

The Impact of Unknown Correlated Rayleigh Fading on the Information Rates of Dual-Hop Amplify and Forward Relay Channels

Mohammadali Mohammadi[†], Parastoo Sadeghi[‡], Tharaka A. Lamahewa[‡], and Mehrdad Ardebilipour[†]

[†]K. N. Toosi University of Technology, Tehran, Iran,

[‡]The Australian National University, Canberra, Australia, ACT 0200.

Email: M.A.Mohammadi@ee.kntu.ac.ir

Abstract—We consider a dual-hop amplify-and-forward relaying system where the channel state is unknown to both the transmitter and the receiver. Our objective is to derive the penalty of the information rate due to the unknown channel state information. We focus on determining the compound effects of *fixed-gain* relaying and unknown time-varying channels on the achievable information rates of the system. We show that penalty of the information rate due to the unknown channels decays with block length and signal-to-noise ratio while it increases with normalized Doppler shift. We also derive a closed-form expression for the information penalty when the block length is large (asymptotic behavior). This derivation shows that the penalty of not knowing the mobile-fixed-mobile dual-hop relay channel can be equal to or higher than that in the direct mobile-to-mobile source-destination transmission. The additional uncertainty should be taken into account when designing cooperative systems with many relays.

I. INTRODUCTION

A. Motivation

It is well known that the availability and accuracy of channel state information (CSI) at the transmitting or receiving ends of a wireless communication channel can fundamentally impact the achievable information rates [1]. In practical systems, CSI is not available prior to transmission and can only be obtained by sacrificing a fraction of resources, such as time and power, for channel estimation or feedback [2]. From a theoretical and practical viewpoint, it is therefore of great interest to find out the ultimate limits on reliable information transfer when the channel is unknown a priori at any node, without restricting oneself to a particular form of channel estimation or feedback.

A large body of work has been devoted to the capacity analysis of channels in the absence of CSI for a number of *direct* source-to-destination channel models with no relays [1], [3]–[5]. As acknowledged in [4], the exact capacity analysis for unknown channels is a challenging problem even for very simple direct channel models. In many cases, only information rate upper and lower bounds have been provided [4].

Nowadays, there is a great interest in cooperative communications as a promising solution for future high data rate ad-hoc networks [6]. As building blocks of cooperative communications, relay channels and relaying techniques have been the subject of intensive research recently [7]. Capacity analysis of relay channels is a non-trivial task even when the

CSI is perfectly available at all nodes and several studies have focused on this problem in the past few decades [8]–[11]. It is interesting to determine the compound effects of relaying and unknown (possibly time-varying) channels on the achievable information rates of cooperative systems. One fundamental question, for example, is whether detrimental effects of additional unknown channels in the system can outweigh well-known advantages of using relays. To the best of our knowledge, this is still an open problem.

B. Contributions and Distinctions Compared to Related Work

In this paper, we study information rate bounds for communication between a mobile source node (S) to a mobile destination node (D) via a fixed relay node (R) using amplify and forward (AF) relaying [7], where the S-R and R-D channels are unknown at all nodes and vary in time according to a correlated complex-Gaussian process. This is a first step towards better understanding the impact of relaying in unknown channels on achievable information rates.

In this context, three key contributions of our paper are summarized as follows:

- We derive bounds on the information rates of AF dual-hop relaying over unknown and correlated time-varying Rayleigh fading channels. While our approach resembles the one proposed in [12], we highlight that due to non-Gaussianity of the compound S-R-D channel and destination noise, certain non-trivial difficulties arise in the computation of entropies that are dealt with numerically in this paper. The lower bound is obtained for the case of finite-dimensional modulations, while the upper bound is limited to finite-dimensional modulations with uncorrelated inputs.
- We study the penalty on information rates due to unknown CSI and present the results in terms of the block length, signal-to-noise ratio (SNR), and normalized Doppler shift.
- We identify the asymptotic behavior of the penalty on information rate due to unknown CSI for high SNR and large block lengths. This analysis reveals that the penalty of not knowing the dual-hop channel is equal to or higher than that in the direct channel. This additional uncertainty can adversely impact the benefits of cooperative diversity.

II. SYSTEM MODEL

We consider the simplest non-trivial scenario of time-varying relaying networks, in which a mobile source node (S) communicates to a mobile destination node (D) via a fixed relay (R). Assuming flat fading, the received signal by the relay at time k , y_k , is given by

$$y_k = h_{1k}s_k + n_{1k}, \quad (1)$$

where s_k is the transmitted signal, h_{1k} is the S-R channel modeled as a zero mean complex Gaussian (ZMCG) random process with power $\sigma_1^2 = 1/d_{sr}^a$, where d_{sr} is the S-R distance and a is the path-loss exponent [13]. n_{1k} is complex additive white Gaussian noise (AWGN) with power σ_n^2 at the relay node. Without trying to estimate the channel or to decode the transmitted signal, the relay amplifies y_k and transmits it to the destination which receives

$$r_k = Ah_{2k}y_k + n_{2k} = h_k s_k + n_k, \quad (2)$$

where h_{2k} is the R-D channel which is modeled as a ZMCG process with power $\sigma_2^2 = 1/d_{rd}^a$, where d_{rd} is the R-D distance. n_{2k} is a zero mean AWGN with power σ_n^2 at the destination. The overall dual-hop channel is $h_k = Ah_{2k}h_{1k}$ and $n_k = Ah_{2k}n_{1k} + n_{2k}$ represents the overall noise at the destination. A is the amplification factor by the relay, which we assume is given by

$$A = \sqrt{\frac{\mathcal{E}_r}{E[|y_k|^2]}} = \sqrt{\frac{\mathcal{E}_r}{\mathcal{E}_s \sigma_1^2 + \sigma_n^2}}, \quad (3)$$

where \mathcal{E}_r and \mathcal{E}_s are the relay and source transmitting power, respectively. Relays using such A are called *fixed-gain* (FG) [14], because A is constant during transmission and is determined using long-term channel statistics. This is consistent with the assumption that the relay does not have CSI. Although systems with FG relays may not perform as well as the systems with variable gain relays [14], their low complexity makes them attractive from a practical point of view.

We consider transmission of a length- L block of symbols, $\mathbf{s} = [s_1, s_2, \dots, s_L]^T$. The corresponding length- L block of received signal, denoted by $\mathbf{r} = [r_1, r_2, \dots, r_L]^T$, is of the form

$$\mathbf{r} = \mathbf{S}\mathbf{h} + \mathbf{n}, \quad (4)$$

where $\mathbf{S} = \text{diag}(s_1, s_2, \dots, s_L)$, $\mathbf{h} = A[h_{11}h_{21}, \dots, h_{1L}h_{2L}]^T$, $\mathbf{n} = A\mathbf{H}_2\mathbf{n}_1 + \mathbf{n}_2 = [n_1, n_2, \dots, n_L]^T$, $\mathbf{H}_2 = \text{diag}(h_{21}, h_{22}, \dots, h_{2L})$, $\mathbf{n}_1 = [n_{11}, n_{12}, \dots, n_{1L}]^T$, and $\mathbf{n}_2 = [n_{21}, n_{22}, \dots, n_{2L}]^T$. Since receiver noises are white, noise vectors \mathbf{n}_1 and \mathbf{n}_2 are complex Gaussian distributed in L dimensions with zero mean and covariance matrices $\mathbf{R}_{\mathbf{n}_i} = \sigma_n^2 \mathbf{I}$ for $i = 1, 2$. Throughout the paper, we assume that all wireless fading channels and all receiver noises are independent of each other.

The channel vectors \mathbf{h}_1 and \mathbf{h}_2 represent mobile to fixed channels. Using the results in [15] for two dimensional isotropic scattering (Clarke's model), the (k, k') entry of $\mathbf{R}_{\mathbf{h}_i}$, the auto-correlation coefficient between h_{ik} and $h_{ik'}$ for $i = 1, 2$, can be expressed as

$$[\mathbf{R}_{\mathbf{h}_i}]_{kk'} = E[h_{ik}h_{ik'}^*] = \sigma_i^2 J_0(2\pi f_{d_i}(k - k')T_s), \quad (5)$$

where J_0 is the zeroth order Bessel function of the first kind, f_{d_i} is the Doppler frequency shift in the channel i , and $1/T_s$ is the sampling rate. Therefore, $f_{D_i} = f_{d_i}T_s$ is the normalized Doppler shift. Finally, the (k, k') entry of $\mathbf{R}_{\mathbf{h}}$, the auto-correlation coefficient between h_k and $h_{k'}$ is [15]

$$[\mathbf{R}_{\mathbf{h}}]_{kk'} = A^2 \sigma_1^2 \sigma_2^2 J_0(2\pi f_{d_1}(k - k')T_s) \times J_0(2\pi f_{d_2}(k - k')T_s). \quad (6)$$

III. BOUNDS ON INFORMATION RATE AND THE PENALTY DUE TO UNKNOWN CSI

The mutual information between \mathbf{r} and $\{\mathbf{s}, \mathbf{h}\}$ can be expressed by using the chain rule as follows:

$$\begin{aligned} \mathcal{I}(\mathbf{r}; \mathbf{s}, \mathbf{h}) &= \mathcal{I}(\mathbf{r}; \mathbf{s}) + \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) \\ &= \mathcal{I}(\mathbf{r}; \mathbf{h}) + \mathcal{I}(\mathbf{r}; \mathbf{s}|\mathbf{h}). \end{aligned} \quad (7)$$

Thus, $\mathcal{I}(\mathbf{r}; \mathbf{s})$, the mutual information between channel input \mathbf{s} and channel output \mathbf{r} without CSI, can be expressed as

$$\mathcal{I}(\mathbf{r}; \mathbf{s}) = \mathcal{I}(\mathbf{r}; \mathbf{s}, \mathbf{h}) - \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) \quad (8a)$$

$$= \mathcal{I}(\mathbf{r}; \mathbf{h}) + \mathcal{I}(\mathbf{r}; \mathbf{s}|\mathbf{h}) - \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}). \quad (8b)$$

Note that $\mathcal{I}(\mathbf{r}; \mathbf{s}|\mathbf{h})$ in (8b) is the mutual information with perfect CSI. Therefore, $P_{\Delta_{dh}} = \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) - \mathcal{I}(\mathbf{r}; \mathbf{h})$ is the information rate penalty due to unknown CSI. Since $\mathcal{I}(\mathbf{r}; \mathbf{h})$ is nonnegative, $\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s})$ can be interpreted as an upper bound on the penalty due to unknown CSI, which is given as

$$\begin{aligned} P_{\Delta_{dh}} &\leq \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) = \mathcal{H}(\mathbf{r}|\mathbf{s}) - \mathcal{H}(\mathbf{r}|\mathbf{h}, \mathbf{s}) \\ &= \mathcal{H}(\mathbf{r}|\mathbf{s}) - \mathcal{H}(\mathbf{n}), \end{aligned} \quad (9)$$

where $\mathcal{H}(\cdot)$ represents the differential entropy function. In the sequel, we study $\mathcal{H}(\mathbf{r}|\mathbf{s})$ in more detail for the purpose of deriving an upper bound on the information rate penalty $P_{\Delta_{dh}}$ and also for obtaining upper and lower bounds on the information rate without CSI in dual-hop AF relay channels. We will highlight main differences with [12].

A. Properties of $\mathcal{H}(\mathbf{r}|\mathbf{s})$

Using the fact that the entropy of a random vector with a given covariance matrix is upper bounded by the entropy of a Gaussian random vector with the same covariance matrix [16], we have

$$\mathcal{H}(\mathbf{r}|\mathbf{s}) \leq E_s \left\{ \log \det \left(\pi e (\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H + \mathbf{R}_{\mathbf{n}}) \right) \right\}, \quad (10)$$

where the elements of $\mathbf{R}_{\mathbf{h}}$ were defined in (6) and the (k, k') entry of $\mathbf{R}_{\mathbf{n}}$, the auto-correlation coefficient between n_k and $n_{k'}$, can be expressed as

$$\begin{aligned} [\mathbf{R}_{\mathbf{n}}]_{kk'} &= E \{ (Ah_{2k}n_{1k} + n_{2k})(Ah_{2k'}n_{1k'} + n_{2k'})^* \} \\ &= (A^2\sigma_2^2 + 1)\sigma_n^2 \delta_{k,k'} = \eta\sigma_n^2 \delta_{k,k'}, \end{aligned} \quad (11)$$

where $\eta \triangleq (A^2\sigma_2^2 + 1)$ and $\delta_{k,k'} = 1$ if $k = k'$ and $\delta_{k,k'} = 0$ otherwise. Therefore, the received noise vector entries are uncorrelated and consequently $\mathbf{R}_{\mathbf{n}} = \eta\sigma_n^2 \mathbf{I}$, which results in

$$\mathcal{H}(\mathbf{r}|\mathbf{s}) \leq E_s \left\{ \log \det \left(\pi e (\mathbf{S}\mathbf{R}_{\mathbf{h}}\mathbf{S}^H + \eta\sigma_n^2 \mathbf{I}) \right) \right\}. \quad (12)$$

Before proceeding, we highlight one main difference for the evaluation of $\mathcal{H}(\mathbf{r}|\mathbf{s})$ with that for a direct S-D time-varying Rayleigh fading channel, which was studied in [12]. In

particular in [12], \mathbf{r} given \mathbf{s} was complex Gaussian distributed with the covariance matrix $(\mathbf{S}\mathbf{R}_{h_1}\mathbf{S}^H + \mathbf{R}_{n_1})$. Therefore, the inequalities shown in (10) and (12) were in fact equalities in [12] and only Jensen's inequality was used to upper bound $\mathcal{H}(\mathbf{r}|\mathbf{s})$. In this paper, however, due to non-Gaussianity of the overall dual-hop channel and destination noise, we have first upper bounded $\mathcal{H}(\mathbf{r}|\mathbf{s})$ using the entropy of a Gaussian random vector and now apply Jensen's inequality to arrive at

$$\mathcal{H}(\mathbf{r}|\mathbf{s}) \leq L \log(\pi e \eta \sigma_n^2) + \log \det \left(\mathbf{I} + \frac{\mathbf{R}_h}{\eta \sigma_n^2} E \{ \mathbf{S}\mathbf{S}^H \} \right),$$

where the equality is obtained for constant modulus (CM) signalling, *i.e.*, $\|s_\ell\|^2 = \mathcal{E}_s$ for $\ell = 1, 2, \dots, L$. Finally, the upper bound on $\mathcal{H}(\mathbf{r}|\mathbf{s})$ is

$$\mathcal{H}(\mathbf{r}|\mathbf{s}) \leq L \log(\pi e \eta \sigma_n^2) + \sum_{\ell=1}^L \log(1 + \varrho \lambda_\ell), \quad (13)$$

where $\varrho = \frac{\mathcal{E}_s}{\eta \sigma_n^2}$ and λ_ℓ for $\ell = 1, 2, \dots, L$ are the eigenvalues of \mathbf{R}_h . Normalizing the mutual information to the number of symbols L and using (9), the rate of penalty on information rate due to unknown CSI is given by

$$\frac{1}{L} P_{\Delta_{ah}} \leq \log(\pi e \eta \sigma_n^2) + \frac{1}{L} \sum_{\ell=1}^L \log(1 + \varrho \lambda_\ell) - \frac{1}{L} \mathcal{H}(\mathbf{n}). \quad (14)$$

We now highlight another important difference between the penalty rate presented here and that presented in [12]. More specifically, the penalty rate in [12] was upper bounded as

$$\frac{1}{L} P_{\Delta} \leq \frac{1}{L} \sum_{\ell=1}^L \log \left(1 + \frac{\mathcal{E}_s}{\sigma_n^2} \lambda_{1\ell} \right), \quad (15)$$

where $\lambda_{1\ell}$ for $\ell = 1, 2, \dots, L$ were the eigenvalues of \mathbf{R}_{h_1} . This was again due to the Gaussian channel and noise models at the destination which resulted in $\log(\pi e \sigma_n^2)$ and $\mathcal{H}(\mathbf{n})/L$ in (14) to cancel each other out. However, here, \mathbf{n} is not Gaussian and hence $\mathcal{H}(\mathbf{n})$ remains in the formulation of the penalty rate. This will be one main source of difficulty for information rate analysis in this paper. Computation of $\mathcal{H}(\mathbf{n})$ will be described in Section III-C.

B. Information Rate Lower and Upper Bounds

Let us first write $\mathcal{I}(\mathbf{r}, \mathbf{s})$ as

$$\mathcal{I}(\mathbf{r}; \mathbf{s}) = \mathcal{H}(\mathbf{r}) - \mathcal{H}(\mathbf{r}|\mathbf{s}). \quad (16)$$

We note that (13) provides us with an upper bound for $\mathcal{H}(\mathbf{r}|\mathbf{s})$. Therefore, if $\mathcal{H}(\mathbf{r})$ can be somehow evaluated numerically, we arrive at the following lower bound on the information rate

$$\frac{1}{L} \mathcal{I}(\mathbf{r}, \mathbf{s}) \geq \frac{1}{L} \mathcal{H}(\mathbf{r}) - \log(\pi e \eta \sigma_n^2) - \frac{1}{L} \sum_{\ell=1}^L \log(1 + \varrho \lambda_\ell). \quad (17)$$

In Section III-C, we will discuss how $\mathcal{H}(\mathbf{r})$ can be evaluated numerically.

Referring to (16), we note if we can somehow evaluate $\mathcal{H}(\mathbf{r}|\mathbf{s})$ numerically and then analytically upper bound $\mathcal{H}(\mathbf{r})$, we obtain an upper bound on the information rate. In Section III-C, we will discuss how $\mathcal{H}(\mathbf{r}|\mathbf{s})$ can be evaluated numerically. One upper bound for $\mathcal{H}(\mathbf{r})$ is as follows:

$$\mathcal{H}(\mathbf{r}) \leq \log \det(\pi e \mathbf{R}_r) \leq \log \prod_{i=1}^L \pi e [\mathbf{R}_r]_{ii} \quad (18)$$

where \mathbf{R}_r is the covariance matrix of the received vector \mathbf{r} and according to (2) its elements are given by

$$[\mathbf{R}_r]_{kk'} = E[r_k r_{k'}^*] = E[s_k s_{k'}^*] E[h_k h_{k'}^*] + E[n_k n_{k'}^*]. \quad (19)$$

The first inequality in (18) holds because the entropy of a random vector is upper bounded by the entropy of a Gaussian random vector with the same covariance matrix and the second inequality follows from Hadamard's inequality. Hadamard's inequality is achieved with equality if $[\mathbf{R}_r]$ is a diagonal matrix. According to (11) the noise n_k is uncorrelated. By further assuming that the input symbols are uncorrelated ($E[s_k s_{k'}^*] = 0$ for $k \neq k'$) the second equality is ensured to obtain tighter upper bounds on information rate. In summary, the upper bound on information rate is written as

$$\frac{1}{L} \mathcal{I}(\mathbf{r}, \mathbf{s}) \leq \log(\pi e \eta \sigma_n^2 (1 + \rho)) - \frac{1}{L} \mathcal{H}(\mathbf{r}|\mathbf{s}). \quad (20)$$

where $\rho = \frac{\mathcal{E}_s}{\sigma_n^2} c_1$ and $c_1 = A^2 \sigma_1^2 \sigma_2^2 / (A^2 \sigma_2^2 + 1)$

C. Numerical Evaluation of $\mathcal{H}(\mathbf{n})$, $\mathcal{H}(\mathbf{r})$, and $\mathcal{H}(\mathbf{r}|\mathbf{s})$

Based on the derivations of the previous two subsections, we note that computation of the information rate penalty due to not knowing the channel or the information lower and upper bounds cannot be performed without evaluating $\mathcal{H}(\mathbf{n})$, $\mathcal{H}(\mathbf{r})$, and $\mathcal{H}(\mathbf{r}|\mathbf{s})$, respectively. Equation (2) indicates that forwarding the received signal by the relay introduces non-Gaussian compound noises and signals at the destination. Unfortunately, finding exact analytical expression for the probability distributed function (pdf) of these multivariate random vectors is very involved, if not impossible. As a result, the entropy functions cannot be evaluated analytically. Here, we focus on numerical methods, borrowed from statistical machine learning and physics literature, to approximate entropies of arbitrary high-dimensional multivariate random vectors.

Most entropy estimation schemes are based on the estimation of the pdf. This is followed by the substitution of these estimates to the expression for entropy. This method is called "plug-in" estimation [17]. In this paper we use the Gaussian mixture (GM) model [18] to estimate the pdf and consequently the entropy because of its high precision.

1) *Entropy Estimation Using GM Model*: Given a set of N samples in d dimensions and a family \mathcal{F} of pdfs on \mathcal{R}^d , GM model estimates the pdf $f(\mathbf{x}) \in \mathcal{F}$ as a linear combination of normal Gaussian pdfs (kernels), that is most likely to have generated the given samples [19]. So, for the multivariate random vectors, GM model can be written as

$$\hat{f}(\mathbf{x}) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k), \quad (21)$$

where $0 \leq \pi_i \leq 1$ for $i = 1, \dots, K$, $\sum_{i=1}^K \pi_i = 1$, $\{\pi_i\}_{i=1}^K$ are the a priori probabilities of each kernel (mixing coefficients), K is the number of kernels, and $\{\mu_i, \Sigma_i\}_{i=1}^K$ are the mean vector and covariance matrix of the Gaussian kernels. Obtaining the optimal set of the parameters is usually posed in terms of maximizing the log-likelihood of the pdf

to be estimated. It is well known that maximum likelihood estimation cannot be determined analytically but it can be approximated by the expectation-maximization (EM) algorithm [18]. In summary, given an initial estimate $\pi_k^{(0)}$, $\mu_k^{(0)}$, $\Sigma_k^{(0)}$, EM algorithm iteratively converges to a local maximum of the likelihood function. We refer the reader to [18] for more detailed description of GM. The justification of this approximation is provided in Section V.

IV. ASYMPTOTIC BEHAVIOR OF INFORMATION RATE PENALTY

In this section, we investigate the asymptotic behavior of (14) and (17) for large block lengths L . The log term of the (14) and (17) is independent of block length L . Moreover, Table II indicates that the estimated entropies $\mathcal{H}(\mathbf{n})$ and $\mathcal{H}(\mathbf{r})$ become nearly constant for long blocks. By Szegő theorem [20], we have the asymptotic expression for the L -dependent term of these equations as

$$I = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L \log(1 + \varrho \lambda_l) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log(1 + \varrho H(w)) dw, \quad (22)$$

where $H(w)$ is the power spectral density (PSD) function of the fading process \mathbf{h} . Considering the complexity of the integral, it is difficult to derive closed-form expressions for (22), when the corresponding PSD of $H(w)$ follows that given in [15]. However, [12] showed that in direct S-D transmission, using a uniform spectrum with the same fading rate as that in Clarke's model yields nearly the same information rate penalty. Thus, considering the ideal uniform spectrum with the same cutoff frequency as in Clarke's model for each link ($H_i(w) = \frac{\sigma_i^2}{2f_{D_i}}, |w| < 2\pi f_{D_i}$), the overall PSD of the relay channel is

$$H(w) = \begin{cases} 2\gamma T_1 & -\alpha \leq w < \alpha \\ \gamma(\beta - |w|) & \alpha \leq |w| < \beta \\ 0 & |w| > \beta \end{cases} \quad (23)$$

where $T_i = 2\pi f_{D_i}$, $\beta = T_1 + T_2$, $\alpha = T_2 - T_1$, and $\gamma = \eta c_1 / 4f_{D_1} f_{D_2}$. Without loss of generality, by assuming $f_{D_2} \geq f_{D_1}$ the asymptotic expression of (14) is

$$\begin{aligned} \tilde{P}_{\Delta_{dh}} &= \log(\pi \eta \sigma_n^2) + 2(f_{D_2} - f_{D_1}) \log\left(1 + \frac{\pi \rho}{f_{D_2}}\right) + \\ &4f_{D_1} \left[\left(1 + \frac{f_{D_2}}{\pi \rho}\right) \log\left(1 + \frac{\pi \rho}{f_{D_2}}\right) - 1 \right] - \frac{1}{L} \mathcal{H}(\mathbf{n}). \end{aligned} \quad (24)$$

The asymptotic expression for (17) is similarly derived.

It is interesting to compare the asymptotic information rate penalty of direct S-D transmission, \tilde{P}_{Δ} derived in [12], with that of dual-hop AF scheme derived here. For this purpose and to avoid the need for numerical computation of $\mathcal{H}(\mathbf{n})$ in (24), we use a simplifying high-SNR assumption and assume that the relay is close to the source. Therefore, the noise amplification factor at relay is small (*i.e.*, $\sigma_2^2/\sigma_1^2 < 1$) and the relay noise is negligible ($n_{1k} = 0$ for all k). (24) becomes

$$\begin{aligned} \tilde{P}_{\Delta_{dh}} &\approx 4f_{D_1} \left[\left(1 + \frac{f_{D_2}}{\pi \eta \rho}\right) \log\left(1 + \frac{\pi \eta \rho}{f_{D_2}}\right) - 1 \right] + \\ &2(f_{D_2} - f_{D_1}) \log\left(1 + \frac{\pi \eta \rho}{f_{D_2}}\right), \end{aligned} \quad (25)$$

TABLE I
RMSE OF GM METHOD FOR d -COMPLEX-VARIATE NORMAL DISTRIBUTION

d	10	20	40	60
GM	5.65E-8	6.25E-6	3.76E-5	8.26E-5

TABLE II
ESTIMATED ENTROPY OF THE RECEIVED NOISE $\mathcal{H}(\mathbf{n})/L$ IN AF SYSTEM ($\mathcal{E}/\sigma_n^2 = 10$ dB)

f_{D_2}	d_{sr}	L				
		10	30	50	70	100
0.006	0.25	2.1795	2.1711	2.1645	2.1609	2.1484
0.006	0.75	5.0980	5.0972	5.0971	5.0954	5.0953
0.01	0.25	2.1787	2.1713	2.1657	2.1602	2.1534
0.01	0.75	5.0977	5.0972	5.0970	5.0963	5.0960

We further assume equal Doppler shifts for S-R and R-D links *i.e.*, $f_{D_1} = f_{D_2} = f_D$ and equal power allocation to S and R ($\mathcal{E}_r = \mathcal{E}_s = \mathcal{E}/2$), where \mathcal{E} is the total transmission power.

Since in the dual-hop scenario both source and destination are mobile, we need to consider a direct mobile-to-mobile S-D transmission, for a fair comparison. However, only the penalty for fixed-to-mobile direct transmission was provided in [12]. To overcome this, using similar arguments as in [12], we use a uniform Doppler power spectrum for the S-D transmission with the maximum Doppler shift of $f_{D_{dir}} = f_{D_1} + f_{D_2} = 2f_D$, corresponding to S and D moving towards each other. By allocating the total power \mathcal{E} to the source, the penalty becomes

$$\tilde{P}_{\Delta} \approx 4f_D \log\left(1 + \frac{1}{4f_D} \frac{\mathcal{E}}{\sigma_n^2}\right) \quad (26)$$

Taking the ratio of (25) (with $f_{D_1} = f_{D_2} = f_D$ and $\mathcal{E}_r = \mathcal{E}_s = \mathcal{E}/2$) and (26) as $\frac{\mathcal{E}}{\sigma_n^2} \rightarrow \infty$ results in equal asymptotic penalty of not knowing the channel in direct and dual-hop transmission. In reality, $\tilde{P}_{\Delta_{dh}}$ is bigger when the relay noise is not negligible.

It is important to note that the benefits of using relays in providing diversity is adversely (and often non-negligibly) affected by the increased penalty of not knowing additional time-varying channels and amplified noise at the relay, which should be taken into account when designing cooperative relaying in mobile wireless systems.

V. NUMERICAL RESULTS

In this section we study the normalized information rate penalty and information rate bounds in an AF system. Since the entropy need to be computed numerically, we first justify the proposed entropy estimation method through the simulations. By assuming a normalized distance between the source and the destination ($d_{sd} = 1$), we consider three different relay positions. In position-1, the relay is closer to the source ($d_{sr} = 0.25, d_{rd} = 0.75$). In position-2, the relay is in the middle of the source and destination ($d_{sr} = d_{rd} = 0.5$). In position-3, we have $d_{sr} = 0.75$ and $d_{rd} = 0.25$. Path-loss exponent is $a = 3$. We consider a carrier frequency of 900 MHz and a sampling rate of 8000 Hz (*i.e.*, GSM system). Moreover, in our scenarios we assume that source and destination have the same speeds and we use the normalized Doppler shifts of 0.006 and 0.01 corresponding to mobile speeds of 60 km/h and 100 km/h, respectively. In the simulations, we generate the correlated mobile-fixed channel coefficients according to Clarke's model [21].

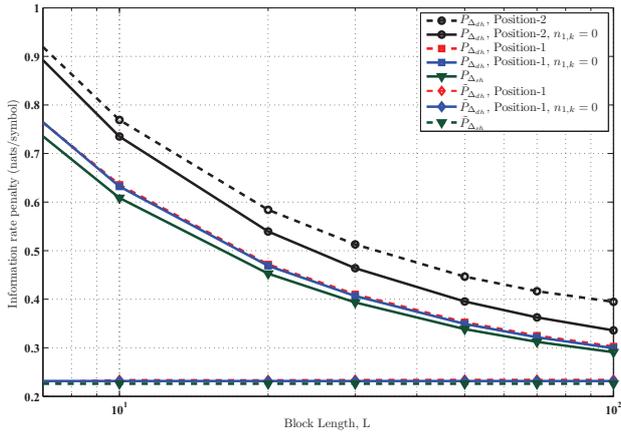


Fig. 1. Comparison of the exact and approximated bounds on $P_{\Delta_{dh}}$ versus L for CM signaling ($\mathcal{E}/\sigma_n^2 = 10$ dB, $f_D = 0.01$, $f_{D_{dir}} = 0.02$).

A. Entropy Estimation Results

To evaluate the accuracy of entropy estimation techniques, we consider a finite sample size $N = 10^5$ of a d -complex-variate Gaussian random distribution with zero mean and identity covariance matrix, which results in the normalized analytical entropy 2.1447¹. We compare the outputs of GM with the analytical result under root mean square error (RMSE) criterion. The results are reported in Table I for GM method with $K = 10$ kernels.

By using the GM method with $K = 10$, Table II shows the normalized estimated entropy of the received noise $\mathcal{H}(\mathbf{n})/L$ for an AF system versus L for SNR per symbol $\mathcal{E}/\sigma_n^2 = 10$ dB, two different relay positions and $f_{D1} = f_{D2} = 0.006$ and 0.01. The following conclusions are drawn from Table II:

- 1) The closer the relay to the destination, the higher the entropy of the received noise. This can be interpreted as follows: For high SNR, the amplification factor at relay node is $A^2\sigma_2^2 \approx \sigma_2^2/\sigma_1^2$. So, when relay is closer to the destination this factor becomes higher and according to Theorem 8.6.4 of [16] the entropy increases. Note that when the relay is closer to the source, the amplification factor is small and the noise entropy is nearly the same as the AWGN noise entropy (*i.e.*, 2.1447).
- 2) The entropy decreases as the block length L increases and becomes nearly constant for long blocks.

The details of numerical evaluation of $\mathcal{H}(r|s)$ are deferred to the appendix A.

B. Information Rate Penalty and Information Rate

Fig. 1 depicts a comparison between the exact penalty of the information rate (14) and the approximated one at which $n_1 = 0$, for $\mathcal{E}/\sigma_n^2 = 10$ dB for position-1 and position-2. It is clear that the higher distance between the source and relay, the higher the penalty of the information rates due to the unknown CSI. From the block length view, Fig. 1 confirms the intuition that increasing the sequence of observations improves the learning of channel dynamics, leading to a reduced loss in information rate and consequently alleviates the information

¹Normalized entropy is the entropy normalized by the number of dimensions.

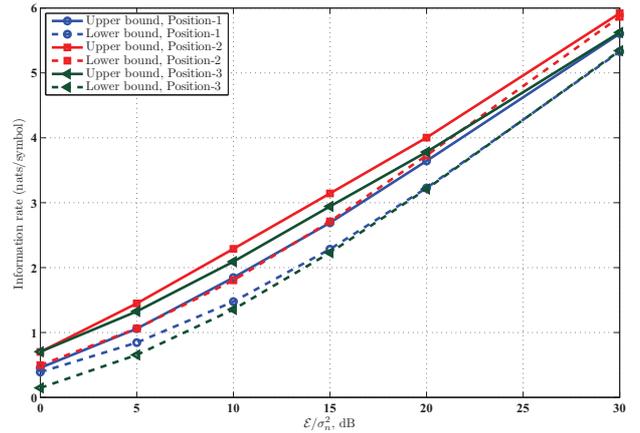


Fig. 2. Information rate bounds versus \mathcal{E}/σ_n^2 for CM signaling ($L = 8$)

rate penalty. Moreover, as expected, the penalty of not knowing CSI approaches the asymptotic bounds in (24) as the block length L increases.

Note that assuming zero noise at the relay changes the value of η which appears in the (24). But here, we are only interested in the impact of relay noise on the difference between $\log(\pi e \eta \sigma_n^2)$ and $\frac{1}{L} \mathcal{H}(\mathbf{n})$. Hence, $\eta = 1$ is used in the first term in (24) and the same $\eta = (A^2\sigma_2^2 + 1)$ as in the case where relay noise is non-zero is used in the second and third term in (24). This enables us to compare the impact of relay noise on the information rate penalty in fair fashion, *i.e.*, for the same effective SNR at the destination.

The difference between the exact and approximated values of the penalty is negligible in position-1. The reason is that the lower the distance between the source and the relay, the lower the noise amplification factor at relay and thus the relay noise is negligible. However, in position-2, where the amplification factor is higher, the actual penalty is considerably higher than the case where $n_{1,k}$ is assumed to be zero. Bounds on the information rates penalty due to the unknown CSI, in a direct transmission system are also presented in Fig. 1 for $f_{D_{dir}} = 2f_D$. To obtain a fair comparison between the penalty of the direct transmission system and that of the AF relaying system, the sum of the transmit power at source and relay nodes in the AF system is set equal to the direct system. Comparison of direct system versus relaying system reveals that the information rate penalty of the dual-hop relaying system is more than that of the direct link only and that the penalty depends on the relay position.

Fig. 2 shows the achievable information rate of AF dual-hop relaying over unknown channels, versus SNR for CM BPSK and $f_D = 0.01$ for the three relay positions described above. We use Clarke's spectrum and block length $L = 8$. It is observed that the information rate bounds become tighter for high SNRs. The tightest bounds are obtained for position-2 that are converging to the same values for high SNRs. We attribute this to the equal power allocation to the source and relay which is optimal for this symmetric case.

VI. CONCLUSION

The information rate penalty and the bounds of the information rates for a dual-hop AF system with unknown CSI have

been studied. A GM model have been used to approximate the entropy of the non-Gaussian received signal and noise at the destination. A closed-form expression for the asymptotic behavior of the penalty in terms of the Doppler shift and SNR was derived. The analysis pointed out to non-negligible penalty of not knowing the dual-hop channel and relay noise amplification.

VII. ACKNOWLEDGEMENTS

This work has been supported by Iran Telecommunication Research Center (ITRC), project no. 18243/500 and the Australian Research Council's Discovery Projects funding scheme (Project no. DP0984950). This work was performed while the co-author Tharaka Lamahewa was at the Australian National University.

REFERENCES

- [1] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [2] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [3] I. Abou-Faycal, M. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.
- [4] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.
- [5] F. Rusek, A. Lozano, and N. Jindal, "Mutual information of IID complex Gaussian signals on block Rayleigh-faded channels," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, Austin, TX, USA, June 2010, pp. 300–304.
- [6] A. Scaglione, D. Goeckel, and J. Laneman, "Cooperative communications in mobile ad hoc networks," *IEEE Signal Process. Mag.*, vol. 23, no. 5, pp. 18–29, Sept. 2006.
- [7] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [8] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, no. 1, pp. 120–154, 1971.
- [9] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [10] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, Mar. 2005.
- [11] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.
- [12] X. Deng and A. Haimovich, "Achievable rates over time-varying Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 55, no. 7, pp. 1397–1406, July 2007.
- [13] T. S. Rappaport, *Wireless Communications, Principles and Practice*. 2nd edition. Upper Saddle River, NJ: Prentice Hall, 2002.
- [14] M. Hasna and M. S. Alouini, "A performance study of dual-hop transmissions with fixed gain relays," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1963–1968, Nov. 2004.
- [15] C. S. Patel, G. L. Stüber, and T. G. Pratt, "Statistical properties of amplify and forward relay fading channels," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 1–9, Jan. 2006.
- [16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 2006.
- [17] J. Beirlant, E. J. Dudewicz, L. Györfi, and E. van der Meulen, "Non-parametric entropy estimation: an overview," *Intern. J. Math. Stat. Sci.*, vol. 6, no. 1, pp. 17–39, June 1997.
- [18] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [19] P. Hall and S. C. Morton, "On the estimation of the entropy," *Ann. Inst. Statist. Math.*, vol. 45, pp. 69–88, 1993.
- [20] U. Grenander and G. Szegő, *Toeplitz Forms and their Applications*. University of California Press, Berkeley and Los Angeles, 1958.
- [21] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, vol. 47, pp. 957–1000, 1968.

TABLE III
ESTIMATED $\mathcal{H}(\mathbf{r}|\mathbf{s})/L$ IN AF SYSTEM FOR $\mathcal{E}/\sigma_n^2 = 10$ dB

d_{sr} \ L	2	5	7	10
0.25	3.0807	2.7095	2.6218	2.5527
0.5	3.8000	3.2462	3.1276	3.0363
0.75	5.8126	5.2222	5.0880	4.9867

APPENDIX A

NUMERICAL EVALUATION OF $\mathcal{H}(\mathbf{r}|\mathbf{s})$

In an AF dual-hop relay system characterized by (4), the conditional entropy of \mathbf{r} given \mathbf{s} , $\mathcal{H}(\mathbf{r}|\mathbf{s})$, for a general M -ary modulation scheme can be numerically found as follows:

- 1) generate correlated channel coefficients \mathbf{h}_1 and \mathbf{h}_2 , according to Clarke's model [21]
- 2) generate AWGN noise vectors at the relay and destination nodes with independent entries
- 3) generate M^L different realizations of the length- L input signal, denoted by \mathbf{s}'
- 4) for each realization of the input signal, obtain the received signal r_k using (2) and then using the GM model, $\mathcal{H}(\mathbf{r}|\mathbf{s} = \mathbf{s}')$ is approximated.
- 5) calculate the normalized entropy by dividing the sum of the $\mathcal{H}(\mathbf{r}|\mathbf{s} = \mathbf{s}')$ from the pervious step, by LM^L .

Table III shows the normalized estimated entropy $\mathcal{H}(\mathbf{r}|\mathbf{s})/L$ for a dual-hop AF system versus block length L , for $\mathcal{E}/\sigma_n^2 = 10$ dB, $f_D = 0.01$ and three relay positions. Note that we have specialized our simulation results in Table III for CM BPSK signaling. The increase in $\mathcal{H}(\mathbf{r}|\mathbf{s})/L$ with respect to the relay position is interpreted in a similar way as for $\mathcal{H}(\mathbf{n})/L$ in Table (II).