PSK bit mappings with good minimax error probability

Erik Agrell
Department of Signals and Systems
Chalmers University of Technology
41296 Göteborg, Sweden
Email: agrell@chalmers.se

Erik G. Ström
Department of Signals and Systems
Chalmers University of Technology
41296 Göteborg, Sweden
Email: erik.strom@chalmers.se

***Abstract*— Labelings of multilevel PSK constellations are analyzed with respect to the error probability for individual bit positions, when symbols are transmitted over additive noise channels. When the maximum bit error probability is minimized instead of the average, the best labeling is in general not a Gray code. We develop a new class of labelings by modifying balanced Gray codes into better, but non-Gray, labelings.**

I. Introduction

It is standard engineering practice to label multilevel phase-shift keying (PSK), pulse-amplitude modulation (PAM), and rectangular quadrature-amplitude modulation (QAM) signal constellations with a Gray code, usually the binary reflected Gray Code (BRGC). The rationale is that maximum likelihood symbol detection for additive white Gaussian noise (AWGN) channels is a minimum distance decision rule, and thus the most likely symbol error is to a neighboring signal point, which will result in only one bit error when a Gray code is used. If we accept this argument, we may be led to the false conclusion that all Gray codes are equally good. However, the number of Gray codes that give different average bit error probabilities increases rapidly with constellation size M (or, equivalently, with $m = \log_2 M$, the number of bits per symbol) [1] [3, p. 15]. The natural question of which Gray code is optimum with respect to the average bit error probability has only recently received a partial answer. In [4], the BRGC was shown to be optimal for AWGN channels when the signal-to-noise ratio (SNR) is larger than a finite threshold for PSK, PAM, and rectangular QAM constellations.

In general, bits that are transmitted at different bit positions in the label will experience unequal bit error probabilities. We will consider an optimality criterion based on minimizing the maximum error probability for any bit position (see Section III for details), and it will be shown that the BRGC is no longer optimal. Other Gray codes are better, and for most m, a non-Gray labeling is shown to be better than all Gray codes. The minimax optimality criterion is useful when we want to guarantee that the bit error probability for all positions in the label does not exceed a given threshold. Hence, the criterion is quite reasonable, perhaps more reasonable than the more common average error probability criterion.

In the literature on this topic, it is usually assumed that the transmitted bits are equally likely and statistically independent, and that the receiver finds the bit decisions from the label of a maximum likelihood (ML) decision on the transmitted symbol. We will make the same assumption in this paper.

TABLE I
The binary reflected Gray code (BRGC), the natural binary code (NBC), and the 2-minimax optimal labeling (2-MML) for $m = 3$.

BRGC	**NBC**	**2-MML**
0 0 0	0 0 0	0 0 0
0 0 1	0 0 1	0 0 1
0 1 1	0 1 0	0 1 0
0 1 0	0 1 1	0 1 1
1 1 0	1 0 0	1 1 1
1 1 1	1 0 1	1 0 1
1 0 1	1 1 0	1 1 0
1 0 0	1 1 1	1 0 0

II. Definitions and Notation

A *binary labeling* of order $m \in \mathbb{Z}^+$ is a sequence of $M = 2^m$ distinct labels (or codewords), $\lambda = (\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{M-1})$, where the labels are represented as binary row vectors $\mathbf{c}_l \in \{0,1\}^m$. The ith component (counted from the right) of $\mathbf{c}$ is denoted $[\mathbf{c}]_i$. For notational convenience, we will consider $\mathbf{c}_l$ to be periodic in l with period M, i.e., $\mathbf{c}_l = \mathbf{c}_{l+M}$ for all l. The well-known BRGC and the natural binary code (NBC) for $m = 3$ are listed in Table I, along with a new labeling to be discussed later.[1]

A *cyclic binary Gray code* is a labeling $(\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{M-1})$ for which $\mathbf{c}_l$ and $\mathbf{c}_{l+1}$ for $l = 0, 1, \ldots, M-1$ differ in one bit position only. From now on, we will say simply Gray codes when meaning cyclic binary Gray codes.

We will assume, without loss of generality, that the first label $\mathbf{c}_0$ of any labeling is the all-zero label. With this convention, a labeling $(\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{M-1})$ of order m is defined from its *transition sequence* $s = (s_0, s_1, \ldots, s_{M-1})$, where $s_k = \{i : [\mathbf{c}_k]_i \neq [\mathbf{c}_{k+1}]_i\}$. For example, the transition sequence of the NBC in Table I is

$$(\{0\}, \{0,1\}, \{0\}, \{0,1,2\}, \{0\}, \{0,1\}, \{0\}, \{0,1,2\}).$$

[1] We distinguish a labeling from a *code*, which is an (unordered) set $\{\mathbf{c}_0, \ldots, \mathbf{c}_{M-1}\}$. Hence, a Gray code is a labeling and not a code.

For compactness, we will from now on omit the braces around sets with a single element.

The *k-spaced transition count* for the ith bit position of a labeling $(\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{M-1})$ of order m is defined for all integers k as

$$e_k(i) \triangleq \sum_{l=0}^{M-1} [\mathbf{c}_l]_i \oplus [\mathbf{c}_{l+k}]_i, \qquad i = 0, 1, \ldots, m-1,$$

where the operator $\oplus$ is defined such that $x \oplus y$ equals the integer 0 if the binary digits x and y are equal and 1 otherwise. Note that the commonly used term *transition count* for the ith bit position, see, e.g., [2], is the 1-spaced transition count in our terminology, i.e., $e_1(i)$.

Now, since $e_k(i)$ is an even and periodic sequence in k with period M and $e_k(i) \in \mathcal{F}_M$, where $\mathcal{F}_M \triangleq \{0, 2, 4, \ldots, M\}$, we can summarize all relevant values in the *transition count vector*

$$\mathbf{e}_i = \begin{bmatrix} e_1(i) & e_2(i) & \cdots & e_{M/2}(i) \end{bmatrix} \in \mathcal{F}_M^{M/2}.$$

The *transition count matrix* $\mathbf{E} \in \mathcal{F}_M^{m \times (M/2)}$ is defined as

$$\mathbf{E} = \begin{bmatrix} \mathbf{e}_0 \\ \cdots \\ \mathbf{e}_{m-1} \end{bmatrix}.$$

As an example, the transition count matrix for the NBC labeling in Table I is

$$\begin{bmatrix} 8 & 0 & 8 & 0 \\ 4 & 8 & 4 & 0 \\ 2 & 4 & 6 & 8 \end{bmatrix}. \tag{1}$$

The *f-lexicographic value* of a vector $\mathbf{v} \in \mathcal{F}_M^N$ is defined for all integers $1 \le f \le N$ as

$$\ell_f(\mathbf{v}) = \sum_{k=0}^{f-1} [\mathbf{v}]_{N-f+k} (M+1)^k.$$

For convenience, we will sometimes write $\ell_N(\mathbf{v})$ as $\ell(\mathbf{v})$, where $\mathbf{v} \in \mathcal{F}_M^N$. We note that the vectors $\mathbf{x}$ and $\mathbf{y}$ are in lexicographic order if $\ell(\mathbf{x}) \le \ell(\mathbf{y})$.

Let $\mathbf{E}$ be a transition count matrix for a labeling of order m. By $\max \mathbf{E}$ and $\min \mathbf{E}$, we denote the rows of $\mathbf{E}$ with the largest and smallest lexicographic value, i.e., $\ell(\max \mathbf{E}) \ge \ell(\mathbf{e}_i) \ge \ell(\min \mathbf{E})$ for $i = 0, 1, \ldots, m-1$.

A labeling of order m is said to be *totally balanced* if its 1-spaced transition count is equal for all bit positions, i.e., $e_1(i) = e_1(j)$ for $0 \le i < j \le m-1$. For a Gray code λ, we have $\sum_{i=0}^{m-1} e_1(i) = 2^m$, and λ can therefore be totally balanced only if m is a power of 2. For all other m, since $e_1(i)$ even, the most uniform distribution of the 1-spaced transition counts we can hope for is when $|e_1(i) - e_1(j)| \le 2$ for $0 \le i < j \le m-1$. If this condition is fulfilled, λ is said to be *balanced*. It has been shown that there exist balanced Gray codes for all $m \ge 1$, see [2] and [3, pp. 14-15][2] and that totally balanced Gray codes exist for all $m = 2^r$ for integers $r \ge 0$ [7].

[2]The perhaps earliest proof of the existence of balanced Gray codes for all m is attributed to T. Bakos [6].

III. Computing the Bit Error Probability

Suppose a labeling with k-spaced transition count $e_k(i)$ and order m is used to label an M-ary signal constellation. The bit error probability for the ith bit is defined as

$$P_b(i) \triangleq \Pr\{[\mathbf{c}_{\mathrm{ML}}]_i \ne [\mathbf{c}]_i\},$$

where $\mathbf{c}$ is the transmitted label and $\mathbf{c}_{\mathrm{ML}}$ is the label corresponding to the maximum likelihood decision on the transmitted symbol. The average bit error probability is defined as

$$P_b \triangleq \frac{1}{m} \sum_{i=0}^{m-1} P_b(i),$$

the maximum bit error probability as

$$\hat{P}_b \triangleq \max_{i \in \{0,\ldots,m-1\}} P_b(i),$$

and the minimum bit error probability as

$$\check{P}_b \triangleq \min_{i \in \{0,\ldots,m-1\}} P_b(i).$$

A binary labeling is said to *optimal in the minimax sense* if it has the smallest $\hat{P}_b$ of all possible labelings λ of the same order. This optimality criterion makes sense when a certain quality of the transmission for the bits in all positions should be guaranteed.

The average bit error probability for M-ary PSK over additive noise channels (with rotationally invariant noise distributions) can be written as [1]

$$P_b = \frac{1}{m} \sum_{k=1}^{M-1} \bar{d}(k) P(k),$$

where $\bar{d}(k)$ is the *average distance spectrum*, defined as

$$\bar{d}(k) \triangleq \frac{1}{M} \sum_{i=0}^{m-1} e_k(i).$$

and $P(k)$ is the *crossover probability*, defined as the probability that the ML symbol decision is k steps clockwise away from the transmitted symbol along the PSK circle. It follows from the channel assumptions above that $P(k) = P(M-k)$ for all k. Clearly, we can write

$$P_b = \frac{1}{m} \sum_{i=0}^{m-1} \frac{1}{M} \sum_{k=1}^{M-1} e_k(i) P(k)$$

and it can easily be shown that the bit error probability for the ith bit is

$$P_b(i) = \frac{1}{M} \sum_{k=1}^{M-1} e_k(i) P(k) = \frac{1}{M} \mathbf{e}_i \mathbf{p}^T, \tag{2}$$

where

$$\mathbf{p} \triangleq \begin{bmatrix} 2P(1) & 2P(2) & \cdots & 2P(M/2-1) & P(M/2) \end{bmatrix}.$$

In general, there is no labeling that is optimal for all channels. However, for many noise distributions at high SNR's, $P(k)$ decreases fast with k for small values of k. We will therefore approach the problem by considering such channels; to be precise, we will consider f*-decreasing* channels for $1 \leq f \leq M/2-1$, which are channels such that

$$P(1) \gg P(2) \gg \cdots \gg P(f) \gg \frac{1}{2}\sum_{k=f+1}^{M-f-1} P(k),$$

where $\gg$ means larger by at least a factor $M+1$. We note that the standard AWGN channel is $M/4$-decreasing for sufficiently large SNR's.

It can be shown that for an f-decreasing channel, the condition $\ell_f(\mathbf{e}') \leq \ell_f(\mathbf{e})$ is necessary for $\mathbf{e}'\mathbf{p}^T \leq \mathbf{e}\mathbf{p}^T$, for any two transition count vectors $\mathbf{e}'$ and $\mathbf{e}$. Hence, the labeling of order m with minimum $\hat{P}_b$ must have a transition count matrix $\mathbf{E}'$ such that

$$\ell_f(\max \mathbf{E}') \leq \ell_f(\max \mathbf{E}), \tag{3}$$

where $\mathbf{E}$ is the transition count matrix for any labeling of order m. We call a labeling whose transition count matrix $\mathbf{E}'$ satisfies (3) an f*-minimax optimal* labeling (f-MML). It is reasonable to search for labelings that are optimal in the minimax sense by first finding the set of 1-MML's and then finding the subset of the 1-MML's that are 2-MML's, 3-MML's, and so on. In fact, a similar approach has been used to prove that the BRGC gives the smallest possible average bit error probability for AWGN channels when the SNR is greater than a finite threshold [4].

Intuitively, a good labeling in the minimax sense should (a) be as balanced as possible, meaning that $|e_k(i)-e_k(j)|$ should be small for all $0 \leq i < j \leq m-1$ and $k = 1, 2, \ldots, M/2$, and (b) have a slowly growing k-step transition count, i.e., $\ell(\max \mathbf{E})$ should be as small as possible. Many labelings that are "balanced," in some sense, can be found in the literature, see [2], [5] and references therein.

IV. Properties of MML's

In the following, $\lceil x \rceil$ and $\lfloor x \rceil$ denote the smallest integer not smaller than x and the closest integer to x, respectively. Ties can be broken arbitrarily in the latter case. In addition, we use the shorthand notation $\lceil x \rceil_i$ for $i\lceil x/i \rceil$ and similarly for $\lfloor x \rceil_i$. For example, $\lfloor x \rceil_2$ is the closest even integer to x.

Theorem 1: A labeling of order $m \geq 1$ is a 1-MML if and only if its 1-spaced transition count satisfies

$$\max_i e_1(i) = \left\lceil \frac{M}{m} \right\rceil_2. \tag{4}$$

Proof: It is well known that the transition count of balanced Gray codes satisfies (4). To show that no other labeling of the same order can do better, consider the transition count $e_1(i)$ for an arbitrary labeling of order m. The number $\sum_i e_1(i)$ is equal to the total number of elements in the M sets that make up the transition sequence. Hence, $\sum_i e_1(i) \geq M$, with equality if the labeling is a Gray code. At least one $e_1(i)$ for $i = 0, 1, \ldots, m-1$, say $e_1(j)$, must be larger than or equal to the average $m^{-1}\sum_i e_1(i) \geq M/m$. Now, since $e_1(j)$ is even, it follows that $\max_i e_1(i) \geq e_1(j) \geq \lceil M/m \rceil_2$, and the theorem follows. ∎

As we will see, there exist balanced labelings that are not Gray labelings.

A Gray code satisfies $e_2(i) = 2e_1(i)$ for all bits i. Better codes in the minimax sense can be designed by trading $e_1(i)$ for $e_2(i)$. Specifically, it can be shown that

$$2\sum_{i=0}^{m-1} e_1(i) + \sum_{i=0}^{m-1} e_2(i) \geq 4M.$$

This relation indicates that for every two units of increase in $\sum_i e_1(i)$, one might be able to reduce $\sum_i e_2(i)$ by four units. We have been able to modify balanced Gray codes in this manner at low orders (see Section V) to obtain labelings with the following $\max_i e_2(i)$, which we believe is the lowest possible value for any order m.

Conjecture 1: A labeling of order $m > 1$ is a 2-MML if and only if its k-spaced transition count satisfies (4) and

$$\max_i e_2(i) = 2\left\lfloor \frac{M}{m} \right\rceil_2.$$

It is known that balanced Gray codes exist for all orders $m \geq 1$, and if m is a power of two, a balanced Gray code is totally balanced [2], [7], [3, Ch. 7.2.1.1]. Since totally balanced Gray codes satisfy $\max_i e_2(i) = 2M/m = 2\lfloor M/m \rceil_2$, they are 2-minimax optimal and prove the Conjecture for power-of-two orders m. For some other orders (but surprisingly few, namely, $m = 12, 18, 25, 36, 42, \ldots$) it also holds that $\lfloor M/m \rceil_2 = \lceil M/m \rceil_2$, which means that a balanced Gray code satisfies the Conjecture. For all other orders, the 2-MML's, if the Conjecture holds, are not Gray codes. In Section V, we exemplify 2-MML's of orders $m = 3, 5,$ and 6 that satisfy the Conjecture. Finding a general construction method for 2-MML's (or even better, f-MML's for some $f > 2$) remains an open question.

V. Performance of some 2-MML's

In this section, we compare the bit error probabilities of M-PSK with various labelings for communication over an AWGN channel, as a function of E_b/N_0, where E_b is the energy per bit and $N_0/2$ is the double-sided noise spectral density. We are mainly interested in the maximum bit error probability $\hat{P}_b$; however, to indicate how balanced (or unbalanced) the labeling is, and how equal error protection the labeling provides, we also plot the minimum error probability $\check{P}_b$.

In Figure 1, we plot $\hat{P}_b$ and $\check{P}_b$ for a 2-MML of order $m = 3$, compared with two classical labelings, the NBC and the BRGC (see Table I). The codewords of the 2-MML are also shown in Table I. Alternatively, it can be defined through its transition sequence, which is $\{0, \{0,1\}, 0, 2, 1, \{0,1\}, 1, 2\}$. The BRGC is the only Gray code of order $m = 3$, and it is balanced. It is nevertheless not the best labeling in the minimax sense, which is apparent from the figure. The 2-MML labeling is better, although the BRGC performance approaches it at high E_b/N_0.

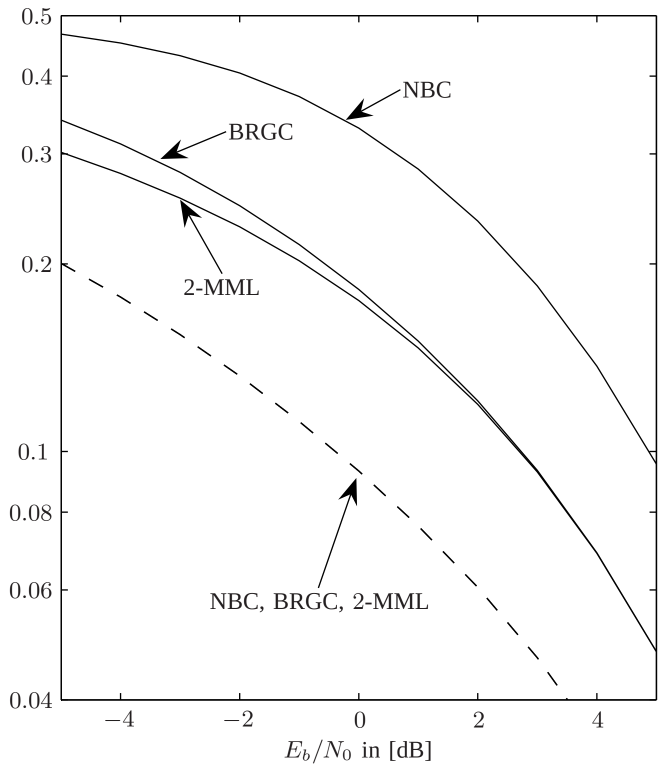

Fig. 1. Plots of $\hat{P}_b$ (solid) and $\check{P}_b$ (dashed) for three labelings of order $m = 3$: the NBC, the BRGC, and a 2-MML.

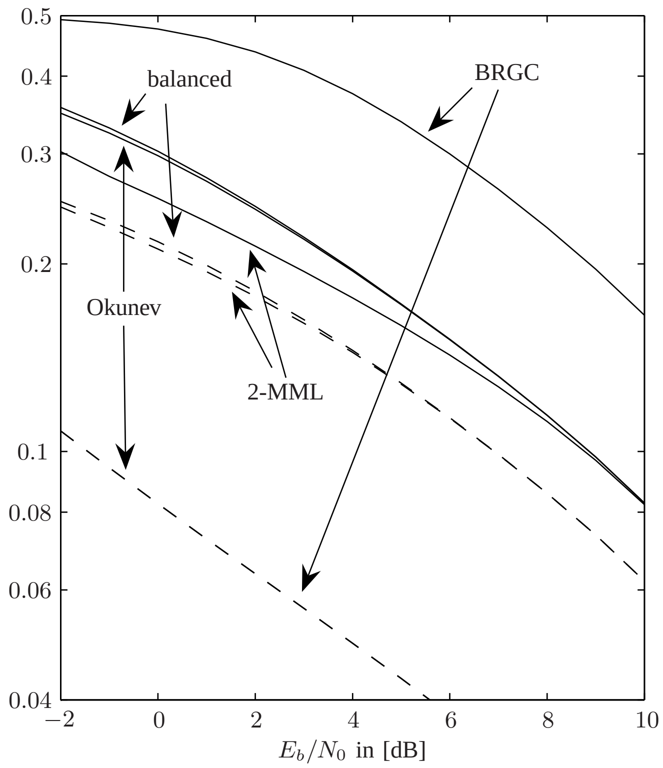

Fig. 2. Plots of $\hat{P}_b$ (solid) and $\check{P}_b$ (dashed) for four labelings of order $m = 5$: the BRGC, a balanced Gray code, a Gray code by Okunev, and a 2-MML.

The transition count matrices for the BRGC and the 2-MML labeling are, respectively,

$$\begin{bmatrix} 4 & 8 & 4 & 0 \\ 2 & 4 & 6 & 8 \\ 2 & 4 & 6 & 8 \end{bmatrix} \qquad \begin{bmatrix} 4 & 4 & 6 & 4 \\ 4 & 4 & 6 & 4 \\ 2 & 4 & 6 & 8 \end{bmatrix}. \tag{5}$$

These matrices have the same maximum value in the first column, which is why the labelings are equivalent asymptotically, but the superiority of the 2-MML is seen in the second column, $e_2(i)$. The corresponding matrix for the NBC is given in (1). It differs from the two other labelings already in the first column, which explains its inferior performance in Figure 1.

The first two columns of the transition count matrix of the 2-MML labeling satisfy Theorem 1 and Conjecture 1. Observe that the elements of the second column, $e_2(i)$, has a constant value, which is characteristic for all 2-MML's in this paper, although there exist 2-MML's without this property.

If we look at the rows of the transition count matrices in (5), we note that the 2-MML has two equally bad rows and one good row (bad and good in the lexicographic sense). This is in contrast to the BRGC which has one bad row and two equally good. As a matter of fact, it can be shown that the BRGC is the *worst* possible Gray code at large SNR when the *maximum* bit error is considered, whereas it in contrast is the *best* possible Gray code (and also the best labeling) for *average* bit error performance, as mentioned in Section I.

Proceeding to higher orders, we skip $m = 4$, where a totally balanced Gray code exists (which is 2-minimax optimal), and exemplify 2-MML's of orders 5 and 6.

For $m = 5$, we study a non-Gray labeling defined by the transition sequence

$$(0, \{0, 1\}, 0, 2, 1, \{0, 1\}, 1, 3, 4, 3, 2, 3, 0, 4,$$
$$2, 4, 1, 4, 2, 4, 0, 3, 0, 1, 2, 1, 0, 3, 4, 2, 1, 3).$$

Its transition count matrix is

$$\begin{bmatrix} 8 & 12 & 16 & 18 & 20 & 20 & \cdots \\ 8 & 12 & 16 & 16 & 18 & 16 & \cdots \\ 6 & 12 & 18 & 24 & 26 & 26 & \cdots \\ 6 & 12 & 14 & 16 & 18 & 20 & \cdots \\ 6 & 12 & 12 & 12 & 16 & 18 & \cdots \end{bmatrix}. \tag{6}$$

It satisfies Theorem 1 and Conjecture 1 and is therefore a 2-MML, assuming that the Conjecture holds.

The maximum and minimum bit error probabilities of this 2-MML are plotted in Figure 2. The corresponding curves are shown for three Gray codes: the BRGC, a balanced code from [2, Fig. 2], and a Gray code designed by Okunev to have the "least nonuniformity" in the sense that all bits except one are as balanced as possible [5, pp. 79–80]. We see that the 2-MML has lower $\hat{P}_b$ for all E_b/N_0 values in the diagram. This can again be explained by the transition count matrix (6). The 2-MML, the balanced Gray code, and Okunev's Gray code all satisfy $\max_i e_1(i) = 8$, but $\max_i e_2(i)$ equals 12 for the 2-MML and 16 for the two others. The BRGC differs from

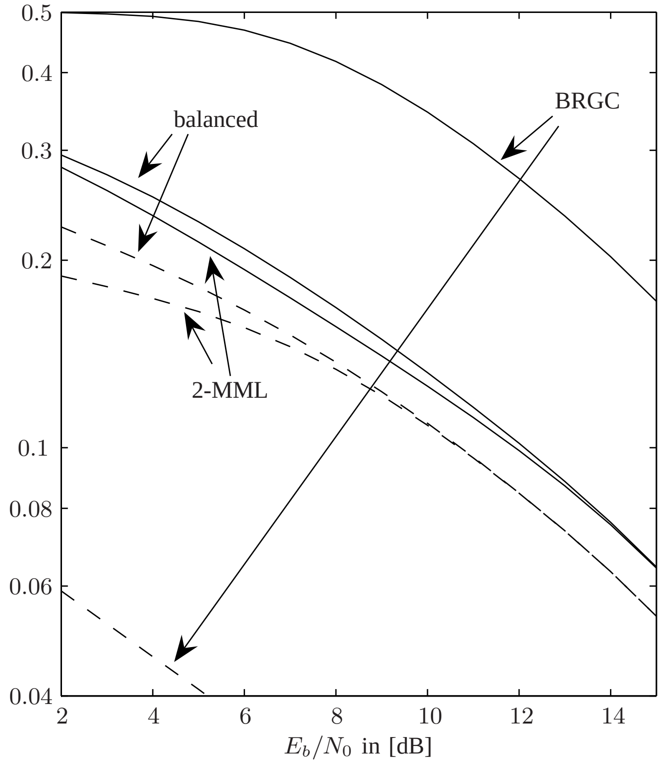

Fig. 3. Plots of $\hat{P}_b$ (solid) and $\check{P}_b$ (dashed) for three labelings of order $m = 6$: the BRGC, a balanced Gray code, and a 2-MML.

the other labelings already in the first column, where it has $\max_i e_1(i) = 16$.

Figure 3 shows a similar comparison of some labelings of order $m = 6$. A (non-Gray) 2-MML is compared with two Gray codes, the BRGC and a balanced Gray code, and the same conclusions as before hold.

The 2-MML of order 6 in Figure 3 is generated by the transition sequence

$$\begin{gathered}(0, \{0,1\}, 0, 2, 1, \{0,1\}, 1, 4, 3, 1, 3, 2, 3, 0\\ 3, \{2,3\}, 3, 5, 4, 0, 4, 2, 4, 0, 4, 1, 4, 0, 4, 3\\ 1, 2, 0, 2, 1, 2, 0, 3, 4, 3, 5, 1, 5, 2, 5, 1, 5, 0,\\ 5, 1, 5, 2, 5, 1, 5, 3, 4, 5, 2, \{2,3\}, 2, 0, 3, 4)\end{gathered}$$

and its transition count matrix is

$$\begin{bmatrix} 12 & 20 & 30 & 36 & 40 & 38 & \cdots \\ 12 & 20 & 30 & 36 & 38 & 34 & \cdots \\ 12 & 20 & 26 & 32 & 38 & 44 & \cdots \\ 12 & 20 & 22 & 24 & 26 & 26 & \cdots \\ 10 & 20 & 20 & 20 & 28 & 36 & \cdots \\ 10 & 20 & 16 & 10 & 16 & 24 & \cdots \end{bmatrix}. \tag{7}$$

The $m = 6$ balanced Gray code is obtained by applying Theorem D in [3, p. 14] twice: first on the $m = 2$ balanced Gray code (underlined transitions are the j_k-transitions in Theorem D)

$$(\underline{0}, 1, \underline{0}, \underline{1})$$

and then on the resulting $m = 4$ labeling

$$(3, \underline{2}, \underline{0}, \underline{1}, 2, \underline{1}, \underline{3}, \underline{1}, \underline{0}, \underline{3}, 2, \underline{3}, \underline{0}, 1, 0, \underline{2}),$$

which produces the $m = 6$ balanced Gray code with transition sequence

$$\begin{gathered}(3, 5, 3, 4, 3, 2, 4, 5, 0, 5, 4, 1, 2, 4, 2, 5,\\ 2, 1, 5, 4, 3, 4, 5, 1, 5, 4, 0, 4, 5, 3, 2, 5,\\ 2, 4, 2, 3, 4, 5, 0, 1, 0, 5, 0, 1, 4, 1, 0, 5,\\ 0, 1, 0, 3, 2, 3, 0, 1, 3, 1, 2, 1, 0, 2, 3, 4),\end{gathered}$$

and the transition count matrix is

$$\begin{bmatrix} 10 & 20 & 22 & 24 & 26 & 28 & \cdots \\ 10 & 20 & 24 & 28 & 30 & 32 & \cdots \\ 10 & 20 & 22 & 22 & 26 & 30 & \cdots \\ 10 & 20 & 22 & 22 & 26 & 32 & \cdots \\ 12 & 24 & 32 & 34 & 30 & 26 & \cdots \\ 12 & 24 & 32 & 36 & 34 & 32 & \cdots \end{bmatrix}.$$

Comparing the second column with (7) shows that this Gray code is not minimax optimal, which is also clear from Figure 3.

VI. Conclusions

The individual bits in M-ary transmission are, in general, subject to unequal error protection. We have argued that it is more natural to design communication systems to minimize the *maximum* rather than the *average* bit error probability. Systems designed according to the minimax principle will be efficient when a certain quality of transmission for all bits must be guaranteed. We have also shown that the ubiquitous binary reflected Gray code is, in general, not optimum in this sense for labeling M-ary PSK constellations. Indeed, there exist labelings that are better than all Gray codes, and we have designed such labelings for a few orders.

References

[1] E. Agrell, J. Lassing, E. G. Ström, and T. Ottosson, “On the optimality of the binary reflected Gray code,” *IEEE Transactions on Information Theory*, vol. 50, pp. 3170–3182, Dec. 2004.

[2] G. S. Bhat and C. D. Savage, “Balanced Gray codes,” *The Electronic Journal of Combinatorics*, vol. 3, no. 1, 1996. paper no. R25.

[3] D. E. Knuth, *The Art of Computer Programming*, vol. 4, fascicle 2. Addision-Wesley, 2005.

[4] J. Lassing, E. Agrell, E. G. Ström, and T. Ottosson, “Optimality of the binary reflected Gray code for the Gaussian channel,” Tech. Rep. R001/2005, Chalmers University of Technology, Dept. of Signals and Systems, Communication Systems Group, 2005.

[5] Y. Okunev, *Phase and Phase-Difference Modulation in Digital Communications*. Artech House Publishers, 1997.

[6] C. Savage, “Comments by the second author on R25 of volume 3 (1), 1996,” *The Electronic Journal of Combinatorics*, 2005.

[7] D. G. Wagner and J. West, “Construction of uniform Gray codes,” *Congressus Numerantium*, vol. 80, pp. 217–223, 1991.