Permutation Codes

Emanuele Viterbo

July 15, 2002

Abstract

Permutation codes are special spherical codes designed for the bandlimited Gaussian channel. A Variant I permutation code is the set of codewords obtained by taking all permutations of an initial vector in the n-dimensional Euclidean space. A Variant II permutation code is the set of codewords obtained by taking all permutations and sign changes of the components of an initial vector in the n-dimensional Euclidean space. These codes may be attractive because of their simple maximum likelihood decoding algorithm.

Keywords: Codes for bandlimited Gaussian channel, Permutation modulation, Spherical codes, Group codes.

1 Introduction

Permutation codes have been proposed by David Slepian in 1965 [12]. In the quest for efficient codes for the bandlimited Gaussian channel, permutation codes are among the first attempts to solve the problem taking into account both coding gains and decoding efficiency. Permutation codes are multidimensional spherical signal constellations with the desirable property that they posses a very simple maximum likelihood (ML) decoding algorithm. Slepian used the term permutation modulation for his permutation codes.

A Variant I permutation modulation is the set of codewords obtained by taking all permutations of an initial vector in the n-dimensional Euclidean space. A Variant II permutation modulation is the set of codewords obtained by taking all permutations and sign changes of the components of an initial vector in the n-dimensional Euclidean space. Trivial examples of Variant I and Variant II modulations are orthogonal and biorthogonal codes, respectively. Also pulse code modulation (PCM), pulse position modulation (PPM) and simplex codes can be viewed as permutation modulations.

Good permutation modulations may be designed by appropriately selecting the initial vector. Permutation modulations may be very efficiently decoded by essentially applying a sorting algorithm to the received signal vector. Permutation modulations are very special cases of group codes for the bandlimited channel proposed by Slepian [10, 13].

Karlof later proposed the term permutation code for a generalization of permutation modulations [6]. In particular, he studied the group codes obtained from subgroups of the symmetric group (the group of permutations of n objects). The resulting spherical codes may be seen as particular subsets of the corresponding permutation modulation. We can think of permutation modulation as a "full rate" permutation code. In this case the initial vector selection problem and the decoding algorithm are much harder (see [6, 7, 8]).

Here, we will focus on permutation codes according to Slepian's definition. This article is organized as follows. The next section introduces the notation and gives detailed definition of permutation modulation. Performance in terms of error probability and the ML decoding algorithm are also presented. Section 3 gives some examples of permutation codes and discusses some application issues.

2 Theory

Digital transmission over the bandlimited Additive White Gaussian Noise (AWGN) channel is commonly modeled in the n-dimensional Euclidean space $\mathbf{R}^n$ as

$$\mathbf{y} = \mathbf{x} + \mathbf{n}$$

where $\mathbf{y}$ is the received signal vector, $\mathbf{x}$ is the transmitted signal vector (or codeword) taken from a finite signal constellation (or codebook) $\mathcal{S}$ and $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$ is a real Gaussian random vector with i.i.d. components. The space dimension n is related to the time-bandwidth product through the sampling theorem: if T is the signal duration and W the occupied bandwidth then $n = 2WT$.

Let $M = |\mathcal{S}|$ be the number of points in the constellation, we define the spectral efficiency as

$$\frac{R}{W} = \frac{2}{n} \log_2 M \quad \text{bit/s/Hz} \tag{1}$$

Let $r = \log_2 M$. If we are interested in transmitting binary information we simply label $2^{[r]}$ codewords by distinct binary vectors of length $[r]$ and disregard the remaining $M - 2^{[r]}$ codewords.[1]

The average signal power of $\mathcal{S} = \{\mathbf{x}_i\}_0^{M-1}$ is given by

$$\mathcal{P} = \frac{1}{nM} \sum_{i=0}^{M-1} \|\mathbf{x}_i\|^2 \tag{2}$$

where $\|\cdot\|$ denotes the Euclidean norm in $\mathbf{R}^n$.

The maximum likelihood (ML) receiver gives an estimate $\hat{\mathbf{x}}$ of the transmitted codeword $\mathbf{x}$ according to the minimum distance criterion

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}_i \in \mathcal{S}} \|\mathbf{y} - \mathbf{x}_i\|^2 \ . \tag{3}$$

We note that the complexity of the ML receiver greatly depends on the structure of the code $\mathcal{S}$. In the worst case a total of M Euclidean distances must be computed. For large values of M this may be impractical, hence it is common to trade some of the performance for a reduced decoding complexity. Many classical forward error correcting codes have been selected for applications because they have simple decoding algorithms.

The average codeword error probability with ML detection is given by

$$P(e) = \frac{1}{M} \sum_{i=0}^{M-1} P(e|\mathbf{x}_i) = \frac{1}{M} \sum_{i=0}^{M-1} \int_{\overline{\mathcal{R}_i}} \frac{e^{-\|\mathbf{x}-\mathbf{x}_i\|^2/(2\sigma^2)}}{(2\pi\sigma^2)^{n/2}} \, d\mathbf{x} \tag{4}$$

[1] $[x]$ denotes the greatest integer smaller than x.

where $\overline{\mathcal{R}_i} = \mathbf{R}^n \backslash \mathcal{R}_i$ is the complement of the decision region corresponding to the codeword $\mathbf{x}_i$, defined as

$$\mathcal{R}_i = \{\mathbf{z} \in \mathbf{R}^n \,:\, \|\mathbf{z} - \mathbf{x}_i\| < \|\mathbf{z} - \mathbf{x}_j\|, \quad \forall j \neq i\} \ . \tag{5}$$

These regions are also known as minimum distance or ML regions.

Good codes should be designed in order to minimize $P(e)$, given the parameters M, n, $\mathcal{P}$ and σ. Shannon showed that the codeword error probability decreases exponentially with n and gave the famous asymptotic result that, for given R, W and $\mathcal{P}/\sigma^2$, the $P(e)$ can be made arbitrarily small as $n \to \infty$, provided that $R < C$, where

$$C = W \log_2(1 + \mathcal{P}/\sigma^2) \ . \tag{6}$$

On the contrary, if $R > C$ then $P(e) \to 1$ as $n \to \infty$.

Explicit construction of optimal codes is an open problem and has been analyzed in some very special cases only [1, 15].

2.1 Definitions

Let $\{\mu_1, \ldots, \mu_k\}$ be a set of distinct real numbers with $\mu_1 < \mu_2 < \cdots < \mu_k$ and let $\{m_1, \ldots, m_k\}$ be a set of positive integers such that

$$n = \sum_{j=1}^{k} m_j \ . \tag{7}$$

Consider the initial vector with components sorted in ascending order

$$\mathbf{x}_0 = (\underbrace{\mu_1, \ldots, \mu_1}_{m_1}, \underbrace{\mu_2, \ldots, \mu_2}_{m_2}, \ \cdots \ , \underbrace{\mu_k, \ldots, \mu_k}_{m_k}) \ . \tag{8}$$

A Variant I permutation code consists of the set of vectors obtained by permuting the components of the initial vector $\mathbf{x}_0$. The total number of codewords in such a code is

$$M_I = \frac{n!}{m_1!\, m_2! \cdots m_k!} \ . \tag{9}$$

The Variant I code with $k = 2$, $m_1 = n - 1$, $m_2 = 1$ and $\mu_1 = 0$ is the well known PPM or orthogonal modulation.

A Variant II permutation code consists of the set of vectors obtained by permuting and applying all possible sign changes to the components of the initial vector $\mathbf{x}_0$. Without loss of generality we may assume

$$0 \leq \mu_1 < \mu_2 < \cdots < \mu_k \ . \tag{10}$$

The total number of codewords in this code is

$$M_{II} = \frac{2^h n!}{m_1!\, m_2! \cdots m_k!} \tag{11}$$

where $h = n - m_1$, if $\mu_1 = 0$ and $h = n$, if $\mu_1 > 0$.

The Variant II code with $k = 2$, $m_1 = n - 1$, $m_2 = 1$ and $\mu_1 = 0$ results in the well known biorthogonal modulation. The Variant II code with $k = 1$, $m_1 = n$, $\mu_1 \neq 0$ yields an n-bit PCM. In this case the points of $\mathcal{S}$ correspond to the 2^n vertices of an n dimensional hypercube of edge length $2\mu_1$.

It is clear that all codewords of both Variant I and II codes lie on a hypersphere of radius $\sqrt{n\mathcal{P}}$ centered at the origin and

$$\mathcal{P} = \frac{1}{n} \sum_{j=1}^{k} m_j \mu_j^2 \ . \tag{12}$$

2.2 Decoding

Let us consider ML decoding of Variant I codes. We need to find the minimum of the the quantities

$$\|\mathbf{y} - \mathbf{x}_i\|^2 = \|\mathbf{y}\|^2 + \|\mathbf{x}_i\|^2 - 2(\mathbf{y}, \mathbf{x}_i) = \|\mathbf{y}\|^2 + n\mathcal{P} - 2(\mathbf{y}, \mathbf{x}_i) \quad i = 0, \ldots, M-1 \tag{13}$$

where $(\mathbf{y}, \mathbf{x}_i)$ denotes the scalar product of the two vectors. Since $\|\mathbf{y}\|$ is independent of i the ML decoder may simply maximize the scalar product between the received vector and the codewords, i.e.

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x} \in \mathcal{S}} \sum_{k=1}^{n} x_k y_k \ . \tag{14}$$

This maximization problem may be solved as follows. Given the received vector $\mathbf{y}$, replace the smallest m_1 components by the values μ_1, replace the smallest m_2 remaining components with μ_2, etc. Until all the components have been replaced.

It is interesting to show the very simple and elegant proof of the optimality of this decoding algorithm given by Slepian. In particular, we want to show that the sum

$$x_{i_1}y_1 + x_{i_2}y_2 + \cdots + x_{i_n}y_n \tag{15}$$

is maximized by the permutation of indexes $(i_1, i_2, \ldots, i_n)$ which pairs the largest x to the largest y, the second largest x to the second largest y, etc.

For $n = 1$ it is trivially true, then we proceed by induction. For some $n > 1$, let $\bar{x}$ and $\bar{y}$ denote the largest x and the largest y. If $\bar{x}$ is not paired with $\bar{y}$ in (15) then the sum contains the two terms $\bar{x}y' + \bar{y}x'$ for some $x' \leq \bar{x}$ and $y' \leq \bar{y}$. If we swap x' with $\bar{x}$ the sum (15) will decrease, in fact

$$(\bar{x}\bar{y} + x'y') - (\bar{x}y' + \bar{y}x') = (\bar{x} - x')(\bar{y} - y') \geq 0 \tag{16}$$

Hence, pairing $\bar{x}$ with $\bar{y}$, maximizes the sum (15). We now delete $\bar{x}\bar{y}$ from the sum and proceed by induction on the $n-1$ terms: pairing the second largest x to the second largest y, does not reduce the $n-1$ term sum etc.

For Variant II codes ML decoding can be performed as follows.

1. Take the absolute value of the components of the received vector $\mathbf{y}$, i.e. let

$$\mathbf{y}' = (|y_1|, |y_2|, \ldots, |y_n|)$$

2. Apply the decoder of Variant I codes to $\mathbf{y}'$ to make a first decision $\mathbf{x}'$.

3. The final decision is given by

$$\hat{\mathbf{x}} = (\mathrm{sgn}(y_1)x_1',\ \mathrm{sgn}(y_2)x_2',\ \ldots,\ \mathrm{sgn}(y_n)x_n')$$

where $\mathrm{sgn}(x) = +1$, if $x \geq 0$ and $\mathrm{sgn}(x) = -1$, if $x < 0$.

It can be shown that the above algorithm is equivalent to solving the maximization problem (14).

The complexity of these decoding algorithms is rather small if compared to the brute force exhaustive search. In particular, it is enough to perform a sorting algorithm on the

n components of the received vector and to keep track of the final index permutation. This permutation uniquely identifies the ML decoded codeword and the corresponding information bit label may be easily recovered. Sorting can be performed with a complexity of $O(n \log(n))$, whereas exhaustive decoding requires Mn multiplications and $M(n-1)$ additions.

2.3 Decision regions and error probability

The evaluation of the average codeword error probability (4) for permutation codes can be simplified by the following argument. Consider the collection $\mathcal{C}$ of all $n \times n$ permutation matrices, i.e., matrices having a single entry equal to one in each row and column and zeros in the remaining positions. When a permutation matrix is applied to the vectors of the codebook of a Variant I code it simply maps the codebook back into itself. In $\mathcal{C}$ we can find a permutation matrix A_{ij} that maps any codeword $\mathbf{x}_i$ into the codeword $\mathbf{x}_j$.

The permutation matrices are also orthogonal matrices so that, when they operate on $\mathcal{S}$, they preserve the distances between the points. Since the decision regions are defined in (5) in terms of distances, the permutation matrix A_{ij} also sends $\mathcal{R}_i$ into $\mathcal{R}_j$. Thus all the decision regions of a Variant I code are congruent and (4) reduces to $P(e) = P(e|\mathbf{x}_i)$, which is independent of i.

For Variant II codes, a similar argument also enables to conclude that $P(e) = P(e|\mathbf{x}_i)$ is independent of i. In particular, it is enough to replace the collection $\mathcal{C}$ by the collection $\mathcal{O}$ of $2^n n!$, $n \times n$ matrices having a single nonzero entry equal to $+1$ or -1 in each row and column.

We note that the above simplification is similar to the one that can be used in evaluating the codeword error probability of linear codes, where it is convenient to consider the case where the all zero codeword is transmitted. For permutation codes we will focus on $P(e|\mathbf{x}_0)$.

Let us now consider in detail the average codeword error probability of Variant I codes. First observe that the received vector components have independent Gaussian distributions. The first m_1 components have mean μ_1 and variance σ^2, the next m_2 components have mean μ_2 and variance σ^2, etc.

Assume the codeword $\mathbf{x}_0$ was transmitted. To understand when a decoding error appears, let us split the received vector components into k runs of length m_j each

$$\mathbf{y} = (y_1^{(1)}, y_2^{(1)}, \ldots, y_{m_1}^{(1)}, y_1^{(2)}, y_2^{(2)}, \ldots, y_{m_2}^{(2)}, \quad \cdots \quad, y_1^{(k)}, y_2^{(k)}, \ldots, y_{m_k}^{(k)}) \ . \tag{17}$$

The correct decision will be made if the first m_1 components are smaller than the following m_2 components, the next m_2 components are smaller than the following m_3 components, etc. Then we can write

$$P_I(e) = P_I(\mu_1, \mu_2, \ldots, \mu_k) = 1 - P\{\eta_1 \leq \xi_2 \leq \eta_2 \leq \xi_3 \leq \cdots \leq \eta_{k-1} \leq \xi_k\} \tag{18}$$

where, for $j = 1, \ldots, k$

$$\begin{aligned} \xi_j &= \min(y_1^{(j)}, y_2^{(j)}, \ldots, y_{m_j}^{(j)}) \\ \eta_j &= \max(y_1^{(j)}, y_2^{(j)}, \ldots, y_{m_j}^{(j)}) \end{aligned} \tag{19}$$

and the n independent Gaussian random variables $y_i^{(j)}$ have mean μ_j and variance σ^2.

Note that $P_I(e)$ only depends on the differences of the μ's, i.e., for all δ

$$P_I(\mu_1 + \delta, \mu_2 + \delta, \ldots, \mu_k + \delta) = P_I(\mu_1, \mu_2, \ldots, \mu_k) \tag{20}$$

Let $\beta_1 = \mu_1$ and $\beta_i = \mu_i - \mu_{i-1}$, for $i = 2, \ldots, k$. Let

$$\phi(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2/(2\sigma^2)} \tag{21}$$

be the probability distribution function of a zero-mean Gaussian random variable with variance σ^2 and

$$\Phi(x) = \int_{-\infty}^{x} \phi(z)\, dz \tag{22}$$

the corresponding cumulative distribution function.

In [12], it is shown that for both Variant I and II it is possible to bound $P(e)$ as

$$B - B^2/2 \leq P(e) \leq B \tag{23}$$

For Variant I codes

$$B = B_I = \sum_{i=2}^{k} P_{m_i,m_{i-1}}(\beta_i) \tag{24}$$

and

$$P_{m,n}(\alpha) = P_{n,m}(\alpha) = n \int_{-\infty}^{\infty} \phi(z)[1-\Phi(z)]^{n-1}[1-\Phi^m(z+\alpha)]\, dz\ . \tag{25}$$

For Variant II, when $\mu_1 = 0$, we have

$$B = B_{II} = \overline{P}_{m_1,m_2}(\beta_i) + \sum_{i=3}^{k} P_{m_i,m_{i-1}}(\beta_i) \tag{26}$$

where

$$\overline{P}_{m,n}(\alpha) = 2m \int_{0}^{\infty} \phi(z)[1-2\Phi(-z)]^{m-1}\{1-[1-\Phi(z-\alpha)]^n\}\, dz \tag{27}$$

and when $\mu_1 \neq 0$

$$B = B_{II'} = 1 - [1-\Phi(-\beta_1)]^{m_1} + \sum_{i=2}^{k} P_{m_i,m_{i-1}}(\beta_i)\ . \tag{28}$$

3 Evaluation

We are now able to consider the problem of selecting good permutation codes for a fixed dimension n. We want to optimize the choice of k, the μ's and the m's. From (1), (9) and (11) we see that R is independent of the μ's and depends only on k and the m's.

It is natural to fix k and the m's (i.e., fix R) and choose the μ's to minimize $\mathcal{P}$ for some fixed value of $P(e)$.

For Variant I codes, the first optimization step to reduce the average signal power is to center the signal constellation $\mathcal{S}$ around its barycenter, by selecting

$$\sum_{j=1}^{k} m_j \mu_j = 0. \tag{29}$$

By imposing the above condition to PPM, we obtain the simplex modulation. For Variant II codes, $\mathcal{S}$ is already centered around its barycenter.

The optimization problem was solved numerically in [12] and some optimal codes are presented for various n, m's and k for two values of $P(e) = 10^{-3}$ and $P(e) = 10^{-5}$. A simplified version of the code optimization problem was solved analytically by Biglieri and Elia in [2] and independently by Ingemarson [9]. They selected the μ's in order to maximize the minimum Euclidean distance $d_{\min}$, among the points of $\mathcal{S}$, for a fixed average power $\mathcal{P}$. Here, we report the optimal codes for $P(e) = 10^{-5}$ in Table 1.

Figures 1, 2, 3 and 4 show the codeword error probability of the codes in Table 1 as a function of E_b/N_0, where $E_b = n\mathcal{P}/r$ and $N_0 = 2\sigma^2$. These figures may be interpreted as follows. Given the system constraints T (maximum acceptable delay) and W (available bandwidth), we have n, then we can chose among the codes of different rate the one that satisfies our packet error rate $P(e)$ requirement.

For example Code 3 for $n = 5$ enables to transmit at the same rate of the 5 bit PCM with a lower $P(e)$, corresponding to an asymptotic gain of 1.2dB (see Fig. 1). Code 10 enables to transmit at 0.9 the rate of PCM with an asymptotic gain of 2.2dB (see Fig. 2). Code 23 enables to transmit at almost at the same rate of PCM with an asymptotic gain of 2.9dB (see Fig. 3).

Figure 5 shows the performance of the optimal codes given in [12] at $P(e) = 10^{-3}$. The comparison with Figure 1 shows that their performance in almost identical in the two cases, so we may conclude that the optimization is quite insensitive to the value $P(e)$.

Given an n dimensional code we define its asymptotic coding gain with respect to an n bit PCM as

$$\gamma = \frac{d^2_{\min}/E_b}{d^2_{\min,PCM}/E_{b,PCM}} \tag{30}$$

We report γ in dB in the last column of Table 1. These asymptotic values can be approximately verified in all Figures at $P(e) = 10^{-8}$, with an accuracy of about 0.5 dB.

We conclude with a few comments on the possible application of permutation codes. In order to implement the transmitter we need to define an orthonormal basis $\{\psi_j(t)\}_1^n$ of the signal space so that each transmitted signal is given by

$$x(t) = \sum_{j=1}^{n} x_j \psi_j(t) \tag{31}$$

The basis functions must be approximately time- and bandlimited. For example, we

No.	n	m's	μ's	$[r]$	γ_{dB}
1	5	I (4,1)	-1.2908, 5.1631	2	1.6
2	5	II (4,1)	0, 6.6558	3	2.2
3	5	II (3,2)	0, 6.7720	5	1.2
4	5	II (2,3)	0, 6.7720	6	0.2
5	5	II (4,1)	4.6540, 11.4152	7	-1.4
6	5	II (2,2,1)	0, 6.7748, 13.3491	7	-2.0
7	5	II (1,3,1)	0, 6.6712, 13.4039	8	-2.3
8	10	I (9,1)	-0.6691, 6.0220	3	2.7
9	10	II (9,1)	0, 6.8907	4	3.3
10	10	II (7,3)	0, 7.1240	9	2.2
11	10	II (4,5,1)	0, 7.1747, 14.1284	16	-0.6
12	10	II (3,5,2)	0, 7.1293, 14.1828	18	-1.6
13	10	II (3,3,3,1)	0, 7.0814, 14.0512, 21.0215	21	-3.7
14	10	II (2,4,2,2)	0, 7.0672, 14.0365, 20.9950	22	-4.4
15	7	II (6,1)	0, 6.7720	3	2.8
16	7	II (5,2)	0, 6.9169	6	2.0
17	25	II (14,11)	0, 7.6409	33	1.8
18	50	II (25,22,3)	0, 8.0046, 15.7001	83	0.6
19	50	II (20,25,5)	0, 7.9966, 15.7951	92	0.0
20	51	II (19,24,6,2)	0, 8.0255, 15.8067, 23.2291	105	-1.5
21	50	II (17,23,8,2)	0, 7.9926, 15.8064, 23.3280	108	-1.7
22	7	II (4,3)	0, 6.9694	8	1.3
23	31	II (23,8)	0, 7.6873	30	2.9

Table 1: Optimized codes at $P(e) = 10^{-5}$ found by Slepian [12].

can select the strictly bandlimited functions

$$\psi_j(t) = \frac{\sin(2\pi Wt - j\pi)}{2\pi Wt - j\pi} \qquad j = 1, \ldots, n \tag{32}$$

or the strictly time-limited functions for $0 \leq t \leq T$

$$\begin{aligned} \psi_{2j-1}(t) &= \sin(2\pi jt/T) \\ \psi_{2j}(t) &= \cos(2\pi jt/T) \qquad j = 1, \ldots, n/2\ . \end{aligned} \tag{33}$$

Other choices are also possible see for example [11, 4, 5, 3]. The receiver can be implemented using a bank of matched filters, matched to the basis functions in order to obtain the components of the received vector $\mathbf{y}$.

Although permutation codes have a very long history and some very promising coding gains, to the author's knowledge, they have never been used in applications. Nevertheless, we can be expect that they may be exploited in future for high speed transmission, due to their very simple decoding algorithm.

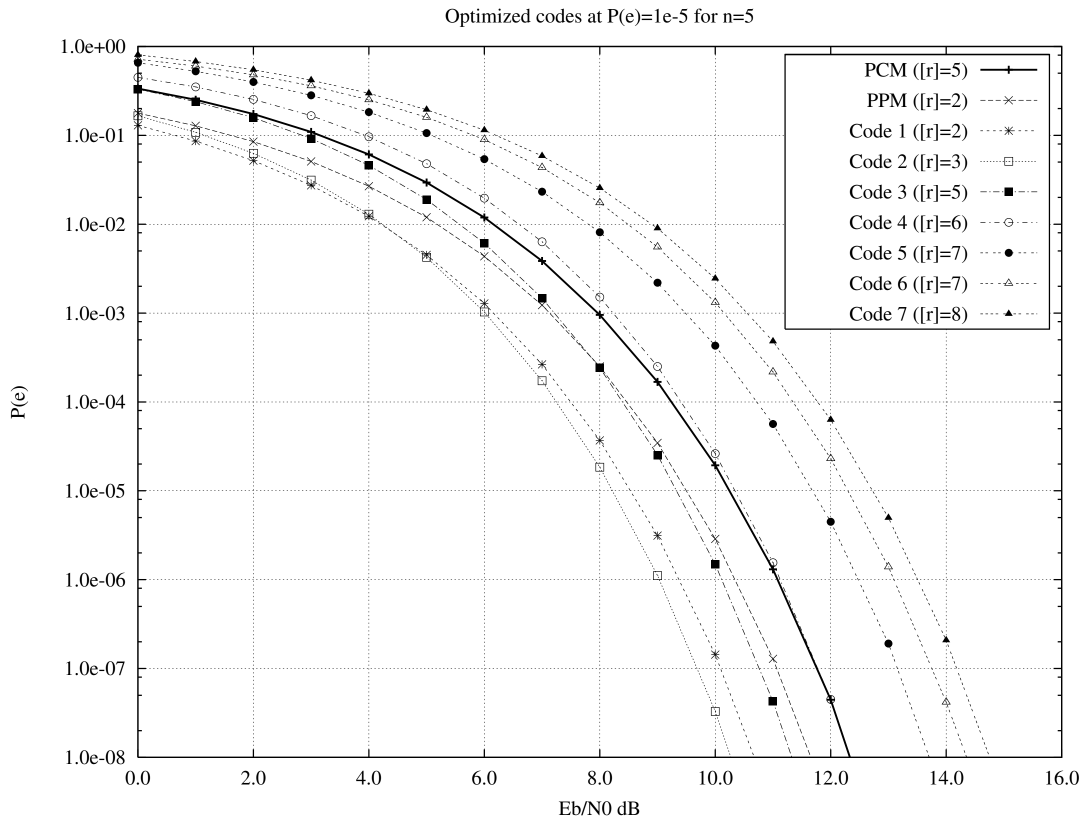

Figure 1: Optimized codes at $P(e) = 10^{-5}$ for $n = 5$.

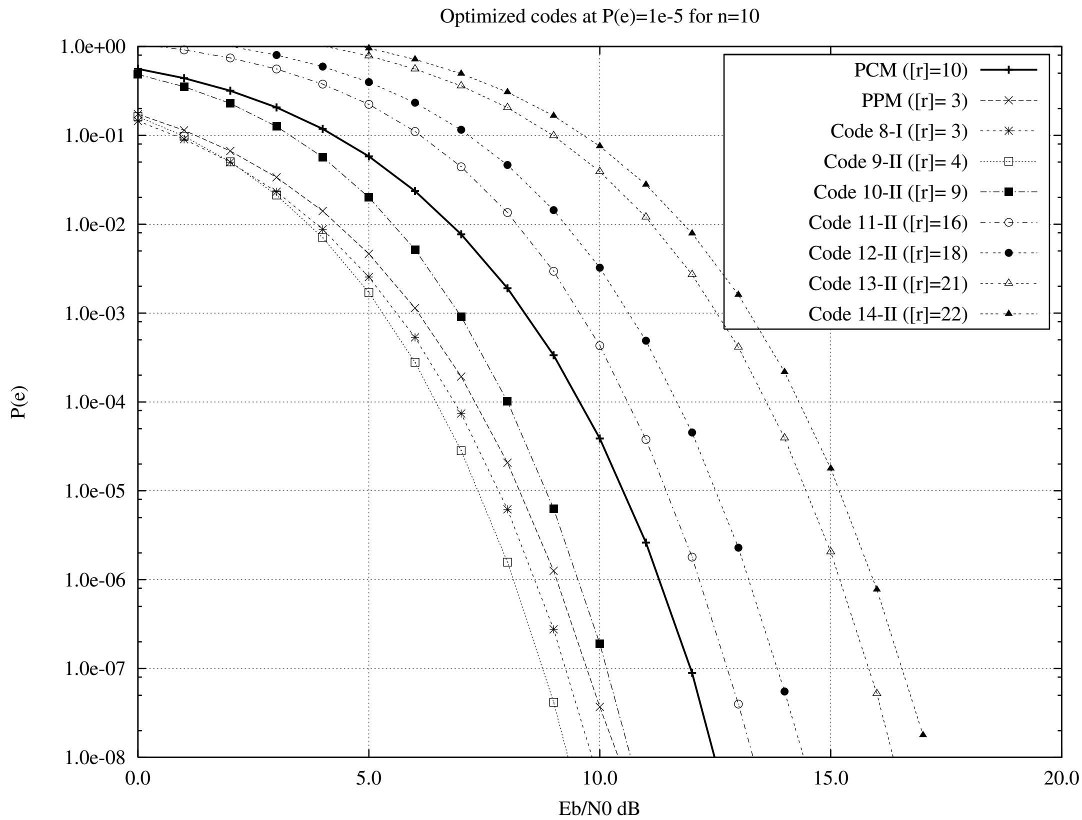

Figure 2: Optimized codes at $P(e) = 10^{-5}$ for $n = 10$.

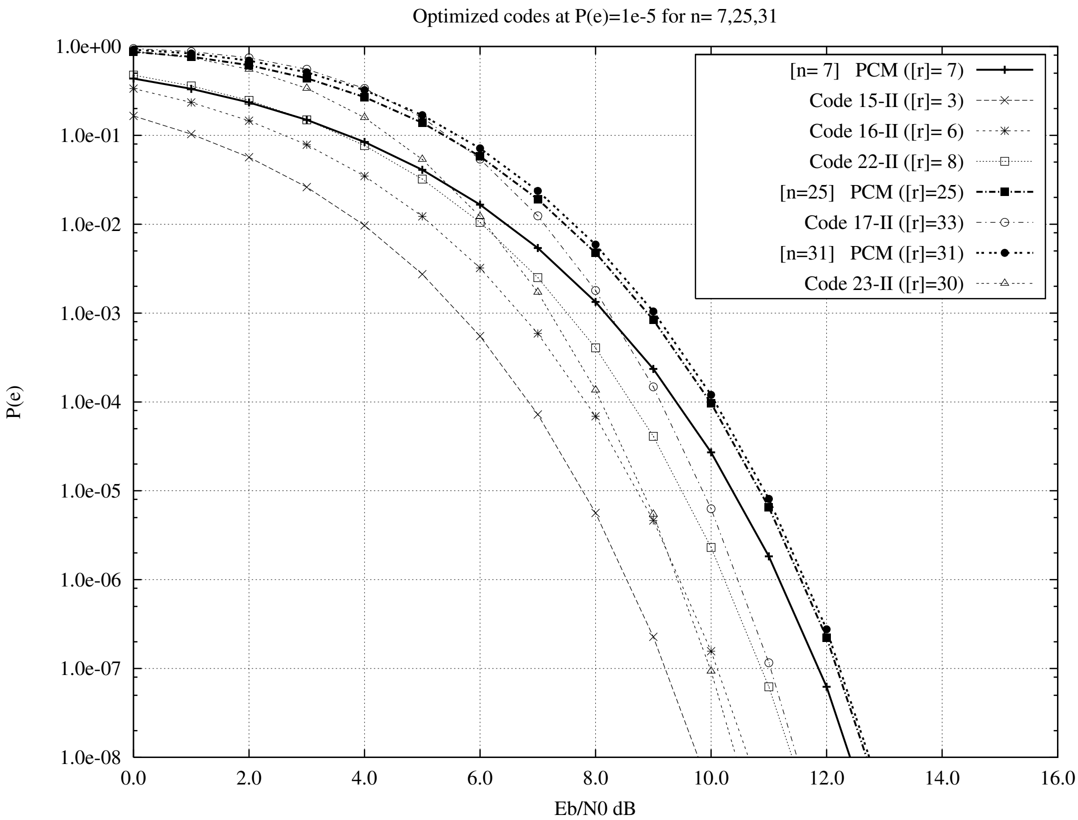

Figure 3: Optimized codes at $P(e) = 10^{-5}$ for $n = 7, 25, 31$.

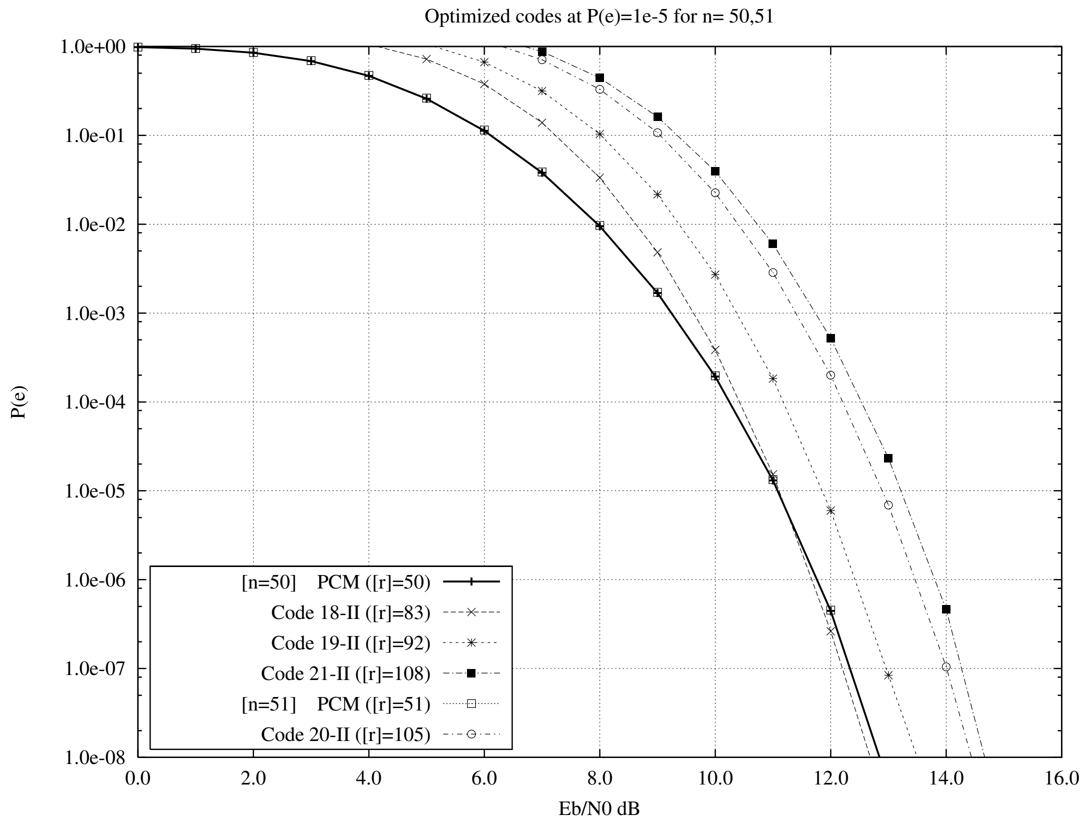

Figure 4: Optimized codes at $P(e) = 10^{-5}$ for $n = 50, 51$.

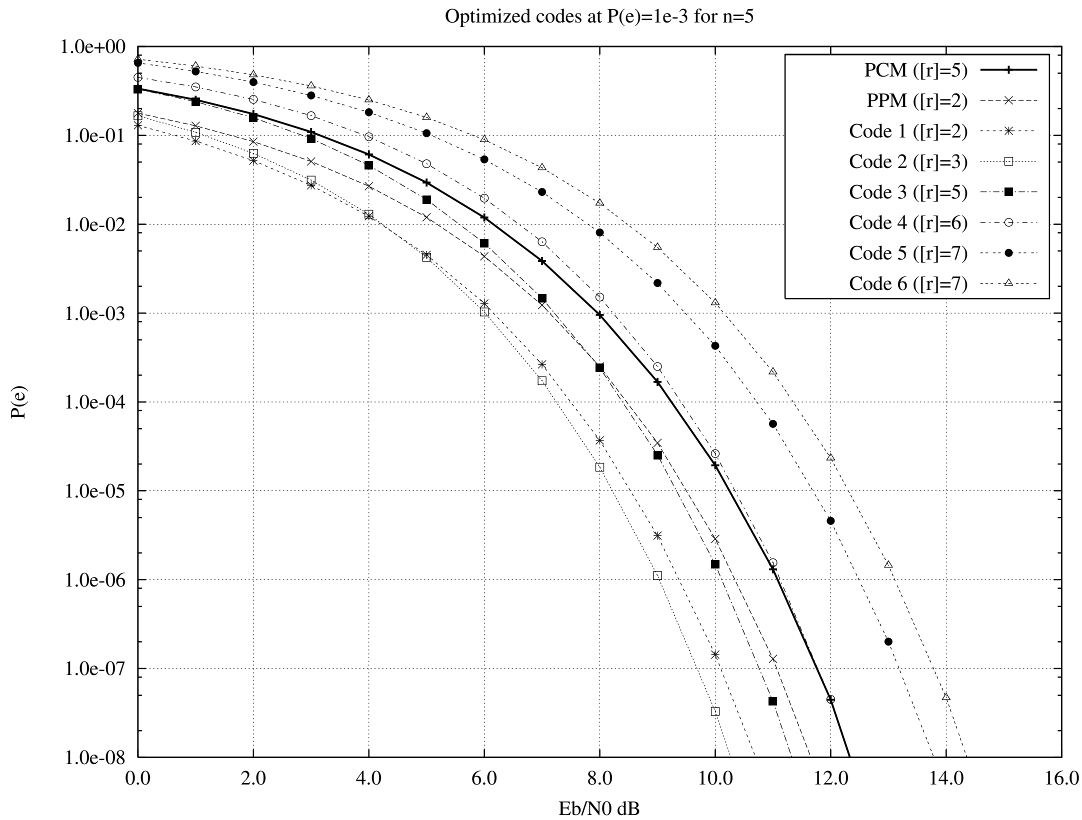

Figure 5: Optimized codes at $P(e) = 10^{-3}$ for $n = 5$.

References

[1] A. V. Balakrishnan, A contribution to the sphere-packing problem of communication theory, *J. Math. Anal. Appl.*, vol. 3, Dec. 1961, pp. 485–506.

[2] E. Biglieri, M. Elia, Optimum permutation modulation codes and their asymptotic performance, *IEEE Transactions on Information Theory*, Nov. 1976, pp. 751–753.

[3] M. Elia, G. Taricco, and E. Viterbo, Optimal Energy Transfer over Bandlimited Communication Channels, *IEEE Transactions on Information Theory*, vol. 45, n. 6, Sept. 1999, pp. 2020–2029.

[4] H.J. Landau and H.O. Pollack, Prolate spheroidal wave functions, Fourier analysis and uncertainty II, *Bell Syst. Tech. J.*, 1961, Vol. 40, pp. 65–84.

[5] H.J. Landau and H.O. Pollack, Prolate spheroidal wave functions, Fourier analysis and uncertainty III, *Bell Syst. Tech. J.*, 1962, Vol. 41, pp. 1295–1336.

[6] J. K. Karlof, Permutation codes for the Gaussian channel, *IEEE Transactions on Information Theory*, n. 4, vol. 35, July 1989, pp. 726–732.

[7] J. K. Karlof, Decoding spherical codes for the Gaussian channel, *IEEE Transactions on Information Theory*, n. 1, vol. 39, Jan. 1993, pp. 60–65.

[8] J. K. Karlof and Y. O. Chang, Optimal permutation codes for the Gaussian channel, *IEEE Transactions on Information Theory*, n. 4, vol. 35, July 1989, pp. 726–732.

[9] I. Ingemarsson, Optimized permutation modulation, *IEEE Transactions on Information Theory*, n. 5, vol. 36, Sept. 1990, pp. 1098–1100.

[10] D. Slepian, Some further theory of group codes, *Bell System Technical Journal*, Sept. 1960, pp. 1219–1252.

[11] D. Slepian and H.O. Pollack, Prolate spheroidal wave functions, Fourier analysis and uncertainty I, *Bell Syst. Tech. J.*, 1961, Vol. 40, pp. 43-64.

[12] D. Slepian, Permutation modulation, *Proceedings of the IEEE*, March 1965, pp. 228–236.

[13] D. Slepian, Group codes for the Gaussian channel, *Bell System Technical Journal*, n. 4, vol. 47, April 1968, pp. 575–602.

[14] D. Slepian, On neighbor distances and symmetry in group codes, *IEEE Transactions on Information Theory*, Sept. 1971, pp. 630–632.

[15] M. Steiner, The strong simplex conjecture is false, *IEEE Transactions on Information Theory*, May 1994, pp. 721–731.