# On the Power Allocation for Relay Networks with Finite-Alphabet Constraints

Weiliang Zeng*, Mingxi Wang†, Chengshan Xiao† and Jianhua Lu*
*State Key Laboratory on Microwave and Digital Communications
Tsinghua National Laboratory for Information Science and Technology
Dept. of Electronic Engineering, Tsinghua University, Beijing, China
†Dept. of Electrical & Computer Engineering, Missouri University of Science and Technology, MO, USA
Email: {zengwl, lujh}@wmc.ee.tsinghua.edu.cn, {xiaoc, mw9zd}@mst.edu

***Abstract*—In this paper, we investigate the optimal power allocation scheme for relay networks with finite-alphabet constraints. It has been shown that the previous work utilizing various design criteria with the Gaussian inputs assumption may lead to significant loss for a practical system with finite constellation set constraint, especially when signal-to-noise ratio (SNR) is in medium-to-high regions, or when the channel coding rate is medium to high. An optimal power allocation scheme is proposed to maximize the mutual information for the relay networks under discrete-constellation input constraint. Numerical examples show that significant gain can be obtained compared to the conventional counterpart for nonfading channels and fading channels. At the same time, we show that the large performance gain on the mutual information will also represent the large gain on the bit-error rate (BER), i.e., the benefit of the power allocation scheme predicted by the mutual information can indeed be harvested and can provide considerable performance gain in a practical system.**

## I. Introduction and Related Work

Cooperative relaying has been shown to provide reliable high data rate services in wireless networks without the need of multiple antennas at each node. These benefits can be further exploited by utilizing judicious cooperative strategies, see [1]–[11] and the references therein.

The existing design methods may be categorized into two groups: i) diversity oriented designs; and ii) transmission rate oriented designs. The first group usually achieves the steepest asymptotic slope (the highest diversity order) on the outage probability versus SNR curve, however, it may not obtain the highest possible coding gain, such as the distributed space-time coding (DST) in [2], [3] and the relaying selection scheme in [4], [5]. The second group often optimizes the performance with the Gaussian inputs assumption, for example, maximizing output SNR [6]–[9], minimizing mean square error (MSE) [6], [10] and maximizing channel capacity [6], [7], [11].

Although Gaussian inputs are capacity-achieving signaling, they can never be realized in practice. Rather, the inputs must be drawn from a finite constellation set (such as pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM) and phase shift keying (PSK) modulation) in a practical communication systems, which may significantly depart from the Gaussian idealization [12]–[14]. Yet, no solution has been found in the above work for the power allocation that maximizes the potential transmission rate, i.e., mutual information, with non-Gaussian inputs over the relay networks, and this is exactly the concern of this study.

For multiple-input multiple-output (MIMO) systems, it has been shown that the design from the standpoint of finite alphabet can result in significant performance improvement [12]. In this paper, we will see similar performance gains achieved in relay networks over the existing schemes utilizing the design criteria such as SNR, MSE, and channel capacity. At the same time, it has been validated that the large performance gain predicted by the mutual information with finite-alphabet constraints can indeed be harvested and will lead to considerable performance improvement in a practical system.

The rest of this paper is organized as follows. In Sec. II, we introduce the relay network model and the main problem. In Sec. III, we analyze the solution structure of the power allocation problem from the information theoretic and optimization theoretic point of view. We also utilize the receiver structure that has the near-capacity performance to validate the effectiveness of the proposed scheme. Sec. IV presents numerical results, followed by the conclusions in Sec. V.

*Notation*: Throughout this paper, we use boldface upper-case letters to denote matrices, boldface lower-case letters to denote column vectors, and italics to denote scalars. The superscripts $(\cdot)^T$ and $(\cdot)^H$ stand for transpose and conjugate transpose, respectively; $[\mathbf{A}]_{i,j}$ and $[\mathbf{A}]_{:,j}$ denote the $(i\text{th}, j\text{th})$ element and $j$th column of matrix $\mathbf{A}$, respectively; and $\|\mathbf{d}\|$ denotes Euclidean norm of vector $\mathbf{c}$. $\mathbf{I}$ denotes the identity matrix with the appropriate dimensions; diag$(\mathbf{c})$ denotes the diagonal matrix with diagonal elements given by the vector $\mathbf{c}$; Tr$(\mathbf{A})$ denotes the trace operation; $\Re$ denotes the real part of complex number; and $\mathbb{E}$ denotes statistical expectation. Likewise, all logarithms are to the base 2.

## II. System Model and Preliminaries

Consider a relay network with one transmit-and-receive pair, where the source node (s) attempts to communicate to the destination node (d) with the assistance of $k$ relays $(r_1, r_2, \cdots, r_k)$. The link from the source to the $i$th relay is denoted as $h_i$, the link from the $i$th relay to the destination is denoted as $g_i$, and the direct link from the source to the destination is denoted as $h_0$.

We model the source-relay (S-R), source-destination (S-D) and relay-destination (R-D) links as the quasi-static flat-fading channels, which are applicable to the scenarios of narrow-band transmissions in a low-mobility environment. We assume that the $i$th relay knows its own channels $h_i$ and $g_i$, and the destination obtains full knowledge of S-D, S-R and R-D channels.

We consider the average power constraint of each node for each time slot, e.g., the power used at the source and the $i$th relay should be less than $P_s$ and $P_r$, respectively.

The data transmission is over two time slots using two hops. The symbols transmitted by the source node in the first and second time slot are denoted as $x_1$ and $x_2$, respectively. They may be chosen from some complex-valued finite constellation $\mathcal{C}$. We assume that $\mathbb{E}[x_i] = 0$ and $\mathbb{E}[|x_i|^2] = 1$ for $i = 1, 2$. During the first time slot, the source node sends $\alpha_0\sqrt{P_s}x_1$. Let $y_{r_i}$ and $y_{d,1}$ be received signals at the $i$th relay node and the destination, respectively, which are given by

$$y_{r_i} = \alpha_0\sqrt{P_s}h_i x_1 + n_{r_i}, \tag{1}$$

$$y_{d,1} = \alpha_0\sqrt{P_s}h_0 x_1 + n_{d,1}, \tag{2}$$

where $n_{r_i}$ and $n_{d,1}$ are complex additive white Gaussian noise at the $i$th relay and the destination with zero mean and unit variance $\sim \mathcal{CN}(0, 1)$.

The $i$th relay node normalizes the received signal by a factor of $\sqrt{\mathbb{E}\left[|y_{r_i}|^2\right]}$ (so that the average energy is unity) and retransmits the signal

$$t_i = \sqrt{\frac{P_r}{\mathbb{E}\left[|y_{r_i}|^2\right]}}\alpha_i y_{r_i},\ i = 1, 2, \cdots, k \tag{3}$$

during the second time slot. At the same time, the source node sends $\alpha_{k+1}\sqrt{P_s}x_2$. Then the destination node receives a superposition of the relay transmissions and the source transmission during the second time slot according to

$$\begin{aligned} y_{d,2} &= \sum_{i=1}^{k} g_i t_i + \alpha_{k+1}\sqrt{P_s}h_0 x_2 + n_{d,2} \\ &= \sum_{i=1}^{k}\sqrt{\frac{P_s P_r}{1 + P_s|\alpha_0 h_i|^2}}\alpha_0\alpha_i h_i g_i x_1 + \sqrt{P_s}\alpha_{k+1}h_0 x_2 + v, \end{aligned} \tag{4}$$

where the effective noise $v \sim \mathcal{CN}(0, N_d)$ with $N_d = 1 + \sum_{i=1}^{k}\frac{P_r|\alpha_i g_i|^2}{1+P_s|\alpha_0 h_i|^2}$. We normalize $y_{d,2}$ by a factor $w = N_d^{1/2}$ in order to simplify the presentation. Finally, the effective input-output relation for the two-hop transmission can be summarized as

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{n}, \tag{5}$$

where $\mathbf{y} = [y_{d,1}\ y_{d,2}/w]^T$ is the received signal vector, and $\mathbf{G}$ is the effective channel matrix given by

$$\mathbf{G} = \left[\begin{array}{cc} \sqrt{P_s}\alpha_0 h_0 & 0 \\ \sum_{i=1}^{k}\sqrt{\frac{P_s P_r}{(1+P_s|\alpha_0 h_i|^2)w^2}}\alpha_0\alpha_i h_i g_i & \sqrt{\frac{P_s}{w^2}}\alpha_{k+1}h_0 \end{array}\right]. \tag{6}$$

$\mathbf{x} = [x_1\ x_2]^T$ is the transmitted signal vector, and $\mathbf{n} \in \mathbb{C}^{2\times 1}$ is the channel noise vector, assumed independent and identically distributed (i.i.d.) complex Gaussian with zero mean and unit variance, i.e., $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Equivalently, we can rewrite the effective channel $\mathbf{G}$ as $\mathbf{HP}$ with the channel related matrix

$$\mathbf{H} = \left[\begin{array}{ccccc} \sqrt{P_s}h_0 & 0 & \cdots & 0 & 0 \\ 0 & \sqrt{P_s P_r}h_1 g_1 & \cdots & \sqrt{P_s P_r}h_k g_k & \sqrt{P_s}h_0 \end{array}\right], \tag{7}$$

and the power allocation related matrix

$$\mathbf{P} = \left[\begin{array}{ccccc} \gamma_0 & \gamma_1 & \cdots & \gamma_k & 0 \\ 0 & 0 & \cdots & 0 & \gamma_{k+1} \end{array}\right]^T, \tag{8}$$

where $\gamma_0 = \alpha_0$; $\gamma_i = \alpha_0\alpha_i/\sqrt{\left(1 + P_s|\alpha_0 h_i|^2\right)w^2}$, $\forall i = 1, \cdots, k$; and $\gamma_{k+1} = \alpha_{k+1}/w$.

We should note that the coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k+1}$ are complex value, and the rationale of introducing them in the model is in fact quite intuitive. First, they can be used to control the average transmit power of each node at each time slot, which requires that

$$\alpha_i\alpha_i^* \le 1 \quad \forall i = 0, \cdots, k+1. \tag{9}$$

Hence the average power used at the $i$th relay node is $|\alpha_i|^2 P_r$. Second, the choice of the angles can be used to cancel the phases introduced from the channels and ensure that the signal components are added constructively at the receiver, i.e., $\arg\alpha_i = -(\arg h_i + \arg g_i)$, $i = 1, \cdots, k$, [7], [8], [10]. We also set $\arg\alpha_0 = \arg\alpha_{k+1} = -\arg h_0$, since the optimal choice of both angles can be realized by rotating the input constellations equivalently. What is left is the choice of their magnitude, so we will treat the power allocation coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k+1}$ as a set of real numbers in the sequel.

Our power allocation scheme is thus the design of coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k+1}$ to maximize the mutual information with finite-alphabet constraints. Note that for the proposed algorithm to be effectively implemented in practice, a low-rate feedback should be allowed from the destination to the source and relay nodes. The feedback is needed in the cooperative system since antennas are not located at a single terminal as in a MIMO system. This may result in small penalty on system performance, but the cost is often compensated by a significant performance gain at high SNR [1], [14].

## III. Optimal Power Allocation for Finite Alphabet Inputs

We consider the conventional equiprobable discrete signaling constellations such as $M$-ary PSK, PAM, or QAM, where $M$ is the number of points in the signal constellation. The mutual information between $\mathbf{x}$ and $\mathbf{y}$, with $\mathbf{H}$ and $\mathbf{P}$ known at the receiver, is $\mathcal{I}(\mathbf{x}; \mathbf{y})$ given by (10) at the top of next page [12], where $\mathbf{x}$ contains two symbols, taken independently from the $M$-ary signal constellation.

The problem that we pose is the determination of the coefficients $\alpha_0, \alpha_1, \cdots, \alpha_{k+1}$ that maximizes the mutual information $\mathcal{I}(\mathbf{x}; \mathbf{y})$ with given input distributions while satisfying individual power constraints, i.e.,

$$\max_{\alpha_0, \cdots, \alpha_{k+1}} \mathcal{I}(\mathbf{x}; \mathbf{y}) \tag{11}$$

subject to:

$$\alpha_i \le 1 \quad \forall i = 0, \cdots, k+1. \tag{12}$$

$$\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)=\log M-\frac{1}{2M^{2}}\sum_{m=1}^{M^{2}}\mathbb{E}_{\mathbf{n}}\left\{\log\sum_{k=1}^{M^{2}}\exp\left[-\left\|\mathbf{HP}\left(\mathbf{x}_{m}-\mathbf{x}_{k}\right)+\mathbf{n}\right\|^{2}+\left\|\mathbf{n}\right\|^{2}\right]\right\}. \tag{10}$$

$$\begin{aligned}\mathbf{E}&=\mathbb{E}\left\{\left[\mathbf{x}-\mathbb{E}\left(\mathbf{x}|\mathbf{y},\mathbf{H},\mathbf{P}\right)\right]\left[\mathbf{x}-\mathbb{E}\left(\mathbf{x}|\mathbf{y},\mathbf{H},\mathbf{P}\right)\right]^{H}\right\}\\&=\frac{\mathbf{I}}{2}-\frac{1}{2\left(\pi M\right)^{2}}\int_{\mathbf{y}}\frac{\left[\sum_{l=1}^{M^{2}}\mathbf{x}_{l}\exp\left(-\left\|\mathbf{y}-\mathbf{HPx}_{l}\right\|^{2}\right)\right]\left[\sum_{k=1}^{M^{2}}\mathbf{x}_{k}^{H}\exp\left(-\left\|\mathbf{y}-\mathbf{HPx}_{k}\right\|^{2}\right)\right]}{\sum_{m=1}^{M^{2}}\exp\left(-\left\|\mathbf{y}-\mathbf{HPx}_{m}\right\|^{2}\right)}d\mathbf{y}.\end{aligned} \tag{23}$$

### *A. Analysis from the Information Theoretical Point of View*

Applying the chain rule for mutual information [15], we have

$$\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)=\mathcal{I}\left(x_{2};\mathbf{y}\right)+\mathcal{I}\left(x_{1};\mathbf{y}|x_{2}\right), \tag{13}$$

where $\mathcal{I}\left(x_{2};\mathbf{y}\right)$ is the mutual information between $x_2$ and $\mathbf{y}$, and $\mathcal{I}\left(x_{1};\mathbf{y}|x_{2}\right)$ is the conditional mutual information between $x_1$ and $\mathbf{y}$ given $x_2$. The vector $\mathbf{y}$ is defined in (5), which can also be written as:

$$\mathbf{y}=\left[\mathbf{G}\right]_{:,1}x_{1}+\left[\mathbf{G}\right]_{:,2}x_{2}+\mathbf{n}. \tag{14}$$

From (14), we can verify that

$$\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)=\mathcal{I}\left(x_{2};\mathbf{y}\right)+\mathcal{I}\left(x_{1};\mathbf{y}|\alpha_{k+1}=0\right), \tag{15}$$

where $\mathcal{I}\left(x_{1};\mathbf{y}|\alpha_{k+1}=0\right)$ is the mutual information between $x_1$ and the received signal $\mathbf{y}$ if the source node does not transmit at the second time-slot.

Since $\mathcal{I}\left(x_{2};\mathbf{y}\right)>0$ for $h_0,\alpha_{k+1}\neq 0$, it follows that to maximize the mutual information $\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)$, the source should always transmit at the second time slot [11]. For the same reason, the source should transmit at the first time slot. Based on the above discussions, we can state the following lemma:

*Lemma 1:* The power allocation in the two time slots at the source node is nonzero, i.e., $\alpha_0,\alpha_{k+1}\neq 0$, if the channel between the source and the destination node $h_0$ is nonzero.

### *B. Analysis from the Optimization Theoretical Point of View*

We should note that the constraint (12) is convex in the coefficient matrix $\mathbf{C}=\mathrm{diag}([\alpha_0\,\alpha_1\,\cdots\alpha_{k+1}])$. The cost function (11), however, is nonconcave in the power allocation coefficient $\alpha_i$ for the general case [16]. In the sequel, we capitalize on the relationship between the mutual information and the minimum mean square error (MMSE) matrix to obtain the power allocation scheme for arbitrary input distributions.

*Theorem 1:* The optimal power allocation coefficients $\left[\alpha_0^\star\,\alpha_1^\star\,\cdots\alpha_{k+1}^\star\right]$ that solves (11) subject to (12) satisfy:

$$\left.\frac{\partial\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial\alpha_{i}}\right|_{\alpha_i=\alpha_i^\star}=\lambda_{i} \tag{16}$$

$$\lambda_{i}\left(\alpha_{i}^{\star}-1\right)=0 \tag{17}$$

$$\lambda_{i}\geq 0 \tag{18}$$

with

$$\frac{\partial\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial\alpha_{i}}=\sum_{j=0}^{k+1}\frac{\partial\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial\gamma_{j}}\frac{\partial\gamma_{j}}{\partial\alpha_{i}}, \tag{19}$$

and

$$\frac{\partial\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial\gamma_{i}}=\Re\left[\mathbf{H}^{H}\mathbf{HPE}\right]_{i+1,1},\quad\forall i=0,\cdots,k \tag{20}$$

$$\frac{\partial\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)}{\partial\gamma_{k+1}}=\Re\left[\mathbf{H}^{H}\mathbf{HPE}\right]_{k+2,2} \tag{21}$$

where $\mathbf{E}$ is the MMSE matrix given by (23) at the top of this page.

*Proof:* The possible solution to (11) subject to (12) is characterized by the *Karush-Kuhn-Tucker* theorem [17], which gives necessary conditions, known as the KKT or first order conditions. To investigate stationary points of the problem (11) we formulate the Lagrangian

$$\mathcal{L}\left(\mathbf{P},\lambda\right)=-\mathcal{I}\left(\mathbf{x};\mathbf{y}\right)+\sum_{i=0}^{k+1}\lambda_{i}\left(\alpha_{i}-1\right), \tag{24}$$

in which the Lagrangian multipliers $\lambda_i$, $i=1,\cdots,k+1$, are chosen to satisfy the power constraints. Then the first order conditions are given by (16) to (18). The partial derivative $\partial\mathcal{I}/\partial\alpha_i$, $i=0,\cdots,k+1$, can be proved by employing the techniques developed in [18], [19] for derivatives of mutual information, the techniques developed in [20] for matrix differentiation, and the chain rule for multiple variables. □

Typically, it is involved to calculate the MMSE matrix (23), especially for large input dimensions $M$. But we have been able to estimate the matrix $\mathbf{E}$ using Monte Carlo methods. Hence, we can solve this problem using gradient-based methods according to the gradient of mutual information (19). Since the cost function (11) is nonconcave for the general case, it is possible that (11) has local maxima. Therefore, we should perform the algorithm with multiple initial points and keep the power allocation coefficients offering the largest mutual information.

Finally, the destination node will notify the source and each relay node of its assigned transmission power. In this way, the instantaneous mutual information is maximized for each set of channel realizations. Notice that the resulting optimal power allocation scheme is significantly different from the existing ones in the conventional setting both due to the presence of finite-alphabet constraints and the multiplexing structure capitalized in the relay networks. The results in section IV show the significant gains obtained compared to the existing methods.

### *C. Iterative Detection and Decoding at the Destination Node*

To evaluate the advantage of the proposed method in a more practical may, we utilize the "turbo principle" at the

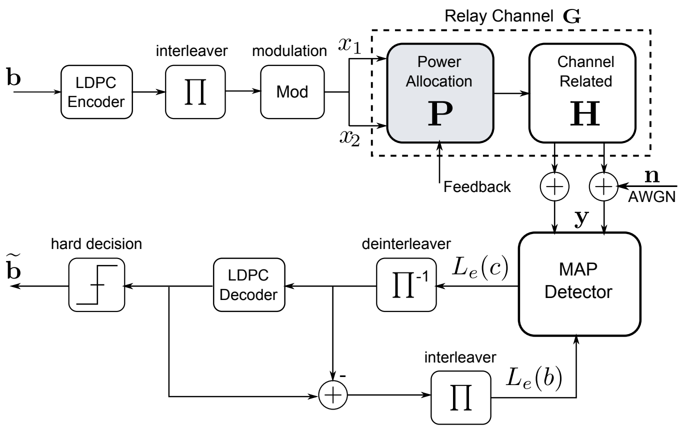


Fig. 1. Block diagram of the relay network with iterative detection and decoding at the destination node.

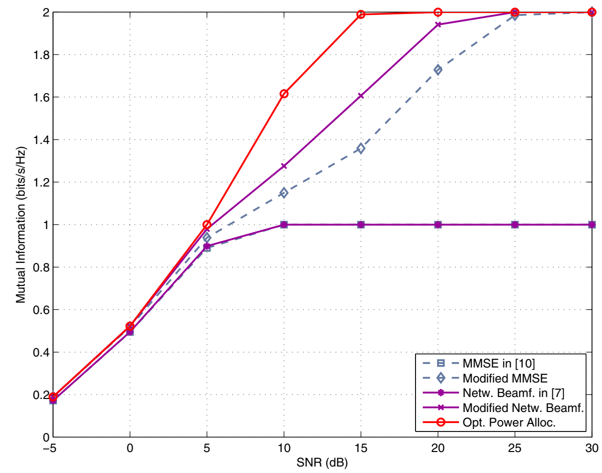


Fig. 2. Mutual Information of the relay network for a fixed channel with QPSK inputs.

destination node [21], [22], which is illustrated in Fig. 1. The signal sequence $\mathbf{b}$ at the source node is encoded by the capacity achievable codes, e.g., low-density parity-check (LDPC) codes, interleaved, and mapped according to the conventional equiprobable discrete signaling constellations. Then it is divided into $x_1$ and $x_2$, and transmitted at the two time slots, respectively.

At the receiver, the maximum a posteriori (MAP) detector takes channel observations $\mathbf{y}$ and *a priori* knowledge $L_e(b)$ from the decoder and computes new information $L_e(c)$ for each of the coded bits. In this way, the extrinsic information between the MAP detector and decoder is exchanged in an iterative fashion until desired performance is achieved. It has been shown that the iterative processing is very effective that can achieve near-capacity performance.

## IV. Numerical Results

Computer simulation was carried out to validate the performance of the proposed scheme. For the sake of completeness, in all the figures, we show the performance corresponding to MMSE strategy with local power constraint in [10] and network beamforming in [7]. To ensure a fair comparison, we also show the performance of modified MMSE strategy and modified network beamforming (sending different symbols at the second time slot, rather than sending the same symbols or sleeping). We consider a three-relay network with the same transmit power at the source and each relay node, i.e., $P_s = P_r = P$, which is indicated by the horizontal axis in the following figures.

We look first at a fixed (non-fading) system with the channel coefficient $h_0 = 0.5$, $h = [0.7, -0.7, 1j]$ and $g = [0.9j, 2.1, 0.3]$. The instantaneous mutual information that can be achieved by different schemes is shown in Fig. 2, in which the information symbol $\mathbf{x}$ is modulated as quadrature phase shift keying (QPSK). From Fig. 2, we have several observations. First, the performance loss of the MMSE and network beamforming is small in the low SNR region, and large in the high SNR region. This is because both schemes maximize the power gain, which is much impressive compared to the degree-of-freedom gain at low SNR [23]. At high SNR, however, a degree-of-freedom gain is much more important, which is not provided by the original MMSE and network beamforming schemes [7], [10], since the source node transmits the same symbols for the two time slots. Hence, the mutual information is bounded by 1 bit/s/Hz. Second, the modified MMSE and modified network beamforming are not bounded by 1 bit/s/Hz, since we remove the constraint that sends the same symbols at the two time slots. Moreover, the proposed power allocation method results in significant gain on mutual information when SNR is in medium-to-high regions, or when the channel coding rate is medium to high. For example, it is about 4dB and 8dB better than the modified network beamforming and modified MMSE scheme when the channel coding rate is 3/4.

In Fig. 3, we show that the large performance gain on the mutual information will also represent the large performance gain on the BER. To validate the advantage of the power allocation scheme, we realize the simulation model illustrated in Sec. III-C. The coding length is 1800; the coding rate is 3/4; and the iteration between the MAP detector and the LDPC decoder is 5. Then we compare the optimal power allocation with modified MMSE and modified network beamforming. It is worth noting that the benefit of the power allocation scheme predicted by the mutual information can indeed be realized, and the power allocation coefficients that are "blessed" by the mutual information formula (10) can provide considerable performance gain in a practical system.

Then we work on the Rayleigh fading channel, and consider the average mutual information achieved by different methods. We assume the channels of S-R and R-D have the same average SNR, and are 10dB better than the S-D channel, considering the practical deployment of the relay nodes. Fig. 4 depicts the average mutual information of the relay network with QPSK inputs. The MMSE and network beamforming saturates very quickly, while the modified ones perform much better. However, they still have about 3dB to 15dB loss compared to the optimal power allocation when the channel coding rate is 3/4.

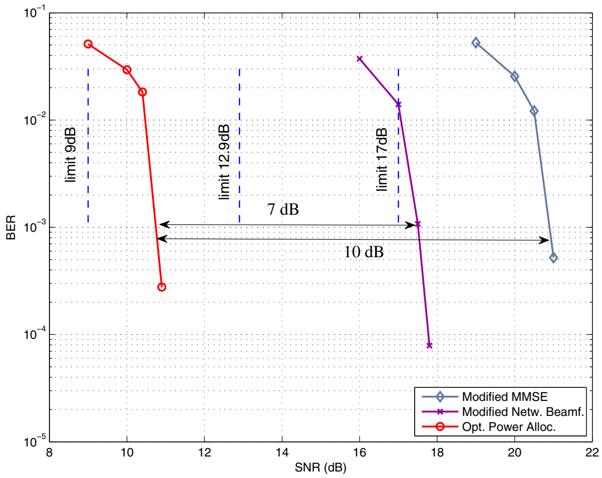


Fig. 3. Bit-error performance comparison for optimal power allocation, modified MMSE and modified network beamforming. The coding rate is 3/4; the coding length is 1800; and the iteration between the MAP detector and the LDPC decoder is 5.

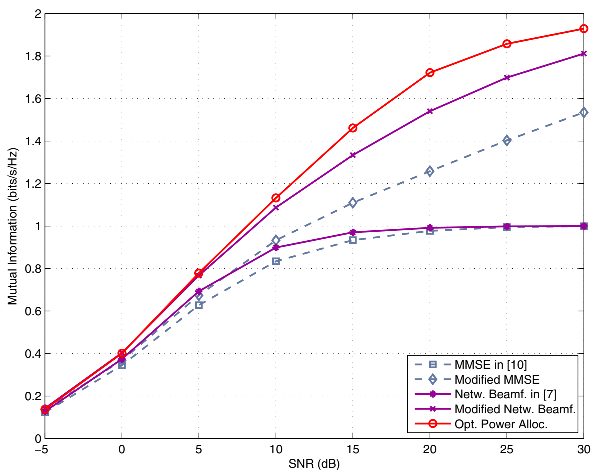


Fig. 4. Average mutual Information of the relay network over Rayleigh fading channels with QPSK inputs.

## V. Conclusion

In this paper, we have studied the optimal power allocation for dual-hop wireless relay networks. In contrast with the previous work utilizing various design criteria with the unrealistic Gaussian inputs assumption, the proposed scheme attempts to maximize the mutual information for the relay networks from the standpoint of discrete-constellation inputs. To determine the optimal power allocation policy, we capitalized on the relationship between mutual information and MMSE. Numerical examples have shown that significant gains can be obtained compared to the conventional counterpart for nonfading channels and fading channels, especially when SNR is in medium-to-high regions, or when the channel coding rate is medium to high. Likewise, it has been shown that the large performance gain on the mutual information can represent the large gain on the bit-error rate, i.e., the benefit of the power allocation scheme predicted by the mutual information can indeed be realized and can provide considerable performance gain in a practical system.


## Acknowledgment

This work was supported in part by the National Basic Research Program of China (2007CB310601) and the U.S. National Science Foundation under Grant CCF-0915846. This work has been carried out while W. Zeng is a visiting scholar at Missouri University of Science and Technology.


## References


[1] Y. Hong, W. Huang, F. Chiu, and C. Kuo, "Cooperative communications in resource-constrained wireless networks," *IEEE Signal Process Mag.*, vol. 24, no. 3, p. 47, 2007.

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

[3] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, p. 3524, 2006.

[4] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, March 2006.

[5] W. Zeng, C. Xiao, Y. Wang, and J. Lu, "Opportunistic relaying for multi-antenna cooperative decode-and-forward relay networks," in *Proc. IEEE ICC*, 2010.

[6] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5062–5073, 2008.

[7] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2499–2517, Jun. 2009.

[8] G. Zheng, K. Wong, A. Paulraj, and B. Ottersten, "Collaborative-relay beamforming with perfect CSI: Optimum and distributed implementation," *IEEE Signal Process Lett.*, vol. 16, no. 4, pp. 257–260, 2009.

[9] A. Gershman, N. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming: From receive to transmit and network designs," *IEEE Signal Process Mag., Special Issue on Convex Optim. for Signal Process*, 2010.

[10] N. Khajehnouri and A. Sayed, "Distributed MMSE relay strategies for wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 7, p. 3336, 2007.

[11] R. Nabar, H. Bolcskei, and F. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, 2004.

[12] C. Xiao and Y. R. Zheng, "On the mutual information and power allocation for vector gaussian channels with finite discrete inputs," in *Proc. IEEE Globecom*, Nov. 2008, pp. 1–5.

[13] A. Lozano, A. Tulino, and S. Verdu, "Optimum power allocation for parallel gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, 2006.

[14] G. Caire and K. Kumar, "Information theoretic foundations of adaptive coded modulation," *Proc. IEEE*, vol. 95, no. 12, pp. 2274–2298, 2007.

[15] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, 2006.

[16] F. Pérez-Cruz, M. Rodrigues, and S. Verdú, "Optimal precoding for digital subscriber lines," in *Proc. IEEE ICC*, 2008.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[18] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

[19] D. Palomar and S. Verdú, "Gradient of mutual information in linear vector gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, 2006.

[20] J. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 2nd ed. Wiley, 1999.

[21] B. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, 2003.

[22] R. Koetter, A. Singer, and M. Tuchler, "Turbo equalization," *IEEE Signal Process Mag.*, vol. 21, no. 1, pp. 67–80, 2004.

[23] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.