Multiuser Cooperative Diversity
for Wireless Networks

Guest Editors: George K. Karagiannidis, Chintha Tellambura, Sayandev Mukherjee, and Abraham O. Fapojuwo
Multiuser Cooperative Diversity for Wireless Networks
Multiuser Cooperative Diversity for Wireless Networks

Guest Editors: George K. Karagiannidis, Chintha Tellambura, Sayandev Mukherjee, and Abraham O. Fapojuwo
Contents

**Multiuser Cooperative Diversity for Wireless Networks**, George K. Karagiannidis, Chintha Tellambura, Sayande Mukherjee, and Abraham O. Fapojuwo
Volume 2006 (2006), Article ID 17202, 3 pages

**Routing and Power Allocation in Asynchronous Gaussian Multiple-Relay Channels**, Zigui Yang and Anders Host-Madsen
Volume 2006 (2006), Article ID 56914, 11 pages

**Joint MIMO-OFDM and MAC Design for Broadband Multihop Ad Hoc Networks**, Dandan Wang and Uf Tureli
Volume 2006 (2006), Article ID 60585, 9 pages

**A Novel Cluster-Based Cooperative MIMO Scheme for Multi-Hop Wireless Sensor Networks**, Yong Yuan, Min Chen, and Taekyoung Kwon
Volume 2006 (2006), Article ID 72493, 9 pages

**Relay Techniques for MIMO Wireless Networks with Multiple Source and Destination Pairs**, Tetsushi Abe, Hui Shi, Takahiro Asai, and Hitoshi Yoshino
Volume 2006 (2006), Article ID 64159, 9 pages

**Nonregenerative Dual-Hop Cooperative Links with Selection Diversity**, Theodoros A. Tsiftsis, George K. Karagiannidis, P. Takis Mathiopoulos, and Stavros A. Kotsopoulos
Volume 2006 (2006), Article ID 17862, 8 pages
Editorial

Multiuser Cooperative Diversity for Wireless Networks

George K. Karagiannidis,1 Chintha Tellambura,2 Sayandev Mukherjee,3 and Abraham O. Fapojuwo4

1Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece
2Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6G2V4
3Marvell Semiconductor, Santa Clara, CA 95054, USA
4Department of Electrical and Computer Engineering, The University of Calgary, 2500 University Drive N.W., Calgary, AB, Canada T2N1N4

Received 24 September 2006; Accepted 24 September 2006

Copyright © 2006 George K. Karagiannidis et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Multihop relaying technology is a promising solution for future cellular and ad hoc wireless communications systems in order to achieve broader coverage and to mitigate wireless channels impairment without the need to use high power at the transmitter. Recently, a new concept that is being actively studied in multihop-augmented networks is multiuser cooperative diversity, where several terminals form a kind of coalition to assist each other with the transmission of their messages. In general, cooperative relaying systems have a source node multicasting a message to a number of cooperative relays, which in turn resend a processed version to the intended destination node. The destination node combines the signal received from the relays, possibly also taking into account the source’s original signal. Cooperative diversity exploits two fundamentals features of wireless medium: its broadcast nature and its ability to achieve diversity through independent channels. There are three advantages from this.

(1) Diversity. This occurs because different paths are likely to fade independently. The impact of this is expected to be seen in the physical layer, in the design of a receiver that can exploit this diversity.

(2) Beamforming gain. The use of directed beams should improve the capacity on the individual wireless links. The gains may be particularly significant if space-time coding schemes are used.

(3) Interference mitigation. A protocol that takes advantage of the wireless channel and the antennas and receivers available could achieve a substantial gain in system throughput by optimizing the processing done in the cooperative relays and in the scheduling of retransmissions by the relays so as to minimize mutual interference and facilitate information transmission by cooperation.

In response to the demand for novel ideas and results, this special issue presents a sample of current activities and up-to-date efforts in design, implementation, and performance analysis of cooperative diversity systems. A brief summary of each paper is listed as follows.

In the first paper by Z. Yang and A. Høst-Madsen, the cooperation efficiency of the multiple-relay channel when carrier-level synchronization is not available is investigated, assuming that all nodes use a decode-forward scheme. It is shown that by using decode-forward relay signaling, the transmission is interference-free, even when all communications share one common physical medium. Furthermore, for any channel realization, there always exist a sequential path and a corresponding simple power-allocation policy, which are optimal. To illustrate the efficiency of cooperation and provide prototypes for practical implementation of relay-channel signaling, the authors propose two heuristic algorithms. Finally, the numerical results show that in the low-rate regime, the gain from cooperation is limited, while the gain is considerable in the high-rate regime.

In the second paper, by D. Wang and U. Tureli, the authors try to face the inefficiencies caused due to the existing medium-access control (MAC) schemes, when multiple-input multiple-output (MIMO) transmit/receive schemes and orthogonal frequency-division multiplexing (OFDM) are used in broadband multihop ad hoc networks. A new transceiver architecture with MIMO-OFDM and MAC scheme is proposed, named multiple-antennas receiver-initiated busy-tone medium access (MARI-BTMA), which
is based on receiver-initiated busy-tone medium access (RI-BTMA) and uses multiple out-of-band busy tones to avoid the collision of nodes on the same channel. With the proposed MAC scheme, multiple users can transmit simultaneously in the same neighborhood. Although the proposed MARI-BTMA shows good performance at high traffic load, to improve the performance at low traffic loads, 1-persistent MARI-BTMA is proposed so that users can choose different MAC schemes according to the statistical traffic load in the system. In the same paper, both theoretical and numerical analyses of the throughput and delay are presented, while analytical and simulation results show the improved performance of MARI-BTMA compared with RI-BTMA and carrier sensing medium access/collision avoidance (CSMA/CA).

In the third paper, by Y. Yuan et al., a cluster-based cooperative MIMO scheme is proposed to reduce the adverse impacts caused by radio irregularity and fading in multihop wireless sensor networks. This is an extension of the LEACH protocol, enabling the multihop transmissions among clusters by incorporating a cooperative MIMO scheme into hop-by-hop transmissions. The proposed scheme can gain effective performance improvement through the adaptive selection of cooperative nodes and the coordination between multihop routing and cooperative MIMO transmissions. Moreover, the optimal parameters which minimize the overall energy consumption, such as the number of clusters and the number of cooperative nodes, are derived. Simulation results exhibit that the proposed scheme can effectively save energy and prolong the network lifetime.

In the fourth paper, by T. Abe et al., the MIMO relay scheme is proposed where each of the multiple-relay nodes performs QR decompositions of the backward and forward channel matrices in conjunction with phase control (QR-P-QR). A group nulling approach is used to decompose a multiple source-destination (SD) MIMO relay channel into parallel independent SD MIMO relay channels, and then apply the QR-P-QR scheme to each of the decomposed MIMO relay links. Numerical examples show that the proposed relay scheme offers higher capacity than other existing relay schemes.

In the last paper by T. A. Tsiftsis et al., the end-to-end performance of dual-hop cooperative selective diversity links, equipped with nonregenerative relays and operating over nonidentical Nakagami-$m$ fading channels, is studied. Closed-form expressions are presented for the cumulative distribution function and the probability density function of the end-to-end signal-to-noise ratio (SNR), while analytical formulae are derived for the moments and the moment-generating function. The proposed mathematical analysis is complemented by numerical examples, including the effects on the overall performance of the SNRs unbalancing as well as the fading severity.

**ACKNOWLEDGMENTS**

We would like to thank all the authors of the papers submitted (either accepted or not) for considering this issue as a means of publication of their own work, the reviewers who allowed us to make our editorial decisions in a timely manner, and the Editor-in-Chief, Phillip Regalia, for giving us the opportunity and the support to achieve this special issue. We hope that these five papers will contribute to the literature of this very exciting research area and also motivate further research.

George K. Karagiannidis
Chintha Tellambura
Sayandev Mukherjee
Abraham Fapojuwo

George K. Karagiannidis was born in Pithagorion, Samos Island, Greece. He received his university degree in 1987 and his Ph.D. degree in 1999, both in electrical engineering, from the University of Patras, Patras, Greece. From 2000 to 2004, he was a Researcher at the Institute for Space Applications and Remote Sensing, National Observatory of Athens, Greece. In June 2004, he joined the faculty of Aristotle University of Thessaloniki, Greece, where he is currently an Assistant Professor at the Electrical and Computer Engineering Department. His major research interests include wireless communications theory, digital communications over fading channels, cooperative diversity systems, satellite communications, and free-space optical communications. He has published and presented more than 70 technical papers in scientific journals and international conferences, he is a coauthor in two chapters in books and also coauthor in a Greek edition book on mobile communications. He acts as a reviewer for several international journals and he served as a Technical Program Committee Member for ICC ’04 and ICC ’05. He is Member of the editorial boards of IEEE Communications Letters and EURASIP Journal on Wireless Communications and Networking.

Chintha Tellambura received the B.S. degree (with first-class honors) from the University of Moratuwa, Moratuwa, Sri Lanka, in 1986, the M.S. degree in electronics from the University of London, London, UK, in 1988, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 1993. He was a Postdoctoral Research Fellow with the University of Victoria (1993–1994) and the University of Bradford (1995–1996). He was with Monash University, Melbourne, Australia, from 1997 to 2002. Presently, he is a Professor with the Department of Electrical and Computer Engineering, University of Alberta, Canada. His research interests include coding, communication theory, modulation, equalization, and wireless communications. He is an Associate Editor for both the IEEE Transactions on Communications and the IEEE Transactions on Wireless Communications. He was Chair of the Communication Theory Symposium in Globecom ’05 held in St. Louis, Mo, USA.
Sayandev Mukherjee was born in Bangalore, India, in 1970. He received the Bachelor of Technology degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1991, and the M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 1994 and 1997, respectively. From 1996 to October 2006, he was a member of technical staff in the Wireless Research Laboratory, Bell Labs, Lucent Technologies, NJ, USA. He is now a Senior Staff Design Engineer with Marvell Semiconductor Inc., Santa Clara, Calif, USA. His research interests include stochastic models, wireless system simulations, and connectivity issues in ad hoc wireless networks.

Abraham O. Fapojuwo received the B.Eng. degree (first-class honors) from the University of Nigeria, Nsukka, in 1980, and the M.S. and Ph.D. degrees in electrical engineering from the University of Calgary, Calgary, AB, Canada, in 1986 and 1989, respectively. From 1990 to 1992, he was a Research Engineer with NovAtel Communications Ltd., where he performed numerous exploratory studies on the architectural definition and performance modeling of digital cellular systems and personal communications systems. From 1992 to 2001, he was with Nortel Networks, where he conducted, led, and directed system-level performance modeling and analysis of wireless communication networks and systems. In January 2002, he joined the Department of Electrical and Computer Engineering, University of Calgary, as an Associate Professor. He is also an Adjunct Scientist at TRLabs, Calgary. His current research interests include protocol design and analysis for future generation wireless communication networks and systems, and best practices in software reliability engineering and requirements engineering. He is a registered professional Engineer in the Province of Alberta.
Routing and Power Allocation in Asynchronous Gaussian Multiple-Relay Channels

Zigui Yang and Anders Høst-Madsen

Department of Electrical Engineering, University of Hawaii at Manoa, Honolulu, HI 96822, USA

Received 31 October 2005; Revised 28 April 2006; Accepted 2 May 2006

We investigate the cooperation efficiency of the multiple-relay channel when carrier-level synchronization is not available and all nodes use a decode-forward scheme. We show that by using decode-forward relay signaling, the transmission is effectively interference-free even when all communications share one common physical medium. Furthermore, for any channel realization, we show that there always exist a sequential path and a corresponding simple power allocation policy, which are optimal. Although this does not naturally lead to a polynomial algorithm for the optimization problem, it greatly reduces the search space and makes finding heuristic algorithms easier. To illustrate the efficiency of cooperation and provide prototypes for practical implementation of relay-channel signaling, we propose two heuristic algorithms. The numerical results show that in the low-rate regime, the gain from cooperation is limited, while the gain is considerable in the high-rate regime.

Copyright © 2006 Z. Yang and A. Høst-Madsen. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

A wireless ad hoc network is an infrastructureless network, in which the communications between two nodes are typically maintained by the cooperation of other nodes. The traditional multihopping operation lets each intermediate node receive information only from its immediate predecessor and then send it to its immediate successor. A more advanced operation is to use relay-channel signaling. The essential difference between the traditional multihopping and the relay-channel signaling is that in the latter, a node uses the information from all its upstream nodes instead of the information from the closest one.

The relay channel was first introduced by van der Meulen [1, 2]. In a simplest case, a relay channel has only one relay to assist the transmission between the source and the destination. The relay channel can be denoted by \((X_1, X_2, Y_2, Y_3)\), where \(X_1, X_2\) are the transmitter alphabets of the source and the relay, respectively, \(Y_2, Y_3\) are the receiver alphabets of the relay and the destination, respectively, and a collection of probability \(p(\cdot, \cdot | x_1, x_2)\) on \(Y_2, Y_3\), one for each \((x_1, x_2) \in X_1, X_2\). Here \(x_1, x_2\) are the channel inputs by the source and the relay and \(y_2, y_3\) are the outputs of the relay and the destination, respectively. The relay channel was extensively studied in [3], where two cooperation schemes, decode-forward and compress-forward, were proposed. Inspired by a renewed interest in ad hoc networks and network information theory, much research has been done recently on relay channels and cooperative diversity [4–14].

We assume that every node uses a decode-forward scheme. Although the other two relaying schemes, amplify-forward and compress-forward, can achieve higher rates under certain channel realizations [4, 6, 7], they are difficult to scale to large networks. In an amplify-forward scheme, the relays essentially act as analog repeaters, and therefore enhance the system noise. Another challenge in using the amplify-forward scheme in large networks is the difficulty of implementing routing algorithms. Compress-forward requires complex Wyner-Ziv coding, which is difficult to be implemented in practice [15], especially when scaled to large networks. Decode-forward has its own drawback in that it requires full decoding at each relay, and therefore may cause error propagation. However, this can be compensated by strong channel coding.

The achievable rate of a one-relay channel using a decode-forward scheme is [3]

\[
R \leq \max_{p(x_1, x_2)} \min \{ I(X_1; Y_2 | X_2), I(X_1, X_2; Y_3) \}. \tag{1}
\]

The interpretation is that the relay first fully decodes the message from the inputs of the source, which results in the first term in the \(\min \{ \cdot \} \) function, and then the destination decodes the messages from the inputs of both the source and the relay, and thus gives the second term in the \(\min \{ \cdot \} \)
function. An adaptive transmission scheme will allow the source to communicate directly with the destination if the relay has a poor link to the source—one form of routing. It then gives the following achievable rate:

$$R \leq \max_{P(x, z)} \max \left\{ \min \left\{ I(X_1; Y_2 | X_2), I(X_1, X_2; Y_3) \right\}, I(X_1; Y_3 | X_2) \right\}.$$  \hfill (2)

For a physically degraded channel, that is, when $X_1 \rightarrow (X_2, Y_2) \rightarrow Y_3$ forms a Markov chain, (1) achieves the capacity. However, for a general relay channel, the capacity is unknown even for one-relay case. Therefore, most of the research on multiple-relay channels concentrated on achievable rates and capacity bounds [5, 6, 8, 16] or on the capacity for some special type of multiple-relay channels such as the degraded multiple-relay channel [17]. A multiple-relay channel is generally a multilevel structure, in which each level contains one or more nodes and the nodes in the same level decode a message at the same time.

The wireless communication broadcast property is referred to as “wireless multicast advantage” (WMA) or “wireless broadcast advantage” (WBA) and may be used in the routing algorithm in wireless networks to reduce power consumption and improve reliability [18, 19]. If different transmitters can be synchronized at carrier level and thus are able to coordinate to use beamforming techniques, it is shown that cooperation achieves significant gain in reducing total power consumption [20]. Relay-channel signaling further exploits the broadcast transmission and multiaccess reception properties by allowing a node to accumulate the soft information of all its received signals, that is, a node’s decoding may depend on multiple received signals. Although it is obvious that relay-channel signaling can further improve the performance, it is at the cost of higher complexity. One fundamental question is whether it pays off using relay-channel signaling or not. In this paper, we will investigate the cooperation efficiency in the multiple-relay-channel setting. Specifically, we consider the quasistatic Gaussian wireless multiple-relay channel.

A quasistatic channel here means that the channel realization remains unchanged during the transmission of one message and goes to another independent realization in the transmission period of the next message. One useful measure of the performance in this scenario is outage probability, which is the probability that the channel cannot support a particular communication rate under certain constraints. The quasistatic model is suitable for delay-sensitive services that have strict delay requirements. For delay-insensitive services, the source and the relay may choose to adjust their transmission rate according to the channel condition [7]. In many applications such as sensor networks, the nodes typically operate on limited-energy batteries, which are usually not rechargeable or replaceable, and thus results in severe energy constraints. A main concern is therefore optimizing energy consumption in the network.

Consider at first a simple point-to-point channel in Rayleigh fading with a channel gain $h$. If it is desired to transmit at a certain constant rate $R$, the required power is proportional to $h^{-2}$ and the average power is proportional to $E[h^{-2}]$, which can be shown to be infinite. Thus, it is impossible to transmit in all channel conditions, and a threshold $h_0$ has to be chosen so that if $h < h_0$, no transmission is done and an outage is declared. Equivalently, a threshold power $P_0$ can be set so that if the required power $P$ for transmission at rate $R$ is larger than $P_0$, transmission is given up and an outage is declared. The average power consumption is an increasing function of $P_0$, while the outage probability is a decreasing function of $P_0$, which should therefore be chosen as a compromise between power consumption and acceptable outage probability. Notice that $P_0$ is not related to the physical power constraint of the transmission circuit of the terminal, although of course $P_0$ must be chosen less than this.

Generalizing this to networks, we consider a total power constraint, that is, at any time, the overall power consumption cannot exceed a particular amount of power $P_0$. This seems the most reasonable point of view: if the total power (energy) needed in the network exceeds a certain threshold, transmission is given up. A precise statement of this is as follows. Assuming that the source-destination pair in the multiple-relay channel wants to maintain a constant communication rate $R$, we define an outage event for a given transmission scheme $T$ and the channel realization $H$ as

$$E_1 : R_T (P_0, H) \leq R,$$

where $R_T (P, H)$ is the maximal rate that the transmission scheme $T$ can achieve for the channel realization $H$ with a total power consumption of at most $P_0$. For all reasonable transmission schemes $T$, $R_T (P, H)$ is a nondecreasing function of $P$. We define $P_T (R, H)$ as the minimum total power required by transmission scheme $T$ to achieve the rate $R$ for the channel realization $H$. Then the outage event is equivalent to the event

$$E_2 : P_T (R, H) > P_0.$$  \hfill (4)

Therefore, we can minimize outage probability by minimizing the total power needed to achieve the target rate $R$ for each channel realization. This problem was investigated for parallel (two-hop) relay channels in [21, 22]. Here we generalize this to multihop channels where arbitrary interrelay communication is allowed. The problem then becomes more complicated, requiring finding both an optimal arrangement of nodes and a corresponding power allocation policy.

Apart from the above overall power constraint, individual node power constraints may also be relevant. Firstly, the power allocation can result in uneven power consumption among the nodes. However, with channel variations, this is averaged out; furthermore, if all nodes at sometime or other act as source-destination pairs, the power consumption can be expected to be fairly distributed. Secondly, the power allocation could result in a solution where an individual node power consumption is above what the node is physically capable of. However, taking this into account would just complicate the solution without giving further insight.

The rest of the paper is organized as follows. In Section 2, we give the model for the Gaussian multiple-relay channel,
for which we will find an optimum arrangement of nodes, shown to be a sequential path, and its corresponding optimal power allocation policy is given in Section 3. To investigate the performance of the relay-channel signaling and to provide some prototype algorithms for practical implementation of relay-channel signaling, we provide two heuristic algorithms for the cooperative relay-channel signaling problem in Section 4. In Section 5, we extend our discussion to the case when nodes have only limited signal processing capability. The numerical results are provided in Section 6 and a brief summarization is given in Section 7.

2. CHANNEL MODEL

In this paper, we consider a quasistatic multiple-relay channel with \(N\) nodes, numbered from 1 to \(N\). Without loss of generality, we assume that 1 and \(N\) is the source-destination pair and that the other nodes act as relays. We assume that all nodes operate in full-duplex mode, and thus they can receive and transmit in the same frequency band at the same time. Full-duplex communication is generally regarded as difficult to achieve in practice, but there are techniques that make it possible [23].

Another important assumption is on synchronization among nodes. There are three levels of synchronization: frame, symbol, and carrier. We assume that the receivers are completely synchronized at all levels. For transmitters, it is realistic to assume that frame- and symbol-level synchronizations are available. The contentious point is on carrier-level synchronization, which requires that separate microwave oscillators at different nodes are synchronized. This seems highly unrealistic. Left by themselves, the drift of the oscillators makes synchronization impossible. It might be possible to couple oscillators, and very closely spaced nodes could even autocouple, but this requires nontrivial microwave innovation, and in general this seems quite improbable especially for sensor networks with simple nodes. We will therefore assume that there is no carrier synchronization. The link between any pair of nodes \((i, j)\) can be parameterized by a complex channel gain \(h_{ij}\), which is assumed to be symmetric, that is, \(h_{ij} = h_{ji}\). The channel gains \(h_{ij}\) are independent random variables as a result of the random movement of nodes and (or) fading. They are assumed to be fixed during one-message transmission period and go to another independent realization in the next-message transmission period.

The source wants to send a message \(w\) to the destination during the duration of each channel realization \(\mathcal{H} = \{h_{ij} : i, j \in \{1, \ldots, N\}, i \neq j\}\). Let \(X_i(k)\), \(i \in \{1, \ldots, N-1\}\), be the channel input of node \(i\) at time \(k\) and let \(Y_j(k)\), \(j \in \{2, \ldots, N\}\), be the channel output of node \(j\) at time \(k\), we have

\[
Y_j(k) = \sum_{j \in \{1, \ldots, N-1\}, j \neq i} h_{ij}X_i(k) + Z_i(k), \quad i \in \{2, \ldots, N\},
\]

where \(Z_i(k) \sim \mathcal{CN}(0, 1)\) are i.i.d. unit power white Gaussian noises for all \(i, k\). We assume that full channel state information is available noncausally to all nodes. While this may not be realistic in fast-changing channels, it is possible if the channel is not varying too quickly. Furthermore, this gives a bound on performance as for the case when less knowledge is available.

3. ACHIEVABLE RATES OF THE GAUSSIAN MULTIPLE-RELAY CHANNEL

In [5], Gupta and Kumar demonstrated an achievable region for a multiple-relay channel, and later Xie and Kumar [16] established an explicit formula for the achievable rate, which, in general, exceeds the rate in [5]. Here we restate the theorem in [16] as follows.

**Theorem 1** (see [16, Theorem 3.4]). For a discrete memoryless multiple-relay channel with source node 1, destination node \(N\), and the other nodes arranged into \(L - 1\) levels with each level \(k\) consisting of a set of nodes \(\Gamma_k\), \(k = 1, \ldots, L - 1\), the following rate is achievable:

\[
R < \max_{P(X_0, \ldots, X_{N-1})} \min_{1 \leq k \leq L} \min_{\iota \in \Gamma_k} I(X_0, \ldots, X_{k-1} ; Y_{\iota} | X_k, \ldots, X_L),
\]

(6)

where boldface characters denote vectors for inputs of the nodes in each group. Here \(\Gamma_0 := \{1\}\) and \(\Gamma_L := \{N\}\).

For an asynchronous Gaussian multiple-relay channel, we have the following corollary of Theorem 1.

**Proposition 1.** Assume that node \(j\) uses transmission power \(P_j\). For an asynchronous Gaussian multiple-relay channel with \(L - 1\) levels of relay nodes, the following rate is achievable:

\[
R \leq \min_{1 \leq l \leq L} \min_{\iota \in \Gamma_l} \frac{1}{2 \log (1 + \sum_{m=0}^{l-1} \sum_{\iota \in \Gamma_m} P_j | h_{ij} |^2)}.
\]

(7)

**Proof.** The message \(w\) is first split into \(B\) blocks \(w_1, \ldots, w_B\) of \(nR\) bits each. Each node \(i\) generates a codebook with \(2^{nR}\) i.i.d. \(n\)-sequences with i.i.d. Gaussian components and index them as \(x_i(w_i)\), \(w_i \in \{1, \ldots, 2^{nR}\}\). The whole transmission is performed in \(B + L - 1\) time slots, and thus the overall rate is \(R \cdot B/(B + L - 1)\) bits per channel use. By making \(B\) large, we can get the rate arbitrarily close to \(R\). In each of the first \(B\) time slots, the source node 1 transmits the codeword \(x_1(w_1)\) for each \(w_1\), \(i \in \{1, \ldots, B\}\), and in the remaining time slots, it transmits constant signals \(x_1(1)\). A node \(i\) in level \(k\), \(1 \leq k \leq L\), starts the decoding of \(w_i\) at the end of \(k\)th time slot and sends out \(x_i(w_i)\) in time slot \(k + 1\). It continues the same decoding and encoding procedure in each time slot thereafter until it has decoded and sent out all the messages. It transmits some constant signals \(x_1(1)\) in the remaining slots. To illustrate the encoding scheme, we give an example of a relay channel of 5 nodes, in which (1, 5) is the source-destination pair. Nodes 2 and 3 are assigned to level 1 and node 4 is in level 2. The message \(w\) is split into 6 message blocks. The encoding scheme is shown in Figure 1.

The relays and the destination decode each \(w_i\), \(i \in \{1, \ldots, B\}\), using similar sliding-window decoding technique.
A node $i$ in level $l$ can decode $w_1$ at the end of $l$th time slot using a window of the first $l$ received blocks. After decoding the first message, the window is shifted by one and the part due to the transmission of the first message is subtracted from the received signals in the new window and then the second message is decoded. It continues until all messages are decoded. For each message, node $i$ is actually receiving information from $l$ independent parallel channels [25]. Thus for node $i$ to successfully decode the message, we have

$$R < \max_{P(X_0, X_1, \ldots, X_{l-1})} I(X_0, X_1, \ldots, X_l; Y_1 | X_1, \ldots, X_{l-1})$$

$$\leq \frac{1}{2} \log \left( 1 + \sum_{m=0}^{l-1} \sum_{j \in \Gamma_m} P_j |h_{ij}|^2 \right),$$

where $P_j$ is the power assigned to node $j$. Since each node except for the source needs to fully decode each message block, we have

$$R \leq \min_{1 \leq l \leq L} \min_{\epsilon_l \in \Gamma_l} \frac{1}{2} \log \left( 1 + \sum_{m=0}^{l-1} \sum_{j \in \Gamma_m} P_j |h_{ij}|^2 \right).$$

Here for simplification of notation, we assume that

$$\sum_{k=0}^{l-1} \sum_{j \in \Gamma_k} P_j |h_{ij}|^2 = 0, \quad \forall i.$$  \hfill (11)

**Remark 1.** Note that we do not introduce any correlation between the inputs of the nodes as it will not produce any gain if no carrier-level synchronization between transmitters is available [6, 7].

**Remark 2.** To achieve the rate in (7), all $X_i$'s are Gaussian distributed and mutually independent.

**Remark 3.** As can be seen from (9), the interference from other nodes is effectively cancelled out after a node subtracts from its received signals the part contributed from the messages it knows. From (9), we obtain an equivalent form of (7).

**Corollary 1.** Fix a rate $R$ and define $d_{ij} = (2^R - 1)/|h_{ij}|^2$. For the rate $R$ to be achievable, the powers $P_1, P_2, \ldots, P_N$ have to satisfy

$$\sum_{m=0}^{l-1} \sum_{j \in \Gamma_m} P_j \frac{d_{ij}}{d_j} \geq 1$$

for all $l$.

### 3.1. The optimal multilevel structure and power allocation policy

As we have shown, in order to minimize the outage probability in a quasi-static channel, we need to find a multilevel structure $S$ and a corresponding power allocation policy $T(S)$ such that the total power to achieve the rate requirement $R$ is minimized. Assuming that a multilevel structure $S$ has $L + 1$ levels, we denote the nodes in each level $0 \leq l \leq L$ by $\Gamma_l$ and the size of $\Gamma_l$ by $|\Gamma_l|$. We have $\Gamma_0 = \{1\}$ and $\Gamma_L = \{N\}$. Denote the level of a node $i$ as $\ell(i)$. Note that $S$ may not include all the nodes, that is, some nodes may be chosen not to participate in the transmission. Denote the power assigned to node $i \in S$ by a power allocation policy $T(S)$ for $S$ as $P_i(T, S)$. We then need to solve the following optimization problem:

$$\min_{T(S)} \sum_{i \in S} P_i(T, S) \text{ such that } R$$

$$\leq \min_{1 \leq l \leq L} \min_{\epsilon_l \in \Gamma_l} \frac{1}{2} \log \left( 1 + \sum_{m=0}^{l-1} \sum_{j \in \Gamma_m} P_j(T, S) |h_{ij}|^2 \right).$$  \hfill (13)

Since a Gaussian multiple-relay channel in general is not a degraded channel as the one studied in [17], it does not have a natural arrangement of nodes that is optimal. However, it does have some special properties for an optimal multilevel structure $S$ and its corresponding optimal power allocation policy $T(S)$ as stated in the next two theorems.

**Theorem 2.** For any channel realization $\mathcal{H}$ and rate requirement $R$, the overall power allocation is minimized by a sequential-path multilevel structure $S$, that is, one with $|\Gamma_l| = 1$ for all $l$.

**Proof.** We need to show the existence of a sequential path $P$ that is optimal. For any channel realization $\mathcal{H}$, there always exist a multilevel structure $S$ and a corresponding power allocation policy $T(S)$ that are optimal. Assuming that $S$ has $L + 1$ levels, we prove by induction that it can always be converted to an equivalent path $P$ without increasing total power consumption by properly removing some nodes in $S$ and adjusting transmission power of the remaining nodes.

First, for level $L$, $\Gamma_L = \{N\}$, thus $|\Gamma_L| = 1$. Suppose that for decoding orders $l \geq T + 1$, we have $|\Gamma_l| = 1$. We will then show that we can always make $|\Gamma_T| = 1$ without violating the constraints. For convenience of presentation, we denote the only node in $\Gamma_l, l \geq T + 1$, by $\zeta_l$.

If $|\Gamma_T| = 1$, we are done. Otherwise, assume that $|\Gamma_T| = M (M \geq 2)$ and $\Gamma_T = \{\zeta_m : 1 \leq m \leq M\}$. Without risk of confusion, we simplify the notation of $P_i(T, S)$ to $P_i$ and we have $P_i > 0$, for all $i \in S$. [End of proof]
We consider two cases.

**Case 1.** There exists \( \{t_1, t_2\} \in \Gamma_T \) such that \( d_{t_1, t_2}^* > 0 \), \( i = 1, 2 \).

We perform the following recursive power updating procedure.

1. Fix the transmission power of all nodes that reside in level \( T \) or higher except for \( t_1 \) and \( t_2 \). Adjust the transmission power \( P_{t_1} \) of \( t_1 \) to \( P_{t_1}^{\text{new}} = P_{t_1} + \delta \), where \( \delta \) is a small value.
2. Adjust the transmission power \( P_{t_2} \) of \( t_2 \) such that the left-hand side of the constraint (12) for \( \zeta_{T+1} \) remains unchanged. Therefore, we have \( P_{t_2}^{\text{new}} = P_{t_2} - \frac{d_{t_2, t_2}^*}{d_{t_1, t_2}^*} \delta \).
3. Adjust the transmission power of \( \zeta_{T+1} \) such that the left-hand side of constraint (12) is kept the same for \( \zeta_{T+2} \) to get
   \[
   P_{\zeta_{T+1}}^{\text{new}} = P_{\zeta_{T+1}} + \left( \frac{d_{t_1, t_2}}{d_{t_1, t_2}} d_{t_2, t_2}^* - \frac{d_{t_2, t_2}^*}{d_{t_1, t_2}^*} \delta \right). \tag{14}
   \]
4. Recursively update the transmission power of node \( i, i = \zeta_{T+2}, \ldots, \zeta_{L-1} \), such that the left-hand side of the constraint (12) is kept the same for the node right behind it.

This recursive updating procedure guarantees that the constraint (12) is still satisfied at all relay nodes and the destination. Since we vary the transmission power of only one node at each step, the total amount of power change is proportional to \( \delta \). Denote the total transmission power for the multilevel structure \( S \) and the corresponding power allocation policy \( T(S) \) as \( \xi(S, T(S)) \), that is, \( \xi(S, T(S)) = \sum_{i \in S} P_i(T, S) \).

Then
   \[
   \xi(S, T^{\text{new}}(S)) = \xi(S, T(S)) + f(S)\delta, \tag{15}
   \]
where \( T^{\text{new}} \) is the new power allocation policy after the power updating procedure and \( f(S) \) is a constant that does not depend on \( \delta \) but only on the multilevel structure \( S \) if \( |\delta| \) is small enough. Obviously, \( \delta \) is allowed to be either positive or negative, that is, we can either increase or decrease the transmission power of \( t_1 \). Thus, if \( f(S) \neq 0 \), we can always choose the sign of \( \delta \) such that the total amount of power change \( f(S)\delta < 0 \), and hence
   \[
   \xi(S, T^{\text{new}}(S)) < \xi(S, T(S)). \tag{16}
   \]

This contradicts the fact that the original multilevel structure and power allocation policy pair \( (S, T(S)) \) is optimal. Therefore we must have \( f(S) = 0 \), and thus \( (S, T^{\text{new}}(S)) \) is also optimal. In this case, we can repeatedly perform the same updating procedure by decreasing the transmission power of \( t_1 \) (or \( t_2 \)) and increasing the transmission power of \( t_2 \) (or \( t_1 \)) until either the transmission power of node \( i, P_{t_i}^{\text{new}} = 0, i \in \{ \zeta_{T+1}, \ldots, \zeta_{L-1} \} \) or \( P_{t_i}^{\text{new}} = 0, i = 1, 2 \). If \( P_{t_i}^{\text{new}} = 0, i \in \{ \zeta_{T+1}, \ldots, \zeta_{L-1} \} \), then node \( i \) can be removed from the relaying structure and we can continue the updating procedure above. If \( P_{t_i} = 0, i = 1, 2 \), it means that we can remove \( t_i \) from the structure \( S \).

If there still exist two or more nodes with decoding order \( T \), we can always take out two of them and repeat the same procedure above to remove one node each time until only one node is kept.

**Case 2.** \( d_{t_i, t_i}^{*} = 0 \), for all \( t_i \in \Gamma_T \setminus \{t_1\} \), and \( d_{t_1, t_2}^{*} > 0 \).

In this case, there is only one node \( t_1 \) in level \( T \) that has finite-length link to node \( \zeta_{T+1} \). This case is actually essentially the same as in Case 1. Pick a node \( t_2 \) in \( \Gamma_T, t_2 \neq t_1 \), and a node \( \zeta_{T+1} \in \{ \zeta_{T+2}, \ldots, \zeta_L \} \) such that \( \ell(\zeta_{T+1}) < \ell(k), \) for all \( k \in \{ \zeta_{T+2}, \ldots, \zeta_L \}, k \neq \zeta_{T+1}, \) and \( d_{t_2, t_i}^* > 0 \). Then we can perform the same recursive power updating algorithm as in Case 1. The only difference is that node \( \zeta_{T+1} \) takes the place of \( t_1 \) in Case 1. Thus we can always reduce the power of \( t_2 \) to 0 and thus remove it from the multilevel relaying structure.

Combining our discussions of Cases 1 and 2, we can conclude that we are always able to keep only one node at decoding order \( T \) without increasing the total power consumption. By induction, for all \( 1 \leq l \leq L \), we may have \( |\Gamma_l| = 1 \) and this establishes the proof.

Note that the new relaying path \( P \) does not necessarily have the same number of levels as \( S \).

The implication of Theorem 2 is that we can restrict our search to sequential paths without loss of optimality. In doing so, we greatly reduce the search space. The following theorem shows how power is optimally allocated given a sequential relaying path.

**Theorem 3.** For a sequential relaying path \( P \), the optimal power allocation policy \( T(P) \) can be implemented by a recursive power-filling procedure, that is, along path \( P \), starting from the source, each node \( i \) adjusts its transmission power such that the constraint (12) is satisfied with equality sign at its immediate successor \( j, \ell(j) = \ell(i) + 1 \).

**Proof.** Let the relaying path be \( P = (\zeta_0, \zeta_1, \ldots, \zeta_L) \), where \( \zeta_0 = 1, \zeta_L = N \). Initially we set the power of all nodes to 0. Since node \( \zeta_i \) only receives information from the source \( \zeta_0 \), we must let the source transmit at a power level such that constraint (12) (12) is exactly satisfied at node \( \zeta_i \). Now the message is known to \( \zeta_0 \) and \( \zeta_1 \) and only they are eligible to transmit. With the objective to save transmission power, at any time we always let the node whose transmission is most efficient (results in less total transmission power) increase its transmission power. Now the transmission of node \( \zeta_i \) will be more efficient. Otherwise, if the transmission of \( \zeta_0 \) is more efficient, it will increase its transmission power until a node other than node \( \zeta_i \) satisfies constraint (12). That node can then decode in the same decoding order as node \( \zeta_i \) and it contradicts the fact that there is only one node in each level. Thus the source has to stop increasing its transmission power as long as node \( \zeta_i \) satisfies (12). Node \( \zeta_i \) then adjusts its power level such that \( \zeta_i \) satisfies constraint (12) with equality sign. This procedure proceeds until the destination meets condition (12) exactly.

Here we do not need to know how to exactly determine the efficiency of the transmission of a particular node. What
we only need to know is that it depends on the structure of the relaying path and the state of the relaying path, that is, whether constraint (12) is satisfied at the nodes in the path or not. Therefore, before the state of the relaying path changes, the transmission efficiency of any node that has satisfied (12) remains unchanged. Theorem 3 implies that every node except for the destination transmits with certain level of positive power and every node except for the source receives exactly enough information from its upstream nodes.

3.2. Example

Now we give a simple example to illustrate the benefit of cooperative relay signaling. Figure 2 shows a multiple-relay channel network with 4 nodes in which (1, 4) is the source-destination pair. The label attached to the link (i, j) is the value $d_{ij}$ as defined before. All 4 possible sequential relaying paths and their corresponding total power consumption are presented in Table 1. The path $1 \rightarrow 3 \rightarrow 2 \rightarrow 4$ is not an eligible relaying path as by the power allocation policy, node 2 cannot decode after node 3. The total power consumption is calculated using the recursive power-filling procedure. For example, for the path $1 \rightarrow 2 \rightarrow 4$, in order to make node 2 able to decode, we have $P_1 = 10$. To make node 4 able to decode, we have $P_1/2 + P_2/30 = 1$, and thus $P_2 \approx 22.86$. The overall power consumption is then $P_1 + P_2 = 32.86$. A traditional multihop operation that uses the shortest path algorithms will find $1 \rightarrow 2 \rightarrow 4$ as the optimal path with overall power consumption 40. However, the transmission from node 1 to node 2 will give rise to interference to the communications between node 2 and node 4. Therefore, the actual power consumption will be larger than 40. From Table 1, it is interesting to see that the best relaying path $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$ is the worst one from the point of view of traditional multihopping algorithms.

4. HEURISTIC ALGORITHMS

From Theorems 2 and 3, we have shown that for any channel realization $\mathcal{H}$, there exist an optimal relaying path $P$ and a corresponding simple power allocation policy $T(P)$. Thus limiting our search to sequential paths can greatly reduce the search space for optimal solutions. There have been some elegant shortest path algorithms to find a shortest path in a network [26]. However, the Bellman principle used in these traditional shortest path algorithms is not satisfied here. For example, consider a relay network with 4 nodes $V = \{1, 2, 3, 4\}$ and costs $d_{21} = 3, d_{32} = 4, d_{31} = 6, d_{41} = 7, d_{42} = 12, d_{43} = 0.1$. We may verify that the optimal relaying path is $1 \rightarrow 3 \rightarrow 4$. By the Bellman principle, the optimal cooperative relaying path from 1 to 3 should be the direct link from 1 to 3, which requires a total power consumption of 34. However, from 1 to 3 we can find that the path $1 \rightarrow 2 \rightarrow 3$ actually requires a smaller total power consumption of $(10 + (34 - 10)/34 \times 25) \approx 27.65$. This shows that the Bellman principle does not apply to the cooperative routing problem.

Another difference between the optimal relaying path problem in this paper and the traditional shortest path problem is that in the former we have to use a node-based metric instead of a link-based metric since we want to minimize the total power consumption of all nodes. Therefore, we cannot expect using standard shortest path algorithms to find an optimal relaying path. An exhaustive search algorithm that searches through all multilevel structures has a complexity of $O((N-2)^{N-2})$. Theorem 2 reduces this complexity to $O((N-1)!$. We may improve on this using the property of an optimal relaying path in Theorem 3 to remove many unqualified candidates. As implied in Theorem 3, when selecting the node for a particular level, it is not necessary to consider those nodes that have already satisfied condition (12). Otherwise, they will receive more information than necessary. This reduces the worst case complexity to $2^{N-2}$ candidate paths, which makes it possible to find the optimum solution for small networks (i.e., less than 20 nodes). Still, for larger networks, the complexity is too high. We therefore consider heuristic algorithms for finding relaying paths and the corresponding power allocation policies for general multiple-relay channels. The algorithms provide achievable rates which might not be optimal for the given coding scheme, but simulation results show that one of the heuristic algorithms is essentially equal to the optimum solutions for small networks where the optimal solution can be found. Furthermore, the heuristic algorithms provide prototype algorithms for practical (central) implementation of relay-channel signaling.

The following heuristic algorithms are based on Theorems 2 and 3. From Theorem 2, although it is still difficult to find an optimal path, we may try to search for a path that is close to optimum. We then enforce the optimal power allocation policy in Theorem 3 on the path selected.
4.1. Heuristic algorithm 1: CTNCR

A traditional noncooperative multihopping algorithm finds a shortest path assuming no interference from upstream nodes and, in general, it generates a suboptimal path. However, it might be a starting point for finding a good relay-signaling cooperative path. In this heuristic, we first find a shortest noncooperative path using standard Dijkstra’s algorithm based on the link-based metric and then use the power allocation policy in Theorem 3 to determine the overall power consumption and possibly remove some nodes from the path. The algorithm works as follows.

Step 1 (initialization). Find a noncooperative path $P$ using Dijkstra’s algorithm. Set the transmission power of all nodes in $P$ to 0. Set the source as the active node, which is the only one that can adjust transmission power.

Step 2. Among the active nodes’ downstream nodes that have not satisfied (12), find node $K$ such that it requires the least transmission power of the active node to decode the message (satisfying condition (12)). Remove the nodes between the active node and $K$ from $P$. Set $K$ as the active node.

Step 3 (stop criterion). Stop if $K$ is the destination and the new $P$ is the final path with the transmission power of nodes as determined in Step 2; otherwise go to Step 2.

The computational complexity of Dijkstra’s algorithm is $O(N^2)$ [26]. In Step 2, we note that there are $|P| - 1$ iterations and the number of operations in each iteration is proportional to $|P|$. Therefore in the worst case, the computation in Step 2 is $O(|P|^2)$. Thus the computation of CTNCR is $O(N^2 + |P|^2)$. Since $|P| \leq N$, in the worst case, the computation of CTNCR is $O(N^2)$.

4.2. Heuristic algorithm 2: SNER

This heuristic algorithm is essentially a greedy algorithm similar to the Prim-Dijkstra spanning-tree algorithm but it stops whenever the destination is included in the tree. The algorithm works as follows.

Step 4 (initialization). Form a set of nodes $\Xi_d$, which is called the decoded set, with only the source node included and a nondecoded set $\Xi_n = V - \Xi_d$, where $V$ is the set of all nodes.

Step 5. For each node $K \in \Xi_n$, find a node $T$ in $\Xi_d$ as its predecessor that requires the least total power consumption for $K$ to satisfy (12) using the recursive power-filling procedure. Record the path and the corresponding overall power allocation for $K$ to satisfy (12). Among all $K \in \Xi_n$, find the node that requires the least overall power, denote it by $K_{min}$. Add $K_{min}$ to $\Xi_d$ and remove it from $\Xi_n$.

Step 6 (stop criterion). If $K_{min}$ is the destination, stop; otherwise, go to Step 5.

To estimate the computation required by SNER algorithm, we note that in the worst case there are $N - 1$ iterations. In each iteration, for each node $K \in \Xi_n$, we need to do $N - |\Xi_n|$ comparisons. Hence in each iteration, the computation is $|\Xi_n|(N - |\Xi_n|)$. In the worst case, the computation of SNER is $\sum_{i=1}^{N-1} i(N - i) = N^3/6 - N^2 + N/6$. Thus the computation complexity of SNER is $O(N^3)$. However, since we only need $|P| - 1$ iterations, the actual computation of SNER is then $\sum_{i=1}^{|P|-1} i(N - i) = (N|P|^2 - N|P|)/2 - (2|P|^3 - 3|P|^2 + |P|)/6$. Since $N > |P|$, the computational complexity of SNER algorithm is $O(N|P|^2)$.

5. Complexity-Constrained Networks

In our previous discussion, every node is assumed to be able to store and process all related received signals to decode a message. In some applications, the relays may have only limited memory and signal processing capability, and thus cannot combine all these signals, especially if the path is long. On the other hand, the signals received from remote upstream nodes bring insignificant information or interference to the decoding of the message and it may not pay off to include these signals in the decoding of the message. Therefore we may treat them as pure noise with possibly only a slight increase of the overall power consumption. We hence consider a variation of decode-forward relaying path problem by adding a constraint that the relays and the destination decode each message only based on the most current received signals. The encoding scheme is the same as in Section 2. The difference lies in the decoding of the relays and the destination in that the sizes of their decoding windows are at most $F$. Notice that the relays in level $i, i \leq F$, can use all the related received signals. We still assume that a node can subtract all interferences from downstream nodes. Since a node has already decoded the message downstream nodes are transmitting, it also knows precisely what signal downstream nodes are transmitting. This interference subtraction is much less complex than the joint decoding required to handle the signal transmitted by upstream nodes, so the algorithm is complexity constrained. However, in practice, the complexity could be reduced more by only subtracting the signal from the first few nodes downstream.

Again using the parallel channels argument [25], for node $i$ with decoding order $l \geq 1$ in a path $P$ to decode a message at rate $R$, we have

$$R \leq \sum_{m=1}^{l-1} \log \left( 1 + \frac{P_m |h_{lm}|^2}{1 + \sum_{k=0}^{m-1} P_k |h_{lk}|^2} \right)$$

(17)

where $|x| = \max(0, x)$ and again for notation simplification, we assume that $\sum_{k=0}^{l-1} P_k |h_{lk}|^2 = 0$. Notice that there is no interference from downstream nodes in (17) in accordance with the assumption of interference subtraction for downstream nodes.

To reduce the complexity of signal processing at the relays and the destination, it is always desirable to keep $F$ small. On the other hand, to be more power efficient, it is desirable
to choose a larger $F$. Therefore, there is a tradeoff in properly selecting the value of $F$. Again, as in the unlimited signal processing case, any optimal multilevel relaying structure can be converted to a relaying path without increasing power consumption.

**Theorem 4.** For any channel realization $\mathcal{H}$, rate requirement $R$, and signal processing length $F$, there always exist a sequential path $P$ and a corresponding power allocation policy $T(P)$ that minimize overall power consumption.

**Proof.** The proof is essentially the same as for Theorem 2. The only difference is that $f(S)$ is changed to $f(S, T(S))$, that is, it also depends on the original power allocation policy.

Similarly, the optimal power allocation policy $T(P)$ for any limited data processing path $P$ is still the recursive power-filling procedure as before.

**Theorem 5.** For a sequential relaying path $P$ with limited signal processing capability, the optimal power allocation policy can be implemented by a recursive power-filling procedure as stated in Theorem 3.

The proof is similar to the proof of Theorem 3.

The two heuristic algorithms CTNCR and SNER can be easily adapted to the limited signal processing capability case. Here we only consider the variation of SNER algorithm and we denote the SNER algorithm with signal processing length $L$ as SNERvL.

## 6. NUMERICAL RESULTS

In this section, we illustrate the performance of the relay-channel signaling by simulation. Since our results only depend on the amplitude of channel gains $h_{ij}$, we consider only the theoretical model of $|h_{ij}|$, the model of which we use in our simulations is

$$|h_{ij}| = \frac{\alpha_{ij}}{(\ell_{ij})^{n/2}}, \quad (18)$$

where $\ell_{ij}$ is the distance between $i$ and $j$, $n$ is the path loss exponent, and $\alpha_{ij}$ is a constant or a random variable. We consider two cases.

1. $\alpha_{ij} = 1$, for all $i, j$. In this case, a signal is attenuated only by path loss. The randomness of the channel realization comes from the random movement of the nodes.

2. $\alpha_{ij}$ is a unit-variance Rayleigh distributed random variable. A signal is then attenuated not only by path loss but also by small scale fading characterized by the parameters $\alpha_{ij}$. All $\alpha_{ij}$’s are assumed to be mutually independent.

A typical value of the path loss exponent $n$ is between 2 and 5. In our simulations, we consider the cases when $n = 2$, a low attenuation regime; and $n = 4$, a high attenuation regime.

To simulate the random movement of nodes, for each channel realization we randomly place all the nodes, in our simulations 20 or 50 nodes, in a $100 \times 100$ grid and randomly pick two of them as the source-destination pair. For $n = 2$, we consider a desired rate of either $R = 0.5$ or $R = 1$; and for $n = 4$ a desired rate of either $R = 0.5$ or $R = 2$. The results are based on 100000 simulation runs for each case. The noncooperative multihopping routes are found by the Bellman-Ford algorithm using the link-based metric. As in Section 5, we assume that nodes can subtract interference from all downstream nodes. Traditional multihopping systems most likely do not have this ability, and the curves for performance of noncooperative multihopping should therefore be seen as a lower bound for the performance of practical multihopping. Multihopping is therefore identical to SNERv1, except that SNERv1 uses an interference sensitive routing. The optimal solution for network size 20 is found by exhaustive search over all paths according to Theorems 2 and 3. The simulation results are presented in Figures 3, 4, 5, 6, 7, and 8, which show the outage performance of various algorithms under different total power constraints.

The first that can be noticed is that in all the 20 node cases, the heuristic optimization algorithm SNER gives a performance which is essentially identical to the optimal performance, while the less complex CTNCR has a performance slightly worse. We do not present the optimal solutions for network size 50 due to the overwhelming computational task, but based on the results for network size 50 we can expect SNER to be representative also of the optimal solution.

The second remarkable result is the qualitative difference between the low-rate case ($R = 0.5$) and the high-rate case ($R = 1$ or $R = 2$). In the low-rate case, the gain from cooperation is limited—at most 5 dB$^1$ for $n = 2$ and network size 50, and for the high-attenuation case $n = 4$, no gain at all. On the other hand, for high rate, the gain from cooperation is very large, up to 18 dB in Figure 5. Recall that the noncooperation curve is actually a lower bound for practical multihopping, so the gain could very well be even larger. This indicates that a main advantage of cooperation is interference avoidance, as interference increases with rate for traditional multihopping, while relay-channel signaling completely avoids interference. The results for $n = 4$ confirm the results in [16, 27, 28] that multihopping is a reasonable choice, but only in the high-attenuation/low-rate regime.

The results for SNERvL show that it is not necessary to use the full relay-channel signaling to get significant gains. In all cases considered, SNERv4 gets very close to the optimal relay-channel signaling, so that it would be enough to decode the transmission of the 4 “nearest” neighbors upstream.

## 7. CONCLUSIONS

In this paper, we show that the optimal operation of an asynchronous Gaussian multiple-relay channel with decode-forward signaling is given by a path with a corresponding simple power allocation policy. This reduces the complexity

---

$^1$ All dB gains discussed are for outage probability $10^{-3}$. 
of finding the optimal solution, although the complexity is still exponential. We therefore propose heuristic polynomial-time algorithms for path finding, and numerical results show that these heuristic algorithms give solutions very close to the optimal solution.

Our numerical results show that in the low-attenuation regime, both with and without Rayleigh fading, cooperation through relay-channel signaling shows significant gains over traditional noncooperative operation. The gains increase as the rate increases because of the interference explosion for a noncooperative algorithm. In the high-attenuation regime, however, for low rate, more traditional multihopping operation that uses single-signal-based decoding can be a quite reasonable choice as cooperation brings little gain. For high rate, the cooperative algorithms still show significant gain because of the poor performance of the traditional multihopping algorithm, which, however, may be greatly improved by carefully choosing paths to try to avoid heavy interference.
The heuristic algorithms developed here for calculating rate can be used as a starting point for developing practical routing algorithms for relay channels. In challenge, however, is the assumption of full network information at each node. This requirement can be mitigated considering further simplification to the proposed heuristic algorithm. For example, we may consider further simplification of SNERvL by finding the path using some rough channel state information, for example, the positions of nodes, and cancelling only the interference from the transmissions of the most immediate downstream nodes. In this case, a node only needs to know the positions of other nodes and the perfect channel gains between itself and its 2L closest nodes in the path selected. The heuristic algorithms can also be adapted to distributed (distance-vector or link-state-based) versions.

Another basic assumption is that nodes use full duplex. It will be interesting to extend the results to half-duplex case, which, however, is not trivial as it involves an additional complicated scheduling problem of time slots or frequency bands. Another interesting problem that we may consider in future work is the optimization problem when nodes have individual power constraints in addition to a global power constraint.

ACKNOWLEDGMENT

This work was supported in part by NSF Grant CCR03-29908.

REFERENCES

[1] E. C. van der Meulen, “Three-terminal communication channels,” Advances in Applied Probability, vol. 3, no. 1, pp. 120–154, 1971.
[2] E. C. van der Meulen, Transmission of information in a T- terminal discrete memoryless channel, Ph.D. thesis, Department of Statistics, University of California, Berkeley, Calif, USA, 1968.
[3] T. M. Cover and A. A. El Gamal, “Capacity theorems for the relay channel,” IEEE Transactions on Information Theory, vol. 25, no. 5, pp. 572–584, 1979.
[4] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: efficient protocols and outage behavior,” IEEE Transactions on Information Theory, vol. 50, no. 12, pp. 3062–3080, 2004.
[5] P. Gupta and P. R. Kumar, “Towards an information theory of large networks: an achievable rate region,” IEEE Transactions on Information Theory, vol. 49, no. 8, pp. 1877–1894, 2003.
[6] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” IEEE Transactions on Information Theory, vol. 51, no. 9, pp. 3037–3063, 2005.
[7] A. Høst-Madsen and J. Zhang, “Capacity bounds and power allocation for the wireless relay channel,” IEEE Transactions on Information Theory, vol. 51, no. 6, pp. 2020–2040, 2005.
[8] A. Høst-Madsen, “Capacity bounds for cooperative diversity,” IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1522–1544, 2006.
[9] B. Schein and R. Gallager, “The Gaussian parallel relay network,” in Proceedings of IEEE International Symposium on Information Theory, p. 22, Sorrento, Italy, June 2000.
[10] G. K. Karagiannidis, T. A. Tsiftsis, and R. K. Mallik, “Bounds for multihop relayed communications in Nakagami-m fading,” IEEE Transactions on Communications, vol. 54, no. 1, pp. 18–22, 2006.
[11] G. K. Karagiannidis, “Performance bounds of multihop wireless communications with blind relays over generalized fading channels,” IEEE Transactions on Wireless Communications, vol. 5, no. 2, pp. 498–503, 2006.
[12] K. Azarian, H. El Gamal, and P. Schniter, “On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels,” IEEE Transactions on Information Theory, vol. 51, no. 12, pp. 4152–4172, 2005.
A. Reznik, S. R. Kulkarni, and S. Verdú, “Degraded Gaussian L.-L. Xie and P. R. Kumar, “A network information theory for S. Cheng and Z. Xiong, “Successive refinement for the Wyner-D. Bertsekas and R. Gallager, Data Networks, Prentice-Hall, Englewood Cliffs, NJ, USA, 1987. F. Xue, L.-L. Xie, and P. R. Kumar, “The transport capacity of wireless networks over fading channels,” IEEE Transactions on Information Theory, vol. 51, no. 3, pp. 834–847, 2005.

Anders Høst-Madsen was born in Denmark in 1966. He received the M.S. degree in electrical engineering in 1990 and the Ph.D. degree in mathematics in 1993, both from the Technical University of Denmark. From 1993 to 1996, he was with Dan-tec Measurement Technology A/S, Copenhagen, Denmark, from 1996 to 1998, he was an Assistant Professor at Kwangju Institute of Science and Technology, Kwangju, Korea, and from 1998 to 2000, he was an Assistant Professor at Department of Electrical and Computer Engineering, University of California, Berkeley, through 1992. His research interests are in statistical signal processing, information theory, and wireless communications, including multiuser detection, equalization, ad hoc networks, and wireless sensor networks. He currently serves as the Editor for Multituser Communications for the IEEE Transactions on Communications and as an Associate Editor for Detection and Estimation for the IEEE Transactions on Information Theory.
Joint MIMO-OFDM and MAC Design for Broadband Multihop Ad Hoc Networks

Dandan Wang and Uf Tureli

Received 1 November 2005; Revised 4 June 2006; Accepted 13 June 2006

Multiple-input multiple-output (MIMO) and orthogonal frequency division multiplexing (OFDM) are very promising techniques to exploit spatial diversity and frequency diversity in the physical layer of broadband wireless communications. However, the application of these techniques to broadband multihop ad hoc networks is subject to inefficiencies since existing medium access control (MAC) schemes are designed to allow only one node to transmit in a neighborhood. Therefore, adding more relays to increase the transmission range decreases the throughput. With MIMO-OFDM, multiple transmissions can coexist in the same neighborhood. A new transceiver architecture with MIMO-OFDM and MAC scheme is proposed in this paper. The new MAC scheme, multiple-antennas receiver-initiated busy-tone medium access (MARI-BTMA), is based on receiver-initiated busy-tone medium access (RI-BTMA) and uses multiple out of band busy tones to avoid the collision of nodes on the same channel. With the proposed MAC scheme, multiple users can transmit simultaneously in the same neighborhood. Although basic MARI-BTMA shows good performance at high traffic load, to improve the performance at low traffic loads, 1-persistent MARI-BTMA is proposed so that users can choose different MAC scheme according to the statistical traffic load in the system. In this paper, both theoretical and numerical analysis of the throughput and delay are presented. Analysis and simulation results show the improved performance of MARI-BTMA compared with RI-BTMA and carrier sensing medium access/collision avoidance (CSMA/CA).

Copyright © 2006 D. Wang and U. Tureli. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

In the recent years, multihop relaying ad hoc networking has attracted a lot of interest for its flexibility to achieve broad coverage without any infrastructure. Many new techniques have been adopted in ad hoc networks to improve the performance in the physical layer, that is, multiple-input multiple-output (MIMO) and orthogonal frequency division multiplexing (OFDM). MIMO systems take advantage of the spatial diversity obtained by spatially separated antennas in a multipath scattering environment. Several different ways can be used to obtain either a diversity gain to combat signal fading or a capacity gain, that is, space-time coding (STBC) [1], vertical Bell laboratories layered space-time (V-BLAST) [2], and singular value decomposition (SVD) diversity [3]. Thus, MIMO techniques have shown a great potential to improve the capacity of the system in the physical layer [4, 5]. On the other hand, OFDM has become a popular technique for transmission of signals over broadband wireless channels since it provides a very high spectral efficiency, combats multipath fading, and can be simply implemented by fast Fourier transform (FFT) with a low receiver complexity. OFDM has been adopted in several wireless standards such as IEEE 802.11a wireless local area network (WLAN) standard [6] and IEEE 802.16a [7]. For high data-rate transmission, the multipath characteristic of the environment results in a frequency selective MIMO channel, OFDM can transform such a frequency-selective MIMO channel into a set of parallel frequency-flat MIMO channels, and therefore decrease receiver complexity. The combination of MIMO and OFDM can provide higher data rate, reduce the fading of a single link with space-time-frequency codes, mitigate interference by using extra spatial degrees of freedom, and allow simultaneous communication with different nodes using combinations of spatial multiplexing and interference cancellation [7–10].

In this paper, we propose a new transceiver architecture with the capabilities of signal separation and interference cancellation of MIMO-OFDM, in a virtual MIMO scheme combined with OFDM and space-time coding at the transmit nodes. At the receive nodes, multiple antennas are used to separate the independent data flows from different transmit nodes.
This new transceiver architecture allows multiple independent data flows to be transmitted on the same channel simultaneously, which provides the system the capability of multipacket reception (MPR). MPR presents new challenges for medium access control (MAC) in wireless networks since classical MAC schemes are designed to allow only one user in a neighborhood [6, 11–14]. Currently, most research works on MAC schemes with MPR focus on central controlled systems, for example, [15]. However, multihop ad hoc networks lack the aid of central controllers. In the literature, there are several MAC schemes proposed for distributed ad hoc networks. Stream controlled multiple access (SCMA) proposed in [16] can optimize the selection of the streams at the transmit nodes. However, SCMA requires a lot of information exchange between the different nodes. In [17], mitigating interference using multiple antennas MAC (MIMA-MAC) is proposed to mitigate the interference from the neighboring nodes. In the simulation analysis of [17], fairness and throughput are shown improved over the traditional carrier sensing medium access/collision avoidance (CSMA/CA). However, MIMA-MAC inherits the exposed node problem and hidden node problem associated with CSMA/CA [13, 18, 19]. In [19], Tobagi and Kleinrock proposed a busy-tone multiple access (BTMA) to alleviate the hidden problem in a network with a base station. When a base station senses the transmission of a terminal, it broadcasts a busy-tone signal to all terminals, keeping them (except the current transmitter) from accessing the channel. Based on BTMA, Wu and Li proposed the receiver-initiated busy-tone multiple access scheme (RI-BTMA) in [11] for ad hoc networks. The total spectrum resource is divided into two subbands. One is used to transmit busy-tone signals while the other is used to transmit data. The busy tone is used to acknowledge the channel access request and to prevent transmissions from other nodes. It solves the hidden node problem and the exposed node problem. In this paper, based on RI-BTMA, we propose a new MAC protocol—multiple-antennas receiver-initiated busy-tone multiple access (MARI-BTMA). MARI-BTMA utilizes multiple busy tones to notify the other nodes of the number of transmissions currently ongoing in the system. To improve the performance at the low traffic load, we also propose 1-persistent MARI-BTMA in this paper. An adaptive scheme is introduced based on the traffic load. Using OFDM in the transceiver architecture, subbands of OFDM signals can be used to transmit busy tones. Therefore, the overhead of busy tones is proportionally small. Performance analysis and simulation results show much better performance than RI-BTMA and CSMA/CA.

The paper is organized as follows. In Section 2, the new transceiver architecture is given. The proposed MAC scheme MARI-BTMA is presented in Section 3. Throughput and delay analysis of MARI-BTMA is given in Section 4. In Section 5, simulation results are given. Conclusions are drawn in Section 6.

2. TRANSCEIVER ARCHITECTURE WITH MIMO-OFDM

In this section, the new transceiver architecture with MIMO-OFDM is presented. Suppose there are six nodes in a network shown in Figure 1. Node 3 is the relay node of nodes 1 and 2. The destination of nodes 1 and 2 is node 6. Node 4 has data to send to node 5. In the physical layer, MIMO and OFDM is used to separate signals and cancel interference. Thus, nodes 1 and 2 can transmit to node 3 at the same time by signal separation. The transmission from node 3 to node 6 and the transmission from node 4 to node 5 can also be done simultaneously by interference cancellation. The transceiver architecture in the physical layer with MIMO-OFDM is shown in Figure 2. Suppose each node has $n_a$ antennas. A single space-time (or space-frequency) encoder is employed on these $n_a$ antennas. The space-time encoder takes a single stream of binary input data and transforms it into $n_a$ parallel streams of baseband constellation symbols. Each stream is broken into OFDM blocks. Each OFDM block of constellation symbols is transformed using an inverse fast Fourier transform (IFFT) and transmitted by the antenna for its corresponding stream. Thus, all $n_a$ transmit antennas simultaneously transmit the transformed symbols. At receive nodes, the received signals at each antenna are similarly broken into blocks and processed using an FFT. Then, an interference cancellation scheme is implemented by a space-time processor. The interference cancellation scheme attempts to separate the received signal due to one of the space-time encoders from the received signal due to the other space-time encoder. After this cancellation, maximum-likelihood sequence estimation (MLSE) decoding is employed, followed by successive interference cancellation. The detailed algorithm can be found in [8, 9]. All these algorithms need perfect synchronization. To recover the data flow from independent nodes, the number of transmit nodes must be less than or equal to the number of receive antennas at the receiver nodes. In this paper, we assume that there are two antennas at each node without loss of generality to more than two antennas.

3. MARI-BTMA

In the previous section, multiples nodes were allowed to transmit at the same time thanks to the advanced transceiver architecture in the physical layer. In this section, our proposed MAC protocol—MARI-BTMA—is presented. MARI-BTMA is designed based on the conventional RI-BTMA [11]. In RI-BTMA, the available frequency is divided into two
parts: control channel and data channel. Busy-tone signals are transmitted on the control channel while data is transmitted on the data channel. When a node has data to transmit, it will first sense the busy-tone channel. If the busy-tone channel is free, a packet of preamble containing the identification of the destination nodes will be sent. Once the preamble is received correctly by the intended receiver, the receiver sets up an out-of-band busy tone and waits for the data packet. The transmitter, upon sensing the busy tone, sends the data packet to the destination. It can be seen that RI-BTMA is designed to accept one user in a neighborhood. To access more than one user, we design a multiple-busy-tone scheme-MARI-BTMA. In MARI-BTMA, the total spectrum resource is divided into several control channels and one data channel. The number of control channels is equal to the number of busy-tone signals and the number of nodes transmitted simultaneously in the system. Since we assume there are two antennas in each node and two independent data flows to separate, two control channels are used in the following. Packet transmissions occur in a frame fashion. The structure of MARI-BTMA is shown in Figure 3. One MARI-BTMA frame is divided into two subframes. One sub frame (contention period) is used to transmit preambles to access the system and the other (contention-free period) is used to transmit data. In the contention period, similar to 802.11 MAC protocol and MIMA MAC in [17], a back off scheme is used to avoid the collision of the preambles sent by more than two nodes in a highly loaded system. The contention period is divided into minislots. The length of each minislot depends on the transmission time and detection delay of the busy tones. The larger size of contention period will reduce the probability of collision of the preambles, while the overhead will be higher. Therefore, the optimal length of contention period should achieve a balance. In the following, we give a detailed description of MARI-BTMA. Since the throughput is not stable when the traffic load is very low shown later in the throughput analysis, we also propose 1-persistent MARI-BTMA and an adaptive MARI-BTMA.

3.1. Basic MARI-BTMA

In the basic MARI-BTMA, only the nodes with data to transmit at the beginning of a frame contend to access the system. A node generating data in the middle of a frame has to wait a random interval till it is scheduled to transmit at the beginning of a frame. Then all the nodes with data to transmit at the beginning of a frame select one minislot in the contention period uniformly and sense the control channels.

(i) If a node senses two busy tones, it will not transmit a preamble in this frame and wait a random interval till it is scheduled to transmit at the beginning of a frame.

(ii) If a node senses one busy tone, it will transmit a preamble. If the preamble is successfully received by the intended receiver, the receiver will set up a busy tone on another free control channel.
(iii) If a node senses no busy tone, it will send a preamble and the receiver will set up a busy tone on either of the control channels once it receives the preamble correctly.

Only the nodes receiving the busy tone from their destination nodes can transmit data in the contention-free period. By sensing the number of busy tones, all the destination nodes get to know how many independent data flows to recover.

3.2. 1-persistent MARI-BTMA

It can be seen that when the traffic load is low, the basic protocol does not work well since all the packets generated during a frame have been ignored. According to [14], when the traffic load is very low, 1-persistent CSMA can improve the throughput of the system greatly. Similar to 1-persistent CSMA, we propose 1-persistent MARI-BTMA in this subsection. In 1-persistent MARI-BTMA, instead of waiting random intervals till the beginning of a frame, all the nodes generating data in the middle of a frame will contend to access the system at the beginning of the next frame. Then all the nodes with data to transmit at the beginning of a frame will get a back off minislot in the contention period and sense the busy-tone channels which is the same with the basic MARI-BTMA.

3.3. Adaptive MARI-BTMA

From the previous two subsections, we know that 1-persistent MARI-BTMA is suitable to the low traffic load, while basic MARI-BTMA is appropriate to the high traffic load. Therefore, in this subsection, we present an adaptive MARI-BTMA to combine the performance of basic MARI-BTMA and 1-persistent MARI-BTMA. In adaptive MARI-BTMA, when the traffic load is low, 1-persistent MARI-BTMA is used while when the traffic load is high, basic MARI-BTMA is adopted. The switch point between 1-persistent scheme and basic scheme depends on the frame length of MARI-BTMA and the traffic load.

4. PERFORMANCE ANALYSIS

In this section, the throughput of MARI-BTMA is analyzed using the method developed by Tobagi and Kleinrock in their study of CSMA and BTMA [18, 19]. The network model consists of a large number of terminals communicating with each other over a single channel. All nodes are within the range of each other. We make the following assumptions for MARI-BTMA protocol and the analysis.

(i) There are $N$ nodes in the system.
(ii) Each node has two antennas. If there are more than two nodes transmitting simultaneously, the receiver cannot recover the original signals correctly. Correspondingly in RI-BTMA, only one node can be accessed in the system, that is, there is no capture effect on the channel.
(iii) Packet collisions are the only source of packet errors.
(iv) The busy-tone signals and the data signals have the same transmission range.
(v) The interference between the busy-tone signals and the data signals is negligible.
(vi) The bandwidth consumption of the busy tones is negligible compared to the bandwidth of the data channel.
(vii) The number of minislots in a contention period is $m_1$ and the number of minislots in contention-free period is $m_2$, which is the packet length. Therefore, the frame length is $m_1 + m_2$.
(viii) The arrival of the packets of each node, including newly generated packets and rescheduled packets, constitutes a Bernoulli process with probability $p$ per minislot at each node. Here, the packet will be rescheduled which means that it waits a random interval and tries again.
(ix) The preamble can be successfully received only if there is exactly one preamble transmitted in that minislot.

4.1. Throughput of basic MARI-BTMA

Suppose there are currently $M$ nodes with packets to transmit at the beginning of a frame. These $M$ nodes will first randomly select a minislot in the contention period. Let $E_i$ denote the event that there is only one node choosing the $i$th minislot, that is, there is no collision in the $i$th minislot. We call this minislot then “collision-free” minislot. Then the probability of at least one collision-free minislot in all the $m_1$ minislots in the contention period is

$$P_1(m_1, M) = P\left(\bigcup_{i=1}^{m_1} E_i\right)$$

$$= \sum_{i=1}^{m_1} P(E_i) - \sum_{i_1 < i_2} P(E_{i_1} \cap E_{i_2}) + \cdots$$

$$+ (-1)^{m_1} \sum_{i_1 < i_2 < \cdots < i_n} P(E_{i_1} \cap E_{i_2} \cdots \cap E_{i_n})$$

$$+ \cdots + (-1)^{m_1+1} P\left(E_1 \cap E_2 \cdots \cap E_{m_1}\right), \quad m_1 < M,$$

$$= \sum_{i=1}^{m_1} P(E_i) - \sum_{i_1 < i_2} P(E_{i_1} \cap E_{i_2}) + \cdots$$

$$+ (-1)^{m_1} \sum_{i_1 < i_2 < \cdots < i_n} P(E_{i_1} \cap E_{i_2} \cdots \cap E_{i_n})$$

$$+ \cdots + (-1)^{M+1} P\left(E_1 \cap E_2 \cdots \cap E_M\right), \quad m_1 > M.$$
there are $m_1^M$ ways to put all the $M$ nodes in the $m_1$ minislots:

$$P(E_{i_1} \cap E_{i_2} \cdots \cap E_{i_n}) = \frac{(M!/(M-n)!)(m_1 - n)^{M-n}}{m_1^M}. \quad (2)$$

Therefore,

$$\sum_{i_1 < i_2 < \cdots < i_n} P(E_{i_1} \cap E_{i_2} \cdots \cap E_{i_n}) = \frac{\binom{M_1}{n}(M!/(M-n)!)(m_1 - n)^{M-n}}{m_1^M} \quad (3)$$

When there are more than two collision-free minislots, only the nodes in the first two collision-free minislots can send preambles since there are at most two busy tones in the system. Let $P_2(M)$ be the probability of one successful preamble transmitted in the contention period. Let $P_3(M)$ be the probability of two successful preambles transmitted in the contention period. Therefore $P_3(M)$ is the probability of exactly one collision-free minislot in the contention period. $P_3(M)$ is the probability of at least two collision-free minislots in the contention period. The probability that only one minislot is collision-free is equal to the probability that none of the remaining $m_1 - 1$ minislot is collision-free. First, we fix attention on a particular collision-free minislot and a particular node selecting that minislot. When there are more than two collision-free minislots, only this minislot and this node are collision free. Therefore, $P_2(M) = m_1M(m_1 - 1)^{M-1}[1 - P_1(m_1 - 1, M - 1)]$, and $P_3(M) = P_1(m_1, M) - P_2(M). \quad (4)$

Thus the throughput obtained given $M$ nodes with packets to transmit at the beginning of a frame is

$$S(M) = \frac{P_2(M)m_2 + 2P_3(M)m_2}{m_1 + m_2}. \quad (5)$$

Let $X$ denote the number of packet generated and rescheduled at the beginning of a frame. $X$ is a binomial random variable with parameter $N$ and $p$. Thus

$$P(X = M) = \binom{N}{M} p^M (1 - p)^{N-M}. \quad (6)$$

Therefore, the average throughput of the system is

$$\bar{S} = \sum_{M=1}^{N} P(M)S(M). \quad (7)$$

In the following analysis, we set $m_2 = 100$ and $N = 100$. Figure 4 gives the throughput of basic MARI-BTMA with different contention periods. The throughput of RI-BTMA is also given as a benchmark. In [11], Wu and Li give the calculation of the throughput of RI-BTMA: $\eta = (1 + E(\text{length of data portion})/E(X) + E(\text{length of data portion}))$ and $E(X) = 1/P_s$. $P_s$ is the probability of exactly one arrival in the system in a slot. In [11], the preamble is counted as the useful information. However, in this paper, we treat the preamble as overhead, thus the throughput of RI-BTMA is calculated as

$$\bar{S} = \frac{m_2}{1/q + m_2}, \quad (8)$$

where $q = Np(1 - p)^{N-1}$.

Different contention periods correspond to different performance as shown in Figure 4. With an appropriately designed contention period, basic MARI-BTMA is shown to have a much higher throughput than RI-BTMA when the traffic load is high. If the contention period is very short, for example, $m_1 = 2$, the probability that no successful preamble can be transmitted in the contention period will be very high. Thus, the probability that there will be no data transmitted in the contention-free period will be high and the throughput will be reduced. However, if the contention period is very long, for example, $m_1 = 50$, the probability of successful preamble transmitted is high, but the overhead is too high so that the throughput is still low. From Figure 4, we can also see that the higher peak throughput of MARI-BTMA is associated with the unstable situation of the system when the traffic load is very low or very high. One possible solution to this unstable situation is to use 1-persistent MARI-BTMA when the traffic load is very low and a better back off scheme when the traffic load is very high.

**4.2. Throughput of 1-persistent MARI-BTMA**

Let $Y$ denote the number of packets contending at the beginning of a frame. In 1-persistent MARI-BTMA, $Y$ is a
Throughput of 1-persistent MARI-BTMA with different contention periods.

Throughput $p_{\text{MARI-BTMA}}$, $m_1 = 2$

$1 - p_{\text{MARI-BTMA}}$, $m_1 = 10$

$1 - p_{\text{MARI-BTMA}}$, $m_1 = 32$

$1 - p_{\text{MARI-BTMA}}$, $m_1 = 50$

RI-BTMA

Figure 5: Throughput of 1-persistent MARI-BTMA with different contention periods.

Thus, the average throughput is

$$S'' = \sum_{M=1}^{N(m_1+m_2)} P''(M) S(M),$$

(10)

where $S(M)$ is obtained from (5). The throughput of 1-persistent MARI-BTMA is shown in Figure 5.

From it, we can see that 1-persistent MARI-BTMA improves the throughput greatly when the traffic load is very low. However, when the traffic load is high, it goes down very quickly.

4.3. Throughput of adaptive MARI-BTMA

From the previous Sections 4.1 and 4.2, we know that 1-persistent MARI-BTMA works very well at the low traffic load while basic MARI-BTMA works very well at the high traffic load. In this subsection, we investigate the performance of the system using adaptive MARI-BTMA. We set $m_1 = 32$, $m_2 = 100$, and $N = 100$. From Figures 4 and 5, we know the cross point of basic MARI-BTMA and 1-persistent MARI-BTMA is $p = 0.01$. Therefore, we select $p = 0.01$ as the switch point between 1-persistent scheme and basic scheme, that is, when the statistic packet generation probability is less than 0.01, 1-persistent scheme is used. However, if that probability is larger than 0.01, basic MARI-BTMA is used. Figure 6 shows the performance of adaptive MARI-BTMA.

4.4. Saturation throughput analysis

In this subsection, the saturation throughput of basic MARI-BTMA is analyzed. The saturation throughput is obtained when all the $N$ nodes in the system always have data to transmit. In Section 4.1, if $p = 1$, it is in the saturation situation. Thus, the saturation throughput (ST) is

$$ST = \frac{P_2(N)m_2 + 2P_3(N)m_2}{m_1 + m_2}.$$  

(11)

In Figure 7, both the theoretical result from (11) and simulation results are given when $m_1 = 32$ and $N = 50$. From Figure 7, we can see that simulation results and theoretical results match very well. The slight difference is caused by the limited number of iterations in simulation. Compared with

Figure 6: Throughput of adaptive MARI-BTMA.

Figure 7: Saturation throughput of MARI-BTMA.
the saturation throughput of 802.11 given by [13], which is around 0.9 with the same windows length (32) and the number of nodes (50), the saturation throughput of MARI-BTMA is much higher, around 1.5.

4.5. Delay performance of MARI-BTMA

In this subsection, the delay performance of basic MARI-BTMA is given. Similar to the delay performance of slotted Aloha given in [14], we can get the average delay

\[ D = 1 - \frac{1}{p} + \frac{N}{\bar{S}}, \]  

(12)

where \( p \) is the transmission probability, \( N \) is the number of nodes in the system, and \( \bar{S} \) is the throughput from (7). The average delay obtained from (12) is the average number of frames back logged. If we calculate in minislots, the average delay is

\[ D = (m_1 + m_2) \times \left(1 - \frac{1}{p} + \frac{N}{\bar{S}}\right). \]  

(13)

Similarly, the average delay of basic RI-BTMA is

\[ D = (1 + m_2) \times \left(1 - \frac{1}{p} + \frac{N}{\bar{S}}\right), \]  

(14)

where \( \bar{S} \) is calculated from (8). The relation between delay and number of nodes is given in Figure 8. It can be seen from Figure 8 that delay increases with the increasing of the number of nodes in the system and the delay of MARI-BTMA is much less than that of RI-BTMA.

5. SIMULATION RESULTS

In this section, two scenarios are expressed to compare the performance of the new MAC scheme with the traditional CSMA/CA and RI-BTMA. In the first scenario, the effect of physical layer to the throughput is considered. For the second scenario, we only consider the effect of MAC to the throughput, that is, the collision of more than two packets is the only source of packet error. All the simulations are run in MATLAB.

5.1. Simulation scenario 1

In this subsection, simulation results are given both in the physical layer and MAC layer. We use the same simulation scenario in [17] shown in Figure 9. Node A has constant data packets to transmit to node B. Node C has constant data packets to transmit to node D. The distances between A and B, C, and D are fixed, while the distance between B and C can be changed. For the distances in the simulation, the relative distances are used.

5.1.1. Physical layer performance

In this paper, to simplify the simulation, we use Alamouti space-time coding at the transmitter. At the receiver, we use the signal separation algorithm described in [9]. Improved space-time or space-frequency coding and space-time processors can be used in this scheme directly. Figure 10 gives the performance of packet error rate (PER) versus signal-to-noise-ratio (SNR). In the simulation, all the antennas at the transmitter nodes have the same power. A 4-tap frequency selective channel model is used with the variance equal to the path loss \((1/d^4)\). The channel information is assumed to be known at the receiver side.

5.1.2. MAC layer performance evaluation

In this subsection, the normalized throughputs of CSMA/CA and MARI-BTMA are given. The input signal-to-noise ratio...
Table 1: Comparison of throughput of CSMA/CA and MARI-BTMA.

|                  | Access probability | $D/r = 0.5$ | $D/r = 1$ | $D/r = 2$ | $D/r = 3$ |
|------------------|--------------------|-------------|-----------|-----------|-----------|
|                  | User1              | 0.5         | 0.5       | 1         | 1         |
|                  | User2              | 0.5         | 0.5       | 1         | 1         |
| CSMA/CA          | PER                | 0.0075      | 0.0075    | 0.0566    | 0.0164    |
|                  | User1              |             | 0.0075    | 0.0075    | 0.01      | 0.0082    |
|                  | User2              |             |           |           |           |
|                  | Throughput         | 0.4963      | 0.4963    | 0.9434    | 0.9836    |
|                  | User1              |             | 0.4963    |           | 0.99      | 0.9918    |
|                  | User2              |             |           | 0.4963    | 0.99      | 0.9918    |
| MARI-BTMA        | Access probability | 1           | 1         | 1         | 1         |
|                  | User1              |             | 1         |           |           |           |
|                  | User2              |             |           |           |           |           |
|                  | PER                | 0.0077      | 0.0188    | 0.0566    | 0.0164    |
|                  | User1              |             | 0.0263    | 0.0566    | 0.01      | 0.0082    |
|                  | User2              |             |           |           |           |
|                  | Throughput         | 0.9021      | 0.892     | 0.8576    | 0.8942    |
|                  | User1              |             | 0.8852    |           | 0.9       | 0.9016    |
|                  | User2              |             |           | 0.8576    |           |           |

is 15 dB. The carrier sensing radius is 1.5. The detection radius of RTS and CTS is 1.1. The busy-tone sensing radius is 1.5. Packet error rate (PER) is used in the calculation of throughput. In this simulation, 1/2 rate convolution channel coding is used. Throughput is calculated as

$$S = (1 - \text{PER}) \ast \frac{\text{number of accessed packets}}{\text{total number of packets}} \ast \frac{m_2}{m_1 + m_2}$$

(15)

In the simulation, $m_1 = 10$ and $m_2 = 100$. For CSMA/CA, we ignore the overhead of RTS/CTS.

From Table 1, we can see that when node B and node C are close to each other, only one node is accessed with CSMA/CA. However, with MARI-BTMA, both nodes can access to the system. When node B and node C are far away, for example, larger than 1.5, CSMA/CA and MARI-BTMA both guarantee the access of these two nodes. However, for the reason of the fixed structure, the overhead of MARI-BTMA is slightly higher than CSMA/CA. It is interesting to point out that since PER is in the level of $10^{-2}$, it is access control probability which dominates the throughput.

5.2. Simulation scenario 2

In the above subsection, we can see that MARI-BTMA works well in the simple scenario. In this subsection, we will see that MARI-BTMA also works well in a scalable system. In the simulation scenario given in Figure 11, Node B to node A, node C to node A, node D to node E, node H to node E, and node G to node F have constant data flows to transmit. All the nodes have the same distance $r$ with each other. The carrier sensing range is $r$. The lengths of both contention period are 32 minislots.

Simulation results are given in Table 2. It can be seen from Table 2 that MARI-BTMA can get better performance than MIMA in [17] and much higher throughput than RI-BTMA and CSMA/CA. The problem with MIMA is that it has hidden node problems and exposed node problems associated with CSMA/CA which cause MIMA to have high overhead. Therefore, even though MIMA can access two users simultaneously, its throughput is still less than RI-BTMA.

6. CONCLUSION

In this paper, we propose a new transceiver architecture with MIMO-OFDM in the physical layer and MARI-BTMA in the MAC layer. MARI-BTMA uses multiple out of band signals busy tones to notify the number of users in the system so that to avoid the collision of the nodes on the same channel. In MARI-BTMA, the packet slot is divided into two subframes: contention subframe and contention-free subframe. The contention subframe is used to access the nodes, while the contention-free subframe is used to transmit data for the successfully accessed nodes. Two MARI-BTMA protocols are proposed in this paper. One is basic MARI-BTMA which is suitable to moderate traffic load. The other is 1-persistent MARI-BTMA which is used in the system with low traffic load. By combining basic MARI-BTMA and 1-persistent MARI-BTMA, an adaptive MARI-BTMA is proposed. The throughput analysis of basic MARI-BTMA, 1-persistent MARI-BTMA, and adaptive MARI-BTMA as well as the delay performance of the basic MARI-BTMA are given in this paper. From both the theoretical analysis and simulation results, the performance of MARI-BTMA is shown to be much better than that of CSMA/CA, RI-BTMA or MIMA.
ACKNOWLEDGMENTS
This work was supported in part by the US Army Contract W15QKN-05-p-0261, AFOSR Grant FA9550-05-1-0529, and NSF Grant CNS-0520232.

REFERENCES
[1] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, “Space-time block codes from orthogonal designs,” IEEE Transactions on Information Theory, vol. 45, no. 5, pp. 1456–1467, 1999.
[2] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, “Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture,” Electronics Letters, vol. 35, no. 1, pp. 14–16, 1999.
[3] J. Ha, A. N. Mody, J. H. Sung, J. R. Barry, S. W. McLaughlin, and G. L. Stüber, “LDPC coded OFDM with alamouti/SVD diversity technique,” Wireless Personal Communications, vol. 23, no. 1, pp. 183–194, 2002.
[4] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” IEEE Journal on Selected Areas in Communications, vol. 21, no. 5, pp. 684–702, 2003.
[5] S. Vishwanath, N. Jindal, and A. Goldsmith, “On the capacity of multiple input multiple output broadcast channels,” in IEEE International Conference on Communications (ICC ’02), vol. 3, no. 3, pp. 1444–1450, New York, NY, USA, April–May 2002.
[6] “IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,” November 1997. P802.11.
[7] H. Yang, “A road to future broadband wireless access: MIMO-OFDM-based air interface,” IEEE Communications Magazine, vol. 43, no. 1, pp. 53–60, 2005.
[8] Y. Li, J. H. Winters, and N. R. Sollenberger, “Signal detection for MIMO-OFDM wireless communications,” in IEEE International Conference on Communications (ICC ’01), vol. 10, pp. 3077–3081, Helsinki, Finland, June 2001.
[9] R. S. Blum, Y. Li, J. H. Winters, and Q. Yan, “Improved space-time coding for MIMO-OFDM wireless communications,” IEEE Transactions on Communications, vol. 49, no. 11, pp. 1873–1878, 2001.
[10] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, “Inter-carrier interference in MIMO OFDM,” IEEE Transactions on Signal Processing, vol. 50, no. 10, pp. 2451–2464, 2002.
[11] C. Wu and V. O. K. Li, “Receiver-initiated busy-tone multiple access in packet radio networks,” in Proceedings of the ACM Workshop on Frontiers in Computer Communications Technology (SIGCOMM ’87), pp. 336–342, Stowe, VT, USA, August 1987.
[12] Z. J. Haas and J. Deng, “Dual busy tone multiple access (DBTMA) - a multiple access control scheme for ad hoc networks,” IEEE Transactions on Communications, vol. 50, no. 6, pp. 975–985, 2002.
[13] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function,” IEEE Journal on Selected Areas in Communications, vol. 18, no. 3, pp. 535–547, 2000.
[14] R. Rom and M. Sidi, Multiple Access Protocols Performance and Analysis, Springer, New York, NY, USA, 1990.
[15] Q. Zhao and L. Tong, “A multiqueue service room MAC protocol for wireless networks with multipacket reception,” IEEE/ACM Transactions on Networking, vol. 11, no. 1, pp. 125–137, 2003.
[16] K. Sundaresan, R. Sivakumar, M. A. Ingram, and T.-Y. Chang, “Medium access control in ad hoc networks with MIMO links: optimization considerations and algorithms,” IEEE Transactions on Mobile Computing, vol. 3, no. 4, pp. 350–365, 2004.
[17] T. Tang, M. Park, R. W. Heath Jr., and S. M. Nettles, “A joint MIMO-OFDM transceiver and MAC design for mobile ad hoc networking,” in International Workshop on Wireless Ad-Hoc Networks (IWWAN ’04), pp. 315–319, Oulu, Finland, May–June 2004.
[18] L. Kleinrock and F. A. Tobagi, “Packet switching in radio channels—1. Carrier sense multiple-access modes and their throughput-delay characteristics,” IEEE Transactions on Communications, vol. 23, no. 12, pp. 1400–1416, 1975.
[19] F. A. Tobagi and L. Kleinrock, “Packet switching in radio channels—2. The hidden terminal problem in carrier sense multiple-access and the busy-tone solution,” IEEE Transactions on Communications, vol. 23, no. 12, pp. 1417–1433, 1975.

Dandan Wang received the B.S. and M.S degrees in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2000 and 2003, respectively. She worked at China Radio Research Lab, Ericsson, from 2003 to 2004. From August 2004 to August 2005, she was with Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, New Jersey. She is currently a Ph.D. student at the Department of Electrical and Computer Engineering, University of Texas at Dallas. Her research interests include wireless communications and networking, signal processing for communications, and sensor networks.

Uf Tureli received the B.S. degree in 1994 from Bogazici University, Istanbul, Turkey, and the M.S. and Ph.D. degrees in 1998 and 2000, respectively, from the University of Virginia, Charlottesville, all in electrical engineering. Since July 2000, he has been an Assistant Professor at the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, New Jersey. He is the Director of the Wireless Research Laboratory at the Department of Electrical and Computer Engineering and Associate Director of Wireless Network Security Center (WiNSeC) at Stevens Institute of Technology. His research interests include wireless communications and networking, signal processing with application to broadband wireless networks, estimation and detection for scalable, adaptive, and robust communications and propagation studies. He has published numerous journal and conference articles in detection and estimation for scalable, adaptive, and robust broadband wireless communications.
A Novel Cluster-Based Cooperative MIMO Scheme for Multi-Hop Wireless Sensor Networks

Yong Yuan,1 Min Chen, 2 and Taekyoung Kwon3

1 Department of Electronics and Information, Huazhong University of Science and Technology, Wuhan 430074, China
2 Department of Electrical and Computer Engineering, University of British Columbia, BC, Canada V6T 1Z4
3 School of Computer Science and Engineering, Seoul National University, Seoul 151742, South Korea

Received 4 November 2005; Revised 11 April 2006; Accepted 26 May 2006

A cluster-based cooperative multiple-input-multiple-output (MIMO) scheme is proposed to reduce the adverse impacts caused by radio irregularity and fading in multi-hop wireless sensor networks. This scheme extends the LEACH protocol to enable the multi-hop transmissions among clusters by incorporating a cooperative MIMO scheme into hop-by-hop transmissions. Through the adaptive selection of cooperative nodes and the coordination between multi-hop routing and cooperative MIMO transmissions, the scheme can gain effective performance improvement in terms of energy efficiency and reliability. Based on the energy consumption model developed in this paper, the optimal parameters to minimize the overall energy consumption are found, such as the number of clusters and the number of cooperative nodes. Simulation results exhibit that the proposed scheme can effectively save energy and prolong the network lifetime.

Copyright © 2006 Yong Yuan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Due to the limited energy and difficulty to recharge a large number of sensors, energy efficiency and maximizing network lifetime have been the most important design goals for wireless sensor networks (WSNs). However, channel fading, interference, and radio irregularity pose big challenges on the design of energy efficient communication and routing protocols in the multi-hop WSNs.

As the MIMO technology has the potential to dramatically increase the channel capacity and reduce transmission energy consumption in fading channels [1], cooperative MIMO schemes have been proposed for WSNs to improve communication performance [2–5]. In those schemes, multiple individual single-antenna nodes cooperate on information transmission and/or reception for energy-efficient communications. Cui et al. [2] analyzed a cooperative MIMO scheme with Alamouti code for single-hop transmissions in WSNs. Li [3] proposed a delay and channel estimation scheme without transmission synchronization for decoding for such cooperative MIMO schemes. Li et al. [4] also proposed a STBC-encoded cooperative transmission scheme for WSNs without perfect synchronization. Jayaweera [5] considered the training overhead of such schemes.

However, in the above proposals, the multi-hop routing and distributed operations in WSNs are not taken into consideration, which limits the practical use of the cooperative MIMO schemes in WSN. In this paper we study the feasibility of a cooperative MIMO scheme in multi-hop WSNs. Radio irregularity of wireless communications and multi-hop routing is considered with the cooperative MIMO scheme. On the other hand, due to its ability of frequency reuse and efficiency in processing highly correlated data, clustering is efficient in the design of WSNs. Therefore, we incorporate the cooperative MIMO scheme with the LEACH protocol, which is an efficient clustering protocol due to its energy-efficient, randomized, adaptive, and self-configuring cluster formation. As only single-hop communications from cluster heads to the sink are considered in the original LEACH protocol, we modify the LEACH protocol to allow cluster heads to form a multi-hop backbone and incorporate the cooperative MIMO scheme into each single-hop transmission. Based on the proposed scheme, we investigate the energy consumption of each transmission/reception. Then, the overall energy consumption model is developed, and the optimal parameters of the scheme are found such as the number of clusters and the number of cooperative nodes.
The overall energy consumption of the proposed scheme is analyzed in Section 3. Section 4 presents simulation results and discussions. Section 5 concludes the paper.

2. THE MULTI-HOP MIMO-LEACH SCHEME

In this section, we will discuss the proposed multi-hop MIMO-LEACH scheme, which is illustrated in Figure 1. First, the strategy to find appropriate cooperative nodes in the single-hop communications between cluster heads is proposed in Section 2.1. Based on the strategy, the multi-hop MIMO-LEACH scheme is presented in Section 2.2.

2.1. Strategy to choose cooperative nodes

To maximize the performance of single-hop communications between cluster heads, an appropriate strategy should be taken to choose the optimal cooperative nodes. Suppose that the current cluster head will use \( J \) cooperative nodes to transmit data to its neighboring cluster head \( t \) by the cooperative MIMO scheme. An AWGN channel with squared power path loss is assumed for intracluster communications. For the intercluster communications, we assume the transmission from each cooperative node experiences frequency-nonsellective and slow Rayleigh fading. Furthermore, the long distance between any two nodes in the network with respect to the wavelength gives rise to independent fading coefficients for the cooperative nodes. The rationale behind such channel assumptions is that the inter-cluster transmission distance is much larger than the intra-cluster transmission distance and the transmission environments are more complex in the inter-cluster communication.

Denote the distance between node \( j \) and its current cluster head by \( d_{jt} \). Also, denote the distance and path loss for node \( j \) to communicate with \( t \) as \( d_{jt} \) and \( k_{jt} \), respectively. For each single-hop transmission, the current cluster head will broadcast a data packet to the cooperative nodes. Then, the cooperative nodes will encode and transmit the transmission sequence according to the orthogonal space-time block codes (STBC) to cluster head \( t \) toward the sink node. The energy consumption for these two operations in the single-hop transmission will be modeled in the remainder of this section. Then, a novel strategy will be developed to find the optimal set of cooperative nodes to minimize the overall energy consumption. In developing the strategy, we assume BPSK is adopted as the modulation scheme and the bandwidth is \( B \) Hz.

1. The energy consumption for the intracluster transmission

Denote by \( E_{bt}(1) \) the energy consumption for the current cluster head to broadcast one bit to the cooperative nodes. \( E_{bt}(1) \) can be broken down into two main components, the transmit energy consumption \( E_{bt}(1) \) and the circuit energy consumption \( E_{ct}(1) \).

The BER performance for BPSK is \( P_b = Q(\sqrt{2r}) \). Here \( r \) is the signal-to-noise ratio (SNR), which is defined as \( r = P_t/(2B\sigma^2 N_f) \) [6] under the assumption of AWGN channel, where \( P_t \) is the received signal power, \( \sigma^2 \) is the power density of the AWGN, and \( N_f \) is the receiver noise figure.

In the high SNR regime, we can approximate the BER performance as \( P_b = e^{-r} \) by the Chernoff bound [6]. Hence, we obtain \( P_r = -2BN_f \sigma^2 \ln(P_b) \). As the assumption of squared power path loss, \( E_{bt}(1) \) can be modelled by

\[
E_{bt}(1) = E_{bt}(1) + E_{ct}(1) = -2(1 + \alpha)N_f \sigma^2 \ln(P_b)G_1d_{max}^2M_1 + \frac{P_{ct} + JP_{cr}}{B},
\]

where \( d_{max} \) is the maximum distance from the cooperative nodes to the cluster head, \( \alpha \) is the efficiency of the RF power amplifier, \( G_1 \) is the gain factor at \( d_{max} = 1 \), \( M_1 \) is the link margin, \( N_f \) is the receiver noise figure, and \( P_{ct} \) and \( P_{cr} \) are the circuit power consumption of the transmitter and receiver, respectively [2].

Let \( f_1(P_b) = -2N_f \sigma^2 \ln(P_b) \) and \( H(d_{max}) = G_1M_1d_{max}^2 \). Then, (1) can be rewritten as

\[
E_{bt}(1) = (1 + \alpha)f_1(P_b)H(d_{max}) + \frac{P_{ct} + JP_{cr}}{B}.
\]

According to the definition, \( H(d_j) \) can be measured as follows. Let the current cluster head transmit a signal with transmit power \( P_{out} \). Then, the power of the received signal at its cluster member, node \( j \), is \( P_{jt} = P_{out}/H(d_j) \). Therefore, \( H(d_j) \) can be measured as

\[
H(d_j) = \frac{P_{out}}{P_{jt}}.
\]
From (2), we can find that the energy consumption in the intra-cluster transmission, $E_{bt}(1)$, can be reduced by choosing the nearer cooperative nodes.

(2) The energy consumption for the intercluster transmission

To analyze the energy consumption for inter-cluster transmissions based on the cooperative scheme, denoted by $E_{bt}(2)$, we refine the results in [2]. In [2] an equal transmit power allocation scheme is used as the channel state information (CSI) is not available at the transmitter. If the average attenuation of the channel for each cooperative node pair can be estimated, we can use an equal signal-to-noise (SNR) policy [7] to allocate the transmit power for its effectiveness and simplicity. The average energy consumption per bit transmission by BPSK in such a scheme can be approximated by

$$E_{bt}(2) = (1 + \alpha) \frac{N_0}{P_b} \sum_{j=1}^{J} \frac{(4\pi)^2 d_{jt}^4}{G_t G_r \lambda^2} M_l N_f + \frac{(J P_{ct} + P_{cr})}{B},$$

where $N_0$ is the single-sided noise power spectral density, $P_b$ is the desired BER performance, $G_t$ and $G_r$ are the transmitter and receiver antenna gains, respectively, also, $\lambda$ is the carrier wavelength [2]. The training overhead and transmission rate are not considered in (4), which will be considered in Section 3.

The average attenuation of the channel for node $j$ can be estimated as follows. Assume the channel is symmetric, and $t$ transmits a signal with transmit power $P_{out}$, then the power of the received signal at node $j$, $P_{jt}$, can be given by

$$P_{jt} = P_{out} \frac{G_t G_r 1^2}{(4\pi)^2 d_{jt}^4 M_l N_f} = \frac{P_{out}}{G(d_{jt}, k_{jt})},$$

where $G(d_{jt}, k_{jt}) = P_{out}/P_{jt} = ((4\pi)^2 d_{jt}^4/G_t G_r 1^2) M_l N_f$. Therefore, (4) can be reformulated as

$$E_{bt}(2) = (1 + \alpha) \frac{N_0}{P_b} \sum_{j=1}^{J} G(d_{jt}, k_{jt}) + \frac{(J P_{ct} + P_{cr})}{B}$$

$$= (1 + \alpha) f_2(P_b) \sum_{j=1}^{J} G(d_{jt}, k_{jt}) + \frac{(J P_{ct} + P_{cr})}{B}. \quad (6)$$

According to (6), the transmit power of node $j$ to communicate with cluster head $t$ can be described by

$$P_{outj} = G(d_{jt}, k_{jt}) \frac{N_0 B}{P_b}. \quad (7)$$

(3) The strategy to choose cooperative nodes

Based on (2) and (6), the overall energy consumption for the single-hop transmission can be written as (8)

$$E_{bt} = E_{bt}(1) + E_{bt}(2)$$

$$= (1 + \alpha) \left[ f_1(P_b) H(d_{max}) + f_2(P_b) \sum_{j=1}^{J} G(d_{jt}, k_{jt}) \right]$$

$$+ \frac{(J + 1)(P_{ct} + P_{cr})}{B}. \quad (8)$$

From (8), the energy consumption for the intracluster transmission $E_{bt}(1)$ and intercluster transmission $E_{bt}(2)$ should be traded off to minimize $E_{bt}$. $E_{bt}$ can be minimized by choosing an appropriate set of cooperative nodes, which can minimize $f_1(P_b) H(d_{max}) + f_2(P_b) \sum_{j=1}^{J} G(d_{jt}, k_{jt})$. In order to simplify the distributed strategy design, the cooperative nodes should be chosen as the nodes whose $f_1(P_b) H(d_{jt}) + f_2(P_b) G(d_{jt}, k_{jt})$ are minimal. In addition, in order to balance the energy consumption, the selection criterion is defined as

$$\beta_{jt} = \frac{E_j}{f_1(P_b) H(d_{jt}) + f_2(P_b) G(d_{jt}, k_{jt})}, \quad (9)$$

where $E_j$ is the remaining energy in the current round for node $j$. The rationale behind definition of $\beta_{jt}$ is that the node, which has a good tradeoff between $E_{bt}(1)$ and $E_{bt}(2)$ and has more remaining energy, should have a larger chance to be selected as cooperative node. Therefore, $J$ nodes with maximum $\beta_{jt}$ will be chosen as the cooperative nodes to communicate with cluster head $t$.

2.2. Scheme design

In this section, we will discuss how to enable cluster heads to form a multi-hop backbone by incorporating the cooperative MIMO scheme into the LEACH protocol for each single-hop transmission. As assumed in the LEACH protocol, each node has a unique identifier (ID). The transmit power of each node can be adjusted, and the nodes are assumed to be always synchronized. Similarly, the operations of the proposed scheme are broken into rounds. Each round consists of three phases: (i) cluster formation phase, during which the clusters are organized and cooperative MIMO nodes are selected; (ii) routing phase, during which a routing table in each selected node is constructed; and (iii) transmission phase, during which data are transferred from the nodes to the cluster heads and forwarded to the sink according to the routing table.

1) Cluster formation phase

In this phase, each node will elect itself to be a cluster head with a probability $p$ as specified in the original LEACH protocol. After the cluster heads are elected, each cluster head will broadcast an advertisement message (ADV) by transmit power $P_{out}$ using a nonpersistent CSMA MAC protocol. The
message contains the head’s ID. If a cluster head receives the advertisement message from another head \( t \) and the received signal strength (RSS) exceeds a threshold \( t_h \), it will take cluster head \( t \) as a neighboring cluster head and record \( t \)'s ID. As for the noncluster head, node \( j \), it will record all the RSSs of the received advertisement messages, and choose the cluster head whose RSS is the maximum. Then, it will calculate and save \( H(d_j), G(d_{jt}, k_{jt}), \beta_{jt}, \) and \( P_{out_{jt}} \) by (3), (5), (7), and (9). Then node \( j \) will join the cluster by sending a join-request message (Join-REQ) to the chosen cluster head. This message contains the information of the node’s ID, the chosen cluster head’s ID, and the corresponding values of \( \beta_{jt} \). After a cluster head has received all join-request messages, it will set up a TDMA schedule and transmit this schedule to its members as in the original LEACH protocol. If the sink receives the advertisement message, it will find the cluster head with the maximum RSS, and send the sink-position (Sink-POS) message to the cluster head and mark the cluster head as the target cluster head (TCH).

After the clusters are formed, each cluster head will select corresponding optimal \( J \) cooperative nodes for cooperative MIMO communications with each of its neighboring cluster heads. As stated in Section 2.1, \( J \) nodes with maximum \( \beta_{jt} \) will be chosen to communicate with a neighboring cluster head \( t \). If no such \( J \) nodes can be found for \( t \), \( t \) will be removed from the neighbor list, since too much energy is consumed for communicating with \( t \). After selecting the cooperative nodes, the total energy per bit transmission for communications with \( t \), \( E_{b_t} \), can be derived by (4). Then, \( E_{b_t} \), the ID set of the cooperative nodes for each neighboring cluster head, will be stored. At the end of this phase, the cluster head will broadcast a cooperate-request message (COOPERATE-REQ) to each cooperative node, which contains the ID of the cluster itself, the ID of the neighboring cluster head \( t \), the IDs of the cooperative nodes, and the index of the cooperative nodes in the cooperative nodes set for each cluster head \( t \). Each cooperative node that receives the cooperate-request message (COOPERATE-REQ) will store the ID of \( t \), the index, and the transmit power \( P_{out_{jt}} \) and send back a cooperate-ACK message (COOPERATE-ACK) to the cluster head.

(2) Routing table construction

To construct the routing table, the basic ideas of distance-vector-based routing will be used. Each cluster head will maintain a routing table, in which each entry contains destination cluster ID, next hop cluster ID, IDs of cooperative nodes, and mean energy per bit. Initially, only the neighboring cluster head will have a record in the routing table. Then each cluster head will simply inform its neighboring cluster heads of its routing table. After receiving route advertisements from neighboring cluster heads, the cluster head will update its routing table according to the route cost and advertise to its neighboring cluster heads the modified routes. After several rounds of route exchange and update, the routing table of each cluster head will be converged to the optimal one. Then, TCH will flood a target announcement message (TARGET-ANNOUNCEMENT) containing its ID to each cluster head to enable the creation of paths to the sink.

(3) Data transmission

In this phase, cluster members will transmit first their data to the cluster head by multiple frames as in the traditional LEACH protocol. In each frame, each cluster member will transmit its data during its allocated transmission slot specified by the TDMA schedule in cluster formation phase, and it will be sleep in other slots to save energy. The duration of a frame and the number of frames are the same for all clusters. Thus the duration of each slot depends on the number of members in the cluster. After a cluster head receives data frames from its cluster members, it will perform data aggregation to remove the redundancy in the data. After aggregating received data frames, the cluster head will forward the data packet to the TCH by multiple hops routing. In each single-hop communication, if there exist \( J \)-cooperative MIMO nodes, the cluster head will add a packet header to the data packet, which includes the information of source cluster ID, next-hop cluster ID, and destination cluster ID. Then the data packet is broadcasted. Once the corresponding cooperative nodes receive the data packet, they will encode the data packet by orthogonal STBC, and transmit the data as an individual antenna with transmission power \( P_{out_{jt}} \) in the MIMO antenna array. In the cooperative MIMO scheme, the transmission delay and channel estimation scheme proposed in [3] can be used to solve the problem of imperfect synchronization in decoding.

3. THE ENERGY CONSUMPTION MODEL OF THE SCHEME

In this section, we will analyze the energy consumption of the scheme. Based on the result, we will develop an optimization model to find the optimal parameters in the scheme, including the number of clusters \( k_c \), and the number of cooperative nodes \( J \).

In analysis, we make the following assumptions. (1) There are \( N \) nodes distributed uniformly in an \( M \times M \) region. (2) An AWGN channel with squared power path loss is assumed for the intracluster communication. (3) A flat Rayleigh fading channel with \( k_\text{th} \)-power path loss is assumed for the intercluster communication. (4) BPSK is used as the modulation scheme and the bandwidth is \( B \) Hz. (5) In each frame every node will send a packet with size \( s \) to the cluster head by probability \( P \). The number of frames in each round is denoted by \( F_n \). (6) In maintaining the routing table in each round, each cluster head will broadcast the routing table, whose size is denoted by \( R_{by} \) for \( R_{by} \) times. (7) The energy consumption for data processing is ignored.

Now, we are ready to model the overall energy consumption in each round, denoted by \( E(k_c,J) \). There are four energy consuming operations in each round. (1) The cluster members transmit data to the cluster head, whose energy consumption is denoted by \( E_t(k_c) \). (2) The cluster heads construct the routing tables, whose energy consumption is
denoted by $E_s(k_c)$. (3) The cluster heads transmit aggregated data to the cooperative nodes in each single-hop transmission, whose energy consumption is denoted by $E_s\theta(k_c, J)$. (4) The cooperative nodes transmit the data to the next cluster head in each single-hop transmission; whose energy consumption is denoted by $E_{c\alpha}(k_c, J)$.

### 3.1. $E_t(k_c)$

In order to model $E_t(k_c)$, we will first analyze the energy consumption for the source nodes to transmit one bit to the cluster head, denoted by $E_{tb}(k_c)$.

Under the assumption of BPSK modulation and AWGN channel with squared power path loss, $E_{tb}(k_c)$ can be modelled in the same manner as $E_{tb}(1)$ in Section 2.1(1),

$$E_{tb}(k_c) = -2(1 + \alpha)N_f\sigma^2\ln (P_b)G_t |d_{tc}^2|M_1 + \frac{P_{at} + P_{cr}}{B}$$

$$= -\frac{1}{\pi k_c}(1 + \alpha)N_f\sigma^2\ln (P_b)G_t M_1 M_2 + \frac{P_{at} + P_{cr}}{B},$$

(10)

where $d_{tc}$ is the distance from the node to the cluster head, $G_t$ is the gain factor at $d_{tc} = 1$ m. In (10), we use the result in [8] that $E(d_{tc}^2) = M^2/2\pi k_c$.

On the other hand, when the number of clusters is $k_c$, the average number of members for each cluster is $[N/k_c]$. Hence, the total number of bits transmitted to the cluster head for each cluster by each round is $S_1(k_c) = [N/k_c]F/k_cP_s$. Therefore, $E_t(k_c) = k_cS_1(k_c)E_{tb}(k_c)$.

### 3.2. $E_s(k_c)$

In this section, we will model the energy consumption in constructing the routing table, denoted by $E_s(k_c)$. When the number of clusters is $k_c$, the radius of each cluster can be approximated as radius $= M\sqrt{\pi k_c}$ [8]. Therefore, the distance between each pair of direct neighboring clusters can be approximated as $d_{nnoc} = 2\text{radius} = 2M\sqrt{\pi k_c}$. We also assume the number of direct neighbors of each cluster is 4. Under the assumption of flat Rayleigh fading channel, $E_s(k_c)$ can be approximated by [2]

$$E_s(k_c) = k_cR_{ct}R_{bt}\left(1 + \alpha\frac{N_0}{P_b}G_tGr\gamma\frac{(4\pi)^2(2M)^k}{\pi k_c}M_1N_f + \frac{P_{at} + 4P_{cr}}{B}\right).$$

(11)

### 3.3. $E_{\theta}(k_c, J)$

In this section, we will analyze the energy consumption for the cluster head to transmit aggregated data to the cooperative nodes, denoted by $E_{\theta}(k_c, J)$. When the cluster head broadcasts the data, $J$ cooperative nodes will receive it. Similar to the analysis of $E_{tb}(k_c)$, the energy per bit for this operation, denoted by $E_{\theta\theta}(k_c, J)$, can be described by

$$E_{\theta\theta}(k_c, J) = -\frac{1}{\pi k_c}(1 + \alpha)N_f\sigma^2\ln (P_b)G_t M_1 M_2 + \frac{P_{at} + JP_{cr}}{B}.$$

(12)

We adopt the aggregation model in [9] to describe the aggregation operation. The amount of data after aggregation for each round is $S_2(k_c) = S_1(k_c)/([N/k_c]\text{agg} + 1)$, where agg is the aggregation factor. Therefore, $E_{\theta}(k_c, J) = k_cS_2(k_c)E_{\theta\theta}(k_c, J)$.

### 3.4. $E_{c\alpha}(k_c, J)$

According to Section 2.1, $J$ cooperative nodes of the current cluster will encode and transmit the transmission sequence according to the orthogonal STBC to the cluster head. In modelling the energy consumption of such operation, we need to consider the impacts of training overhead and transmission rate. Suppose that the block size of the STBC code is $F$ symbols and in each block we include $pJ$ training symbols, and the block will be transmitted in $L$ symbols duration. $F/L$ is called the transmission rate, denoted by $R$. Then, the actual amount of data to transmit the $S_1(k_c)$ bits is $S_2(k_c, J) = F/S_2(k_c)/R(F - pJ)$. Therefore, $E_{c\alpha}(k_c, J)$ can be described by

$$E_{c\alpha}(k_c, J) = S_1(k_c, J)\left((1 + \alpha)\frac{N_0}{P_b}G_t (4\pi)^2(2M)^k\gamma\pi k_c^2M_1N_f + \frac{J\pi k_c^2 + P_{at}}{B}\right).$$

(13)

Based on the above analysis, the overall energy consumption in each round, $E(k_c, J)$ can be described as

$$E(k_c, J) = E_t(k_c) + E_s(k_c) + \pi k_cE_{\theta}(k_c, J) + \pi k_cE_{c\alpha}(k_c, J),$$

(14)

where $\pi k_c$ is the average number of hops. In order to simplify the analysis, we assume $\pi k_c = \sqrt{k_c}$, which is just the number of clusters along each edge of the sensed region.

Based on (14), we can formulate the optimization model to choose the optimal $k_c$ and $J$ as

$$(k_c^*, J^*) = \text{argmin} E(k_c, J) \text{ s.t. } J \leq 10, \ k_c \leq \frac{N}{3},$$

(15)

where the first constraint comes from the fact that more cooperative nodes will not improve the transmission energy efficiency but cost much circuit energy, and the rationale behind the second constraint is that the size of the cluster should not be too small to make efficient aggregation. Since the search space is not large, we can use exhaustive search method to solve (15).

### 4. SIMULATION RESULTS

In the simulations, 400 nodes are randomly deployed on a 200 x 200 field. The location of the sink is randomly chosen...
in each round. The system parameters are summarized in Table 1.

The meanings of the entries in Table 1 are summarized as follows. $\alpha$ is the efficiency of the RF power amplifier, $M_l$ is the link gain, $G_l$ is the gain factor at $1m$, $k$ is the path loss, $\sigma^2$ is the power density of the AWGN channel in the intrachannel communication, $N_f$ is the receiver noise figure, $f_c$ is the carrier frequency, $B$ is the bandwidth, $P_b$ is the desired BER performance, $P_{ct}$ and $P_{cr}$ are the circuit power consumption of the transmitter and receiver, respectively, $F_n$ is the number of frames per round, $G_t, G_r$ are the antenna gains of the transmitter and receiver, $s$ is the packet size, $P$ is the transmit probability of each node, $R$ is the transmission rate, $F$ is the number of symbols in each block, $p$ is the number of required training symbols for each cooperative node, $R_{tr}$ is the times for exchanging the routing table for each round, and $R_{ts}$ is the routing table size.

To simulate the phenomena of radio irregularity, the path loss of the communication between each pair of nodes is distributed randomly from 3 to 5.

Each node begins with 400 J of energy and an unlimited amount of data to send to the sink. When the nodes use up their limited energy during the course of the simulation, they can no longer transmit or receive data.

During the simulation, we tracked the overall number of packets transferred to the sink, the amount of energy and duration required to get the data to the sink, and the percentage of nodes alive. We are interested in the transmission quality and energy saving performance of the proposed scheme. The performance of the proposed multi-hop MIMO-LEACH scheme is compared with the original LEACH and the multi-hop LEACH scheme, in which cooperative MIMO communication is not implemented. The optimal value of $k_c$ for the original LEACH is determined by the model in [8]. We also develop a similar model to find the optimal $k_c$ for the multi-hop LEACH scheme, which will not be discussed here due to the limited space. In the investigated scenario, it is found that the optimal $k_c$ for the original LEACH protocol, the multi-hop LEACH scheme, and the proposed scheme are 3, 41, and 27, respectively. The optimal $J$ for the proposed scheme is found to be 3.

Due to the aggregation operation, the number of effective received packets by sink [8] is a good application-independent indication of the transmission quality. The effective received packets refer to the “real” packets represented by the aggregated packets. If no aggregation carries out, the number of effective received packets equals to the number of actual received packets. If the aggregation operation in transmission is information lossless, the number of effective received packets is just the number of total packets transferred by the source nodes.

Figures 2 and 3 show the total number of effective packets received at the sink over time and the total number of effective packets received at the sink for a given amount of energy.

Figure 2 shows that during its lifetime the LEACH protocol can obtain better latency performance compared to the multi-hop LEACH scheme and the proposed MIMO LEACH scheme. The reason is that the multi-hop operation in the multi-hop LEACH scheme and the multi-hop MIMO-LEACH scheme will increase the latency, and thus result in a less number of data packets sent to the sink for a given period of time. However, the better latency performance of the LEACH protocol comes from the more energy consumption compared to the other two schemes. Especially, in the fading channel environment, LEACH protocol will consume much more energy due to its single-hop transmission from the cluster heads to the sink, which will result in less network lifetime and less total number of transmitted packets. Figure 3 shows that, with the same amount of energy consumption, the multi-hop MIMO-LEACH scheme can transmit much more data packets compared to the LEACH protocol and the multi-hop LEACH scheme. From these simulation results, we can find that the multi-hop MIMO-LEACH scheme is more suitable for the application scenario which has large requirements on network lifetime but little requirements on latency.

Figure 4 shows the percentage of nodes alive over time. From Figure 4, we can find that the proposed multi-hop MIMO-LEACH scheme can improve the network lifetime greatly. If we define the network lifetime of WSN as the duration of more than 70% of network nodes are alive, then we can observe that the network lifetime of WSN with the original LEACH protocol, the multi-hop LEACH scheme, and the proposed multi-hop MIMO-LEACH scheme is about $0.7 \times 10^4, 8.2 \times 10^4$, and $11 \times 10^4$ s. The improvement on network lifetime obtained by the multi-Hop MIMO-LEACH scheme is significant.

However, the percentage of nodes alive over time is not always a good indication to the energy saving performance of a protocol. For example, during the same time, one protocol transmits less packets than other protocols. Then, though the energy saving performance of the protocol is worse than other protocols, it will still consume less energy. In order to further investigate the energy saving performance, we also simulate the performance in terms of the percentage of nodes alive per amount of effective data packets received at the sink, which is shown in Figure 5.

From Figure 5, we find that the proposed multi-hop MIMO-LEACH scheme needs significantly less energy to transmit the same amount of data packets. Therefore, the

| Table 1: The system parameters. |
|--------------------------------|
| $\alpha = 0.4706$            |
| $M_l = 40$ dB                |
| $G_l = 30$ dB                |
| $k \in [3, 5]$               |
| $\sigma^2 = \frac{N_f}{2} = -134$ dBm/Hz |
| $N_f = 10$ dB                |
| $f_c = 2.5$ GHz              |
| $B = 20$ kHz                 |
| $P_b = 10^{-3}$              |
| $P_{ct} = 98.2$ mw           |
| $P_{cr} = 112.6$ mw          |
| $F_n = 2$                    |
| $G_tG_r = 5$ dB              |
| $s = 2$ kbits                |
| $P = 0.8$                    |
| $R = 0.75$                   |
| $F = 200$                    |
| $p = 2$                      |
| $R_{tr} = 5$                 |
| $R_{ts} = 100$               |
improvement on network lifetime obtained by the multi-hop MIMO-LEACH scheme is significant.

On the other hand, the impacts of the parameters, including the number of cluster heads \( k_c \) and the number of cooperative nodes \( J \), are also investigated in the simulation. Figures 6 and 7 show the percentage of nodes alive over time in different settings of \( k_c \) and \( J \).

From the simulation results including those shown in Figures 6 and 7, we can find that the energy saving performance of the proposed scheme is impacted by the parameters. As for the number of cluster heads, too many cluster heads will reduce the distance for each single hop transmission, which will reduce the transmit energy consumption. More cluster heads will also generate a larger search space.
for the routing table construction, which will also reduce the transmit energy consumption further. However, more cluster heads will result in more number of hops in transmission to the sink, which will consume more circuit energy for relaying the data packets. Therefore, the number of cluster heads should be chosen by trading off the transmit energy consumption and circuit energy consumption. As for the number of cooperative nodes, a certain number of cooperative nodes can form the effective independent multi-path transmission so as to energy-efficiently combat the fading effects. However, too many cooperative nodes will result in large circuit energy consumption, which will cause large overall energy consumption. Therefore, the number of cooperative nodes should also be chosen to trade off the transmit energy consumption and the circuit energy consumption.

5. CONCLUSION

In this paper, we proposed a cluster based cooperative MIMO scheme to reduce energy consumption and prolong the network lifetime. A cooperative MIMO scheme is adopted to mitigate the adverse impacts of fading while clustering is used to facilitate network control and coordination. In the proposed scheme, the original LEACH protocol is extended by incorporating the cooperative MIMO communications and multi-hop routing. An adaptive cooperative nodes selection strategy is also designed. Based on the scheme, we investigated the energy consumption of each operation. Then, the overall energy consumption model of the scheme is developed, and the optimal parameters of the scheme are found such as the number of clusters and the number of cooperative nodes. Simulation results exhibit that the proposed scheme minimizes energy consumption.

ACKNOWLEDGMENTS

The authors thank the editors and the anonymous reviewers for their valuable suggestions. This work was supported in part by KOSEF Grant no. R01-2004-000-10372-0.

REFERENCES

[1] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, “Space-time block codes from orthogonal designs,” IEEE Transactions on Information Theory, vol. 45, no. 5, pp. 1456–1467, 1999.
[2] S. Cui, A. J. Goldsmith, and A. Bahai, “Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks,” IEEE Journal on Selected Areas in Communications, vol. 22, no. 6, pp. 1089–1098, 2004.
[3] X. Li, “Energy efficient wireless sensor networks with transmission diversity,” IEEE Electronics Letters, vol. 39, no. 24, pp. 1753–1755, 2003.
[4] X. Li, M. Chen, and W. Liu, “Application of STBC-encoded cooperative transmissions in wireless sensor networks,” IEEE Signal Processing Letters, vol. 12, no. 2, pp. 134–137, 2005.
[5] S. K. Jayaweera, “Energy analysis of MIMO techniques in wireless sensor networks,” in Proceedings of 38th Annual Conference on Information Sciences and Systems (CISS ‘04), Princeton, NJ, USA, March 2004.
[6] S. Cui, A. J. Goldsmith, and A. Bahai, “Energy-constrained modulation optimization,” IEEE Transactions on Wireless Communications, vol. 4, no. 5, pp. 2349–2360, 2005.
[7] C. S. Park and K. B. Lee, “Transmit power allocation for BER performance improvement in multicarrier systems,” IEEE Transactions on Communications, vol. 52, no. 10, pp. 1658–1663, 2004.
[8] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, “An application-specific protocol architecture for wireless microsensor networks,” IEEE Transactions on Wireless Communications, vol. 1, no. 4, pp. 660–670, 2002.
[9] Y. Yu, B. Krishnamachari, and V. K. Prasanna, “Energy-latency tradeoffs for data gathering in wireless sensor networks,” in Proceedings of 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '04), vol. 1, pp. 244–255, Hong Kong, March 2004.

Yong Yuan received the B.E. and M.E. degrees from the Department of Electronics and Information, Yunnan University, Kunming, China, in 1999 and 2002, respectively. Since 2002, he has been studying at the Department of Electronics and Information, Huazhong University of Science and Technology, China, as a Ph.D. candidate. His current research interests include wireless sensor network, wireless ad hoc network, wireless communication, and signal processing.

Min Chen was born on December 1980. He received the BS, MS, and Ph.D degrees from the Department of Electronic Engineering, South China University of Technology, in 1999, 2001, and 2004, respectively. He is a postdoctoral fellow in the Communications Group, Department of Electrical and Computer Engineering, University of British Columbia. He was a postdoctoral Researcher in the Multimedia & Mobile Communications Lab., School of Computer Science and Engineering, Seoul National University, in 2004 and 2005. His current research interests include wireless sensor network, wireless ad hoc network, and video transmission over wireless networks.

Taekyoung Kwon is an Assistant Professor in the School of Computer Science and Engineering, Seoul National University (SNU), since 2004. Before joining SNU, he was a postdoctoral Research Associate at UCLA and at City University New York (CUNY). He obtained the B.S., M.S., and Ph.D. degrees from the Department of Computer Engineering, SNU, in 1993, 1995, 2000, respectively. During his graduate program, he was a visiting student at IBM T. J. Watson Research Center and at the University of North Texas. His research interests lie in sensor networks, wireless networks, IP mobility, and ubiquitous computing.
Relay Techniques for MIMO Wireless Networks with Multiple Source and Destination Pairs

Tetsushi Abe,¹ Hui Shi,² Takahiro Asai,² and Hitoshi Yoshino²

¹DoCoMo Communications Laboratories Europe GmbH, 312 Landsbergerstreet, Munich 80687, Germany
²NTT DoCoMo, Inc., Japan

Received 1 November 2005; Revised 17 May 2006; Accepted 16 August 2006

A multiple-input multiple-output (MIMO) relay network comprises source, relay, and destination nodes, each of which is equipped with multiple antennas. In a previous work, we proposed a MIMO relay scheme for a relay network with a single source and destination pair in which each of the multiple relay nodes performs QR decompositions of the backward and forward channel matrices in conjunction with phase control (QR-P-QR). In this paper, we extend this scheme to a MIMO relay network employing multiple source and destination pairs. Towards this goal, we use a *group nulling* approach to decompose a multiple S-D MIMO relay channel into parallel independent S-D MIMO relay channels, and then apply the QR-P-QR scheme to each of the decomposed MIMO relay links. We analytically show the logarithmic capacity scaling of the proposed relay scheme. Numerical examples confirm that the proposed relay scheme offers higher capacity than existing relay schemes.

Copyright © 2006 Tetsushi Abe et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

A wireless network comprises a number of nodes connected by wireless channels. Using internode transmission (relaying) is an important technique to widen network coverage. Network information theory has shown that the use of multiple relay nodes in source and destination (S-D) communications increases the capacity of the S-D system logarithmically with the number of relay nodes [1].

The use of multiple antennas at each node provides additional degrees of freedom to improve further the capacity per S-D pair in the relay network. A significant capacity improvement achieved with multiple-input multiple-output (MIMO) transmission was revealed in [2–5] for a point-to-point wireless link, and in [6–9] for multiple-access and broadcast channels. The capacity bounds of the MIMO relay network have recently been derived in [10, 11] where the capacity of the MIMO relay network was analyzed in terms of *distributed array gain*, which offers logarithmic capacity scaling, *spatial multiplexing gain*, and *receive array gain*. In [12], we proposed a MIMO relay scheme for a relay network comprising a single S-D pair and multiple relay nodes. The relay technique in [12], called QR-P-QR, performs the QR decomposition (QRD) in the backward and forward channels in conjunction with employing phase control at each relay node, and successive interference cancellation (SIC) at the destination node to detect multiple data streams. This architecture achieves both *distributed array gain* and *receive array gain* while maintaining the maximum *spatial multiplexing gain*, which leads to higher capacity than the existing zero-forcing (ZF) and amplify and forward (AF) relaying techniques [11].

In this paper, we consider a relay network of multiple S-D pairs and multiple relay nodes, and provide a new relaying technique. The proposed relay architecture employs (1) a *group nulling* (GN) technique, which is applied to the backward and forward MIMO relay channels to decompose the multiple S-D MIMO relay channel into parallel independent S-D MIMO relay channels, and (2) the QR-P-QR scheme, which is applied to each of the decomposed S-D relay links. The *group nulling* technique separates multiple S-D pairs via unitary transforms that project both received and transmitted signal vectors at a relay node onto the null space of the signals of undesired S-D pairs. Thus, the *group nulling* technique retains a higher degree of freedom than the ZF-based stream-wise nulling in MIMO relay channels. Furthermore, the QR-P-QR scheme achieves both *distributed array gain* and *receive array gain* while maintaining the maximum *spatial multiplexing gain* at each of the decomposed MIMO relay links. We analyze the asymptotic capacity of the proposed relay technique and through numerical examples show that the proposed relay...
technique achieves higher capacity than other existing relay schemes.

The rest of this paper is organized as follows. Section 2 shows a system model and the upper bound for the capacity of the MIMO relay network. We describe the proposed and existing relay schemes in Section 3. Numerical examples are given in Section 4. Finally, Section 5 concludes this paper.

Notation

\(E\{\cdot\}\) and \(\text{tr}\{\cdot\}\) denote the expectation and trace operation, respectively. \(\|a\|\) stands for the norm of vector \(a\), and superscripts \(T, H, \) and \(\ast\) represent the transpose, the conjugate transpose, and the conjugate operation, respectively. \((\mathbf{A})_{i,j}\) and \((\mathbf{A})_{i}\) denote the \(i\)th row and \((i, j)\)th entry of matrix \(\mathbf{A}\), respectively. \(\mathbf{I}\) is the \(i \times i\) identity matrix.

2. MIMO RELAY NETWORK

The MIMO relay network used in this paper is illustrated in Figure 1. This paper assumes a one-hop relay network comprising \(L\) source and destination nodes, each of which has \(M\) antennas, and \(K\) relay nodes, each of which has \(N\) antennas. In addition, we assume that the relay nodes do not transmit and receive simultaneously. In other words, two time slots are required to send a message from the source to the destination as shown in Figure 1.

First, \(M \times 1\) vector \(s_l (l = 1, \ldots, L)\), destined for the \(l\)th destination node, is sent to all relay nodes from the \(l\)th source node without using any channel state information (CSI). The \(N \times 1\) vector received at the \(k\)th relay node is expressed as \(y_k = \sum_{l=1}^{L} \mathbf{H}_{k,l} s_l + \mathbf{n}_k\), where \(\mathbf{H}_{k,l} (k = 1, \ldots, K)\) is the \(N \times M\) MIMO channel matrix between the \(l\)th source node and the \(k\)th relay node (backward channel), and \(\mathbf{n}_k\) refers to the \(N \times 1\) noise vector at the \(k\)th relay node with zero mean and covariance matrix \(\mathbf{E}\{\mathbf{n}_k\mathbf{n}_k^H\} = \sigma_k^2 \mathbf{I}_N\). We constrain the transmitted signal power at the source node to \(\mathbf{E}\{s_l s_l^H\} = \frac{P}{M} \mathbf{I}_M\), where \(P\) is the total transmit power. A relay operation is performed at the \(k\)th relay node by using \(N \times N\) relay matrix \(\mathbf{W}_k\) to obtain \(N \times 1\) transmitted signal vector \(\mathbf{x}_k = E_k \mathbf{W}_k y_k\), where \(E_k\) is a power coefficient resulting from total power constraint \(\mathbf{E}\{\mathbf{x}_k\mathbf{x}_k^H\} = P\). This can be expressed as

\[
E_k = \sqrt{\frac{P}{\text{tr}\left\{ (\mathbf{W}_k \mathbf{H}_k)^H (\mathbf{W}_k \mathbf{H}_k) \right\} + M \sigma_k^2 \text{tr}\left\{ (\mathbf{W}_k)^H (\mathbf{W}_k) \right\}}}.
\]

where \(N \times LM\) matrix \(\mathbf{H}_k = [\mathbf{H}_{k,1}, \ldots, \mathbf{H}_{k,L}]\). Finally, the \(M \times 1\) receive vector given by

\[
r_l = \sum_{k=1}^{K} \mathbf{G}_{k,l} \mathbf{x}_k + \mathbf{z}_l
\]

is obtained at the \(l\)th destination node, where \(\mathbf{G}_{k,l}\) and \(\mathbf{z}_l\) are the \(M \times N\) channel matrix between the \(k\)th relay node and the \(l\)th destination node (forward channel), and the \(M \times 1\) noise vector added at the \(l\)th destination node with zero mean and covariance matrix \(\mathbf{E}\{\mathbf{z}_l \mathbf{z}_l^H\} = \sigma_l^2 \mathbf{I}_M\), respectively.

Using the cut-set theorem [13], the upper bound for the capacity of the MIMO relay network is derived in [10] as

\[
C_{\text{upper}} = \mathbf{E}_{\{\mathbf{H}_k\}} \left\{ \frac{1}{2} \log \left[ \det \left( \mathbf{I}_L + \frac{P}{M \sigma_k^2} \sum_{k=1}^{K} \mathbf{H}_k^H \mathbf{H}_k \right) \right] \right\}.
\]

3. MIMO RELAY TECHNIQUES

In this paper, we assume that each node knows the CSI of its own backward and forward channels. However, we do not allow source nodes, relay nodes, and destination nodes to exchange their CSI with other nodes.

3.1. ZF RELAYING SCHEME [11]

The ZF relaying scheme computes backward and forward ZF matrices \(\mathbf{H}_k^*_k\) and \(\mathbf{G}_k^*_k\) that satisfy \(\mathbf{H}_k^*_k \mathbf{H}_k^* = \mathbf{I}_{LM}\) and \(\mathbf{G}_k^*_k \mathbf{G}_k^* = \mathbf{I}_{LM}\) with \(LM \times N\) matrix \(\mathbf{G}_k^* = [\mathbf{G}_{k,1}^T, \ldots, \mathbf{G}_{k,L}^T]^T\). Relay matrix \(\mathbf{W}_k\) for the ZF scheme is then written as \(\mathbf{W}_k = \mathbf{G}_k^* \mathbf{H}_k^*\).
Note here that the ZF scheme requires that \( N \geq LM \). In this case, the effective signal-to-noise ratio (SNR) for the \( m \)th data stream, \( \lambda_{m}^{ZF} \), at the \( l \)th destination node is

\[
\lambda_{m}^{ZF} = \frac{(P/M)\left(\sum_{k=1}^{K} E_{k}\right)^{2}}{\sigma_{d}^{2} \left(\sum_{k=1}^{K} E_{k}|| (H^{l}_{k})_{m}||^{2} \right) + \sigma_{d}^{2}}.
\]  

(4)

From (4), we find that due to the transmit and receive ZF operations the signals from \( K \) relay nodes are coherently combined at the destination node, which leads to distributed array gain [11].

3.2. GN/QR-P-QR relaying scheme

The first step of the GN/QR-P-QR scheme is to compute a pre-group nulling filter at a relay node to suppress the signal component from all source nodes except from the \( l \)th source node. To accomplish this, we define \( N \times M(L-1) \) matrix \( H_{k}^{(l)} \equiv [H_{k,1}^{(l)}, \ldots, H_{k,l-1}^{(l)}, H_{k,l+1}^{(l)}, \ldots, H_{k,L}^{(l)}] \). Note that the channel matrix between the \( l \)th source and \( k \)th relay node, \( H_{k,l} \), is removed. Next, we perform the singular value decomposition (SVD) of \( H_{k}^{(l)} \) as

\[
H_{k}^{(l)} = \begin{bmatrix}
U_{k,1}^{(l)} & \ldots & U_{k,l-1}^{(l)} & U_{k,l}^{(l)}
\end{bmatrix}
\begin{bmatrix}
\Lambda_{k,1}^{(l)} & & & \\
& \ddots & & \\
& & \Lambda_{k,l-1}^{(l)} & \\
& & & 1
\end{bmatrix}
\begin{bmatrix}
V_{k,1}^{(l)H} \\
\vdots \\
V_{k,l-1}^{(l)H} \\
O
\end{bmatrix},
\]  

(5)

where \( M \times M \) matrices \( \Lambda_{k,1}^{(l)}, \ldots, \Lambda_{k,l-1}^{(l)} \) are diagonal matrices, and \( N \times M \) matrices \( U_{k,1}^{(l)}, \ldots, U_{k,l-1}^{(l)} \) and \( M(L-1) \times M \) matrices \( V_{k,1}^{(l)}, \ldots, V_{k,l-1}^{(l)} \) have orthonormal columns. \( N \times N-M(L-1) \) matrix \( U_{k,l}^{(l)H} \) spans the null space of \( H_{k}^{(l)} \). Matrix \( U_{k,l}^{(l)} \) is then multiplied to \( y_{k,l} \) to obtain \( N-M(L-1) \times 1 \) vector \( y_{k,l} \) as

\[
y_{k,l} = U_{k,l}^{(l)H} y_{k,l} = U_{k,l}^{(l)H} H_{k,l} s_{l} + U_{k,l}^{(l)H} n_{k,l}.
\]  

(6)

From (6), we see that \( U_{k,l}^{(l)} \) removes the signal contribution from all source nodes except that from the \( l \)th source node due to the projection of the received signal vector onto the null space of undesired source nodes. A null space-based method was also employed in [14] for the precoding in a MIMO down link transmission.

The second step of the GN/QR-P-QR scheme is the transformation of \( y_{k,l} \) using \( N-M(L-1) \times N-M(L-1) \) matrix \( \Phi_{k,l} \) to obtain vector \( y'_{k,l} = \Phi_{k,l} U_{k,l}^{(l)H} y_{k,l} \). The computation of \( \Phi_{k,l} \) will be described later in this section.

The third step is to compute the post-group nulling filter to suppress the transmitted signal to all destination nodes except that to the \( l \)th destination node. Toward this goal, we define \( N \times M(L-1) \) matrix \( G_{k}^{(l)} \equiv [G_{k,1}^{l}, \ldots, G_{k,l-1}^{l}, G_{k,l+1}^{l}, \ldots, G_{k,L}^{l}] \). Next, we perform the SVD of \( G_{k}^{(l)} \) as

\[
G_{k}^{(l)} = \begin{bmatrix}
A_{k,1}^{(l)} & \ldots & A_{k,l-1}^{(l)} & A_{k,l}^{(l)}
\end{bmatrix}
\begin{bmatrix}
\Omega_{k,1}^{(l)} & & & 0 \\
& \ddots & & \\
& & \Omega_{k,l-1}^{(l)} & \\
& & & \Omega_{k,l}^{(l)}
\end{bmatrix}
\begin{bmatrix}
B_{k,1}^{(l)H} \\
\vdots \\
B_{k,l-1}^{(l)H} \\
0
\end{bmatrix},
\]  

(7)

where \( M \times M \) matrices \( \Omega_{k,1}^{(l)}, \ldots, \Omega_{k,l}^{(l)} \) are diagonal matrices, and \( N \times M \) matrices \( A_{k,1}^{(l)}, \ldots, A_{k,l-1}^{(l)} \) and \( M(L-1) \times M \) matrices \( B_{k,1}^{(l)}, \ldots, B_{k,l}^{(l)} \) have orthonormal columns. \( N \times N-M(L-1) \) matrix \( A_{k,l}^{(l)} \) spans the null space of \( G_{k}^{(l)} \). Matrix \( A_{k,l}^{(l)} \) is then multiplied to \( y'_{k,l} \) to obtain \( N \times 1 \) vector \( y''_{k,l} = A_{k,l}^{(l)} Q_{k,l} U_{k,l}^{(l)H} y_{k,l} \). Note here that similar to the ZF scheme, the group nulling scheme also requires that \( N \geq LM \) in order to obtain null space matrices \( U_{k,l}^{(l)} \) and \( A_{k,l}^{(l)} \).

The above three-step procedure is performed for all \( L \) source and destination pairs \( (l = 1, \ldots, L) \) at the \( k \)th relay node. Finally, the \( N \times 1 \) signal vector transmitted from the \( k \)th relay node is \( x_{k} = E_{k} \sum_{l=1}^{L} y''_{k,l} = E_{k} \sum_{l=1}^{L} A_{k,l}^{(l)} Q_{k,l} U_{k,l}^{(l)H} y_{k,l} \). In this case, the relaying matrix is written as \( W_{k} = \sum_{l=1}^{L} A_{k,l}^{(l)} Q_{k,l} U_{k,l}^{(l)H} \), and the received signal vector at the \( l \)th destination is written from (2) as

\[
r_{l} = \sum_{k=1}^{K} E_{k} G_{k,l} A_{k,l}^{(l)} Q_{k,l} U_{k,l}^{(l)H} H_{k,l} s_{l} + \sum_{k=1}^{K} E_{k} G_{k,l} A_{k,l}^{(l)} Q_{k,l} U_{k,l}^{(l)H} n_{k,l} + z_{l}.
\]  

(8)

Equation (8) shows that at the \( l \)th destination node, the signal contribution from all source nodes is removed except from the \( l \)th source node. Namely, we can establish an independent MIMO relay link between the \( l \)th source and destination nodes that is characterized by \( M \times M \) MIMO channel matrix \( E_{k} G_{k,l} A_{k,l}^{(l)} Q_{k,l} U_{k,l}^{(l)H} H_{k,l} \).

To compute the intermediate filter \( Q_{k,l} \), we use the QR-P-QR scheme [12]. The QR-P-QR relaying scheme first performs the QRD of \( N-M(L-1) \times M \) matrices \( U_{k,l}^{(l)H} H_{k,l} \) and \( (G_{k,l} A_{k,l}^{(l)})^{H} \) as \( U_{k,l}^{(l)H} H_{k,l} = Q_{ik,l} R_{ik,l} \) and \( (G_{k,l} A_{k,l}^{(l)})^{H} = Q_{2k,l} R_{2k,l} \), where \( N-M(L-1) \times M \) matrices \( Q_{ik,l} \) and \( Q_{2k,l} \) have orthonormal columns, and \( M \times M \) matrices \( R_{ik,l} \) and \( R_{2k,l} \) are upper triangular matrices. By using these results, the intermediate filter is computed as \( Q_{k,l} = Q_{2k,l} D_{k,l}^{-1} \) \( Q_{ik,l}^{H} \) where the \( M \times M \) matrix, \( D_{k,l} \), is a diagonal matrix whose \( m \)th diagonal entry is \( d_{k,l,m} = (R_{ik,l}^{H} \Pi_{R_{ik,l}})_{m,m-1} ||(R_{2k,l}^{H} \Pi_{R_{2k,l}})_{m,M-1} || \) and \( \Pi_{R} \) is an \( M \times M \) exchange matrix (see [12] for details). We can see that \( Q_{k,l} \) consists of two orthogonal matrices, \( Q_{ik,l} \) and \( Q_{2k,l} \), obtained by the QRD in the backward and forward channels with phase control matrix \( D_{k,l} \) in between (for this reason this scheme is called QR-P(Phase)-QR). Finally, by using the
computed $\Phi_{k,l}$, (8) is rewritten as
\[
\mathbf{r}_{l} = \sum_{k=1}^{K} E_k \mathbf{R}_k^H \mathbf{D}_{k,l} \mathbf{R}_{ik,l} \mathbf{s}_l + \sum_{k=1}^{K} E_k \mathbf{R}_k^H \mathbf{D}_{k,l} \mathbf{Q}_{ik,l} \mathbf{U}_{k,l}^H \mathbf{n}_k + \mathbf{z}_l.
\]

An important note here is that $E_k \mathbf{R}_k^H \mathbf{D}_{k,l} \mathbf{R}_{ik,l}$ takes the lower triangular form with positive scalars in diagonal entries. The triangular structure provides the receive array gain by using the SIC at the destination node to detect each data stream. The positive diagonal entries achieved by the phase control matrix enable the diagonal elements transmitted from $K$ relay nodes to be coherently combined at the destination node, which obtains the distributed array gain.

The $l$th destination node simply performs SIC by using the CSI of compound triangular channel $\sum_{k=1}^{K} E_k \mathbf{R}_k^H \mathbf{D}_{k,l} \mathbf{R}_{ik,l}$ to detect each of the multiple streams. The effective signal-to-interference-plus-noise ratio (SINR) for the $m$th data stream at the $l$th destination node can be expressed as
\[
\lambda^Q_{lm} = \frac{\left( P/M \right) \left( \sum_{k=1}^{K} E_k \mathbf{R}_k^H \mathbf{D}_{k,l} \mathbf{R}_{ik,l} \right)_{m,M-m+1}}{\sigma^2 \left( \sum_{k=1}^{K} \mathbf{E}_k \left( \mathbf{R}_k^H \mathbf{D}_{k,l} \right)_{m,m} \right) + \sigma_d^2}.
\]

Consequently, the ergodic capacity of the relay network with total $L$ S-D pairs is
\[
C_{QR} = E\{H_{l,m}\} \left\{ \frac{1}{2} \sum_{l=1}^{L} \sum_{m=1}^{M} \log_2 \left( 1 + \lambda^Q_{lm} \right) \right\}.
\]

### 3.3. Achievable gains in the relay schemes

To evaluate the achievable gains of the GN/QR-P-QR relay technique, we investigate its asymptotic capacity when $K$ approaches infinity. From (10) and (11), when $K$ approaches infinity, the capacity becomes
\[
C_{QR} = \frac{1}{2} \sum_{l=1}^{L} \sum_{m=1}^{M} \log_2 \left( 1 + \left( P/K \right) \left( \sum_{k=1}^{K} E_k \mathbf{R}_k^H \mathbf{D}_{k,l} \mathbf{R}_{ik,l} \right)_{m,M-m+1} \right)
\]
\[
+ \sigma^2 \left( \sum_{k=1}^{K} \mathbf{E}_k \left( \mathbf{R}_k^H \mathbf{D}_{k,l} \right)_{m,m} \right) + \left( 1/K \right) \sigma_d^2
\]
\[
- \frac{ML}{2} \log_2 (K)
\]
\[
+ \frac{1}{2} \sum_{l=1}^{L} \sum_{m=1}^{M} \log_2 \left( \frac{\left( P/M \right) \left( \sum_{k=1}^{K} E_k \mathbf{R}_k^H \mathbf{D}_{k,l} \mathbf{R}_{ik,l} \right)_{m,M-m+1}}{\sigma^2 \mathbf{E}_k \left( \mathbf{R}_k^H \mathbf{D}_{k,l} \right)_{m,m} + \sigma_d^2} \right)^2.
\]

where we use the approximation $\log_2 (1+x) \approx \log_2 x (x \gg 1)$.

From (12), we see that the capacity of the GN/QR-P-QR scheme scales with $(LM/2) \log_2 (K)$ asymptotically in $K$. The term $\log_2 (K)$ indicates that the distributed array gain of the GN/QR-P-QR scheme is $K$. In addition, the prelog term $LM/2$ implies that the multiplexing gain is $LM/2$, where $1/2$ represents the loss when using two time slots in each transmission. Furthermore, it was shown in [11] that the upper bound of the capacity in (3) and the capacity of the ZF scheme asymptotically scale with $(LM/2) \log_2 (K)$. Thus, we see that the GN/QR-P-QR scheme as well as the ZF scheme exhibit the optimum capacity scaling for a large $K$ value.

The difference between the GN/QR-P-QR scheme and the ZF scheme is the available degrees of freedom remaining after interference suppression among multiple S-D pairs. The ZF scheme performs complete stream-wise nulling in both the backward and forward channels. At each channel the ZF scheme separates $LM$ streams, which requires $LM - 1$ degrees of freedom. Thus, the degrees of freedom that remain after the ZF relaying are $N = (LM - 1)$. On the other hand, since the proposed scheme performs group-wise nulling, it preserves a higher degree of freedom than the ZF scheme. To be more specific, we define the $N = (LM - 1) \times M$ decomposed forward MIMO channel for the $l$th S-D pair from (6) as $\mathbf{H}_{kl} \equiv \mathbf{U}_{k,l}^H \mathbf{H}_{k,l}$. Assuming $(\mathbf{H}_{k,l})_{ij}$ is i.i.d. complex random variables with zero mean and unit variance, $(\mathbf{H}_{k,l})_{ij}$ has the following statistical property:
\[
E\{\mathbf{H}_{k,l}^* (\mathbf{H}_{k,l})_{i,j}\} = \begin{cases} 1, & i = i', j = j', \\ 0, & \text{otherwise} \end{cases}
\]

**Proof.** When $i = i'$ and $j = j'$, $E\{\mathbf{H}_{k,l}^* (\mathbf{H}_{k,l})_{i,j}\} = 1$ because $E\{\mathbf{H}_{k,l}^* \mathbf{H}_{k,l}\} = \mathbf{I}_M$ and the norm of each column in $\mathbf{U}_{k,l}^H$ is one. When $i = i'$ and $j \neq j'$, $E\{\mathbf{H}_{k,l}^* (\mathbf{H}_{k,l})_{i,j}\} = 0$ because $(\mathbf{H}_{k,l})_{ij}$ are mutually uncorrelated. When $i \neq i'$ and $j = j'$, $E\{\mathbf{H}_{k,l}^* (\mathbf{H}_{k,l})_{i,j}\} = 0$ because the columns of $\mathbf{U}_{k,l}^H$ are mutually orthogonal. Equation (13) is then proven.

We can see from (6) and (13) that the group nulling transforms $N \times M$ i.i.d. matrix $\mathbf{H}_{k,l}$ to an $N - (LM - 1) \times M$ i.i.d. matrix $\mathbf{H}_{k,l}$. This shows that due to the group nulling, $M(L-1)$ degrees of freedom are lost for the $l$th S-D pair, but $\mathbf{H}_{k,l}$ still holds $N - M(L-1)$ degrees of freedom. Furthermore, it is straightforward that the same discussion holds for the backward decomposed channel $\mathbf{G}_{k,l} \mathbf{A}_{k,l}^H$. Thus, after the group nulling operations, the proposed scheme holds $N - M(L-1)$ degrees of freedom, which are higher than that of ZF by $M - 1$. This additional degree of freedom is converted as the receive array gain through the channel triangulation in (9) using the QR-P-QR technique and the following SIC at the destination node.

### 3.4. Other simple schemes

For GN-based relaying, we could simply employ an AF relay scheme instead of the QR-P-QR scheme, which gives the intermediate filter $\Phi_{k,l} = \mathbf{I}_{N - M(L-1)}$. In this case, however, we cannot obtain the distributed array gain because signals from $K$ relay nodes are randomly combined at the destination.
node. In addition, [10, 15] describe another simple matched filter (MF) relaying scheme in which each relay node performs receive and transmit MF operations. For the MF relaying, the relay matrix is expressed as $W_k = G_k^H H_k^H$. Unlike the ZF and the proposed schemes, this scheme does not require that $N \geq LM$, and the capacity still scales logarithmically with the number of relay nodes [10].

4. NUMERICAL RESULTS

The ergodic capacities of the relaying schemes presented in the previous section were evaluated. We obtained the capacity plots of the upper bound, ZF, GN/QR-P-QR, GN/AF, and MF. In addition, we evaluated as a reference the capacity of QR-P-QR when all relay and destination nodes fully cooperate. To be more specific, we calculated the capacity of the QR-P-QR scheme in a network comprising a source node with $LM$ transmit antennas, a relay node with $KN$ antennas, and a destination node with $LM$ antennas. In this case, the power constraints at the source and relay are $LP$ and $KP$, respectively. We assumed a flat fading channel in which each component of $H_k$ and $G_k$ is an i.i.d. complex random variable with zero mean and unit variance. We set $\sigma^2 = \sigma^2_0$ and identical transmit power $P$ for all source and relay nodes. We did not take into account path loss.

4.1. Capacity versus the number of relay nodes

Figure 2 shows the capacity versus the number of relay nodes $K$ for $L = 2, M = 4$, and $N = 8$. The total transmit power-to-noise ratio (PNR) $P/\sigma^2$ was set to 20 dB. The graph shows that the capacity of the GN/AF scheme is saturated when $K$ becomes large. This is because although the separation of multiple S-D pairs is accomplished by the group nulling, the signals relayed from multiple relay nodes are randomly combined at each destination node due to the simple AF relay operation, and thus the distributed array gain is not obtained. On the other hand, we can see that the GN/QR-P-QR scheme, ZF scheme, and MF scheme exhibit logarithmic capacity scaling as does the upper bound of the capacity. This is due to the fact that signal components from multiple relay nodes are coherently combined at the destination node. Furthermore, the GN/QR-P-QR scheme offers higher capacity than the ZF scheme due to the higher degree of freedom converted to the receive array gain at the destination node as described in Section 3.3. The capacity of the MF scheme is lower than that of the others due to its inability to suppress actively the interference among S-D pairs. The capacity gap between GN/QR-P-QR and the upper bound is due to the imperfect cooperation among nodes. As mentioned in [10], the capacity upper bound in (3) can be achieved if all the relay nodes perform joint decoding and encoding. To examine this, we obtained the capacity of QR-P-QR when all the relay nodes and all destination nodes cooperate. Note that in this case, there is no need for GN. We can see that the capacity of the QR-P-QR scheme with perfect node cooperation approaches the upper bound. Furthermore, when $K$ becomes larger the gap between the two becomes narrower. This can be briefly explained as follows. The capacity upper bound in (3) only depends on the backward channel. On the other hand, the capacity expressions of QR-P-QR in (10) with (11) show that the noise power at destination node $\sigma^2_d$ becomes less significant when $K$ becomes large. Thus, the capacity depends more on the backward channel and thus approaches closer to the upper bound. Therefore, if we allow relay nodes to perform the joint relay operation, we could approach closer to the bound. However, this requires all relay nodes and all the destination nodes to exchange their CSI. In addition, the joint relay operation requires the QRD of $KN \times LM$ matrix, which might be practically demanding in terms of complexity. Figure 3 shows capacity plots for $L = 2$, $M = 2$, and $N = 4$. A similar tendency is observed, but the gap between GN/QR-P-QR and ZF is decreased. This is because the number of antennas at each node is reduced by half, and thus the receive array gain obtained in the GN/QR-P-QR scheme is decreased. Figure 4 shows capacity plots for $L = 4$, $M = 2$, and $N = 8$. In this case, the total number of antennas in the network is the same as in the case in Figure 2, but the capacity obtained by each relay scheme is higher than that in Figure 2 except for MF. This is because the total transmit power in the network is increased due to the increased number of the S-D pairs.

4.2. Capacity versus PNR

Figures 5 and 6 show the capacity versus the PNR for $L = 2$, $M = 4$, and $N = 8$ for $K = 2$ and 8, respectively. The figures show that the GN/QR-P-QR and the GN/AF schemes offer similar capacity for $K = 2$. However, Figure 6 shows that when $K = 8$, GN/QR-P-QR outperforms GN/AF due to the distributed array gain. In both figures, the capacity of the
MF scheme is better than the other schemes in a low PNR region due to the SNR gain of the matched filtering. However, the capacity saturates in a high PNR region due to the interference among S-D pairs.

### 4.3. Effectiveness of spatially multiplexing multiple S-D pairs

Figure 7 shows the capacity curves of the GN/QR-P-QR scheme for \( L = 2, M = 4, \) and \( N = 8 \) with \( K = 2 \) and 8. Here, we measured the capacity for two cases: time-division multiplexing (TDM) and spatial-division multiplexing (SDM) for the two S-D pairs. Note that in the former case, only one S-D pair is active at any instant, and thus group nulling is not needed. Figure 7 shows that in a low PNR region, TDM provides higher capacity, but in higher PNR regions, SDM offers significantly higher capacity, which matches results of conventional studies on the trade-off between spatial multiplexing and beam-forming. Furthermore, the figure shows that when \( K \) increases, the crosspoint of SDM and TDM is shifted to lower PNR regions. This is because the effective SNR at the destination node increases as \( K \) increases. Thus, it is clear that it is more advantageous to multiplex spatially multiple S-D pairs in a situation, where the PNR is relatively high or the number of relay nodes is relatively large.

### 4.4. Capacity versus the number of antennas at the relay node

Figure 8 shows the capacity of the GN/QR-P-QR and the ZF schemes with various \( N \) for \( L = 2 \) and \( M = 4 \). \( K \) is set to 2 and 8. We can see that when the number of antennas per relay node, \( N \), increases, the capacity gap between the GN/QR-P-QR and the ZF schemes becomes smaller. This is because as \( N \) becomes larger, both the GN and the ZF operations retain enough degrees of freedom after the interference suppression as shown in Section 3.3.

### 4.5. Complexity

Finally, Table 1 shows the computational complexity of the relaying schemes. The complexities were measured as the
number of required complex multiplications at each relay node. We approximated the complexity by computing only matrix inversion, multiplication, SVD, and QRD parts and evaluated only terms with the highest order (cubic) in terms of matrix size. First, we observe that the complexity of the MF scheme is much lower than that of others due to its simple operations. The ZF scheme needs only one matrix inversion for both the backward and forward channel matrices ($H_k^{(l)}$ and $G_k^{(l)}$), but the matrix size $N \times M(L - 1)$ is smaller than that in ZF. The GN/QR-P-QR scheme further requires QRD for every S-D pair of both equivalent backward and forward channels $U_{k,l}^{(l)}H_k^{(l)}$ and $(G_k^{(l)}U_{k,l}^{(l)})^H$, and their matrix size, $N - M(L - 1) \times M$, is smaller than that in ZF. Thus, when the number of S-D pairs is small, such as when $(L,M,N) = (2,2,4)$ and $(2,4,8)$, the GN-based relay schemes offer lower complexity than the ZF due to the matrix size reduction. On the other hand, when the number of S-D pairs becomes larger, such as when $(L,M,N) = (4,2,8)$, the ZF scheme offers lower complexity due to fewer matrix operations. Therefore, when the number of S-D pairs is small, the GN/QR-P-QR scheme achieves higher capacity with lower complexity than the ZF scheme.

5. CONCLUDING REMARKS

In this paper, we proposed a relay technique for a MIMO relay network with multiple S-D pairs. The group nulling technique projects the receive and transmitted signal vectors at the relay node onto the null space of the signals of nondesired S-D pairs, so the multiple S-D MIMO relay channel is decomposed into parallel independent MIMO channels. To each decomposed MIMO relay link, the QR-P-QR technique is applied. This relaying architecture preserves a higher degree of freedom in the MIMO relay channel than the ZF scheme and enables coherent combination of the signals at the destination to achieve distributed array gain. We analyzed the asymptotic capacity of the proposed relay technique and clarified its achievable gains. Numerical examples confirmed that the proposed relay scheme achieves higher capacity than other existing relay schemes. It should be mentioned,
however, that the requirement for the number of antennas, $N \geq LN$, in the proposed scheme as well as in the ZF relay scheme could still be a limiting factor in some application scenarios. In addition, since the relay techniques described in this paper assume perfect CSI knowledge for both the backward and forward MIMO channels at each relay terminal, investigation of their capacity with imperfect CSI is an important future research topic.

**ACKNOWLEDGMENT**

The authors thank Mr. Katsutoshi Kusume for his helpful discussion regarding the complexity issues.

**REFERENCES**

[1] M. Gastpar and M. Vetterli, “On the capacity of wireless networks: the relay case,” in *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM ’02)*, vol. 3, pp. 1577–1586, New York, NY, USA, June 2002.

[2] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[3] 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, no. 3, pp. 311–335, 1998.

[4] G. G. Raleigh and J. M. Cioffi, “Spatio-temporal coding for wireless communication,” *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 357–366, 1998.

[5] J. H. Winters, “On the capacity of radio communication systems with diversity in a Rayleigh fading environment,” *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 5, pp. 871–878, 1987.

[6] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, 2003.

[7] G. Caire and S. Shamai, “On the achievable throughput of a multiantenna Gaussian broadcast channel,” *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[8] S. Vishwanath, N. Jindal, and A. Goldsmith, “Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.

[9] W. Yu and J. M. Cioffi, “Sum capacity of Gaussian vector broadcast channels,” *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1875–1892, 2004.

[10] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, “Capacity scaling laws in MIMO relay networks,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 6, pp. 1433–1444, 2006.

[11] R. U. Nabar, Ö. Oyman, H. Bölcskei, and A. J. Paulraj, “Capacity scaling laws in MIMO wireless networks,” in *Proceedings of Allerton Conference on Communication, Control, and Computing*, pp. 378–389, Monticello, Ill, USA, October 2003.

[12] H. Shi, T. Abe, T. Asai, and H. Yoshino, “A relaying scheme using QR decomposition with phase control for MIMO wireless networks,” in *Proceedings of IEEE International Conference on Communications (ICC ’05)*, vol. 4, pp. 2705–2711, Seoul, Korea, May 2005.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.

[14] L.-U. Choi and R. D. Murch, “A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 20–24, 2004.

[15] H. Bölcskei and R. U. Nabar, “Realizing MIMO gains without user cooperation in large single-antenna wireless networks,” in *Proceedings of IEEE International Symposium on Information Theory*, p. 18, Chicago, Ill, USA, June-July 2004.

---

**Tetsushi Abe** received his B.S. degree and M.S. degree in electrical and electronic engineering from Tokyo Institute of Technology, Tokyo, Japan, in 1998 and 2000, respectively. During 1998–1999, he studied in the Department of Electrical and Computer Engineering in University of Wisconsin, Madison, USA, under the scholarship exchange student program offered by the Japanese Ministry of Education. He joined NTT DoCoMo, Inc., in 2000. Since 2005, he has been with DoCoMo Euro-Labs. He has conducted researches on signal processing for wireless communications: multiple-input and multiple-output (MIMO) transmission, space-time turbo equalization, relay transmission, and OFDM transmission. He is a Member of IEEE and IEICE.
Hui Shi received his B.S. degree in mechanical engineering from Dalian University of Technology, Dalian, China, in 1998 and M.S. degree in electrical and electronic engineering from Nagoya University, Nagoya, Japan, in 2002. Since 2002, he has been with the Research Laboratories at NTT DoCoMo, Inc. His research interests cover the wireless network systems, relay networks, multiple-input and multiple-output (MIMO) transmission, and information theory issues. He is a Member of IEEE and IEICE.

Takahiro Asai received the B.E. and M.E. degrees from Kyoto University, Kyoto, Japan, in 1995 and 1997, respectively. In 1997, he joined NTT Mobile Communications Network, Inc. (now NTT DoCoMo, Inc.). Since joining NTT Mobile Communications Network, Inc., he has been engaged in the research of signal processing for mobile radio communication. He is a Member of IEEE.

Hitoshi Yoshino received the B.S. and M.S. degrees in electrical engineering from the Science University of Tokyo, Tokyo, Japan, in 1986 and 1988, respectively, and the Dr.Eng. degree in communications and integrated systems from the Tokyo Institute of Technology, Tokyo, Japan, in 2003. From 1988 to 1992, he was with Radio Communication Systems Laboratories, Nippon Telegraph and Telephone Corporation (NTT), Japan. Since 1992, he has been with NTT Mobile Communications Network, Inc. (currently, NTT DoCoMo, Inc.), Japan. Since joining NTT DoCoMo, he has been engaged in the areas of mobile radio communication systems and digital signal processing. From 1998 to 1999, he was at the Deutsche Telekom Technologiezentrum, Darmstadt, Germany, as a Visiting Researcher. He is currently an Executive Research Engineer in Wireless Laboratories, NTT DoCoMo, Inc. He received the Young Engineer Award and the Excellent Paper Award from the Institute of Electronics, Information, and Communication Engineers (IEICE) of Japan both in 1995. He is a Member of IEEE.
Nonregenerative Dual-Hop Cooperative Links with Selection Diversity

Theodoros A. Tsiftsis,1 George K. Karagiannidis,2 P. Takis Mathiopoulos,3 and Stavros A. Kotsopoulos1

1Department of Electrical & Computer Engineering, University of Patras, Rion, 26500 Patras, Greece
2Department of Electrical & Computer Engineering, Aristotle University of Thessaloniki, 54 124 Thessaloniki, Greece
3Institute for Space Applications and Remote Sensing (ISARS), National Observatory of Athens, Metaxa & Vas. Pavlou Street, Palaia Penteli, 152 36 Athens, Greece

Received 1 November 2005; Revised 11 March 2006; Accepted 17 May 2006

The end-to-end performance of dual-hop cooperative diversity systems equipped with nonregenerative relays and a selection combining receiver at the destination terminal over independent and nonidentical Nakagami-\(m\) fading channels is studied. Closed-form expressions for the cumulative distribution function and the probability density function of the end-to-end signal-to-noise ratio (SNR) are presented, while analytical formulae are derived for the moments and the moment generating function. Using these statistical results, closed-form expressions for the outage probability are presented for both channel state information and fixed gain relays. Furthermore, for the case of fixed gain relay, the average end-to-end SNR, the amount of fading, and the average bit error rate can be numerically evaluated. The proposed mathematical analysis is complemented by numerical examples, including the effects on the overall performance of the SNRs unbalancing as well as the fading severity.

Copyright © 2006 Theodoros A. Tsiftsis et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Relaying technology is a promising solution for the high throughput/data-rate coverage, as required in future cellular and ad hoc wireless and satellite communications systems. There are two main advantages of this relaying technology:

1. very low transmit RF power requirements,
2. use of spatial/multiuser diversity to combat fading.

Recently, the concept of cooperative diversity, where the mobile users relay signals for each other to emulate an antenna array and exploit the benefits of spatial diversity, has gained great interest [1–9]. More specifically, Emamian et al. [1] have studied the multiuser spatial diversity systems with channel-state-information- (CSI-) based relays, as was first proposed in [10] (and later in [6]), and have derived closed-form expressions for the outage probability over Rayleigh fading channels. CSI-based relays use the instantaneous CSI of the incoming signal to control the output gain and as a result limit the power of the retransmitted signal. In another contribution, Sendonaris et al. [3, 4] have proposed the user cooperation concept and considered practical issues related to its implementation. Moreover, Anghel and Kaveh have presented tight bounds for the outage and error probability of a distributed spatial diversity wireless system in the presence of Rayleigh fading [2]. Note that the relay considered in [2] is an ideal type of CSI-based relay, where the noise figure has been ignored from the relay gain. Also, efficient lower bounds for the average error probability in dual-hop cooperative diversity systems, especially in low average signal-to-noise ratio (SNR) region, have been presented in [5]. Moreover, Laneman et al. have proposed a variety of low-complexity cooperative protocols using the three-terminal case [6]. These protocols have applied on different relaying modes as amplify-and-forward (i.e., nonregenerative relays) and decode-and-forward (i.e., regenerative relays) and also the outage probability, using high-SNR approximations, has been analyzed. The work of Laneman et al. has been extended by Nabar et al. [7], where a new cooperative protocol is presented realizing maximum degree of broadcasting and exhibits no receive collision. Furthermore, Zimmermann et al. [8] have presented an overview of cooperating relaying protocols and compared their performance with that of direct transmission and conventional relaying. Recently, an interesting work using selection diversity among cooperative users was presented by Bletsas et al. [9].
In this work, a method of opportunistic relaying as an efficient cooperative diversity scheme has been proposed. This scheme selects the “best” relay between source and destination based on instantaneous channel measurements and neither topology information nor communication among the relays is needed.

Based upon the above, our paper presents for the first time a completely analytical approach in obtaining the end-to-end performance of dual-hop cooperative links over Nakagami-\(m\) fading channels. In doing so, we first present closed-form expressions for the probability density function (PDF) and cumulative distribution function (CDF) of the end-to-end SNR. Moreover, analytical expressions for the moments and moment generating function (MGF) can be easily obtained. These new results are then applied to study the end-to-end performance of dual-hop cooperative links such as outage probability, average end-to-end SNR, amount of fading (AoF), and average bit error rate (BER) when a selection combining (SC) receiver is assumed at the destination terminal. It will be shown that our general expressions for Nakagami-\(m\) fading reduce to previously published results for \(m = 1\), that is, Rayleigh fading.

The remainder of this paper is organized as follows. Section 2 introduces the system and channel model under consideration. In Section 3, the performance analysis of the system is presented. In particular, closed-form expressions are derived for the outage probability for both types of relays. The analysis is complemented by presenting the moments and the average BER in closed form when fixed gain relay is considered. Numerical results are presented in Section 4 and concluding remarks are given in Section 5.

2. SYSTEM AND CHANNEL MODEL

A dual-hop relaying system operating over independent and nonidentical Nakagami-\(m\) fading channels is illustrated in Figure 1. The source terminal \(S\) communicates with the destination terminal \(D\) not only directly but also via the cooperative diversity link through terminal \(R\), which acts as a non-regenerative gain relay. Each transmission period is divided into two signaling intervals: in the first signaling interval, terminal \(S\) communicates with the relay and the destination terminal, while in the second one, only the relay communicates with terminal \(D\). The above transmission protocol was originally proposed in [10]. The destination terminal combines the received signals using a SC. Assuming that \(S\) is transmitting a signal with an average power normalized to unity, the instantaneous equivalent end-to-end SNR of the dual-hop path can be expressed as in [11, 12]

\[
y_{\text{end}} = \frac{(\alpha_i^1/N_0,1) (\alpha_i^2/N_0,2)}{(\alpha_i^2/N_0,2) + (1/g^2 N_0,1)}.
\]

In the above equation, \(\alpha_i\) is the fading amplitude of the \(i\)th path, \(i = 0, 1, 2\) (\(a_0\) is the fading amplitude of the direct path), \(N_0,2\) is the one-sided power spectral density of the additive white Gaussian noise at the input of \(R\) and \(D\), respectively, and \(g\) is the gain of the relay. Since \(\alpha_i\) is modeled as Nakagami-\(m\) random variable (RV), the instantaneous SNR, \(y_i = \alpha_i^2/N_0,1,\) is a gamma distributed RV with PDF given by

\[
f_{y_i}(y) = \frac{m_i}{\gamma_i^m} y^{m-1} \exp \left( -\frac{m_i y}{\gamma_i} \right),
\]

where \(\Gamma(\cdot)\) is the gamma function [13, equation (8.310/1)], \(m_i \geq 1/2\) is a parameter describing the fading severity, and \(\gamma_i\) the average SNR of the \(i\)th path. Hence, its CDF can be written as

\[
F_{y_i}(y) = 1 - \frac{\Gamma(m_i, (m_i/\gamma_i)y)}{\Gamma(m_i)},
\]

where \(\Gamma(\cdot, \cdot)\) is the incomplete gamma function defined in [13, equation (8.350.2)].

When \(R\) has available CSI from the first hop and its gain aims to limit the output power of the relay, one kind of gain proposed by Laneman and Wornell [10] is given by

\[
g_i^2 = \frac{1}{\alpha_i^2 + N_0,1}.
\]

Therefore, the instantaneous equivalent end-to-end SNR of the dual-hop path can be expressed as in [1, 2, 11, 12]

\[
y_{\text{eq}1} = \frac{y_1 y_2}{y_1 + y_2 + 1}.
\]

When \(R\) introduces fixed gain to the received signal given by [14]

\[
g_i^2 = \frac{1}{CN_0,1},
\]

where \(C\) is a positive constant, the instantaneous end-to-end SNR of the dual-hop path can be expressed as in [14]

\[
y_{\text{eq}2} = \frac{y_1 y_2}{C + y_2}.
\]

3. PERFORMANCE ANALYSIS OF DUAL-HOP COOPERATIVE LINKS

3.1. CSI-based gain relay

3.1.1. Outage probability

When a SC receiver is employed in terminal \(D\), the end-to-end outage probability at \(D\), \(P_{\text{out}}\), is defined as the probability that the equivalent output SNR, \(y_{\text{eq,csi}} = \max(y_{\text{eq}1}, y_0)\),
falls below a given threshold, $\gamma_{th}$. Since it is assumed independent fading channels, the end-to-end outage probability can be written as

$$P_{\text{out,csi}} = \Pr \left[ y_{0\text{,csi}} \leq \gamma_{th} \right] = F_{\gamma_{th}}(\gamma_{th}),$$

(8)

where $F_{\gamma_{th}}(\gamma_{th})$ and $F_{\gamma}(\gamma_{th})$ are the CDFs of the instantaneous end-to-end SNRs of the relay fading channel and direct path, evaluated at $y = \gamma_{th}$, respectively. Since $F_{\gamma_{th}}(\gamma)$ is given by (3), only $F_{\gamma_{th}}(\gamma)$ has to be evaluated. The CDF, $F_{\gamma_{th}}(\gamma)$, using (5), can be expressed as

$$F_{\gamma_{th}}(\gamma) = \int_{0}^{\infty} \left[ y_{1}y_{2} \leq y \right] f_{\gamma_{th}}(y) \, dy_{2}$$

$$= \int_{0}^{\gamma} \left[ y_{1} \geq \frac{y(y+1)}{\gamma_{th}} \right] f_{\gamma_{th}}(y) \, dy_{2} + \int_{\gamma}^{\infty} \left[ y_{1} \leq \frac{y(y+1)}{\gamma_{th}} \right] f_{\gamma_{th}}(y) \, dy_{2}$$

$$= I_{1} + I_{2},$$

(9)

where

$$I_{1} = \int_{0}^{\gamma} f_{\gamma_{th}}(y) \, dy_{2} = F_{\gamma_{th}}(\gamma) = 1 - \frac{\Gamma(m_{1}, (m_{2}/\gamma_{th}))}{\Gamma(m_{2})},$$

$$I_{2} = \int_{0}^{\gamma} f_{\gamma_{th}}\left( \frac{y(y+1)}{\gamma_{th}} \right) f_{\gamma_{th}}(y) \, dy_{2}$$

$$= \int_{0}^{\gamma} \left[ 1 - \frac{\Gamma(m_{1}, (m_{2}/\gamma_{th}))}{\Gamma(m_{1})} \right] f_{\gamma_{th}}(y) \, dy_{2}$$

$$= \frac{\Gamma(m_{2}, m_{2}/\gamma_{th})}{\Gamma(m_{2})} - \frac{m_{2}^{m_{2}}}{\Gamma(m_{1}) \Gamma(m_{2})} I_{3},$$

(10)

$$I_{3} = \int_{\gamma}^{\infty} y_{2}^{m_{2}-1} \exp \left( -\frac{m_{2}y_{2}}{\gamma_{th}} \right) \Gamma(\frac{m_{1}m_{2}y_{2}^{1}}{\gamma_{th}}) \, dy_{2}.$$

(11)

Since the integral $I_{3}$ is not a tabulated one, it can be solved in closed form (see the appendix) and by using (10) and after some straightforward mathematical manipulations, $F_{\gamma_{th}}(\gamma)$ can be expressed as

$$F_{\gamma_{th}}(\gamma) = 1 - \frac{2m_{2}^{m_{2}}(m_{1} - 1)! \exp \left[ -\left( m_{1}y/\gamma_{th} + m_{2}y/\gamma_{th} \right) \right]}{\gamma_{th}^{m_{2}} \Gamma(m_{1}) \Gamma(m_{2})}$$

$$\times \sum_{k=0}^{m_{1}-1} \sum_{\ell=0}^{m_{2}-1} \frac{1}{k!} \frac{1}{l!} \left( \frac{m_{1} - 1}{r} \right) \left( \frac{m_{2}}{\gamma_{th}} \right)^{l}$$

$$\times \frac{m_{1}}{\gamma_{th}^{k+1}} \times y_{2}^{(2k+2m_{2}-\ell)\gamma_{th}^{1}} \left( y + 1 \right)^{l+1/2}$$

$$\times K_{\ell-r-1} \left( 2 \sqrt{\frac{m_{1}m_{2}y(1+y)}{\gamma_{th}}}, \frac{\Gamma(m_{1})}{\Gamma(m_{2})} I_{1} \right),$$

(12)

where $K_{\ell}(-)$ is the $\ell$th-order modified Bessel function of the second kind, defined in [13, equation (8.432)]. Note, that for the case of $m_{i} = 1$, that is, Rayleigh fading channels, (12) reduces to the previously published result presented in [1, equation (14)].

It is clear now that by using (3) and (12), the outage probability of cooperative dual-hop links with CSI-based relays when SC diversity is considered in the destination terminal, can be expressed in closed form as

$$P_{\text{out,csi}}$$

$$= \left[ 1 - \frac{\Gamma(m_{0}, (m_{0}/\gamma_{th}))}{\Gamma(m_{0})} \right]$$

$$\times \left\{ 1 - \frac{2m_{2}^{m_{2}}(m_{1}-1)! \exp \left[ -\left( m_{1}y/\gamma_{th} + m_{2}y/\gamma_{th} \right) \right]}{\gamma_{th}^{m_{2}} \Gamma(m_{1}) \Gamma(m_{2})} \right\}$$

$$\times \sum_{k=0}^{m_{1}-1} \sum_{\ell=0}^{m_{2}-1} \left( \frac{1}{k!} \frac{1}{l!} \left( \frac{m_{1} - 1}{r} \right) \left( \frac{m_{2}}{\gamma_{th}} \right)^{l} (\ell-r-1)^{2} \right)$$

$$\times K_{\ell-r-1} \left( 2 \sqrt{\frac{m_{1}m_{2}y(1+y)}{\gamma_{th}}}, \frac{\Gamma(m_{1})}{\Gamma(m_{2})} I_{1} \right).$$

(13)

### 3.2. Fixed gain relay

#### 3.2.1. Outage probability

For the case of fixed gain relay, the outage probability is given by

$$P_{\text{out,fg}} = F_{\gamma_{th}}(\gamma_{th}) F_{\gamma_{fg}}(\gamma_{th}),$$

(14)

where $F_{\gamma_{fg}}(\gamma_{th})$ is the CDF of the instantaneous end-to-end SNR for the dual-hop path.

Using (7), $F_{\gamma_{fg}}(\gamma_{th})$ can be written as

$$F_{\gamma_{fg}}(\gamma)$$

$$= \int_{0}^{\infty} \left[ \frac{y_{1}}{y_{2} + C} \leq y \right] f_{\gamma_{fg}}(y) \, dy_{2}$$

$$= \int_{0}^{\infty} \left[ y_{1} \leq \frac{y(y_{2} + C)}{y_{2}} \right] f_{\gamma_{fg}}(y) \, dy_{2}$$

$$= \int_{0}^{\infty} \left[ 1 - \frac{\Gamma(m_{1}, (m_{2}/y_{2} + C)/(y_{2} + C))}{\Gamma(m_{1})} \right] f_{\gamma_{fg}}(y) \, dy_{2}$$

$$= 1 - \frac{m_{2}^{m_{2}}}{\gamma_{th}^{m_{2}} \Gamma(m_{1}) \Gamma(m_{2})} I_{1}$$

(15)

with

$$I_{1} = \int_{0}^{\infty} y_{2}^{m_{2}-1} \exp \left( -\frac{m_{2}y_{2}}{\gamma_{th}} \right) \Gamma\left( \frac{m_{1}y(y + C)}{y_{2}}, \frac{\Gamma(m_{1})}{\Gamma(m_{2})} I_{1} \right).$$

(16)
Since the integral $I_1$ is not a tabulated one, it can be solved in closed form by setting $y = 0$ in the integral $I_3$ (see the appendix) as

$$I_1 = \sum_{k=0}^{m_1-1} \sum_{\ell=0}^{k} \frac{2(m_1 - 1)!}{k!} \exp \left( -\frac{m_1 y}{\bar{y}_1} \right) (\ell/\ell) e^{k-\ell} \times \left( \frac{m_1 y}{\bar{y}_1} \right)^{\ell} \left( \frac{m_2}{\bar{y}_2} \right)^{(k-\ell-m_2)/2} \left( \frac{\bar{y}_1}{C m_1 y} \right)^{(k-\ell-m_2)/2} \times K_{k-m_2-\ell} \left( 2 \sqrt{\frac{m_1 m_2 C y}{\bar{y}_1 \bar{y}_2}} \right).$$

Using (15) and (17), $F_{\gamma|\ell}(y)$ can be written as

$$F_{\gamma|\ell}(y) = 1 - \sum_{k=0}^{m_1-1} \sum_{\ell=0}^{k} \Xi[k, \ell] \exp \left( -\frac{m_1 y}{\bar{y}_1} \right)^{y^{k+\ell+m_2}/2} \times K_{k-m_2-\ell} \left( 2 \sqrt{\bar{y}} \right),$$

where

$$\Xi[k, \ell] = 2 \left( \frac{k}{\ell} \right) (m_1 - 1)! \left( \frac{C^{k+m_2-\ell}}{\Gamma(m_1) \Gamma(m_2) k!} \right) \times \left( \frac{m_1 y}{\bar{y}_1} \right)^{(k+m_2+\ell)/2} \left( \frac{m_2}{\bar{y}_2} \right)^{(k+\ell-m_2)/2},$$

$$\tau = \frac{m_1 m_2 C}{\bar{y}_1 \bar{y}_2}.$$

Note that for the case of Rayleigh fading channels, (18) reduces to the previously published result presented in [14, equation (9)].

Therefore, using (3) and (18), the outage probability of cooperative dual-hop links with fixed gain relays when SC diversity is considered in the destination terminal can be expressed in closed form as

$$P_{out,fg} = \left[ 1 - \frac{\Gamma(m_0, (m_0/\bar{y}_0) \gamma_{th})}{\Gamma(m_0)} \right] \times \left[ 1 - \sum_{k=0}^{m_1-1} \sum_{\ell=0}^{k} \Xi[k, \ell] \exp \left( -\frac{m_1 y_{th}}{\bar{y}_1} \right)^{y^{k+\ell+m_2}/2} \times K_{k-m_2-\ell} \left( 2 \sqrt{\bar{y}_1} \right) \right].$$

3.2.2. Moments of the end-to-end SNR

The first- and the second-order moments of the end-to-end SNR are statistical parameters which can be efficiently used to evaluate important performance system measures, such as average output SNR and variance. The higher-order moments are also useful in signal processing algorithms for signal detection, classification, and estimation and they play a fundamental role in understanding the performance of wideband communication systems in the presence of fading [15].

The $n$th-order moment of the end-to-end SNR at terminal $D$ is given by

$$\mu_n = E[y_{eq,fg}^n] = \int_0^\infty y^n f_{\gamma|\ell}(y) dy,$$

where $E[\cdot]$ denotes the statistical average operator, $y_{eq,fg}$ is max($y_{eq2,0}$ and $y_{eq,fg}$) is the PDF of the SC output. This PDF can be found by taking the derivative of (14) with respect to $y_{th}$, yielding

$$f_{\gamma|\ell}(y) = \frac{m_0^m}{\bar{y}_0^m \Gamma(m_0)} \exp \left( -\frac{m_0 y}{\bar{y}_0} \right)^{y_{th}^{m-1} F_{\gamma|\ell}(y)}.$$

The PDF of the dual-hop path, $f_{\gamma|\ell}(y)$, can be evaluated by taking the derivative of (18) with respect to $y$, yielding

$$f_{\gamma|\ell}(y) = \sum_{k=0}^{m_1-1} \sum_{\ell=0}^{k} \Xi[k, \ell] \exp \left( -\frac{m_1 y}{\bar{y}_1} \right)^{y^{k+\ell+m_2}/2} \times \left[ \frac{m_1 y}{\bar{y}_1} \right] \times K_{k-m_2-\ell} \left( 2 \sqrt{\bar{y}_1} \right)$$

By substituting (23) into (21), the resulting integral can be expressed as

$$\mu_n = \frac{m_0^m}{\bar{y}_0^m \Gamma(m_0)} \left\{ \int_0^\infty y^n \exp \left( -\frac{m_0 y}{\bar{y}_0} \right)^{y_{th}^{m-1} - \sum_{k=0}^{m_1-1} \sum_{\ell=0}^{k} \Xi[k, \ell] \exp \left( -\frac{m_1 y_{th}}{\bar{y}_1} \right)^{y^{k+\ell+m_2}/2} \times K_{k-m_2-\ell} \left( 2 \sqrt{\bar{y}_1} \right) \right\} \int_0^\infty y^n f_{\gamma|\ell}(y) dy.$$

The first integral can be solved in closed form using [15, equation (3.381.4)]. The second and the third integrals, after a simple transformation, are in the form of $\int_0^\infty x^r \exp(-bx^r)K_n(x) dx$, which can be evaluated using [15, equation (6.631.3)] yielding a complicated closed-form mathematical expression involving the Whittaker function. This expression can be simplified further by applying the identities [13, equations (9.220.3), (9.220.4)] of the well-known confluent hypergeometric function, $_1F_1(x, y; z)$, [13, equations (9.210.1) and (9.220.4)]. Hence, the result is
a finite nested sum of $iF_1(x, y; z)$. The last integral is not a tabulated one since it involves the incomplete gamma function. By applying [13, equation (8.354.2)], incomplete gamma function can be written as an infinite series sum (i.e.,
\[ \Gamma(x, y) = \frac{1}{\Gamma(x)} \sum_{n=0}^{\infty} \frac{(-1)^n x^{y+n}}{n!(a+n)} \] It is obvious, after the previous expansion, that the last integral has similar form as the previous ones yielding to an infinite nested sums of $iF_1(x, y; z)$. The analytical expression for the moments\(^1\) is very useful since it can be used to directly obtain the average end-to-end SNR (i.e., $\mu_1$) or the amount of fading (AoF) defined as
\[ \text{AoF} = \frac{\mu_2}{\mu_1} - 1 \] (25)
under various fading channel conditions.

It is worth to derive a closed-form expression for the third integral, which represents the moments of the dual-hop relay fading channel using a fixed gain relay over Nakagami-$m$ environment, given by

\[
\mu_n = \sum_{k=0}^{m-1} \sum_{\ell=0}^{k} \Xi[k, \ell] \\
\times \left\{ \begin{array}{l}
\xi_1[k, \ell, n]\left[ iF_1(k + n, m + n, k - m_2 - \ell; \frac{\eta}{m_1}) \Gamma(1 + k + m_2 - \ell)^{-1} \\
- \frac{1}{\Gamma(-k + m_2 - \ell)} \left( \frac{1}{2} \right) \right] \\
+ \xi_2[k, \ell, n]\left[ \frac{1}{m_1} \frac{1}{\Gamma(1 + k + m_2 - \ell)} \left( \frac{1}{2} \right) \right] \\
+ \xi_3[k, \ell, n]\left[ \frac{1}{m_1} \frac{1}{\Gamma(1 + k + m_2 - \ell)} \left( \frac{1}{2} \right) \right] \\
+ \xi_4[k, \ell, n]\left[ \frac{1}{m_1} \frac{1}{\Gamma(1 + k + m_2 - \ell)} \left( \frac{1}{2} \right) \right] \end{array} \right. \\
\right. \\
(26a)
\]
where
\[
\xi_i[k, \ell, n] = \frac{1}{2} \left( \frac{m_1}{\gamma} \right)^{-k-n} i^{(k+m_2-\ell)/2} \Gamma(k + n),
\xi_2[k, \ell, n] = \frac{m_1}{\gamma} i^{(k-m_2-\ell)/2} \Gamma(1 + k + n),
\xi_3[k, \ell, n] = \frac{1}{2} \left( \frac{m_1}{\gamma} \right)^{-m_2-n-\ell} i^{(k-m_2-\ell)/2} \Gamma(m_2 + n + \ell),
\xi_4[k, \ell, n] = \frac{1}{2} \left( \frac{m_1}{\gamma} \right)^{-m_2-n-\ell} i^{(k-m_2-\ell)/2} \Gamma(m_2 + n + \ell).
\] (26b)

3.2.3. Average BER

The MGF of the end-to-end SNR at terminal $D$ is defined as
\[
\mathcal{M}_{\gamma_{eq}(s)} \triangleq \mathbb{E}[e^{\gamma_{eq}}] = \int_0^\infty e^{\gamma} f_{\gamma_{eq}(\gamma)}(\gamma) d\gamma.
\] (27)
Following the same procedure as for the moments, the MGF for the dual-hop cooperative links can be derived in terms of nested sums of $iF_1(x, y; z)$. Again, the analytical result of $\mathcal{M}_{\gamma_{eq}(s)}$ is not presented here for the same reason as for the moments. However, with the aid of $\mathcal{M}_{\gamma_{eq}(s)}$, the performance of a variety of modulation schemes can be easily obtained [16]. Examples of such modulation schemes include $M$-ary quadrature amplitude modulation ($M$-QAM), $M$-ary phase-shift keying ($M$-PSK), and noncoherent binary schemes such as binary frequency-shift keying (BFSK) and differential phase-shift keying (DPSK). For example, the average BER of binary DPSK can be given by
\[
P_b(E) = 0.5 \mathcal{M}_{\gamma_{eq}(x)}(-1).
\]

$\mu_n$ \(^1\)

4. NUMERICAL EXAMPLES AND SIMULATIONS

In this section, numerical examples are presented, which also verify the accuracy of our mathematical analysis. Figure 2 depicts the outage probability as a function of the average SNR of the direct link, $\sqrt{\gamma_{eq}}$, for both types of relays and for different outage threshold values. Comparing the performance between CSI-based relay with the equivalent, in terms of average introduced gain, fixed gain relay (i.e., $g_0^2 = \mathbb{E}[(1/(\alpha^2 + N_{0i}))])$, it can be seen that for medium-to-large SNR values the CSI-based relay outperforms those with fixed gain. However, for the low-SNR region, dual-hop systems employed with fixed gain relays outperform those with variable gain. This happens due to the fact that the maximum value of the CSI-based gain relays, $g_1$, is $1/N_0$ when $\alpha_i \rightarrow 0$, which is very possible in the low average SNR regime. Considering the higher complexity nature of CSI-based relays, our results show that fixed gain relays may serve as an efficient solution for relayed transmissions. Similar observations have been made for the Rayleigh fading channel in [14]. Furthermore, it is observed that as $\sqrt{\gamma_{eq}}$ increases, the SNR range where fixed gain outperforms CSI-based relays also increases.

\(^1\)Although the analytical expression for the moments of the cooperative dual-hop link with fixed gain relay has been derived, it is not presented here since it is very complicated and does not provide much insight to the performance analysis. However, since $iF_1(x, y; z)$ is a built-in standard function to the most well-known mathematical software packages, the results can be easily extracted.
Figure 2: Outage probability of dual-hop cooperative links with CSI-based relay and SC diversity receiver versus average SNR $\gamma_0$, for different relay types ($m_1 = m_2 = 2, m_0 = 1.5, \gamma_1 = 2\gamma_0$, and $\gamma_2 = 5\gamma_0$).

Figure 3: AoF for a dual-hop cooperative link with fixed gain relay and SC diversity receiver versus fading severity ($m_1 = m_2 = m_0 = m, \gamma_1 = 1.5\gamma_0, \gamma_2 = 3\gamma_0$, and $\gamma_0 = 0$ dB).

In Figure 3 the AoF of a dual-hop cooperative link with fixed gain relay over integer-order Nakagami-$m$ fading parameters is depicted. It is clear from Figure 3 that the overall AoF is reduced when a relay is used. Also, it is interesting to note that AoF deteriorates for low values of $m$ even for high amplification gains. This can be explained by the fact that for strong fading conditions the fixed gain relay does not amplify only the received signal but also the noise of the end-to-end link. Moreover, as $m$ increases, the overall difference between AoFs of a direct link and a dual-hop cooperative transmission decreases. Finally, in Figure 4 the average BER for DPSK is showed versus $\gamma_0$. As it was expected, an increase in fading parameter upgrades the error performance of the dual-hop cooperative link.

5. CONCLUSION

In this paper, the performance of dual-hop cooperative diversity systems with nonregenerative relays and a SC receiver at the destination terminal over Nakagami-$m$ fading channels was studied. Specifically, closed-form expressions for the outage probability for both CSI-based and fixed gain relays were presented. Regarding the outage probability, the analysis showed that the fixed gain relays may serve as an efficient solution for relayed transmissions in contrast to the high complexity nature of CSI-based relays. Moreover, analytical expressions for the AoF and the average BER when
a fixed gain relay is considered were derived. These expressions were presented as infinite nested sums of the confluent hypergeometric function, which can be easily evaluated numerically.

**APPENDIX**

**EVALUATION OF THE INTEGRAL** \( I_3 \)

The integral \( I_3 \) in (11), can be written in its general form as

\[
I_3 = \int_{-\infty}^{\infty} x^{a-1} \exp(-bx) \Gamma \left( c, d \left( \frac{x + e}{x - y} \right) \right) dx, \tag{A.1}
\]

where \( a, c \) are positive integers and \( b, d, e \) real numbers. Then, the incomplete gamma function of the integrand in (A.1), using [13, equations (8.352/2), (1.111)], can be expressed as

\[
\Gamma \left( c, d \left( \frac{x + e}{x - y} \right) \right) = (c - 1)! \exp(-d) \times \exp \left[ -d \left( \frac{y + e}{x} \right) \right] \times \sum_{k=0}^{c-1} \sum_{\ell=0}^{k} \frac{d^\ell}{k!} \frac{(k + e)}{c - 1!} \left( \frac{y + e}{x} \right)^{\ell}, \tag{A.2}
\]

where \( \binom{\cdot}{\cdot} \) denotes the binomial coefficient. Using (A.1) and (A.2) after a simple transformation of the integration variable, \( I_3 \) can be expressed as

\[
I_3 = (c - 1)! \exp \left[ -(d + by) \right] \times \sum_{k=0}^{c-1} \sum_{\ell=0}^{k} \frac{d^\ell}{k!} \frac{(k + e)}{c - 1!} \left( \frac{y + e}{x} \right)^{\ell} \int_{0}^{\infty} x^{-\ell} \exp(-bx) \exp \left[ -d \left( \frac{y + e}{x} \right) \right] dx. \tag{A.3}
\]

The inner integral of (A.3) can be solved by using [13, equation (3.471.9)], resulting in

\[
I_3 = (c - 1)! \exp \left[ -(d + by) \right] \times \sum_{k=0}^{c-1} \sum_{\ell=0}^{k} \frac{d^\ell}{k!} \frac{(k + e)}{c - 1!} \left( \frac{y + e}{x} \right)^{\ell} \times 2 \left( \frac{d}{y + e} \right)^{(r+\ell+1)/2} K_{\ell+1/2} \left( 2 \sqrt{bd(y + e)} \right). \tag{A.4}
\]

**ACKNOWLEDGMENTS**

This work was performed within the framework of the Satellite Communications Network of Excellence (SatNEx) project (IST-507052) and its Phase-II, SatNEx-II (IST-27393), funded by the European Commission (EC) under its FP6 program.
Theodoros A. Tsiftsis was born in Lamia, Greece, in 1970. He received the B.S. degree in physics from Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1993, and the M.S. degree in electrical engineering from the Heriot-Watt University, Edinburgh, Scotland, UK, in 1995. Also, he received the M.S. degree in decision sciences from the Athens University of Economics and Business, Greece, in 2000. He is currently working towards the Ph.D. degree in electrical engineering at the University of Patras, Greece. Since September 2002, he has been a research assistant at the Electrical and Computer Engineering Department, University of Patras, Greece, and a tenure-track research assistant at the Technological Educational Institute of Lamia, Greece. His current research interests include topics as digital communications over fading channels, multihop wireless communications, and free-space optical communication systems.

George K. Karagiannidis was born in Pathagorion, Samos Island, Greece. He received his university degree in 1987 and his Ph.D. degree in 1999, both in electrical engineering, from the University of Patras, Patras, Greece. From 2000 to 2004 he was a Researcher at the Institute for Space Applications & Remote Sensing, National Observatory of Athens, Greece. In June 2004, he joined the faculty of Aristotle University of Thessaloniki, Greece, where he is currently an Assistant Professor at the Electrical & Computer Engineering Department. His major research interests include wireless communications theory, digital communications over fading channels, satellite communications, mobile radio systems, and free-space optical communications. He has published and presented more than 70 technical papers in scientific journals and international conferences, he is a Coauthor in 3 chapters in books and also a Coauthor in a Greek edition book on mobile communications. He is a Member of the Editorial Boards of IEEE Communications Letters and EURASIP Journal on Wireless Communications and Networking.

P. Takis Mathiopoulos is currently Director of Research at the Institute for Space Applications and Remote Sensing (ISARS) of the National Observatory of Athens (NOA), where he has established the Wireless Communications Research Group. As ISARS’ Director he has lead the institute to a significant expansion, R&D growth, and international scientific recognition. For these achievements, ISARS has been selected as one of the national centers of excellence for the years 2005–2008. He was a faculty member at the University of British Columbia (UBC) in Canada for 14 years, last holding the rank of Full Professor of electrical and computer engineering. Maintaining his ties with academia, he is an Adjunct Professor at UBC and is teaching part-time at the Department of Informatics and Telecommunications, University of Athens. Over the years, he has supervised university- and industry-based R&D groups and has successfully acted as a Technical Manager for large R&D Canadian and European projects. He has supervised the theses of more than 30 graduate students, and has published close to 150 papers in journals and conference proceedings. He serves on the Editorial Board of several scientific journals, including the IEEE Transactions on Communications. He has regularly acted as a consultant for various governmental and private organizations, including the European Commission. He has delivered numerous invited presentations, including plenary lectures, and has taught many short courses all over the world.

Stavros A. Kotsopoulos was born in Argos-Argolidos (Greece) in the year 1952. He received his B.S. degree in physics in the year 1975 from the Aristotle University of Thessaloniki, and in the year 1984 he received his Diploma in electrical and computer engineering from the University of Patras. He did his postgraduate studies in the University of Bradford, UK, and he is an M.Phil and Ph.D. holder since 1978 and 1985 correspondingly. Currently he is a Member of the academic staff of the Department of Electrical and Computer Engineering of the University of Patras and holds the position of Professor. Since 2004, he is the Director of the Wireless Telecommunications Laboratory and develops his professional life teaching and doing research in the scientific area of telecommunications, with interest in cellular mobile communications, wireless networks, interference, satellite communications, telematics applications, communication services, and antennae design. Moreover, he is the Coauthor of the book titled Mobile Telephony. The research activity is documented by more than 160 publications in scientific journals and proceedings of international conferences. Associate Professor Kotsopoulos has been the leader of several international and many national research projects. Finally, he is a Member of the Greek Physicists Society and a Member of the Technical Chamber of Greece.