Modulation Diversity Design in Cooperative Relay-HARQ Network

Wenhao Wu, Hans Mittelmann, and Zhi Ding

***Abstract*—Modulation diversity (MoDiv) is a practically useful diversity enhancement technique that utilizes different modulation mappings from bits to constellation symbols. This approach 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 the bit error rate (BER). We formulate the MoDiv design optimization problem into a quadratic three-dimensional assignment problem (Q3AP) solved with a modified iterated local search (ILS) method. Numerical results demonstrate its performance gain, and show that when the channels are heavily fading, a comparable BER reduction is achieved with a heuristic MoDiv scheme.**

***Index Terms*—Modulation diversity, relay, HARQ, Q3AP.**

I. Introduction

In modern wireless communication, high rate data transmission often leads to reception errors. To overcome packet loss due to transmission errors, Automatic Repeat reQuest (ARQ) or Hybrid ARQ (HARQ) are important mechanism for better reliability. HARQ in conjunction with relay networks has attracted great deal of research interest in recent years [1]. 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) [2], in which each group of $\log_2 Q$ bits are mapped to different constellation points across different links and in different transmissions.

There is a number of research works on MoDiv for relay network [3], [4] and relay-HARQ system [5], [6]. Despite the encouraging performance gain, the afore-mentioned works all assume a low bandwidth efficiency scheme where (re)transmissions on the Source-Destination (S-D) link and the Relay-Destination (R-D) links occupy orthogonal bands or time slots. From this point of view, we study the MoDiv design in a relay-HARQ network in Fig. 1 where the S-D and R-D links are additive as in [7], [8], practically forming a 2-by-1 multiple-input single-output (MISO) diversity system. In this context, we formulate the bit error rate (BER) minimization MoDiv design into a quadratic three-dimensional assignment problem (Q3AP). Though Q3AP is NP-hard so that the MoDiv design problem can not be solved exactly in practice even for a middle size constellation such as 16-QAM, there are a number of heuristic algorithms [9] for its reduced form, Quadratic Assignment Problem (QAP), that results in high-quality solutions over the QAPLIB dataset [10]. In this work, we adopt an efficient modified Iterated Local Search (ILS) method to solve the Q3AP formulated from BER-minimization MoDiv design problem. Our numerical results demonstrate significant BER reduction over non-MoDiv HARQ schemes. Moreover, we also show that when the channels are heavily fading, a comparable performance gain is offered by an effortless heuristic MoDiv scheme.

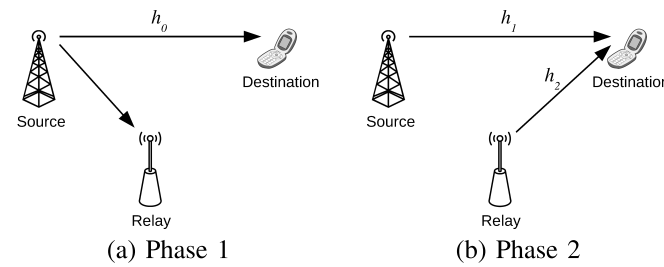


Fig. 1. Cooperative relay-HARQ networks.

II. System Model

We consider the cooperative relay-HARQ network shown in Fig. 1. There are two phases in this transmission system. In phase 1, source node broadcasts its packet that can be received by both the destination and the relay. In phase 2 of this HARQ setup, upon packet loss, both the relay and the source node cooperatively retransmit the lost packet information to the destination. Our goal is to design the H-ARQ constellation mapping for modulation diversity such that the packet BER at the receiver is minimized.

Denote $\mathcal{C}$ as the 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 is indexed by its decimal equivalence $p \in \{0, \ldots, Q-1\}$. The source transmits $\psi_0[p]$ to the destination via channel h_0 which is also overheard by the decode-and-forward (DF) relay. We assume that this relay is placed strategically such that it has negligible decoding error rate as in [6], [5]. 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, $\psi_1 \neq \psi_0$ 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 the channels h_0, h_1 and h_2 follow independent Rician distribution.

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

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

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 the 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) [11]

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

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

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

In other words, given indices p, q and the remapping scheme ψ_1, ψ_2, the probability of the random variable $\delta < 0$ being evaluated over the random variables h_0, h_1, h_2, v_1, v_2, and δ is defined as

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

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

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

and $\mathcal{I}$ and $\mathcal{J}$ which are defined as in (6) by replacing the summation index p with i and j, respectively. Then from (3)(4)(5), the BER minimization MoDiv scheme $\min_{\psi_1,\psi_2} P_{\mathrm{BER}}$ can be reformulated as

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

in which

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

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

B. Computation of the Pair-wise Symbol Error Rate

In this section we focus on the computation of the parameters $\{c_{pijqkl}\}$ of the Q3AP problem. According to (8), the key to compute 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 $\delta(p,i,j,q,k,l)$ as in (5). Define the moment generating function (MGF)

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

Under the well known Rician channel model, $h_m \sim \mathcal{CN}(\mu_{h_m}, \sigma_{h_m}^2), m = 0,1,2$, we can extend the method in [11], [12] to compute $P_{h_0,h_1,h_2,v_1,v_2}\{\delta(p,i,j,q,k,l) < 0\}$:

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

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 [12]. 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 $\Phi_\delta(\omega)$, denote the 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

$$\boldsymbol{\mu}_1 = [\mu_{h_0}, 0]^T, \boldsymbol{\mu}_2 = [\mu_{h_1}, \mu_{h_2}, 0]^T, \quad (11)$$
$$\boldsymbol{\Sigma}_1 = \mathrm{diag}\left(\sigma_{h_0}^2, \sigma_v^2\right), \boldsymbol{\Sigma}_2 = \mathrm{diag}\left(\sigma_{h_1}^2, \sigma_{h_2}^2, \sigma_v^2\right). \quad (12)$$

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

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

here $e_{ab} = \psi_0[a] - \psi_0[b]$. Then 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 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 too high. We will address in future works to reduce this complexity by adding a few rules to restrict the remapping schemes.

C. Q3AP Solution

For practical-sized constellation such as 16-QAM and 32-QAM, it is impractical to apply the exact branch-and-bound algorithm [13]. Also, our tests show that they do not have enough symmetry to exploit for faster solution as does the 16-PSK constellation [14]. Consequently, the MoDiv problem is solved with the ILS method [13] extended from its QAP version [15]. Starting from two random initial mappings $\psi_1^{(0)}, \psi_2^{(0)}$, the algorithm executes a local search by attempting to exchange the mapping of exactly 2 indices in order to lower the objective function and whenever a reduction in the objective function is made the mapping is updated. When it 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.

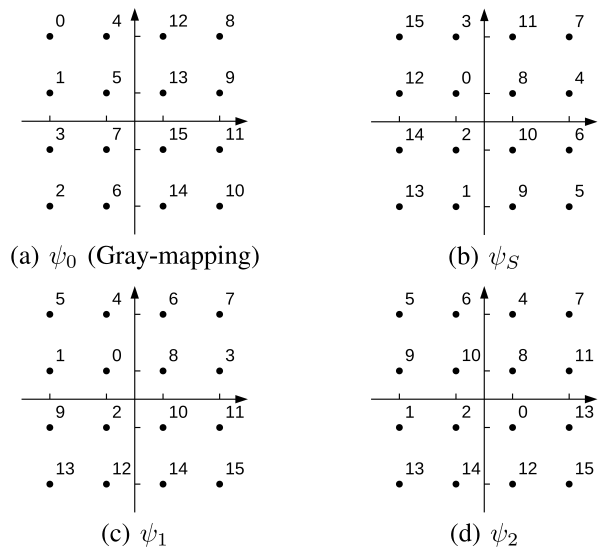
D. A Heuristic MoDiv Scheme

The MoDiv design problem can be simplified if we force $\psi_1 = \psi_2$. Also, from (5), a simple heuristic is 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. A simple remapping scheme based on this heuristic is given by Seddik in [3]. We will show in the next section that when the channels are heavily fading, this heuristic remapping offers comparable performance gain as our Q3AP-based MoDiv.

IV. Numerical Results

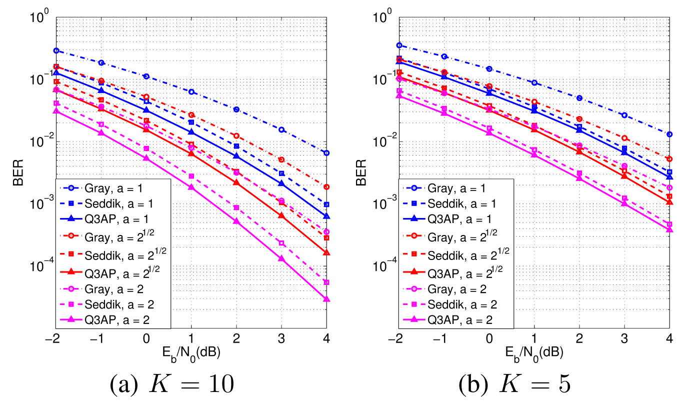
In our simulation, all Rician channels are assumed to have the same Rician parameter K. Also we assume that during the second phase, the phases of the line of sight (LOS) components of channels h_1 and h_2 can be aligned at the source and relay, respectively. Consequently, we define $\mu_{h_0} = \mu_{h_1} = \sqrt{K/(K+1)}$, $\mu_{h_2} = a\sqrt{K/(K+1)}$, $\sigma_{h_0}^2 = \sigma_{h_1}^2 = 1/(K+1)$ and $\sigma_{h_2}^2 = |a|^2/(K+1)$, where a represents the ratio between the amplitude of the LOS component of the relay-to-destination and the S-D link. Throughout the simulation we consider 16-QAM constellation, i.e., $Q = 16$. The noise power is parameterized with E_b/N_0 of the S-D link.

First, we provide an example of Q3AP optimized MoDiv. When $E_b/N_0 = 2$dB, $K = 10$ and $a = 1$, the remapping scheme of ψ_0 (using Gray-mapping), ψ_1 and ψ_2 is depicted in Fig. 2. For comparison we also plot Seddik's heuristic remapping ψ_S. The results justify the heuristics mentioned in Section III-D and both ψ_1 and ψ_2 are essentially very close to ψ_S. However, we note that in general $\psi_1 \neq \psi_2$.



(a) ψ_0 (Gray-mapping) (b) ψ_S (c) ψ_1 (d) ψ_2

Fig. 2. Q3AP optimized MoDiv schemes.



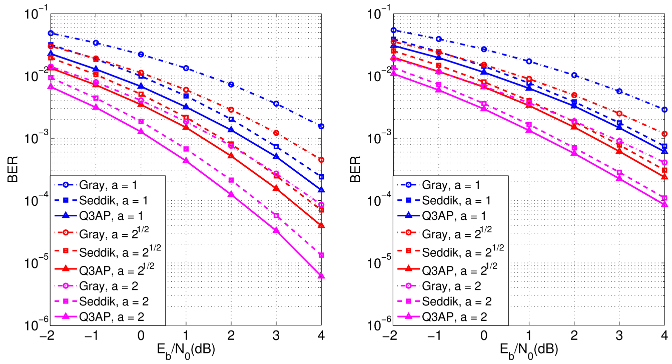
(a) $K = 10$ (b) $K = 5$

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

Next we compare the BER upper bound P_{BER}, which is also the Q3AP objective function, for three mapping schemes: the Q3AP optimized MoDiv, the heuristic MoDiv $\psi_1 = \psi_2 = \psi_S$, and no MoDiv at all, i.e. $\psi_1 = \psi_2 = \psi_0$. 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 transmission can further boost the BER performance. However, as K decreases, heuristic MoDiv becomes a good approximation as the its performance gap with the Q3AP optimized MoDiv diminishes.

The performance gain of MoDiv for the cooperative relay-HARQ system is further verified with the actual BER evaluated with Monte-Carlo simulation in Fig. 4. We implement the ML demodulator in (2) and evaluate the average BER of $M = 10^7$ randomly generated indices p. A similar trend as in Fig. 3 is demonstrated.

Finally, we test the robustness of the Q3AP MoDiv solution



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

Fig. 4. Monte-Carlo simulated BER of (1) Q3AP optimized MoDiv (2) $\psi_1 = \psi_2 = \psi_S$ (3) $\psi_1 = \psi_2 = \psi_0$ for $K = 5, 10$, $a = 1, \sqrt{2}, 2$.

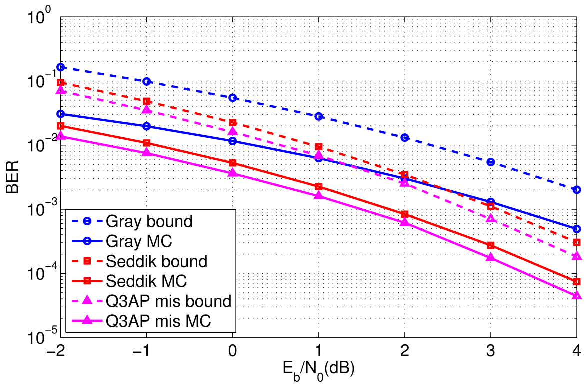


Fig. 5. BER performance under channel mismatch. Q3AP MoDiv solution for CSI $a = 1$, $K = 5$ while the actual CSI is $a = \sqrt{2}\exp(j\pi/12)$, $K = 10$.

against mismatch in channel information. The intuition is that in practical systems, the CSI may not be perfectly accurate and the phase of the LOS component of the S-D and R-D links may not be perfectly aligned. Also, the Q3AP MoDiv scheme can only be computed off-line for a limited number of channel settings a priori, instead of for all possible combinations. We deliberately test the BER performance of the Q3AP MoDiv solution for channel parameters $a = 1$, $K = 5$ while the actual channel setting is $a = \sqrt{2}\exp(j\pi/12)$, $K = 10$. The BER upper bound and Monte-Carlo simulated BER are plotted in Fig. 5. Apparently, the Q3AP MoDiv solution is not sensitive to the relative amplitude and the phase alignment error between S-D and R-D links.

V. Conclusion

In this work, we investigated the modulation diversity (MoDiv) design problem in a three-node cooperative relay-HARQ systems featured by the 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 have demonstrated the performance advantage of the Q3AP-based MoDiv design over simply repeating the use of Gray mapping, and when the channel is heavily fading, a comparable performance gain can be achieved with a heuristic MoDiv design.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

[10] R. E. Burkard, S. E. Karisch, and F. Rendl, "QAPLIB–a quadratic assignment problem library," *Journal of Global optimization*, vol. 10, no. 4, pp. 391–403, 1997.

[11] 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.

[12] 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.

[13] 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.

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

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