EURASIP Journal on
Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Achievable data rates and power allocation for frequency-selective fading relay channels with imperfect channel estimation

Yanwu Ding[1*] and Murat Uysal[2]

Abstract

In this article, we investigate the information-theoretical performance of a cooperative orthogonal frequency division multiplexing (OFDM) system with imperfect channel estimation. Assuming the deployment of training-aided channel estimators, we derive a lower bound on the achievable rate for the cooperative OFDM system with amplify-and-forward relaying over frequency-selective Rayleigh fading channels. The bound is later utilized to allocate power among the training and data transmission phases. Numerical results demonstrate that the proposed power allocation scheme brings between 5 and 19% improvement depending on the level of signal-to-noise ratio and relay locations.

Keywords: Achievable rate, Amplify-and-forward relaying, Channel estimation, Power allocation, OFDM

1 Introduction

Cooperative transmission has been proposed as a powerful method to overcome the degrading effects of fading in wireless channels [1-3]. Exploiting the broadcasting nature of the wireless channel, cooperative transmission builds upon the idea of a number of nodes helping each other through relaying. It extracts spatial diversity advantages in a distributed manner and brings significant improvements in link reliability, spectral efficiency, and coverage area. Two popular relaying schemes are decode-and-forward (DF) and amplify-and-forward (AF), which are sometimes referred to as regenerative and non-regenerative relaying, respectively. In AF relaying, the relay node retransmits a scaled version of the received message without any attempt to decode it. In DF relaying, the relay node decodes the received message, re-encodes, and transmits to the destination.

Information-theoretical aspects of cooperative communications have been investigated by several authors [4-8]. Gastpar and Vetterli [4] have examined the asymptotic capacity as the number of relay nodes goes to infinity. In their derivations, they have assumed arbitrarily complex network coding over Gaussian relay channels and ignored the effects of fading. Wong et al. [7] have derived upper and lower bounds on the capacity for both deterministic (i.e., fixed channel coefficients) and Rayleigh fading channels. Optimum resource allocation has been proposed in [9,10] to optimize the capacity of AF networks. Specifically, Maric and Yates [9] have investigated power and bandwidth allocations for a large number of relay nodes assuming that the channel state information is available at the transmitter. Deng and Haimovich [10] have developed power allocation strategies to optimize the outage performance for a single-relay AF cooperative system. Zheng and Gursoy [11] have derived achievable rates for AF and DF relaying with imperfect channel estimation.

A common assumption in the aforementioned works is frequency-flat fading channel model. Although this model is sufficient to model narrowband systems, it becomes unrealistic for broadband communication systems where the transmission bandwidth is larger than the coherence bandwidth of the channel. This, in return, results in a frequency-selective channel, which causes intersymbol interference (ISI) at the receiver. A widely used approach to overcome the degrading effects of ISI is orthogonal frequency division multiplexing (OFDM). OFDM has already been adopted by various industry standards such as IEEE 802.11 (WiFi) and 802.16 (WiMax).

*Correspondence: yanwu.ding@wichita.edu
[1] Department of Electrical Engineering and Computer Science, Wichita State University, 1845 Fairmount, Wichita, KS 67260, USA
Full list of author information is available at the end of the article

Currently, there has been a growing interest in the application of OFDM to cooperative communication systems [12-16]. In [12], a space–time cooperative protocol with the transmitter and receiver architecture, frame structure, and synchronization algorithms are designed for an OFDM relay system. In [13], power loading is considered in the frequency and time domains to maximize an instantaneous rate, assuming channel knowledge is available at the transmitter. In [14], equalization methods for cooperative diversity schemes over frequency-selective channels have been investigated. Ma et al. [15] have proposed a margin-adaptive bit and power loading approach for an OFDM single-relay system. Ibrahimi and Liang [16] have investigated joint power allocation among the source, relays, and OFDM subchannels for coherent reception. These works are based on the assumption that perfect channel knowledge is available at the receiver and/or transmitter. In practice, the channel coefficients need to be estimated and made available to the receiver. Recent research efforts have focused on the analysis and design of OFDM relay systems with imperfect channel estimations. Amin and Uysal [17] have investigated bit and power loading for an AF OFDM relay systems using bit error rate as the performance measure. Wang et al. [18] have considered the resource allocation and relay selection in a DF orthogonal frequency division multiple access-based downlink network. However, few of the current works address the achievable rates for an OFDM system with imperfect channel estimation.

In this article, we study the achievable rate for a single-relay OFDM system with AF relaying and a training-aided channel estimator at the receivers. We assume no knowledge of channel state information at the transmitter side realizing an open-loop scheme. Minimum mean square error (MMSE) estimators are applied to obtain the channel estimates. In the derivation of the achievable rate of the OFDM relay system, the channel estimation errors are considered together with the noise forwarded from the relay node and noise at the destination. Since the statistical distributions of the channel estimation errors are difficult to characterize, a closed-form analytical expression for the achievable rate is mathematically intractable. We, therefore, resort to find a lower bound on the achievable rate. Then, we use this bound to allocate power between the training and data transmission phases.

The rest of the article is organized as follows: Section 2 introduces the system model and describes the training and data transmission phases. In Section 3, we derive a lower bound on the achievable rate, assuming the distintegrated estimation of source-to-relay and relay-to-destination links. Section 4 presents a power allocation scheme by maximizing the derived bound. In Section 5, we investigate the problem of cascaded channel estimation for the overall relaying link. Numerical results are provided in Section 6. Finally, Section 7 concludes the article.

Notation: Matrices and column vectors are denoted by uppercase and lowercase boldface characters, respectively (e.g., $\mathbf{A}, \mathbf{a}$). The transpose of $\mathbf{A}$ is denoted by $\mathbf{A}^{\mathrm{T}}$, and the conjugate and transpose of $\mathbf{A}$ by $\mathbf{A}^{\mathrm{H}}$. A vector $\mathbf{s}$ of length N is denoted by $\mathbf{s} = [s(1), s(2), \ldots, s(N)]$. $\mathbf{I}_K$ denotes a $K \times K$ identity matrix and $\mathbf{0}$ stands for an all-zero matrix of appropriate dimensions. $X(i,j)$ denotes the (i,j)th element in matrix $\mathbf{X}$. The ith diagonal element in diagonal matrix $\mathbf{D}$ is denoted by $D(i)$. $\mathbb{E}[\cdot]$ is the expectation operator, and $\log(\cdot)$ represents a logarithm of base 2. The notation $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \Sigma)$ means that $\mathbf{n}$ is a circularly symmetric complex Gaussian (CSCG) random vector with zero mean and covariance matrix $\mathbf{\Sigma}$. Matrix $\mathbf{V}_{\mathrm{AB}}$ denotes a $Q \times L_{\mathrm{AB}}$ matrix whose (k,m)th element is given by $V_{\mathrm{AB}}(k,m) = \exp\left(-j2\pi(k-1)(m-1)/Q\right)$, $1 \le k \le Q$, $1 \le m \le L_{\mathrm{AB}}$.

2 System model

A three-node cooperative system is illustrated in Figure 1. The relay node R assists transmission from the source node S to the destination D. Each node is equipped with a single antenna and operates in a half-duplex mode. The transmissions and receptions are not carried out simultaneously. An orthogonal AF relaying strategy is applied, whereby the source node first transmits to the destination and the relay node (broadcasting phase), and then the relay node forwards a scaled noisy version of the signal received and the source node is silent (relaying phase).

The underlying channels are modeled as frequency-selective Rayleigh fading with a uniform delay profile. To overcome the ISI in frequency-selective channels, we apply the OFDM scheme to the relay system, which converts the frequency-selective fading channel into a number of parallel frequency-flat channels free of ISI. An aggregate channel model consisting of both long-term path loss and short-term fading effects is considered. The path loss is proportional to $d^{-\gamma}$, where d is the propagation distance and γ is the path loss exponent. By normalizing the path loss in the source-to-destination (S→D) link to be unity, the relative gains from source-to-relay (S→R), and from the relay-to-destination (R→D)

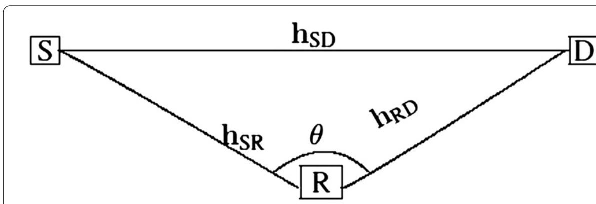

Figure 1 Relay-assisted transmission model.

links are defined, respectively, as $K_{\mathrm{SR}} = (d_{\mathrm{SD}}/d_{\mathrm{SR}})^{\gamma}$ and $K_{\mathrm{RD}} = (d_{\mathrm{SD}}/d_{\mathrm{RD}})^{\gamma}$ [19]. The channel impulse responses (CIRs) for S $\rightarrow$ R, R $\rightarrow$ D, and S $\rightarrow$ D links are given, respectively, by $\mathbf{h}_{\mathrm{SR}} = [h_{\mathrm{SR}}(1),\ h_{\mathrm{SR}}(2),\ \ldots,\ h_{\mathrm{SR}}(L_{\mathrm{SR}})]$, $\mathbf{h}_{\mathrm{RD}} = [h_{\mathrm{RD}}(1),\ h_{\mathrm{RD}}(2), \ldots, h_{\mathrm{RD}}(L_{\mathrm{RD}})]$, and $\mathbf{h}_{\mathrm{SD}} = [h_{\mathrm{SD}}(1),\ h_{\mathrm{SD}}(2), \ldots, h_{\mathrm{SD}}(L_{\mathrm{SD}})]$. The entries in $\mathbf{h}_{\mathrm{SR}}, \mathbf{h}_{\mathrm{RD}}$, and $\mathbf{h}_{\mathrm{SD}}$ are independent identically distributed (i.i.d) zero-mean CSCG random variables with variances of $1/L_{\mathrm{SR}}$, $1/L_{\mathrm{RD}}$, and $1/L_{\mathrm{SD}}$, respectively. The underlying channels are modeled as quasi-static Rayleigh fading, whereas the CIRs remain constant in the duration of one OFDM block and change to independent values that hold for another block.

To avoid the inter-block interference for the OFDM system, a cyclic prefix of length $\max(L_{\mathrm{SD}}, L_{\mathrm{SR}}, L_{\mathrm{RD}}) - 1$ is applied. After the cyclic prefix is removed at the receivers, the length of the OFDM block is denoted by Q (which is also the number of subcarriers in one OFDM block). We assume that both relay and destination nodes are equipped with channel estimators. The relay node obtains an estimate of the CIR from the S $\rightarrow$ R link through training symbols and feed-forwards this information to the destination. The relay node also transmits "clean" training symbols so that the CIR for the R $\rightarrow$ D link can be obtained at the destination.[a] In [20], it is proven that the minimum length of training symbols required for a non-cooperative OFDM system equals the channel length, and the optimal placement is that the training symbols are periodically inserted in each OFDM block. In this article, we adopt a similar channel training strategy. The number of training symbols is chosen as the maximum channel length among the links, i.e., $N = \max(L_{\mathrm{SD}}, L_{\mathrm{SR}}, L_{\mathrm{RD}})$, where N is the number of training symbols. The number of subcarriers in an OFDM block is chosen as $Q = (M+1)N$, with $M \geq 1$ being an integer. The training symbols are placed periodically at positions $i_\ell = 1 + (\ell - 1)(M+1), \ell = 1, \ldots, N$ in the OFDM block.

Let the vectors $\mathbf{x}_{\mathrm{S}} = [x_{\mathrm{S}}(1), x_{\mathrm{S}}(2), \ldots, x_{\mathrm{S}}(N)]^T$ and $\mathbf{x}_R = [x_R(1), x_R(2), \ldots, x_R(N)]^T$ denote, respectively, the training symbols transmitted from the source and relay nodes. The data symbols are collected in vector $\mathbf{y} = \left[y(1), y(2), \ldots, y(MN)\right]^T$. With the training symbols periodically inserted, an OFDM block transmitted from the source node is expressed as $\left[x_{\mathrm{S}}(1), y(1), y(2), \ldots, y(M), x_{\mathrm{S}}(2), y(M+1), \ldots, y(2M), \ldots, x_{\mathrm{S}}(N), y\left((N-1)M+1\right) \ldots, y(MN)\right]$. Let P_{S} and P_R denote, respectively, the available power at the source and relay nodes. Assuming that the training symbols are independent of the data symbols, we define $P_{\mathrm{S}} = \frac{1}{(M+1)N}(\mathbf{x}_{\mathrm{S}}^H\mathbf{x}_{\mathrm{S}} + \mathbb{E}[\mathbf{y}^H\mathbf{y}])$, and $P_R = \frac{1}{(M+1)N}(\mathbf{x}_{\mathrm{R}}^H\mathbf{x}_{\mathrm{R}} + \mathbb{E}[\mathbf{w}_{\mathrm{R}}^H\mathbf{w}_{\mathrm{R}}])$, where $\mathbf{w}_{\mathrm{R}}^H$ is the signal vector forwarded from the relay node. The power allocated in training and data transmission phases at the source and relay nodes can individually be written as $\mathbf{x}_{\mathrm{S}}^H\mathbf{x}_{\mathrm{S}}/N = \alpha_{\mathrm{t}}P_{\mathrm{S}}, \mathbb{E}[\mathbf{y}^H\mathbf{y}]/(MN) = \alpha_{\mathrm{d}}P_{\mathrm{S}}, \mathbf{x}_{\mathrm{R}}^H\mathbf{x}_{\mathrm{R}}/N = \beta_{\mathrm{t}}P_{\mathrm{R}}$, and $\mathbb{E}[\mathbf{w}_{\mathrm{R}}^H\mathbf{w}_{\mathrm{R}}]/(MN) = \beta_{\mathrm{d}}P_{\mathrm{R}}$, where $\alpha_{\mathrm{t}}, \alpha_{\mathrm{d}}, \beta_{\mathrm{t}}$, and β_{d} are, respectively, the power allocation factors deployed at the source and the relay node, and they are related by $\alpha_{\mathrm{t}} + M\alpha_{\mathrm{d}} = M + 1$ and $\beta_{\mathrm{t}} + M\beta_{\mathrm{d}} = M + 1$.

2.1 Training phase

Let the diagonal matrix $\mathbf{G}_{\mathrm{AB}}$ denote the frequency response in the link A$\rightarrow$B. The qth diagonal entry is given by $\mathbf{G}_{\mathrm{AB}}(q) = \sum_{k=1}^{L_{\mathrm{AB}}} h_{\mathrm{AB}}(k)\exp(-j2\pi(k-1)(q-1)/Q)$, $q = 1, 2, \ldots, Q$, where L_{AB} is the channel length. Since the training symbols are placed periodically in each OFDM block at positions $i_\ell = 1 + (\ell - 1)(M+1), \ell = 1, \ldots, N$, the ℓth frequency response in the training phase for link A$\rightarrow$B is equal to $G_{\mathrm{AB}}(i_\ell)$, and the received vector in the S$\rightarrow$D link during the training phase can be expressed as

$$\begin{pmatrix} z_{\mathrm{SD}}(1) \\ z_{\mathrm{SD}}(2) \\ \vdots \\ z_{\mathrm{SD}}(N) \end{pmatrix} = \begin{pmatrix} G_{\mathrm{SD}}(1_\ell) & 0 & \cdots & 0 \\ 0 & G_{\mathrm{SD}}(2_\ell) & & 0 \\ & & \ddots & \\ 0 & & & G_{\mathrm{SD}}(N_\ell) \end{pmatrix} \begin{pmatrix} x_{\mathrm{S}}(1) \\ x_{\mathrm{S}}(2) \\ \vdots \\ x_{\mathrm{S}}(N) \end{pmatrix} + \begin{pmatrix} n(1_\ell) \\ n(2_\ell) \\ \vdots \\ n(N_\ell) \end{pmatrix} \quad (1)$$

where $n(i_\ell), \ell = 1, \ldots, N$ is the independent CSCG noise random variable at the destination in the i_ℓth received frequency response. The frequency response matrix in (1) is also written as, $\mathrm{diag}[G_{\mathrm{AB}}(1_\ell), G_{\mathrm{AB}}(2_\ell), \ldots, G_{\mathrm{AB}}(N_\ell)] = \Phi\mathbf{G}_{\mathrm{AB}}$, and Φ is a selection matrix of size $N \times Q$ [20] with the ith row equal to the ith row of the identity matrix $\mathbf{I}_Q$. Let the diagonal matrix $\mathbf{X}_{\mathrm{S}} = \mathrm{diag}(\mathbf{x}_{\mathrm{S}})$ consist of the training symbols at the source node; then the received signals in (1) are further written in a matrix form as

$$\mathbf{z}_{\mathrm{SD}} = \mathbf{X}_{\mathrm{S}}\Phi\mathbf{V}_{\mathrm{SD}}\mathbf{h}_{\mathrm{SD}} + \mathbf{n}_{\mathrm{t}} \quad (2)$$

where $\mathbf{z}_{\mathrm{SD}} = [z_{\mathrm{SD}}(1), z_{\mathrm{SD}}(2), \ldots, z_{\mathrm{SD}}(N)]^T$, and $\mathbf{n}_{\mathrm{t}} = [n(1_\ell), n(2_\ell), \ldots, n(N_\ell)]^T \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$ is the noise vector at the destination in the channel training for S$\rightarrow$D link. Similarly, the received vectors in the channel training for S$\rightarrow$R and R$\rightarrow$D links have the following form

$$\mathbf{z}_{\mathrm{SR}} = \sqrt{K_{\mathrm{SR}}}\mathbf{X}_{\mathrm{S}}\Phi\mathbf{V}_{\mathrm{SR}}\mathbf{h}_{\mathrm{SR}} + \mathbf{u}_{\mathrm{t}} \quad (3)$$

$$\mathbf{z}_{\mathrm{RD}} = \sqrt{K_{\mathrm{RD}}}\mathbf{X}_{\mathrm{R}}\Phi\mathbf{V}_{\mathrm{RD}}\mathbf{h}_{\mathrm{RD}} + \mathbf{v}_{\mathrm{t}} \quad (4)$$

where $\mathbf{X}_{\mathrm{R}} = \mathrm{diag}(\mathbf{x}_{\mathrm{R}})$, $\mathbf{u}_{\mathrm{t}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$ is the noise vector at the relay in the channel training for S$\rightarrow$R link, and $\mathbf{v}_{\mathrm{t}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$ is the noise vector at the destination in the channel training for R→D link. We assume that the MMSE estimator is deployed in the channel estimation. Inserting $K_{\rm SD}=1$, the estimated CIRs for the A → B link is given by [21]

$$\widehat{\mathbf{h}}_{\rm AB}=\sqrt{K_{\rm AB}}\mathbf{V}_{\rm AB}^H\Phi^T\mathbf{X}_{\rm A}^H \times\left(K_{\rm AB}\mathbf{X}_{\rm A}\Phi\mathbf{V}_{\rm AB}\mathbf{V}_{\rm AB}^H\boldsymbol{\Phi}^T\mathbf{X}_{\rm AB}^H+\sigma^2L_{\rm AB}\mathbf{I}\right)^{-1}\mathbf{z}_{\rm AB} \quad (5)$$

where {AB} takes the values of {SD, SR, or RD}, and we have used the fact $\mathbb{E}[\mathbf{h}_{\rm AB}\mathbf{h}_{\rm AB}^H]=\mathbf{I}_{L_{\rm AB}}/L_{\rm AB}$. Let $\widetilde{\mathbf{h}}_{\rm AB}=\mathbf{h}_{\rm AB}-\widehat{\mathbf{h}}_{\rm AB}$ be the estimation error for the channel gains in the A→B link. The error covariance matrix is thus given by

$$\mathbb{E}[\widetilde{\mathbf{h}}_{\rm AB}\widetilde{\mathbf{h}}_{\rm AB}^H]=\frac{1}{L_{\rm AB}}\left(\mathbf{I}+\frac{K_{\rm AB}\mathbf{V}_{\rm AB}^H\boldsymbol{\Phi}^T\mathbf{X}_{\rm A}^H\mathbf{X}_{\rm A}\boldsymbol{\Phi}\mathbf{V}_{\rm AB}}{\sigma^2L_{\rm AB}}\right)^{-1} \quad (6)$$

2.2 Data transmission phase

In the data transmission phase, the information symbols are first broadcasted to the destination and the relay node (broadcasting phase), and then the relay node forwards a scaled received signals to the destination (relaying phase). The received vectors at the destination during the broadcasting and relaying phases are given, respectively, as

$$\mathbf{r}_1=\mathbf{G}_{\rm SDd}\mathbf{y}+\mathbf{n}_{\rm d} \quad (7)$$

$$\mathbf{r}_2=A\sqrt{K_{\rm RD}}\mathbf{G}_{\rm RDd}\mathbf{r}_{\rm R}+\mathbf{g}_{\rm d} \quad (8)$$

where $\mathbf{r}_{\rm R}=\sqrt{K_{\rm SR}}\mathbf{G}_{\rm SRd}\mathbf{y}+\mathbf{w}_{\rm d}$ is the received vector at the relay node during the broadcasting phase, $\mathbf{w}_{\rm d}\sim(\mathbf{0},\sigma^2\mathbf{I}_{MN})$ is the CSCG noise vector at the relay, diagonal matrices $\mathbf{G}_{\rm SDd}=\bar{\boldsymbol{\Phi}}\mathbf{G}_{\rm SD}$, $\mathbf{G}_{\rm SRd}=\bar{\boldsymbol{\Phi}}\mathbf{G}_{\rm SR}$, and $\mathbf{G}_{\rm RDd}=\bar{\boldsymbol{\Phi}}\mathbf{G}_{\rm RD}$ designate, respectively, the frequency response for S → D, S → R, and R → D links in the data transmission phase, $\bar{\boldsymbol{\Phi}}$ is a selection matrix of size $MN\times Q$ obtained by removing the rows in $\boldsymbol{\Phi}$ from $\mathbf{I}_Q$, A is the amplification coefficient at the relay node to guarantee that the power of the signal forwarded from the relay does not exceed the available power, and $A=\sqrt{\frac{MN\beta_{\rm d}P_{\rm R}}{||\mathbf{r}_{\rm R}||^2}}=\sqrt{\frac{\beta_{\rm d}P_{\rm R}}{K_{\rm SR}\alpha_{\rm d}P_{\rm S}+\sigma^2}}$, and $\mathbf{n}_{\rm d},\mathbf{g}_{\rm d}\sim(\mathbf{0},\sigma^2\mathbf{I}_{MN})$ are the CSCG noise vectors at the destination during the broadcasting and relaying phases. From (7) and (8), the overall received vector at the destination in the data transmission phase is

$$\mathbf{r}=\begin{pmatrix}\mathbf{G}_{\rm SDd}\\ \lambda\sqrt{K_{\rm SR}}\mathbf{G}_{\rm SRd}\mathbf{G}_{\rm RDd}\end{pmatrix}\mathbf{y}+\begin{pmatrix}\mathbf{n}_{\rm d}\\ \lambda\mathbf{G}_{\rm RDd}\mathbf{w}_{\rm d}+\mathbf{g}_{\rm d}\end{pmatrix} \quad (9)$$

where $\mathbf{r}=[\mathbf{r}_1^T\ \mathbf{r}_2^T]^T$, $\lambda=\sqrt{K_{\rm RD}}A$, $J_{\rm S}=\frac{E_{\rm S}}{\sigma^2}$, and $J_{\rm R}=\frac{E_{\rm R}}{\sigma^2}$. Let the diagonal entries in $\widehat{\mathbf{G}}_{\rm ABC}$ collect the estimated frequency response in data transmission phase for the link A → B. The associated estimation error is therefore obtained as $\widetilde{\mathbf{G}}_{\rm ABC}=\mathbf{G}_{\rm ABC}-\widehat{\mathbf{G}}_{\rm ABC}$, where $\widehat{\mathbf{G}}_{\rm ABC}=\bar{\boldsymbol{\Phi}}\widehat{\mathbf{G}}_{\rm AB}$, and $\widehat{\mathbf{G}}_{\rm AB}$ is of size $Q\times Q$, its diagonal elements designated as the estimated frequency response for the link A → B. From the channel estimate given in (5), the qth diagonal entry in $\widehat{\mathbf{G}}_{\rm AB}$ is given by $\widehat{G}_{\rm AB}(q)=\sum_{k=1}^{L_{\rm AB}}\widehat{\mathbf{h}}_{\rm AB}(k)\exp(-j2\pi(k-1)(q-1)/Q)$, $q=1,2,\ldots,Q$. Organizing the estimation errors and noise components into one vector, the input–output relation in (9) is rewritten as

$$\mathbf{r}=\widehat{\mathbf{G}}_{\rm d}\mathbf{y}+\mathbf{v} \quad (10)$$

where $\widehat{\mathbf{G}}=\begin{pmatrix}\widehat{\mathbf{G}}_{\rm SDd}\\ \lambda\sqrt{K_{\rm SR}}\widehat{\mathbf{G}}_{\rm SRd}\widehat{\mathbf{G}}_{\rm RDd}\end{pmatrix}$, vector $\mathbf{v}$ collects the channel estimation errors, additive Gaussian noise at the destination, and noise forwarded from the relay node, and it has the following form:

$$\mathbf{v}=\begin{pmatrix}\mathbf{v}_1\\ \mathbf{v}_2\end{pmatrix}=\begin{pmatrix}\widetilde{\mathbf{G}}_{\rm SDd}\\ \lambda\sqrt{K_{\rm SR}}\left(\widetilde{\mathbf{G}}_{\rm SRd}\widetilde{\mathbf{G}}_{\rm RDd}+\widetilde{\mathbf{G}}_{\rm SRd}\widehat{\mathbf{G}}_{\rm RDd}+\widehat{\mathbf{G}}_{\rm SRd}\widetilde{\mathbf{G}}_{\rm RDd}\right)\end{pmatrix}\mathbf{y}+\begin{pmatrix}\mathbf{n}_{\rm d}\\ \lambda\mathbf{G}_{\rm RDd}\mathbf{w}_{\rm d}+\mathbf{g}_{\rm d}\end{pmatrix} \quad (11)$$

3 Lower bound on the achievable rate

In this section, we derive a lower bound on the achievable rate for the cooperative OFDM system under consideration. The achievable rate for the training-aided system is given by [20,22],

$$C=\max_{p(\mathbf{y})}I(\mathbf{r},\widehat{\mathbf{G}}_{\rm d};\mathbf{y}) \quad (12)$$

where $p(\mathbf{y})$ is the probability distribution of $\mathbf{y}$, and $I(\mathbf{r},\widehat{\mathbf{D}};\mathbf{y})$ denotes the mutual information between the observation $\mathbf{r}$, channel estimates $\widehat{\mathbf{D}}_{\rm SDd}$, $\widehat{\mathbf{G}}_{\rm SRd}$, and $\widehat{\mathbf{G}}_{\rm RDd}$, and data vector $\mathbf{y}$. Since the statistical distribution of $\mathbf{v}$ in the received vector is difficult to characterize, we seek a lower bound on the achievable rate. Following similar steps in obtaining a tight lower bound on the achievable rate for a non-cooperative multiple-input multiple-output system [22,23], a lower bound for (12) can be found by imposing assumptions that the input data $\mathbf{y}$ is i.i.d zero-mean Gaussian, and vector $\mathbf{v}$ is Gaussian distributed with the same first- and second-order statistics that are specified by the random vector in (11). The lower bound can be evaluated as $C_{\rm lb}=\frac{1}{Q}\mathbb{E}\left[\log\det\left(\mathbf{I}+\mathbf{R}_{\mathbf{v}}^{-1}\widehat{\mathbf{G}}_{\rm d}\mathbf{R}_{\mathbf{y}}\widehat{\mathbf{G}}_{\rm d}^H\right)\right]$, where $\mathbf{R_y} = \mathbb{E}[\mathbf{yy}^H] = \alpha_d P_S \mathbf{I}_{MN}$, and $\mathbf{R_v} = \mathbb{E}[\mathbf{vv}^H]$. The bound is further evaluated as

$$C_{\text{lb}} = \frac{1}{Q}\mathbb{E}\left[\log\det\left(\mathbf{I} + \alpha_{\text{d}}P_{\text{S}}\mathbf{R}_{\mathbf{v}_1}^{-1}\widehat{\mathbf{G}}_{\text{SDd}}\widehat{\mathbf{G}}_{\text{SDd}}^H + \alpha_{\text{d}}P_{\text{S}}\lambda^2 K_{\text{SR}}\mathbf{R}_{\mathbf{v}_2}^{-1}\widehat{\mathbf{G}}_{\text{SRd}}\widehat{\mathbf{G}}_{\text{SRd}}^H\widehat{\mathbf{G}}_{\text{RDd}}\widehat{\mathbf{G}}_{\text{RDd}}^H\right)\right] \tag{13}$$

where the expectation is with respect to the random variables in $\widehat{\mathbf{G}}_{\text{SDd}}, \widehat{\mathbf{G}}_{\text{SRd}}$, and $\widehat{\mathbf{G}}_{\text{RDd}}$, matrices $\mathbf{R}_{\mathbf{v}_1}$ and $\mathbf{R}_{\mathbf{v}_i}$ are, respectively, the autocorrelation of $\mathbf{v}_1$ and $\mathbf{v}_2$,

$$\begin{aligned}\mathbf{R}_{\mathbf{v}_1} &= \mathbb{E}[\widetilde{\mathbf{G}}_{\text{SDd}}\mathbf{yy}^H\widetilde{\mathbf{G}}_{\text{SDd}}^H] + \mathbb{E}[\mathbf{n}_{\text{d}}\mathbf{n}_{\text{d}}^H]\\ &= \alpha_{\text{d}}P_{\text{S}}\mathbb{E}[\widetilde{\mathbf{D}}_{\text{SDd}}\widetilde{\mathbf{G}}_{\text{SDd}}^H] + \sigma^2\mathbf{I}_{MN} \qquad (14)\\ \mathbf{R}_{\mathbf{v}_2} &= \lambda^2 K_{\text{SR}}\alpha_{\text{d}}P_{\text{S}}\left(\mathbb{E}\left[\widetilde{\mathbf{G}}_{\text{SRd}}\widetilde{\mathbf{G}}_{\text{SRd}}^H\widetilde{\mathbf{G}}_{\text{RDd}}\widetilde{\mathbf{G}}_{\text{RDd}}^H\right]\right.\\ &\quad + \mathbb{E}\left[\widetilde{\mathbf{G}}_{\text{SRd}}\widetilde{\mathbf{G}}_{\text{SRd}}^H\widehat{\mathbf{G}}_{\text{RDd}}\widehat{\mathbf{G}}_{\text{RDd}}^H\right]\\ &\quad \left. + \mathbb{E}\left[\widehat{\mathbf{G}}_{\text{SRd}}\widehat{\mathbf{G}}_{\text{SRd}}^H\widetilde{\mathbf{G}}_{\text{RDd}}\widetilde{\mathbf{G}}_{\text{RDd}}^H\right]\right) + (\lambda^2+1)\sigma^2\mathbf{I}_{MN} \qquad (15)\end{aligned}$$

The matrices in (14) and (15) are diagonal. The lower bound on the training-based achievable rate in (13) becomes

$$C_{\text{lb}} = \frac{1}{N(M+1)}\sum_{i=1}^{MN}\mathbb{E}\left[\log\left(1 + \frac{\alpha_{\text{d}}P_{\text{S}}|\widehat{\mathbf{G}}_{\text{SDd}}(i)|^2}{R_{\mathbf{v}_1}(i)} + \frac{\alpha_{\text{d}}P_{\text{S}}K_{\text{SR}}A^2|\widehat{\mathbf{D}}_{\text{SRd}}(i)|^2|\widehat{\mathbf{G}}_{\text{RDd}}(i)|^2}{R_{\mathbf{v}_2}(i)}\right)\right] \tag{16}$$

The ith diagonal entry in $\widetilde{\mathbf{G}}_{\text{ABd}}$ is equal to the ith element in vector $\bar{\mathbf{\Phi}}\mathbf{V}_{\text{AB}}\widetilde{\mathbf{h}}_{\text{AB}}, i = 1,\ldots,MN$. From (6), the variance of the diagonal entries in $\widetilde{\mathbf{G}}_{\text{ABd}}$ is identical for a specified A→B link. Assuming each training symbol has identical power, the variance of the ith diagonal entry in $\widetilde{\mathbf{G}}_{\text{SDd}}, \widetilde{\mathbf{G}}_{\text{SRd}}$, and $\widetilde{\mathbf{G}}_{\text{RDd}}$ are therefore expressed, respectively, as

$$\mathbb{E}[|\widetilde{G}_{\text{SDd}}(i)|^2] = \frac{L_{\text{SD}}}{\alpha_{\text{t}}J_{\text{S}}N + L_{\text{SD}}} \tag{17}$$

$$\mathbb{E}[|\widetilde{G}_{\text{SRd}}(i)|^2] = \frac{L_{\text{SR}}}{\alpha_{\text{t}}J_{\text{S}}K_{\text{SR}}N + L_{\text{SR}}} \tag{18}$$

$$\mathbb{E}[|\widetilde{G}_{\text{RDd}}(i)|^2] = \frac{L_{\text{RD}}}{\beta_{\text{t}}J_{\text{R}}K_{\text{RD}}N + L_{\text{RD}}} \tag{19}$$

By revoking $\mathbb{E}[|\widehat{G}_{\text{ABd}}(i)|^2] = 1 - \mathbb{E}[|\widetilde{G}_{\text{ABd}}(i)|^2]$, $\widehat{G}_{\text{ABd}}(i)$ are thus i.i.d. CSCG random variables for all i, due to the deployed MMSE estimation. We have

$$\widehat{G}_{\text{SDd}}(i) \sim \mathcal{CN}\left(0, \frac{\alpha_{\text{t}}J_{\text{S}}N}{\alpha_{\text{t}}J_{\text{S}}N + L_{\text{SD}}}\right) \tag{20}$$

$$\widehat{G}_{\text{SRd}}(i) \sim \mathcal{CN}\left(0, \frac{\alpha_{\text{t}}J_{\text{S}}G_{\text{SR}}N}{\alpha_{\text{t}}J_{\text{S}}K_{\text{SR}}N + L_{\text{SR}}}\right) \tag{21}$$

$$\widehat{G}_{\text{RDd}}(i) \sim \mathcal{CN}\left(0, \frac{\beta_{\text{t}}J_{\text{R}}K_{\text{RD}}N}{\beta_{\text{t}}J_{\text{R}}K_{\text{RD}}N + L_{\text{RD}}}\right) \tag{22}$$

Recall $\lambda^2 = \frac{\beta_{\text{d}}J_{\text{R}}}{G_{\text{SR}}\alpha_{\text{d}}J_{\text{S}}+1}$. Inserting (17), (18), (19), (21), and (22) into (15), the covariance matrices of $\mathbf{v}_1$ and $\mathbf{v}_2$ are given, respectively, by

$$\mathbf{R}_{\mathbf{v}_1} = \frac{\alpha_{\text{d}}P_{\text{S}}L_{\text{SD}} + (\alpha_{\text{t}}J_{\text{S}}N + L_{\text{SD}})\sigma^2}{(\alpha_{\text{t}}J_{\text{S}}N + L_{\text{SD}})}\mathbf{I}_{MN}$$

$$\mathbf{R}_{\mathbf{v}_2} = \frac{\Upsilon}{(\alpha_{\text{t}}J_{\text{S}}K_{\text{SR}}N + L_{\text{SR}})(\beta_{\text{t}}J_{\text{R}}K_{\text{RD}}N + L_{\text{RD}})\,(K_{\text{SR}}\alpha_{\text{d}}J_{\text{S}}+1)}\mathbf{I}_{MN}$$

where

$$\begin{aligned}\Upsilon = {}&(K_{\text{SR}}K_{\text{RD}}\beta_{\text{d}}J_{\text{R}}\alpha_{\text{d}}P_{\text{S}})(L_{\text{SR}}L_{\text{RD}} + \alpha_{\text{t}}J_{\text{S}}K_{\text{SR}}L_{\text{RD}}N\\ &+ \beta_{\text{t}}J_{\text{R}}K_{\text{RD}}L_{\text{SR}}N) + \sigma^2(K_{\text{RD}}\beta_{\text{d}}J_{\text{R}} + K_{\text{SR}}\alpha_{\text{d}}J_{\text{S}} + 1)\\ &(\alpha_{\text{t}}J_{\text{S}}K_{\text{SR}}N + L_{\text{SR}})(\beta_{\text{t}}J_{\text{R}}K_{\text{RD}}N + L_{\text{RD}})\end{aligned}$$

The following normalized CSCG random variables are introduced as

$$\bar{\mathbf{G}}_{\text{SDd}}(i) = \frac{\widehat{\mathbf{G}}_{\text{SDd}}(i)}{\sqrt{\frac{\alpha_{\text{t}}J_{\text{S}}N}{(\alpha_{\text{t}}J_{\text{S}}N + L_{\text{SD}})}}} \sim \mathcal{CN}(0,1)$$

$$\bar{\mathbf{G}}_{\text{SRd}}(i) = \frac{\widehat{\mathbf{G}}_{\text{SRd}}(i)}{\sqrt{\frac{\alpha_{\text{t}}J_{\text{S}}K_{\text{SR}}N}{(\alpha_{\text{t}}J_{\text{S}}K_{\text{SR}}N + L_{\text{SR}})}}} \sim \mathcal{CN}(0,1)$$

$$\bar{\mathbf{G}}_{\text{RDd}}(i) = \frac{\widehat{\mathbf{G}}_{\text{RDd}}(i)}{\sqrt{\frac{\beta_{\text{t}}J_{\text{R}}K_{\text{RD}}N}{(\beta_{\text{t}}J_{\text{R}}K_{\text{RD}}N + L_{\text{RD}})}}} \sim \mathcal{CN}(0,1)$$

Since $\bar{\mathbf{G}}_{\text{ABd}}(i)$ has the same distribution $\forall i$, then C_{lb} in (16) is further evaluated by,

$$C_{\text{lb}} = \frac{M}{M+1}\mathbb{E}\left[\log\left(1 + B_1 x + B_2 yz\right)\right] \tag{23}$$

where x, y, and z are exponentially distributed, with probability density functions given by e^{-x}, e^{-y}, and e^{-z}, respectively, B_1 and B_2 are expressed as

$$B_1 = \frac{\alpha_{\text{t}}\alpha_{\text{d}}J_{\text{S}}^2 MN}{(g_1\alpha_{\text{t}} + g_2)}, \tag{24}$$

$$B_2 = \frac{\alpha_{\text{t}}\alpha_{\text{d}}J_{\text{S}}^2 MN^2 J_{\text{R}}^2 K_{\text{SR}}^2 K_{\text{RD}}^2 \beta_{\text{t}}\beta_{\text{d}}}{g_3\beta_{\text{t}}^2 + g_4\beta_{\text{t}} + g_5}, \tag{25}$$

and $g_k, k = 1, \ldots, 5$ are given as

$$\begin{aligned} g_1 =&(MN - L_{\rm SD})J_{\rm S}, g_2 = (M + MJ_{\rm S} + J_{\rm S})L_{\rm SD} \\ g_3 =&- K_{\rm SR}K_{\rm RD}^2\alpha_{\rm d}J_{\rm R}^2J_{\rm S}L_{\rm SR}N - K_{\rm RD}^2J_{\rm R}^2N(\alpha_{\rm t}J_{\rm S}K_{\rm SR}N + L_{\rm SR}) \\ g_4 =&- K_{\rm SR}K_{\rm RD}J_{\rm R}\alpha_{\rm d}J_{\rm S}\left(L_{\rm SR}L_{\rm RD} + \alpha_{\rm t}J_{\rm S}K_{\rm SR}L_{\rm RD}N\right) \\ &+ K_{\rm SR}K_{\rm RD}^2\alpha_{\rm d}J_{\rm R}^2J_{\rm S}L_{\rm SR}N(M+1) \\ &+ (\alpha_{\rm t}J_{\rm S}K_{\rm SR}N + L_{\rm SR})\left(K_{\rm RD}^2J_{\rm R}^2N(M+1)\right. \\ &\left.+MNK_{\rm RD}J_{\rm R}K_{\rm SR}\alpha_{\rm d}J_{\rm S} + MNK_{\rm RD}J_{\rm R} - L_{\rm RD}K_{\rm RD}J_{\rm R}\right) \\ g_5 =&K_{\rm SR}K_{\rm RD}J_{\rm R}\alpha_{\rm d}J_{\rm S}\left(L_{\rm SR}L_{\rm RD} + \alpha_{\rm t}J_{\rm S}K_{\rm SR}L_{\rm RD}N\right)(M+1) \\ &+ (\alpha_{\rm t}J_{\rm S}K_{\rm SR}N + L_{\rm SR})\left(K_{\rm RD}J_{\rm R}(M+1)\right. \\ &\left.+MK_{\rm SR}\alpha_{\rm d}J_{\rm S} + M\right)L_{\rm RD} \end{aligned} \quad (26)$$

4 Power allocation

In this section, we aim to find the power allocation factors between the training and data transmission phases at the source and relay nodes to maximize the derived bound on the achievable rate. Since the power allocation factors in the data transmission phases (α_d and $\beta_{\rm d}$) can be expressed as functions of the power allocation factors in training phase (α_t and $\beta_{\rm t}$), we can obtain optimal α_t and β_t by solving the following optimization problem

$$(\breve{\alpha}_t, \breve{\beta}_t) = \arg\max_{\breve{\alpha}_t, \breve{\beta}_t} C_{\rm lb} \quad (27)$$

It seems difficult to obtain a closed-form expression for the integral in (27). However, by noting that the integrand required for expectation in (23) is monotonically increasing with B_1 for fixed B_2 and vice versa, we can resort to a suboptimal solution by maximizing B_1 and B_2 separately. From (24), B_1 is not a function of $\beta_{\rm t}$ and is convex with respect to α_t, which can readily be determined by checking $\frac{d^2B_1}{d\alpha_{\rm t}^2} < 0$. The optimal value for α_t to maximize B_1 is then given by

$$\breve{\alpha}_{\rm t} = \frac{1}{g_1}\left(-g_2 + \sqrt{g_2^2 + g_1g_2(M+1)}\right) \quad (28)$$

On the other hand, B_2 is convex with respect to β_t, i.e., $\frac{d^2B_2}{d\beta_{\rm t}^2} < 0$. Inserting (28) into (25), a suboptimal β_t to maximize B_2 can be obtained by

$$\beta_{\rm t0} = \left.\frac{-g_5 + \sqrt{g_5^2 + g_5(M+1)\left(g_3(M+1) + g_4\right)}}{g_3(M+1) + g_4}\right|_{\alpha_{\rm t} = \breve{\alpha}_{\rm t}} \quad (29)$$

5 Lower bound on the achievable rate and power allocation for the case of cascaded channel estimation

So far, we have assumed that both relay and destination nodes are equipped with channel estimators and the channel estimates in S→R and R→D links are obtained, respectively, by the training symbols sent at the source and the relay nodes. In this section, we assume that only the destination is equipped with a channel estimator. Therefore, it is the duty of the destination to obtain an estimate of the overall relaying S→R→D link using the training symbols sent from the source. In describing this alternative scheme, we try to use the same variables, whenever possible, as in prior sections, or we use $\breve{(\cdot)}$ whenever necessary.

Inserting the training symbols periodically (similar placement as in prior discussion), the OFDM block transmitted at the source node is written as $\mathbf{d} = \left[x_{\rm S}(1), y(1), y(2), \ldots, y(M), x_{\rm S}(2), y(M+1), \ldots, y(2M), \ldots, x_{\rm S}(N), y\left((\breve{N}-1)M+1\right), \ldots, y(M\breve{N})\right]$, where $x_{\rm S}(i), i = 1, \ldots, \breve{N}$ are the training symbols. The power constraint at the source is $P_{\rm S} = \frac{1}{(M+1)\breve{N}}(\mathbf{x}_{\rm S}^H\mathbf{x}_{\rm S} + \mathbb{E}[\,\mathbf{y}^H\mathbf{y}]\,)$, $\mathbf{x}_{\rm S}^H\mathbf{x}_{\rm S}/\breve{N} = \alpha_{\rm t}P_{\rm S}, \mathbb{E}[\,\mathbf{y}^H\mathbf{y}]\,/(M\breve{N}) = \alpha_{\rm d}P_{\rm S}$. To remove inter-block interference, a cyclic prefix is added at the beginning of the transmitted vectors at the source node. Let $\mathbf{d}_{\rm wcp}$ denote the transmitted vector with cyclic prefix, the length of which is $\breve{Q} + L_{\rm cp}$, where $\breve{Q} = (M+1)\breve{N}$ is the total length of training symbols and data information (length of vector $\mathbf{d}$), $L_{\rm cp}$ is the length of the cyclic prefix, and "wcp" stands for "with cyclic prefix." The received signal at the relay can be expressed as $\sqrt{K_{\rm SR}}\mathbf{h}_{\rm SR} \otimes \mathbf{d}_{\rm wcp}$, in addition to the noise at relay, where $\otimes$ denotes the operation of convolution. Scaling the received signal by factor $\breve{A} = \sqrt{\frac{P_r}{(\breve{Q}+L_{\rm cp})\left(K_{\rm SR}(\breve{N}\alpha_t + M\breve{N}\alpha_d)P_S + \sigma^2\right)}}$, the relay forwards the signal to the destination. The received signal at the destination is, therefore, $\breve{A}\sqrt{K_{\rm SR}K_{\rm RD}}\mathbf{h}_{\rm RD} \otimes \mathbf{h}_{\rm SR} \otimes \mathbf{d}_{\rm wcp}$ and subject to the noise forwarded by the relay and noise at the destination.

The overall convolution channel in the relaying link S→R→D is $\mathbf{h}_{\rm SRD} = \mathbf{h}_{\rm SR} \otimes \mathbf{h}_{\rm RD}$, the length of which is $L_{\rm SRD} = L_{\rm SR} + L_{\rm RD} - 1$. The minimum length for the cyclic prefix in order to avoid inter-block interference is $\min(L_{\rm cp}) = \max(L_{\rm SD}, L_{\rm SRD}) - 1$. Removing the cyclic prefix and taking FFT, the received OFDM block (including training and data transmission phases) at the destination is $\breve{A}\sqrt{K_{\rm SR}K_{\rm RD}}\mathbf{G}_{\rm SRD}\mathbf{d}$, plus the noise forwarded from the relay and noise at the destination, where $\mathbf{G}_{\rm SRD}$ is the diagonal matrix of frequency response, and the qth diagonal entry $G_{\rm SRD}(q) = \sum_{k=1}^{L_{\rm SRD}} h_{\rm SRD}(k)\exp(-j2\pi(k-1)(q-1)/\breve{Q})$, $q = 1, 2, \ldots, \breve{Q}$.The estimate of the overall relaying channel is

$$\begin{aligned} \widehat{\mathbf{h}}_{\rm SRD} =& \eta\sqrt{K_{\rm SR}}\mathbf{R}_{h_{\rm SRD}}\mathbf{V}_{\rm SRD}^H\mathbf{\Psi}^T\mathbf{X}_{\rm S}^H \\ &\times\left(\eta^2K_{\rm SR}\mathbf{X}_{\rm S}\mathbf{\Psi}\mathbf{V}_{\rm SRD}\mathbf{R}_{\rm SRD}\mathbf{V}_{\rm SRD}^H\mathbf{\Psi}^T\mathbf{X}_S^H + \sigma^2\mathbf{I}\right)^{-1}\mathbf{z}_{\rm SRD} \end{aligned} \quad (30)$$

where $\eta = \sqrt{K_{\mathrm{RD}}}\breve{A}$, $\mathbf{R}_{\mathbf{h}_{\mathrm{SRD}}} = \mathbb{E}[\,\mathbf{h}_{\mathrm{SRD}}\mathbf{h}_{\mathrm{SRD}}^H]$, $\mathbf{z}_{\mathrm{SRD}} = \mathbf{\Psi d}$, $\mathbf{\Psi}$ is the selection matrix collecting $\ell_i = 1 + (\ell - 1)(M + 1), i = 1, \ldots, \breve{N}$ rows from $\mathbf{I}_{\breve{Q}}$. Let $\widetilde{\mathbf{h}}_{\mathrm{SRD}} = \mathbf{h}_{\mathrm{SRD}} - \widehat{\mathbf{h}}_{\mathrm{SRD}}$ denote the channel estimate error for the relaying link. The error covariance matrix $\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SRD}}} = \mathbb{E}[\,\widetilde{\mathbf{h}}_{\mathrm{SRD}}\widetilde{\mathbf{h}}_{\mathrm{SRD}}^H]$ is diagonal with the ith diagonal entry given by

$$\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SRD}}}(i) = \frac{R_{h_{\mathrm{SRD}}}(i)}{1 + \eta^2 K_{\mathrm{SR}} N \alpha_t J_s \sigma^2 R_{h_{\mathrm{SRD}}}(i)} \tag{31}$$

where $R_{h_{\mathrm{SRD}}}(i)$ is the ith diagonal entry in $\mathbf{R}_{\mathbf{h}_{\mathrm{SRD}}}$. Organizing the channel estimate errors and noise components into vector $\mathbf{f}$, the received signals at the destination is written as

$$\mathbf{r} = \begin{pmatrix} \widehat{\mathbf{G}}_{\mathrm{SDd}} \\ \eta\sqrt{K_{\mathrm{SR}}}\widehat{\mathbf{G}}_{\mathrm{SRDd}} \end{pmatrix} \mathbf{y} + \mathbf{f} \tag{32}$$

where

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{pmatrix} = \begin{pmatrix} \widetilde{\mathbf{G}}_{\mathrm{SD}} \\ \eta\sqrt{K_{\mathrm{SR}}}\widetilde{\mathbf{G}}_{\mathrm{SRDd}} \end{pmatrix} \mathbf{y} + \begin{pmatrix} \mathbf{n}_{\mathrm{d}} \\ \eta \mathbf{G}_{\mathrm{RDd}}\mathbf{w}_{\mathrm{d}} + \mathbf{g}_{\mathrm{d}} \end{pmatrix}. \tag{33}$$

Since the overall cascaded channel is no longer Gaussian, the derivation of a lower bound on the achievable rate for the system in (32) is intractable, as discussed in [24]. However, an approximate lower bound can be found by imposing similar assumptions as in [22,23]. Specifically, assuming that the input data $\mathbf{y}$ are i.i.d zero-mean Gaussian, and vector $\mathbf{f}$ is Gaussian distributed with the same first- and second-order statistics specified by the random vector in (33), we obtain an approximation of the bound as follows: $\breve{C}_{\mathrm{lb}} \approx \frac{1}{\breve{Q}}\mathbb{E}\left[\log\det\left(\mathbf{I} + \mathbf{R}_{\mathbf{f}}^{-1}\widehat{\mathbf{G}}_{\mathrm{d}}\mathbf{R}_{\mathbf{y}}\widehat{\mathbf{G}}_{\mathrm{d}}^H\right)\right]$, where $\mathbf{R}_{\mathbf{y}} = \mathbb{E}[\,\mathbf{y}\mathbf{y}^H] = \alpha_d P_{\mathrm{S}}\mathbf{I}_{MN}$, and $\mathbf{R}_{\mathbf{f}} = \mathbb{E}[\,\mathbf{f}\mathbf{f}^H]$. Expanding the determinant, the approximation of the bound is further expressed as

$$\breve{C}_{\mathrm{lb}} \approx \frac{1}{\breve{Q}}\mathbb{E}\left[\log\det\left(\mathbf{I} + \alpha_{\mathrm{d}}P_{\mathrm{S}}\mathbf{R}_{\mathbf{f}_1}^{-1}\widehat{\mathbf{G}}_{\mathrm{SDd}}\widehat{\mathbf{G}}_{\mathrm{SDd}}^H + \alpha_{\mathrm{d}}P_{\mathrm{S}}\eta^2 K_{\mathrm{SR}}\mathbf{R}_{\mathbf{f}_2}^{-1}\widehat{\mathbf{G}}_{\mathrm{SRDd}}\widehat{\mathbf{G}}_{\mathrm{SRDd}}^H\right)\right]. \tag{34}$$

The matrices in (34) are all diagonal, with $\mathbf{R}_{\mathbf{f}_1} = \mathbf{R}_{\mathbf{v}_1}, \mathbf{R}_{\mathbf{f}_2} = \eta^2 K_{\mathrm{SR}}\alpha_{\mathrm{d}}P_S\mathbb{E}[\,\widetilde{\mathbf{G}}_{\mathrm{SRDd}}\widetilde{\mathbf{G}}_{\mathrm{SRDd}}^H] + (\eta^2 + 1)\sigma^2\mathbf{I}$. The bound thus becomes

$$\breve{C}_{\mathrm{lb}} = \frac{1}{\breve{Q}}\sum_{i=1}^{MN}\mathbb{E}\left[\log\left(1 + \frac{\alpha_{\mathrm{d}}P_{\mathrm{S}}|\widehat{G}_{\mathrm{SDd}}(i)|^2}{R_{\mathbf{f}_1}(i)} + \frac{\alpha_{\mathrm{d}}P_{\mathrm{S}}\eta^2 K_{\mathrm{SR}}|\widehat{G}_{\mathrm{SRDd}}(i)|^2}{R_{\mathbf{f}_2}(i)}\right)\right] \tag{35}$$

where $R_{\mathbf{f}_2}(i) = \alpha_{\mathrm{d}}P_S\mathrm{tr}(\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SD}}})$, $R_{\mathbf{f}_2}(i) = \eta^2 K_{\mathrm{SR}}\alpha_{\mathrm{d}}P_S\mathrm{tr}(\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SRD}}}) + (\eta^2 + 1)\sigma^2$, and $\mathrm{tr}(\cdot)$ denotes the trace operation. Note that $\widehat{G}_{\mathrm{SRDd}}(i)$ is Gaussian and $\mathbb{E}[\,|\widehat{G}_{\mathrm{SRDd}}(i)|^2] = \mathrm{tr}(\mathbf{R}_{\mathbf{h}_{\mathrm{SRD}}}) - \mathrm{tr}(\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SRD}}})$. Introducing random variables $\bar{G}_{\mathrm{SRDd}}(i) = \frac{1}{\sqrt{\mathrm{tr}(\mathbf{R}_{\mathbf{h}_{\mathrm{SRD}}}) - \mathrm{tr}(\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SRD}}})}}\widehat{G}_{\mathrm{SRDd}}(i)$, we have

$$\breve{C}_{\mathrm{lb}} = \frac{1}{Q}\sum_{i=1}^{MN}\mathbb{E}\left[\log\left(1 + \frac{\alpha_{\mathrm{d}}P_{\mathrm{S}}\left(\mathrm{tr}(\mathbf{R}_{\mathbf{h}_{\mathrm{SD}}}) - \mathrm{tr}(\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SD}}})\right)|\bar{G}_{\mathrm{SDd}}(i)|^2}{\alpha_{\mathrm{d}}P_{\mathrm{S}}\mathrm{tr}(\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SD}}}) + \sigma^2} + \frac{\alpha_{\mathrm{d}}P_S\eta^2 K_{\mathrm{SR}}\left(\mathrm{tr}(\mathbf{R}_{\mathbf{h}_{\mathrm{SRD}}}) - \mathrm{tr}(\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SRD}}})\right)|\bar{G}_{\mathrm{SRDd}}(i)|^2}{\alpha_{\mathrm{d}}P_S\eta^2 K_{\mathrm{SR}}\mathrm{tr}(\mathbf{R}_{\widetilde{\mathbf{h}}_{\mathrm{SRD}}}) + (\eta^2 + 1)\sigma^2}\right)\right] \tag{36}$$

There seems no analytical solutions to the power allocation factors α_{t} and α_{d} to maximize (36). The solutions can be obtained by maximizing the bound numerically.

6 Numerical results

In this section, we present numerical results to elaborate the derived bound on the achievable rate and the possible improvements through the proposed power allocation schemes. Unless specified otherwise, the plots are obtained for the disintegrated channel estimation scheme described in Section 2.1. For simplicity, the angle between links $\mathrm{S} \to \mathrm{R}$ and $\mathrm{R} \to \mathrm{D}$ in the relay system is chosen as $\theta = 60°$, and the relay is located at equal distance from the source and the destination, except for the case where the effect of relay locations is considered. The following power allocation schemes for the training and data transmission are considered:

- The proposed power allocation (Proposed-PA) strategy, with the values of $\breve{\alpha}_{\mathrm{t}}$ and $\breve{\beta}_{\mathrm{t}}$ given, respectively, in (28) and (29).
- Uniform power allocation (Uniform-PA) with $\alpha_{\mathrm{t}} = \beta_{\mathrm{t}} = 1$.
- Numerical power allocation (Numerical-PA), with the values of α_{t} and β_{t} obtained by a numerically exhaustive search to maximize the lower bound on the training-based achievable rate in (23).

6.1 Performance of power allocation schemes

Table 1 lists the power allocation factors at the source and relay nodes by the Proposed-PA and Numerical-PA for different SNRs. The channel lengths are as follows: $L_{\mathrm{SD}} = 4, L_{\mathrm{SR}} = L_{\mathrm{RD}} = 3$. The OFDM block length is $Q = N(M + 1) = 164$, with $N = 4, M = 40$. Figure 2 plots the bounds on the achievable rate for different power allocation schemes. As a benchmark, we also plot the achievable rate of a genie-aided coherent OFDM relay system with perfect channel state information, which is obtained from (13) by inserting $\alpha_{\mathrm{t}} = 0, \lambda_{\mathrm{coh}}^2 = \lambda^2\big|_{\alpha_{\mathrm{t}}=\beta_{\mathrm{t}}=0}, \mathbf{R}_{\mathbf{v}_1} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$, and $\mathbf{R}_{\mathbf{v}_2} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2(\lambda_{\mathrm{coh}}^2\mathbf{G}_{\mathrm{RDd}}\mathbf{G}_{\mathrm{RDd}}^H + 1)\mathbf{I}\right)$.

Table 1 Comparison of the power allocation factors for Proposed-PA and Numerical-PA

SNR in dB	Proposed-PA $\breve{\alpha}_t$	Numerical-PA α_t	Proposed-PA $\breve{\beta}_t$	Numerical-PA β_t
0	7.42	7.36	6.6	6.31
6	6.15	6.31	5.43	5.26
12	5.74	5.26	5.07	5.26
18	5.63	5.26	4.97	5.26
24	5.60	5.26	4.95	5.26
30	5.60	5.26	4.94	5.26

The achievable rate of coherent OFDM relay system is thus given by

$$C_{\text{coh}} = \mathbb{E}\Bigg[\log\det\Bigg(\mathbf{I} + \frac{(M+1)P_{\text{S}}}{M\sigma^2}\mathbf{G}_{\text{SDd}}\mathbf{G}_{\text{SDd}}^H + \frac{(M+1)K_{\text{R}}DP_{\text{S}}\lambda_{\text{coh}}^2}{M\sigma^2} \times\mathbf{G}_{\text{SRd}}\mathbf{G}_{\text{SRd}}^H(\lambda_{\text{coh}}^2\mathbf{G}_{\text{RDd}}\mathbf{G}_{\text{RDd}}^H + \mathbf{I})^{-1}\mathbf{G}_{\text{RDd}}\mathbf{G}_{\text{RDd}}^H\Bigg)\Bigg] \quad (37)$$

where a ratio $(M+1)/M$ is applied, because it is not necessary for the genie-aided coherent relay system to allocate power to the channel training. From Figure 2, we observe that the SNR gap between the Uniform-PA and coherent case is about 2.8–3 dB, and the achievable rate of the coherent transmission is 0.7–1.1 bits/s/Hz or 12–37% higher than that of the Uniform-PA. The Numerical-PA scheme reduces the gap by about 1.6 dB and provides 0.3–0.6 bits/s/Hz or 5–16% improvement for the bound on the achievable rate over the Uniform-PA scheme. It is also observed that the performance of the Proposed-PA scheme is almost identical to that of the Numerical-PA scheme.

To demonstrate the performance in the case of cascaded channel estimation described in Section 5, we obtain the power allocation factors by numerically maximizing the bound derived in (36) and plot the maximized bound. As shown in Figure 2, the curve is marked by "Proposed-PA (cascaded estimation)." In this case, the number of training symbols is $N = \max(L_{\text{SD}}, L_{\text{SR}} + L_{\text{RD}}) - 1 = 5$, and the OFDM block length is chosen as 205. The bound in (36) with uniform power allocation $\alpha_t = 1$ is also included, marked by "Uniform-PA (cascaded estimation)." It can been observed that, with appropriate power allocation factors, the lower bounds on the achievable rates remain almost the same, regardless of whether the channel gains in the relaying link are estimated as two separate channels or as one overall channel.

In the following, we demonstrate the impact of some practical parameters such as OFDM block length, relay location, and channel lengths on the bound of the achievable rate.

6.2 OFDM block length

In Figure 3, we investigate how the length of OFDM block impacts the derived bound on the achievable rate

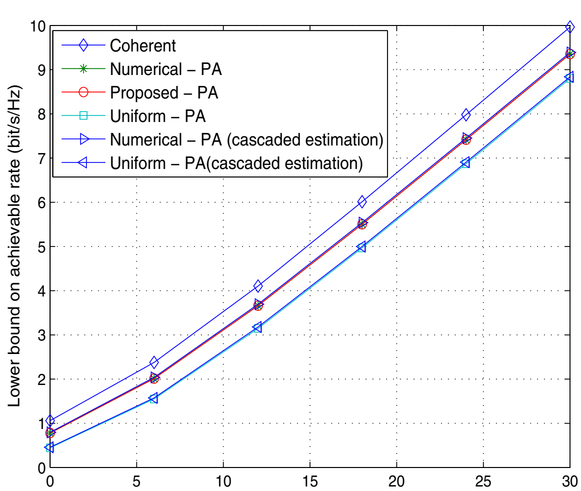

Figure 2 Comparison of Numerical-PA, Proposed-PA, and Uniform-PA schemes ($Q = 164$).

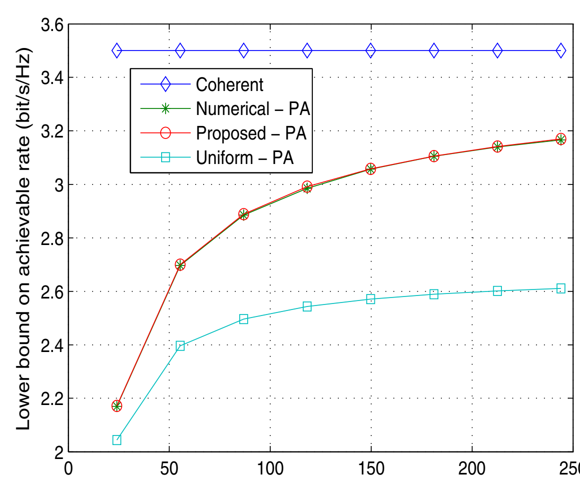

Figure 3 Effect of block length on the bound of achievable rate (SNR= 10 dB).

for SNR = 10 dB. Plots indicate that the Proposed-PA and Numerical-PA schemes are consistently better than the Uniform-PA scheme regardless of the block lengths.

6.3 Relay location

In Figure 4, we investigate the effect of relay location on the derived bound. The level of SNR is fixed at 10 dB. The ratio $K_{SRD} = K_{SR}/K_{RD}$ (in dB) is introduced to reflect the relative relay locations. The more negative this ratio in dB, the closer the relay is placed to the destination. On the other hand, positive dB values of this ratio indicate that the relay node is closer to the source. The plots in Figure 4 suggest that the cooperative system achieves higher lower bounds when the relay is close to the destination. The Proposed-PA scheme can provide an improvement of 17–19% on the bound, in comparison to that of the Uniform-PA scheme depending on the relay location.

6.4 Channel lengths

In Figure 5, we investigate the effect of channel lengths on the achievable rate. We consider the following three scenarios:

- Scenario 1: $L_{SD} = 30,\ L_{SR} = 3,\ L_{RD} = 3$
- Scenario 2: $L_{SD} = 3,\ L_{SR} = 30,\ L_{RD} = 3$
- Scenario 3: $L_{SD} = 3,\ L_{SR} = 3,\ L_{RD} = 30$

For these channel configurations, the number of pilot symbols increased to $N = \max(L_{RD}, L_{RD}, L_{RD}) = 30$. We choose the block length $Q = 30(M+1) = 150$ with $M = 4$, and SNR =10 dB. It can be observed that the system achieves almost the same bounds for Scenarios 2 and 3, and a little less for Scenario 1. These can be explained through the values of B_1 and B_2. One can observe from (24) and (25) that B_1 and B_2 capture, respectively, the equivalent SNR in the direct link and in the relaying link with imperfect channel gains. In Scenarios 2 and 3, B_1 and B_2 yield identical values, resulting in a similar performance for the mutual information. On the other hand, for Scenario 1, the value of B_1 drops by 89% while B_2 increases by 32%. The larger drop in B_1 results in a little more decrease of the bound in Scenario 1.

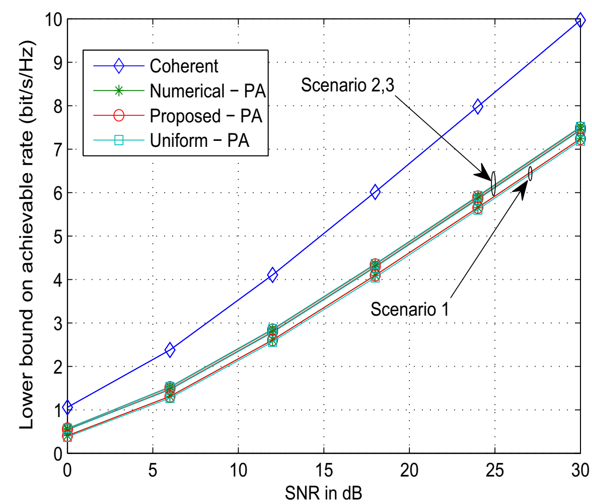

Figure 5 Bound on achievable rate for different channel length combinations.

6.5 Channel estimate and power allocation

To illustrate the performance of the Proposed-PA scheme on channel estimates, we examine the mean square error (MSE) in the channel estimate with and without power allocations. Figure 6 plots the MSE of channel estimate

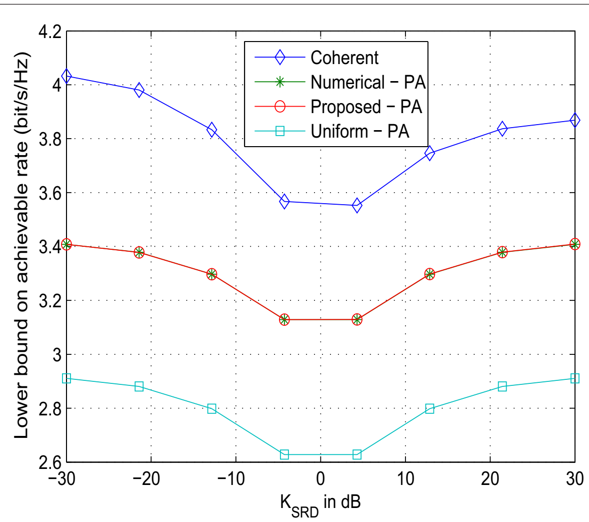

Figure 4 Effect of relay location on the bound of achievable rate.

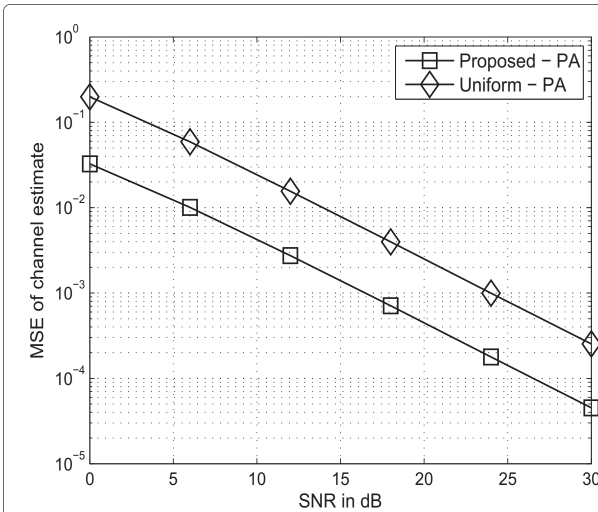

Figure 6 MSE of channel estimate.

for the S→D link using the Proposed-PA and Uniform-PA schemes. The number of training symbols is $N = 4$, and the OFDM block length is $Q = 164$. Plots indicate that the Proposed-PA scheme provides about 7 dB SNR gain over the Uniform-PA scheme. Smaller errors in the channel estimate help improve the training-based achievable rates for the Proposed-PA scheme.

7 Conclusion

In this article, we have investigated the achievable rates for a single-relay OFDM system with imperfect channel estimation. We first obtained a lower bound on the achievable rate and then used this bound to optimally allocate power between the training and data transmission phases. Since the optimum solution does not yield a closed-form expression, we proposed a suboptimal scheme by sequentially maximizing the terms in the integrand of the lower bound. Monte Carlo simulations demonstrate that the proposed power allocation scheme brings improvements of 5–19% depending on the SNR and relay locations in the bound on achievable rate, depending on the relay location and level of SNR.

8 Endnote

[a] In Section 5, we will further consider an alternative scheme in which only the destination is equipped with a channel estimator and therefore the overall cascaded channel is estimated.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Electrical Engineering and Computer Science, Wichita State University, 1845 Fairmount, Wichita, KS 67260, USA. [2]Faculty of Engineering, Ozyegin University Istanbul, Istanbul, Turkey.

Received: 5 September 2011 Accepted: 21 November 2012
Published: 31 December 2012

References

1. A Sendonaris, E Erkip, B Aazhang, User cooperation diversity—Part I: system description. IEEE Trans. Commun. **51**, 1927–1938 (2003)
2. K Azarian, H El Gamal, P Schniter, On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels. IEEE Trans. Inf. Theory. **51**, 4152–4172 (2005)
3. JN Laneman, DNC Tse, GW Wornell, Cooperative diversity in wireless networks: Efficient protocols and outage behavior. IEEE Trans. Inf. Theory. **50**, 3062–3080 (2004)
4. M Gastpar, M Vetterli, *On the capacity of wireless networks: the relay case*, vol. 3, (New York, June 2002), pp. 1577–1586
5. RU Nabar, H Bölcskei, FW Kneubühler, Fading relay channel: performance limits and space-time signal design. IEEE J. Sel. Areas Commun. **22**, 1099–1109 (2004)
6. A El Gamal, S Zahedi, Capacity of a class of relay channels with orthogonal components. IEEE Trans. Inf. Theory. **51**(5), 1815–1817 (2005)
7. B Wong, J Zhang, A Host-Madsen, On the capacity of MIMO relay channels. IEEE Trans. Inf. Theory. **51**(1), 29–43 (2005)
8. H Böckei, RU Nabar, O Oyman, AJ Paulraj, Capacity scaling laws in MIMO relay networks. IEEE Trans. Wirel. Commun. **5**, 1433–1443 (2006)
9. I Marić, RD Yates, in *The 38th Asilomar Conference on Signals, Systems and Computers*. bandwidth and power allocaiton for cooperative startegies in Gaussian realy networks, (Pacific Grove, CA, USA, November 2004), pp. 1907–1911
10. X Deng, AM Haimovich, Power allocation for cooperative relaying in wireless networks. IEEE Commun. Lett. **9**(11), 994–996 (2005)
11. J Zhang, MC Gursoy, Achievable rates and resource allocation strategies for imperfectly-known fading relay channels. EURASIP J. Wirel. Commun. Netw. **2009** (2009)
12. O-S Shin, A Chan, HT Kung, V Tarokh, Design of an OFDM co-operative diversity system. IEEE Trans. Veh. Technol. **56**(4), 2203–2215 (2007)
13. I Hammerstrom, A Wittneben, Power allocation schemes for amplify-and-forward MIMO-OFDM relay links. IEEE Trans. Wirel. Commun. **8**(8), 4507–4512 (2007)
14. H Mheidat, M Uysal, N Al-Dhahir, Equalization techniques for distributed space-time block codes with amplify-and-forward relaying. IEEE Trans. Signal Process. **55**(5), 1839–1852 (2007)
15. Y Ma, N Yi, R Tafazolli, Bit and power loading for OFDM-based three-node relaying communications. IEEE Trans. Signal Process. **56**(7), 3236–3247 (2008)
16. M Ibrahimi, B Liang, in *Proceedings of the IEEE International Conference on Communication (ICC)*. Efficient power allocation in cooperative OFDM system with channel variation, (Beijing, China, May 2008), pp. 3022–3028
17. O Amin, M Uysal, Optimal bit and power loading for amplify-and-forward cooperative OFDM systems. IEEE Trans. Wirel. Commun. **10**(3), 772–781 (2011)
18. J-H Wang, Y-B Lin, YT Su, in *Proceedings of the IEEE Signal Processing Workshop in Signal Advances in Wireless Communication (SPAWC)*. Resource allocation and relay selection for cooperative OFDMA networks with imperfect channel information, (San Francisco, June 2011), pp. 286–290
19. H Ochiai, P Mitran, V Tarokh, Variable-rate two-phase collaborative communication protocols for wireless networks. IEEE Trans. Inf. Theory. **52**(9), 4299–4313 (2006)
20. S Adireddy, L Tong, H Viswanathan, Optimal placement of training for frequency-selective block-fading channels. IEEE Trans. Inf. Theory. **48**(8), 2338–2353 (2002)
21. SM Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. (Prentice-Hall, Inc. , Upper Saddle River, NJ 07458, 1993)
22. B Hassibi, BM Hochwald, How much training is needed in multiple-antenna wireless links?. IEEE Trans, Inf. Theory. **49**(4), 951–963 (2003)
23. TM Cover, JA Thomas, *Elements of Information Theory*. (Wiley, New York, 1991)
24. TA Lamahewa, P Sadeghi, X Zhou, On lower bounding the information capacity of amplify and forward wireless relay channels with channel estimation errors. IEEE Trans. Wirel. Commun. **10**(7), 2075–2079 (2011)

doi:10.1186/1687-1499-2012-372
Cite this article as: Ding and Uysal: **Achievable data rates and power allocation for frequency-selective fading relay channels with imperfect channel estimation.** *EURASIP Journal on Wireless Communications and Networking* 2012 **2012**:372.