the National Technique Invention of China in 2011. His research interests include communication theory, coding/decoding, mobile communications, and VLSI design.