Mapping Design for General Multidimensional Communication Systems [*]

Nan Guo and Laurence B. Milstein [†]
Department of Electrical and Computer Engineering
University of California, San Diego, La Jolla, CA 92093
E-mail: nguo@ece.ucsd.edu, milstein@ece.ucsd.edu
Phone: (619) 534-3096, Fax: (619) 534-2486

Abstract **In communication systems, the mapping between the data space and the signal constellation needs to be carefully designed. The mapping based on the "Gray" code (or the Karnaugh map) is a well known scheme for constellations with uniformly-distributed points. However, for some systems with an irregular constellation, like the permutation modulation system and the parallel combinatory DS/SS commnication system which, essentially, are multidimensional systems, the common method of mapping is invalid. In this paper, we propose a practical algorithm to find a suboptimal mapping for a general constellation. For a constellation with M points, there are a total of $M!$ different mappings. To achieve computational efficiency, the algorithm is designed based on an incomplete search. A comparison of several mappings shows that very good mappings can be found by this systematic method.**

1 Introduction

In communication systems, the mapping between the data space and the signal constellation needs to be carefully designed to achieve low bit error rate (BER). For some minimum-distance-based detectors, the mapping based on the "Gray" code (or the Karnaugh map) [1] is a well known scheme for one- and two-dimensional constellations with uniformly-distributed points within a plane (like QAM) or, in particular, on a circle (like MPSK). Actually, for H×MPSK [2] and H×MQAM [3] (where H is a positive integer referring to multiplicity), the use of the Gray code is still effective. For example, a constellation of 2×MQAM consists of two MQAM complex planes; the input data stream is converted into two parallel substreams, corresponding to the two complex planes; for each complex plane and corresponding substream, the mapping can be designed based on the Gray code, independently of the other complex plane. Note that the above designs are such that any pair of points separated by the minimum distance correspond to the minimum Hamming distance (i.e., one bit difference). To guarantee that the use of the Gray-code-based mapping is good, we need, at least, (1) that the higher-dimensional constellation can be divided into one- or two-dimensional uniform subconstellations, each having 2^n points for some integer n, and (2) that each subconstellation has its own input data substream, whereby the assignment of its points only depends on its own input data. However, for any system with an irregular constellation, such as those mentioned in [4]-[8], the above method of mapping is invalid, because the two basic conditions cannot be satisfied simultaneously. In this paper, we will develop a practical algorithm to find good mappings for general constellations.

Suppose M is the number of points in the signal constellation, $m = log_2 M$ is number of bits per symbol, and the duration of a symbol is T. Our analysis is based on a low-pass complex equivalent model (shown in Fig. 1) of a general multidimensional communication system, where, for a given time interval, say $[0, T), D_i, 1 \le i \le M$, is the ith symbol at the transmitter, and $D_{\hat{i}}, 1 \le \hat{i} \le M$ is the corresponding estimated symbol at the receiver.

$$\phi(t) = (\phi_1(t), \phi_2(t), \phi_3(t), \cdots, \phi_H(t)) \tag{1}$$

is a vector of H orthogonal functions, and for any $h, h' = 1, 2, 3, \cdots, H$, $\phi_h(t)$ and $\phi_{h'}(t)$ satisfy

$$\frac{1}{T}\int_0^T \phi_h(t)\phi_{h'}^*(t)dt = \begin{cases} 1, & h = h' \\ 0, & h \ne h' \end{cases}. \tag{2}$$

An $M \times H$ matrix

$$Q = \begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_M \end{pmatrix} \tag{3}$$

is defined as a mapping code with codewords

$$Q_i = (q_{i1}, q_{i2}, q_{i3}, \cdots, q_{iH}), 1 \le i \le M, \tag{4}$$

and code elements

$$q_{ih} = \alpha_{ih} exp(j\Omega_{ih}), \tag{5}$$

---

[*] Acknowledgement: This work was partially supported by both TRW and Rockwell Corporation.

[†] corresponding author

where $\alpha_{ih} \geq 0$. The M codewords one-to-one correspond to the M symbols through the "mapping code encoder", which equivalently maps the data to the constellation. In this paper,

$$E_i[Q_i(Q_i^*)^T] = 1 \tag{6}$$

is assumed, where $E_i[\ \ ]$ is the expectation over D_i (or i), $i = 1, 2, \cdots, M$; and the distance between any pair of signal points for D_i and $D_{i'}$ is given by

$$\begin{aligned} d_0(i, i') &= \sqrt{\sum_{h=1}^{H} (q_{ih} - q_{i'h})(q_{ih} - q_{i'h})^*} \\ &= \sqrt{\Delta Q_{i,i'} (\Delta Q_{i,i'}^*)^T}, \end{aligned} \tag{7}$$

where

$$\Delta Q_{i,i'} = Q_i - Q_{i'}. \tag{8}$$

With the assumption of (6), the average power of transmission signal is P (see Fig. 1). By introducing the mapping code, our system becomes sufficiently general so that any case in [4], [5], [6] and [7] can be considered a special case. As shown in some previous literature, constant weight codes (CWC) are very useful [5][6], and we say that a mapping code has weight w if every codeword in this mapping code includes exactly w non-zero elements.

2 Mapping design

In order to introduce the algorithm, consider the following. Since the smaller Euclidean distances play a more important role than the larger ones, we want the small Hamming distances to correspond to the small Euclidean distances. In addition, for a constellation with M points, there are a total of $M!$ different mappings. Thus, the computational efficiency of the algorithm is very important. To make the algorithm computationally efficient, we turn to an incomplete search, outputting the mapping step-by-step, which may not result in optimal mappings, but mappings which nevertheless perform well.

For simplicity, we use the terms "symbol" and "point" to stand for a data symbol and a signal point in the constellation, respectively. A symbol can be viewed as an m-bit binary number with value between 0 and $M-1$, and we label those M symbols $1, 2, 3, \cdots, M$. Therefore, for any k, $k = 1, 2, 3, \cdots, M/2$, the Hamming distance between symbol k and symbol $k + M/2$ is always equal to 1. To label the points in the constellation, first partition the constellation into two subconstellations with $M/2$ points in each one, such that the minimum intra-subconstellation distance is maximum (if partitioning cannot increase the minimum intra-subconstellation distance, try to make the occurence of the minimum intra-subconstellation distance in each subconstellation as rare as possible); then, label the points $1, 2, 3, \cdots, M/2$ in one subconstellation, and label the points $M/2+1, M/2+2, M/2+3, \cdots, M$ in the other subconstellation, such that for any given $l, l = 1, 2, 3, \cdots, M/2$, point $l + M/2$ is nearest (in Euclidean distance) to point l. Note that this labelling may not be unique. For example, two possible labellings for the same bipolar CWC (BCWC) are shown in Table 1, where $M = 16$, $q_{ih} \in \{0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\}$ (from (6)), and "+" and "-" refer to $\frac{1}{\sqrt{2}}$ and $\frac{-1}{\sqrt{2}}$, respectively; for each of the two labellings, the eight pairs of points (1,9), (2,10), $\cdots$, (8,16) have the smallest distance (i.e., 1) between the two points in the same pair.

Suppose $\mathcal{S}$ is a set of the pairs of points, and $\mathcal{S}^*$ and $\mathcal{D}$ are a vector of the pairs of points and a vector of pairs of symbols, respectively. Define a reversed pair of points $\bar{s} = (y, x)$, if the original pair of points is $s = (x, y)$. Also, define the operator "$\oplus$" to refer to augmenting the size of a vector by adding an element to the right-hand side of that vector. For example, if $\mathcal{S}^* = (s_1, s_2)$, then $\mathcal{S}^* \oplus s_3 = (s_1, s_2, s_3)$. Let $MATCH(X, Y)$ be a metric that can indicate how well the pairwise symbol vector X matches with the pairwise point vector Y [9].

Algorithm

```
$Initialization:$
$\mathcal{S} = \{\ (l, l+M/2) \mid l = 2, 3, 4, \cdots, M/2\ \};$
$\mathcal{S}^* = (s_1),\ \ s_1 = (1, 1+M/2);$
$\mathcal{D} = (\ (1, 1+M/2)\ );$

$for\ \ k = 2\ \ to\ \ M/2$
    $for\ \ \forall\ s \in \mathcal{S}\ \ compute$
        $MATCH(\mathcal{D} \oplus (k, k+M/2), \mathcal{S}^* \oplus s)$
        $and\ \ MATCH(\mathcal{D} \oplus (k, k+M/2), \mathcal{S}^* \oplus \bar{s});$
        $if\ \ MATCH(\mathcal{D} \oplus (k, k+M/2), \mathcal{S}^* \oplus s^*),$
            $s^* \in \mathcal{S},\ is\ the\ maximum$
            $then\ \ s_k = s^*;$
                $\mathcal{S}^* \Leftarrow \mathcal{S}^* \oplus s_k;$
                $\mathcal{S} \Leftarrow \mathcal{S} - \{s_k, \bar{s}_k\};$
    $end$
    $\mathcal{D} \Leftarrow \mathcal{D} \oplus (k, k+M/2);$
$end$
```

Each k-loop operation corresponds to an incomplete search, and sets a correspondence between a pair of symbols and a pair of points. After k incomplete searches, we obtain $\mathcal{S}^* = (s_1, s_2, \cdots, s_{M/2})$, and the relation between $\mathcal{D}$ and $\mathcal{S}^*$ gives a suboptimal mapping, i.e., the pair of symbols $(k, k+M/2)$ corresponds to the pair of points s_k, $k = 1, 2, 3, \cdots, M/2$. A practical version of the algorithm is provided in [9]. Note that the above labelling and algorithm together guarantee that the $M/2$ smallest-Hamming-distance pairs of symbols match the first $M/2$ smallest-Euclidean-distance pairs of points [i.e., a symbol pair $(k, k+M/2)$ with the smallest Hamming distance, 1, always corresponds to a point pair $(l, l+M/2)$ or $(l+M/2, l)$, where $k,\ l \in \{1, 2, 3, \cdots, M/2\}$]. For instance, the mapping given in Table 2 shows that the eight pairs of symbols (1,9), (2,10), $\cdots$, (8,16) match the eight pairs of points (1,9), (11,3), $\cdots$, (12,4), respectively.

3 Results and Conclusion

Let us consider two examples.
System 1: The mapping code is a BCWC with $H = 4$, $w = 2$, $M = 16$ and $q_{ih} \in \{0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\}$. The dimensionality of the constellation is 4, and a 4-dimensional (4-D) point in the constellation can be represented by a vector of four elements, e.g., $(0, 0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$, etc. One can check that the set of Euclidean distances is $\{1,\ \sqrt{2},\ \sqrt{3},\ 2\}$, by using Equation (7), under the conditions of (6) and an equally likely data source. Two specific labellings for the same mapping code are shown in Table 1. The suboptimal mapping based on the labelling given in Table 1(a) is shown in Table 2, where the "k's"refer to the labels of the symbols, the " $map(k)$'s" refer to the labels of the 4-D points, and the "$map(k)$'s" are obtained by running the algorithm.
System 2: $H = w = 1$ and $M = 16$. This is the well-known two-dimensional constellation of 16-QAM shown in Fig. 2, where the Gray code mapping is given in [10, P.209]. From Equation (7), under the conditions of (6) and an equally likely data source, we find that the set of Euclidean distances is $\{0.6325, 0.8944, 1.2649, 1.4142, 1.7889, 1.8947, 2.0000, 2.2804\}$. The labelled points and the corresponding symbols provided by our algorithm are illustrated in Fig. 3, where the white points represent one subconstellation and are labelled $1, 2, 3, \cdots, M/2$, and the black points represent another subconstellation and are labelled $M/2+1, M/2+2, M/2+3, \cdots, M$.

BER performances are given in Fig. 4 and Fig. 5 for *Systems* 1 and 2, respectively. In Fig. 4, the upper and lower bounds are equal to P_s and P_s/m, respectively, where P_s is the symbol error rate and m is the number of bits embeded in each symbol; the average BER is obtained by averaging over all of $M!$ mappings [11]; and M1 and M2 refer to two randomly chosen mappings for *System* 1. In Fig. 5, we compare the results of using both the Gray code and our algorithm for *System* 2. It can be seen that the two BER curves overlap, so that for this regular constellation, the proposed mapping technique can yield as good a result as does the Gray-code mapping.

Although the labelling and mapping are typically not unique, our study based on a large number of observations shows that the proposed method works very well.

From the above results, we can conclude that: (1) the proposed mapping scheme is a systematic and computer-based method; (2) it is computationally efficient; the computational load approximates $\frac{M^2}{8}$ for a constellation with M signal points; (3) it can be applied to any kind of constellation; and (4) very good mappings—even the optimal result—can be obtained.

References

[1] Edwin Jones, *Digital Transmission*, McGraw-Hill, UK, 1993.

[2] S. S. Pietrobon, R. B. Deng, A. Lafanechere, *et al*, "Trellis-coded multidimensional phase modulation," *IEEE Trans. on Inform. Theory*, vol.36, No.1, pp.63-89, 1990.

[3] S. S. Pietrobon and D. J. Costello, Jr., "Trellis coding with multidimensional QAM signal sets," *IEEE Trans. on Inform. Theory*, vol.39, No.2, pp.325-336, 1993.

[4] D. Slepian, "Permutation modulation," *Proc. IEEE*, vol.53, pp.228-236, 1965.

[5] J. F. Pieper, J. G. Proakis, R. R. Reed, and J. K. Wolf, "Design of efficient coding and modulation for a Rayleigh Fading Channel," *IEEE Trans. on Inform. Theory*, vol.IT-24, No.4, pp.457-468, 1978.

[6] G. E. Atkin and H. P. Corrales, "An efficient modulation/coding scheme for MFSK systems on bandwidth constrained channels," *IEEE Journal on Selected Areas in Commun.*, vol.7, No.9, pp.1396-1401, 1989.

[7] S. Sasaki, H. Kikuchi, H. Watanabe and J. Zhu, "Performance evaluation of parallel combinatory SSMA systems in Rayleigh fading channel," Proceedings of IEEE 3rd International Symposium on Spread Spectrum Techniques and Applications (ISSSTA'94), Oulu, Finland, 4-6 July 1994, vol.1, pp.198-202.

[8] N. Guo and L. B. Milstein, "On rate-variable multidimensional DS/SSMA with dynamic sequence sharing," *IEEE Journal on Selected Areas in Commun.*, vol.17, No.5, May 1999.

[9] N. Guo and L. B. Milstein, "ML detector with imperfect CSI for multidimensional communication over Rayleigh fading channel," Submitted to *IEEE Trans. on Vehicular Technology.*

[10] R. E. Ziemer and R. L. Peterson, *Digital Communications and Spread Spectrum Systems*, Macmillan, New York, 1985.

[11] N. Guo and L. B. Milstein, Performance Evaluation of an ARQ System Based on Multi-Code DS/CDMA. To be submitted.

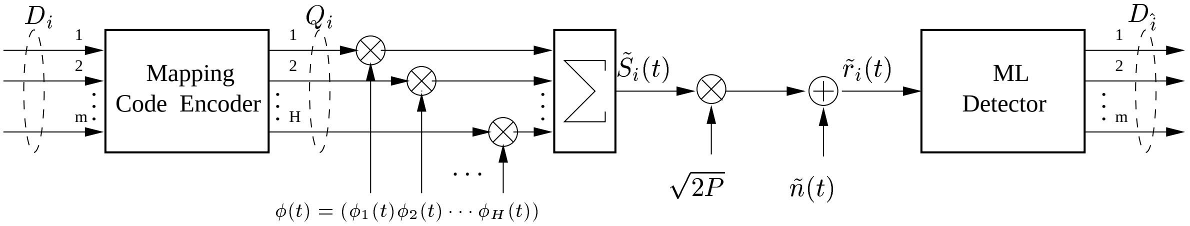

Figure 1: Low-pass complex equivalent model of multidimensional communication system

Table 1. Two examples of labelling for a 16-point constellation

(a)

labels	vectors	labels	vectors	labels	vectors	labels	vectors
1	$+ + 0\ 0$	5	$0\ 0 + +$	9	$0 + + 0$	13	$+ 0\ 0 +$
2	$+ - 0\ 0$	6	$0\ 0 + -$	10	$0 - + 0$	14	$+ 0\ 0 -$
3	$- + 0\ 0$	7	$0\ 0 - +$	11	$0 + - 0$	15	$- 0\ 0 +$
4	$- - 0\ 0$	8	$0\ 0 - -$	12	$0 - - 0$	16	$- 0\ 0 -$

(b)

labels	vectors	labels	vectors	labels	vectors	labels	vectors
1	$+ + 0\ 0$	5	$0 + + 0$	9	$+ 0\ 0 +$	13	$0\ 0 + +$
2	$+ - 0\ 0$	6	$0 - + 0$	10	$+ 0\ 0 -$	14	$0\ 0 + -$
3	$- + 0\ 0$	7	$0 + - 0$	11	$- 0\ 0 +$	15	$0\ 0 - +$
4	$- - 0\ 0$	8	$0 - - 0$	12	$- 0\ 0 -$	16	$0\ 0 - -$

Table 2. Suboptimal mapping of *System* 1 ($d_{E0} = 1$, $d_{ET} = 1.5$, $d_{HT} = 3$, with labelling given in Table 1(a))

symbols k	points $map(k)$	symbols k	points $map(k)$	symbols k	points $map(k)$	symbols k	points $map(k)$
1	1	5	14	9	9	13	6
2	11	6	8	10	3	14	16
3	13	7	2	11	5	15	10
4	7	8	12	12	15	16	4

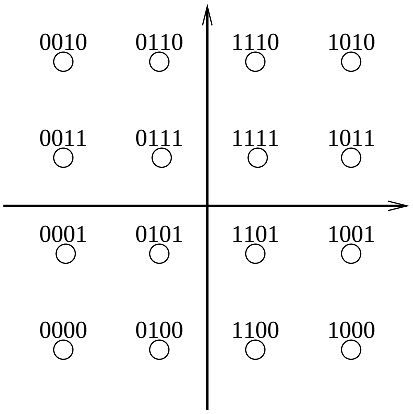

Figure 2: 16-QASK constellation and the Gray code mapping

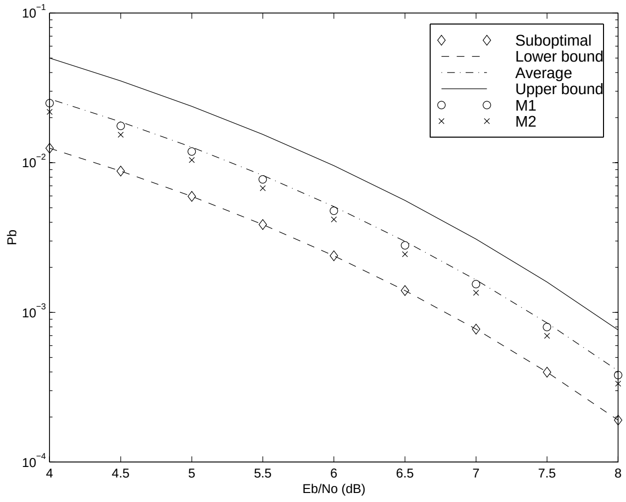

Figure 4: P_b vs. E_b/N_0 for the mappings of $System$ 1

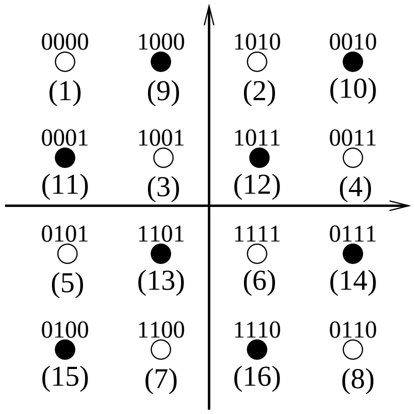

Figure 3: Labelled points and the corresponding data symbols resulting from the algorithm for the 16-QAM constellation

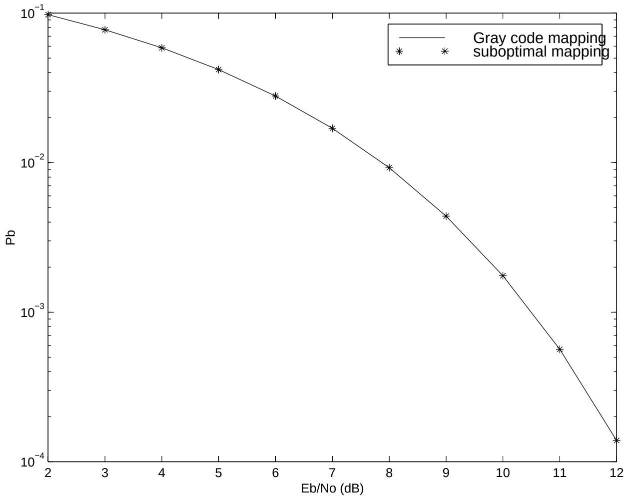

Figure 5: P_b vs. E_b/N_0, resulting from using both the Gray code and the algorithm for $System$ 2