Space-Time Diversity Enhancements Using Collaborative Communications

Patrick Mitran, *Student Member, IEEE*, Hideki Ochiai, *Member, IEEE*, and Vahid Tarokh, *Senior Member, IEEE*

Abstract

The use of the spatial dimension is known to greatly increase the reliability of quasi-static (i.e., non-ergodic) wireless channels. In this paper, it is demonstrated that most of this gain can also be achieved through collaborative communications with single-antenna/multiple-antenna nodes when there is one receiving agent. In particular, for the single antenna case, communication is considered to take place between clusters of nearby nodes. The existence of collaborative codes for which the intra-cluster negotiation penalty is in principle small (and almost all the diversity gain of traditional space-time codes may be realized) is shown. For example, for a single transmitter node with two collaborators and one receiver node, if the collaborators have as little as 10 dB path loss advantage over the receiver, the penalty for collaboration over traditional space-time systems is negligible.

Index Terms

Collaborative communications, compound channels, MIMO, relay channels, spatial diversity, wireless sensor networks.

I. Introduction

Due to the current interest in wireless sensor networks [1], recent research in wireless ad hoc networks has focused on the idea of collaborative (or cooperative) communication. In a collaborative communication

Manuscript received June 8, 2004; revised November 1, 2004. This work was presented in part at the 22nd Biennial Symposium on Communications, Kingston, ON, Canada, June 2004 and the Intel Workshop on Antenna Arrays, San Jose, CA, February 2004.

Patrick Mitran and Vahid Tarokh are with the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA. (email: mitran@deas.harvard.edu; vahid@deas.harvard.edu).

Hideki Ochiai is with the Division of Electrical and Computer Engineering, Yokohama National University, Yokohama, Japan 240-8501. (email: hideki@ynu.ac.jp).

framework, communication nodes which would have otherwise remained silent in a traditional point-to-point source/destination paradigm, collaborate with the source and destination to increase communication capacity and reliability.

Arguably, the initial work on collaborative communications stretches as far back as the pioneering papers by van der Meulen [28] and Cover *et al.* [4] on the relay channel. However, the results obtained there do not appear to be directly applicable to inexpensive relays for wireless networks. This is because in realistic wireless models, it is not practically feasible to transmit and receive on the same antenna simultaneously (half-duplex constraint), since the intensity of the near field of the transmitted signal is much higher than that of the far field of the received signal. In the context of wireless communications, a quasi-static fading model is often assumed. This may be reasonable for slow fading channels with moderate frame length. However, except for a special class of degenerate channels, the probability of a decoding error cannot be made arbitrarily small with any finite frame length. It is thus traditional to assume that the fading coefficients remain fixed for the entire duration of the frame, regardless of its length. Furthermore, the channel fading coefficients are usually not known to the transmitting nodes; only the receiving nodes have knowledge of the channel, i.e., realistic wireless channels are also compound channels [7], [32]. Finally, while the degraded relay channel has been completely solved [4], [23], in wireless systems most noise is due to thermal noise in the receiver frontend. While it may be reasonable to assume that the relay has a better signal to noise ratio (SNR) than the ultimate receiver, it is unrealistic to assume that the receiver is a degraded version of the relay. Various extensions of the non-compound relay channel may be found in [9]–[12], [16]–[18], [20], [24], [29], [33]. In particular, [16]–[18] consider a strategy whereby, at any given time, the network may only be in one of a finite set of modes (i.e., the relay is in reception mode or transmission mode).

Some more recent work on collaborative communications with emphasis on treating the wireless channel as a compound channel may be found in [13], [14], [21], [22], [25], [26]. In [22], the authors consider a two stage communications approach (where the source transmits for a fixed amount of time followed by a fixed duration relaying phase) to solve the half-duplex constraint and consider repetition and space-time based cooperative diversity algorithms. This is extended in [21] with the consideration of adaptive protocols such as selection relaying and incremental relaying. In [13], [14] a similar time-division (TD) approach is employed where the relay is permitted to transmit its own information during the second phase if it is unable to collaborate. In [25], [26], the authors assume two dedicated orthogonal subchannels between two mobile users, derive an achievable region for communication to a base station and consider Code Division Multiple Access (CDMA) implementation aspects. These results are derived by employing

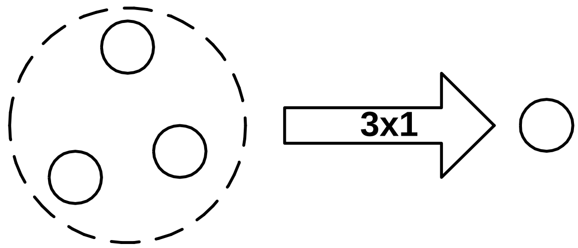

Fig. 1. Is an ideal 3×1 space-time gain achievable with three separate transmit nodes and one receive node?

coding techniques [30], [31] similar to those used for multiple access channels with generalized feedback [3].

In this work, we investigate a bandwidth efficient decode and forward approach that does not employ predetermined phase durations or orthogonal subchannels to resolve the half-duplex constraint: each relay determines based on its own receive channel when to listen and when to transmit. Furthermore, the transmitters are not aware of the channel and we make no assumption of degradedness: the noise at the relays is independent of that at the destination. Also, as opposed to previous collaborative communications literature, we show that our results still hold under a bounded asynchronous model. Finally, in the case of multiple relays assisting the source, our approach permits one relay to assist another in receiving the message, a feature not present in much of the early work on collaborative communications over compound channels. However, more recent work along this line may be found in [2], [15].

Of primary interest in this paper is to determine if we can achieve the genie bound on diversity: the diversity gain that would be achieved if all the transmit antennas of the source and relay nodes were in fact connected to a single node (in [21] this is referred to as the transmit diversity bound). For example, suppose we consider the three transmit collaborators and one receiver node scenario (each equipped with a single antenna) as illustrated in Fig. 1. If all the collaborators were aware of the message *a priori*, we could in principle achieve the ideal performance of a 3×1 space-time system between the transmit cluster and the receiver node. However, only the source node in the transmit cluster is aware of the message *a priori*. The other two nodes in the cluster must serve as relays and are not aware of the message *a priori*. There will be a loss in performance (as measured by the probability of outage) compared to the idealized 3×1 space-time system. In particular, we shall be interested in determining sufficient conditions on the geometry and signal path loss of the transmitting cluster for which performance close to the genie bound can be guaranteed.

To determine an upperbound on this loss, in Section II we will derive a novel approach to the compound relay channel. This approach is best summarized as follows. In a traditional compound channel, we are given a set of possible channel realizations and we seek to prove the existence of a code (with maximal rate) which is simultaneously good on all channel realizations. In this work, we frame the problem in the opposite direction. We fix a rate and ask ourselves how large can we make the set of compound channels while guaranteeing that the code is still good. Section III of this paper summarizes our results. In particular, in Section III-A, we define what we mean by achievable rates for a compound relay channel and state our main result, Theorem 1, with the proof relegated to Appendix I. Section III-B generalizes this result to bounded asynchronous relay channels (Theorem 2). In Section IV, we analyze the performance of this scheme in a quasi-static Rayleigh fading environment. This is accomplished by simulation and, in the case of a two collaborator scenario, the derivation of a tight lower bound on outage probability for our scheme. Our analysis demonstrates that in most cases if the channel path loss of the relay is 10 dB better than that of the ultimate receiver, the genie bound can essentially be obtained[1]. From a geometric viewpoint, this corresponds to transmit clusters whose radius is permitted to be as large as 1/3 the distance between the source and the destination node. Section V concludes this work while Appendix I contains the proofs of Theorems 1 and 2 and Appendix II contains the more technical aspects of the derivation of the tight lower bound.

II. Overview

For simplicity, we consider three nodes denoted as source (s), relay (r) and destination (d) as illustrated in Fig. 2 and each equipped with N_s, N_r and N_d antennas respectively (the results readily generalize to multiple relay nodes).

We assume that while listening to the channel, the relay may not transmit. Hence, the communications protocol we propose is as follows. The source node wishes to transmit one of 2^{nR} messages to the destination employing n channel uses. While not transmitting, the relay node listens. Due to the relay node's proximity to the source, after n_1 samples from the channel (a number which the relay determines on its own and for which the source has no knowledge), it may correctly decode the message. After decoding the message, it then proceeds to transmit for the remaining $n - n_1$ transmissions in an effort to improve the reception of the message at the destination. The destination is assumed to be made aware of n_1 before attempting to decode the message. This may be achieved by an explicit low-rate transmission

[1]It has been shown that a similar result holds for fast (ergodic) fading wireless networks, – if the relay has a certain SNR advantage then there is no loss at all in capacity [19].

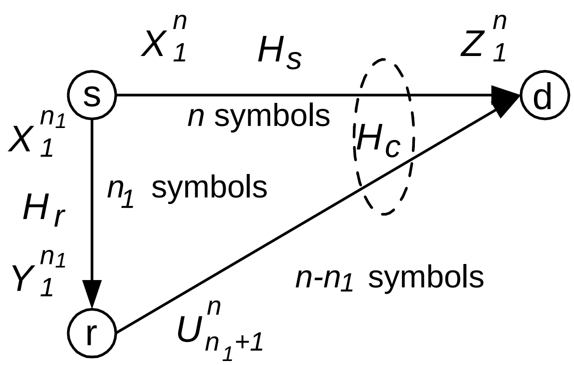

Fig. 2. The collaborative communications problem for two transmit collaborators and one receiver.

from the relay to the destination. Alternatively, if the value of n_1 is constrained to some integer multiple of a fundamental period n_0 (say $n_0 \sim \sqrt{n}$), then the destination may estimate n_1 accurately using power detection methods. We denote the first phase of the n_1 transmissions as the *listening* phase while the last $n - n_1$ transmissions as the *collaboration* phase.

We assume that all channels are modeled as additive white Gaussian noise (AWGN) with quasi-static fading. In particular, X and U are column vectors representing the transmission from the source and relay nodes respectively and we denote by Y and Z the received messages at the relay and destination respectively. Then during the listening phase we have that

$$Z = H_s X + N_Z \tag{1}$$

$$Y = H_r X + N_Y, \tag{2}$$

where the N_Z and N_Y are column vectors of statistically independent complex AWGN with variance 1/2 per row per dimension, H_s is the the fading matrix between the source and destination nodes and likewise, H_r is the fading matrix between the source and relay nodes. During the collaboration phase, we have that

$$Z = H_c[X^T, U^T]^T + N_Z, \tag{3}$$

where H_c is a channel matrix that contains H_s as a submatrix (see Fig. 2).

We further assume that the source has no knowledge of the H_r and H_c matrices (and hence the H_s matrix too). Similarly, the relay has no knowledge of H_c but is assumed to know H_r. Finally, the destination knows H_c.

Without loss of generality, we will assume that all transmit antennas have unit average power during their respective transmission phases. Likewise, the receive antennas have unit power Gaussian noise. If this is not the case, the respective H matrices may be appropriately scaled row-wise and column-wise.

Under the above unit transmit power per transmit antenna and unit noise power per receive antenna constraint, it is well known that a Multiple Input Multiple Output (MIMO) system with Gaussian codebook and with rate R bits/channel use can reliably communicate over any channel with transfer matrix H such that $R < \log_2 \det(I + HH^\dagger) \stackrel{\triangle}{=} C(H)$ [2] [8], [27], where I denotes the identity matrix and $H^\dagger$ is the conjugate transpose of H.

Intuition for the above problem then suggests the following. During the listening phase, the relay knowing H_r listens for an amount of time n_1 such that $nR < n_1 C(H_r)$. During this time, the relay receives at least nR bits of information and may reliably decode the message. The destination, on the other hand, receives information at the rate of $C(H_s)$ bits/channel use during the listening phase and at the rate of $C(H_c)$ bits/channel use during the collaborative phase. It may reliably decode the message provided that $nR < n_1 C(H_s) + (n - n_1) C(H_c)$. In the limit as $n \to \infty$, the ratio n_1/n approaches a fraction f and we may conjecture that there exists a "good" code of rate R for the set of channels (H_r, H_c) which satisfy

$$R \leq f C(H_s) + (1 - f) C(H_c) \tag{4}$$

$$R \leq f C(H_r), \tag{5}$$

for some $f \in [0, 1]$. We note that if the channel between the source and the relay is particularly poor, we may fall back on the traditional point-to-point communications paradigm and add the following region to that given in (4) – (5)

$$R \leq C(H_s). \tag{6}$$

The above intuition is not a proof of achievability but it does provide an upper bound on the performance of the protocol. The essential difficulty in proving that there exists a code which is "good" for any such pair of channels (H_r, H_c) is two-fold. The problem we are dealing with is a *relay channel* which is also a *compound channel*: we seek to prove the existence of a code which performs well over an entire set of channels (unknown to the transmitters). The key will be to show the existence of a code that may essentially be refined. Regardless of the actual value of n_1, there exists a codebook for the source which,

[2]Here, $C(H)$ does not, in general, designate the capacity of each link as is witnessed by the fact that only for a special subset of matrices is capacity achieved by placing an equal transmit power on each antenna.

starting at time n_1+1, may be layered with the transmission of the relay and perform just as well as if the value of n_1 had been known to the source.

III. Definitions and Main Results

In this section, we define the notion of achievability for compound relay channels (with synchronism first and then for bounded asynchronous relay channels later) and present our main results.

A. Synchronous Communication

The codebook for the source will be denoted by $C_s^{(n)}$ and consists of $K2^{nR}$ codewords for some constant $K > 0$. The wth codeword of the source node codebook will be denoted by $x_1^n(w)$. If the source node has N_s transmit antennas, then each codeletter consists of a column vector with dimension N_s and each codeword is in fact an $N_s \times n$ matrix. For the relay, we will denote by $C_r^{(n)}$ a family of n codebooks $C_r^{(n,n_1)}$ indexed by $1 \leq n_1 \leq n$ where $C_r^{(n,n_1)}$ is a codebook with $K2^{nR}$ codewords of length $n-n_1$. The wth codeword of $C_r^{(n,n_1)}$ will be denoted by $u_{n_1+1}^n(w)$. Finally, we will denote by $C^{(n)} = \{C_s^{(n)}, C_r^{(n)}\}$.

Before explaining the encoding procedure, it will help to explain the decoder. With each message $W = w$, pair of channels (H_r, H_c) and value of n_1, we associate some disjoint (over w) subsets of $\mathbb{C}^N$ as follows: $S_{w,H_c,n_1} \subset \mathbb{C}^{N_d \times n}$ and $R_{w,H_r,n_1} \subset \mathbb{C}^{N_r \times n_1}$. We shall refer to $C^{(n)}$ as the encoder or codebook and the sets S_{w,H_c,n_1} and R_{w,H_r,n_1} as the decoder.

Encoding and decoding procedure: The source wishes to transmit message $W = w$ to the destination. To that end, the source looks up the wth codeword in its codebook and proceeds to transmit it to the destination and the relay. The relay, knowing the channel H_r, decides upon the smallest value of n_1 for which $nR+\delta < n_1 C(H_r)$ (for some fixed $\delta > 0$) and for which $\delta < n_1/n < 1-\delta$. If no such n_1 exists, the relay takes $n_1 = n$, makes no attempt to decode the message and remains silent. If $n_1 < n$, the relay listens to the channel for this duration and lists all the $\hat{w}$ for which $Y_1^{n_1} \in R_{\hat{w},H_r,n_1}$. If $\hat{w}$ exists (and is hence unique), the relay looks up the $\hat{w}$th codeword in the $C_r^{(n,n_1)}$ codebook and proceeds to transmit it for the remaining $n-n_1$ channel uses. Otherwise, the relay declares an error.

After the last transmission, the destination has now received Z_1^n where

$$Z_i = \begin{cases} H_s X_i + N_{Z,i} & i \leq n_1 \\ H_c[X_i^T, U_i^T]^T + N_{Z,i} & i > n_1, \end{cases} \tag{7}$$

and is informed of the value of n_1. The destination then proceeds to list all $\hat{w}$ such that $Z_1^n \in S_{\hat{w},H_c,n_1}$. If $\hat{w}$ exists (and is hence unique), the destination declares the transmitted message as $\hat{W} = \hat{w}$. Otherwise an error is declared. We shall abuse notation and denote by the event $\hat{W} \neq W$ the case where either the relay or the destination declares an error or decodes an incorrect message (if the relay makes no attempt at decoding the message, it cannot produce an error).

Since the source, relay and destination nodes each have N_s, N_r and N_d antennas respectively, we note that $H_r \in \mathbb{C}^{N_r \times N_s}$ and $H_c \in \mathbb{C}^{N_d \times (N_s+N_r)}$. We denote by $\mathcal{H}$ a subset of compound relay channels, i.e., $\mathcal{H} \subset \mathbb{C}^{N_r \times N_s} \times \mathbb{C}^{N_d \times (N_s+N_r)}$. Also, for a codebook $C^{(n)}$, we denote by λ_n^s (where the superscript s denotes synchronism)

$$\lambda_n^s = \max_w \sup_{(H_r,H_c)\in\mathcal{H}} P[\hat{W} \neq W | W = w, H_r, H_c]. \tag{8}$$

Definition 1: (Achievability for a compound relay channel) A rate R is said to be achievable for a set of pairs $(H_r, H_c) \in \mathcal{H}$ if for any $\epsilon > 0$, there exists a sequence of encoders and decoders $C^{(n)}$, S_{w,H_c,n_1} and R_{w,H_r,n_1} in n such that $\lambda_n^s \to 0$ as $n \to \infty$ and each codeword in each sub-codebook of $C^{(n)}$ has average power at most $1 + \epsilon$. □

Before stating a theorem on the existence of good codes, we introduce a norm on the space of complex matrices. Let H be a complex matrix with entries $H_{i,j}$. Then, we define $||H|| \stackrel{\triangle}{=} \max_{i,j}\{|H_{i,j}|\}$.

Theorem 1: (Synchronous collaborative communications) Consider the set $\mathcal{H}_{\delta,L}(R)$ of matrices (H_r, H_c) such that $||H_r|| \leq L$ and $||H_c|| \leq L$ and which satisfy either both of the following

$$R + \delta \leq fC(H_s) + (1 - f)C(H_c) \tag{9}$$

and

$$R + \delta \leq fC(H_r), \tag{10}$$

or

$$R + \delta \leq C(H_s), \tag{11}$$

for some $\delta \leq f \leq 1 - \delta$ (each f may depend on H_r). Then, the rate R is achievable for the compound relay channel $\mathcal{H}_{\delta,L}(R)$ for any $\delta > 0$ and $L > 0$. □

The above theorem essentially states that the region in equations (4) – (6) may be arbitrarily approximated by taking $\delta > 0$ sufficiently small and $L > 0$ sufficiently large.

B. Asynchronous Collaborative Communications

In this subsection, we consider the case of bounded asynchronism similar to that in [5] between the various nodes. In particular, we shall consider the asynchronous case where the relay and destination each have (but are unaware of) an *integer* delay of d_r and d_d symbols with respect to the source where $0 \leq d_r \leq D$ and $0 \leq d_d \leq D$ and D is a known upperbound to the delay. The delay between the relay and the destination is absorbed into d_r and d_d. For simplicity, the model we deal with reflects asynchronism due to timing offsets between various nodes of the sensor network. More general asynchronous models that include propagation delay may be dealt with similarly (with results identical to that of Theorem 2).

For example, if the relay has a delay d_r, then the sequence of observations $Y_1^{n_1}$ depends on $X_{1+d_r}^{n_1+d_r}$ and on no other terms in the sequence X_1^n. We denote this relationship by the expression

$$Y_1^{n_1} \sim X_{1+d_r}^{n_1+d_r}. \tag{12}$$

Likewise, we have that

$$Z_1^{n_1-D} \sim X_{1+d_d}^{n_1-D+d_d} \tag{13}$$

$$Z_{n_1+1+D}^{n-D} \sim (X_{n_1+1+D+d_d}^{n-D+d_d}, U_{n_1+1+D+d_d-d_r}^{n-D+d_d-d_r}). \tag{14}$$

In the case of asynchronism, the encoder $\mathcal{C}^{(n)}$ will remain identical to the one in the synchronous case. However, the decoder will now be described by sets of the form R_{w,H_r,n_1,d_r} and S_{w,H_c,n_1,d_r,d_d} such that for any $w_1 \neq w_2$, H_c, H_r, n_1, d_r, d_r', d_d and d_d',

$$R_{w_1,H_r,n_1,d_r} \cap R_{w_2,H_r,n_1,d_r'} = \varnothing \tag{15}$$

$$S_{w_1,H_c,n_1,d_r,d_d} \cap S_{w_2,H_c,n_1,d_r',d_d'} = \varnothing. \tag{16}$$

We do not require that either $R_{w_1,H_r,n_1,d_r} \cap R_{w_1,H_r,n_1,d_r'} = \varnothing$ or $S_{w_1,H_c,n_1,d_r,d_d} \cap S_{w_1,H_c,n_1,d_r',d_d'} = \varnothing$.

Encoding and decoding procedure: The source wishes to transmit message $W = w$ to the destination. To that end, the source looks up the wth codeword in its codebook and proceeds to transmit it to the destination and relay. The relay, knowing H_r (but not d_r), decides upon the smallest value of n_1 for which $nR + \delta < n_1 C(H_r)$ (for some fixed $\delta > 0$) and for which $\delta < n_1/n < 1 - \delta$. If no such n_1 exists, the relay takes $n_1 = n$, makes no attempt to decode the message and remains silent. If $n_1 < n$, the relay listens to the channel for this duration and lists all the $\hat{w}$ for which there exists a $\hat{d}_r$ such that $Y_1^{n_1} \in R_{\hat{w},H_r,n_1,\hat{d}_r}$. If $\hat{w}$ exists (and is hence unique), the relay looks up the $\hat{w}$th codeword in the $C_r^{(n,n_1)}$ codebook and proceeds to transmit it for $n - n_1$ channel uses. Otherwise, the relay declares an error.

After the last transmission, the destination has now received Z_1^n (and is informed of n_1) where

$$Z_i = \begin{cases} H_s X_{i+d_d} + N_{Z,i} & i \leq n_1 + d_r \\ H_c [X_{i+d_d}^T, U_{i+d_d-d_r}^T]^T + N_{Z,i} & i > n_1 + d_r, \end{cases} \tag{17}$$

and we have employed the convention that $X_j = U_j = 0$ for $j > n$ and $U_j = 0$ for $j \leq n_1$. (If symbol synchronism is not achieved then the channel between the source and the destination is frequency selective, however this is outside the scope of the current study.) The destination then proceeds to list all $\hat{w}$ such that for some $\hat{d}_r$ and $\hat{d}_d$, $Z_1^n \in S_{\hat{w},H_c,n_1,\hat{d}_r,\hat{d}_d}$. If $\hat{w}$ exists (and is hence unique), the destination declares the transmitted message as $\hat{W} = \hat{w}$. Otherwise an error is declared.

In the case of symbol asynchronism, we have the following performance measure of a code,

$$\lambda_n^a = \max_{w} \max_{0 \leq d_r, d_d \leq D} \sup_{(H_r,H_c)\in\mathcal{H}} P[\hat{W} \neq W | W = w, H_r, H_c, d_r, d_d]. \tag{18}$$

Definition 2: (Achievability for a compound channel with bounded asynchronism) A rate R is said to be achievable for a set of pairs $(H_r, H_c) \in \mathcal{H}$ and bounded asynchronism $0 \leq d_r, d_d \leq D$ if for any $\epsilon > 0$, there exists a sequence of encoders and decoders $C^{(n)}$, S_{w_1,H_r,n_1,d_r,d_d}, R_{w_1,H_c,n_1,d_r} in n such that $\lambda_n^a \to 0$ as $n \to \infty$ and each codeword in each sub-codebook of $C^{(n)}$ has average power at most $1 + \epsilon$. □

Our main result is that the performance of Theorem 1 is still achievable in the asynchronous case.

Theorem 2: (Asynchronous collaborative communications) Consider the set $\mathcal{H}_{\delta,L}(R)$ of matrices (H_r, H_c) such that $||H_r|| \leq L$ and $||H_c|| \leq L$ and which satisfy either both of the following

$$R + \delta \leq fC(H_s) + (1 - f)C(H_c) \tag{19}$$

and

$$R + \delta \leq fC(H_r), \tag{20}$$

or

$$R + \delta \leq C(H_s), \tag{21}$$

for some $\delta \leq f \leq 1 - \delta$ (each f may depend on H_r). Then, the rate R is achievable for $\mathcal{H}_{\delta,L}(R)$ with bounded symbol asynchronism for any $\delta > 0$ and $L > 0$. □

IV. Performance Analysis

A. Two Collaborators

In this section we evaluate numerically the theoretical performance of a code that achieves the compound channels in Theorems 1 and 2 for single antenna nodes when the fading is quasi-static Rayleigh distributed. In particular, since L was an arbitrarily large constant, for the purposes of this section, we shall take $L = \infty$. Similarly, since δ was an arbitrarily small positive number, we take $\delta = 0$.

Furthermore, we will relax our restrictions on unit power per transmit antenna (as stated earlier, this was allowable since the respective H matrices could be appropriately scaled to compensate). In this section, it will be more convenient to keep the H matrices fixed and show the explicit dependence of the outage probability on the receive signal power at the destination node (during the listening phase) per transmit antenna at the source node, E_s, and the noise power at each receive antenna, σ^2. Under these conditions, we have that

$$C(H, \gamma) \stackrel{\triangle}{=} \log_2 \det(I + \gamma H H^\dagger), \tag{22}$$

where $\gamma \stackrel{\triangle}{=} \frac{E_s}{\sigma^2}$ and the expression holds regardless of the number of transmit antennas (as E_s is defined as the normalized receive power per transmit antenna).

Furthermore, in a practical situation, the node that the source collaborates with will typically be near the source node. We model this proximity by a reduction $G \in \mathbb{R}$ in path loss, or equivalently, an increase in the achievable rate between collaborator nodes as expressed by $C(H, G\gamma)$.

With these conventions, we will assume that the code successfully transmits the message from the source to the destination in a two collaborator scenario provided that either

$$R \leq fC(H_r, G\gamma) \tag{23}$$

and

$$R \leq fC(H_s, \gamma) + (1 - f)C(H_c, \gamma), \tag{24}$$

for some $0 < f < 1$, or

$$R \leq C(H_s, \gamma), \tag{25}$$

holds.

We note that the fraction f is determined by the relay and depends only on the realization of H_r according to

$$\hat{f} \stackrel{\triangle}{=} \min\{1, R/C(H_r, G\gamma)\}. \tag{26}$$

Since $C(H_c, \gamma) \geq C(H_s, \gamma)$, this is the optimal choice of f to minimize the outage probability of our scheme. Even if the relay knew H_c, it could not do better.

Furthermore, given f, the effective receive power at the destination is $(2 - f)E_s$ as the relay was only transmitting for a fraction $1 - f$ of the total transmission time. The effective receive SNR for the duration of the transmission is then $(2-f)E_s/\sigma^2$. We may thus rewrite the set of outage events *for our proposed scheme*, E_o, compactly as

$$\mathsf{E}_o = \left\{ R > \hat{f}C(H_s,\gamma) + (1-\hat{f})C(H_c,\gamma) \right\}. \tag{27}$$

Let $v \in [0,1]$ denote the *maximum* fraction of time that the relay transmits for a given realization of H_r. Then, we have

$$v = 1 - \hat{f} \tag{28}$$

$$= \max\left(0, 1 - \frac{R}{C(H_r, G\gamma)}\right). \tag{29}$$

Substituting this into (27), we may write the outage probability for our proposed scheme as

$$P_{\text{out}} = P\left[R > (1-v)\, C(H_s,\gamma) + vC(H_c,\gamma)\right]. \tag{30}$$

In the case of a two transmit collaborator system with single antenna nodes, H_c is a 1×2 matrix with $H_c = [H_s, H_{r,d}]$ where $H_{r,d}$ denotes the complex channel coefficient of the link between the relay and destination nodes. The corresponding mutual informations are given by

$$C(H_s,\gamma) = \log_2\left(1 + \gamma|H_s|^2\right) \tag{31}$$

$$C(H_c,\gamma) = \log_2\left(1 + \gamma\left(|H_s|^2 + |H_{r,d}|^2\right)\right). \tag{32}$$

Note the assumption that the relay transmits the same instantaneous energy as the source node and suffers the same amount of path loss (if this were not the case, the difference could be absorbed into the G coefficient). Therefore, the transmit energy per symbol *for a given realization of the relay channel* is expressed as $(1+v)E_s$ and the corresponding averaged receive SNR is given by $(1+E[v])E_s/\sigma^2$. Furthermore, $E[v]$ can be obtained analytically in many cases of interest, such as quasi-static Rayleigh fading (see Appendix II).

Substituting (31) and (32) into (30), we obtain

$$P_{\text{out}} = P[R > (1-v)\log_2\left(1+\gamma|H_s|^2\right) + v\log_2\left(1+\gamma\left(|H_s|^2 + |H_{r,d}|^2\right)\right)]. \tag{33}$$

Evaluation of the above probability is difficult to carry out even numerically, since exact analysis typically yields double integrals. Alternatively, in the following we derive a tight lower bound for the above probability and for which numerical evaluation with high accuracy is easy.

By Jensen's inequality, we have for $0 \leq v \leq 1$

$$(1-v)\log_2\left(1+\gamma|H_s|^2\right)+v\log_2\left(1+\gamma\left(|H_s|^2+|H_{r,d}|^2\right)\right)$$
$$\leq \log_2\left(1+\gamma\left(|H_s|^2+v|H_{r,d}|^2\right)\right), \tag{34}$$

where the equality holds iff $v=0$ or 1. Thus, the bound may be expected to become tight in the high SNR region ($v \approx 1$) as well as the low SNR region ($v \approx 0$).

Since (33) can be seen as a complementary cumulative distribution function (CCDF) with respect to R and the CCDF is a monotonically decreasing function, we have

$$P_{\text{out}} \geq P\left[R > \log_2\left(1+\gamma\left(|H_s|^2+v|H_{r,d}|^2\right)\right)\right] \tag{35}$$
$$= P\left[|H_s|^2+v|H_{r,d}|^2 < \frac{2^R-1}{\gamma}\right]. \tag{36}$$

The right hand side of (36) is the CDF of the random variable $\omega \triangleq |H_s|^2+v|H_{r,d}|^2$. Since ω is a combination of products and sums of two statistically independent random variables, its CDF can be found by standard algebra of random variables. The explicit expression of (36) is derived in Appendix II for the case that all fading channel coefficients $H_r, H_s, H_{r,d}$ are i.i.d. complex Gaussian random variables.

Fig. 3 illustrates the simulated outage probability (for $R = 0.5$ and $R = 2$) and the corresponding lower bound (36) for this idealized code for various values of G versus the averaged receive SNR for quasi-static Rayleigh fading channels, i.e., channels where each of the H matrices have independent circularly symmetric Gaussian distributed entries with total variance 1. Exact results were obtained by Monte-Carlo simulation of equations (26) and (27). All simulated outage probabilities were calculated by simulating enough channel realizations to generate at least 10^4 outage events. These simulation results are confirmed by the tightness of the lower bound (36).

Also illustrated in Fig. 3 is the outage probability of an idealized traditional 1×1 and 2×1 space-time system. From the figure, we see that even with as little gain as $G = -5$ dB, for an outage probability of 0.01, the loss in performance (difference in receive SNR for identical outage probabilities) is only 3.5 dB compared to the genie 2×1 bound. With $G = 10$ dB, the genie 2×1 bound is closely approached by our collaborative scheme. If we assume a channel path decay exponent of 2, this corresponds to a cluster size at the transmitter side with a radius 1/3 the distance between the source and the destination.

Finally, it is instructive to compare the performance of this scheme (where the relay listens for the smallest fraction of time f that is necessary) to a scheme where f is not allowed such flexibility. In Fig. 3, the performance of a scheme where f is constrained to 0.5 or 1.0 is also illustrated. (We use the notation TD for this scheme. Whereas $f = 0.5$ corresponds to a half listening/half collaboration protocol, $f = 1.0$ is equivalent to no collaboration.)

For a geometric gain of $G = 10$ dB and a rate $R = 0.5$, we see that this TD scheme performs as well as the proposed scheme with $G = 5$ dB at an outage probability of 0.01. Hence, in this case, the penalty for employing a predetermined TD scheme is equivalent to a 5 dB penalty in geometric gain. In terms of node geometry, a 5 dB penalty is equivalent to a cluster whose radius is at most 0.178 times the distance between the source and the destination node. For higher rates such as $R = 2$, the penalty increases.

B. Three Transmit Collaborators

The results in Theorems 1 and 2 generalize in a straight forward manner to multiple transmit collaborators. In particular, we now consider three transmit collaborators and one receiver, all equipped with a single antenna (as illustrated in Fig. 4). There, the channel between two nodes, say nodes s and r_1, will be denoted by H_{s,r_1}. Likewise, the channel between the pair of nodes (s, r_0) and the node r_1 will be denoted H_{sr_0,r_1} with H_{s,r_1} a submatrix of H_{sr_0,r_1}. Let us suppose without loss of generality that $C(H_{s,r_0}, G\gamma) \geq C(H_{s,r_1}, G\gamma)$. If we denote by f_0 the fraction of time that relay r_0 listens and by f_1 the fraction of *additional* time that r_1 listens beyond f_0 (i.e., r_1 listens for the total fraction of time $f_0 + f_1$), then communication is successful provided that either

$$R \leq f_0 C(H_{s,r_0}, G\gamma) \tag{37}$$

$$R \leq f_0 C(H_{s,r_1}, G\gamma) + f_1 C(H_{sr_0,r_1}, G\gamma) \tag{38}$$

$$R \leq f_0 C(H_{s,d}, \gamma) + f_1 C(H_{sr_0,d}, \gamma), \tag{39}$$

$$+ (1 - f_0 - f_1) C(H_{sr_0r_1,d}, \gamma)$$

for some $f_0 > 0$, $f_1 > 0$ with $f_0 + f_1 < 1$ or

$$R \leq f_0 C(H_{s,r_0}, G\gamma) \tag{40}$$

$$R \leq f_0 C(H_{s,d}, \gamma) + (1 - f_0) C(H_{sr_0,d}, \gamma), \tag{41}$$

for some $0 < f_0 < 1$ or

$$R \leq C(H_{s,d}, \gamma). \tag{42}$$

If in fact $C(H_{s,r_0}, G\gamma) < C(H_{s,r_1}, G\gamma)$, then we must add two more regions to the achievable compound channel (which may be obtained by symmetrically interchanging r_0 and r_1 in the above regions). Furthermore, f_0 is greedily chosen by node r_0 to be the smallest value which satisfies (37).

Likewise, f_1 is then chosen as the smallest value which satisfies (38). Similar to the two collaborator scenario, if a solution to (37) – (42) exists, this greedy approach will also find a solution.

We note a feature prevalent in this approach that is not present in much of the early collaborative communications literature for compound channels. If the relay r_0 has a better channel from the source than the relay r_1, relay r_1 may receive information not only from the source node s, but from the relay node r_0 as soon as r_0 has finished listening (see Fig. 4 and (38)). By symmetry, a similar situation is possible if relay r_1 has a better channel than r_0. If the number of collaborative nodes were further increased to m, we would see a cascade effect by which the relays would quickly share among themselves the message by way of $m(m-1)/2$ possible paths (more recent literature that allows for this sort of information sharing strategy may be found in [2], [15]).

Fig. 5 shows the simulated outage probability of this scheme for rates $R = 0.5$ and $R = 2$. Again, in these two cases, a gain of 10 dB in path loss results in performance that closely approaches the idealized 3×1 bound. Even with a gain of 0 dB, we achieve the same performance as the idealized 2×1 system at an outage probability of 0.01. Furthermore, a comparison to the TD approach (f_1 and f_2 are constrained to 0.0, 0.5 or 1.0) shows that for a geometric gain of $G = 10$ dB, the TD approach again demonstrates a 5 dB geometric penalty compared to the proposed scheme at a rate $R = 0.5$ and the penalty is larger for rate $R = 2$.

The result is readily extended to three single antenna transmit collaborators and one receiver with two antennas by "adding a second row" to the H matrices. This could correspond to a network of three nodes communicating to a base station. Fig. 6 shows the outage probability of our proposed scheme in this case. Here, we observe that a gain G of 10 dB performs within 1.5 dB of the genie 3×2 system for a rate $R = 0.5$ while a similar result holds with a gain of 15 dB for a rate $R = 2$. Comparisons to the equivalent TD scheme show that for $R = 0.5$, a geometric loss of about 2.5 dB is observed for $G = 10$ dB. For $R = 2$, a geometric loss of about 5 dB is observed at both $G = 10$ and $G = 15$ dB.

Finally, in Fig. 7, the outage probability versus receiver SNR is illustrated for three single antenna transmit collaborators, and a two antenna receiver with antenna selection at the receiver side, i.e., provided one of the receiving antennas allows for correct decoding based only on its received signal, communication is deemed successful. In traditional MIMO systems, receiver side antenna selection is a low complexity approach to obtain good diversity. Here we see that for the rate $R = 0.5$ and $R = 2$ cases, performance is within 1 dB of the genie 0.01 outage probability of error (with antenna selection) with path loss gains of 10 dB and 15 dB respectively. An equivalent TD approach exhibits similar performance losses to that of the three single antenna collaborator and receiver with two antennas.

V. Conclusion

We have considered a novel approach to compound Gaussian relay channels and shown the existence of a collaborative code which is good over a wide range of relay channels. Our approach does not employ predetermined phase durations or orthogonal subchannels to resolve the half-duplex constraint. Each relay, based on knowledge of its receive channel determines on its own when to listen and when to transmit. In addition to bandwidth efficiency, this approach has the advantage that each relay can receive information not only from the source, but via other relays. We have also shown that the method may be applied to a bounded asynchronous scenario where each relay and destination node has an unknown integer delay with respect to the source.

Numerical and simulation results have shown that if the intra-cluster communication has a 10 dB path loss advantage over the receiver at the destination node, in most cases there is essentially no penalty for the intra-cluster communication. Physically, in a two collaborator scenario, this corresponds to a transmit cluster whose radius is 1/3 the distance between the source and destination nodes. By comparison, for a TD scheme with a 5 dB geometric penalty, the allowable cluster size is at most 0.178 times the distance between the source and the destination.

One possible extension of this work is to a fully asynchronous model where the delays between the various nodes need not be constrained to integer multiples of the symbol period. Another extension is to investigate more refined collaboration on the part of the relays. In particular, in this work the relays were constrained to transmit only after having decoded the complete message. Since the relays know their receive channels, in some cases a relay may be aware that this is impossible given the time constraint n. In such a case, it may be beneficial for a relay to start transmitting based on a partial decoding of the message. This requires an interesting tradeoff. On the one hand, the more the relay listens, the more efficiently it may collaborate. On the other hand, the more it listens, the less time it has to collaborate. In terms of minimizing outage probabilities, the optimal tradeoff will depend heavily on the distribution of the quasi-static fading parameters.

Appendix I
Proof of Theorems

A. Proof of Theorem 1

Before proving Theorem 1, we require a simple lemma relating to typical sequences. Recall that the set of ϵ-typical sequences x_1^n, which we denote by $A_\epsilon^n(X)$, consists of all n length sequences such that $|\frac{1}{n}\log p(x_1^n) - h(X)| \leq \epsilon$. We also recall that if H is a complex matrix then $||H|| \stackrel{\triangle}{=} \max_{i,j} |H_{i,j}|$.

Lemma 1: Let $Y_1^n = HX_1^n + N_1^n$ where X_1^n and N_1^n are $M \times n$ matrices with i.i.d. complex Gaussian entries with mean 0 and variance 1/2 per dimension per row and H is any $N \times M$ complex matrix such that $||H|| \leq L$ for some $L \geq 0$. Then, for each integer $l \geq 0$, there exist constants K_l^X, K_l^Y and $K_l^{X,Y}$ (which depend on L) such that

$$P[X_1^n \notin A_\epsilon^n(X)] \leq \frac{K_l^X}{n^l} \tag{43}$$

$$P[Y_1^n \notin A_\epsilon^n(Y|H)|H] \leq \frac{K_l^Y}{n^l} \tag{44}$$

$$P[(X_1^n, Y_1^n) \notin A_\epsilon^n(X,Y|H)|H] \leq \frac{K_l^{X,Y}}{n^l} \tag{45}$$

for all such $||H|| \leq L$.

Proof: We only prove the result for K_l^Y in the case $N = 1$. The other cases are similar. Now, by a variant of Chebychev's inequality,

$$\begin{aligned} &P[Y_1^n \notin A_\epsilon^n|H] \\ &\quad = P\left[\frac{1}{n}\left|\sum_{i=1}^n \log p(Y_i|H) - h(Y_i|H)\right| > \epsilon\right] &(46)\\ &\quad \leq \frac{E[\sum_{i=1}^n \log p(Y_i|H) - h(Y_i|H)]^{2l}}{n^{2l}\epsilon^{2l}} &(47)\\ &\quad \leq \frac{2^l(2l)!n^l K(H)}{n^{2l}\epsilon^{2l}}, &(48)\end{aligned}$$

where $K(H)$ is continuous function of the first $2l$ moments of Y. Since Y is Gaussian, these moments exist for all H and are continuous functions of H. Since the set of allowable matrices $||H|| \leq L$ is compact, $K(H)$ has a maximal value, say K' and the lemma is satisfied with $K_l^Y = 2^l(2l)!K'/\epsilon^{2l}$. ■

Proof of Theorem 1: We will prove the result for $N_s = N_r = N_d = 1$, though the proof may easily be generalized if this is not the case. Without loss of generality, we assume that $\delta = 5\epsilon$. If δ was greater than 5ϵ then reducing δ such that $\delta = 5\epsilon$ merely increases the set $\mathcal{H}_{\delta,L}(R)$ of compound channels. On the other hand, if $\delta < 5\epsilon$, decreasing ϵ results in stricter power limitations which would still satisfy the original prescribed limitations.

The remainder of the proof is broken into four parts which are summarized as follows: 1) The relay strategy for the choice of n_1, 2) A sequence $\mathcal{H}_{\delta,L}^n(R)$ of "asymptotically" dense subsets in $\mathcal{H}_{\delta,L}(R)$, 3) Coding arguments for the compound relay channel $\mathcal{H}_{\delta,L}^n(R)$ and 4) Extension to $\mathcal{H}_{\delta,L}(R)$.

Part 1) The relay strategy for the choice of n_1.

Since either

$$R + 5\epsilon \leq fC(H_s) + (1-f)C(H_c) \tag{49}$$

$$R + 5\epsilon \leq fC(H_r), \tag{50}$$

for some $\delta \leq f \leq 1-\delta$ or

$$R + 5\epsilon \leq C(H_s), \tag{51}$$

and $||H_c|| \leq L$ and $||H_r|| \leq L$, it is possible to find an integer N^* such that for any $n > N^*$, there exists an integer $n_1(n)$ such that either

$$R + 4\epsilon \leq fC(H_s) + (1-f)C(H_c) \tag{52}$$

$$R + 4\epsilon \leq fC(H_r), \tag{53}$$

or

$$R + 4\epsilon \leq C(H_s), \tag{54}$$

with $\delta \leq f = n_1(n)/n \leq 1-\delta$. From now on, we assume that $n > N^*$. Note, the relay chooses the smallest such n_1 which satisfies (53). By virtue of the fact that $C(H_c) \geq C(H_s)$ (since H_s is a submatrix of H_c), if there exists a value of n_1 that satisfies both (52) and (53), then this greedy approach by the relay will find a value of n_1 that satisfies these constraints. If no such value of n_1 exists, the relay chooses $n_1 = n$ (i.e., the relay chooses to remain silent during the entire communication). Either way, we may thus denote this choice of n_1 as $n_1(n, H_r)$.

Part 2) Description of a sequence $\mathcal{H}^n_{\delta,L}(R)$ of "asymptotically" dense subsets in $\mathcal{H}_{\delta,L}(R)$.

The next two parts show the existence of a code that does well on an "asymptotically" dense subset of $\mathcal{H}_{\delta,L}(R)$. Let $\mathcal{H}^c_{\delta,L}(R)$ be the subset of $\mathcal{H}_{\delta,L}(R)$ which satisfies equations (52) and (53). Likewise, let $\mathcal{H}^s_{\delta,L}(R) = \mathcal{H}_{\delta,L}(R) \backslash \mathcal{H}^c_{\delta,L}(R)$, i.e., the subset of $\mathcal{H}_{\delta,L}(R)$ which satisfies (54) but not (52) and (53).

Let $\mathcal{H}^{n,c}_{\delta,L}(R)$ be a finite subset of $\mathcal{H}^c_{\delta,L}(R)$ such that for each pair $(H_r, H_c) \in \mathcal{H}^c_{\delta,L}(R)$, there is a pair $(H'_r, H'_c) \in \mathcal{H}^{n,c}_{\delta,L}(R)$ such that

1) $n_1(n, H_r) = n_1(n, H'_r)$
2) $||H_r - H'_r|| \leq 1/n^2$
3) $||H_c - H'_c|| \leq 1/n^2$.

Due to the boundedness of $\mathcal{H}^c_{\delta,L}(R)$, this may be accomplished using a polynomial in n sized set of pairs (H'_r, H'_c). In particular, since $\mathcal{H}_{\delta,L}(R)$ is a subset of a six dimensional real Euclidean space, the last two conditions can be satisfied using $O(n^{12})$ pairs. Then, all three conditions may be satisfied with $O(n^{13})$ pairs.

The last statement is best argued as follows. Let us denote a ball of relay channels of radius $1/n^2$ about the pair (H_r', H_c') by $B((H_r', H_c'), 1/n^2)$ with norm $||(H_r, H_c)|| = \max\{||H_r||, ||H_c||\}$. From the set $\mathcal{H}_{\delta,L}^{n,c}(R)$ let us incrementally construct a new set $\hat{\mathcal{H}}_{\delta,L}^{n,c}(R)$ starting with $\hat{\mathcal{H}}_{\delta,L}^{n,c}(R) = \varnothing$. For each $(H_r', H_c') \in \mathcal{H}_{\delta,L}^{n,c}(R)$, consider all pairs $(H_r'', H_c'') \in B((H_r', H_c'), 1/n^2)$. For all such pairs, $n_1(n, H_r'')$ takes on at most $n+1$ values. Hence, for each $(H_r', H_c') \in \mathcal{H}_{\delta,L}^{n,c}(R)$, add to $\hat{\mathcal{H}}_{\delta,L}^{n,c}(R)$ as many pairs $(H_r'', H_c'') \in B((H_r', H_c'), 1/n^2)$ as possible, in such a way that no two pairs added from $B((H_r', H_c'), 1/n^2)$ have the same $n_1(n, H_r'')$. Then $\hat{\mathcal{H}}_{\delta,L}^{n,c}(R)$ is a set of cardinality $O(n^{13})$ which by the triangle inequality satisfies conditions 1), 2) and 3) with $1/n^2$ replaced by $2/n^2$. If the initial balls of $\mathcal{H}_{\delta,L}^{n,c}(R)$ had radius $1/2n^2$, then conditions 1), 2) and 3) would be satisfied with a set of cardinality $O(n^{13})$.

Likewise, we denote by $\mathcal{H}_{\delta,L}^{n,s}(R)$ a similar approximation to $\mathcal{H}_{\delta,L}^{s}(R)$ and we let $\mathcal{H}_{\delta,L}^{n}(R) = \mathcal{H}_{\delta,L}^{n,c}(R) \cup \mathcal{H}_{\delta,L}^{n,s}(R)$.

Part 3) Coding arguments for the compound relay channel $\mathcal{H}_{\delta,L}^{n}(R)$.

Now, we consider employing a variant of the random coding argument on the compound channel $\mathcal{H}_{\delta,L}^{n}(R)$. In particular, we choose each codeword in each sub-codebook of $C^{(n)} = \{C_s^{(n)}, C_r^{(n)}\}$ randomly according to an i.i.d. generated complex Gaussian with mean 0 and variance 1/2 per dimension. Define the following two decoding sets:

$$\begin{aligned} R_{w,H_r,n_1} = \{y_1^{n_1} :& \\ &(x_1^{n_1}(w), y_1^{n_1}) \in A_\epsilon^{n_1}(X,Y|H_r), \\ &\forall w' \neq w, (x_1^{n_1}(w'), y_1^{n_1}) \notin A_\epsilon^{n_1}(X,Y|H_r)\}. \end{aligned} \tag{55}$$

$$\begin{aligned} S_{w,H_c,n_1} = \{z_1^{n} :& \\ &(x_1^{n_1}(w), z_1^{n_1}) \in A_\epsilon^{n_1}(X,Z|H_c), \\ &(x_{n_1+1}^{n}(w), u_{n_1+1}^{n}(w), z_{n_1+1}^{n}) \in A_\epsilon^{n-n_1}(X,U,Z|H_c), \\ &\forall w' \neq w, \\ &(x_1^{n_1}(w'), z_1^{n_1}) \notin A_\epsilon^{n_1}(X,Y|H_c) \\ &(x_{n_1+1}^{n}(w'), u_{n_1+1}^{n}(w'), z_{n_1+1}^{n}) \notin A_\epsilon^{n-n_1}(X,U,Z|H_c)\}. \end{aligned} \tag{56}$$

With this choice of decoding sets (i.e., the decoding sets are the typical decoding sets less any intersections with another typical decoding set), we *mimic* the random coding argument. Without any loss, we may thus assume that the relay looks for codewords which are jointly typical with the observed sequence and if there is one and only one, decides that this codeword was the one that was transmitted. Likewise, we may equivalently assume that the destination node looks for the codewords which are jointly typical with the received sequence during *both* phases of communication.

Let $N = K2^{nR}$. If we show that

$$P_e = E_{C^{(n)}}\Big[\frac{1}{N}\sum_{w}\Big\{ \tag{57}$$

$$\sum_{(H_r,H_c)\in\mathcal{H}^n_{\delta,L}(R)} P[\hat{W}\neq W|W=w,H_r,H_c]\Big\}\Big]$$

$$= \sum_{(H_r,H_c)\in\mathcal{H}^n_{\delta,L}(R)} E_{C^{(n)}} P[\hat{W}\neq W|H_r,H_c] \tag{58}$$

$$< \lambda, \tag{59}$$

where the expectation is over all randomly chosen codebooks, then we know that there is one code for which the average probability of error (over all codewords) is no more than λ. Using standard arguments, by purging the worst half of the messages from the codebook, we can produce a new code (with K one half that of the previous code) for which

$$\max_{w}\left\{\sum_{(H_r,H_c)\in\mathcal{H}^n_{\delta,L}(R)} P[\hat{W}\neq W|W=w,H_r,H_c]\right\} < \lambda. \tag{60}$$

However, this implies that,

$$\max_{w}\max_{(H_r,H_c)\in\mathcal{H}^n_{\delta,L}(R)} P[\hat{W}\neq W|W=w,H_r,H_c]\} < \lambda. \tag{61}$$

Without loss of generality, we assume that $W=1$ was transmitted. The expression in (58) may then be upperbounded by

$$P_e \leq \sum_{(H_r,H_c)\in\mathcal{H}^n_{\delta,L}(R)} \Big\{P[\mathsf{E}_{0,s}|H_r,H_c]\Big\} \tag{62}$$

$$+ \sum_{(H_r,H_c)\in\mathcal{H}^{n,c}_{\delta,L}(R)} \Big\{$$

$$P[\bar{\mathsf{E}}_{1,r}|H_r,H_c]$$

$$+ P[\mathsf{E}_{2,r}\cup\cdots\cup\mathsf{E}_{N,r}|H_r,H_c]$$

$$+ P[\mathsf{E}_{0,r}|H_r,H_c]$$

$$+ P[\bar{\mathsf{E}}_{1,d}|\mathsf{E}_{1,r},\bar{\mathsf{E}}_{2,r}\cap\cdots\cap\bar{\mathsf{E}}_{N,r},H_r,H_c]$$

$$+ P[\mathsf{E}_{2,d} \cup \cdots \cup \mathsf{E}_{N,d} | \mathsf{E}_{1,r}, \bar{\mathsf{E}}_{2,r} \cap \cdots \cap \bar{\mathsf{E}}_{N,r}, H_r, H_c]\Big\}$$

$$+ \sum_{(H_r,H_c)\in\mathcal{H}^{n,s}_{\delta,L}(R)} \Big\{$$

$$P[\bar{\mathsf{E}}_{1,d} | H_r, H_c]$$

$$+ P[\mathsf{E}_{2,d} \cup \cdots \cup \mathsf{E}_{N,d} | H_r, H_c]\Big\},$$

where $\bar{\mathsf{E}}$ denotes the complement of the event E and we made the following definitions,

$\mathsf{E}_{0,s} = \{$ codeword sent by the source has power greater than $1+\epsilon$, i.e., $\sum_{i=1}^{n} |X_i(1)|^2 > n(1+\epsilon)\}$.

$\mathsf{E}_{0,r} = \{$ codeword sent by the relay has power greater than $1+\epsilon$, i.e., $\sum_{i=n_1+1}^{n} |U_i(1)|^2 > (n-n_1)(1+\epsilon)\}$.

$\mathsf{E}_{w,r} = \{$ codeword w jointly typical with the received message at the relay after time n_1, i.e., $(X_1^{n_1}(w), Y_1^{n_1}) \in A_\epsilon^{n_1}(X, Y|H_r)\}$.

$\mathsf{E}_{w,d} = \{$ codeword w jointly typical with both parts of the received message at the destination after time n, i.e., $(X_1^{n_1}(w), Z_1^{n_1}) \in A_\epsilon^{n_1}(X, Z|H_s)$ and $(X_{n_1+1}^{n}(w), U_{n_1+1}^{n}(w), Z_1^{n_1}) \in A_\epsilon^{n-n_1}(X, U, Z|H_c)\}$.

Now, by Lemma 1, we know that

$$P[\mathsf{E}_{0,s}|H_r, H_c] \le \frac{K_l^X}{n^l}, \tag{63}$$

and by choosing l sufficiently large, we can make the sum $\sum P[\mathsf{E}_{0,s}|H_r, H_c]$ in (62) go to 0 as $n \to \infty$. Likewise, we have the following bound

$$P[\bar{\mathsf{E}}_{1,r}|H_r, H_c] \le \frac{K_l^{X,U}}{n_1^l} \tag{64}$$

$$\le \frac{K_l^{X,U}}{(\delta n)^l}, \tag{65}$$

and the sum of all these also goes to 0. Employing the union bound for the random coding argument (i.e., [6], section 8.7), we have that

$$P[\mathsf{E}_{2,r} \cup \cdots \cup \mathsf{E}_{N,r}|H_r, H_c] \le K 2^{nR} 2^{-n_1(I(X;Y)-3\epsilon)} \tag{66}$$

$$\le K 2^{nR} 2^{-n_1(C(H_r)-3\epsilon)} \tag{67}$$

$$\le K 2^{n(R-n_1/nC(H_r)+3\epsilon)} \tag{68}$$

$$\le K 2^{-n\epsilon}, \tag{69}$$

where the last equality follows from (53). The fourth term is bounded as

$$P[\mathsf{E}_{0,r}|H_r, H_c] \le \frac{K_l^U}{(n-n_1)^l} \tag{70}$$

$$\le \frac{K_l^U}{(\delta n)^l}, \tag{71}$$

and the sum of these terms also goes to zero for l sufficiently large. We now consider the next to last term inside the summation over $\mathcal{H}^{n,c}_{\delta,L}(R)$. This may be bounded as

$$P[\bar{\mathsf{E}}_{1,d}|\mathsf{E}_{1,r}, \bar{\mathsf{E}}_{2,r} \cap \cdots \cap \bar{\mathsf{E}}_{N,r}, H_r, H_c] \leq M(n)\left[\frac{K_l^{X,Z}}{n_1^l} + \frac{K_l^{X,U,Z}}{(n-n_1)^l}\right] \quad (72)$$

$$\leq M(n)\frac{K_l^{X,Z} + K_l^{X,U,Z}}{(\delta n)^l}, \quad (73)$$

where $M(n) = [1 - K_1^{X,U}/n - K2^{-n\epsilon}]^{-1}$ and we have made use of the inequality $P[A|B,C] \leq P[A]/(P[B] - P[\bar{C}])$ which holds when $P[\bar{C}] < P[B]$. For an appropriate choice of l, the sum of all these terms goes to 0 as $n \to \infty$. Finally, employing the fact that the probability of joint typicality between $X_1^{n_1}$ and $Z_1^{n_1}$ during the listening phase is independent of the probability of joint typicality between $X_{n_1+1}^n$, $U_{n_1+1}^n$ and $Z_{n_1+1}^n$ during the collaborative phase,

$$P[\mathsf{E}_{2,d} \cup \cdots \cup \mathsf{E}_{N,d}|\mathsf{E}_{1,r}, \bar{\mathsf{E}}_{2,r} \cap \cdots \cap \bar{\mathsf{E}}_{N,r}, H_r, H_c]$$

$$\leq KM(n)2^{nR}2^{-n_1(I(X;Z)-3\epsilon)} \times 2^{-(n-n_1)(I(X,U;Z)-3\epsilon)} \quad (74)$$

$$\leq KM(n)2^{nR}2^{-n_1(C(H_s)-3\epsilon)} \times 2^{-(n-n_1)(C(H_c)-3\epsilon)} \quad (75)$$

$$\leq KM(n)2^{-n\epsilon} \quad (76)$$

where we have again made use of $P[A|B,C] \leq P[A]/(P[B] - P[\bar{C}])$ and the last inequality follows from (52).

Finally, the sum of the two terms inside the summation over $\mathcal{H}^{n,s}_{\delta,L}(R)$ also vanishes as $n \to \infty$ for similar reasons.

Part 4) Extension to $\mathcal{H}_{\delta,L}(R)$.

We have now shown the existence of a sequence of codes for which the maximum probability of error over a progressively finer approximation $\mathcal{H}^{n}_{\delta,L}(R)$ to the true compound relay channel $\mathcal{H}_{\delta,L}(R)$ goes to 0 as $n \to \infty$. We shall complete the proof by showing that such a code may easily be extended to a good code over the entire compound relay channel $\mathcal{H}_{\delta,L}(R)$. We shall illustrate the case of the relay only (the case of the destination node is similar, whether the relay collaborated or remained silent). In particular, consider any realization $(H_r, H_c) \in \mathcal{H}^{c}_{\delta,L}(R)$. The relay looks for an (H'_r, H'_c) in the compound channel $\mathcal{H}^{n,c}_{\delta,L}(R)$ such that

1) $n_1(n, H'_r) = n_1(n, H_r)$
2) $||H_r - H'_r|| \leq 1/n^2$.

The decoder then discards the H'_c matrix and applies the decoding algorithm for H'_r even though the actual channel is H_r, i.e., we take $R_{w,n_1,H_r} = R_{w,n_1,H'_r}$. We know that for any w and H'_r,

$$P[Y_1^{n_1} \in R_{w,n_1,H'_r} | X_1^{n_1}(w), H'_r] \geq 1 - \lambda_n \tag{77}$$

with $\lambda_n \to 0$. We now claim that as $n \to \infty$, the approximation H'_r approaches the true H_r fast enough so that the penalty to the probability of correct detection for any w and H_r vanishes uniformly as $n \to \infty$. To demonstrate this, it will suffice to show that the ratio of the density functions

$$\frac{p(y_1^{n_1} | x_1^{n_1}(w), H_r, y_1^{n_1} \in R_{w,n_1,H'_r})}{p(y_1^{n_1} | x_1^{n_1}(w), H'_r, y_1^{n_1} \in R_{w,n_1,H'_r})} \tag{78}$$

approaches 1 uniformly over all w and H_r and $y_1^{n_1} \in R_{w,n_1,H'_r}$ as $n \to \infty$. In particular, if we let

$$1 - \beta_n \triangleq \inf_{H_r, w} \left[\inf_{y_1^{n_1} \in R_{w,n_1,H'_r}} \frac{p(y_1^{n_1} | x_1^{n_1}(w), H_r)}{p(y_1^{n_1} | x_1^{n_1}(w), H'_r)} \right], \tag{79}$$

then the probability of error on the compound channel increases from λ_n on $\mathcal{H}_{\delta,L}^{c,n}(R)$ to at most $\lambda_n + \beta_n - \lambda_n \beta_n$ on $\mathcal{H}_{\delta,L}^{c}(R)$. We observe that the ratio in (79) may be expressed as

$$\frac{p(y_1^{n_1} | x_1^{n_1}(w), H_r)}{p(y_1^{n_1} | x_1^{n_1}(w), H'_r)} = \frac{e^{-(y_1^{n_1} - H_r x_1^{n_1})^\dagger (y_1^{n_1} - H_r x_1^{n_1})/2}}{e^{-(y_1^{n_1} - H'_r x_1^{n_1})^\dagger (y_1^{n_1} - H'_r x_1^{n_1})/2}} \tag{80}$$

$$= e^{[2(y_1^{n_1})^\dagger (H_r - H'_r) x_1^{n_1} + (x_1^{n_1})^\dagger ({H'_r}^\dagger H'_r - {H_r}^\dagger H_r) x_1^{n_1}]/2}. \tag{81}$$

However, we note that $||H'_r - H_r|| < 1/n^2$ and $||{H'_r}^\dagger H'_r - H_r^\dagger H_r|| \leq 2(N_s + N_r)L/n^2$. Furthermore, by construction $R_{w,n_1,H'_r} \subset A_\epsilon^{n_1}(Y|H'_r)$, and hence we have that $||y_1^{n_1}||_2 \leq n_1(L^2 + 2\epsilon) \leq n(L^2 + 2\epsilon)$ and likewise $||x_1^{n_1}||_2 \leq n(1 + 2\epsilon)$. It then follows that (since $N_s = N_r = 1$),

$$\begin{aligned} &|2(y_1^{n_1})^\dagger (H_r - H'_r) x_1^{n_1} + (x_1^{n_1})^\dagger ({H'_r}^\dagger H'_r - {H_r}^\dagger H_r) x_1^{n_1}| \\ &\quad \leq 2\sqrt{(L^2 + 2\epsilon)(1 + 2\epsilon)}/n + 4L(1 + 2\epsilon)/n. \end{aligned} \tag{82}$$

It then follows that $1 - \beta_n \to 1$ as $n \to \infty$ and that the convergence is uniform in w and H_r and $y_1^{n_1} \in R_{w,n_1,H'_r}$.

The destination node employs a similar procedure with the actual H_c being approximated by some H'_c. ∎

B. Proof of Theorem 2

Proof of Theorem 2: The proof largely follows that of Theorem 1 and is similar to the approach in [5]. We will only highlight the differences from the proof of Theorem 1. Intuitively, we choose the set R_{w,H_r,n_1,d_r} to be all Y sequences of length n_1 jointly typical with the received message w given that there was a delay of d_r. When the decoder compares $y_1^{n_1}$ with R_{w,H_r,n_1,d_r} where w is the correct message and d_r the correct delay, $y_1^{n_1} \in R_{w,H_r,n_1,d_r}$ with high probability. If the correct w is employed with an incorrect $\tilde{d}_r$, then regardless of the outcome, no harm is done. Finally, when the decoder compares $y_1^{n_1}$ with either $R_{\tilde{w},H_r,n_1,\tilde{d}_r}$ or $R_{\tilde{w},H_r,n_1,d_r}$, the received sequence $y_1^{n_1}$ is unlikely (with probability of error decaying exponentially) to find a jointly typical match. Since in the latter case, compared to synchronous decoding, we have increased the number of ways that an error can be declared by a fixed factor, $D+1$, the total contribution still decays exponentially. Likewise, the set S_{w,H_c,n_1,d_r,d_d} is similarly defined with guard bands of length D on each side of the phase transition from listening to collaboration. The destination does not base its decoding decision on the contents of these guard bands. This guarantees that the destination node will not confuse any part of a transmission during the listening phase for one during the collaboration phase for any d_r and d_d. Hence, at the destination the probability of error compared to synchronous decoding increases by at most a constant factor.

In particular, we choose R_{w,H_r,n_1,d_r} and S_{w,H_c,n_1,d_r,d_d} as follows,

$$
\begin{aligned}
R_{w,H_r,n_1,d_r} = \{y_1^{n_1} :& \\
& (x_{1+d_r}^{n_1+d_r}(w), y_1^{n_1}) \in A_\epsilon^{n_1}(X,Y|H_r), \\
& \forall w' \neq w, \forall d_r' \\
& (x_{1+d_r'}^{n_1+d_r'}(w'), y_1^{n_1}) \notin A_\epsilon^{n_1}(X,Y|H_r)\}
\end{aligned} \tag{83}
$$

$$
\begin{aligned}
S_{w,H_c,n_1,d_r,d_d} = \{z_1^n :& \\
& (x_{1+d_d}^{n_1+d_d-D}(w), z_1^{n_1-D}) \in A_\epsilon^{n_1-D}(X,Z|H_c), \\
& (x_{n_1+1+D+d_d}^{n-D+d_d}(w), u_{n_1+1+D-d_r+d_d}^{n-D-d_r+d_d}(w), z_{n_1+1+D}^{n-D}) \\
& \quad \in A_\epsilon^{n-n_1-2D}(X,U,Z|H_c), \\
& \forall w' \neq w, \forall d_r', d_d' \\
& (x_{1+d_d'}^{n_1+d_d'-D}(w'), z_1^{n_1-D}) \notin A_\epsilon^{n_1-D}(X,Z|H_c)
\end{aligned}
$$

$$(x_{n_1+1+D+d'_d}^{n-D+d'_d}(w'), u_{n_1+1+D-d'_r+d'_d}^{n-D-d'_r+d'_d}(w'), z_{n_1+1+D}^{n-D})$$
$$\notin A_\epsilon^{n-n_1-2D}(X,U,Z)\}.$$

Similar to (57), it will suffice to consider the expression

$$P_e = E_{C^{(n)}}\Big[\frac{1}{N}\sum_w \Big\{ \sum_{0\le d_r, d_d\le D} \sum_{(H_r,H_c)\in\mathcal{H}_{\delta,L}^n(R)} P[\hat{W}\neq W|W=w,H_r,H_c,d_r,d_d]\Big\}\Big]. \tag{84}$$

Then, the bound for P_e in (62) is essentially still valid (if we condition on d_r and d_d and sum over d_r and d_d) provided we make the following changes to the events $\mathsf{E}_{w,r}$ and $\mathsf{E}_{w,d}$,

$$\mathsf{E}_{w,r} = \{\exists d'_r : (X_{1+d'_r}^{n_1+d'_r}, Y_1^{n_1}) \in A_\epsilon^{n_1}(X,Y|H_r)\}$$
$$\mathsf{E}_{w,d} = \{\exists d'_r, d'_d :$$
$$(X_{1+d'_d}^{n_1+d'_d-D}(w), Z_1^{n_1-D}) \in A_\epsilon^{n_1-D}(X,Z|H_s),$$
$$(X_{n_1+1+D+d'_d}^{n-D+d'_d}(w), u_{n_1+1+D-d'_r+d'_d}^{n-D-d'_r+d'_d}(w), z_{n_1+1+D}^{n-D})$$
$$\in A_\epsilon^{n-n_1-2D}(X,U,Z|H_c)\}.$$

Similar to the proof of Theorem 1, the sum of all terms in the expression for P_e goes to zero. For example, consider the sum

$$\sum_{0\le d_r,d_d\le D}\ \sum_{(H_r,H_c)\in\mathcal{H}_{\delta,L}^{n,c}(R)} P[\mathsf{E}_{2,d}\cup\cdots\cup\mathsf{E}_{N,d}|\mathsf{E}_{1,r},\bar{\mathsf{E}}_{2,r}\cap\cdots\cap\bar{\mathsf{E}}_{N,r},H_r,H_c,d_r,d_d]. \tag{85}$$

Since we have that message $W=1$ was transmitted by assumption, from (13) and (14) we can union bound the summand as follows

$$P[\mathsf{E}_{2,d}\cup\cdots\cup\mathsf{E}_{N,d}|\mathsf{E}_{1,r},\bar{\mathsf{E}}_{2,r}\cap\cdots\cap\bar{\mathsf{E}}_{N,r},H_r,H_c]$$
$$\le K(D+1)^2M'(n)2^{nR}2^{-(n_1-D)(I(X;Z)-3\epsilon)} \times 2^{-(n-n_1-2D)(I(X,U;Z)-3\epsilon)} \tag{86}$$
$$\le K(D+1)^2M'(n)2^{nR}2^{-(n_1-D)(C(H_s)-3\epsilon)} \times 2^{-(n-n_1-2D)(C(H_c)-3\epsilon)} \tag{87}$$
$$\le K(D+1)^2M'(n)2^{-n\epsilon}2^{3DC(H_c)}, \tag{88}$$

where we have used the fact that $C(H_s)\le C(H_c)$ since H_s is a submatrix of H_c. However, $C(H_c)$ is continuous on the set of matrices such that $||H_c||\le L$. Hence, we have that $C(H_c)$ may be universally bounded above for all channels $||H_c|| \leq L$. Similar to $M(n)$ in (72), $M'(n) \to 1$ uniformly and the summand is also seen to go to 0 uniformly. Hence, the sum in (85) also goes to 0.

Finally, similar to the proof of Theorem 1, the extension to the compound relay channel $\mathcal{H}_{\delta,L}$ from $\mathcal{H}^n_{\delta,L}$ increases the probability of error from λ_n to at most $\lambda_n + \beta_n - \lambda_n\beta_n$ with $\beta_n \to 0$. ■

APPENDIX II
DERIVATION OF THE OUTAGE PROBABILITY LOWER BOUND OF THE PROPOSED SCHEME FOR QUASI-STATIC RAYLEIGH FADING CHANNELS

In this appendix, we derive the CDF of the random variable $\omega = |H_s|^2 + v|H_{r,d}|^2$ in (36), assuming all the fading coefficients H_r, H_s, $H_{r,d}$ are i.i.d. zero-mean unit-variance complex Gaussian random variables. Consequently, $\alpha_r = |H_r|^2$, $\alpha_s = |H_s|^2$, $\alpha_{r,d} = |H_{r,d}|^2$ are i.i.d. random variables with exponential distribution, i.e., $f_\alpha(a) = e^{-a}$, and $\omega = \alpha_s + v\alpha_{r,d}$.

A. Probability Density and Mean of v

We begin with the derivation of the probability density of the relay transmission interval v defined in (29). Recall that

$$C(H_r, G\gamma) = \log_2\left(1 + G\gamma|H_r|^2\right) \tag{89}$$

$$= \log_2\left(1 + G\gamma\alpha_r\right), \tag{90}$$

where G is a geometrical gain achieved by the relay node over the destination node. Then v can be rewritten as

$$v = \begin{cases} 0 & \text{for} \quad C(H_r, G\gamma) < R \\ 1 - \frac{R}{\log_2(1+\gamma G|H_r|^2)} & \text{for} \quad C(H_r, G\gamma) \geq R \end{cases}$$
$$= \begin{cases} 0 & \text{for} \quad \alpha_r \leq \frac{2^R-1}{\gamma G} \\ 1 - \frac{R}{\log_2(1+\gamma G\alpha_r)} & \text{for} \quad \alpha_r > \frac{2^R-1}{\gamma G}. \end{cases} \tag{91}$$

Changing the random variable through standard probability operations yields,

$$f_v(v) = \frac{\ln 2^R}{\gamma G} \frac{2^{\frac{R}{1-v}} e^{\frac{1}{\gamma G}\left(1-2^{\frac{R}{1-v}}\right)}}{(1-v)^2} + \left(1 - e^{-\frac{2^R-1}{\gamma G}}\right)\delta(v), \tag{92}$$

for $0 \leq v \leq 1$, where $\delta(x)$ is a Dirac delta function.

The mean value of v is given by

$$E[v] = \int_0^1 v f_v(v) dv = \int_{\frac{2^R}{\gamma G}}^{\infty} \left(1 - \frac{\ln 2^R}{\ln(\gamma G t)}\right) e^{\frac{1}{\gamma G} - t} dt. \tag{93}$$

The above integration can be easily computed numerically.

B. Probability Density of $\nu = v\,\alpha_{r,d}$

Making the change of random variables $\nu = v\,\alpha_{r,d}$ and $\eta = (1-v)\,\alpha_{r,d}$ and integrating the auxiliary random variable η out, we get

$$\begin{aligned} f_\nu(\nu) = \frac{\ln 2^R}{\gamma^2 G} \int_0^\infty \frac{\nu+\eta}{\eta^2} 2^R \exp\Big\{ & \frac{1 - 2^{R(\frac{\nu}{\eta}+1)}}{\gamma G} \\ & - \Big(\frac{1}{\gamma} - \frac{\ln 2^R}{\eta}\Big)\nu - \frac{\eta}{\gamma}\Big\} d\eta + \Big(1 - e^{-\frac{2^R-1}{\gamma G}}\Big)\delta(\nu), \end{aligned} \tag{94}$$

for $\nu \geq 0$.

C. Cumulative Distribution of $\omega = \alpha_s + \nu$

Since ω is now a sum of the two independent random variables, we have

$$f_\omega(\omega) = \int_0^\omega f_\nu(\nu) e^{-(\omega-\nu)} d\nu. \tag{95}$$

Setting $\eta = \nu/t$ we get

$$\begin{aligned} f_\omega(\omega) = \frac{\ln 2^R}{\gamma^3 G} \int_0^\omega \int_0^\infty & \left(1 + \frac{1}{t}\right) 2^{R(1+t)} \times \\ & \times \exp\left\{ \frac{1}{\gamma}\left[\frac{1 - 2^{R(1+t)}}{G} - \omega - \frac{\nu}{t} \right] \right\} dt\, d\nu + T(\omega), \end{aligned} \tag{96}$$

where

$$T(\omega) = \frac{1}{\gamma} \int_0^\omega \left(1 - e^{-\frac{2^R-1}{\gamma G}}\right) \delta(x) e^{-\frac{\omega-\nu}{\gamma}} d\nu \tag{97}$$

$$= \frac{1}{\gamma}\left(1 - e^{-\frac{2^R-1}{\gamma G}}\right) e^{-\frac{\omega}{\gamma}}. \tag{98}$$

Carrying out the outer integral of (96), we get

$$\begin{aligned} f_\omega(\omega) = \frac{\ln 2^R}{\gamma^2 G} e^{-\frac{\omega}{\gamma}} & \int_0^\infty (1+t) e^{g(t)} \left(1 - e^{-\frac{\omega}{\gamma t}}\right) dt \\ & + \frac{1}{\gamma}\left(1 - e^{-\frac{2^R-1}{\gamma G}}\right) e^{-\frac{\omega}{\gamma}}, \end{aligned} \tag{99}$$

where

$$g(t) \triangleq \frac{1 - 2^{R(1+t)}}{\gamma G} + \ln 2^{R(1+t)}. \tag{100}$$

Finally, the CDF of ω is given by

$$P[\omega < \zeta] = \int_0^{\zeta} f_\omega(\omega) d\omega. \tag{101}$$

The above integral results in the following form

$$P[\omega < \zeta] = 1 - e^{-\frac{\zeta}{\gamma}} \left(1 + V(\zeta)\right), \tag{102}$$

where

$$V(\zeta) = \frac{\ln 2^R}{\gamma G} \int_0^{\infty} t e^{g(t)} \left(1 - e^{-\frac{\zeta}{\gamma t}}\right) dt. \tag{103}$$

The integral in $V(\zeta)$ can be calculated numerically. Letting $\zeta = \frac{2^R - 1}{\gamma}$ in (102) results in the desired lower bound for the outage probability.

References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey of sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–115, Aug. 2002.

[2] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," Submitted to *IEEE Trans. Inform. Theory*, July 2004.

[3] A. B. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 841–850, Nov. 1982.

[4] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[5] T. M. Cover, R. J. McEliece, and E. C. Posner, "Asynchronous multiple-access channel capacity," *IEEE Trans. Inform. Theory*, vol. 27, no. 4, pp. 409–413, July 1981.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[7] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1981.

[8] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[9] M. Gastpar, G. Kramer, and P. Gupta, "The multiple relay channel: coding and antenna-clustering capacity," in *Proc. IEEE Int. Symp. Inform. Theory*, Lausanne, Switzerland, July 2002, p. 136.

[10] M. Gastpar and M. Vetterli, "On the asymptotic capacity of Gaussian relay networks," in *Proc. IEEE Int. Symp. Inform. Theory*, Lausanne, Switzerland, July 2002, p. 195.

[11] ——, "On the capacity of wireless networks: the relay case," in *Proc. IEEE INFOCOM*, New York, NY, June 2002, pp. 1577–1586.

[12] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: an achievable rate region," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1877–1894, Aug. 2003.

[13] T. Hunter, A. Hedayat, M. Janani, and A. Nosratinia, "Coded cooperation with space-time transmission and iterative decoding," in *WNCG Wireless Networking Symposium*, Oct. 2003.

[14] T. Hunter and A. Nosratinia, "Coded cooperation under slow fading, fast fading, and power control," in *Asilomar Conference on Signals, Systems, and Computers*, Nov. 2002.

[15] M. Katz and S. Shamai, "Communicating to co-located ad-hoc receiving nodes in a fading environment," in *Proc. IEEE Int. Symp. Inform. Theory*, Chicago, IL, July 2004, p. 115.

[16] M. Khojastepour, B. Aazhang, and A. Sabharwal, "Bounds on achievable rates for general multi-terminal networks with practical constraints," in *Information Processing in Sensor Networks*, Apr. 2003.

[17] M. Khojastepour, A. Sabharwal, and B. Aazhang, "On capacity of Gaussian 'cheap' relay channel," in *Proc. IEEE Global Telecommun. Conf.*, Apr. 2003, pp. 1776–1780.

[18] ——, "On the capacity of 'cheap' relay networks," in *Conference on Information Sciences and Systems*, Apr. 2003.

[19] G. Kramer, M. Gastpar, and P. Gupta, "Capacity theorems for wireless relay channels," in *Proc. Allerton Conf. Commun., Control and Comp.*, 2003, pp. 1074–1083.

[20] ——, "Cooperative strategies and capacity theorems for relay networks," Submitted to *IEEE Trans. Inform. Theory*, 2004.

[21] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[22] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[23] A. Reznik, S. Kulkarni, and S. Verdú, "Capacity and optimal resource allocation in the degraded Gaussian relay channel with multiple relays," in *Proc. Allerton Conf. Commun., Control and Comp.*, Monticello, IL, Oct. 2002.

[24] B. Schein and R. G. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Int. Symp. Inform. Theory*, Sorrento, Italy, June 2000, p. 22.

[25] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity–part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[26] ——, "User cooperation diversity–part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[27] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm.*, vol. 10, no. 6, pp. 585–595, 1999.

[28] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.

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

[30] F. M. J. Willems, "Information theoretic results for the discrete memoryless multiple access channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Oct. 1982.

[31] F. M. J. Willems, E. C. van der Meulen, and J. P. M. Schalkwijk, "An achievable rate region for the multiple access channel with generalized feedback," in *Proc. Allerton Conf. Commun., Control and Comp.*, Oct. 1983, pp. 284–293.

[32] J. Wolfowitz, *Coding Theorems of Information Theory*. New York: Springer-Verlag, 1978.

[33] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inform. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

PLACE
PHOTO
HERE

Patrick Mitran (S'01) received the Bachelor's and Master's degrees in electrical engineering, in 2001 and 2002, respectively, from McGill University, Montreal, PQ, Canada. He is currently working toward the Ph.D. degree in the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA.

His research interests include iterative decoding theory, joint source-channel coding, detection and estimation theory as well as information theory.

PLACE
PHOTO
HERE

Hideki Ochiai (S'97, M'01) received the B.E. degree in communication engineering from Osaka University, Osaka, Japan, in 1996, and the M.E. and Ph.D. degrees in information and communication engineering from The University of Tokyo, Tokyo, Japan, in 1998 and 2001, respectively. From 2001 to 2003, he was with the Department of Information and Communication Engineering, The University of Electro-Communications, Tokyo, Japan. Since April 2003, he has been with the Division of Physics, Electrical and Computer Engineering, Yokohama National University, Yokohama, Japan, where he is currently an Assistant Professor. From 2003 to 2004, he was a Visiting Scientist at the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA.

Dr. Ochiai was a recipient of a Student Paper Award from the Telecommunications Advancement Foundation in 1999 and the Ericsson Young Scientist Award in 2000.

PLACE
PHOTO
HERE

Vahid Tarokh (M'95, SM'03) received the PhD in Electrical Engineering from the University of Waterloo, Ontario, Canada in 1995. He then worked at AT&T Labs-Research and AT&T wireless services, where his last position was the head of the Department of Wireless Communications and Signal Processing. In 2000, he joined Department of EECS at MIT as an associate professor and in 2002, he joined Harvard University as a professor of electrical engineering and a Vinton Hayes senior research fellow. He is a co-founder of Center for Communications and Networking (CNC) (jointly with Professor H.T. Kung) and a member of Harvard Broadband Communications Laboratory .

Dr. Tarokh is responsible for a number of inventions, most notably his invention of Space-Time Coding (jointly with N. Seshadri and A.R. Calderbank). He has been ranked in the list of "Top 10 Most Cited Researchers in Computer Science" by the ISI Web of Science and has received a number of awards including the Gold Medal of the Governor General of Canada 1995 , the IEEE Information Theory Society Prize Paper Award 1999, The Alan T. Waterman Award 2001. In 2002, he was selected as one of the Top 100 Inventors of Year by the Technology Review Magazine. In 2003, he received an honourary D.Sc. from the University of Windsor.

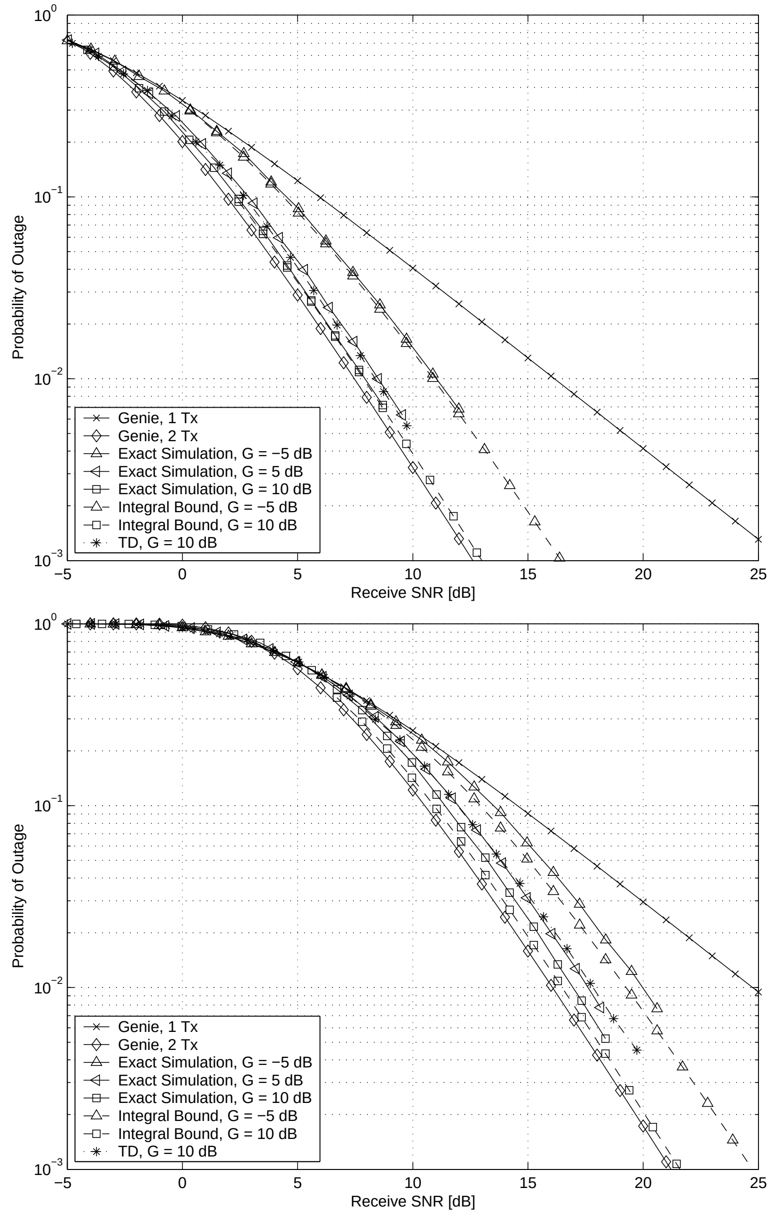

Fig. 3. Outage probability of two transmit collaborators, one receiver with R = 0.5 (top) and R = 2 (bottom) for various geometric gain factors G.

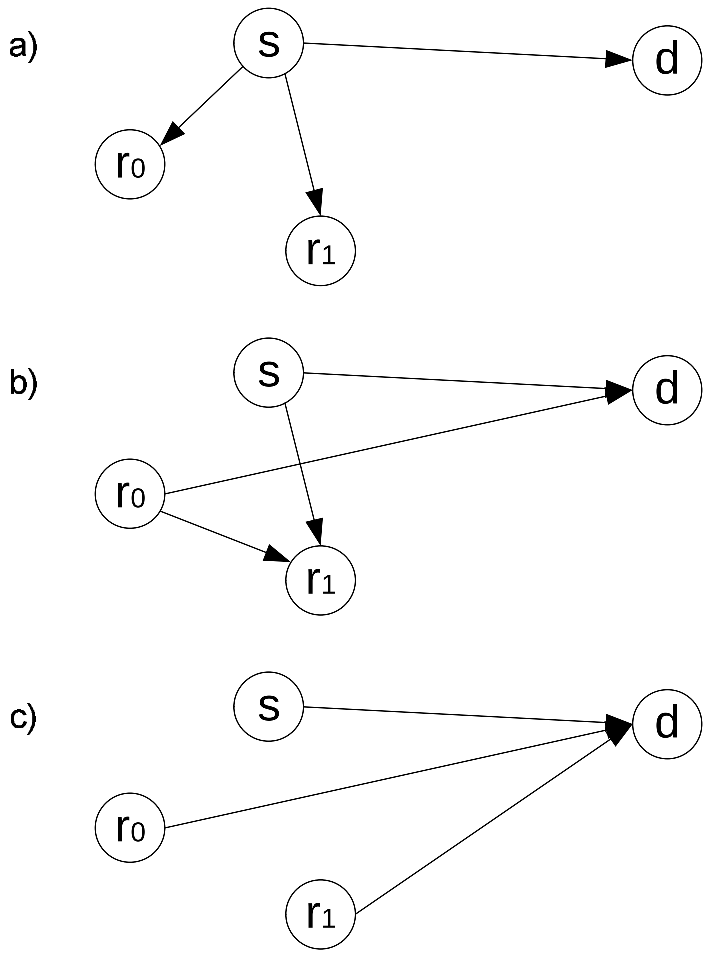

Fig. 4. a) The three collaborators problem with one receiver. b) Node r_0 has stopped listening and started collaborating. It transmits to nodes d *and* r_1. c) Node r_1 has stopped listening and started collaborating.

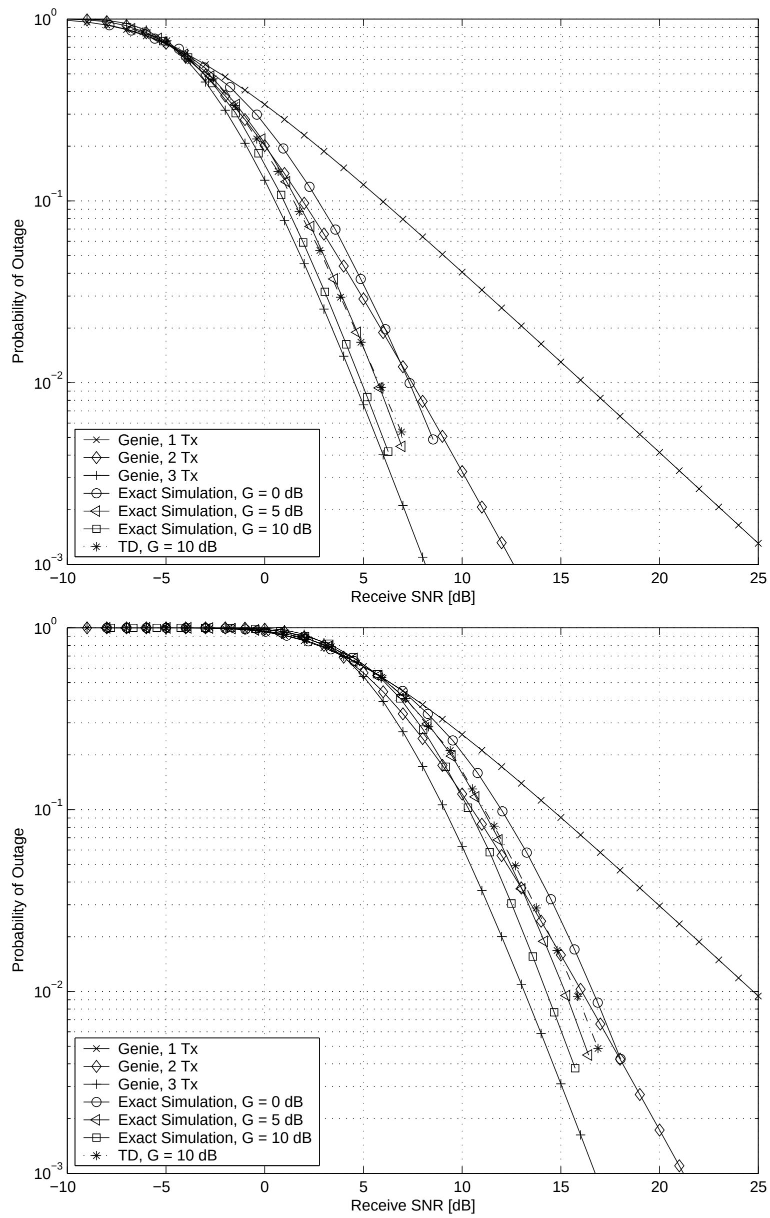

Fig. 5. Outage probability of three transmit collaborators, one receiver with R = 0.5 (top) and R = 2 (bottom) for various geometric gain factors G.

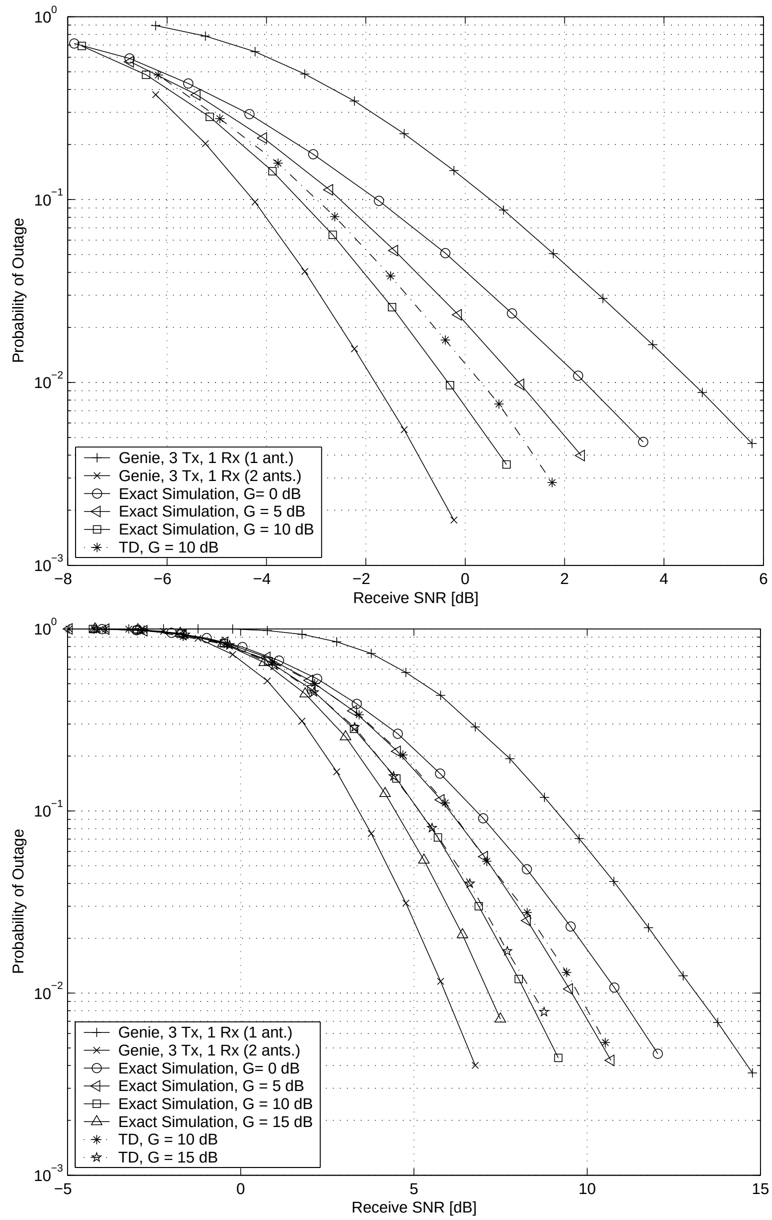

Fig. 6. Outage probability of three transmit collaborators, one receive collaborator (with two antennas) with R = 0.5 (top) and R = 2 (bottom) for various geometric gain factors G.

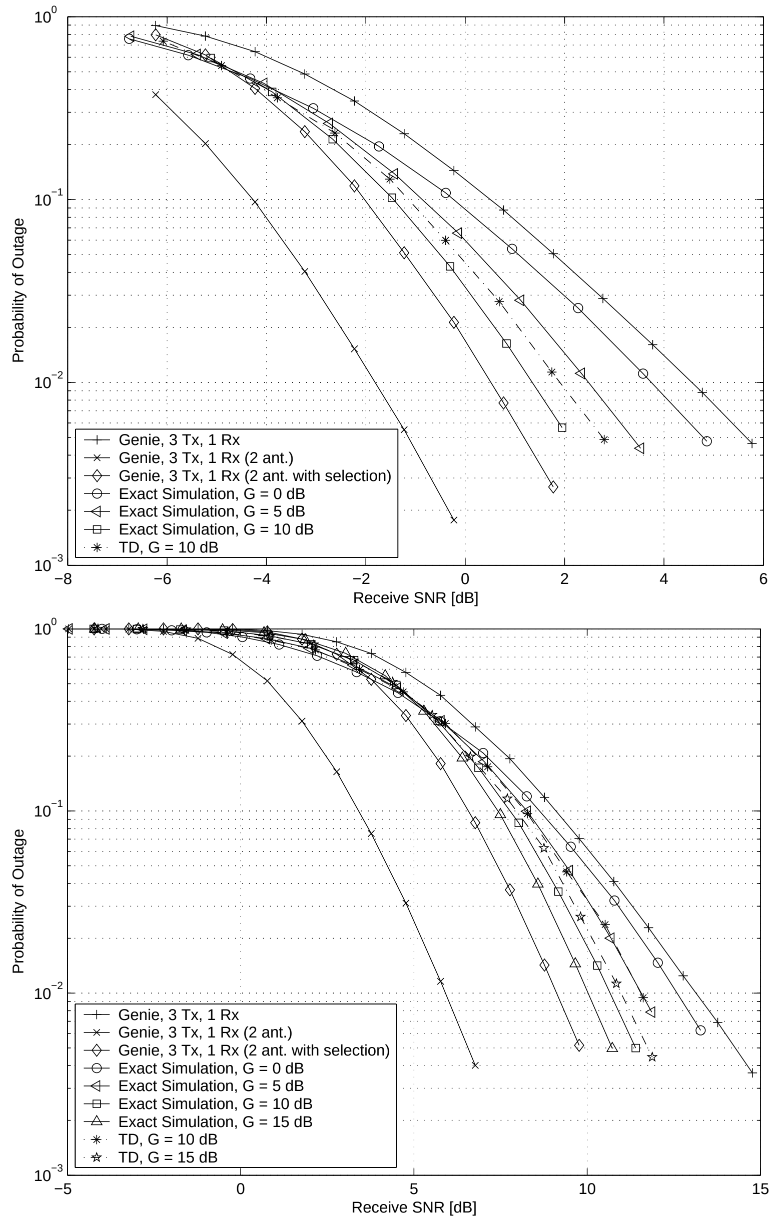

Fig. 7. Outage probability of three transmit collaborators, one receive collaborator (with two antennas) using antenna selection with R = 0.5 (top) and R = 2 (bottom) for various geometric gain factors G.