Fading-resistant Modulation Using Several Transmitter Antennas*

Victor M. DaSilva and Elvino S. Sousa
Department of Electrical and Computer Engineering
University of Toronto
email: sousa@comm.utoronto.ca
(416)978-3651

Abstract

This paper proposes a bandwidth-efficient fading-resistant transmission scheme which implements transmitter diversity using L antennas at the base station. When the antennas are spaced sufficiently far apart, the transmission from each antenna undergoes a different degree of fading. These transmissions are coordinated to mitigate the effects of Rayleigh fading and the mobile receiver can recover the entire L-dimensional transmitted vector signal as long as the signal energy of at least one coordinate is large enough. L-dimensional fading-resistant signal constellations are generated by maximizing a figure of merit for the Rayleigh fading channel. This scheme offers a significant performance improvement over a conventional single-antenna BPSK scheme when coding is ineffective due to slow fading.

*This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Canadian Space Agency, and the Communications Research Centre in the form of scholarships, and in part by the Information Technology Research Center of Ontario (ITRC). The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Canada.

1 Introduction

Diversity schemes are normally considered when fading severely affects the transmission quality of a communication link. When a receiver has several antennas, space, polarization, or angle diversity can be achieved by combining the different signals from the antennas. But these receiver diversity methods have drawbacks. For example, space diversity requires that the antennas be spaced sufficiently far apart in order for the signals at the different antennas to be uncorrelated, which is difficult to achieve when the receiver is a small hand-held unit. Simple transmitter diversity schemes such as time and frequency diversity techniques normally involve retransmission in different time slots or frequency bands. These methods of diversity involve bandwidth expansion in order that the retransmitted signals do not interfere with one another. An example of a bandwidth-efficient method of transmitter diversity is selection diversity, where a transmitter can transmit using one of several antennas. If both uplink and downlink carrier frequencies are the same, as in a time division duplex (TDD) system, the base station can transmit using the antenna which had the largest signal level during the last transmission from the mobile [1] or it may predict which antenna will have the strongest signal level [2]. In this paper we assume that the uplink and downlink frequencies are different. Hence selection diversity based on fading on the reverse link will perform poorly. The transmission of extra bits on the reverse channel to control antenna switching adds complexity to the system and also has an inherent delay amounting to several symbols.

Another method of achieving diversity is to use coding. Codes with large Hamming distance are preferable because they offer larger coding gain over uncoded BPSK (binary phase shift keying). However, binary convolutional or block codes reduce bandwidth efficiency. Trellis-coded modulation can achieve coding gain and have the same bandwidth efficiency as uncoded BPSK because the number of modulation levels is increased to compensate for the reduction in information bit rate due to coding [3]-[7]. Both of these approaches assume that the fading process varies with time sufficiently fast enough for an interleaver to spread the symbols of a codeword over time so that the fading amplitudes of the symbols are uncorrelated. Thus, performance gain is achieved through time diversity. However, the interleaver has a finite interleaving depth so that the delay due to interleaving is not too large. If the mobile receiver is moving slowly, the fading amplitude will vary too slowly over time for the interleaver to be effective and the performance will be poor.

In this paper we propose a bandwidth-efficient fading-resistant modulation scheme in which the base station transmitter has L antennas that transmit simultaneously. These antennas are spaced sufficiently far apart so that the fading amplitudes of the different links are uncorrelated. Each component of an L-dimensional signal point is transmitted on one antenna and the receiver makes a decision based on the entire L-dimensional received vector. The advantage of this scheme over one which uses coding is that L^{th}-order diversity performance is obtained even when the mobile receiver is moving very slowly. This fading-resistant technique was proposed in [8].

The proposed scheme is effective for any link which has fading, although it requires transmitter diversity which may not be convenient to implement at the mobile. Hence it is mostly applicable to the forward link of a cellular system. In the analysis we assume that fading is sufficiently slow so that the fading amplitudes may be estimated in order to perform optimum detection. This is the case in mobile cellular communications where the fading amplitudes are typically constant over a few symbols.

Fading-resistant constellations are generated by optimizing a distance measure for the Rayleigh fading channel. These L-dimensional constellations are compared to constellations proposed in [9] for L-link channels with different noise powers in each of the links. Fading-resistant lattices have also been proposed in [10], but constellations drawn from these lattices have worse performance than the constellations considered in this paper when the channel has no fading, that is, when the channel is the additive white Gaussian noise (AWGN) channel.

The outline of the paper is as follows. First, the channel model and the receiver used are described. The baseline scheme, which has the same performance as uncoded BPSK, and a simple fading-resistant scheme, which uses BPSK and is very bandwidth costly, are described. The performance of bandwidth-efficient fading-resistant schemes is discussed in terms of a figure of merit for the Rayleigh fading channel. This figure of merit is maximized by rotating the baseline constellations in L dimensions and the performance of these optimally rotated constellations is compared with that of other fading-resistant constellations.

2 Channel Model

The forward link channel model is depicted in Fig. 1. The base station has L transmitter antennas, the mobile receiver has a single antenna, and each of the L links has a different fading amplitude. The received signal is

$$r(t) = \sum_{i=1}^{L} \alpha_i m_i s_i(t) \cos(w_c t + \theta_i) + n(t), \quad 0 \leq t \leq T. \tag{1}$$

The signal from the transmitter's i^{th} antenna is a pulse amplitude modulated (PAM) signal and m_i is the signal amplitude, $s_i(t)$ is the pulse shape, α_i is the fading amplitude of link i, and $n(t)$ is a white Gaussian noise process with power spectral density $N_0/2$. It is assumed that the fading amplitude for a given link is constant over the signalling interval $[0, T]$ and that the receiver uses coherent detection. The signals $s_i(t), s_j(t), i \neq j$, are assumed to be orthogonal and all of the energy of $s_i(t), 1 \leq i \leq L$, is contained in $[0, T]$.

The optimum receiver[1] consists of a bank of L correlators, as shown in Fig. 2. The output of the i^{th} correlator is

$$\begin{aligned} y_i &= 2\int_0^T r(t) s_i(t) \cos(w_c t + \theta_i) dt \\ &= \alpha_i m_i E_s + \eta_i, \quad 1 \leq i \leq L \end{aligned} \tag{2}$$

where $E_s = \int_0^T s_i^2(t) dt$ is the baseband pulse energy and $\eta_i = 2\int_0^T n(t) s_i(t) \cos(w_c t + \theta_i) dt$. η_i, $1 \leq i \leq L$, are uncorrelated zero mean Gaussian random variables with variance $N_0 E_s$. The received vector $y = (y_1, y_2, \ldots, y_L)$ is fed into a decision device which estimates the transmitted vector $m = (m_1, m_2, \ldots, m_L)$. It is assumed that the receiver can estimate the fading amplitudes α_i, $1 \leq i \leq L$. The receiver finds the L-dimensional constellation point in C, with coordinates suitably amplified, that has the closest Euclidean distance to the received vector y. That is, the receiver picks the symbol $\hat{m} = (\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L) \in C$ that minimizes $\sum_{i=1}^{L} (y_i/E_s - \alpha_i \hat{m}_i)^2$. A symbol detection error occurs when $\hat{m} \neq m$.

The fading amplitudes α_i are modelled as independent and identically distributed Rayleigh

[1]We assume perfect estimation of the fading amplitudes.

random variables with the common probability density function

$$f_\alpha(\alpha) = 2\alpha e^{-\alpha^2}, \quad \alpha \geq 0. \tag{3}$$

The assumption of independent fading is valid if the transmitter antennas are spaced sufficiently far apart, which is relatively easy to do when the transmitter is the base station.

The fading-resistant transmission schemes discussed in this paper assume that the receiver is capable of estimating the fading amplitude of each link. This is possible if the fading amplitudes vary slowly over time. If the fading amplitudes vary quickly over time, the performance of the receiver will degrade due to estimation errors.

3 Baseline scheme: independent BPSK signals

The baseline scheme consists of a transmitter with L antennas which sends either a $+1$ or -1 bit on each antenna, independently of the rest, and the output of the i^{th} correlator is given by (2) with $m_i \in \{1, -1\}$. This corresponds to sending independent BPSK signals on each antenna. There is an L-fold expansion in bandwidth over the $L = 1$ case in order to have L orthogonal transmissions but the overall data rate also increases by a factor of L so that there is no bandwidth penalty.

For optimal detection, the correlator output y_i is fed into a threshold device which outputs a 1 if the input is positive, and a -1 otherwise. The probability of bit error is (e.g. see [11, p.717])

$$\mathrm{P(error)} = \frac{1}{2}\left(1 - \sqrt{\frac{E_b/N_0}{E_b/N_0 + 1}}\right) \tag{4}$$

where $E_b = \overline{\alpha^2} E_s/2 = E_s/2$ is the average received bit energy. Eq. (4) applies to each of the L links, and so the overall bit error rate is also given by (4). For this scheme the overall bit error rate is independent of L and there is no advantage over single-antenna BPSK.

4 Identical BPSK signals

If the transmitter sends the same bit stream on each of the L antennas, and if the receiver can estimate the fading amplitudes α_i, maximal ratio combining can be used by the receiver.

This scheme is resistant to fading because all of the L signals have to fade severely before a bit is lost. The drawback of this scheme is that an L-fold increase in bandwidth is needed to make the L signals orthogonal, while the data rate remains the same as in the $L = 1$ case.

The output of the i^{th} correlator is given by (2) where $m_i \in \{1, -1\}$ and $m_1 = m_2 = \ldots = m_L$. In order to compare this scheme with the baseline scheme on the basis of the same energy per bit, we have $E_s = 2E_b/L$. Although the decision device in Fig. 2 is optimal, an equivalent but simpler implementation is to compute the decision variable $\sum_{i=1}^{L} y_i \alpha_i$ and feed the result into a threshold device. If it is positive, the threshold device outputs a 1, otherwise it outputs a –1. The probability of bit error is (e.g. see [11, p.723])

$$\mathrm{P(error)} = \left(\frac{1-\mu}{2}\right)^L \sum_{k=0}^{L-1} \binom{L-1+k}{k} \left(\frac{1+\mu}{2}\right)^k, \quad L \geq 2 \tag{5}$$

where $\mu = \sqrt{\frac{(E_b/N_0)/L}{1+(E_b/N_0)/L}}$.

5 Upper bound on performance in a Rayleigh fading channel

Consider an L-dimensional constellation C with $|C|$ points. Assuming that all points in C are transmitted with equal probability, an upper bound on the average probability of symbol detection error can be obtained from the union bound,

$$P(\text{error}) \leq \frac{1}{|C|} \sum_{m} \sum_{\substack{\hat{m} \\ \hat{m} \neq m}} P(m \rightarrow \hat{m}) \tag{6}$$

where $P(m \rightarrow \hat{m})$ is the probability that the received symbol is closer to the L-dimensional symbol $\hat{m} = (\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L)$ than to $m = (m_1, m_2, \ldots, m_L)$ given that m was transmitted. If the receiver estimates the fading amplitudes in (2) perfectly, and if a minimum-distance decoding rule is used by the receiver, then the pairwise error probability, conditioned on the fading amplitude vector $(\alpha_1, \alpha_2, \ldots, \alpha_L)$, can be upper-bounded using the Chernoff [12] bound as follows,

$$P(m \to \hat{m}|\alpha_1, \ldots, \alpha_L) \leq \prod_{i=1}^{L} \exp\{-\alpha_i^2 (m_i - \hat{m}_i)^2 E_s/(8N_0)\}. \tag{7}$$

Assuming the fading amplitudes are independent, (7) is averaged over the probability density function for α_i to obtain (for example, see [13])

$$P(m \to \hat{m}) \leq \prod_{i=1}^{L} \frac{1}{1 + (m_i - \hat{m}_i)^2 E_s/(8N_0)} \leq \prod_{\substack{i=1 \\ m_i \neq \hat{m}_i}}^{L} \frac{8N_0}{E_s(m_i - \hat{m}_i)^2}. \tag{8}$$

At sufficiently high SNR, (6) is dominated by the largest $P(m \to \hat{m})$ term, and assuming $m_i \neq \hat{m}_i, 1 \leq i \leq L$, then $P(\text{error}) \propto 1/(E_s/N_0)^L$. We refer to constellations which have a probability of error that varies inversely proportional to the L^{th} power of the signal-to-noise ratio as fading-resistant constellations.

From (8), the following quantity, which we call the constellation figure of merit for the Rayleigh fading channel, gives an indication of the performance of a signal constellation at high SNR,

$$\mathrm{CFM}_{\mathrm{Rayleigh}}(C) = \min_{\substack{m, \hat{m} \in C \\ m \neq \hat{m}}} \prod_{\substack{i=1 \\ m_i \neq \hat{m}_i}}^{L} (m_i - \hat{m}_i)^2. \tag{9}$$

In the next section we describe a method for constructing L-dimensional fading-resistant constellations which have a large $\mathrm{CFM}_{\mathrm{Rayleigh}}$. We are interested only in constellations where $m_i \neq \hat{m}_i, 1 \leq i \leq L$.

In [3] trellis-coded modulation codes for the Rayleigh fading channel are designed so as to maximize the minimum Hamming distance between transmitted sequences, l_m. A secondary criterion is to maximize the minimum squared product distance $d_p^2 = \min_{m, \hat{m} \in C} \prod_{\substack{i=1 \\ m_i \neq \hat{m}_i}}^{l_m} (m_i - \hat{m}_i)^2$. In the modulation scheme we are considering, L-dimensional symbols can be viewed as being equivalent to trellis-coded modulation sequences of length L. The Hamming distance l_m in this case is made to equal L and the design problem is to maximize d_p^2, and thus $\mathrm{CFM}_{\mathrm{Rayleigh}}$.

6 Construction of fading-resistant constellations

Given an L-dimensional constellation of points, we would like to apply a transformation to the constellation which preserves the Euclidean distances between points but improves the constellation's resistance to fading. We impose the restriction that the transformation preserve Euclidean distances and norms because we do not want to degrade the performance of the constellation in the AWGN channel. Such transformations are called isometries.

We represent our original constellation as a $|C| \times L$ matrix C, where each row of the matrix corresponds to a point in the L-dimensional constellation. One possible distance-preserving transformation is to multiply this matrix by an orthogonal $L \times L$ matrix A. The optimal matrix A maximizes the fading-resistance of the transformed constellation CA, that is, it maximizes $\mathrm{CFM}_{\mathrm{Rayleigh}}(CA)$.

In Appendix I we show how an $L \times L$ orthogonal matrix A can be written as the product of $\binom{L}{2}$ rotation matrices and a reflection matrix. From (23), we see that multiplication of the constellation matrix C by an arbitrary orthogonal matrix A has the following geometrical interpretation. The constellation is rotated with respect to the (i,j)-plane by an amount θ_{ij}, $1 \leq i \leq L-1,\ i+1 \leq j \leq L$, and there are $\binom{L}{2}$ such rotations. Then the constellation is reflected in the i^{th} axis, where the matrix $\hat{I}$ has (i,i) entry equal to -1, and the number of such reflections is equal to the number of -1 elements on the main diagonal of $\hat{I}$. Writing $A = Q\hat{I}$, where Q is the product of $\binom{L}{2}$ rotation matrices in (23), we see that

$$\mathrm{CFM}_{\mathrm{Rayleigh}}(CA) = \mathrm{CFM}_{\mathrm{Rayleigh}}(CQ\hat{I}) = \mathrm{CFM}_{\mathrm{Rayleigh}}(CQ) \tag{10}$$

where the second identity in (10) follows because the matrix $CQ\hat{I}$ is just the matrix CQ with several of its columns negated, and negating the columns of a constellation matrix does not affect the $\mathrm{CFM}_{\mathrm{Rayleigh}}$ of the constellation. Rather than look for an optimal constellation CA it is sufficient to look for an optimal constellation CQ.

Our method of obtaining an optimal constellation $C_{\mathrm{opt}} = CQ$, that is, one with maximum $\mathrm{CFM}_{\mathrm{Rayleigh}}(CQ)$, given a starting constellation C, is to vary $\binom{L}{2}$ rotation angles according to a numerical optimization or search algorithm. The constellation C is rotated with respect to the (i,j) plane, $1 \leq i \leq L-1,\ i+1 \leq j \leq L$. Note that Q does not need to be computed explicitly. There are $\binom{L}{2}$ degrees of freedom and for $L = 2, 3$ and for a large angle discretization interval (for example, 1°) the search for the optimal angles can be made exhaustively but for $L \geq 4$ and for reasonably small discretization intervals (for example, less than 5°) the exhaustive search method takes too long. For these larger values of L, we pick at random many different starting rotation vectors and use a gradient descent method each time to converge to a (possibly local) maximum; then a possibly sub-optimal rotation vector is obtained corresponding to the maximum of all these trials.

In general, we can start with constellations drawn from L-dimensional packings. For a given number of constellation points, points are drawn from the minimum-energy shells of these packings. These constellations are then rotated in L-space so as to maximize $\mathrm{CFM}_{\mathrm{Rayleigh}}$. For example, in two dimensions two possible packings are $\mathbb{Z}^2$ and the hexagonal lattice. In three dimensions, possible packings are the hexagonal close packing, tetrahedral packing, face centered cubic lattice, body centered cubic lattice, and the $\mathbb{Z}^3$ lattice . These lattices are defined, for example, in [14]. Denser packings are preferred because large constellations drawn from these packings have better performance in the additive white Gaussian noise channel.

For the same bit rate and AWGN channel performance as the baseline scheme, we start off with the L-cube, which is the constellation for the baseline scheme. (This also corresponds to a single energy shell of the $\mathbb{Z}^L$ lattice.) For example, when L=2, the baseline constellation matrix is

$$C = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}. \tag{11}$$

In this case maximization of $\mathrm{CFM}_{\mathrm{Rayleigh}}$ is done by varying only one rotation angle. The optimal angle of rotation for this constellation can be found using an exhaustive search to be $\theta_{\mathrm{opt}} = 31.7^\circ$, assuming a discretization interval of 0.1°. This optimal angle can be derived analytically as $\theta_{\mathrm{opt}} = \frac{1}{2}\tan^{-1}(2)$ but the derivation is omitted for the sake of brevity. The optimally rotated constellation is

$$C_{\text{opt}} = \begin{bmatrix} -0.325 & -1.376 \\ -1.376 & 0.325 \\ 1.376 & -0.325 \\ 0.325 & 1.376 \end{bmatrix}. \tag{12}$$

These two constellations are shown in Fig. 3. Each row in the constellation matrix corresponds to a point (m_1, m_2). The Kerpez constellation shown in Fig. 3 will be discussed in the next section.

When $L = 3$ the baseline constellation C consists of the vertices of a three-dimensional cube and optimization is done over three rotation angles. The optimal rotation vector can be found using an exhaustive search to be $\theta_{\text{opt}} = [\theta_{12}, \theta_{13}, \theta_{23}] = [24^o, 36^o, 66^o]$, using a discretization interval of 1^o. Both the baseline constellation and the optimally rotated constellation are shown below,

$$C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ -1 & -1 & -1 \end{bmatrix}, \quad C_{\text{opt}} = \begin{bmatrix} 0.177 & 0.474 & -1.656 \\ -0.997 & -1.003 & -0.998 \\ -0.480 & 1.654 & -0.181 \\ -1.655 & 0.176 & 0.476 \\ 1.655 & -0.176 & -0.476 \\ 0.480 & -1.654 & 0.181 \\ 0.997 & 1.003 & 0.998 \\ -0.177 & -0.474 & 1.656 \end{bmatrix}. \tag{13}$$

These constellations are shown in Fig. 4.

For $L \geq 4$, an exhaustive search over the $\binom{L}{2}$ rotation angles in order to maximize $\text{CFM}_{\text{Rayleigh}}$ for the L-cube proved to be too time consuming. For these constellations, many rotation vectors were picked at random, and a gradient based approach was used to vary the rotation angles so as to converge to a local maximum. For L=4 there are six degrees of freedom and the optimal rotation vector found was $\theta_{\text{opt}} = [\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}] = [206^o, 15^o, 306^o, 42^o, 213^o, 31^o]$. For L=5 there are ten degrees of freedom and the optimal rotation vector found was $\theta_{\text{opt}} = [\theta_{12}, \theta_{13}, \theta_{14}, \theta_{15}, \theta_{23}, \theta_{24}, \theta_{25}, \theta_{34}, \theta_{35}, \theta_{45}] = [294^o, 349^o, 18^o,$

$340^o, 103^o, 184^o, 114^o, 275^o, 212^o, 25^o]$.

7 Kerpez constellations

In [9], L-dimensional constellations were proposed for L-link nonfading channels where the links have different noise powers. These constellations were derived heuristically and we note that they should also perform well in the case we are considering where the noise powers are equal but the fading amplitudes are different.

For these constellations, there are M^L constellation points $(m_1, m_2, \ldots, m_L)$. Specification of $i_1 \in \{1, 2, \ldots, M^L\}$ uniquely identifies the L-dimensional symbol, which has j^{th} coordinate given by

$$m_j = -2i_j + M^L + 1, \quad i_j \in \{1, 2, \ldots, M^L\} \tag{14}$$

where $i_j, 2 \leq j \leq L$, is related to i_1 by

$$i_j = [(i_1 - 1)M^{L-j+1}] \bmod M^L + M^{L-j+1} + 1 - \left\lceil \frac{i_1}{M^{j-1}} \right\rceil . \tag{15}$$

We consider only the case $M = 2$ so as to have the same bit rate as the baseline scheme. For example, when $L = 2$, the Kerpez constellation is

$$C = \begin{bmatrix} 3 & 1 \\ -1 & 3 \\ -3 & -1 \\ 1 & -3 \end{bmatrix} . \tag{16}$$

This constellation, normalized by $\sqrt{5}$ so that the symbol energy is the same as that of the baseline two-dimensional constellation, is shown in Fig. 3. As another example, when $L = 3$, the set of constellation points is shown in the first column in Table 1. Note that, except for the $L = 2$ case, the Kerpez constellations are not rotated versions of the baseline L-cube constellations. In fact, the points in a Kerpez constellation with $L > 2$ do not all have the same energy.

From [9], the average symbol energy of the Kerpez constellation is

$$E = \frac{L}{3}(2^{2L} - 1) \tag{17}$$

and the minimum squared distance between constellation points is $d^2_{\min} = 4(2^{2L} - 1)/3$. The normalized constellation figure of merit for the additive white Gaussian noise channel is thus

$$\mathrm{CFM}_{\mathrm{AWGN}} = d^2_{\min}/E = 4/L. \tag{18}$$

For the baseline constellation, $d^2_{\min} = 4$ and the constellation symbol energy is $E = L$ so that $\mathrm{CFM}_{\mathrm{AWGN}}$ for the baseline constellation is also given by (18). Thus, for the AWGN channel the probability of symbol error of both the Kerpez and baseline schemes is the same at high signal-to-noise ratio. However, unlike the baseline scheme, the Kerpez constellation is resistant to fading because no two L-dimensional vectors have the same j^{th} component, $1 \leq j \leq L$. Even if all but one of the L components have a fading amplitude of zero, the receiver will be able to determine which of the 2^L vectors was sent as long as the signal energy in the nonfaded component is strong enough.

The output of the i^{th} correlator is given by (2) where $m_i, 1 \leq i \leq L$, is given by (14). The total energy in the signalling interval $[0, T]$ is, from (17), $LE_s\overline{m_i^2} = L^2 E_s(2^{2L} - 1)/3$ and L bits are transmitted during this interval. In order to compare the different schemes on the basis of the same energy per bit, we assume that $E_s = 6E_b/(L(2^{2L} - 1))$.

As in the previous section, the receiver requires knowledge of the fading amplitudes α_i, $1 \leq i \leq L$. The receiver decides which L-dimensional symbol in C was transmitted using the same decision rule as in the previous section. The probability of bit error depends on how bit patterns of length L are assigned to constellation points. Given an L-dimensional constellation, we may wish to assign bit patterns to constellation points such that points separated by $d_{\min}$ differ in only one bit. This is because in an AWGN channel, symbol detection errors are more likely to be associated with nearest-neighbour symbol pairs than with symbol pairs that are separated by a larger distance. For the Kerpez constellation it turns out that there is a simple method for achieving this as follows.

Method for bit pattern assignment: There are 2^L constellation points. For a given constellation point with indexes $(i_1, i_2, \ldots, i_L)$, $i_1 \in \{1, 2, 3, \ldots, 2^L\}$, we let $i_1 = 1 + \sum_{n=1}^{L} 2^{n-1} a_n$, where $a_n \in \{0,1\}$. We compute i_v using (15), for $2 \le v \le L$. If we assign the bit pattern $(a_1, a_2, a_3, \ldots, a_L)$ to the constellation point $(i_1, i_2, i_3, \ldots, i_L)$ then bit patterns that differ in only one location correspond to constellation points separated by $d_{\min}$. The proof is given in Appendix II.

As an example, consider the case $L = 3$. In Table 1 the mapping between constellation points and bit patterns is given. Note that i_1 is just the decimal representation of the bit pattern $a_1a_2a_3$ plus one. For this constellation $d^2_{\min} = 84$ and all pairs of constellation points separated by this squared distance (there are twelve of these pairs) differ in only one bit.

(m_1, m_2, m_3)	(i_1, i_2, i_3)	$a_1a_2a_3$
(7,1,5)	(1,4,2)	000
(5,-7,1)	(2,8,4)	001
(3,3,-3)	(3,3,6)	010
(1,-5,-7)	(4,7,8)	011
(-1,5,7)	(5,2,1)	100
(-3,-3,3)	(6,6,3)	101
(-5,7,-1)	(7,1,5)	110
(-7,-1,-5)	(8,5,7)	111

Table 1: Mapping of bit patterns to constellation points for $L = 3$ Kerpez constellation.

8 Performance comparisons

Simulations were performed to determine the performance of the optimally-rotated L-cube constellations and the Kerpez constellations. For a given constellation, bit patterns that differed in only one bit were assigned to nearest neighbour constellation points. (The method for bit pattern assignment for the Kerpez constellations is described in section 7.) The bit error rate of the constellation was simulated as follows. An L-dimensional point $(m_1, m_2, \ldots, m_L)$ was chosen at random and the L decision variables $(y_1, y_2, \ldots, y_L)$ were computed according to (2) for a given signal energy and noise variance. The decision rule used to map this vector to the constellation point $(\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L)$ was described in section 2. The process was repeated several million times at several values of E_b/N_0.

Figures 5–8 show the probability of bit error for the different schemes. In each figure, two curves are theoretical plots, corresponding to the analytical expressions (4) and (5). The curves corresponding to the Kerpez constellations and optimally-rotated L-cube constellations were obtained by simulation. As L is increased from 2 to 5, the performance of all schemes at high SNR improves except that of the baseline transmission scheme, since the bit error rate varies as $1/(E_b/N_0)^L$ for these fading-resistant schemes. As L is increased, the SNR gap between the Kerpez and optimal constellations widens. When $L = 2$, the Kerpez constellation is slightly better than the optimally-rotated constellation, although the difference is negligible, because optimal rotation of the L-cube only minimizes an *upper bound* on the probability of *symbol* error. In fact, the $\mathrm{CFM}_{\mathrm{Rayleigh}}$ for these two constellations are nearly identical, and it is hard to predict which constellation will perform better when this is the case without resorting to simulation. When $L = 5$, the optimally-rotated constellation has a performance advantage of 2.5 dB over the Kerpez constellation, at a bit error rate of 10^{-4}. As L increases beyond $L = 5$, the rate of performance improvement of the diversity schemes over the baseline scheme diminishes, as expected. A significant performance advantage with as little as two antennas can be achieved[2]

9 Conclusion

We have assumed a base station transmitter with L different antennas and have considered several different transmission schemes. The simplest scheme, in which transmissions are independent, has the same performance as uncoded BPSK. The three fading-resistant schemes discussed assume that the receiver can estimate the fading amplitudes in the different links. Optimally rotated baseline constellations and Kerpez constellations are shown to provide performance comparable to that of the bandwidth-costly method of transmitting identical bit streams on each of the antennas. As the number of antennas increases, the performance advantage of the optimally rotated constellations over the Kerpez constellations increases. When there are $L = 2$ antennas, the signal-to-noise ratio advantage of the Kerpez and optimal constellations over uncoded BPSK is 13 dB, assuming a bit error rate of 10^{-4}. When $L = 5$, the performance advantage of the optimally rotated constellation over the Kerpez constellation is 2.5 dB, and the rotated constellation is 19 dB better than uncoded BPSK, at a bit error rate of 10^{-4}.

[2]Note that this scheme is different than rotated QPSK with Rayleigh fading where the bit error rate is invariant with rotations of the constellation. In our case each of two components undergoes Rayleigh fading.

Although it was assumed that there are L transmitter antennas, the results presented can still be applied to the case where a single transmitter antenna is used and where diversity is achieved using other means, for example, L different frequency bands or time slots which are separated far enough apart so that the assumption of uncorrelated fading in the different links still applies.

Appendix I

Theorem (QR factorization)
Given an $n \times m$ matrix A with $n \geq m$, there is an $n \times m$ matrix Q with orthonormal columns and an upper triangular $m \times m$ matrix R such that $A = QR$. If $m = n$, Q is orthogonal. (e.g. see [15, p. 112])

There are several algorithms that can be used to compute Q and R. For example, methods based on Householder and Givens transformations can be used. The method based on Givens transformations is of particular interest to us (e.g. see [16, p.214]). If A is an $n \times n$ matrix, then $A = QR \Rightarrow Q^T A = R$. The QR factorization method based on Givens transformations gives us a method for writing Q^T as the product of $n(n-1)/2$ Givens matrices.

$$G(i,k,\theta) = \begin{bmatrix} 1 & & & & & & \\ & \cdots & & & & & \\ & & c & \cdots & s & & \\ & & \vdots & & \vdots & & \\ & & -s & \cdots & c & & \\ & & & & & \cdots & \\ & & & & & & 1 \end{bmatrix} \tag{19}$$

is an $n \times n$ matrix called the Givens (or Jacobi) matrix, with $c = \cos\theta$, $s = \sin\theta$. It consists of 1's on the main diagonal except for the two elements c in rows (and columns) i and k. All off-diagonal elements are zero except the two elements s and $-s$ in rows (and columns) i and k. Post-multiplication of a vector by G rotates the vector counter-clockwise by θ degrees with respect to the (i,k) plane.

Proposition 1

If an $n \times n$ matrix A is orthogonal, then a QR factorization algorithm yields $Q = A\hat{I}$, where $\hat{I}$ denotes a matrix in which each main diagonal element is either 1 or -1 and all off-diagonal elements are zero.

Proof

Q^T is orthogonal because Q is orthogonal. If A is orthogonal then $Q^T A = R$ is also orthogonal since the product of two orthogonal matrices is an orthogonal matrix. Since R is an upper triangular orthogonal matrix, each main diagonal element must be either 1 or -1 and all off-diagonal elements must be zero. We have $A = Q\hat{I} \;\Rightarrow\; Q = A\hat{I}^{-1} = A\hat{I}^T = A\hat{I}$.

Proposition 2

Any orthogonal $n \times n$ matrix A can be written as the product of $n(n-1)/2$ $n \times n$ Givens matrices and an $n \times n$ $\hat{I}$ matrix.

Proof

The proof follows directly from the proof of the previous proposition. Given an arbitrary orthogonal $n \times n$ matrix A we have a method for constructing Q, the product of $n(n-1)/2$ Givens matrices, such that $A = Q\hat{I}$. Note that the identity matrix is a Givens matrix as well as an $\hat{I}$ matrix.

For example, all 2×2 orthogonal matrices have the form

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{20}$$

or

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}. \tag{21}$$

This is a well known fact (for example, see [17, p. 282]). The first form corresponds to $G(1,2,\theta)I_{2\times2}$ and the second form corresponds to $G(1,2,\theta)(-I_{2\times2})$. Proposition 2, together with the QR factorization method using Givens transformations, implies that all 3×3 orthogonal matrices have the form

$$A \;=\; G(1,2,\theta_{12})G(1,3,\theta_{13})G(2,3,\theta_{23})\hat{I}_{3\times3}$$

$$= \begin{bmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta_{13} & 0 & \sin\theta_{13} \\ 0 & 1 & 0 \\ -\sin\theta_{13} & 0 & \cos\theta_{13} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{bmatrix}$$

$$\cdot \begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \end{bmatrix} \tag{22}$$

In general, any $L \times L$ orthogonal matrix A can be factored into $L(L-1)/2 = \binom{L}{2}$ rotation matrices and an $\hat{I}$ matrix as follows

$$A = \Big(\prod_{\substack{1 \le i \le L-1, \\ i+1 \le j \le L}} G(i, j, \theta_{ij}) \Big) \cdot \hat{I}_{L \times L} \tag{23}$$

for suitable θ_{ij}.

Appendix II

Assume that the bit patterns corresponding to the distinct constellation points $(i_1, i_2, \ldots, i_L)$ and $(j_1, j_2, \ldots, j_L)$ are $(a_1, a_2, \ldots, a_L)$ and $(b_1, b_2, \ldots, b_L)$, respectively. Since the points are distinct, a property of the Kerpez constellation is that $i_v \neq j_v$ for all $v \in \{1, 2, \ldots, L\}$. Define $d_{v_{ij}} = |i_v - j_v|$. From [9], if $d^2_{\text{min}} = \sum_{v=1}^{L} d^2_{v_{ij}}$ then the distance vector between these two points, $[d_{1_{ij}}, d_{2_{ij}}, \ldots, d_{L_{ij}}]$, is a permutation of $[2^{L-1}, 2^{L-2}, \ldots, 2^2, 2, 1]$. According to the method in Section 7, we have $i_1 = 1 + \sum_{n=1}^{L} 2^{n-1} a_n$ and $j_1 = 1 + \sum_{n=1}^{L} 2^{n-1} b_n$. Assume that these two bit patterns differ in the m^{th} location, $1 \le m \le L$, that is,

$$b_n = \begin{cases} a_n, & n \neq m \\ a_n \oplus 1, & n = m \end{cases} \tag{24}$$

We want to show that the distance vector between these two constellation points is a permutation of $[2^{L-1}, 2^{L-2}, \ldots, 2^2, 2, 1]$.

We have

$$\begin{aligned} j_1 - i_1 &= \sum_{n=1}^{L} 2^{n-1}(b_n - a_n) \\ &= 2^{m-1}(a_m \oplus 1 - a_m) \\ &= \pm 2^{m-1} \end{aligned} \tag{25}$$

Also, for $2 \leq v \leq L$, we get

$$j_v - i_v = (j_1 - 1)2^{L-v+1} \mod 2^L - (i_1 - 1)2^{L-v+1} \mod 2^L - \left\lceil \frac{j_1}{2^{v-1}} \right\rceil + \left\lceil \frac{i_1}{2^{v-1}} \right\rceil \tag{26}$$

and

$$\begin{aligned} (i_1 - 1)2^{L-v+1} \mod 2^L &= \left(\sum_{n=1}^{L} 2^{n-1} a_n \right) 2^{L-v+1} \mod 2^L \\ &= \sum_{n=1}^{v-1} 2^{L+n-v} a_n \mod 2^L \\ &= \sum_{n=1}^{v-1} 2^{L+n-v} a_n \end{aligned} \tag{27}$$

where (27) follows from the relation $\sum_{n=1}^{v-1} 2^{L+n-v} a_n \leq 2^{L-v} \sum_{n=1}^{v-1} 2^n < 2^L$. Also,

$$\begin{aligned} \left\lceil \frac{i_1}{2^{v-1}} \right\rceil &= \left\lceil \frac{1 + \sum_{n=1}^{L} 2^{n-1} a_n}{2^{v-1}} \right\rceil \\ &= \sum_{n=v}^{L} 2^{n-v} a_n + \left\lceil \frac{1 + \sum_{n=1}^{v-1} 2^{n-1} a_n}{2^{v-1}} \right\rceil \\ &= \sum_{n=v}^{L} 2^{n-v} a_n + 1 \end{aligned} \tag{28}$$

where (28) follows from the relation $1 + \sum_{n=1}^{v-1} 2^{n-1} a_n \leq 1 + \sum_{n=1}^{v-1} 2^{n-1} = 2^{v-1}$. Substituting (27) and (28) in (26), we have

$$j_v - i_v = \sum_{n=1}^{v-1} 2^{L+n-v}(b_n - a_n) + \sum_{n=v}^{L} 2^{n-v}(a_n - b_n), \quad 2 \leq v \leq L. \tag{29}$$

There are two cases to consider. Case *i)* $1 \leq m \leq v-1 \Rightarrow m+1 \leq v \leq L$. Eq. (29) gives $|j_v - i_v| = |2^{L+m-v}(a_m \oplus 1 - a_m)| = 2^{L+m-v}$. Case *ii)* $v \leq m \leq L \Rightarrow 2 \leq v \leq m$. Eq. (29)

gives $|j_v - i_v| = |2^{m-v}(a_m - a_m \oplus 1)| = 2^{m-v}$. Thus, $[d_{1_{ij}}, d_{2_{ij}}, \ldots, d_{m-1_{ij}}, d_{m_{ij}}, d_{m+1_{ij}}, \ldots, d_{L_{ij}}]$ $= [2^{m-1}, 2^{m-2}, \ldots, 2, 1, 2^{L-1}, 2^{L-2}, \ldots, 2^m]$ which is a permutation of the minimum distance vector, as required.

Acknowledgments

We would like to thank Prof. Frank R. Kschischang for helpful discussions.

References

[1] Y. Akaiwa, "Antenna selection diversity for framed digital signal transmission in mobile radio channel", *IEEE Vehicular Technology Conference*, pp. 470–473, May 1989.

[2] Y. Kondo, K. Suwa, "Linear predictive transmitter diversity for microcellular TDMA/TDD mobile radio system", *IEEE Vehicular Technology Conference*, pp. 602–606, 1993.

[3] C. Schlegel, D. J. Costello, Jr., "Bandwidth efficient coding for fading channels: code construction and performance analysis", *IEEE Journal on Selected Areas in Communications*, vol. 7, pp. 1356-1368, Dec. 1989.

[4] D. Divsalar, M. K. Simon, "Trellis coded modulation for 4800–9600 bits/s transmission over a fading mobile satellite channel," *IEEE J. Select. Areas Commun*, vol. SAC-5, pp. 162- 177, Feb. 1987.

[5] Jun Du, Branka Vucetic,, "Trellis coded 16-QAM for fading channel", *European Transactions on Telecommunications*, vol. 4, no. 3, May–June 1993, pp. 335–341.

[6] James K. Cavers, Paul Ho, "Analysis of the error performance of trellis-coded modulations in Rayleigh-fading channels", *IEEE Trans. on Commun.*, vol. 40, pp. 74–83, Jan. 1992.

[7] Américo Brajal, Antoine Chouly, "Optimal trellis-coded 8-PSK and 4-AM modulations for the Rayleigh channel", *ICC '94*, pp. 28–33.

[8] V. M. DaSilva, E. S. Sousa, "Fading-resistant transmission from several antennas", *6th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC '95*, Toronto, Canada, September 27-29, 1995, pp. 1218–1222.

[9] K. J. Kerpez, "Constellations for good diversity performance", *IEEE Trans. Commun.*, vol. 41, pp. 1412-1421, Sept. 1993.

[10] K. Boullé, J. C. Belfiore, "Modulation schemes designed for the Rayleigh channel", *Conf. on Information Sciences and Systems Proc.*, Princeton, New Jersey, Dept. Elec. Eng., Princeton University, pp. 288-293, 1992.

[11] J. Proakis, *Digital Communications*, 2nd edition, McGraw-Hill Book Company, New York, 1989.

[12] R. Gallager, *Information Theory and Reliable Communication*, J. Wiley, New York, 1968.

[13] D. Divsalar, M. K. Simon, "The design of trellis coded MPSK for fading channels: performance criteria", *IEEE Trans. Commun.* vol. 36, no. 9, September 1988, pp. 1004–1012.

[14] J. H. Conway, N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, Springer–Verlag, New York, 1993.

[15] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[16] G. H. Golub, C. F. Van Loan, *Matrix Computations*, 2nd ed., The John Hopkins University Press, 1989.

[17] W. K. Nicholson, *Elementary Linear Algebra with Applications*, PWS–KENT Publishing Co., Boston, 1986.

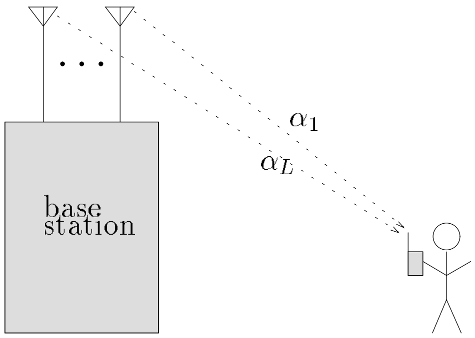

Figure 1: Forward link channel.

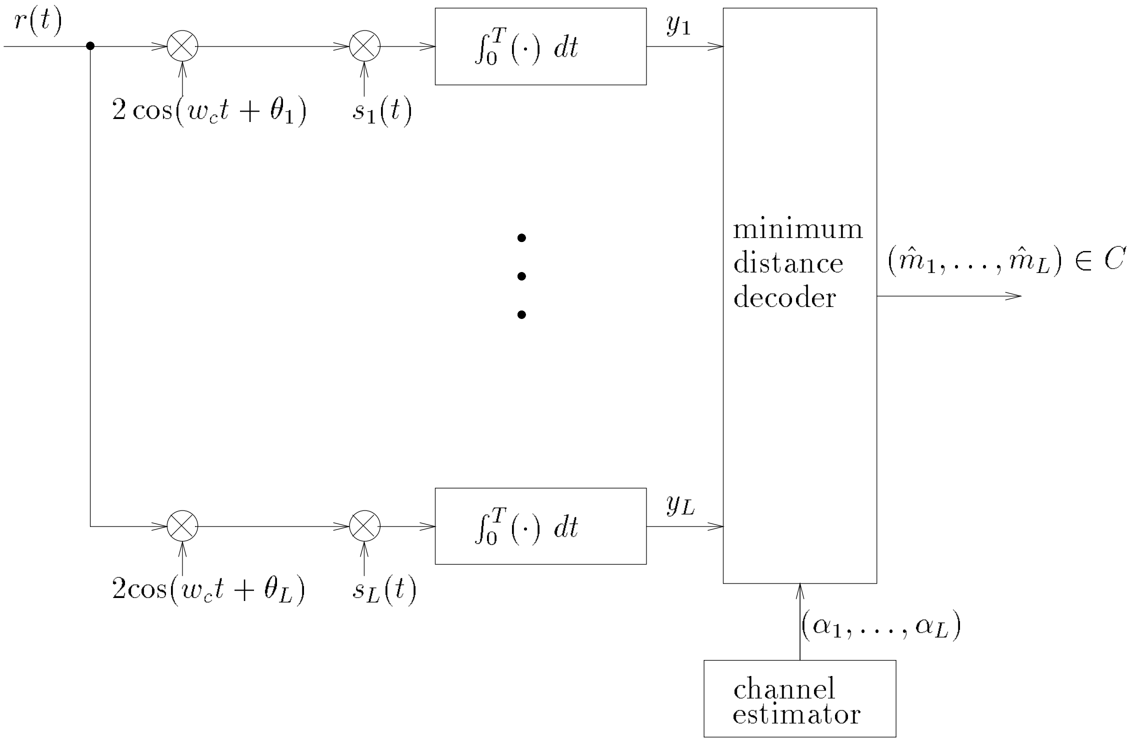

Figure 2: Receiver for L orthogonal transmissions.

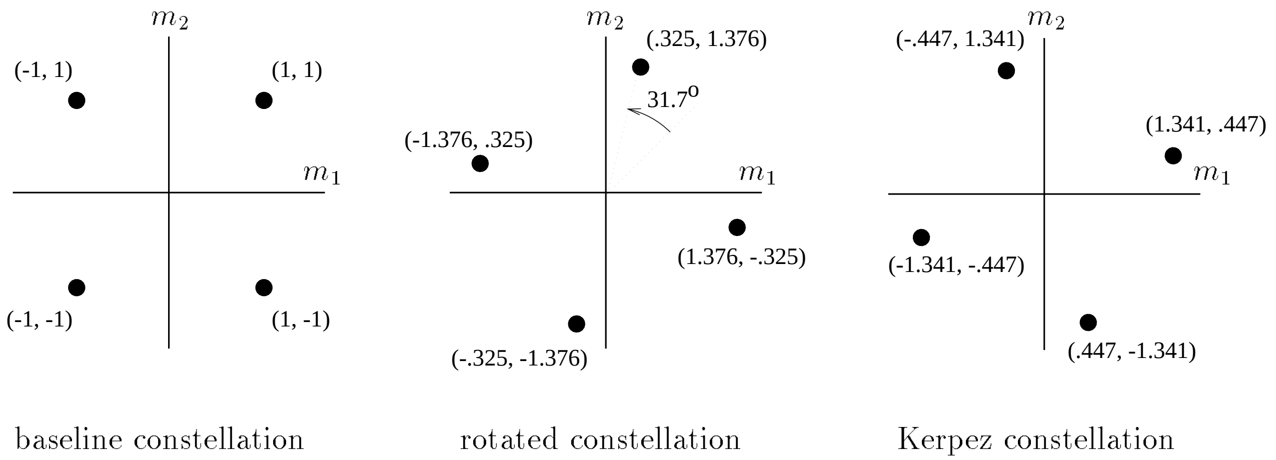

Figure 3: Examples of constellations for $L = 2$.

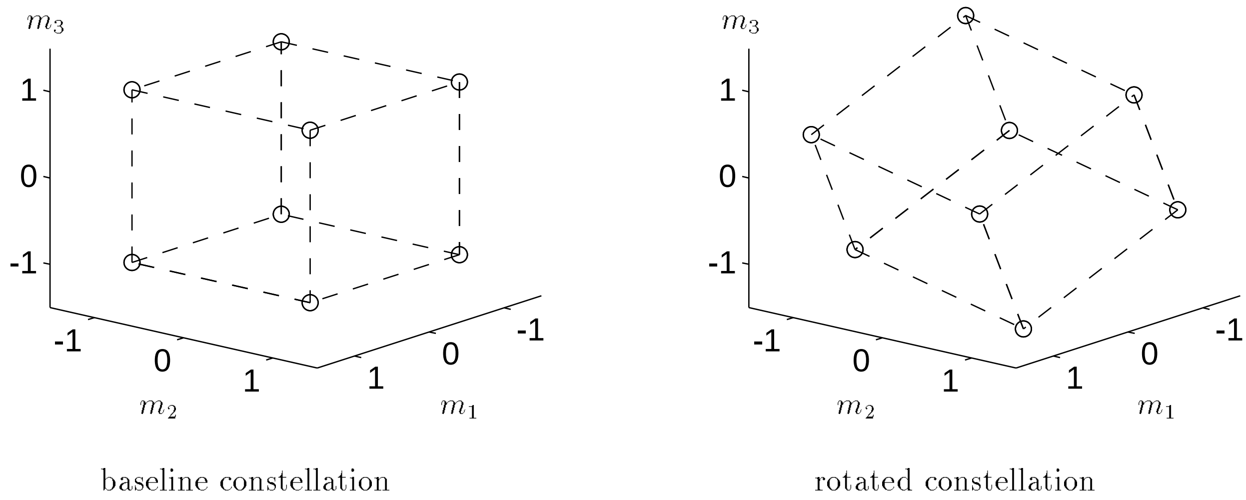

Figure 4: Examples of constellations for $L = 3$.

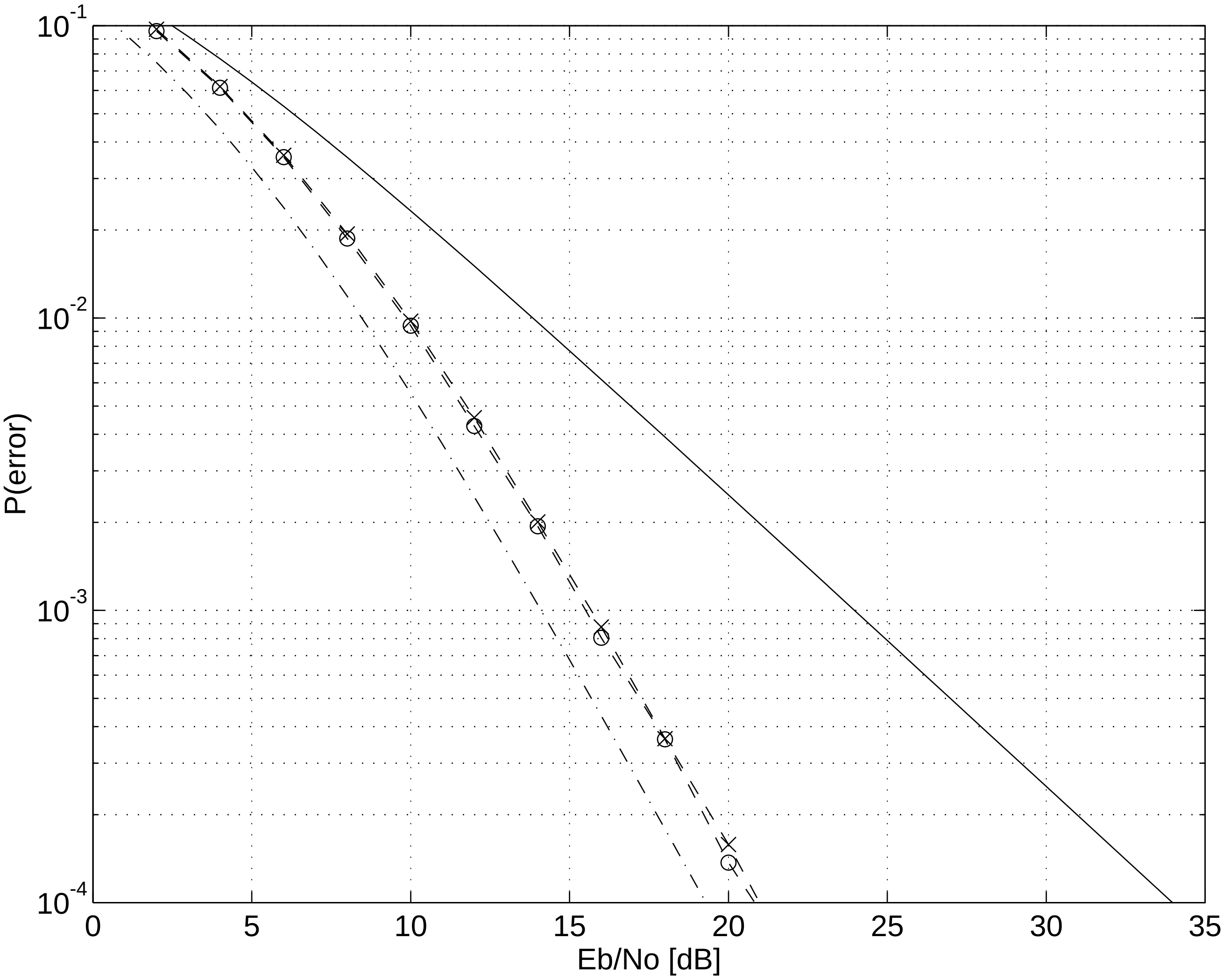

Figure 5: Probability of bit error for several transmission schemes. L=2. (—–) baseline, (-x-) optimal constellation, (-o-) Kerpez constellation, ($- \cdot -$) identical transmissions.

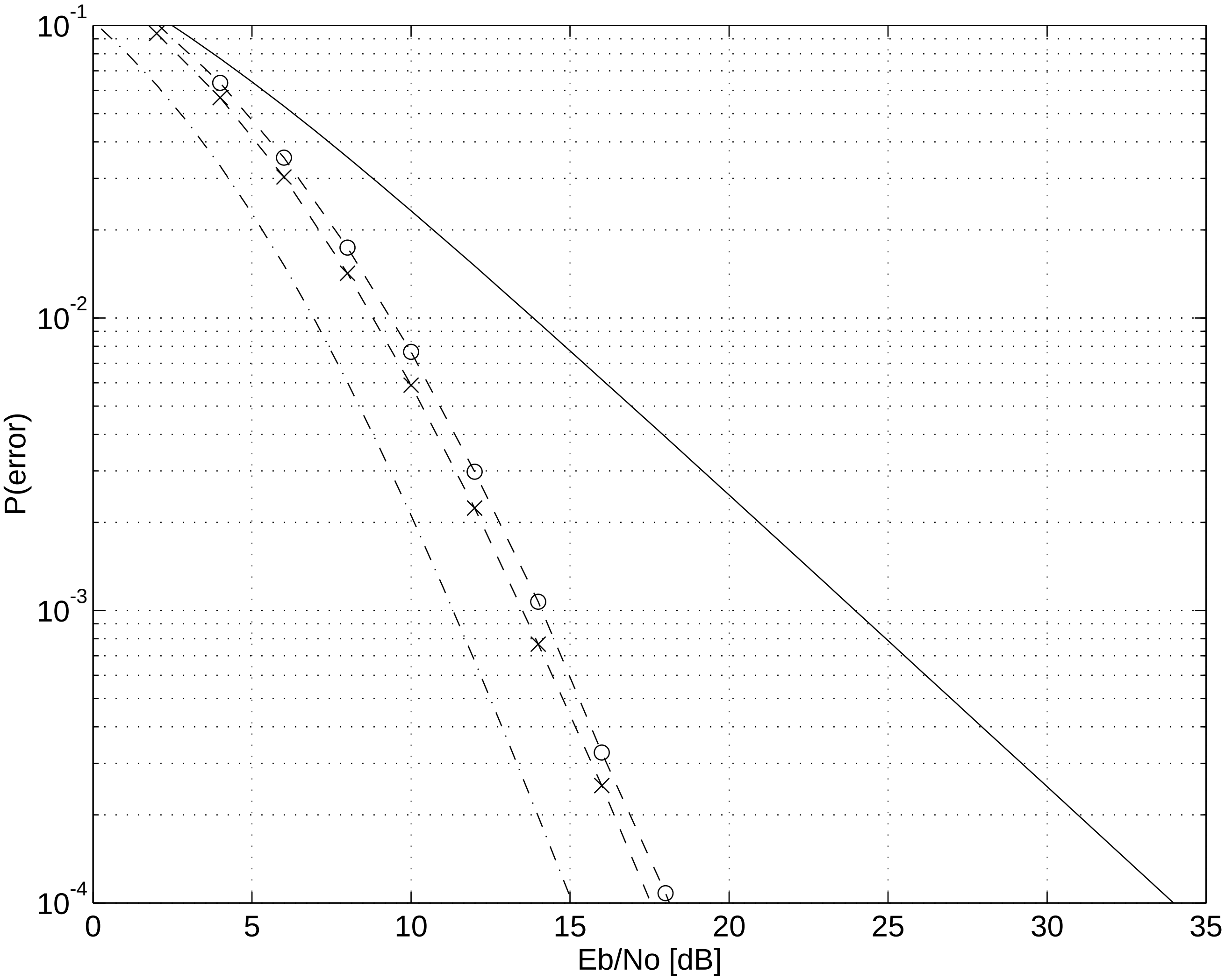

Figure 6: Probability of bit error for several transmission schemes. L=3. (—–) baseline, (-x-) optimal constellation, (-o-) Kerpez constellation, ($- \cdot -$) identical transmissions.

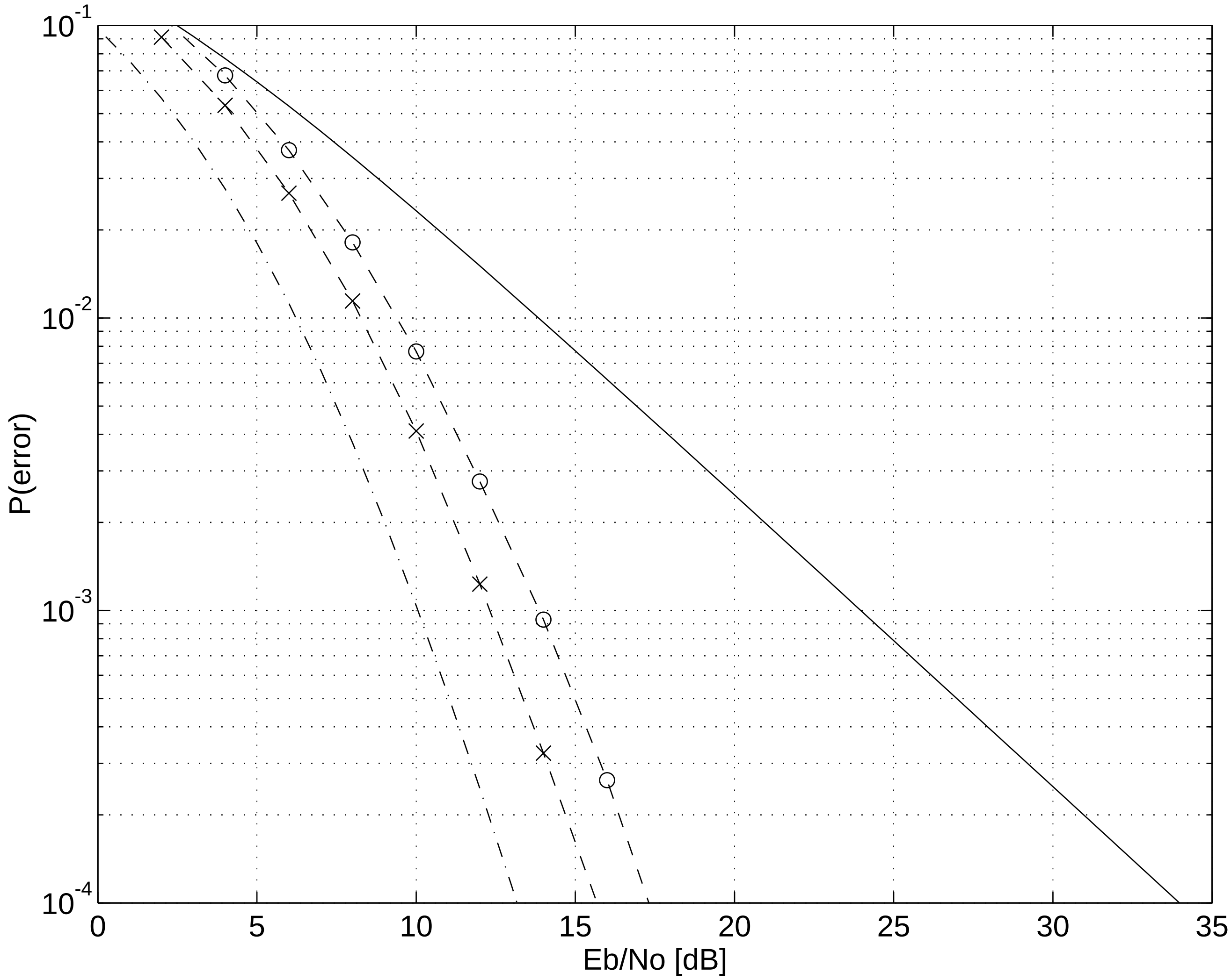

Figure 7: Probability of bit error for several transmission schemes. L=4. (—–) baseline, (-x-) optimal constellation, (-o-) Kerpez constellation, ($- \cdot -$) identical transmissions.

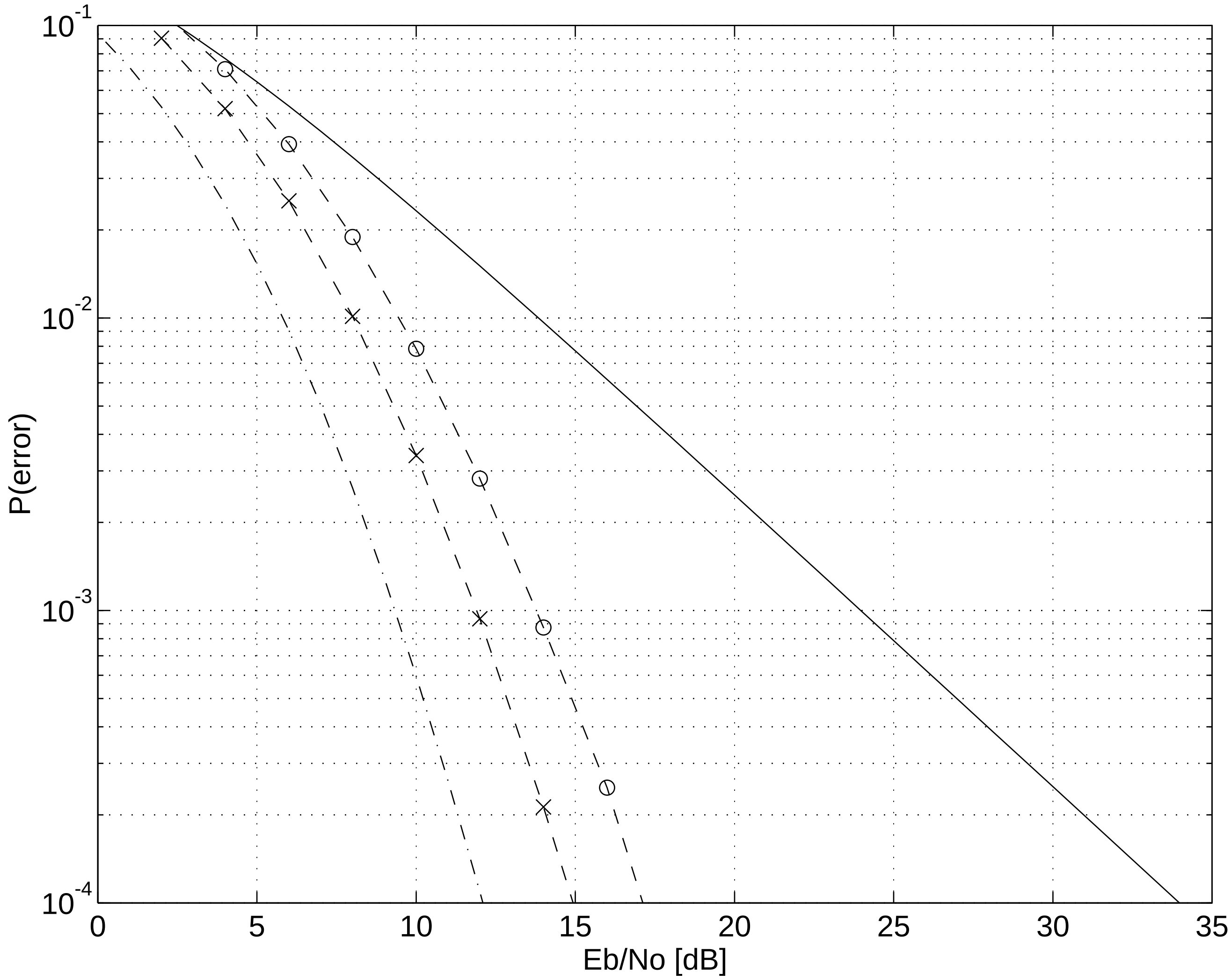

Figure 8: Probability of bit error for several transmission schemes. L=5. (—–) baseline, (-x-) optimal constellation, (-o-) Kerpez constellation, ($- \cdot -$) identical transmissions.