

Optimal Power Allocation for Orthogonal Relay Systems with Quality-of-Service Constraints

F. Ahmed and Timothy N. Davidson
School of Computational Engineering and Science
McMaster University, Hamilton, Ontario, Canada

Abstract—In this paper, we address the problem of minimizing the cost of the power required to achieve reliable communication at a specified rate on a link that is assisted by an orthogonal half-duplex relay. The considered scenario is that of a narrowband channel with a long coherence time, single-antenna nodes, and the availability of perfect channel state information. We consider the regenerative decode-and-forward and amplify-and-forward relaying strategies, and for both of these strategies we obtain closed-form expressions for the optimal powers. These expressions yield insight into the design of relay-assisted communication schemes with quality-of-service requirements.

I. INTRODUCTION

Enabling collaboration among nodes in a wireless communication network, or collaboration with dedicated relaying nodes, offers the potential for substantial gains in the data rate, reliability and coverage of the network; e.g., [1]–[6]. In response to this potential, a large number of collaborative schemes have been proposed. In many of these schemes, a key step towards realizing their potential is the development of computationally-efficient algorithms for determining the power that each node in the network should employ. In this paper we will develop such an algorithm for a simple point-to-point link that is assisted by a half-duplex relay; see Fig. 1. We will focus on orthogonal relaying schemes in which the source and relay transmit in different time slots, as this avoids the need for the relay to receive and transmit simultaneously, and simplifies the decoding process at the destination. We will consider both the regenerative decode-and-forward relaying strategy and the amplify-and-forward strategy.

In some previous work on this class of systems, and some closely-related systems, the objective was to choose the powers of the source and relay nodes so as to optimize a measure of the system performance, such as the achievable data rate, subject to constraints on the allowable powers; e.g., [7]–[10]. In this paper, we will take a different approach and we will seek to minimize the cost of the transmission powers of the source and relay nodes subject to the requirement that a given target rate can be supported. This approach ensures that the required level of quality of service is established at minimum cost. This is of interest in applications such as real time interactive video, in which a certain data rate needs to be maintained, at a given error rate, in order for the service to be viable. We will consider a scenario involving narrowband transmission between nodes with a single antenna over a

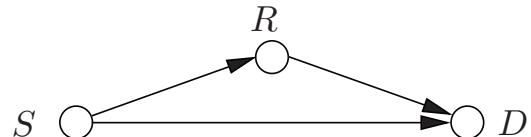


Fig. 1. A simple relay-assisted link.

channel with a long coherence time, and we will study the limiting case in which perfect channel state information is available to the receivers and design node, and the transmitters can be informed of the precise powers that they are to use. For that scenario, we will obtain closed-form expressions for the optimal powers.

II. SYSTEM DESCRIPTION

We consider the simple system illustrated in Fig. 1. This system operates either in a direct transmission mode, or in a relay-assisted mode. In the relay-assisted mode, frames of $2L$ channel uses are partitioned into two equal time slots. In the first time slot of the i th frame, the source, S , transmits a signal block $\mathbf{x}^{(i)}$ and the relay, R , and destination, D , listen. In the second time slot, the relay transmits a function $f(\cdot)$ of the signal block $\mathbf{y}_R^{(i)}$ that it received in the first time slot, and destination listens. In contrast, in the i th frame of the direct transmission mode, the source transmits the signal block $\mathbf{x}^{(2i)}$ in the first time slot and the signal block $\mathbf{x}^{(2i+1)}$ in the second time slot. The relay-assisted and direct transmission modes are illustrated in Figs 2 and 3, respectively.

There are number of strategies that we could consider for the operation of the relay; i.e., for the nature of function $f(\cdot)$ in Fig. 2; e.g., [2], [6]. Since we are considering a system in which each frame is divided into two time slots of equal duration, two natural strategies to consider are regenerative decode-and-forward (RDF) and amplify-and-forward (AF). In the RDF strategy, the relay decodes the message that it received in the first time slot, then re-encodes that message using the same code book as the source and transmits it to the destination in the second time slot. In the AF strategy there is no data regeneration. In the second time slot, the relay simply transmits a scaled version of the signal that it received during the first time slot.

This work was supported in part by the Canada Research Chairs program.

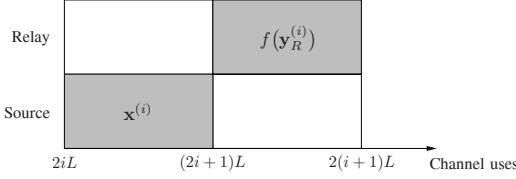


Fig. 2. The i th frame of the relay-assisted transmission mode.

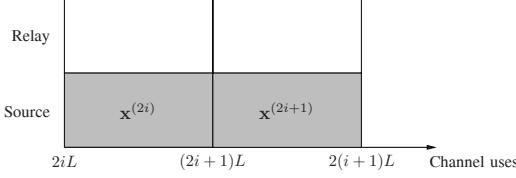


Fig. 3. The i th frame of the direct transmission mode.

A. Channel model

We will consider a scenario in which the nodes each have a single antenna and we consider a narrowband channel model with long coherence times. The noise at the receiver is assumed to be additive, white and Gaussian, and we assume perfect synchronization. Since the coherence time is long, we can assume that perfect channel state information is available at the receivers and hence they can perform coherent detection. We will also assume that channel state information for all three channels is available to the node at which the power allocation takes place, and we will assume that this power allocation can be perfectly communicated to the transmitters.

To construct a mathematical model for the system that we consider, we let h_{ij} , $i, j \in \{S, R, D\}$ denote the complex gain of the channel from node i to node j . We also let n_j denote the additive noise at node $j \in \{S, R, D\}$ and let σ_j^2 denote its variance. If we let $x[k]$ denote the signal transmitted by the source node at an arbitrary instant in the first time slot of a given frame, with the time origin shifted to the start of that frame, then the corresponding signals received at the relay and destination node can be written as

$$y_R[k] = h_{SR}x[k] + n_R[k], \quad (1)$$

$$y_D[k] = h_{SD}x[k] + n_D[k], \quad (2)$$

respectively. If we let $u[\ell]$ denote the signal transmitted by the relay at an arbitrary instant in the second time slot of the given frame, then the corresponding signal at the destination can be written as

$$y_D[L + \ell] = h_{RD}u[\ell] + n_D[L + \ell]. \quad (3)$$

For later convenience we will define the effective power gain for each of the channels as $\gamma_{ij}^2 = |h_{ij}|^2/\sigma_j^2$.

III. PROBLEM FORMULATION

In general terms, the goal of our approach to power allocation is to minimize the cost of the power expended by the source and relay in providing a specified quality of service for communication from the source to the destination. We will

specify the quality of service in terms of a target rate and hence the quality of service constraint will be that reliable communication at the specified rate can be achieved; i.e., the maximum achievable rate of the system is at least as large as the target rate.

In most scenarios, the costs of the power expended by the source and the relay will be different, and we will use the parameter λ to capture the relative cost of the (average) power expended by the relay to that expended by the source. For example, for a dedicated relay station with an external power supply, λ will likely be significantly less than one, whereas when the relaying node is an idle mobile station with a low battery level, λ may be significantly greater than one. Once the appropriate value of λ has been chosen, the objective of the optimization is to minimize the total cost of the power, $\tilde{P}_S + \lambda \tilde{P}_R$, where \tilde{P}_S and \tilde{P}_R denote the average powers of the source and relay, respectively, subject to a specified (source–destination) target rate R_{tar} being achievable. In the relay-active mode, the source and relay are “on” for half of the frame, and hence $\tilde{P}_S = P_S/2$ and $\tilde{P}_R = P_R/2$, where P_S and P_R are the power levels at which the source and relay operate when they are on, respectively. In contrast, in the direct transmission mode, the source transmits over the whole frame, and hence the operating power level is equal to \tilde{P}_S . In other words, if the source operates at a power level P_S in the relay-active mode, then the operating power level in the direct transmission mode that results in the same average power is $P_S/2$.

The set of rates that can be achieved for given values of P_S and P_R is dependent on whether the system is operated in the direct transmission mode or the relay-assisted mode. In the relay-assisted mode, the set of achievable rates also depends on the relaying strategy. As we will consider both the RDF and AF relaying strategies, for the time being we will simply denote the maximum achievable rate in the relay-active mode for given values of P_S and P_R as $g(P_S, P_R)$. Since we have a simple additive white Gaussian noise (AWGN) channel model, the maximum achievable rate in the direct transmission mode is simply

$$g_{\text{dir}}(\tilde{P}_S) = \log(1 + \gamma_{SD}^2 \tilde{P}_S). \quad (4)$$

Throughout this paper, the achievable rates will be expressed in terms of bits per (complex) channel use, and hence the logarithm in (4) is with base 2.

We are now in a position to formally state the design problem. For notational convenience, we will pose that problem in terms of the operating power levels, P_S and P_R , and we will scale the objective by a factor of two. In particular, given a target rate for source–destination communication, R_{tar} , and the relative cost of the relay power, λ , the generic formulation of the design problem is:

$$\min_{P_S, P_R \geq 0} \quad P_S + \lambda P_R \quad (5a)$$

$$\text{subject to} \quad \max\{g(P_S, P_R), g_{\text{dir}}(P_S/2)\} \geq R_{\text{tar}}. \quad (5b)$$

We will denote the optimal solution to (5) by (P_S^*, P_R^*) .

IV. OPTIMAL POWER ALLOCATION

To develop a generic solution procedure for the problem in (5), we first observe from equation (4), that by using the direct transmission mode, reliable communication at a rate R_{tar} can be achieved using

$$P_S = P_{S,\text{dir}}^* = 2(2^{R_t} - 1)/\gamma_{SD}^2. \quad (6)$$

Keeping that in mind, we then calculate the optimal operating power levels in the relay-active mode. That is, if we let $P_{S,\text{ra}}^*$ and $P_{R,\text{ra}}^*$ denote the optimal power allocations in the relay-active mode, we determine the solution of the following problem

$$\min_{P_S, P_R \geq 0} P_S + \lambda P_R \quad (7a)$$

$$\text{subject to } g(P_S, P_R) \geq R_{\text{tar}}. \quad (7b)$$

Using the solution to the problem in (7) and the expression in (6), the optimal solution to the original problem in (5) can be determined as follows:

- If $P_{S,\text{dir}}^* < P_{S,\text{ra}}^* + \lambda P_{R,\text{ra}}^*$, then the optimal solution to (5) is to operate the system in direct transmission mode with power $P_{S,\text{dir}}^*$ at the source.
- If $P_{S,\text{dir}}^* > P_{S,\text{ra}}^* + \lambda P_{R,\text{ra}}^*$ then the optimal solution to (5) is to operate the system in the relay-active mode with powers $P_{S,\text{ra}}^*$ at the source and $P_{R,\text{ra}}^*$ at the relay.

In the case of equality, one would typically choose direct transmission, due to its relative simplicity.

Given the nature of the generic procedure above, the goal of the following sections is to obtain closed-form solutions to the problem in (7) for the cases of regenerative decode-and-forward relaying and amplify-and-forward relaying.

A. Regenerative decode-and-forward

For operating powers P_S and P_R , the maximum achievable rate of the relay-active mode under the regenerative decode-and-forward (RDF) relaying strategy is [4]

$$g(P_S, P_R) = \min\left\{\frac{1}{2}\log(1 + \gamma_{SR}^2 P_S), \frac{1}{2}\log(1 + \gamma_{SD}^2 P_S + \gamma_{RD}^2 P_R)\right\}. \quad (8)$$

Using that expression, the problem in (7) for the RDF case can be written as

$$\min_{P_S, P_R \geq 0} P_S + \lambda P_R \quad (9a)$$

$$\text{subject to } \frac{1}{2}\log(1 + \gamma_{SR}^2 P_S) \geq R_{\text{tar}}, \quad (9b)$$

$$\frac{1}{2}\log(1 + \gamma_{SD}^2 P_S + \gamma_{RD}^2 P_R) \geq R_{\text{tar}}, \quad (9c)$$

and subsequently as

$$\min_{P_S, P_R \geq 0} P_S + \lambda P_R \quad (10a)$$

$$\text{subject to } \gamma_{SR}^2 P_S \geq 2^{2R_{\text{tar}}} - 1, \quad (10b)$$

$$\gamma_{SD}^2 P_S + \gamma_{RD}^2 P_R \geq 2^{2R_{\text{tar}}} - 1. \quad (10c)$$

The formulation in (10) is a linear program and hence can be efficiently solved. In fact, since this linear program has only two variables and two constraints, we have been able to derive

a closed-form solution. Indeed, by examining the geometry of the feasible set, it can be shown that the key points are: (A) the point at which the constraints in (10b) and (10c) intersect, and (B) the point at which the constraint in (10c) intersects with the line $P_R = 0$. The coordinates of these points are $(P_{S,A}, P_{R,A})$ and $(P_{S,B}, 0)$, respectively, where [11]

$$P_{S,A} = (2^{2R_{\text{tar}}} - 1)/\gamma_{SR}^2, \quad (11a)$$

$$P_{R,A} = \frac{(2^{2R_{\text{tar}}} - 1)}{\gamma_{RD}^2} - \left(\frac{\gamma_{SD}^2}{\gamma_{RD}^2}\right) \frac{(2^{2R_{\text{tar}}} - 1)}{\gamma_{SR}^2}, \quad (11b)$$

$$P_{S,B} = (2^{2R_{\text{tar}}} - 1)/\gamma_{SD}^2. \quad (11c)$$

Using these expressions, and the geometry of the feasible set, the optimal solution to the problem in (9) can be written as [11]

$$(P_{S,\text{ra}}^*, P_{R,\text{ra}}^*)$$

$$= \begin{cases} (P_{S,A}, P_{R,A}) & \text{if } \gamma_{SR}^2 > \gamma_{SD}^2 \text{ and } \lambda < \frac{\gamma_{RD}^2}{\gamma_{SD}^2}, \\ (P_{S,B}, 0) & \text{if } \gamma_{SR}^2 \leq \gamma_{SD}^2 \text{ or } \lambda > \frac{\gamma_{RD}^2}{\gamma_{SD}^2}. \end{cases} \quad (12)$$

If $\gamma_{SR}^2 > \gamma_{SD}^2$ and $\lambda = \gamma_{RD}^2/\gamma_{SD}^2$, then any solution of the form $\mu(P_{S,A}, P_{R,A}) + (1 - \mu)(P_{S,B}, 0)$ with $\mu \in [0, 1]$ is optimal.

We observe from (12) that if $\gamma_{SR}^2 \leq \gamma_{SD}^2$ or $\lambda > \gamma_{RD}^2/\gamma_{SD}^2$ then the optimal solution in the relay active mode results in the relay being idle and source communicating directly to destination in first half of each time slot. As intuition would suggest, this requires more power than the direct transmission mode (i.e., $P_{S,B} \geq P_{S,\text{dir}}^*$) and hence, in that case, the optimal solution to the original problem in (5) is to employ direct transmission. If $\gamma_{SR}^2 > \gamma_{SD}^2$ and $\lambda < \gamma_{RD}^2/\gamma_{SD}^2$, then both relaying and direct transmission remain candidate solutions. Although we are continuing to try to determine an *a priori* criterion that indicates which solution will be optimal, there are only two candidate solutions and it is straightforward to compare the values of $P_{S,\text{dir}}^*$ and $P_{S,A} + \lambda P_{R,A}$ and select the scheme corresponding to the smaller of these two values.

B. Amplify-and-forward

For operating powers P_S and P_R , the maximum achievable rate of the relay-active mode under the amplify-and-forward (AF) relaying strategy is [4]

$$g(P_S, P_R) = \frac{1}{2}\log\left(1 + \gamma_{SD}^2 P_S + \frac{\gamma_{SR}^2 P_S \gamma_{RD}^2 P_R}{1 + \gamma_{SR}^2 P_S + \gamma_{RD}^2 P_R}\right). \quad (13)$$

Let us define the following parameters,

$$a_1 = \gamma_{SD}^2 - \gamma_{SR}^2(2^{2R_{\text{tar}}} - 1), \quad (14a)$$

$$a_2 = \gamma_{SD}^2 \gamma_{SR}^2, \quad (14b)$$

$$a_3 = \gamma_{SD}^2 \gamma_{RD}^2 + \gamma_{SR}^2 \gamma_{RD}^2, \quad (14c)$$

$$a_4 = -(2^{2R_{\text{tar}}} - 1)\gamma_{RD}^2, \quad (14d)$$

$$a_5 = (2^{2R_{\text{tar}}} - 1). \quad (14e)$$

Using these definitions, the problem in (7) for the AF case can be written as

$$\min_{P_S, P_R \geq 0} P_S + \lambda P_R \quad (15a)$$

$$\text{subject to } a_2 P_S^2 + a_1 P_S + a_3 P_R^2 + a_4 P_R + a_5 \geq 0. \quad (15b)$$

Our approach to solving this problem will be to first obtain an expression for the optimal value of P_R as a function of P_S , which we will denote by $P_{R,C}(P_S)$. We will then solve the one-dimensional problem

$$\min_{P_S \geq 0} P_S + \lambda P_{R,C}(P_S). \quad (16)$$

The domain on which $P_{R,C}(P_S)$ is well defined is upper bounded by $P_{S,B}$ in (11c), which is the source power level that enables R_{tar} to be achieved using only the first time slot of each frame (and no power from the relay). The domain is strictly lower bounded by

$$P_{S,E} = (2^{2R_{\text{tar}}} - 1)/(\gamma_{SD}^2 + \gamma_{SR}^2), \quad (17)$$

which is the source power at which the required relay power becomes unbounded. On the domain $(P_{S,E}, P_{S,B}]$, we have [11]

$$P_{R,C}(P_S) = \frac{a_5 - a_1 P_S - a_2 P_S^2}{a_4 + a_3 P_S}. \quad (18)$$

Using that expression, it can be shown [11] that the stationary points of the function $P_S + \lambda P_{R,C}(P_S)$ can be expressed as the solutions of a quadratic equation in P_S , and hence closed-form expressions for these stationary points can be obtained in a straightforward way. The critical points of the problem in (16) are, therefore, those stationary points that lie in the interval $(P_{S,E}, P_{S,B})$ and the point $P_{S,B}$. By evaluating the objective at these points, the optimal value of P_S in the relay-active mode, and the corresponding value for P_R can be determined. Since there are at most three critical points, this enumeration approach is rather inexpensive to implement. Actually, since the operating point $(P_S, P_R) = (P_{S,B}, 0)$ is inferior to direct transmission (cf. Section IV-A), this point need not be explicitly considered.

Once $P_{S,ra}^*$ and $P_{R,ra}^* = P_{R,C}(P_{S,ra}^*)$ have been computed in the manner described above, the optimal operating power levels for the original problem in (5), P_S^* and P_R^* , can be determined by selecting the values associated with the smaller of $P_{S,ra}^* + \lambda P_{R,ra}^*$ and $P_{S,dir}^*$, as per the discussion that follows (7).

V. NUMERICAL EXAMPLE

In this section, we examine the average performance of the transmission schemes in Sections IV-A and IV-B. In both of these schemes the transmission is either in the relay-assisted mode or in the direct mode, depending on the channel realization, and hence we will call these schemes the “RDF or direct” and “AF or direct” schemes, respectively. The performance of these schemes is examined in an environment in which each link is modeled as an independent Rician block fading channel with additive white circular complex Gaussian

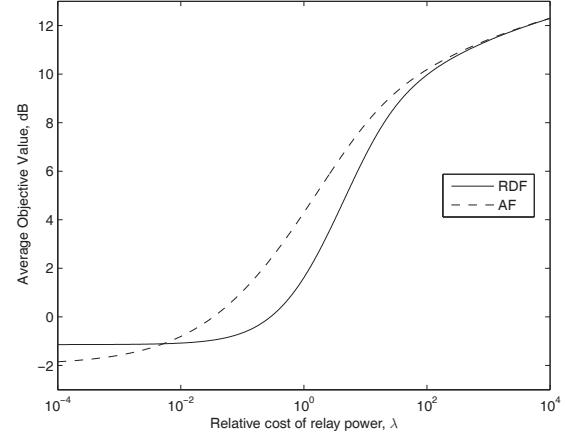


Fig. 4. Average cost of the power, on a decibel scale, against the relative cost of relaying, for the “RDF or direct” and “AF or direct” schemes in Section IV, under the Rician model in Section V. For reference, the average cost of direct transmission is 14.5 dB.

noise of unit variance at the receiver. The Rician “K-factor” on each link was chosen to be $K = 5$. Under a Rician model, the average power gain can be written in the form $\alpha^2 K / (K+1)$, and we consider a scenario in which $\alpha_{SR} = 1$, $\alpha_{RD} = 1$, and $\alpha_{SD} = 0.25$. The target rate for our example is chosen to be $R_{\text{tar}} = 0.25$ bits per (complex) channel use.

In Fig. 4 we plot the average value of the cost, $P_S + \lambda P_R$, of the power required to achieve the target rate against the relative cost of the relay power, λ . In Figs 5 and 6 we provide the corresponding plots for the source and relay powers. (The averages were computed over 10^6 channel realizations.) The qualitative characteristics of these figures is typical of the wide variety of examples that we have examined. In particular, the AF relaying strategy has the advantage that the relay can contribute even if the target rate cannot be supported on the source–relay link. This advantage becomes dominant when the relay power is cheap; see Figs 4 and 6. However, when the relay power is moderately priced, the advantages of the RDF scheme come to the fore. As the relay power becomes expensive, both of the schemes in Section IV resort to direct transmission more often (see Fig. 6), and, as one would expect, the cost of both schemes converges to that of direct transmission.

These observations suggest that for an infrastructure relay with a continuous power source, the amplify-and-forward strategy is not only simpler to implement, but may also reduce, on average, the overall cost of the power required to achieve the target data rate. These observations also suggest that for systems in which the relay is an idle battery-powered mobile station, the investment in terms of the additional implementation cost of employing regenerative decode-and-forward relaying provides a return in terms of reduced power requirements.

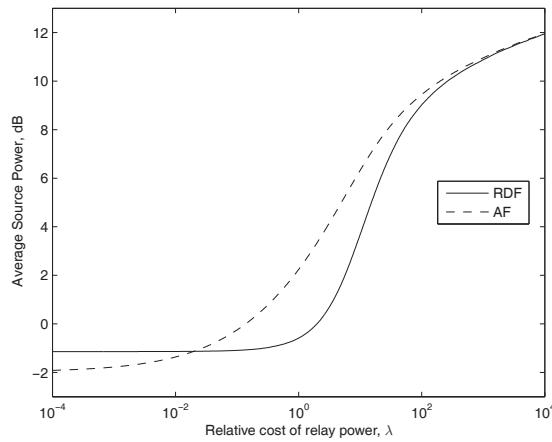


Fig. 5. Average value of the source power, P_S , against the relative cost of relaying, for the “RDF or direct” and “AF or direct” schemes in Section IV, under the Rician model in Section V. For reference, the average value of $P_{S,\text{dir}}^*$ is 14.5 dB.

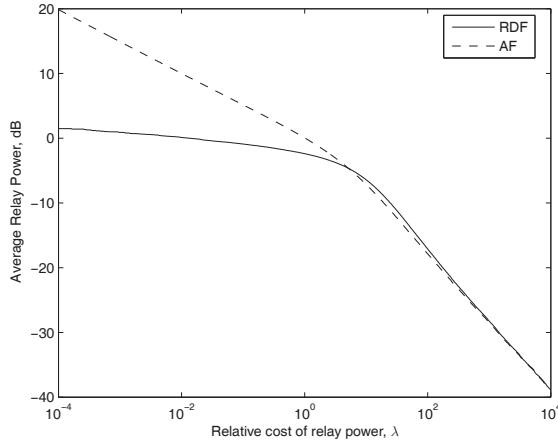


Fig. 6. Average value of the relay power, P_R , against the relative cost of relaying, for the “RDF or direct” and “AF or direct” schemes in Section IV, under the Rician model in Section V.

VI. CONCLUSION

In this paper, we have derived closed-form expressions for the power allocation that minimizes the cost of the power required to enable reliable communication at a specified target rate on a point-to-point link that is assisted by an orthogonal half-duplex relay. The relaying strategies that were considered are regenerative decode-and-forward (RDF) relaying and amplify-and-forward (AF) relaying. The simplicity of the obtained closed-form expressions sheds some light on the nature of quality-of-service problems in relay systems, and also offers the potential for further development. One immediate extension is that in which the relaying strategy is not specified *a priori*. If the relay is able to implement both strategies, then one can jointly select the transmission

mode, the relaying strategy and the power allocation by simply comparing the objective values obtained by each of the closed-form expressions obtained herein. That requires the comparison of at most four quantities and hence is quite straightforward to implement. Other extensions that are currently under investigation include extensions to other relaying strategies, the incorporation of robustness to uncertainty in the channel state information at the design node and to imperfect feedforward of the operating powers from the design node, and the development of power allocations for systems with statistical channel state information.

REFERENCES

- [1] E. C. van der Meulen, “A survey of multi-way channels in Information Theory: 1961–1976,” *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 1–37, Jan. 1977.
- [2] T. M. Cover and A. A. El Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1977.
- [3] A. Sendonaris, E. Erkip and B. Aazhang, “User cooperation diversity—Parts I and II,” *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1948, Nov. 2003.
- [4] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behaviour,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [5] G. Kramer, M. Gastpar and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
- [6] G. Kramer, I. Maric, and R. D. Yates, “Cooperative communications,” *Foundations and Trends in Networking*, vol. 1, no. 3–4, pp. 271–425, 2006.
- [7] A. Høst-Madsen and J. Zhang, “Capacity bounds and power allocation for the wireless relay channel,” *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.
- [8] Y. Liang and V. V. Veeravalli, “Gaussian orthogonal relay channels: Optimal resource allocation and capacity,” *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3284–3289, Sep. 2005.
- [9] S. Serbetli and A. Yener, “Relay assisted F/TDMA ad hoc networks: Node classification, power allocation and relaying strategies,” *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 937–947, June 2008.
- [10] W. Mesbah and T. N. Davidson, “Power and resource allocation for orthogonal multiple access relay systems,” *Eurasip J. Adv. Signal Processing*, vol. 2008, Article ID 476125, 15 pages, 2008. doi:10.1155/2008/476125.
- [11] F. Ahmed, “Optimal power allocation for orthogonal relay systems,” M.Sc. Report, School of Computational Engineering and Science, McMaster University, Hamilton, ON, Canada. May 2009.