Modulation Diversity Design in Cooperative Relay and HARQ Transmission

Wenhao Wu, Hans Mittelmann, and Zhi Ding

Abstract

Modulation diversity (MoDiv) is a practically useful transmission enhancement technique that utilizes different modulation mappings to reduce packet loss rate and achieve higher transmission efficiency. MoDiv is particularly effective in hybrid-ARQ (HARQ) systems. In this paper, we study the optimization of MoDiv in a coordinated relay-HARQ network to reduce packet loss. We formulate the design optimization of MoDiv into a quadratic three-dimensional assignment problem (Q3AP), which we solve using a modified iterated local search (ILS) method. Numerical results demonstrate clear performance gain over simple relay/retransmissions and a heuristic design under fading channels.

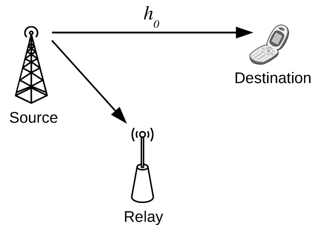
Index Terms

Modulation diversity, relay, HARQ, Q3AP.

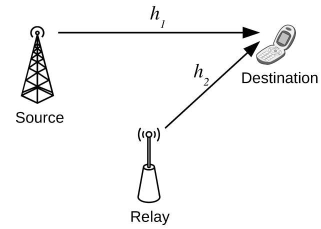
I. Introduction

In wireless data communication systems, high rate transmission and poor channel conditions often lead to reception errors. To recover lost packet due to transmission errors, Automatic Repeat reQuest (ARQ) or Hybrid ARQ (HARQ) are important mechanism for better reliability at network layer [1] and PHY layer [2]. HARQ in conjunction with relay networks has attracted great deal of research interest in recent years [3]. Because symbols transmitted in practice often utilize linear modulations of finite-size constellation (e.g., Q-ary PSK, or Q-ary QAM), the performance of cooperative relay-HARQ systems can benefit from Modulation Diversity (MoDiv) [4], in which each group of $\log_2 Q$ bits are mapped to different constellation points across different links and in different transmissions to better exploit channel diversity.

This work was supported in part by the National Science Foundation under Grants CNS-1443870, ECCS-1307820, and CCF-1321143. The work of the second author was supported in part by AFOSR under grant FA9550-12-1-0153.



(a) Phase 1



(b) Phase 2

Fig. 1. Cooperative relay-HARQ networks.

There are a number of existing research works on MoDiv for HARQ [5], relay network [6], [7] and relay-HARQ system [8], [9]. Despite the encouraging performance gain, these works all consume substantial bandwidth for which (re)transmissions on the Source-Destination (S-D) link and the Relay-Destination (R-D) links occupy orthogonal bands or time slots. For better spectral efficiency, we study MoDiv design in a relay-HARQ scenario of Fig. 1 where the S-D and R-D links are co-channel and additive as in [10], [11], specifically forming a 2-by-1 multiple-input single-output (MISO) diversity transmission. For this system, we formulate the bit error rate (BER) minimization MoDiv design into a quadratic three-dimensional assignment problem (Q3AP). Q3AP is NP-hard, which means that our MoDiv design cannot be solved exactly even for a moderate size constellation such as 16-QAM. Fortunately, there are a number of heuristic algorithms [12] for its reduced form, known as Quadratic Assignment Problem (QAP), that results in high-quality solutions over the QAPLIB dataset. In this work, we adopt a modified Iterated Local Search (ILS) method that is efficient in solving the Q3AP formulated from the MoDiv design problem for minimizing BER. Our numerical results demonstrate significant BER reduction over non-MoDiv (simple) HARQ re-transmissions. Moreover, we also show that when network channels are in heavy fading, comparable performance gain is offered by a low complexity heuristic MoDiv scheme.

In this manuscript, Section II first describes the relay-HARQ cooperation model. Section III presents the optimal MoDiv design problem and propose different solutions. Section IV provides and compares the numerical results to illustrate the benefit of MoDiv. Finally, Section V concludes this work.

II. System Model

We consider the cooperative relay-HARQ network shown in Fig. 1, in which there are two transmission phases. In phase 1, source node broadcasts its packet to both destination and relay. In phase 2 of this HARQ setup, upon packet loss notification, both the relay and the source nodes cooperatively retransmit the lost packet information to the destination. Our goal is to optimize H-ARQ constellation mappings to achieve modulation diversity so as to minimize the packet BER at the receiver.

Denote $\mathcal{C}$ as fixed constellation used by this relay network whose size is $Q = |\mathcal{C}|$. In the first phase, the source converts a bit sequence of length $\log_2 Q$ into symbols with Gray mapping $\psi_0 : \{0, \ldots, Q-1\} \to \mathcal{C}$. The bit sequence can be indexed by its decimal equivalence $p \in \{0, \ldots, Q-1\}$. The source transmits $\psi_0[p]$ to the destination via channel h_0, though $\psi_0[p]$ is simultaneously received by the decode-and-forward (DF) relay. We assume that the relay is placed strategically such that it has negligible decoding error rate as in [9], [8]. Upon receiving a request for retransmission, the second phase begins with the source and the relay remapping p into $\psi_1[p]$ and $\psi_2[p]$, respectively. In general, we have 3 distinct mappings $\psi_1 \neq \psi_0$ and $\psi_2 \neq \psi_0$. The remapped symbols are transmitted simultaneously on the same frequency band to the destination via channels h_1 and h_2. In summary, the received signals at the destination during the two phases are, respectively,

$$y_1 = h_0\psi_0[p] + v_1, \tag{1a}$$

$$y_2 = h_1\psi_1[p] + h_2\psi_2[p] + v_2, \tag{1b}$$

where $v_1, v_2 \sim \mathcal{CN}(0, \sigma_v^2)$ are additive channel noises. Throughout this work, we assume fading wireless channels h_0, h_1 and h_2 to follow independent Rician distribution.

Assuming that the destination acquires perfect channel state information (CSI). Based on the received symbols y_1 and y_2, the destination decodes the data by identifying index p via maximum likelihood (ML) detection:

$$\min_{\hat{p}} |y_1 - h_0\psi_0[\hat{p}]|^2 + |y_2 - h_1\psi_1[\hat{p}] - h_2\psi_2[\hat{p}]|^2. \tag{2}$$

III. Optimal Constellation Mapping for Modulation Diversity

In this section we first formulate the minimum BER design of MoDiv into a Q3AP problem. We then elaborate on the numerical approach for computing the input cost matrix of the Q3AP problem. We then provide an efficient algorithm to obtain numerical Q3AP solution.

A. BER Minimization via Q3AP solution

Assume that the information-bearing index p follows a uniform distribution, the BER can be upper-bounded and approximated using the pair-wise error probability (PEP) [5]

$$P_{\mathrm{BER}} \approx \sum_{p=0}^{Q-1}\sum_{q=0}^{Q-1}\frac{B[p,q]}{Q}P_{\mathrm{PEP}}(q|p), \tag{3}$$

where $B[p,q]$ is the Hamming distance between the binary representation of p and q divided by $\log_2 Q$ and $P_{\mathrm{PEP}}(q|p)$ is the probability for the ML decoder to prefer q over p when p is actually transmitted. According to (2), we have

$$P_{\mathrm{PEP}}(q|p) = P_{h_0,h_1,h_2,v_1,v_2}\{\delta(p,q,\psi_1,\psi_2) < 0\}. \tag{4}$$

in which, given random channels and random noise variables h_0, h_1, h_2, v_1, v_2, δ is defined as

$$\begin{aligned}\delta = &|h_0(\psi_0[p]-\psi_0[q])+v_1|^2 - |v_1|^2 + \\ &|h_1(\psi_1[p]-\psi_1[q]) + h_2(\psi_2[p]-\psi_2[q]) + v_2|^2 - |v_2|^2.\end{aligned} \tag{5}$$

In other words, given indices p, q and the remapping scheme ψ_1, ψ_2, the pairwise error event is equivalent to $\delta < 0$. In order to formulate the Q3AP problem, we introduce binary variables

$$x_{pij} = \begin{cases} 1, & \text{if } \psi_1[p] = \psi_0[i] \text{ and } \psi_2[p] = \psi_0[j] \\ 0, & \text{otherwise.}\end{cases}$$

Denote $\mathbf{x} = \{x_{pij} |\ p, i, j = 0, \ldots, Q-1\}$ and constraint set

$$\mathcal{P} = \left\{\mathbf{x} : \sum_{p=0}^{Q-1} x_{pij} = 1, x_{pij} \in \{0,1\}\right\}. \tag{6}$$

We further denote $\mathcal{I}$ and $\mathcal{J}$ as in (6) by replacing the summation index p with i and j, respectively. Then from (3)(4)(5), the BER minimization MoDiv scheme $\min_{\psi_1,\psi_2} P_{\mathrm{BER}}$ becomes

$$\begin{aligned}&\min_{\mathbf{x}} \sum_{p=0}^{Q-1}\sum_{i=0}^{Q-1}\sum_{j=0}^{Q-1}\sum_{q=0}^{Q-1}\sum_{k=0}^{Q-1}\sum_{l=0}^{Q-1} c_{pijqkl}x_{pij}x_{qkl}, \\ &\text{s.t. } \mathbf{x} \in \mathcal{P}\cap\mathcal{I}\cap\mathcal{J}.\end{aligned} \tag{7}$$

in which

$$c_{pijqkl} = \frac{B[p,q]}{Q}P_{h_0,h_1,h_2,v_1,v_2}\{\delta(p,i,j,q,k,l) < 0\}, \tag{8}$$

$$\begin{aligned}\delta = &|h_1(\psi_0[i]-\psi_0[k]) + h_2(\psi_0[j]-\psi_0[l]) + v_2|^2 \\ &+ |h_0(\psi_0[p]-\psi_0[q]) + v_1|^2 - |v_1|^2 - |v_2|^2.\end{aligned} \tag{9}$$

B. Computation of the Pair-wise Symbol Error Rate

In this section we focus on computing parameters $\{c_{pijqkl}\}$ of the Q3AP problem. According to (8), the key to evaluating c_{pijqkl} lies in the evaluation of $P_{h_0,h_1,h_2,v_1,v_2}\{\delta(p,i,j,q,k,l)$, i.e. the cumulative distribution function (CDF) of the random variable $\delta(p,i,j,q,k,l)$ of (5). Define the moment generating function (MGF)

$$\Phi_\delta(\omega) = \mathbb{E}_\delta[\exp(-\omega\delta)].$$

Under the well known Rician channel model, $h_m \sim \mathcal{CN}(\mu_{h_m}, \sigma^2_{h_m}), m = 0, 1, 2$, we can extend the method proposed in [5], [13] to compute $P_{h_0,h_1,h_2,v_1,v_2}\{\delta(p,i,j,q,k,l) < 0\}$:

$$P_{h_0,h_1,h_2,v_1,v_2}\{\delta(p,i,j,q,k,l) < 0\} \approx \frac{1}{2v}\sum_{t=1}^{v} \mathrm{Re}\left\{\Phi_\delta(\xi + \mathrm{j}\xi\tau_\mathrm{t})\right\} + \tau_\mathrm{t}\mathrm{Im}\left\{\Phi_\delta(\xi + \mathrm{j}\xi\tau_\mathrm{t})\right\}, \quad (10)$$

where $\tau_t = \tan((t-1/2)\pi/v)$ and $\mathrm{Re}\{\cdot\}$, $\mathrm{Im}\{\cdot\}$ denote the real and imaginary parts, respectively. The parameter ξ is selected to ensure convergence of the integration and $\xi = 1/4$ was suggested in [13]. The size v of the expansion (10) needs to be large when $P_{h_0,h_1,h_2,v_1,v_2}\{\delta(p,i,j,q,k,l) < 0\}$ is small in order to maintain an acceptable numerical accuracy.

To compute $\Phi_\delta(\omega)$, let us denote Gaussian random vectors $\mathbf{z}_1 = [h_0, v_1]^T$, $\mathbf{z}_2 = [h_1, h_2, v_2]^T$, such that $\mathbf{z}_m \sim \mathcal{CN}(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$, $m = 1, 2$, where

$$\boldsymbol{\mu}_1 = [\mu_{h_0}, 0]^T, \ \boldsymbol{\mu}_2 = [\mu_{h_1}, \mu_{h_2}, 0]^T, \quad (11)$$

$$\boldsymbol{\Sigma}_1 = \mathrm{diag}\left(\sigma^2_{h_0}, \sigma^2_v\right), \ \boldsymbol{\Sigma}_2 = \mathrm{diag}\left(\sigma^2_{h_1}, \sigma^2_{h_2}, \sigma^2_v\right). \quad (12)$$

Then (5) can be rewritten as $\delta = \mathbf{z}_1^H\mathbf{A}_1\mathbf{z}_1 + \mathbf{z}_2^H\mathbf{A}_2\mathbf{z}_2$, where

$$\mathbf{A}_1 = \begin{bmatrix} |e_{pq}|^2 & e^*_{pq} \\ e_{pq} & 0 \end{bmatrix}, \ \mathbf{A}_2 = \begin{bmatrix} |e_{ik}|^2 & e^*_{ik}e_{jl} & e^*_{ik} \\ e_{ik}e^*_{jl} & |e_{jl}|^2 & e^*_{jl} \\ e_{ik} & e_{jl} & 0 \end{bmatrix}, \quad (13)$$

where $e_{ab} = \psi_0[a] - \psi_0[b]$. Thus, the MGF is

$$\Phi_\delta(\omega) = \sum_{m=1,2} \frac{\exp(-\omega\boldsymbol{\mu}_m^H\mathbf{A}_m(\mathbf{I}+\omega\boldsymbol{\Sigma}_m\mathbf{A}_m)^{-1}\boldsymbol{\mu}_m)}{\det(\mathbf{I}+\omega\boldsymbol{\Sigma}_m\mathbf{A}_m)}. \quad (14)$$

Note that for any instance of Q3AP a total number of Q^6 coefficients must be computed. Fortunately, the MoDiv design is based on statistical CSI which does not require online computation of the Q^6 coefficients in real time. In fact, the optimized MoDiv can be precomputed and stored a priori in our network nodes. In our simulation, we implement the above procedure for 16-QAM and 32-QAM with in C++ on a workstation with 48 cores and finished the computation in several days for a $Q = 32$ case and a few hours for a 16-QAM case. For larger constellation such as 64-QAM, however, the time and spacial complexity may still be impractical. We will address in future works new means to reduce this complexity by imposing rules to restrict the remapping schemes.

C. Q3AP Solution

For modest constellation sizes such as 16-QAM and 32-QAM, it is impractical to apply the exact branch-and-bound algorithm [14]. Also, our tests show that they do not have enough symmetry to exploit for faster solution as does the 16-PSK constellation [15]. Consequently, the MoDiv problem is solved with the ILS method [14] extended from its QAP version [16]. Starting from two random initial mappings $\psi_1^{(0)}, \psi_2^{(0)}$, the algorithm executes a local search by exchanging the mapping of exactly 2 indices whenever a reduction in the objective function is made. This approach can lower the objective function and update the mapping locally. When the process hits a local minimum, it executes a perturbation step by exchanging the mapping of k_p indices, where k_p is adaptively adjusted within $[k_{p,min}, k_{p,max}]$. The perturbation is accepted with a probability defined as in simulated annealing, after which the local search is restarted from the new mappings until the stopping criterion is satisfied.

D. A Heuristic MoDiv Scheme

For Rician fading channels, there is a strong Line Of Sight (LOS) component in h_1 and h_2. Our Q3AP solution benefits from 2 different gains. Firstly, by allowing $\psi_1 \neq \psi_0$ and $\psi_2 \neq \psi_0$, we achieves the signal space diversity gain just as existing MoDiv schemes. Secondly, by jointly designing ψ_1 and ψ_2 and allowing $\psi_1 \neq \psi_2$, we achieve the cooperative gain between the source and the relay. When the channels are independently Rayleigh, however, there is little cooperative gain to exploit and we can greatly simplify the MoDiv design problem by fixing $\psi_1 = \psi_2$. Though we can solve the resulting QAP problem rigorously, from (5), we can take a simple heuristic approach noticing that the two indices mapped to two symbols close to each other in Phase 1 should be mapped to two symbols far apart in Phase 2. Based on such heuristic, $\psi_1 = \psi_2$ can be designed by adapting the trans-modulation scheme of Seddik [6] for 16-QAM and 64-QAM. For 32-QAM constellation, a similar heuristic MoDiv design may be extended by remapping

(a) ψ_1 (b) ψ_2

Fig. 2. Q3AP optimized MoDiv schemes.

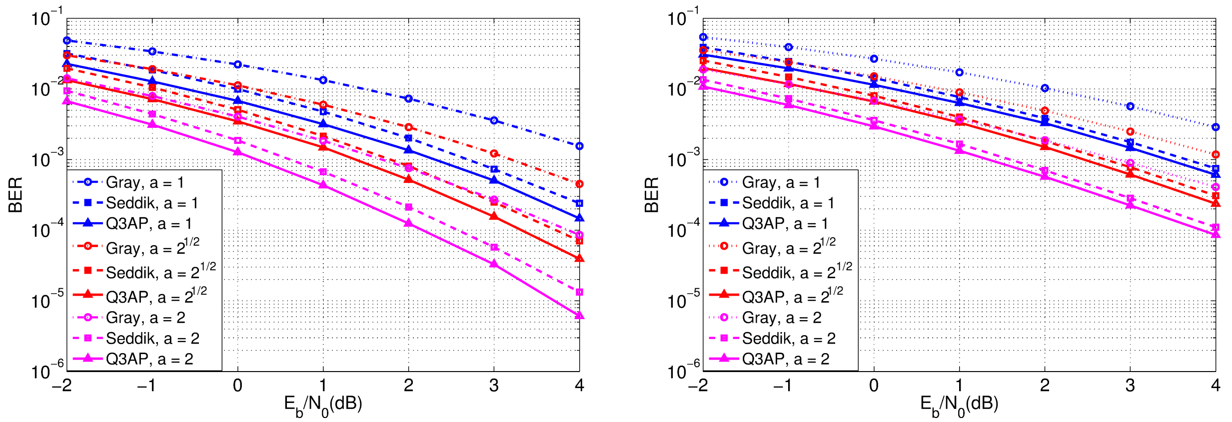
the 3 Most Significant Bits (MSBs) and the 2 Least Significant Bits (LSBs) separately. We will show in the next section that when the channels experience deep fade, such heuristic remapping method offers comparable performance gain as our Q3AP-based MoDiv.

IV. Numerical Results

In our simulation, all Rician fading channels are assumed to have the same Rician parameter K. Also we assume that during the second phase, the phases of the line of sight (LOS) components of channels h_1 and h_2 can be aligned at the source and relay, respectively. Consequently, we define $\mu_{h_0} = \mu_{h_1} = \sqrt{K/(K+1)}$, $\mu_{h_2} = a\sqrt{K/(K+1)}$, $\sigma^2_{h_0} = \sigma^2_{h_1} = 1/(K+1)$ and $\sigma^2_{h_2} = |a|^2/(K+1)$, where a denotes the ratio between the amplitude of the LOS component of the relay-to-destination and the S-D link. The noise power is parameterized with E_b/N_0 of the S-D link.

First, we provide an example of Q3AP optimized MoDiv for 16-QAM. For $E_b/N_0 = 2$dB, $K = 10$ and $a = 1$, the remapping scheme ψ_1 and ψ_2 are depicted in Fig. 2. The Gray mapping ψ_0 and the heuristic remapping $\psi_1 = \psi_2 = \psi_S$ are defined according to [6]. The results justify the use of heuristics discussed in Section III-D as both ψ_1 and ψ_2 are essentially very close to ψ_S. However, we note that in general $\psi_1 \neq \psi_2$.

Next we compare the empirical BER for three mapping schemes: Q3AP-optimized MoDiv, the heuristic MoDiv $\psi_1 = \psi_2 = \psi_S$ and zero MoDiv at all, i.e. $\psi_1 = \psi_2 = \psi_0$. The empirical BER is evaluated by applying the ML demodulator in (2) and over $M = 10^7$ randomly generated channels, noise instances and data symbols. In Fig. 3, there is a considerable performance gain for MoDiv versus no MoDiv. Moreover, the gain of the Q3AP optimized MoDiv over $\psi_1 = \psi_2 = \psi_S$ indicates that different rearrangements at the source and relay during the cooperative



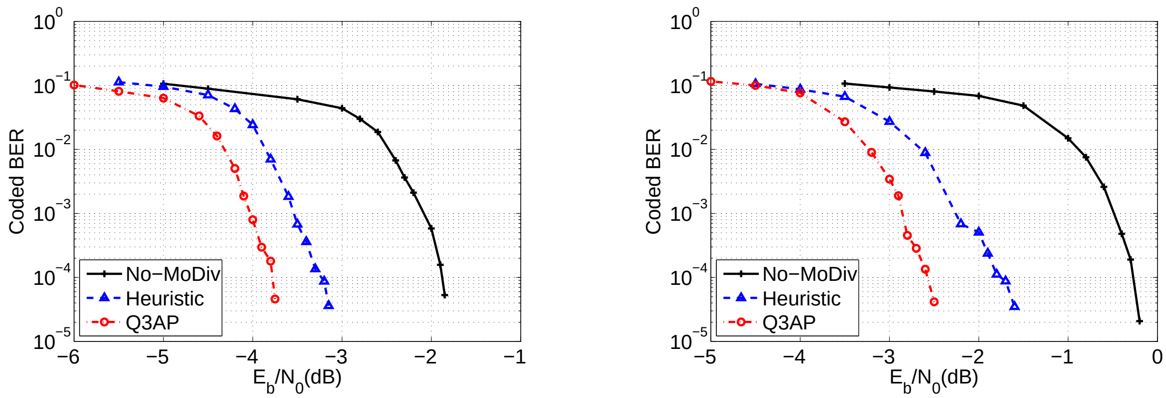
(a) $K = 10$ (b) $K = 5$

Fig. 3. Monte-Carlo simulated BER of (1) Q3AP optimized MoDiv (2) $\psi_1 = \psi_2 = \psi_S$ (3) $\psi_1 = \psi_2 = \psi_0$ for $K = 5, 10$, $a = 1, \sqrt{2}, 2$.

transmission can further reduce BER to boost performance. However, as K decreases, heuristic MoDiv becomes a good approximate solution with diminishing performance gap from Q3AP-optimized MoDiv.

Finally, we test the coded BER performance and the robustness of the Q3AP MoDiv design with a typical LDPC coded system based on [17]. We use a LDPC code with length $L = 2400$, coding rate of 0.75 and a Monte-Carlo run of up to 2000 LDPC frames. Although our Q3AP MoDiv scheme depends on noise power and statistical CSI, in this simulation we only use Q3AP MoDiv schemes designed for the condition of $E_b/N_0 = -2dB$ under AWGN channel. We test the MoDiv design on Rician channels of $K = 10$, $a = 1$, and various E_b/N_0 to demonstrate its robustness. The coded BER results for 16-QAM and 32-QAM are plotted in Fig. 4. We can see that the Q3AP MoDiv solution provides additional gain over the heuristic MoDiv and no MoDiv even for mismatched design conditions. Also we notice that the Q3AP MoDiv provides larger performance gain for 32-QAM than for 16-QAM constellation.

The robustness is important in scalable extensions of the present simple network model. For multiple cooperative relays, our natural extension is to allow different relay nodes to randomly select between the two Q3AP-optimized mappings ψ_1 and ψ_2 to ensure sufficient MoDiv and channel diversity.



(a) 16-QAM (b) 32-QAM

Fig. 4. Monte-Carlo simulated coded BER.

V. Conclusion

In this work, we investigated the modulation diversity (MoDiv) design problem in a three-node cooperative relay-HARQ system driven by coordinated retransmission from both the source and the relay. Aiming to minimize the bit error rate (BER) upper bound, we formulated the MoDiv design into a quadratic three-dimensional assignment problem (Q3AP), and presented an efficient modified iterative local search (ILS) solution. Our numerical tests demonstrate the performance advantage and robustness of the Q3AP-based MoDiv design over simply repeating the use of Gray mapping. When the channel experiences deep fades, a heuristic MoDiv design can also achieve comparable performance gains at low complexity.

References

[1] 3GPP TS36.331, "E-UTRA Radio Resource Control (RRC); Protocol specification (Release 12)," March 2015, v12.5.0.

[2] 3GPP TS36.213, "E-UTRA Physical layer procedures (Release 12)," March 2015, v12.5.0.

[3] H. A. Ngo and L. Hanzo, "Hybrid Automatic-Repeat-reQuest systems for cooperative wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 25–45, First Quarter 2014.

[4] G. Benelli, "A new method for the integration of modulation and channel coding in an ARQ protocol," *IEEE Trans. Commun.*, vol. 40, no. 10, pp. 1594–1606, Oct 1992.

[5] H. Samra, Z. Ding, and P. Hahn, "Symbol mapping diversity design for multiple packet transmissions," *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 810–817, May 2005.

[6] K. Seddik, A. Ibrahim, and K. Liu, "Trans-modulation in wireless relay networks," *IEEE Commun. Lett.*, vol. 12, no. 3, pp. 170–172, Mar 2008.

[7] M. Khormuji and E. Larsson, “Rate-optimized constellation rearrangement for the relay channel,” *IEEE Commun. Lett.*, vol. 12, no. 9, pp. 618–620, Sep 2008.

[8] J. W. Kim, H. Lee, J. Ahn, and C. Kang, “Design of signal constellation rearrangement (CoRe) for multiple relay links,” in *Proc. IEEE GLOBECOM*, Nov 2009, pp. 1–6.

[9] H.-S. Ryu, J.-S. Lee, and C. Kang, “BER analysis of constellation rearrangement for cooperative relaying networks over Nakagami-m fading channel,” in *Proc. IEEE Int. Commun. Conf. (ICC)*, Jun 2011, pp. 1–5.

[10] T. Cover and A. Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep 1979.

[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, Aug 2004.

[12] U. Benlic and J.-K. Hao, “Memetic search for the quadratic assignment problem,” *Expert Systems with Applications*, vol. 42, no. 1, pp. 584–595, 2015.

[13] G. Taricco and E. Biglieri, “Exact pairwise error probability of space-time codes,” *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 510–513, Feb 2002.

[14] P. M. Hahn, B.-J. Kim, T. Stützle, S. Kanthak, W. L. Hightower, H. Samra, Z. Ding, and M. Guignard, “The quadratic three-dimensional assignment problem: Exact and approximate solution methods,” *European Journal of Operational Research*, vol. 184, no. 2, pp. 416 – 428, 2008.

[15] H. D. Mittelmann and D. Salvagnin, “On solving a hard quadratic 3-dimensional assignment problem,” *Mathematical Programming Computation*, 2015, (online).

[16] T. Stützle, “Iterated local search for the quadratic assignment problem,” *European Journal of Operational Research*, vol. 174, no. 3, pp. 1519 – 1539, 2006.

[17] B. Hochwald and S. ten Brink, “Achieving near-capacity on a multiple-antenna channel,” *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, March 2003.